International Conference on Fuzzy Sets and Soft Computing in Economics and Finance
FSSCEF 2006

Proceedings

Saint-Petersburg, Russia
June 28 - July 1, 2006

2

ISSN 1818-6750

Copyright © 2006

Russian Association for Fuzzy Systems and Soft Computing
International Fuzzy Economics Lab (Russia)
Instituto Mexicano del Petroleo, Mexico

http://fsscef.narod.ru

Editors:
Ildar Z. Batyrshin
Alexey O. Nedosekin

This work is subject to copyright. All rights are reserved. Reproduction of this publication in any form by any electronic or mechanical means (including photocopying, recording or information storage and retrieval) is permitted only under the permission of editors.

Printed in Russia
Supported by Russian Foundation for Basic Research

International Conference on Fuzzy Sets and Soft Computing in Economics and Finance, FSSCEF 2006, Saint-Petersburg, Russia, June 28 - July 1, 2006, was organized second time in Saint-Petersburg, in one of most famous Russian cultural and scientific centers. The conference is organized as a platform for exchanging ideas, experiences, and opinions among the academicians, professional engineers and practitioners on the application of fuzzy sets and soft computing methods and techniques to economics and finance.

The proceedings of the conference contain papers presented by researchers from 11 countries. Applications of fuzzy sets, soft computing and uncertainty models in economics and finance are growing each year. Important role in development and applications of uncertainty models plays interval mathematics which is highly related with the theory of fuzzy sets. The methods of interval computations and their applications were discussed in invited lecture of Prof. V. Kreinovich from University of Texas at El Paso, USA. During the conference, two special sessions on “Interval and Fuzzy Techniques in Business Applications” and “Applications of Interval Uncertainty Models” were organized. Also the following sessions were presented in conference program: “Fuzzy Modeling in Economics and Finance”, “Investments and Risk Analysis”, “Neural Networks and Genetic Algorithms in Decision Making and Optimization”, “Information Security and Decision Making in Complex Systems”, “Analysis of Time Series in Economics and Finance”.

We would like to thank all the authors and participants for contributing to the conference and stimulating the technical discussions.

Ildar Z. Batyrshin, Alexey O. Nedosekin, Nadezhda G. Yarushkina
Program and Organizing Committees Co-Chairs

Saint Petersburg
June 2006

TABLE OF CONTENTS

Interval Computations and Their Applications

Vladik Kreinovich

Department of Computer Science,
University of Texas at El Paso,
El Paso, TX 79968, USA
vladik@utep.edu
http://www.cs.utep.edu/vladik

In many areas of science and engineering, uncertainty is an (important but) relatively minor factor. In many areas in science and engineering, the dynamics of different systems is reasonably predictable: when we apply a control, e.g., when we turn a car, we can reasonably well predict where and how fast the car will be going. Of course, due to measurement inaccuracy and incomplete knowledge of factors such as road conditions, there is always some uncertainty (see, e.g., [9]), but the resulting prediction uncertainty is usually relatively small in comparison with the actual values of the predicted quantity.

In economics and finance, uncertainty is extremely important. In contrast to science and engineering, in economics, and especially in finance, uncertainty plays a major role. Many successful investments and financial tools and techniques are based on the fact that it is difficult to accurately predict the results of different investments.

Probabilities: traditional approach to describing uncertainty in economics and finance. Traditionally, the uncertainty in economics and finance is described by statistical models. This description is the basis of the current financial mathematics, that started largely with the Nobel-prize winning work of Black and Scholes; see, e.g., [1, 2].

Limitations of the traditional probabilistic approach. One of the main limitations of the traditional probabilistic approach to uncertainty is that this approach requires that we known the exact values of all the relevant probabilities.

In many situations in science and engineering, the processes are reasonably stationary, so in principle, by making sufficiently many calibration measurements, we can determine – with reasonable accuracy – all the probabilities related to measurement uncertainty; see, e.g., [9].

In economics and finance, however, the situation often changes; so by the time we need to make important decisions, we do not have enough statistics about the current situation to accurately determine all the needed probabilities.

How we can overcome these limitations: the need to consider interval uncertainty. There are two natural approaches to overcoming the above limitations.

The first approach takes into account that expert knowledge is not limited to statistics; experts also use rules and techniques based on their previous expertise. Experts can rarely formulate their rules in precise mathematical terms; most of these rules are formulated by using words from natural language, like "if the price of the stock drastically decreases, sell it". To formulate this knowledge inside a computer, it is reasonable to use a special technique specifically designed for formalizing this knowledge – the techniques of fuzzy sets; see, e.g., [6].

The second approach is accept the fact that we do not know the exact values of the probabilities p, we only the intervals $[\underline{p}, \overline{p}]$ of possible values of these probabilities – or that sometimes, we do not have any information about the probabilities of different values of the quantity at all, we only know the interval $[\underline{x}, \overline{x}]$ of possible values of this quantity x.

These approaches are inter-related – because, in effect, a fuzzy set can be described as a nested family of intervals (α-cuts) corresponding to different degrees α.

Interval computations: a technique for handling interval uncertainty. There exist numerous techniques for handling situations in which we have a combination of probabilistic and interval uncertainty; see, e.g., [8, 10].

All these techniques, in effect, cover the following situation:

- we have an (ideal-case) algorithm $f(x_1, \ldots, x_n)$ that transforms the exact values of the parameters x_i (e.g., values of some quantity, or probabilities of different values of this quantity, or characteristics of the corresponding probability distribution) into the value $y = f(x_1, \ldots, x_n)$ of the desired quantity;
- we know the intervals $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$ of possible values of the parameters x_i;
- we want to find the corresponding range of possible values of y:

$$\mathbf{y} = f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \{f(x_1, \ldots, x_n) \,|\, x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\}.$$

Techniques for solving this problem are called techniques of *interval computations*; see, e.g., [3–5].

This area of research was started largely by works of Ramon Moore who developed the corresponding techniques in the late 1950s and early 1960s – to take into account interval uncertainty when planning spaceflights to the Moon. Since 1950s, these techniques have been widely applied to many areas of science and engineering, as well as to economic and financial problems.

Straightforward interval computations: main idea. Historically the first method for computing the enclosure for the range is the method which is sometimes called "straightforward" interval computations. This method is based on the fact that inside the computer, every algorithm consists of elementary operations (arithmetic operations, min, max, etc.). For each elementary operation $f(a, b)$, if we know the intervals $\mathbf{a}$ and $\mathbf{b}$ for a and b, we can compute the exact range

$f(\mathbf{a}, \mathbf{b})$. The corresponding formulas form the so-called *interval arithmetic*. For example,

$$[\underline{a}, \overline{a}] + [\underline{b}, \overline{b}] = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]; \quad [\underline{a}, \overline{a}] - [\underline{b}, \overline{b}] = [\underline{a} - \overline{b}, \overline{a} - \underline{b}];$$

$$[\underline{a}, \overline{a}] \cdot [\underline{b}, \overline{b}] = [\min(\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b}), \max(\underline{a} \cdot \underline{b}, \underline{a} \cdot \overline{b}, \overline{a} \cdot \underline{b}, \overline{a} \cdot \overline{b})].$$

In straightforward interval computations, we repeat the computations forming the program f step-by-step, replacing each operation with real numbers by the corresponding operation of interval arithmetic. It is known that, as a result, we get an enclosure $\mathbf{Y} \supseteq \mathbf{y}$ for the desired range.

In some cases, this enclosure is exact. In more complex cases (see examples below), the enclosure has excess width.

More sophisticated techniques for interval computations. There exist more sophisticated techniques for producing a narrower enclosure, e.g., a centered form method.

Comment. For each of the known interval computations techniques, there are cases when we get an excess width. Reason: as shown in [7], the problem of computing the exact range is known to be NP-hard even for polynomial functions $f(x_1, \ldots, x_n)$ (actually, even for quadratic functions f).

As a result, new more and more efficient techniques are being developed.

What will be covered in the tutorial. In the tutorial, we will describe the main interval computations techniques, provide examples of applications of interval computations techniques, including applications to economic and financial problems.

We also show how interval computations techniques can be (and are) used to process fuzzy uncertainty – indeed, if we know the fuzzy sets X_i for x_i, then , for every α, the α-cut $Y(\alpha)$ of the resulting fuzzy set Y for y is equal to the range $Y(\alpha) = f(X_1(\alpha), \ldots, X_n(\alpha))$ of the function $f(x_1, \ldots, x_n)$ on the corresponding α-cuts $X_i(\alpha)$.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

References

1. Black, F., Scholes, M.: The pricing of optional and corporate liabilities. Journal of Political Economy **81** (1973) 637–659.

2. Heath, D. C., Swindle, G. (Eds.), *Introduction to Mathematical Finance*, American Mathematical Society, Proveidence, RI, 1997.
3. Interval computations website http://www.cs.utep.edu/interval-comp
4. Jaulin, L., Kieffer, M., Didrit, O., Walter, E.: *Applied Interval Analysis*, Springer Verlag, London, 2001.
5. Kearfott, R. B., Kreinovich, V., Eds.: *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.
6. Klir, G., Yuan, B.: *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, New Jersey, 1995.
7. Kreinovich, V., Lakeyev, A., Rohn, J., Kahl, P. *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1997.
8. Kuznetsov, V. P.: *Interval statistical models*, Radio i Svyaz Publ., Moscow, 1991 (in Russian).
9. Rabinovich, S.: *Measurement Errors: Theory and Practice*, American Institute of Physics, New York, 2005.
10. Walley, P.: *Statistical reasoning with imprecise probabilities*, Chapman and Hall, N.Y., 1991.

The Stock Market Forecasting: An Application of the Interval Measurement and Computation

Ling He[1] and Chenyi Hu[2]

[1]Department of Economics and Finance
[2]Department of Computer Science
University of Central Arkansas
Conway, AR 72035, USA
{linghe, chu}@uca.edu

Abstract : This study forecasts changes in the stock market by using the interval measured data and interval computing approach. Empirical results in this study strongly suggest that the interval forecasting is superior to the traditional point data based OLS forecasting, in terms of significantly smaller mean and standard deviation of forecast errors Moreover, the accuracy ratio indicates that the out-of-sample interval forecasting has an actual forecasting accuracy over 58%. The ratio becomes more than 63% for the in-sample interval forecasting.

1. Introduction

1.1. Major issues in the stock market forecasting

One of modern capital asset pricing theories, the arbitrage pricing theory (APT) of Ross (1976) provides a theoretical framework to identify macroeconomic variables that can significantly or systematically influence stock prices. In an early attempt to discover economic driving forces for the stock market, Chen, Roll and Ross (1986) find five systematic factors that are still widely accepted in the literature: the spread between long- and short-term interest rates; expected and unexpected inflations; industrial production; and the spread between high- and low-grade bonds. All of them represent sources of risk that “are significantly priced" in the stock market. In other words, these macroeconomic variables have predicting power on the stock market. Therefore, they can be used to forecast changes in the stock market. This study adopts these five variables in the stock market forecasting.

There is a consensus in the literature that sensitivities of the stock market to macroeconomic factors vary over time. Many statistical methods have been introduced to deal with the time-varying issue. For instance, Brown, Durbin, and Evans (1975) developed the Cusum and Cusum of Squares tests to find switch points for a multi-factor regression model. Fama and French (1997) use the rolling regressions to produce time-varying regression coefficients and forecasts for excess stock returns of industrial portfolios. In a recent study, He (2005) suggests exponentially weighted rolling regressions to gradually discount old observations kept in the estimation period, under the assumption that earlier observations have less influence on coefficient estimates. He also uses the Flexible Least Squares (FLS) developed by Kalaba and Tesfatsion (1988, 1989, 1990) to generate rolling forecasts. The results are superior to those from the Cusum and Cusum of Squares, normal and exponentially weighted rolling regressions.

Nonetheless, the poor forecasting quality is a persistent problem. Fama and French (1997) report an average monthly forecast error of 2.72% for their out-of-sample rolling forecasts. The number for He’s (2005) sample is 1.88%. The FLS monthly rolling forecasts have a mean error of 1.57%. Apparently, multiplying by 12, these monthly numbers can be turned into big annual ones. A new feasible approach to improve the forecasting quality is to use a more effective data measurement and estimation technique.

The primary measurement in economic and financial data is to quantify points. Take the stock market as an example. A daily stock price index consists of daily closing prices and a monthly index is the average of daily index numbers. They simply reflect a particular number (level) or an average number at a particular time spot. When a growth rate is in need, a current point is then compared with a point at the previous time spot. Only the level change (a point) over the time interval (daily, monthly, or annual) is measured at the end of the period. The primary purposes of various forecasting methods based on the point measured data are in common, that is, to predict points that can match actual ones. A major shortcoming of the point data and point forecasting is that they are unable to use and forecast variability in a time period. In fact, not only regression coefficients are time-varying, macroeconomic variables are time-varying as well. For a given trading day (month or year), there are two price bounds, the highest and lowest. The two prices reflect the variability of the day (month or year). In addition, in the real world forecasts are often in the form of a range or interval. It is not unusual at all for an economist to make the next year's inflation forecast to be, for example, between 2-4%. It is a forecasted inflation interval for the next year. Unfortunately, only the point data and point-based forecasting methods currently exist in the economic and financial forecasting literature.

1.2. Interval forecasting methods

In the past two decades some statisticians developed various point-based interval forecasting methods. However, they are fundamentally "semi" interval forecasts, because they are based on point data and point-based forecasting methods, such as Bayesian, Bootstrapping, Box-Jenkins, GARCH, and Holt-Winters methods. The general forecasting principle of these methods is that the forecasted interval is simply the sum of the point forecast and some percent of positive and negative variance of forecasts (Chatfield, 1993, 1998). There are many reasons why those "semi" interval forecasting methods are never widely used. The most important reason is the poor quality. Normally, those forecasted intervals are so narrow that there is only a 50% chance for a future point lying inside the interval (Gardner, 1988 and Granger, 1996). Evaluating "semi" interval forecasts is another major problem (Christoffersen, 1998 and Clements and Taylor, 2003). This problem stems from the methodology of "semi" interval forecasting. In the process of estimating the point forecast, the aim is to minimize estimation errors. However, some percent of estimation variance is added to and subtracted from the point forecast to form an interval forecast. This contradiction makes traditional forecast error measure impossible.

1.3 The interval computing approach

In the mathematical and computing fields, a new method, interval computing, has theoretically or computationally been developed since the late 1970s (Moore, 1979). In interval computation, both operands and computational results are intervals. By taking interval valued parameters (more information) into considerations, one may obtain computational results that can be reliable both computationally and mathematically with interval arithmetic. Numerous results have been obtained with interval computing especially for some otherwise very difficulty computational problems such as reliable non-linear global optimization and others (Kearfott, 1996; Hu, Xu and Yang, 2002; Kearfott and Hongthong, 2005). By applying interval computing, Stadtherr's group discovered roots of equations never reported previously in chemical engineering (Gau and Stadtherr, 2000; Hua, Brennecke, and Stadtherr, 1996). These findings result in the 1998 Computing in Chemical Engineering Award of American Institute of Chemical Engineers. More recently, the interval computation has been extended to different computing areas such as data mining, decision making, game theory, and others (Korvin, Hu and Chen, 2002 and 2004; Collins, 2005).

The interval computation creates a new forecasting format and perhaps, more accurate forecasts. The major purpose of this study is to employ the interval method to forecast changes in the stock market.

In order to make interval forecasts, a different data measurement, interval measurement, must be used. A mathematical interval [a, b] is the set $\{x \in \Re \mid a \le x \le b\}$ where $a \le b$. Mathematical intervals are represented and operated with machine intervals inside computers. In an interval data set, every observation is measured as an interval, a combination of an upper bound and a lower bound. Contrast to the point measurement, interval measurement automatically contains the (level) information measured by the point number as well as the variability in the time duration. Use stock price as an example. The closing price (Level) measured by the point number must locate in somewhere between the upper and lower bounds of the trading day. Moreover, the range between the upper and lower bounds represents the maximum variation of the day. A variation measurement is a risk measurement. Some traditional statistical measurements, such as range, variance, standard deviation, and coefficient of variation, are examples for risk measurements. In the same sense, intervals also measure risk. Given that more information is included in the interval measurement, it is reasonable to expect a higher quality for the interval forecasting.

The main objective of interval computing procedure is to match the center of two interval vectors, in order to measure their relationships which are essential to the interval forecasting. The predicted interval reveals future variation or risk which is important in determining required returns and capital asset value, there is no doubt that the output of the interval forecasting can provide useful information to investors and economic/financial policy-makers.

The main objective of interval computing procedure is to match the center of two interval vectors, in order to measure their relationships which are essential to the interval forecasting. The predicted interval reveals future variation or risk which is important in determining required returns and capital asset value, it is no doubt that the output of the interval forecasting can provide useful information to investors and economic/financial policy-makers.

The remainder of this paper is organized as follows. Section 2 describes the estimation model and data, Section 3 discusses interval computation procedures, Section 4 presents empirical results, and Section 5 contains concluding comments.

2 Model and data

According to Chen, Roll and Ross (1986), changes in the stock market (SP_t) is linearly determined by the following five macroeconomics factors: growth rate variations of seasonally-adjusted Industrial Production Index (IP), changes in expected inflation (DEI_t) and unexpected inflation (UI_t), default risk premiums (DEF_t), and unexpected changes in interest rates ($TERM_t$):

$$SP_t = a_t + I_t(IP_t) + U_t(UI_t) + D_t(DEI_t) + F_t(DEF_t) + T_t(TERM_t) + e_t \quad (1)$$

This study uses time series data that covers the period of January 1930-December 2004. There is no specific reason, other than the availability of data, for the selected sample period. The data set includes the following basic monthly series:

- *IP*: the growth rate of seasonally-adjusted Industrial Production Index at the beginning of the month. This study uses the one-month lead term of IP in the monthly data.
- *LONG*: the monthly returns on long-term U.S. government bonds.
- *CORP*: the monthly returns on long-term corporate.

- *SHORT*: the monthly returns on one-month U.S. Treasury Bills. According to Fama and French (1993), $SHORT_{t-1}$ is the proxy for "the general level of expected returns on bonds."
- *CPI*: the growth rate of the U.S. Consumer Price Index.
- SP: the growth rate of Standard and Poor's 500 Stock Price Index.

For IP, CPI, and SP, the monthly growth rate is the first difference of natural logarithms in months of t and t-1; the annual growth rate is the first difference of natural logarithms in Decembers of t and t-1. The annual returns for LONG, CORP, and SHORT are compounded monthly returns of January through December. The following additional series are derived from the above basic series:

$DEF_t = CORP_t - LONG_t$. It represents the default risk premium (Fama and French, 1993).

$TERM_t = LONG_t - SHORT_{t-1}$. It measures unexpected changes in interest rates (Fama and French, 1993).

UI_t: unexpected inflation. It is proxied by the residuals from the following regression model (Fama, 1981):
$CPI_t = \alpha_{t-1} + \beta SHORT_{t-1} + \eta_t$

EI_{t-1}: expected inflation at the end of month t-1. It is the difference of CPI_t and UI_t.

DEI_t: the change in expected inflation. It measures the difference of EI_t and EI_{t-1}.

In addition to monthly inflation variables, this study also calculates quarterly expected and unexpected inflation variables. By multiplying 4, the quarterly inflation variables are converted into annual variables (Fama, 1981).

3. Interval algorithm and estimation accuracy

3.1. An interval least squares algorithm

In a given time period, for instance, a month or a year, economic data can be expressed as intervals. An interval measures the greatest variation for a given time period. In this study the maximum monthly number in a given year forms the upper bound of the year's interval; while the minimum monthly number in the year becomes the lower bound of the interval. After construction of the annual interval data, it is necessary to minimize an objective function, in terms of the least squares principle, with the interval data. Different from the ordinary least squares algorithm, an interval valued linear system of equations needs to be solved to estimate the coefficients in (1). Therefore, an interval least squares algorithm has been developed and implemented in C++ in this study to solve the problem. Details of the algorithm and software will be reported in another paper to appear in the field of computer science.

Based on the least squares principle, the interval arithmetic (Moore, 1979) is used to construct an interval valued linear system of equations. In this process, the product of two intervals are used mostly. In interval arithmetic, the product of two intervals is defined as [a, b]* [c, d] = [min {ac, ad, bc, bd}, max{ac, ad, bc, bd}].

In order to estimate the coefficients, it is necessary to solve an interval linear systems of equations $\mathbf{M}a = \mathbf{v}$, where **M** is a 6 by 6 interval matrix; *a* is the vector of the coefficients to be determined; and **v** is an interval vector. It is assumed that the coefficients are scalars initially. By taking M_{mid}, the mid-point matrix of **M,** and v_{mid}, the midpoint vector of **v**, a classic linear system of equations about *a* is constructed. The numerical estimations of the coefficients are obtained by using Gaussian elimination with scaled partial pivoting. This initial approach has the intuition of

matching the center of two interval vectors. However, it has not yet taken the widths into considerations. Therefore, it is essential to adjust the width of the forecasted interval. Consistent with the rolling estimation period, the width is adjusted by a rolling scalar constant which is equivalent to the average width of previous ten years.

The coefficient estimates obtained from Equation (1) may be used to forecast changes of the stock market by estimating the following the out-of-sample forecasting model:

$$SP_t = a_{t-1} + I_{t-1}(IP_t) + U_{t-1}(UI_t) + D_{t-1}(DEI_t) + F_{t-1}(DEF_t) + T_{t-1}(TERM_t) \quad (2)$$

In order to better reflect time-varying relationships between the stock market and other macroeconomic variables, a rolling estimation period of ten consecutive years is used to establish the interval linear system of equations. Then Equation (2) is estimated to forecast the stock market, starting with the eleventh year.

3.2. Accuracy of interval forecasting

The interval forecasting makes an effective assessment measure feasible. Both the input and output of interval forecasting are intervals. The range covered by both the forecasted and actual intervals represents the accurate part of an interval forecast. The quality of forecasting can be easily measured by a ratio of the commonly covered range to the maximum distance reached by the forecasted and actual intervals. This ratio reflects the *accuracy* of the forecast. For example, if the predicted interval is [2, 5] and the actual interval is [1, 3], then the overlapped range is the interval [2, 3]. The width of the overlapped interval is 1 (3-2) and the maximum distance is the width of the interval [1, 5], that is 4 (5-1). Therefore, the accuracy ratio is 25%. If the actual interval is used as the denominator, the ratio increases to 50%, however, it is misleading. The accuracy ratio reflects the real quality of forecasting and does not need any confidence assumption. Consider the following example: the predicted interval is [1, 5] and the actual interval remains the same, [1, 3]. The accuracy ratio should be 40% (2/5). If the actual interval is used as the denominator, the value of the ratio becomes 100%. Obviously, it is wrong. In this study, SP_{est} represents the forecasted SP interval. Then, the concept of the estimation accuracy is defined as: $w(SP \cap SP_{est})/w(SP \cup SP_{est})$, where w is the width function of an interval.

4. Empirical results

Table 1 provides summary statistics for the monthly and annual point data as well as the annual interval data. Compared with the interval data, the point data (monthly and annual) are more volatile, evidenced by higher standard deviations. Given the higher stability for both upper and lower bounds of the interval data, the range between the two bounds should be the major source for the higher variability of the point data, because most points locate somewhere between the upper and lower bounds. The point measurement only quantifies an item at a particular time spot; while an interval measurement reflects the greatest variation in the item between two time spots. Therefore, most of the volatility of the point measurements should be covered in the intervals. In other words, intervals primarily serve as the variability measurement or risk measurement.

The OLS estimates of Equation (1) indicate that the model has a higher explanatory power (R^2=26%) with the annual data, rather than the monthly data (R^2=17%). The better fitness of the data guarantees better forecasting results. In order to compare the forecasting results generated by the OLS and interval methods, this study uses the annual data.

Out-of-sample rolling forecasts of changes in the stock market based on Equation (2) are reported in Table 2. The OLS forecasting uses monthly, quarterly, and annual point data, respectively. The monthly average of forecast errors is about 2.65% with a

standard deviation of 2.44%. By multiplying 12, those two numbers can be converted into annual figures of 31.83% and 29.34% which are much larger than those for annual forecasts. The annual mean of forecast errors is 20.57% and the standard deviation of forecast errors is 19%. The results unequivocally suggests that annual forecasts have the best quality among the three types of OLS forecasts.

Nonetheless, forecast errors of the annual interval forecasting are even smaller and more stable. The mean of forecast errors is only 6.25% for intervals, the combinations of the upper and lower bounds. The interval forecasts are highly consistent over time. The standard deviation of forecast errors is as low as 3.73%. Equality tests provide evidence that the quality of interval forecasts is significantly higher than OLS forecasts. The T-statistic (5.96) suggests that the mean of forecast errors for interval forecasts is significantly lower than that for OLS forecasts. In addition, both the Newbold (1995) F-statistic (25.98) and Bartlett-statistic (123.6) reject the null hypotheses of equality and homogeneity of variances in favour of interval forecasts. All three statistics are significant at the 1% level. More importantly, there is a direct quality measure for the interval forecasting. The ratio of forecasting accuracy is defined as the overlapped range by both the forecasted interval and the actual interval divided by the maximum range stretched by both the predicted and actual intervals. The ratio is 58.21% in this study. It indicates that the interval approach can predict about 58% of changes in the stock market. This accuracy ratio measures the real preciseness of interval forecasts without any confidence assumption, therefore, is only feasible for the interval measurement. When the point measurement is used, it is impossible to find the overlapped range between the forecasted and actual points. As a result, the forecasting quality is traditionally assessed by a negative measure, the forecast error.

In-sample rolling forecasts tell the similar story (Table 3). Since the current coefficient estimates, rather than their lag terms, are used in in-sample forecasting, a better fitness of forecasts to the data is expected. Therefore, it is not surprising that forecast errors for all kinds of forecasts become smaller, compared to the out-of-sample forecasts. The biggest drop is in the annual OLS forecasts, from 20.57% (Table 2) to 6.52% (Table 3). Nevertheless, the interval forecasts still have significantly lower forecast errors and their standard deviation. The ratio of forecasting accuracy for interval forecasts also increases to 63.47%.

Figures 1 and 2 display the OLS and interval rolling forecasts of SP contrasted with the actual growth rates of SP. The scale in the interval forecast graph (Figure 2) ranges between -0.2 and 0.2, due to the stability of the forecasts. The graph of OLS forecasts has a range from -1 to 1.5. It is evident that OLS forecasts are unstable over time and have large forecast errors. In contrast to the OLS forecasts, interval forecasts have smaller errors in general. The lower bound forecasts have about 10 perfect or near perfect matches with the real lower bounds.

5. Concluding comments

Traditionally, forecasting in Economics and Finance uses point measured data and predicts a particular point in future. Point forecasts are normally imprecise. In order to improve forecasting quality, this study uses a different data measurement, interval, to forecast the range of future stock market changes. Several benefits of this new approach are apparent. First, intervals are an effective economic and financial information measurement. An interval contains information about the level as well as variability which is an important factor in asset pricing. Second, the output of the interval forecasting is more useful than the point forecasting. An interval forecast essentially consists of predicted levels and a predicted variability. Theoretically, the predicted variability can reduce uncertainty or risk, therefore, may influence required returns and capital asset prices. Third, interval forecasting can generate high quality results, because interval data contains more information than does point data. Finally, a unique quality assessment measure is feasible for interval forecasts. The ratio of forecasting accuracy compares the correctly predicted range to the maximum range

stretched by the forecasted and actual intervals. Unlike the mean and standard deviation of forecast errors, the accuracy ratio is a direct and simple assessment of forecasting quality. It does not require any statistical assumptions. Nevertheless, the interval forecasting does require interval data as inputs.

Empirical results in this study provide strong evidence that the interval forecasting is superior to the traditional point data based OLS forecasting. Both the mean and standard deviation of forecast errors for interval forecasts are statistically lower than that for the OLS forecasts. Moreover, the accuracy ratio indicates that the out-of-sample interval forecasting has an actual forecasting accuracy over 58%. The ratio becomes more than 63% for the in-sample interval forecasting.

This study is merely the first attempt to use the interval computing approach in financial forecasting. Future research in this area may be fruitful by further exploring the accuracy of interval forecasting by using higher frequency data (monthly or daily) and extending the interval forecasting into other important economic and financial aspects.

References:

Brown R., Durbin, J., and Evans, J. (1975). Techniques for testing the consistency of regression relationships over time. *Journal of the Royal Statistical Society, Series B*, 37, 149-192.

Chatfield, C. (1993). Calculating interval forecasts. *Journal of Business and Economic Statistics* 11, 121-144.

Chatfield, C. (2001). Prediction intervals for time-series forecasting. In *Principles of Forecasting: Handbook for Researchers and Practitioners* edited by J.S. Armstrong. Boston: Kluwer Academic Publisher.

Chen, Nai-fu, Roll R., and Ross S. (1986). Economic forces and the stock market. *Journal of Business* 59, 383-403.

Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review* 39, 841-862.

Clements, M.P. and Taylor N. (2003). Evaluating interval forecasts of high-frequency financial data. *Journal of Applied Econometrics* 18, 445-456.

Collins, D. W. (2005). On interval valued Nash games. Graduate thesis, Computer Science Department, University of Central Arkansas.

Fama, E. (1981). Stock returns, real activity, inflation, and money. *The American Economic Review* 71, 545-565.

Fama, E., and French K. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E., & French, K. (1997). Industry costs of equity. *Journal of Financial Economics* 43, 153-193.

Gardner, E.S. (1988). A simple method of computing prediction intervals for time series forecasts. *Management Science* 34, 541-546.

Granger, C.W.J. (1996). Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics* 11, 455-473.

He, L. (2005). Instability and predictability of factor betas of industrial stocks: The flexible least squares solutions. *The Quarterly Review of Economics and Finance* 45, 619-640.

Hu, C., Xu, S and Yang, X. (2002). A review on interval computation software and applications. *International Journal of Computational and Numerical Analysis and Applications* 1(2), 149-162.

Hu, C., Cardenas, A. Hoogendoorn S., and Selpulveda, P. (1998). An interval polynomial interpolation problem and its Lagrange solution. *Journal of Reliable Computing* 4, 27-38.

Hua, Z., Brennecke, J. F. and Stadtherr, M. A. (1996) Reliable Prediction of Phase Stability Using an Interval-Newton Method. *Fluid Phase Equilibria*, 116, 52-59

Kalaba, R., & Tesfatsion, L. (1988). The flexible least squares approach to time-varying linear regression. *Journal of Economic Dynamics and Control* 12, 43-48.
Kalaba, R., & Tesfatsion, L. (1989). Time-varying linear regression via flexible least squares. *Computers and Mathematics with Applications* 17, 1215-1245.
Kalaba, R., & Tesfatsion, L. (1990). Flexible least squares for approximately linear systems. *IEEE Transactions on Systems, Man, and Cybernetics SMC* 20, 978-989.
Kearfott, R. B. (1996). *Rigorous global search: Continuous problems*. Kluwer Academic Publishes.
Kearfott, R. B. (1996). Interval extensions of non-smooth functions for global optimization and nonlinear systems solvers. *Computing* 57, 149-162.
Kearfott, R. B. and Hongthong S. (2005). Validated linear relaxations and preprocessing: Some experiments. *SIAM Journal on Optimization* 16, 418-433.
Korvin, A., Hu, C. and Chen P. (2002). Association analysis with interval valued fuzzy sets and body of evidence. *Proc. of the 2002 IEEE Int. Conf. on Fuzzy Systems*, pp.518-523.
Korvin, A., Hu, C. and Chen P. (2004). Generating and applying rules for interval valued fuzzy observations. *Lecture Notes on Computer Science* 3177: 279-284.
Moore, R. E. (1979). Methods and applications of interval analysis. *SIAM Studies in Applied Mathematics*. Philadelphia: SIAM.
Newbold, P. (1995). *Statistics for Business and Economics*. New Jersey: Prentice-Hall.
Ross, Stephen. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13: 341-360.

APPENDICES:

Table 1. Summary statistics of monthly and annual data (in percent) and regression coefficients

Monthly (1/1930-12/2004)	SP	IP	UI	DEI	DEF	TERM	
Mean	0.45	0.29	0.002	-0.000	0.04	0.17	
Standard Deviation	4.50	1.90	0.53	0.06	1.18	2.30	
	IP	UI	DEI	DEF	TERM	Con	R^2
OLS coefficients							
With SP as	0.79	0.03	-2.92	0.97	0.37	0.00	0.17
dependent variable	(10.87)[a]	(0.12)	(-1.23)	(7.35)[a]	(5.39)[a]	(0.87)	
Annual (1930-2004)	SP	IP	UI	DEI	DEF	TERM	
Mean	5.37	3.49	0.21	-0.03	0.46	2.04	
Standard Deviation	19.06	9.94	4.57	1.30	3.41	8.84	
	IP	UI	DEI	DEF	TERM	Con	R^2
OLS coefficients							
With SP as	0.93	-0.40	2.53	0.52	0.48	0.01	0.26
dependent variable	(4.20)[a]	(-0.83)	(1.61)	(0.77)	(1.86)[c]	(0.48)	
Annual interval(1930-2004)	SP	IP	UI	DEI	DEF	TERM	
Upper bound mean	6.61	2.36	0.66	0.08	1.82	3.44	
Standard Deviation	2.59	0.08	0.01	0.00	0.05	0.22	
Lower bound mean	-6.30	-1.80	-0.51	-0.08	-1.66	-3.14	
Standard Deviation	0.75	0.24	0.09	0.00	0.15	0.90	

SP=Growth rate in S&P stock index.
IP=Growth rate in industrial production index.
UI=Unexpected inflation.
DEI=Changes in expected inflation.
DEF=Default risk premium.
TERM=unexpected changes in interest rates.
Con=Constant term.

t-values in parentheses.
The 1% significant level represented by a.
The 10% significant level represented by c.

Table 2. Out-of-sample rolling forecasts of changes in the stock market: OLS vs. interval results

	Monthly	Quarterly	Annual	Interval
Rolling window	60 months	20 quarters	10 years	10 years
Forecasting period	1/40-12/04	1Q/40-4Q/04	1940-2004	1940-2004
Forecast error mean(E)	0.026522	0.048455	0.20572	
Std. dev. Of E	0.024446	0.044082	0.18996	
Forecast error mean Of Upper bound(UE)				0.028631
Std. dev. Of UE				0.024216
Forecast error mean Of lower bound(LE)				0.033874
Std. dev. Of LE				0.031837
Forecast error mean Of Interval (IE=UE+LE)				0.062504
Std. dev. Of IE				0.037266
Ratio of forecasting Accuracy (FA)				0.582072
Std. dev. Of FA				0.189395

Equality tests for the following pairs

	Annual	Interval	Test result
Forecast error mean	0.20572	0.062504	
Std dev. Of FE	0.18996	0.037266	
T-statistic			5.96[a]
F-statistic			25.98[a]
Bartlett-statistic			123.60[a]

Forecast error=Absolute value of the difference between the forecasted and actual values.
The T-statistic tests the null hypothesis of equality of means without the assumption of equal population variance.
The F-statistic tests the null hypothesis of equality of variances.
The Bartlett-statistic tests the null hypothesis of homogeneity of variances.
The 1% significant level represented by a.

Table 3. In-sample rolling forecasts of changes in the stock market: OLS vs. interval results

	Monthly	Quarterly	Annual	Interval
Rolling window	60 months	20 quarters	10 years	10 years
Forecasting period	1/39-12/04	1Q/39-4Q/04	1939-2004	1939-2004
Forecast error mean(E)	0.023503	0.031501	0.065210	
Std. dev. Of E	0.021330	0.028751	0.054414	
Forecast error mean Of Upper bound(UE)				0.024819
Std. dev. Of UE				0.020353
Forecast error mean Of lower bound(LE)				0.027135
Std. dev. Of LE				0.025885
Forecast error mean Of Interval (IE=UE+LE)				0.051729

Std. dev. Of IE	0.033623
Ratio of forecasting Accuracy (FA)	0.634659
Std. dev. Of FA	0.195348

Equality tests for the following pairs

	Annual	Interval	Test result
Forecast error mean	0.065210	0.051729	
Std dev. Of FE	0.054414	0.033623	
T-statistic			1.71[c]
F-statistic			2.62[a]
Bartlett-statistic			14.41[a]

Forecast error=Absolute value of the difference between the forecasted and actual values.
The T-statistic tests the null hypothesis of equality of means without the assumption of equal population variance.
The F-statistic tests the null hypothesis of equality of variances.
The Bartlett-statistic tests the null hypothesis of homogeneity of variances.
The 1% significant level represented by a.
The 10% significant level represented by c.

Figure 1. Out-of-sample 10-year rolling OLS forecasts

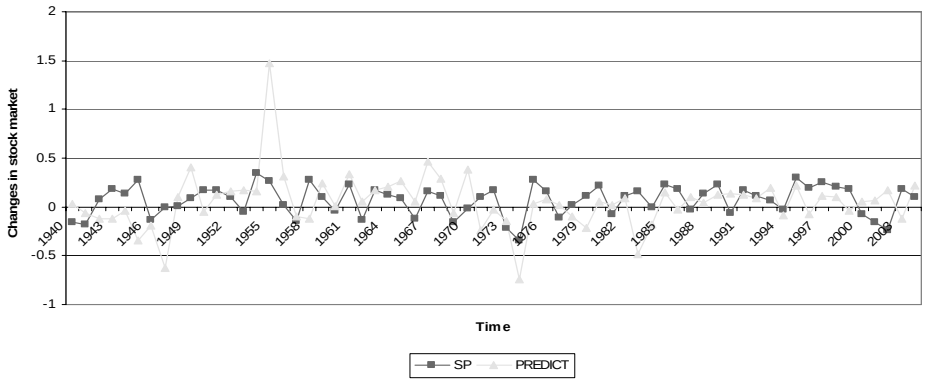
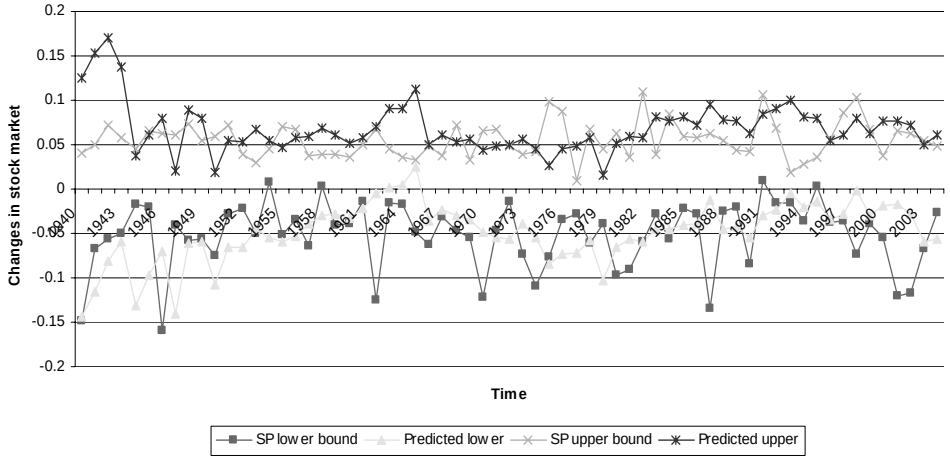


Figure 2. Out-of-sample 10-year rolling interval forecasts



Interval and Fuzzy Techniques in Business-Related Computer Security: Intrusion Detection, Privacy Protection

Mohsen Beheshti[1], Jianchao Han[1], Luc Longpré[2], Scott A. Starks[2], J. Ivan Vargas[2], Gang Xiang[2]

[1] Computer Science Dept., California State University, Dominguez Hills
Carson, CA 90747, USA, mbeheshti@csudh.edu
[2] NASA Pan-American Center for Earth and Environmental Studies
University of Texas, El Paso, TX 79968, USA, gxiang@utep.edu

Abstract. E-commerce plays an increasingly large role in business. As a result, business-related computer security becomes more and more important. In this talk, we describe how interval and fuzzy techniques can help in solving related computer security problems.

1 Interval Techniques in Computer Security: Motivations

Importance of computer security. E-commerce plays an increasingly large role in business. As a result, business-related computer security becomes more and more important; see, e.g., [2, 16].

Why interval techniques. In computer security, interval uncertainty comes from the lack of knowledge. One of the reasons for this lack of knowledge is that the users are reluctant to provide the businesses with the exact information because they do not want this information to be misused. For example, a user may be reluctant to provide his or her exact date of birth but willing to provide an age interval (e.g., 30–40).

To be successful, an electronic business needs to process the user data. It is therefore important to develop efficient algorithms for statistical processing of such interval-valued data.

2 How to Extend Statistical Techniques to Situations with Interval Uncertainty

Traditional approach to data processing: statistical analysis. One of the main objectives of computer security is to predict the user's behavior, so that we will be able to stop malicious intrusions without interfering with the legitimate use of the computer systems.

To be able to make these predictions, we must find the relation between the desired difficult-to-observe characteristics of the user behavior – such as

maliciousness – and the observable characteristics. Situations when we must be able to predict difficult-do-directly-observe characteristics based on easier-to-measure ones are typical in engineering and science. For example, in engineering, we must predict the building's stability based on the observed characteristics; in medicine, it is desirable to check whether a person has a certain disease (such as cancer) based, ideally, only on non-invasive tests such as ultrasonic and X-ray imaging.

The traditional way to find this dependence is (see, e.g., [20]):

- to find the make several observations of different characteristics,
- to compute statistical characteristics of the corresponding measurement results, such as the population mean $E_x = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$, the population variance $V_x = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E_x)^2$, the population covariance

$$C_{xy} = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E_x) \cdot (y_i - E_y),$$

  and then
- to use these statistical characteristics in the design of an (approximate) model (such as linear regression $z = a_0 + a_x \cdot x + a_y \cdot y + \ldots$) that predicts the value of the desired characteristic z as a function of easier-to-measure quantities x, y, etc.

Statistical analysis under interval uncertainty: a problem. We have mentioned that in many real-life situations, we only know the intervals $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ of possible values of x_i. For different possible values $x_i \in \mathbf{x}_i$, we get different values of E_x, V_x, C_{xy}, etc. In such situations, it is desirable to find the ranges of possible values of these characteristics:

$$\mathbf{E}_x = \left\{ \frac{1}{n} \cdot \sum_{i=1}^{n} x_i : x_i \in \mathbf{x}_i, i = 1, \ldots, n \right\},$$

$$\mathbf{V}_x = \left\{ \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E_x)^2 : x_i \in \mathbf{x}_i, i = 1, \ldots, n \right\},$$

$$\mathbf{C}_{xy} = \left\{ \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - E_x) \cdot (y_i - E_y) : x_i \in \mathbf{x}_i, y_i \in \mathbf{y}_i, i = 1, \ldots, n \right\}.$$

The practical importance of the problem of computing population variance under interval uncertainty was emphasized, e.g., in [17, 18] on the example of processing geophysical data and in [6] on the example of processing environmental data.

The problem of computing the range of a given function under interval uncertainty is known as the problem of *interval computations*; see, e.g., [11, 15]. It is known that in general, this problem is computationally difficult (NP-hard). Specifically, even computing the range of the variance is, in general, NP-hard [7, 8].

Statistical analysis under interval uncertainty: privacy case. NP-hardness means that there is no general algorithm for computing $\mathbf{V}$ in all possible cases. As shown in [12], there are practically useful cases when a feasible algorithm for computing $\mathbf{V}$ is possible. One such case is the case of privacy, when we have a fixed partition of the real line, and all intervals come from this partition. In precise terms, we fix values $x_{(1)} < x_{(2)} < \ldots < x_{(m)}$, and we are only allowing intervals of the type $[x_{(k)}, x_{(k+1)}]$ – e.g., the age of 10–20, 20–30, etc.

Statistical analysis in the privacy case: a sample algorithm. The corresponding algorithm for computing the range $[\underline{V}, \overline{V}]$ for the variance V is based on the following idea. Since no two intervals are proper subsets of one another, we can sort them in lexicographic order – i.e., in the order where $\mathbf{x}_i \le \mathbf{x}_j$ if and only if either $\underline{x}_i < \underline{x}_j$ or ($\underline{x}_i = \underline{x}_j$ and $\overline{x}_i \le \overline{x}_j$). After the sorting, it can then be shown that the maximum of V is attained at one of sequences of the type $(\underline{x}_1, \ldots, \underline{x}_k, \overline{x}_{k+1}, \ldots, \overline{x}_n)$. Thus, $\overline{V}$ is equal to the largest of the values V corresponding to $n+1$ such sequences.

For each value k, for the corresponding sequence, computing $V = M_2 - E$, where $M_2 = \dfrac{1}{n} \cdot \sum_{i=1}^{n} x_i^2$, requires linear time. Actually, we only need to spend this time to compute E, M_2, and V for $k = 0$. After we computed M_2 and E for some k, going from the case k to the case $k+1$ means replacing only one term in E and in M_2 ($\overline{x}_{k+1}$ by $\underline{x}_{k+1}$), and thus, requires constant number of steps:

$$E \to E + \frac{1}{n} \cdot (\underline{x}_{k+1} - \overline{x}_{k+1});\ M_2 \to M_2 + \frac{1}{n} \cdot ((\underline{x}_{k+1})^2 - (\overline{x}_{k+1})^2).$$

So, after sorting (which requires time $O(n \cdot \log(n))$; see, e.g., [4]), we only need linear time to compute V for all $n+1$ candidates for an optimal sequence – and after that, linear time to find the largest of these n values as $\overline{V}$. Thus, the overall computation time for computing $\overline{V}$ is equal to

$$O(n \cdot \log(n)) + O(n) = O(n \cdot \log(n)).$$

Similarly, the minimum of V is attained at one of sequences of the type $(\overline{x}_1, \ldots, \overline{x}_k, \underline{x}_{k+1}, \ldots, \underline{x}_n)$. Thus, $\underline{V}$ is equal to the smallest of the values V corresponding to $n+1$ such sequences, and we can also compute $\underline{V}$ in time $O(n \cdot \log(n))$; for details and proofs, see [1, 5, 9, 12, 23].

General case. The problem of computing the range for the variance is particular case of a general class of problems in which we need to combine probabilistic and interval uncertainty. The correspongind problems and algorithms are described, e.g., in [6, 12, 13, 21, 22].

3 Rough Set Techniques – An Alternative to Statistical Techniques

Need for alternative techniques. In the traditional statistical approach, each variable x can, in principle, take an arbitrary value from its continuous range:

e.g., from the entire real line, or from the interval [0,120] for age. In the case of privacy-type interval uncertainty, in effect, for each variable x, we have a *finite* number of possible interval values: 0–10, 10–20, etc. Therefore, instead of trying to adjust continuous techniques to this discrete case, it may make sense to design new discrete techniques for handling the corresponding data.

Rough set approach: main idea. Some such techniques have been designed as part of Pawlak's *rough set theory* – a theory that, since its appearance in the early 1980s, has been successfully used in data analysis, pattern recognition, data mining, and knowledge discovery; see, e.g., [10, 14, 19, 24].

To find the dependence between the discrete characteristics X, ..., Y, and the desired discrete characteristic Z, the rough set theory proposes the following idea. We start with the set U of objects (situations) for which we know the values $(x, \ldots, y, z)$ of all the characteristics X, ..., Y, and the value of $z \in Z$ of the characteristic Z. It may happen that the known characteristics are not sufficient to determine Z, i.e., that we have a combination of values $x \in X$, ..., $y \in Y$, of these characteristics for which there are two different patterns $(x, \ldots, y, z)$ and $(x, \ldots, y, z')$ with $z \neq z'$. It is reasonable to define the *degree of dependency* $\gamma_P(Z)$ of Z on the set $P = \{X, \ldots, Y\}$ as the ratio $\gamma_P(Z) \stackrel{\text{def}}{=} \dfrac{|CL_P(Z)|}{|U|}$, where $CL_P(Z)$ is the set of all patterns for which the values of all variables from P uniquely determine the value of Z, and $|A|$ denotes the total number of patterns in the set A.

Usually, we start by measuring a large number of different characteristics X, ..., Y, most of which may be irrelevant to the problem of predicting Z. One of our objectives is to try to come up with a smaller set of characteristics $C \subset P$ that would enable us to achieve the same predictive power. Ideally, we should find such subsets that enables to classify exactly the same number of patterns as before, i.e., for which $\gamma_C(Z) = \gamma_P(Z)$. Such subsets are called *reducts*.

Rough set approach: computational aspects. It is easy to check whether a given subset is a reduct – it is sufficient to go over all originally classifiable patterns and make sure that we do not have two patterns in which the values of all characteristics from C are the same, but the values of Z are different. However, finding an *optimal* reduct – e.g., a reduct consisting of the smallest possible number of characteristics – is known to be NP-hard. Crudely speaking, it means that the following:

- we can always find the optimal reduct by trying all $2^{|P|}$ possible subsets $C \subseteq P$ – which requires exponential time;
- NP-hard means, crudely speaking, that, in general, it is not possible to find an optimal reduct faster than in the exponential time.

In practice, when we start with a large number of characteristics, we cannot test all subsets, so we have to use heuristic algorithms that provide us with a possible sub-optimal ("good enough") reduct instead of the optimal one. Several such algorithms have been developed and successfully used in applications of

rough set theory. In particular, such heuristics have been successfully used for intrusion detection [3].

Limitations of the interval approach to rough set classification. In the above approach, we implicitly assumed that we know exactly how to classify each pattern, and we are absolutely sure that all the information within each pattern is correct.

In practice, however, some patterns may be more dubious, and some cases of seemingly legitimate computer behavior may actually be intrusions. How can we take this uncertainty into account?

4 Towards a Combination of Fuzzy and Interval Techniques in Computer Security

Why fuzzy techniques in computer security. Usually, a human being can easily recognize a spam. However, in some practical situations, a user may not be sure whether a given message is a spam or a legitimate email. The best a user can do in such a situation is to say that a given email is most probably, legitimate, or that it is somewhat suspicious – but the user usually cannot provide us with a probability or other numerical characteristic of his or her degree of suspicion.

Similarly, system administrators can often detect intrusion, but frequently, they can only indicate that something suspicious is going on, without being 100% sure that what they observing is indeed a computer intrusion. A system administrator usually cannot describe this feeling in numerical terms; at best, he or she can tell that a given situation is somewhat suspicious, or very suspicious, or most probably legitimate.

Fuzzy logic was designed to describe these and similar qualitative terms from natural language. In the fuzzy logic, each term like "somewhat" is translated into a numerical degree – a number from the interval $[0, 1]$. Let us show how this approach can be used in computer security.

In this paper, we will consider the most important practical case when the predicted variable Z has two possible values "positive" and "negative" – such as "intrusion" or "no intrusion".

How to incorporate fuzzy techniques into our problem. In accordance with the above comment, for each object $u \in U$, we have a degree $d(u)$ to which this object is relevant and its data is reliable. Also, instead of the exact value $z \in Z$, we have the degree $g(u)$ to which this object u is a positive example.

In other words, we have the list of tuples each of which has the form $(x, \ldots, y, g(u), d(u))$, where:

- $x \in X$, $\ldots$, $y \in Y$, are the (discrete) values of the corresponding characteristics,
- $g(u)$ is the degree to which this object is a positive example, and
- $d(u)$ is the degree of importance of u.

From the traditional fuzzy logic to interval-valued fuzzy logic. It is worth mentioning that we may want to distinguish between situations in which we have an equal number of arguments for and against positivity of a given object u, and situations when we know nothing about u. In the traditional fuzzy logic, both situations are characterized by the value $g(u) = 0.5$. To distinguish between such situations, it is reasonable to use *intuitionistic fuzzy logic* in which we have tow different degrees: the degree $g^+(u)$ to which u is positive and the degree $g^-(u)$ to which u is negative.

- In the case of equal number of arguments, we have $g^+(u) = g^-(u) = 0.5$;
- in the case of complete ignorance, we have $g^+(u) = g^-(u) = 0$;
- in general, we have $g^+(u) + g^-(u) \le 1$.

In this approach, our degree of belief that an object u is positive is no longer described by a single value $g(u)$. Actually, the only information that we have about this value is that it is somewhere between $g^+(u)$ and $1 - g^-(u)$. So:

- in the case of equal number of arguments, we have $g^+(u) = g^-(u) = 0.5$ and thus, the interval $[g^+(u), 1 - g^-(u)]$ consists of a single value 0.5;
- in the case of complete ignorance, we have $g^+(u) = g^-(u) = 0$, so the interval $[g^+(u), 1 - g^-(u)]$ coincides with the entire interval $[0, 1]$;
- in general, we have an interval $[g^+(u), 1 - g^-(u)] \subseteq [0, 1]$.

How to incorporate fuzzy techniques into the rough set approach. In the traditional rough set approach, any rule with one counter-example is immediately dismissed as invalid. Since we now allow the possibility that tuples may be described wrong or misclassified, we are no longer sure that this counter-example actually disproves the rule.

In such situation, for each rule, we can only talk about the degree to which it is consistent with the given data. Let us assume that we restrict ourselves to a subset $C = \{X, \ldots, Y\} \subseteq P$ of the original set of characteristics. Each combination of values $c = (x, \ldots, y)$ of characteristics from C can be then classified as either positive or negative. Let S_c denote all the tuples with this combination of values. Then, it is reasonable to define the certainty to classify S_c as a positive region as $C_c^+ = \dfrac{\sum\limits_{u \in S_c} d(u) \cdot g^+(u)}{\sum\limits_{u \in S_c} d(u)}$, and the certainty to classify S_c as a negative region as $C_c^- = \dfrac{\sum\limits_{u \in S_c} d(u) \cdot g^-(u)}{\sum\limits_{u \in S_c} d(u)}$. We then:

- classify c as positive if $C_c^+ \ge t^+$ for some pre-defined threshold t^+,
- classify c as negative if $C_c^- \ge t^-$ for some pre-defined threshold t^-.

The set of all classified c will be denoted by CL. For all combinations $c \notin CL$, the information presented in c is not sufficient for the classification.

We can then define the degree of dependency $\gamma_C(Z)$ as the relative weight of all classified tuples, i.e., as the ratio

$$\gamma_C(Z) \stackrel{\text{def}}{=} \frac{\sum\limits_{c(u)\in CL} d(u)}{\sum\limits_{u\in U} d(u)}.$$

A reduct is then defined as a subset $C \subseteq P$ for which $\gamma_C(Z) = \gamma_P(Z)$.

Preliminary results. The existing heuristics for finding sub-optimal reducts can be naturally extended to this fuzzy situation, and our preliminary results of using these heuristics are encouraging.

5 Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

References

1. Aló, R., Beheshti, M., Xiang, G.: Computing Variance under Interval Uncertainty: A New Algorithm and Its Potential Application to Privacy in Statistical Databases. In: *Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU'06*, Paris, France, July 2–7, 2006 (to appear).
2. Bace, R. G.: *Intrusion Detection*, Sams Publishing, 1999.
3. Cai, Z., Guan, X., Shao, P., Peng, A., Sun, G.: A rough set theory based method for anomaly intrusion detection in computer network systems. Expert Systems **20** (2003) 251–260.
4. Cormen, Th. H., Leiserson, C. E., Rivest, R. L., Stein, S.: *Introduction to Algorithms*, MIT Press, Cambridge, MA, 2001.
5. Dantsin, E. Kreinovich, V. Wolpert, A. Xiang, G.: Population Variance under Interval Uncertainty: A New Algorithm. Reliable Computing **12** (2006) 273–280.
6. Ferson S.: *RAMAS Risk Calc 4.0: Risk Assessment with Uncertain Numbers*, CRC Press, Boca Raton, Florida, 2002.
7. Ferson, S., Ginzburg, L., Kreinovich, V., Longpré, L., Aviles, M.: Computing Variance for Interval Data is NP-Hard. ACM SIGACT News **33** (2002) 108–118.
8. Ferson S., Ginzburg, L., Kreinovich, V., Longpré, L., Aviles, M.: Exact Bounds on Finite Populations of Interval Data", Reliable Computing **11** (2005) 207–233.
9. Granvilliers, L., Kreinovich, V., Müller, N.: Novel Approaches to Numerical Software with Result Verfication. In: Alt, R., Frommer, A., Kearfott, R. B., Luther, W., Eds., *Numerical Software with Result Verification*, Proceedings of the International Dagstuhl Seminar, Dagstuhl Castle, Germany, January 19–24, 2003, Springer Lectures Notes in Computer Science, 2004, Vol. 2991, pp. 274–305.

10. Han, J., Hu, X., Cercone, N.: Supervised Learning: A Generalized Rough Set Approach. In: *Proc. of International Conference on Rough Sets and Current Trend in Computing*, Banff, Canada, 2000, pp. 284–291.
11. Kearfott, R. B., Kreinovich, V. Eds.: *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.
12. Kreinovich, V., Xiang, G., Starks, S. A., Longpré, L., Ceberio, M., Araiza, R., Beck, J., Kandathi, R., Nayak, A., Torres, R., Hajagos, J.: Towards combining probabilistic and interval uncertainty in engineering calculations: algorithms for computing statistics under interval uncertainty, and their computational complexity. *Reliable Computing* (to appear).
13. Kuznetsov, V. P.: *Interval statistical models*, Moscow, Radio i Svyaz Publ., 1991 (in Russian).
14. Lin, T. Y., Yao, Y. Y., Zadeh, L. A.: *Data Mining, Rough Sets and Granular Computing*, Physica-Verlag, 2002.
15. Moore, R. E.: *Methods and Applications of Interval Analysis*, SIAM, Philadelphia, 1979.
16. Ning, P.: *Intrusion Detection in Distributed Systems: An Abstraction-Based Approach*, Springer, 2003.
17. Nivlet, P., Fournier, F., Royer, J.: A new methodology to account for uncertainties in 4-D seismic interpretation. In: *Proceedings of the 71st Annual International Meeting of the Society of Exploratory Geophysics SEG'2001*, San Antonio, Texas, September 9–14, 2001, pp. 1644–1647.
18. Nivlet, P., Fournier, F., Royer, J.: Propagating interval uncertainties in supervised pattern recognition for reservoir characterization", *Proceedings of the 2001 Society of Petroleum Engineers Annual Conference SPE'2001*, New Orleans, Louisiana, September 30 – October 3, 2001, paper SPE-71327.
19. Pawlak, Z.: *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, 1991.
20. Rabinovich, S.: *Measurement Errors: Theory and Practice*, American Institute of Physics, New York, 2005.
21. Walley, P.: *Statistical reasoning with imprecise probabilities*, Chapman and Hall, N.Y., 1991.
22. Walster, G. W.: Philosophy and practicalities of interval arithmetic, In: Moore, R. E., Ed.: *Reliability in Computing*, 1988, pp. 307–323.
23. Xiang, G.: Fast algorithm for computing the upper endpoint of sample variance for interval data: case of sufficiently accurate measurements. Reliable Computing **12** (2006) 59–64.
24. Ziarko, W.: Variable Precision Rough Set Model. Journal of Computer and System Sciences **46**, No. 1 (1993).

Growth Rates under Interval Uncertainty

Janos Hajagos[1] and Vladik Kreinovich[2]

[1] Applied Biomathematics,
100 North Country Road
Setauket, New York 11733, USA
logistic@sdf-eu.org
[2] Department of Computer Science,
University of Texas at El Paso,
El Paso, TX 79968, USA
vladik@utep.edu
http://www.cs.utep.edu/vladik

Abstract. For many real-life systems ranging from financial to population-related to medical, dynamics is described by a system of linear equations. For such systems, the growth rate λ can be determined as the largest eigenvalue of the corresponding matrix a_{ij}. In many practical situations, we only know the components of the matrix a_{ij} with interval (or fuzzy) uncertainty. In such situations, it is desirable to find the range of possible values of λ. In this paper, we propose an efficient algorithm for computing λ for a practically important case when all the components a_{ij} of the matrix are non-negative.

1 Growth Rates: A Linear Approximation to the Description of a General System

General description. In general, the state of a real-life complex system can be described by listing the current values of its parameters $x_1, \ldots, x_n$.

- For continuous-time systems, their dynamic can be described as

$$\dot{x}_i = f_i(x_1, \ldots, x_n)$$

  for some functions $f_1, \ldots, f_n$.
- For discrete-time systems, their dynamic can be described as

$$x_i(t+1) = f_i(x_1(t), \ldots, x_n(t))$$

  for some functions $f_1, \ldots, f_n$.

Linearized description. The dependencies $f_i(x_1, \ldots, x_n)$ are usually smooth, so within a reasonable range of values x_i, we can approximate each of these functions by a linear expression:

$$f_i(x_1, \ldots, x_n) = b_i + \sum_{j=1}^{n} a_{ij} \cdot x_j.$$

By applying an appropriate shift $x_i \to x_i - s_i$, we can simplify this system even further, into

$$f_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} a_{ij} \cdot x_j.$$

Thus:

- the dynamic of a continuous-time system can be described by the equation

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij} \cdot x_j;$$

- the dynamic of a discrete-time systems can be described by the equation

$$x_i(t+1) = \sum_{j=1}^{n} a_{ij} \cdot x_j(t).$$

The notion of a growth rate. By using eigenvectors of the matrix $A = (a_{ij})$ as a new base, we get a yet simpler expression for the new variables y_i – the coefficients in the expansion of $x_i(t)$ in this new base.

In the generic case when all eigenvalues are different, the dynamic equations take the simplest possible form:

- for a continuous-time system,

$$\dot{y}_i = \lambda_i \cdot y_i,$$

  where λ_i is the corresponding eigenvalue;
- for a discrete-time system system,

$$y_i(t+1) = \lambda_i \cdot y_i(t).$$

These equations have an explicit solution:

- for a continuous-time system, we get

$$y_i(t) = y_i(0) \cdot \exp(\lambda_i \cdot t);$$

- for a discrete-time system, we get

$$y_i(t) = y_i(0) \cdot \lambda_i^t.$$

This decomposition into simple solutions $y_i(t)$ is one of the main ideas behind the Principle Component Analysis.

A general solution $x_i(t)$ is a linear combination of such terms. Thus, in the general case, asymptotically,

– for a continuous-time system,

$$x_i(t) \sim \exp(\lambda \cdot t),$$

where λ is the largest of these eigenvalues;
– for a discrete-time system,

$$x_i(t) \sim \lambda^t.$$

When the largest eigenvalue is degenerate, we have $x_i(t) \sim x^k \cdot \exp(\lambda \cdot t)$ or $x_i(t) \sim x_k \cdot \lambda^t$ for some integer k, i.e., modulo polynomial terms, still the same asymptotic.

Because of this fact, the largest eigenvalue λ is called the *growth rate* of a system.

This formula indeed describes a growth rate. This asymptotic behavior well describes different types of growth (see, e.g., [3]):

– the population growth,
– the growth in animals and plants,
– the growth rates of number of affected people under an epidemic,
– financial growth,
– etc.

For example, in population growth, different variables x_i describe the number of people of i-th age group, etc.

2 Growth under Interval Uncertainty: A Computational Problem

Computing the growth rate is important. In view of the above applications, it is important to compute the growth rate for a given system.

Idealized case: exactly known coefficients. In general, we never known the *exact* values of parameters of real-life systems, these parameters are always known with some *uncertainty*.

In many real-life situations, however, this uncertainty is small. In such situations, we can safely assume that we know the exact values a_{ij} of all the coefficients. In such situations, we can use known algorithms to find the eigenvalues [3, 7, 14] and thus, find the largest of these eigenvalues – the growth rate.

Often, we cannot ignore the uncertainty. In many real-life situations, however, we cannot ignore the uncertainty. In such situations, we have to take into account the fact that the coefficients a_{ij} are only known with uncertainty.

Case of interval uncertainty. Often, in addition to the approximate values $\widetilde{a}_{ij}$ of the corresponding coefficients, we also know the upper bounds Δ_{ij} on the approximation error $|\widetilde{a}_{ij} - a_{ij}|$. In such situations, we know that the actual (unknown) value of each coefficient a_{ij} belongs to the interval

$$\mathbf{a}_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}] \stackrel{\text{def}}{=} [\widetilde{a}_{ij} - \Delta_{ij}, \widetilde{a}_{ij} + \Delta_{ij}].$$

Computing the growth rate under interval uncertainty: a computational problem. In case of interval uncertainty, different values $a_{ij} \in \mathbf{a}_{ij}$ lead to different growth rates λ. In such situations, it is desirable to find the interval $[\underline{\lambda}, \overline{\lambda}]$ of possible values of λ – or at least an interval that guarantees to contain this interval.

Of special importance is the upper endpoint $\overline{\lambda}$ of the desired interval, because this upper endpoint indicates how fast a population can grow, or how fast a disease can spread.

The need for computing such an interval has been known for a few decades; see, e.g., [2].

Case of small uncertainty: sensitivity analysis. When the uncertainty is relatively small, i.e., when the uncertainty $\Delta a_{ij} \stackrel{\text{def}}{=} \widetilde{a}_{ij} - a_{ij}$ is much smaller than the approximate value $\widetilde{a}_{ij}$, we can linearize the equations for describing the eigenvalues in terms of a_{ij} and use the sensitivity analysis techniques to get reasonable estimates for $[\underline{\lambda}, \overline{\lambda}]$; see, e.g., [2].

General case: the problem is computationally intractable (NP-hard). In many real-life situations, e.g., in many financial and biological systems, the uncertainty is not small, so we can no longer use the linearized techniques to find $\overline{\lambda}$.

In general, we this face a problem of finding the range of possible values of λ for all matrices a_{ij} within a given interval matrix $\mathbf{a}_{ij}$, i.e., for all matrices for which $a_{ij} \in \mathbf{a}_{ij}$. It is known that in general, this problem is NP-hard; see [9] and references therein. This means, crudely speaking, that it is not possible top have an algorithm that would always compute the desired range for λ in physically reasonable time.

Moreover, it is also known [9] that even the problem of computing the eigenvalues with a given accuracy is NP-hard.

This means that while there exist computationally efficient methods of computing an enclosure for the desired interval $[\underline{\lambda}, \overline{\lambda}]$, but these methods sometimes lead to a drastic excess width.

Important case: a non-negative matrix. In many real-life situations, the matrix a_{ij} is *non-negative* in the sense that all its coefficients are non-negative. Such non-negative matrices describe population growth, spread of disease, financial situations, etc. [3].

What we propose. We propose a new algorithm that, for non-negative matrices, exactly computes the upper bound $\overline{\lambda}$ on λ in feasible computation time.

Comment. In this paper, we concentrated on the computation of the largest eigenvalue, a practically useful characteristic of an interval matrix. If, in addition to describing the asymptotic growth rate, we want to find a more detailed description of a growth, then we need to find not only the largest eigenvalue, but also other eigenvalues and the corresponding eigenvectors. Algorithms for solving this problem under interval uncertainty are presented, e.g., in [10].

3 New Algorithm

This algorithm is based on known algorithms for the case of the exact matrix. Our algorithm assumes that already have an algorithm $\mathcal{A}$ for computing the largest eigenvalue $\lambda(A)$ of a given non-negative matrix $A = \|a_{ij}\|$. Such algorithms are described, e.g., in [3, 7, 14].

Input to the new algorithm. Let us assume that instead of the exact non-negative matrix $A = \|a_{ij}\|$, we are given the interval-valued matrix $\mathbf{A} = \|\mathbf{a}_{ij}\|$, where $\mathbf{a}_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}]$.

In other words, for each i and each j, instead of a single value a_{ij}, we have an interval $[\underline{a}_{ij}, \overline{a}_{ij}]$ – i.e., in effect, two values $\underline{a}_{ij}$ and $\overline{a}_{ij}$. We can alternatively describe this situation by saying that:

- for each i and j, we know the lower endpoint $\underline{a}_{ij}$, and
- for each i and j, we know the upper endpoint $\overline{a}_{ij}$.

In other words, instead of a single matrix $A = \|a_{ij}\|$, we are given two matrices: $\underline{A} \stackrel{\text{def}}{=} \|\underline{a}_{ij}\|$ and $\overline{A} \stackrel{\text{def}}{=} \|\overline{a}_{ij}\|$.

Description of the new algorithm.

- First, we apply the algorithm $\mathcal{A}$ for computing $\lambda(A)$ to the matrix $\underline{A}$; the resulting value is returned as $\underline{\lambda}$.
- Then, we apply the algorithm $\mathcal{A}$ for computing $\lambda(A)$ to the matrix $\overline{A}$; the resulting value is returned as $\overline{\lambda}$.

Comment. This simple idea does not work for general interval matrices, only for non-negative ones. In the following text, we provide a proof that for non-negative matrices, this algorithm indeed works well.

Justification of the new algorithm. In order to provide the desired justification, let us introduce some notations. We have defined a matrix A to be non-negative if all its components are non-negative, i.e., as $a_{ij} \geq 0$ for all i and j. It is natural to denote this non-negativity in the usual way, as $A \geq 0$.

We can similarly define a vector $x = (x_1, \ldots, x_n)$ to be non-negative if all its components are non-negative, i.e., if $x_i \geq 0$ for all i. This relation will also be denoted by $x \geq 0$.

We can also define the relation $A \le B$ for two matrices A and B as $B - A \ge 0$. In other words, for matrices $A = \|a_{ij}\|$ and $B = \|b_{ij}\|$, the order $A \le B$ is defined component-wise, as $a_{ij} \le b_{ij}$ for all i and j.

Now, we are ready for the justification. This justification is based on the two known facts (described below in detail):

- the representation of the largest eigenvalue as a solution to an auxiliary optimization problem, and
- the Perron-Frobenius Theorem about the eigenvectors of non-negative matrices.

The first known fact is that the largest eigenvalue $\lambda(A)$ of a matrix A can be described as

$$\lambda(A) = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2},$$

where

$$\|x\|_2 \stackrel{\text{def}}{=} \sqrt{x_1^2 + \ldots + x_n^2}$$

denotes the length of a vector x; see, e.g., [3, 7, 14].

For non-negative matrices A, the Perron-Frobenius Theorem [3, 4, 7, 14] states that at least one of the eigenvectors $x = (x_1, \ldots, x_n)$ corresponding to the largest eigenvalue $\lambda(A)$ is also non-negative: $x \ge 0$. Thus, the maximum in the above definition of $\lambda(A)$ is attained on a non-negative vector. Therefore, when computing this maximum, we can restrict ourselves only to non-negative vectors:

$$\lambda(A) = \max_{x \ge 0 \,\&\, x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}.$$

When $0 \le A$ and $x \ge 0$, then, by definition of the matrix multiplication, we get $Ax \ge 0$. If $0 \le A \le B$ and $x \ge 0$, then similarly $0 \le Ax \le Bx$. When we have two vectors a and b for which $0 \le a_i \le b_i$ for all i, then, of course,

$$a_1^2 + \ldots + a_n^2 \le b_1^2 + \ldots + b_n^2,$$

hence $\|a\|_2 \le \|b\|_2$. Thus, if $0 \le A \le B$, then for every vector $x \ge 0$, we get $\|Ax\|_2 \le \|Bx\|_2$ and therefore,

$$\frac{\|Ax\|_2}{\|x\|_2} \le \frac{\|Bx\|_2}{\|x\|_2}.$$

Since this inequality holds for every vector $x \neq 0$, the maximum $\lambda(A)$ of its left-hand side is smaller than or equal than the maximum $\lambda(B)$ of its right-hand side. In other words, if $0 \le A \le B$, then $\lambda(A) \le \lambda(B)$.

By definition of an interval-valued matrix, all possible matrices $A \in [\underline{A}, \overline{A}]$ satisfy the inequality $\underline{A} \le A \le \overline{A}$. Since we assumed that our matrices are non-negative, we conclude that $0 \le \underline{A} \le A \le \overline{A}$. Thus, for every possible matrix $A \in [\underline{A}, \overline{A}]$, we get $\lambda(\underline{A}) \le \lambda(A) \le \lambda(\overline{A})$.

Hence, all the values $\lambda(A)$ lie within the interval $[\lambda(\underline{A}), \lambda(\overline{A})]$. Since both endpoints of this interval are attained for some matrices from the matrix interval $[\underline{A}, \overline{A}]$, we thus conclude that the interval $[\lambda(\underline{A}), \lambda(\overline{A})]$ is the actual range of $\lambda(A)$. Thus, for non-negative interval matrices, the above algorithm is indeed justified.

Practical applications. In [5, 6], we apply our ideas to the dynamics of real-life ecological systems for which we only know the components a_{ij} with interval uncertainty.

4 Case of Fuzzy Uncertainty

Case of fuzzy uncertainty. Often, knowledge comes in terms of uncertain expert estimates. To describe this uncertainty, for each possible value of a_{ij}, we describe the degree $\mu_{ij}(a_{ij}) \in [0, 1]$ to which this value is possible. These degrees form a *fuzzy set.*

Processing fuzzy uncertainty can be reduced to processing interval uncertainty. For each degree of certainty α, we can determine the set of values of a_{ij} that are possible with at least this degree of certainty – the *α-cut*

$$\{a_{ij} \,|\, \mu_{ij}(a_{ij}) \geq \alpha\}$$

of the original fuzzy set. In most cases, this α-cut is an interval.

Vice versa, if we know α-cuts for every α, then, for each value a_{ij}, we can determine the degree of possibility that a_{ij} belongs to the original fuzzy set [1, 8, 11–13]. A fuzzy set can be thus viewed as a nested family of its α-cuts.

So, if instead of a (crisp) interval $\mathbf{a}_{ij}$ of possible values of the component a_{ij}, we have a fuzzy set $\mu_{ij}(a_{ij})$ of possible values, then we can view this information as a family of nested interval matrices $\mathbf{a}_{ij}(\alpha)$ – α-cuts of the given fuzzy-valued matrices.

Computation under fuzzy uncertainty. Let us consider the case when instead of a (crisp) interval $\mathbf{a}_{ij}$ of possible values of the components, we have a fuzzy set $\mu_{ij}(a_{ij})$ of possible values. In this case, we can view this information as a family of nested interval-valued matrices $\mathbf{a}_{ij}(\alpha)$ – α-cuts of the given fuzzy sets.

Our objective is then to compute the fuzzy number λ – the largest eigenvalue of this fuzzy-valued matrix. In this case, for each level α, to compute the α-cut of this fuzzy number, we can apply interval computations to the α-cuts $\mathbf{a}_{ij}(\alpha)$ of the corresponding fuzzy-valued matrix. The resulting nested intervals form the desired fuzzy set for λ.

So, e.g., if we want to describe 10 different levels of uncertainty, then we must solve 10 interval computation problems – i.e., apply the above algorithm 10 times.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

The authors are thankful to Arnold Neumaier for his valuable advise.

References

1. Bojadziev G., Bojadziev, M.: *Fuzzy Sets, Fuzzy Logic, Applications*, World Scientific, Singapore, 1995.
2. Caswell, H. A.: General formula for the sensitivity of population growth rate to changes in life history parameters. Theoretical Population Biology **14** (1978) 215–230.
3. Caswell, H.: *Matrix Population Models: Construction, Analysis, and Interpretation*, Sinauer Publ., Sunderland, Massachusetts, 2001.
4. Ding, J., Fay, T. H.: The Perron-Frobenius theorem and limits in geometry. American Mathematical Monthly **112**, No. 2 (2005) 171–175.
5. Hajagos, J. G.: Low-dose extrapolation models for reliable human health assessment. In: *Proceedings of the Workshop on Reliable Engineering Computing*, Savannah, Georgia, September 15–17, 2004.
6. Hajagos, J. G.: *Accurately Computing Ecological Risk under Measurement Uncertainty*, Ph.D. Dissertation, Department of Ecology and Evolution, State University of New York at Stony Brook, June 2005.
7. Horn, R. A., Johnson, C. R.: *Matrix Analysis*, Cambridge University Press, Cambridge, 1994.
8. Klir, G., Yuan, B.: *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, New Jersey, 1995.
9. Kreinovich, V., Lakeyev, A., Rohn, J., Kahl, P.: *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1998.
10. Lauro, C. N., Palumbo, F.: Principal componwent analysis of interval data: a symbolic data analysis approach. Computational Statistics **15** (2000) 73–87.
11. Moore, R. E., Lodwick, W. A.: Interval analysis and fuzzy set theory. Fuzzy Sets and Systems **135** (2003) 5–9.
12. Nguyen, H. T., Kreinovich, V.: Nested intervals and sets: concepts, relations to fuzzy sets, and applications, In: Kearfott, R. B., Kreinovich, V., Eds.: *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.
13. Nguyen, H. T., Walker, E. A.: *First Course in Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.
14. Varga, R. S.: *Matrix Iterative Analysis*, Springer Verlag, Berlin, 2000.

Economics of Engineering Design under Interval (and Fuzzy) Uncertainty: Case Study of Building Design

Carlos Ferregut[1], Jan Beck[2], Araceli Sanchez[1], and Vladik Kreinovich[1]

[1] University of Texas, El Paso, TX 79968, USA, vladik@utep.edu
[2] X-L Synergy, 2000 Wyoming Ave., El Paso, TX 79903, USA, janbeck@gmail.com

Abstract. One of the main objectives of engineering design is to find a design that is the cheapest among all designs that satisfy given constraints. Most of the constraints must be satisfied under all possible values within certain ranges. Checking all possible combinations of values is often very time-consuming. In this paper, we propose a faster algorithm for checking such constraints.

1 Formulation of the Problem

General problem of engineering design. In engineering design, e.g., when designing a building, the problem is usually to find the cheapest design among all designs that satisfy given constraints. This is a problem behind the usual bidding process when the constraints (requirements) are announced beforehand, and the contract goes to the lowest bidder among those whose design satisfies the constraints.

Checking constraints is often difficult. Most design constraints require that some condition is satisfied for all possible values of parameters within given ranges. For example, a levee must withstand all Category 4 hurricanes, i.e., all the winds within a certain range. A building must remain stable under all possible distribution of loads in different rooms on different floors.

There are many possible distributions, so checking all of them is not practically possible. It is therefore desirable to come up with efficient algorithms for checking such constraints.

General mathematical description of the problem. Let us denote the number of parameters by n, the i-th parameter by x_i, and the range of possible values of the i-th parameter by $\mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$. For example, for the building stability, x_i is the load in the i-th room.

The desired constraint can be described as a restriction on some characteristics depending on these parameters x_i: e.g., that at every point, the strain does not exceed the critical value. Each constraint of this type can be described as $f(x_1, \ldots, x_n) \le f_0$, where the (known) function $f(x_1, \ldots, x_n)$ describes the

dependence of the critical characteristic (like strain) on the values x_i, and f_0 is the threshold value of this characteristic.

The problem is: given the intervals $\mathbf{x}_i$ and the function $f(x_1, \ldots, x_n)$, we want to check whether $f(_1, \ldots, x_n) \le f_0$ for all combinations of possible values $x_i \in \mathbf{x}_i$ of the corresponding parameters.

What we do. In this paper, we describe a new faster algorithm for checking this constraint.

2 Existing Techniques for Solving The Problem

Relation to interval computations. A natural way to check the desired condition is to compute the range $\mathbf{y} = [\underline{y}, \overline{y}]$ of the function f over the box $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$:

$$\mathbf{y} = [\underline{y}, \overline{y}] = \{f(x_1, \ldots, x_n) \,|\, x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\},$$

and then to check whether the upper endpoint $\overline{y}$ of this range exceeds the threshold f_0.

The process of computing this interval range based on the input intervals $\mathbf{x}_i$ is called *interval computations*; see, e.g., [3, 4].

Nominal values. It is reasonable to take the midpoint $\widetilde{x}_i = \dfrac{\underline{x}_i + \overline{x}_i}{2}$ of the range as the nominal value, and treat the difference $\Delta x_i \stackrel{\text{def}}{=} x_i - \widetilde{x}_i$ as the deviations from the nominal value. The value of f for the nominal values will be denoted by $\widetilde{y} \stackrel{\text{def}}{=} f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

Linearized case. In this paper, we will only consider situations in which the ranges are narrow enough. If the ranges are narrow, then the terms which are quadratic (or of higher order) in Δx_i can be safely neglected. In such situations, the dependence of the desired value

$$y = f(x_1, \ldots, x_n) = f(\widetilde{x}_1 + \Delta x_1, \ldots, \widetilde{x}_n + \Delta x_n)$$

on Δx_i can be safely assumed to be linear.

Comment. There are practical situations when the ranges are not narrow enough, and hence, quadratic terms cannot be safely neglected. In this case, the problem of error estimation for indirect measurements becomes computationally difficult (NP-hard) even when the function $f(x_1, \ldots, x_n)$ is quadratic [7, 10].

When the differences are small, we can simplify the expression for $\Delta y \stackrel{\text{def}}{=} y - \widetilde{y} = f(x_1, \ldots, x_n) - f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ if we expand the function f in Taylor series around the point $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ and restrict ourselves only to linear terms in this expansion. As a result, we get the expression

$$\Delta y = c_1 \cdot \Delta x_1 + \ldots + c_n \cdot \Delta x_n, \tag{1}$$

where by c_i, we denoted the value of the partial derivative $\partial f/\partial x_i$ at the point $(\widetilde{x}_1, \ldots, \widetilde{x}_n)$:

$$c_i = \frac{\partial f}{\partial x_i}|_{(\widetilde{x}_1, \ldots, \widetilde{x}_n)}. \tag{2}$$

The sum (1) attains its largest possible value if each term $c_i \cdot \Delta x_i$ in this sum attains the largest possible value:

- If $c_i \geq 0$, then this term is a monotonically non-decreasing function of Δx_i, so it attains its largest value at the largest possible value $\Delta x_i = \Delta_i$; the corresponding largest value of this term is $c_i \cdot \Delta_i$.
- If $c_i < 0$, then this term is a decreasing function of Δx_i, so it attains its largest value at the smallest possible value $\Delta x_i = -\Delta_i$; the corresponding largest value of this term is $-c_i \cdot \Delta_i = |c_i| \cdot \Delta_i$.

In both cases, the largest possible value of this term is $|c_i| \cdot \Delta_i$, so, the largest possible value of the sum Δy is

$$\Delta = |c_1| \cdot \Delta_1 + \ldots + |c_n| \cdot \Delta_n. \tag{3}$$

Similarly, the smallest possible value of Δy is $-\Delta$.

Hence, the interval of possible values of Δy is $[-\Delta, \Delta]$, with Δ defined by the formula (3), and the range of y is equal to $[\widetilde{y} - \Delta, \widetilde{y} + \Delta]$.

A precise computational formulation of the problem. As a result of the above analysis, we get the following explicit formulation of the problem: given a function $f(x_1, \ldots, x_n)$, n numbers $\widetilde{x}_1, \ldots, \widetilde{x}_n$, and n positive numbers $\Delta_1, \ldots, \Delta_n$, compute the corresponding expression (3).

Let us describe how this problem is solved now.

Textbook case: the function f is given by its analytical expression. If the function f is given by its analytical expression, then we can simply explicitly differentiate it, and get an explicit expression for (3).

A more complex case: automatic differentiation. In many practical cases, we do not have an explicit analytical expression, we only have an *algorithm* for computing the function $f(x_1, \ldots, x_n)$, an algorithm which is too complicated to be expressed as an analytical expression.

When this algorithm is presented in one of the standard programming languages such as Fortran or C, we can let the compute perform an explicit differentiation; for that, we can use one of the existing automatic differentiation tools (see, e.g., [1, 2]). These tools analyze the code of the program for computing $f(x_1, \ldots, x_n)$ and, as they perform their analysis, they produce the "differentiation code", i.e., a program that computes the partial derivatives c_i.

Once we know an algorithm that computes f in time T, automatic differentiation (AD) enables us to compute all partial derivatives in time $\leq 3T$, hence we can compute Δ in time $O(T + n)$.

Case of proprietary software. The software that computes the value y of the desired characteristic is often proprietary: the owners of the program f do not want to disclose its code; instead, they only allow to use f as a black box.

If we do not know the code of f, then we cannot apply AD to compute all n partial derivatives $c_i = \dfrac{\partial f}{\partial x_i}$.

A straightforward method of solving this problem: numerical differentiation. The most straightforward algorithm for solving this problem is to compute the derivatives c_i one-by-one, and then use the corresponding formula (3) to compute the desired Δ. To compute the i-th partial derivative, we change the i-th input x_i to $\widetilde{x}_i + h_i$ for some h_i, and leave other inputs unchanged, i.e., we take $\delta_i = h_i$ for this i and $\delta_j = 0$ for all $j \neq i$. Then, we estimate c_i as

$$c_i = \frac{f\left(\widetilde{x}_1, \ldots, \widetilde{x}_{i-1}, \widetilde{x}_i + h_i, \widetilde{x}_{i+1}, \ldots, \widetilde{x}_n\right) - \widetilde{y}}{h_i}. \tag{4}$$

This algorithm is called *numerical differentiation.*

We want the change h_i to be small (so that quadratic terms can be neglected); we already know that changes of the order Δ_i are small. So, it is natural to take $h_i = \Delta_i$. In other words, to compute c_i, we use the following values: $\delta_1 = \ldots = \delta_{i-1} = 0$, $\delta_i = \Delta_i$, $\delta_{i+1} = \ldots = \delta_n = 0$.

Problem: sometimes, numerical differentiation takes too long. Very often, the program f requires a reasonable time to compute. In this case, applying the function f is the most time-consuming part of this algorithm. So, the total time that it takes us to compute Δ is (approximately) equal to the running time T for the program f multiplied by the number of times N_f that we call the program f.

For numerical differentiation, $N_f = n$ (we call f n times to compute n partial derivatives). Hence, if the program f takes a long time to compute, and n is huge, then the resulting time $T \cdot n$ (which is $\gg T + n$) may be too long. For example, for different loads in a multi-story building, we may get n in the hundreds, and T in minutes. In this case, $T \cdot n$ may take several days. This may be OK for a single design, but too long if we want to compare different designs in order to select the optimal one – i.e., the cheapest of all the designs that satisfy all the constraints.

3 A New Method Based on Cauchy Distribution

Can we use Monte-Carlo simulations in the interval setting? It is well known that for probabilistic uncertainty, Monte-Carlo simulation speeds up computations. It is desirable to use a similar technique in interval setting as well.

There is a problem here. In the interval setting, we do not know the exact distribution, we may have different probability distributions – as long as they

are located within the corresponding intervals. If we only use one of these distributions for simulations, there is no guarantee that the results will be valid for other distributions as well.

In principle, we could repeat simulations for several different distributions, but this repetition would drastically increase the simulation time and thus, eliminate the advantages of simulation as opposed to numerical differentiation.

Yes, we can. Luckily, there is a mathematical trick that enables us to use Monte-Carlo simulation in interval setting as well. This trick is based on using *Cauchy distribution* – i.e., probability distributions with the probability density

$$\rho(z) = \frac{\Delta}{\pi \cdot (z^2 + \Delta^2)}; \tag{5}$$

the value Δ is called the *scale parameter* of this distribution, or simply a *parameter*, for short.

Cauchy distribution has the following property that we will use: if $z_1, \ldots, z_n$ are independent random variables, and each of z_i is distributed according to the Cauchy law with parameter Δ_i, then their linear combination $z = c_1 \cdot z_1 + \ldots + c_n \cdot z_n$ is also distributed according to a Cauchy law, with a scale parameter $\Delta = |c_1| \cdot \Delta_1 + \ldots + |c_n| \cdot \Delta_n$.

Therefore, if we take random variables δ_i which are Cauchy distributed with parameters Δ_i, then the value

$$c = f(\widetilde{x}_1 + \delta_1, \ldots, \widetilde{x}_n + \delta_n) - f(\widetilde{x}_1, \ldots, \widetilde{x}_n) = c_1 \cdot \delta_1 + \ldots + c_n \cdot \delta_n \tag{6}$$

is Cauchy distributed with the desired parameter (3). So, repeating this experiment N times, we get N values $c^{(1)}, \ldots, c^{(N)}$ which are Cauchy distributed with the unknown parameter, and from them we can estimate Δ.

The bigger N, the better estimates we get.

There are two questions to be solved:

- how to simulate the Cauchy distribution;
- how to estimate the parameter Δ of this distribution from a finite sample.

Simulation can be based on the functional transformation of uniformly distributed sample values:

$$\delta_i = \Delta_i \cdot \tan(\pi \cdot (r_i - 0.5)), \tag{7}$$

where r_i is uniformly distributed on the interval $[0, 1]$.

In order to estimate Δ, we can apply the Maximum Likelihood Method

$$\rho(d^1) \cdot \rho(d^2) \cdot \ldots \cdot \rho(d^n) \to \max,$$

where $\rho(z)$ is a Cauchy distribution density with the unknown Δ. When we substitute the above-given formula for $\rho(z)$ and equate the derivative of the product with respect to Δ to 0 (since it is a maximum), we get an equation

$$\frac{1}{1 + \left(\frac{c^{(1)}}{\Delta}\right)^2} + \ldots + \frac{1}{1 + \left(\frac{c^{(N)}}{\Delta}\right)^2} = \frac{N}{2}. \tag{8}$$

The left-hand side of (8) is an increasing function that is equal to $0 (< N/2)$ for $\Delta = 0$ and $> N/2$ for $\Delta = \max\left|c^{(k)}\right|$; therefore the solution to the equation (8) can be found by applying a bisection method to the interval $\left[0, \max\left|c^{(k)}\right|\right]$.

It is important to mention that we assumed that the function f is reasonably linear within the box

$$[\widetilde{x}_1 - \Delta_1, \widetilde{x}_1 + \Delta_1] \times \ldots \times [\widetilde{x}_n - \Delta_n, \widetilde{x}_n + \Delta_n]. \tag{9}$$

However, the simulated values δ_i may be outside the box. When we get such values, we do not use the function f for them, we use a normalized function that is equal to f within the box, and that is extended linearly for all other values (we will see, in the description of an algorithm, how this is done).

As a result, we arrive at the following algorithm (described, for a somewhat different problem, in [5, 6, 8, 9]):

Algorithm.

- Apply f to the results of direct measurements: $\widetilde{y} := f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$;
- For $k = 1, 2, \ldots, N$, repeat the following:
  - use the standard random number generator to compute n numbers $r_i^{(k)}$, $i = 1, 2, \ldots, n$, that are uniformly distributed on the interval $[0, 1]$;
  - compute Cauchy distributed values $c_i^{(k)} := \tan(\pi \cdot (r_i^{(k)} - 0.5))$;
  - compute the largest value of $|c_i^{(k)}|$ so that we will be able to normalize the simulated measurement errors and apply f to the values that are within the box of possible values: $K := \max_i |c_i^{(k)}|$;
  - compute the simulated measurement errors $\delta_i^{(k)} := \Delta_i \cdot c_i^{(k)}/K$;
  - compute the simulated measurement results $x_i^{(k)} := \widetilde{x}_i + \delta_i^{(k)}$;
  - apply the program f to the simulated measurement results and compute the simulated error of the indirect measurement:

$$c^{(k)} := K \cdot \left(f\left(x_1^{(k)}, \ldots, x_n^{(k)}\right) - \widetilde{y}\right);$$

- Compute Δ by applying the bisection method to solve the equation (8).

When is this randomized algorithm better than deterministic numerical differentiation? To determine the parameter Δ, we use the maximum likelihood method. It is known that the error of this method is asymptotically normally distributed, with 0 average and standard deviation $1/\sqrt{N \cdot I}$, where I is Fisher's information:

$$I = \int_{-\infty}^{\infty} \frac{1}{\rho} \cdot \left(\frac{\partial \rho}{\partial \Delta}\right)^2 \mathrm{d}z.$$

For Cauchy probability density $\rho(z)$, we have $I = 1/(2\Delta^2)$, so the error of the above randomized algorithm is asymptotically normally distributed, with a standard deviation $\sigma_e \sim \Delta \cdot \sqrt{2/N}$. Thus, if we use a "two sigma" bound, we conclude that with probability 95%, this algorithm leads to an estimate for Δ

which differs from the actual value of Δ by $\leq 2\sigma_e = 2\Delta \cdot \sqrt{2/N}$. So, if we want to achieve a 20% accuracy in the error estimation, we must use the smallest N for which $2\sigma_e = 2\Delta \cdot \sqrt{2/N} \leq 0.2 \cdot \Delta$, i.e., to select $N_f = N = 200$.

When it is sufficient to have a standard deviation of 20% (i.e., to have a "two sigma" guarantee of 40%), we need only $N = 50$ calls to f. For $n \approx 10^3$, both values N_f are much smaller than $N_f = n$ required for numerical differentiation.

So, if we have to choose between the (deterministic) numerical differentiation and the randomized Monte-Carlo algorithm, we must select:

- a deterministic algorithm when the number of variables n satisfies the inequality $n \leq N_0$ (where $N_0 \approx 200$), and
- a randomized algorithm if $n \geq N_0$.

Comment. If we use fewer than N_0 simulations, then we still get an approximate value of the range, but with worse accuracy – and the accuracy can be easily computed by using the above formulas.

This algorithm is naturally parallelizable. Similarly to the Monte-Carlo algorithm for statistical setting, we can run all N simulations in parallel and thus, speed up the computations.

Remark: the problem of non-linearity. In the above text, we assumed that the intervals $\mathbf{x}_i$ are narrow. In this case, terms quadratic in Δx_i are negligible, and so, we can safely assume that the desired function $f(x_1, \ldots, x_n)$ is linear on the box $\mathbf{x}_1 \times \ldots \times \mathbf{x}_n$. In practice, some intervals $\mathbf{x}_i$ may be wide, so even when restricted to the box, the function $f(x_1, \ldots, x_n)$ is non-linear. What can we do in this case?

Usually, only a few ranges are wide. For each of the corresponding variables, we can *bisect* the corresponding interval $[\underline{x}_i, \overline{x}_i]$ into two smaller subintervals – for which the dependence is approximately linear. Then, we estimate the range of the function f separately on each of the resulting subboxes, and take the union of these two ranges as the range over the entire box.

If one bisection is not enough and the dependence of f on x_i is non-linear over one or several subboxes, we can bisect these boxes again, etc.

This bisection idea has been successfully used in interval computations; see, e.g., [3, 4].

4 Conclusions

In many engineering problems, it is desirable to find the cheapest design among all designs that satisfy given constraints $f(x_1, \ldots, x_n) \leq f_0$. The constraints must be satisfied for all possible values of several parameters $x_1, \ldots, x_n$, and usually, we only know ranges of possible values $[\underline{x}_i, \overline{x}_i]$ for these parameters. Thus, to check whether the constraint is satisfied, we must check whether the upper endpoint $\overline{y}$ of the range $[\underline{y}, \overline{y}]$ of the function $f(x_1, \ldots, x_n)$ over $x_i \in\in [\underline{x}_i, \overline{x}_i]$ satisfies the desired inequality $\overline{y} \leq f_0$.

When we know the code for f, then we can use automatic differentiation (AD) and compute the range of f in time $O(T+n)$. If the owner of the program f only allows to use it as a black box, then we cannot use AD any more. In principle, we can compute each of n derivatives $\partial f/\partial x_i$ by numerical differentiation, but this would require computation time $T \cdot n \gg T + n$.

We have shown that we can also compute Δ in time $O(T)$ by using an artificial Monte-Carlo simulations in which each Δx_i is Cauchy distributed with parameter Δ_i – then simulated Δy is Cauchy distributed with the desired parameter $\Delta = \dfrac{\overline{y} - \underline{y}}{2}$.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

References

1. Berz, M., Bischof, C., Corliss, G., Griewank, A.: *Computational Differentiation: Techniques, Applications, and Tools*, SIAM, Philadelphia, 1996.
2. Griewank, A.: *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, SIAM, Philadelphia, 2000.
3. Jaulin, L., Kieffer, M., Didrit, O., Walter, E.: *Applied Interval Analysis*, Springer Verlag, London, 2001.
4. Kearfott, R. B., Kreinovich, V., Eds.: *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.
5. Kreinovich, V., Bernat, A., Villa, E., Mariscal, Y.: Parallel computers estimate errors caused by imprecise data. Interval Computations No. 2 (1991) 21–46.
6. Kreinovich, V., Ferson, S.: A New Cauchy-Based Black-Box Technique for Uncertainty in Risk Analysis. Reliability Engineering and Systems Safety **85** (2004) 267–279.
7. Kreinovich, V., Lakeyev, A., Rohn, J., Kahl, P.: *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1998.
8. Kreinovich, V., Pavlovich, M. I.: Error estimate of the result of indirect measurements by using a calculational experiment. Measurement Techniques **28** (1985) 201–205.
9. Trejo, R., Kreinovich, V.: Error estimations for indirect measurements: randomized vs. deterministic algorithms for 'black-box' programs. In: Rajasekaran, S., Pardalos, P., Reif, J., Rolim, J., Eds.: *Handbook on Randomized Computing*, Kluwer, 2001, pp. 673–729.
10. Vavasis, S. A.: *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York, 1991.

Towards Combining Probabilistic, Interval, Fuzzy Uncertainty, and Constraints: On the Example of Inverse Problem in Geophysics

G. Randy Keller, Scott A. Starks, Aaron Velasco, Matthew Averill, Roberto Araiza, Gang Xiang, and Vladik Kreinovich

University of Texas at El Paso, El Paso, TX 79968, USA,
keller@utep.edu sstarks@utep.edu vladik@utep.edu

Abstract. In many real-life situations, we have several types of uncertainty: measurement uncertainty can lead to probabilistic and/or interval uncertainty, expert estimates come with interval and/or fuzzy uncertainty, etc. In many situations, in addition to measurement uncertainty, we have prior knowledge coming from prior data processing, prior knowledge coming from prior interval constraints. In this paper, on the example of the seismic inverse problem, we show how to combine these different types of uncertainty.

1 Seismic Inverse Problem: A Brief Description

In evaluations of natural resources and in the search for natural resources, it is very important to determine Earth structure. Our civilization greatly depends on the things we extract from the Earth, such as fossil fuels (oil, coal, natural gas), minerals, and water. Our need for these commodities is constantly growing, and because of this growth, they are being exhausted. Even under the best conservation policies, there is (and there will be) a constant need to find new sources of minerals, fuels, and water.

The only sure-proof way to guarantee that there are resources such as minerals at a certain location is to actually drill a borehole and analyze the materials extracted. However, exploration for natural resources using indirect means began in earnest during the first half of the 20th century. The result was the discovery of many large relatively easy to locate resources such as the oil in the Middle East.

However, nowadays, most easy-to-access mineral resources have already been discovered. For example, new oil fields are mainly discovered either at large depths, or under water, or in very remote areas – in short, in the areas where drilling is very expensive. It is therefore desirable to predict the presence of resources as accurately as possible before we invest in drilling.

From previous exploration experiences, we usually have a good idea of what type of structures are symptomatic for a particular region. For example, oil and gas tend to concentrate near the top of natural underground domal structures.

So, to be able to distinguish between more promising and less promising locations, it is desirable to determine the structure of the Earth at these locations. To be more precise, we want to know the structure at different depths z at different locations (x, y).

Data that we can use to determine the Earth structure. In general, to determine the Earth structure, we can use different measurement results that can be obtained without actually drilling the boreholes: e.g., gravity and magnetic measurements, analyzing the travel-times and paths of seismic ways as they propagate through the earth, etc.

To get a better understanding of the Earth structure, we must rely on *active* seismic data – in other words, we must make artificial explosions, place sensors around them, and measure how the resulting seismic waves propagate. The most important information about the seismic wave is the *travel-time* t_i, i.e., the time that it takes for the wave to travel from its source to the sensor. to determine the geophysical structure of a region, we measure seismic travel times and reconstruct velocities at different depths from these data. The problem of reconstructing this structure is called the *seismic inverse problem.*

2 Known Algorithms for Solving the Seismic Inverse Problem: Description, Successes, Limitations

We want to find the values of the velocity $v(\boldsymbol{x})$ at different 3-D points $\boldsymbol{x}$. Based on the finite number of measurements, we can only reconstruct a finite number of parameters. So, we take a rectangular grid and to reconstruct the velocities v_j at different grid points.

Algorithm for the forward problem: brief description. Once we know the velocities v_j in each cell j, we can then determine the paths which seismic waves take. Seismic waves travel along the shortest path – shortest in terms of time. It can be easily determined that for such paths, within each cell, the path is a straight line, and on the border between the two cells with velocities v and v', the direction of the path changes in accordance with Snell's law $\dfrac{\sin(\varphi)}{v} = \dfrac{\sin(\varphi')}{v'}$, where φ and φ' are the angles between the paths and the line orthogonal to the border between the cells. (If this formula requires $\sin(\varphi') > 1$, this means that this wave cannot penetrate into the neighboring cell at all; instead, it bounces back into the original cell with the same angle φ.)

In particular, we can thus determine the paths from the source to each sensor. The travel-time t_i along i-th path can then be determined as the sum of travel-times in different cells j through which this path passes: $t_i = \sum\limits_j \dfrac{\ell_{ij}}{v_j}$, where ℓ_{ij} denotes the length of the part of i-th path within cell j.

This formula becomes linear if we replace the original unknowns – velocities v_j – by their inverses $s_j \stackrel{\text{def}}{=} \dfrac{1}{v_j}$, called *slownesses*. In terms of slownesses, the formula for the travel-time takes the simpler form $t_i = \sum\limits_j \ell_{ij} \cdot s_j$.

Algorithm for the inverse problem: general description. There are several algorithms for solving this inverse problem; see, e.g., [5, 7, 9]. The most widely used is the following iterative algorithm proposed by John Hole [5].

At each stage of this algorithm, we have some approximation to the desired slownesses. We start with some reasonable initial slownesses, and we hope that after several iterations, we will be able to get slownesses which are much closer to the actual values.

At each iteration, we first use the currently known slownesses s_j to find the corresponding paths from the source to each sensor. Based on these paths, we compute the predicted values $t_i = \sum\limits_j \ell_{ij} \cdot s_j$ of travel-times.

Since the currently known slownesses s_j are only approximately correct, the travel-times t_i (which are predicted based on these slownesses) are approximately equal to the measured travel-times $\widetilde{t}_i$; there is, in general, a discrepancy $\Delta t_i \stackrel{\text{def}}{=} \widetilde{t}_i - t_i \neq 0$. It is therefore necessary to use these discrepancies to update the current values of slownesses, i.e., replace the current values s_j with corrected values $s_j + \Delta s_j$. The objective of this correction is eliminate (or at least decrease) the discrepancies $\Delta t_i \neq 0$. In other words, the objective is to make sure that for the corrected values of the slowness, the predicted travel-times are closer to $\widetilde{t}_i$.

Of course, once we have changed the slownesses, the shortest paths will also change; however, if the current values of slownesses are reasonable, the differences in slowness are not large, and thus, paths will not change much. Thus, in the first approximation, we can assume that the paths are the same, i.e., that for each i and j, the length ℓ_{ij} remains the same. In this approximation, the new travel-times are equal to $\sum \ell_{ij} \cdot (s_j + \Delta s_j)$. The desired condition is then $\sum \ell_{ij} \cdot (s_j + \Delta s_j) = \widetilde{t}_i$. Subtracting the formula $t_i = \sum\limits_j \ell_{ij} \cdot s_j$ from this expression, we conclude that the corrections Δs_j must satisfy the following system of (approximate) linear equations: $\sum \ell_{ij} \cdot \Delta s_j \approx \Delta t_i$.

Solving this system of linear equations is not an easy task, because we have many observations and many cell values and thus, many unknowns, and for a system of linear equations, computation time required to solve it grows as a cube n^3 of the number of variables n. So, instead of the standard methods for solving a system of linear equations, researchers use special faster geophysics-motivated techniques (described below) for solving the corresponding systems. These methods are described, in detail, in the next subsection.

Once we solve the corresponding system of linear equations, we compute the updated values Δs_j, compute the new (corrected) slownesses $s_j + \Delta s_j$, and repeat the procedure again. We stop when the discrepancies become small; usually, we stop when the mean square error $\dfrac{1}{n} \sum\limits_{i=1}^{n} (\Delta t_i)^2$ no longer exceeds a given

threshold. This threshold is normally set up to be equal to the measurement noise level, so that we stop iterations when the discrepancy between the model and the observations falls below the noise level – i.e., when, for all practical purposes, the model is adequate.

Algorithm for the inverse problem: details. Let us describe, in more detail, how the corresponding linear system of equations is usually solved. In other words, for a given cell j, how do we find the correction Δs_j to the current value of slowness s_j in this cell?

Let us first consider the simplified case when there is only path, and this path is going through the j-th cell. In this case, cells through which this path does not go does not need any correction. To find the corrections Δs_j for all the cells j through which this path goes, we only have one equation $\sum\limits_j \ell_{ij} \cdot \Delta s_j = \Delta t_i$. The resulting system of linear equations is clearly under-determined: we have a single equation to find the values of several variables Δs_j. Since the system is under-determined, we have a infinite number of possible solutions. Our objective is to select the most geophysical reasonable of these solutions.

For that, we can use the following idea. Our single observation involves several cells; we cannot distinguish between the effects of slownesses in different cells, we only observe the overall effect. Therefore, there is no reason to assume that the value Δs_j in one of these cells is different from the values in other cells. It is thus reasonable to assume that all these values are close to each other: $\Delta s_j \approx \Delta s_{j'}$. The least squares method enables us to describe this assumption as minimization of the objective function $\sum\limits_{j,j'} (\Delta s_j - \Delta s_{j'})^2$ under the condition that $\sum \ell_{ij} \cdot \Delta s_j = \Delta t_i$. The minimum is attained when all the values Δs_j are equal. Substituting these equal values into the equation $\sum\limits_j \ell_{ij} \cdot \Delta s_j = \Delta t_i$, we conclude that $L_i \cdot \Delta s = \Delta t_i$, where $L_i = \sum\limits_j \ell_{ij}$ is the overall length of i-th path. Thus, in the simplified case in which there is only one path, to the slowness of each cell j along this path, we add the same value $\Delta s_j = \dfrac{\Delta t_i}{L_i}$.

Let us now consider the realistic case in which there are many paths, and moreover, for many cells j, there are many paths i which go through the corresponding cell. For a given cell j, based on each path i passing through this cell, we can estimate the correction Δs_j by the corresponding value $\Delta s_{ij} \stackrel{\text{def}}{=} \dfrac{\Delta t_i}{L_i}$. Since there are usually several paths going through the j-th cell, we have, in general, several different estimates $\Delta s_j \approx \Delta s_{ij}$. Again, the least squares approach leads to $\sum\limits_i (\Delta s_j - \Delta s_{ij})^2 \to \min$, hence to Δs_j as the arithmetic average of the values Δs_{ij}.

Comment. To take into account that paths with larger ℓ_{ij} provide more information, researchers also used weighted average, with weight increasing with ℓ_{ij}.

Successes of the known algorithms. The known algorithms have been actively used to reconstruct the slownesses, and, in many practical situations, they have led to reasonable geophysical models.

Limitations of the known algorithms. Often, the velocity model that is returned by the existing algorithm is not geophysically meaningful: e.g., it predicts velocities outside of the range of reasonable velocities at this depth. To avoid such situations, it is desirable to incorporate the expert knowledge into the algorithm for solving the inverse problem.

In this paper, we describe how to do it.

3 Case of Interval Prior Knowledge

For each cell j, a geophysicist often provides us with his or her estimate of possible values of the corresponding slowness s_j. Sometimes, this estimates comes in the form of an interval $[\underline{s}_j, \overline{s}_j]$ that is guaranteed to contain the (unknown) actual value of slowness.

It is desirable to modify Hole's algorithm in such a way that on all iterations, slownesses s_j stay within the corresponding intervals. Such a modification is described in [1, 2].

Namely, in the original Hole's algorithm, once we know the current approximations $s_j^{(k)}$ to slownesses, then, along each path i, among all corrections Δs_{ij} that provide the desired compensation, i.e.., for which

$$\sum_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} = \Delta t_i, \tag{1}$$

we find the assignment that minimizes the objective function $\sum\limits_{j,j'} (\Delta s_{ij} - \Delta s_{ij'})^2$, i.e., equivalently, that minimizes the variance of the values Δs_{ij} along this path:

$$V \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}^2 - \left(\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij} \right)^2. \tag{2}$$

In the presence of the interval prior information, on each iteration of Hole's algorithm, we must still minimize the objective function (2), but this time, we minimize it under two constraint: the same constraint (1) and the new constraints

$$\underline{s}_j \le s_j^{(k)} + \Delta s_{ij} \le \overline{s}_j. \tag{3}$$

We have found the following efficient $O(c \cdot \log(c))$ time algorithm for solving the corresponding constraint optimization problem. We start with the initial slowness values $s_j^{(0)}$ which are within the given intervals $[\underline{s}_j, \overline{s}_j]$.

On each iteration of the new procedure, we start with the slowness values $s_j^{(k-1)}$ which are within given intervals $[\underline{s}_j, \overline{s}_j]$. Based on these slownesses, we find

the paths from the sources to the sensors, compute the predicted travel-times t_i along each path, and the discrepancies $\Delta t_i = \widetilde{t}_i - t_i$.

We then compute, for each cell j, the values $\underline{\Delta}_j = \underline{s}_j - s_j^{(k-1)}$ and $\overline{\Delta}_j = \overline{s}_j - s_j^{(k-1)}$. We will consider the case when $\Delta t_i > 0$; the case when $\Delta t_i < 0$ is treated similarly. In this case, we first sort all c values $\overline{\Delta}_j$ along the i-th path into a non-decreasing sequence

$$\overline{\Delta}_{(1)} \le \overline{\Delta}_{(2)} \le \ldots \le \overline{\Delta}_{(c)}.$$

Then, for every p from 0 to c, we compute the values A_p and $\mathcal{L}_p$ as follows:

$$A_0 = 0, \quad \mathcal{L}_0 = L_i, \quad A_p = A_{p-1} + \ell_{i(p)} \cdot \overline{\Delta}_{(p)}, \quad \mathcal{L}_p = \mathcal{L}_{p-1} - \ell_{i(p)}.$$

After that, for each p, we compute $S_p = A_p + \mathcal{L}_p \cdot \Delta_{(p+1)}$, and we find p for which $S_{p-1} \le \Delta t_i < S_p$. Once this p is found, we take $\Delta s_{i(j)} = \overline{\Delta}_j$ for $j \le p$, and for $j > p$, we take $\Delta s_{(j)} = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$.

When $\Delta t_i < 0$, we similarly sort the values $\underline{\Delta}_j$ into a decreasing sequence, and find p so that the first p corrections are "maxed out" to $\underline{\Delta}_j$, and the rest $c - p$ corrections are determined from the condition $\Delta s_{i(j)} = \dfrac{\Delta t_i - A_p}{\mathcal{L}_p}$.

Once we have computed these corrections for all the paths, then for each cell j, we take the average (or weighted average) of all the corrections coming from all the paths which pass through this cell.

4 Case of Fuzzy Prior Knowledge

In general, experts are often not 100% sure about the corresponding intervals. They can usually produce a wider interval $[\underline{s}_j, \overline{s}_j]$ of which they are practically 100% certain, but in addition to that, they can also produce narrower intervals about which their degree of certainty is smaller. As a result, instead of a single interval, we have a nested family of intervals corresponding to different levels of uncertainty – i.e., in effect, a fuzzy interval (of which different intervals are α-cuts).

So, instead of simply saying that a given solution to the seismic inverse problem is satisfying or not, we provide a *degree* to which the given solution is satisfying – as the largest α for which the velocity at every point is within the corresponding α-cut intervals.

To solve the seismic inverse problem under such fuzzy uncertainty, we apply the interval algorithm for α-cuts corresponding to $\alpha = 0$, $\alpha = 0.1$, $\alpha = 0.2$, etc., until we reach such a value of α that the process no longer converges. Then, the solution corresponding to the previous value α – i.e., to the largest value α for which the process converged – is returned as the desired solution to the seismic inverse problem.

5 Case of Probabilistic Prior Knowledge

Often, prior information comes from processing previous observations of the region of interest. In this case, before our experiments, for each cell j, we know a prior (approximate) slowness value $\widetilde{s}_j$, and we know the accuracy (standard deviation) σ_j of this approximate value $\widetilde{s}_j$. It is known that this prior information can lead to much more accurate velocity models; see, e.g., [6]. How can we modify Hole's algorithm so that it takes this prior information into account?

Due to the prior knowledge, for each cell j, the ratio $\dfrac{(s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j}{\sigma_j}$ is normally distributed with 0 mean and variance 1. Since each path i consists of a reasonable number of cells, we can thus conclude that the sample variance of this ratio should be close to σ_j, i.e., that

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \frac{((s_j^{(k)} + \Delta s_{ij}) - \widetilde{s}_j)^2}{\sigma_j^2} = 1. \tag{4}$$

So, to find the corrections Δs_{ij}, we must minimize the objective function (2) under the constraints (1) and (4).

By applying the Lagrange multiplier method to this problem, we can reduce this problem to the unconstrained minimization problem

$$\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}^2 - \left(\frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij} \right)^2 + \lambda \cdot \left(\sum_{j=1}^{c} \ell_{ij} \cdot \Delta s_{ij} - \Delta t_i \right) +$$

$$\mu \cdot \frac{1}{n} \cdot \sum_{j=1}^{c} \frac{(s_j^{(k)} + \Delta s_{ij} - \widetilde{s}_j)^2}{\sigma_j^2} \to \min. \tag{5}$$

Differentiating this equation by Δs_{ij} and equating the derivative to 0, we conclude that

$$\frac{2}{n} \cdot \Delta s_{ij} - \frac{2}{n} \cdot \overline{\Delta s} + \lambda \cdot \ell_{ij} + \frac{2\mu}{n \cdot \sigma_j^2} \cdot (s_j^{(k)} + \Delta s_{ij} - \widetilde{s}_j) = 0,$$

where

$$\overline{\Delta s} \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{j=1}^{c} \Delta s_{ij}. \tag{6}$$

Once we fix λ, μ, and $\overline{\Delta s}$, we get an explicit expression for the values Δs_{ij}. Substituting these expressions into the equations (1), (4), and (6), we get an easy-to-solve system of 3 non-linear equations with 3 unknowns, which we can solve, e.g., by using Newton's method.

Now, instead of explicit formulas for a transition from $s_j^{(k)}$ to $s_j^{(k+1)}$, we need a separate iteration process – so the computation time is somewhat larger, but we get a more geophysically meaningful velocity map – that takes prior knowledge into account.

6 Combination of Different Types of Prior Knowledge

In many real-life situations, we have both the prior measurement results – which lead to the probabilistic prior knowledge, and expert estimates – which lead to interval and fuzzy prior knowledge. In the presence of probabilistic and interval prior knowledge, we must minimize (2) under the constraints (1), (3), and (4).

If we replace the equality in (4) by an inequality ≤ 1, then we get a problem of minimizing a convex function under convex constraints, a problem for which there are known efficient algorithms; see, e.g., [8].

For example, we can use a method of alternating projections, in which we first add a correction that satisfy the first constraint, then the additional correction that satisfies the second constraint, etc. In our case, we first add equal values of Δs_{ij} to satisfy the constraint (2), then we restrict the values to the nearest points from the interval $[\underline{s}_j, \overline{s}_j]$ – to satisfy the constraint (3), and after that, find the extra corrections that satisfy the condition (4), after which we repeat the cycle again until the process converges.

Acknowledgments. This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

References

1. Averill, M. G., Miller, K. C., Keller, G. R., Kreinovich, V., Araiza, R., Starks, S. A.: Using Expert Knowledge in Solving the Seismic Inverse Problem. Proceedings of the 24nd International Conference of the North American Fuzzy Information Processing Society NAFIPS'2005, Ann Arbor, Michigan, June 22–25, 2005, pp. 310–314.
2. Averill, M. G., Miller, K. C., Keller, G. R., Kreinovich, V., Araiza, R., Starks, S. A.: Using Expert Knowledge in Solving the Seismic Inverse Problem. International Journal of Approximate Reasoning (to appear).
3. Bardossy, G., Fodor, J.: *Evaluation of Uncertainties and Risks in Geology*, Springer Verlag, Berlin, 2004.
4. Demicco, R., Klir, J., Eds.: *Fuzzy Logic in Geology*, Academic Press, 2003.
5. Hole J. A.: Nonlinear High-Resolution Three-Dimensional Seismic Travel Time Tomography. J. Geophysical Research **97** (1992) 6553–6562.
6. Maceira, M., Taylor, S. R., Ammon, C. J., Yang, X., Velasco, A. A.: High-resolution Rayleigh wave slowness tomography of Central Asia. Journal of Geophysical Research **110** (2005) paper B06304.
7. Parker R. L.: *Geophysical Inverse Theory*, Princeton University Press, Princeton, New Jersey, 1994.
8. Vavasis, S. A.: *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York, 1991.
9. Zelt, C. A., Barton, P. J.: Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin. J. Geophysical Research **103** (1998) 7187–7210.

Bilinear Models from System Approach Justified for Classification, with Potential Applications to Bioinformatics

Richard Aló[1], François Modave[2], Vladik Kreinovich[2],
David Herrera[2], and Xiaojing Wang[2]

[1] Center for Computational Sciences & Advanced Distributed Simulation,
University of Houston-Downtown, One Main Street,
Houston, TX 77002, USA, RAlo@uh.edu
[2] Department of Computer Science, University of Texas at El Paso,
El Paso, TX 79968, USA, vladik@utep.edu

Abstract. When we do not know the dynamics of a complex system, it is natural to use common sense to get a reasonable first approximation – which turns out to be a bilinear dynamics. Surprisingly, for classification problems, a similar bilinear approximation turns out to be unexpectedly accurate. In this paper, we provide an explanation for this accuracy.

1 System Approach: In Brief

For many complex systems, e.g., for large-scale financial or biological systems, we do not know the exact equations that describe the system's evolution. In such situations, a reasonable idea is to use *system approach*, i.e., to describe the dynamics of the parameters $x_1, \ldots, x_n$ (that describe the system) by differential equations

$$\dot{x}_i = f_i(x_1, \ldots, x_n),$$

and build the expressions for f_i based on common sense [1, 2, 4, 5] (see also [3]).

For example, if an increase in x_i slows down the growth of x_j, then the expression f_j for $\dot{x}_j$ should include a term $-k \cdot x_i$ for some $k > 0$.

If the two factors x_j and x_k, when combined, produce an increased positive effect on the growth of x_i, then the expression f_i for $\dot{x}_i$ should include a bilinear term $+k \cdot x_j \cdot x_k$. The values of the corresponding parameters k can be determined empirically.

Common sense rarely goes beyond such simple interaction between parameters, so we end up with linear and bilinear expressions in $\dot{x}_i$.

Starting from the global economic and environmental models developed in the 1960s by the Club of Rome's Limits to Growth project [2], the resulting models provide a reasonable first *qualitative* approximation to the dynamics of complex systems – although, of course, to get good *quantitative* predictions, we need more sophisticated models.

2 Systems Approach in Classification

A similar approach can be used in classification and clustering, where we need to separate, e.g., stocks with a good growth potential from the risky ones, or cancerous cells from the normal ones.

To separate two classes based on the values of the parameters x_i, we can use a discrimination function $f(x_1, \ldots, x_n)$:

- objects for which $f > 0$ belong to the first class, and
- objects for which $f < 0$ belong to the second class.

3 Systems Approach in Classification Works Surprisingly Well: A Mystery and a Related Practical Question

By using common sense, we can also easily come up with a bilinear model for f, and we can empirically find the coefficients of the corresponding bilinear expression.

We applied this approach to bioinformatics data, and surprisingly, the resulting bilinear models provide not only a good qualitative fit, but a good quantitative fit as well. A natural question is: Is it a special feature of bioinformatics data, or we can hope to get a good quantitative fit for, e.g., financial data as well?

4 Mystery Explained

In this paper, we show that the success of bilinear models in classification is a general mathematical phenomenon – caused by the fact that we can change a discrimination function f and keep the same classes, as long as we do not change the set of value for which $f \geq 0$.

5 Approximations of Different Order of Accuracy: General Idea

In accordance with the above text, let us denote the actual (unknown) discrimination function by $f(x_1, \ldots, x_n)$. It is usually reasonable to assume that the function $f(x_1, \ldots, x_n)$ is smooth. Therefore, in the vicinity of every point $\widetilde{x} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$, we can expand this function into the Taylor series and keep only lower order terms in this expansion.

By using this approximate polynomial expression for the function $f(x_1, \ldots, x_n)$, we can get a good approximate description of the classification in the vicinity of the point $\widetilde{x}$.

Which starting points should we consider?

– If we consider a point $\widetilde{x}$ which is completely inside the first class, then all the points in its vicinity belong to the same class – and thus, in this vicinity, the classification problem can be trivially solved.
– Similarly, if a point $\widetilde{x}$ is completely inside the second class, then all the points in its vicinity belong to the same class – and thus, in this vicinity, the classification problem can also be trivially solved.

Thus, the localized classification problem is of interest only when the starting point $\widetilde{x} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$ is located on the border between the two classes, i.e., when

$$f(\widetilde{x}_1, \ldots, \widetilde{x}_n) = 0.$$

We can somewhat simplify the formulas if we introduce new coordinates $x_i \to x_i - \widetilde{x}_i$ in which the starting point takes the form $(0, \ldots, 0)$. In these new coordinates, the condition that the starting points lies on the border between the two classes takes the form

$$f(0, \ldots, 0) = 0.$$

Thus, in the following text, we will assume that the starting point is 0, and that $f(0, \ldots, 0) = 0$.

6 First Approximation: Linear Separation

In the first approximation, we can approximate an arbitrary smooth function $f(x_1, \ldots, x_n)$ in the vicinity of 0 by a linear expression:

$$f(x_1, \ldots, x_n) = a_0 + \sum_{i=1}^{n} a_i \cdot x_i.$$

Since our interest is in classification, we are only interested in functions for which $f(0, \ldots, 0) = 0$. Substituting $x_1 = \ldots = x_n = 0$ into the above general expression for a linear function and equating the result to 0, we conclude that $a_0 = 0$.

Thus, in the first approximation, the general classification-related discrimination function takes the form

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i \cdot x_i.$$

7 Second Approximation: Quadratic Separation

In the next (second) approximation, we can approximate an arbitrary smooth function $f(x_1, \ldots, x_n)$ in the vicinity of 0 by a quadratic expression:

$$f(x_1, \ldots, x_n) = a_0 + \sum_{i=1}^{n} a_i \cdot x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot x_i \cdot x_j.$$

Since we need $f(0,\ldots,0)=0$, we conclude that $a_0=0$ and thus, that

$$f(x_1,\ldots,x_n)=\sum_{i=1}^{n} a_i\cdot x_i+\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\cdot x_i\cdot x_j.$$

This general expression is much more general than bilinear:

- it has linear terms $a_i\cdot x_i$;
- it has bilinear terms $a_{ij}\cdot x_i\cdot x_j$ for $i\neq j$;
- however, the general expression also has purely quadratic (not bilinear) terms $a_{ii}\cdot x_i^2$ (corresponding to $i=j$).

8 For System Dynamics, Bilinear Functions Provide a Rather Crude Approximation

In the original application of systems theory, the function $f(x_1,\ldots,x_n)$ describes the dynamics of the system, i.e., the (first) time derivative of the corresponding coordinates. In such applications, the function $f(x_1,\ldots,x_n)$ can be determined by observing the dynamics of the system, i.e., by observing how the values of the parameters x_i change with time. Since in principle, we can have an arbitrary dynamical system, we can therefore have arbitrary functions $f(x_1,\ldots,x_n)$.

When we approximate a general function $f(x_1,\ldots,x_n)$ by a linear expression, we thus ignore quadratic and higher order terms in the expansion of this function $f(x_1,\ldots,x_n)$. Thus, the approximation error of this approximation is quadratic (in x_i).

Similarly, when we approximate a general function $f(x_1,\ldots,x_n)$ by a quadratic expression, we thus ignore cubic and higher order terms in the expansion of this function $f(x_1,\ldots,x_n)$. Thus, the approximation error of this approximation is cubic (in x_i).

In principle, we can consider systems-theory bilinear approximation, in which we only keep linear and bilinear terms in the quadratic expansion but we ignore purely quadratic terms $a_{ii}\cdot x_i^2$. This approximation is intermediate between the linear and the quadratic ones.

What is the accuracy of this intermediate approximation? Since we ignore some quadratic terms, the accuracy of this approximation is quadratic is x_i. In other words, this intermediate approximation is asymptotically of the same order of accuracy as the much simpler linear approximation.

Comment. Of course, the accuracy of the bilinear model is somewhat better than the accuracy of the linear approximation – since in the intermediate approximation, we keep some quadratic terms. For example, in systems in which all the variables x_i are highly related, it is reasonable assume that all the terms $a_{ij}\cdot x_i\cdot x_j$ are of the same order of magnitude. In this case,

- in the linear approximation, we ignore n^2 terms $a_{ij}\cdot x_i\cdot x_j$ corresponding to all possible pairs (i,j), $i=1,2,\ldots,n$, $j=1,2,\ldots,n$; so, the approximation error is of order $n^2\cdot\delta$, where δ is the average value of each term $a_{ij}\cdot x_i\cdot x_j$;

– in the bilinear approximation, we only ignore n quadratic terms $a_{ii} \cdot x_i^2$ corresponding to $i = 1, 2, \ldots, n$; so, the approximation error is of order $n \cdot \delta$.

Thus, the approximation error of the bilinear model is $\approx n$ times smaller than for the linear model – i.e., for large n, drastically smaller.

9 The Fact that We Are Interested in Classification Applications Allows Further Simplifications

In dynamics applications, the function $f(x_1, \ldots, x_n)$ can be determined from observations. In the classification applications, however, the only think that we can infer from observations is for which points $x = (x_1, \ldots, x_n)$, the function $f(x_1, \ldots, x_n)$ has positive values and for which points $x = (x_1, \ldots, x_n)$, the function $f(x_, \ldots, x_n)$ has negative values.

Thus, if we replace the original function $f(x_1, \ldots, x_n)$ with an "equivalent" new function $f'(x_1, \ldots, x_n)$ – i.e., a new function that leads to the same subdivision of points into the first class and the second class – then both function will be consistent with the same observations. Thus, from observations, we will not be able to tell whether the actual discrimination function is $f(x_1, \ldots, x_n)$ or $f'(x_1, \ldots, x_n)$.

In this paper, we will use this non-uniqueness of the discrimination function $f(x_1, \ldots, x_n)$ to simplify the general quadratic expression for f.

In precise terms, the two functions $f(x_1, \ldots, x_n)$ and $f'(x_1, \ldots, x_n)$ are equivalent if the following two conditions are satisfied:

$$f(x_1, \ldots, x_n) > 0 \text{ if and only if } f'(x_1, \ldots, x_n) > 0$$

and

$$f(x_1, \ldots, x_n) > 0 \text{ if and only if } f'(x_1, \ldots, x_n) > 0.$$

10 Explanation of Bilinear Functions

As the new function $f'(x_1, \ldots, x_n)$, we will consider a function of the type

$$f'(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) \cdot \left(1 + \sum_{j=1}^{n} b_j \cdot x_j\right).$$

In the small vicinity of 0, $\left|\sum\limits_{i=j}^{n} b_j \cdot x_j\right| \ll 1$, hence

$$1 + \sum_{i=1}^{n} b_j \cdot x_j > 0.$$

Thus, in this vicinity, the values $f(x_1, \ldots, x_n)$ and $f'(x_1, \ldots, x_n)$ have the same sign – i.e., these functions are indeed equivalent.

Let us show that for a generic quadratic function

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i \cdot x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot x_i \cdot x_j,$$

we can select the values b_j in such a way that the resulting product $f'(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) \cdot \left(1 + \sum\limits_{j=1}^{n} b_j \cdot x_j\right)$ (or, to be more precise, the quadratic approximation to this product) does not have purely quadratic terms and is, thus, purely bilinear.

Indeed, in this product $f'(x_1, \ldots, x_n)$ the terms coming from multiplying $\sum a_{ij} \cdot x_i \cdot x_j$ and $\sum b_j \cdot x_j$ are cubic and can be thus, in this approximation, safely ignored. Thus, in our quadratic approximation, only the product of $\sum a_i \cdot x_i$ and $\sum b_j \cdot x_j$ must be added. So, in the quadratic approximation, the product function has the following form:

$$f'(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i \cdot x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot x_i \cdot x_j + \left(\sum_{i=1}^{n} a_i \cdot x_i\right) \cdot \left(\sum_{j=1}^{n} b_j \cdot x_j\right).$$

What are the purely quadratic terms in this expression? For each i, we have two terms proportional to x_i^2: the original term $a_{ii} \cdot x_i^2$ and the new term $a_i \cdot b_i \cdot x_i^2$. Thus, if $a_{ii} + a_i \cdot b_i = 0$ for all $i = 1, 2, \ldots, n$, the resulting expression $f'(x_1, \ldots, x_n)$ will be free of purely quadratic terms – i.e., bilinear.

Thus, it is sufficient to take

$$b_i = -\frac{a_{ii}}{a_i}.$$

Of course, this division is only possible when $a_i \neq 0$ for all i – but this is what is happening in the *generic* case, and this is what we planned to prove – that in the generic case, it is possible to use a purely bilinear discrimination function.

The result is proven.

11 Discussion: A Similar Simplification Is Not Always Possible for Higher Order Models

We have shown that in the second approximation, we can reduce an arbitrary discrimination function to an equivalent bilinear one.

Bilinear terms come from common sense analysis of the system. For higher order terms, a similar common sense analysis enable us to come up with trilinear terms $a_{ijk} \cdot x_i \cdot x_j \cdot x_k$, where i, j, and k are different variables. A natural question is: can we reduce a general cubic expression to a trilinear one?

More generally, we can have multi-linear functions, i.e., functions $f(x_1, \ldots, x_n)$ which are linear in each of their variables x_i. A natural general

question is: can we always reduce an arbitrary discrimination function to a multi-linear one?

The answer to this general equation is "no". Indeed, for any fixed number of variables $x_1, \ldots, x_n$, we only have finite many possible multi-linear terms:

- for a single variable x_1, we can only have one term $a_1 \cdot x_1$;
- for two variables x_1 and x_2, we can only have three term $a_1 \cdot x_1 + a_2 \cdot x_2 + a_{12} \cdot x_1 \cdot x_2$;
- etc.

Thus, for each n, we have a finite-parametric family of multi-linear discrimination functions, i.e., a family that can be characterized by finitely many parameters.

On the other hand, the set of possible class is infinitely-parametric: indeed, each class can be described by a smooth separating line $x_n = F(x_1, \ldots, x_{n-1})$, and a general separating line can be described by a general Taylor expansion and thus, require infinitely many parameters.

So, for general higher order approximations, multi-linear functions are not sufficient.

Are multi-linear functions sufficient at least for cubic terms? The answer is again "no", even for $n = 3$ variables. Indeed, a general trilinear discrimination function of 3 variables has:

- three parameters a_1, a_2, and a_3 describing the linear terms;
- three parameters a_{12}, a_{13}, and a_{23} describing the bilinear terms; and
- a single parameter a_{123} describing the only possible trilinear term

$$a_{123} \cdot x_1 \cdot x_2 \cdot x_3.$$

to the total of $3 + 3 + 1 = 7$ parameters. On the other hand, a general cubic discriminating curve $x_3 = F(x_1, x_2)$, with a general cubic dependence

$$F(x_1, x_2) = b_1 \cdot x_1 + b_2 \cdot x_2 + b_{11} \cdot x_1^2 + b_{12} \cdot x_1 \cdot x_2 + b_{12} \cdot x_2^2 +$$

$$b_{111} \cdot x_1^3 + b_{112} \cdot x_1^2 \cdot x_2 + b_{122} \cdot x_1 \cdot x_2^2 + b_{222} \cdot x_2^3$$

requires $9 > 7$ parameters. Thus, it is not possible to describe a general third-order classification by a trilinear discrimination function.

Comment. A similar analysis can answer the following natural question: we have reduced a general quadratic discrimination function to a bilinear one; can we reduce it further, to some class with even fewer parameters?

To answer this question, let us count how many parameters we need to describe a general bilinear function of n variables, and how many parameters we need to describe a general class in the quadratic approximation.

To describe a general bilinear function of n variables, we need to describe:

- n coefficients a_i, $i = 1, 2, \ldots, n$;
- $\dfrac{n \cdot (n-1)}{2}$ coefficients a_{ij} corresponding to all possible pairs (i, j), $i \neq j$.

The total number of parameters needed for this approximation is

$$n + \frac{n \cdot (n-1)}{2} = \frac{1}{2} \cdot (2n + n^2 - n) = \frac{1}{2} \cdot (n^2 + n).$$

A generic second-order classification can be described by a quadratic expression

$$F(x_1, \ldots, x_{n_1}) = \sum_{i=1}^{n} b_i \cdot x_i + \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} b_{ij} \cdot x_i \cdot x_j$$

describing the separating curve $x_n = F(x_1, \ldots, x_{n-1})$. To describe this general quadratic function, we need:

- $n-1$ parameters b_i, $i = 1, 2, \ldots, n-1$;
- $\frac{(n-1) \cdot (n-2)}{2}$ parameters b_{ij} corresponding to all possible pairs (i, j), $i \neq j$; and
- $n-1$ parameters b_{ii}, $i = 1, 2, \ldots, n-1$.

The total number of parameters needed for this approximation is

$$n - 1 + \frac{(n-1) \cdot (n-2)}{2} + n - 1 = \frac{1}{2} \cdot (2(n-1) + (n^2 - 3n + 2) + 2(n-1)) =$$

$$\frac{1}{2} \cdot (n^2 + n - 2) = \frac{1}{2} \cdot (n^2 + n) - 1.$$

By comparing these two values, we can see that there is only one extra parameter in the bilinear expression – and it can be reduced by setting, e.g., $a_1 = \pm 1$. This reduction can be achieved if we divide the original function $f(x_1, \ldots, x_n)$ by a positive number $|a_1|$ – this division does not change the sign of $f(x_1, \ldots, x_n)$ and thus, leads to an equivalent discrimination function. After this simple reduction, we have exactly as many parameters as we need to describe a general quadratic expression – and thus, no further reduction is possible.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, Star Award from the University of Texas System, and Texas Department of Transportation grant No. 0-5453.

References

1. Forrester, J. W.: *Principles of Systems*, Pegasus Communications, 1968.
2. Meadows, D. H.: *Limits to Growth*, Signet Publ., 1972.
3. Nguyen, H. T., Kreinovich, V.: *Applications of Continuous Mathematics in Computer Science*, Kluwer, Dordrecht, 1997.
4. von Bertalanffy, L.: *Perspectives on General System Theory*. G. Braziller, New York, 1975.
5. von Bertalanffy, L.: *General System Theory*. G. Braziller, New York, 1984.

Towards Economics of Education: Optimization under Uncertainty

Olga Kosheleva[1] and Richard Aló[2]

[1] Department of Teacher Education, University of Texas at El Paso,
El Paso, TX 79968, USA, olgak@utep.edu
[2] Center for Computational Sciences & Advanced Distributed Simulation,
University of Houston-Downtown, One Main Street,
Houston, TX 77002, USA, RAlo@uh.edu

Abstract. Since we cannot spend as much time as we would like to on teaching all the topics, it is necessary to optimally distribute the limited amount of time between different topics. In this paper, we explain how general techniques of optimization under uncertainty can be used in education.

1 Optimizing Content Development

Basic assumptions about training. In order to find the optimal training schedule, let us make some (simplifying but realistic) assumptions about training.

In principle, there are several different types of items that we want a student to learn; for example:

- when we teach typing, we want the student to acquire the motor skills of typing all the symbols on the keyboard and all pairs of consequent symbols;
- when we teach words from a foreign language, we want the student to learn all these words;
- when we teach cross-country driving, we want the student to develop motor skills corresponding to different types of terrain: flat surface, rugged terrain, uphill, downhill, narrow bridge, etc.

Let us denote the total number of types to learn by T. For simplicity, we assume that acquiring skills necessary for each of these types takes the same number of training situations s. So, to learn all necessary types, a students needs at least $T \cdot s$ repetitions. If we denote, by T_0, the time necessary for handling each repetition, then the total time for training a student for all necessary types is equal to $T_0 \cdot T \cdot s$.

In many learning situations, the total number T of necessary types is large, so the above time of total training is unrealistically large. Therefore, we cannot expect every single student to be 100% skilled in every possible situation type.

Since we cannot train a student to be skilled in every possible situation, it is therefore necessary to train a student in such a way that the student will be able to handle *the largest possible number* of these types.

In future applications, some of these types are more frequent, some are less frequent. So, if we know that a student can only learn, say, t different words, and we have to choose which of these words the student will learn perfectly well, we should choose t most frequent ones.

A *skill* of a student can be thus characterized by the number t of the types of items in which this student is well skilled.

To estimate frequencies of different types, we can use a general (semi-empirical) law discovered by G. K. Zipf (see, e.g., [3, 5]), according to which, if we order types from the most frequent to the least frequent one, then the frequency f_i of i-th type is proportional to $1/i$: $f_i = c/i$ for some constant c. The value of this constant can be determined from the fact that the sum of all these frequencies should be equal to 1: $f_1 + \ldots + f_T = 1$. Since $1 + 1/2 + \ldots + 1/T \approx \ln(T)$, we thus conclude that $c \cdot \ln(T) = 1$, $c = 1/\ln(T)$, and

$$f_i = \frac{1}{\ln(T) \cdot i}. \tag{1}$$

Traditional learning. In traditional learning of a language, a student in trained on texts from real language. In traditional typing lessons, student learn to type by typing real-life texts. In all these cases, a student is trained on a real-life flow of items.

Let us denote by I the time allocated for training. Since handling each repetition takes time T_0, during this training time, the trainee will see $N = I/T_0$ repetitions. According to our assumption about the training time, the student will be trained only in those types i for which he or she has seen at least s repetitions. Out of the total of N repetitions, the student will see $N \cdot f_i$ repetitions of i-th type; so, the student will be trained in all the types for which $N \cdot f_i \geq s$. Substituting Zipf's expression (1) for f_i, we conclude that the student will learn all the types i for which $\frac{I}{T_0} \cdot \frac{1}{\ln(T) \cdot i} \geq s$, i.e., for which $i \leq \frac{I}{T_0 \cdot \ln(T) \cdot s}$. Therefore, the resulting student's skill level t (i.e., the total number of types in which this student will be skilled), will be equal to

$$t = \frac{I}{T_0 \cdot \ln(T) \cdot s}. \tag{2}$$

This formula describes the skill level acquired during a given training time I.

We can also consider the inverse problem: we want a student to be trained for a certain skill level t, and we need to know the time I required for this training. From the formula (2), we can conclude that

$$I = t \cdot T_0 \cdot \ln(T) \cdot s. \tag{3}$$

Optimal training. We can generate repetitions in arbitrary order, not necessarily with real-life frequencies. If we want a student to be trained on t different types, then we need to generate exactly s repetitions of this type.

If we fix the total training time I, then during this time, we can generate $N = I/T_0$ repetitions. Since learning each type requires s repetitions, the total

amount of different types in which a student can get skilled is equal to $t = N/s = (I/T_0) \cdot s$. Thus, after this training, the student will acquire the skill level

$$t = \frac{I}{T_0 \cdot s}. \tag{4}$$

This formula describes the skill level acquired during a given training time I.

We can also consider the inverse problem: we want a student to be trained for a certain skill level t, and we need to know the time I required for this training. From the formula (4), we can conclude that

$$I = t \cdot T_0 \cdot s. \tag{5}$$

Conclusion: optimal training is faster and better. By comparing the formulas (2) and (4), we conclude that during the same training time, the skill level acquired during the automated training can be much higher ($\ln(T)$ times higher) that the skill level acquired in traditional training.

Similarly, by comparing the formulas (3) and (5), we conclude that the training time necessary to acquire a given skill can be much shorter ($\ln(T)$ times shorter) for the automated training than for traditional training.

How to optimally combine classroom and field training. Usually, after classroom training, a students goes into realistic situations (field training) which solidify his training. How can we best organize this combined training?

Let us denote the time that we can allocate for classroom training by I_{au}, and the training time for the follow-up field training by I_{tr}. During the follow-up training, the student encounters $N_{\text{tr}} = I_{\text{tr}}/T_0$ repetitions. Of these repetitions, $N_{\text{tr}} \cdot f_i$ are of type i.

If this number of repetitions is $\geq s$, then for this type, the student acquires necessary skills during the follow-up training, so there is no need to simulate patients of this type during the automated training. Thus, we get all types from 1 to

$$t_{\text{tr}} = \frac{I_{\text{tr}}}{T_0 \cdot \ln(T) \cdot s} \tag{6}$$

covered.

For each type $i > t_{\text{tr}}$, we get $f_i \cdot N_{\text{tr}} = \dfrac{I_{\text{tr}}}{T_0 \cdot \ln(T) \cdot i} < s$ repetitions covered during traditional training. So, if we want the student to get the necessary skills, we must generate the remaining number of repetitions

$$n_i = s - \frac{I_{\text{tr}}}{T_0 \cdot \ln(T) \cdot i} \tag{7}$$

during the automated training.

We want to learn as many new types as possible. How many situation types can we thus learn? During the time I_{au}, we can only generate $N_{\text{au}} = I_{\text{au}}/T_0$ repetitions. Since learning type i requires n_i repetitions, the skill level t acquired by

a student can be determined by the formula $\frac{I_{\text{au}}}{T_0} = N_{\text{au}} = \sum_{i=t_{\text{tr}}}^{t} n_i$. Substituting the above expression for n_i, we conclude that

$$\frac{I_{\text{au}}}{T_0} = s \cdot (t - t_{\text{tr}}) - \frac{I_{\text{tr}}}{T_0 \cdot \ln(T)} \cdot \sum_{i=t_{\text{tr}}}^{t} \frac{1}{i}.$$

Since $1 + 1/2 + \ldots + 1/i \approx \ln(i)$, we can rewrite this equation as

$$\frac{I_{\text{au}}}{T_0} = s \cdot (t - t_{\text{tr}}) - \frac{I_{\text{tr}} \cdot (\ln(t) - \ln(t_{\text{tr}})}{T_0 \cdot \ln(T)}. \tag{8}$$

So, we can make two conclusions:

- If the training times I_{au} and I_{tr} are given, then the resulting acquired skill t can be determined from the equation (8), where t_{tr} is determined from the equation (6).
- Vice versa, if we know the training time I_{au} for the classroom training, and the required skill level t, then we must find t_{tr} for the equation (8), and then use the formula (6) to determine the necessary traditional training period as $I_{\text{tr}} = t_{\text{tr}} \cdot T_0 \cdot \ln(T) \cdot s$.

In both cases, the number of repetitions of different types $i = t_{\text{tr}}, t_{\text{tr}} + 1, \ldots, t$ generated during the classroom training is determined by the formula (7).

Other applications. In [1], we used a similar idea to optimize the types of virtual patients used by doctors during medical training – specifically, during a training of surgeons for spinal cord stimulation procedures; see, e.g., [2].

2 Optimal Order of Presenting the Material

Formulation of the problem. In the above section, we described the optimal frequencies with which we repeat each of the items that a student has to learn. Once we know the number of repetitions of each item, the next natural question is: in what order should we present these repetitions? Should we first present all the repetitions of item 1, then all the repetitions of item 2, etc., or should we randomly mix these repetitions?

Towards mathematical formulation of the corresponding optimization problem. Each item is characterized by several (n) numerical characteristics, so we can geometrically represent each item as a point in the corresponding n-dimensional space.

- Similar items have close values of these characteristics, so the distance between the points corresponding to similar items is small.

– Vice versa, when the items are different, they at least some of these characteristics have different values on these items, so the resulting distance is large.

Thus, the distance between the corresponding points in a multi-D space can be viewed as a measure of similarity between the items.

In terms of multi-D space, an order in which we present repetitions is described as a function $x(t)$, where x is a multi-D point corresponding to the item presented at moment $t = k \cdot \Delta t$, where Δt is the time between repetitions.

As we have mentioned, when we have a few items to learn, we can easily learn them all, so there is no need for sophisticated optimization. Optimization becomes necessary when there are many items – and thus, many repetitions. In this case, similar to the way we simplify the physical problems if we approximate a collection of atoms by a continuous medium, we can approximate the discrete dependence $x(t)$ on discrete time t by a continuous function $x(t)$ of continuous time t.

What is the optimal trajectory $x(t)$? The experience of learning shows that often, presenting the items in random order is beneficial. To allow for this possibility, instead of looking for a deterministic function $x(t)$, we look for *random* processes $x(t)$. Since a deterministic function is a particular case of a random process, we are thus not restricting ourselves.

Let us consider Gaussian random processes. A Gaussian random process can be uniquely characterized by its mean $m(t) \stackrel{\text{def}}{=} E[x(t)]$ and autocorrelation function $A(t,s) \stackrel{\text{def}}{=} E[(x(t) - x(s))^2]$.

Students come with different levels of preparation. Therefore, a good learning strategy should work not only for a student that comes from 0, but also for a student that comes at moment t_0 with the knowledge that other students have already acquired by this time. From this viewpoint, a student's education starts at the moment t_0. It is therefore natural to require that the random process should look the same whether we start with a point $t = 0$ or with some later point t_0. Hence, the characteristics of the process should be the same, i.e., $m(t) = m(t + t_0)$ and $A(t,s) = A(t + t_0, s + t_0)$ for every t, s, and t_0.

From the first condition, we conclude that $m(t) = \text{const}$. Thus, by changing the origin of the coordinate system, we can safely assume that $m(t) = 0$.

From the second condition, for $t_0 = -s$, we conclude that $A(t,s) = A(t-s, 0)$, i.e., that the autocorrelation function depends only on the difference between the times: $A(t,s) = a(t-s)$, where we denoted $a(t) \stackrel{\text{def}}{=} A(t,0)$. In other words, the random process must be *stationary*.

The final question is: what autocorrelation function $a(t)$ should we use?

We must choose a family of functions, not a single function. The function $a(t)$ depends on how intensely we train. In more intensive training, we present the material faster, and thus, within the same time interval t, we can cover more diverse topics. More diverse topics means that the average change $a(t)$ can be larger. A natural way to describe this increase is by proportionally enlarging

all the distances, which leads from $a(t)$ to $C \cdot a(t)$. In other words, if $a(t)$ is a reasonable function for some training, then a new function $C \cdot a(t)$ should also be reasonable.

We can say that the functions $a(t)$ and $C \cdot a(t)$ describe exactly the same learning strategy, but with different intensities. Since intensity can be different, we cannot select a unique function $a(t)$ and claim it to be the best, because for every function $a(t)$, the function $C \cdot a(t)$ describes exactly the same learning strategy. In view of this, instead of formulating a problem of choosing the best autocorrelation *function*, it is more natural to formulate a problem of choosing the best *family* $\{C \cdot a(t)\}_C$ of autocorrelation functions.

Which family is the best? We may need non-numerical optimality criteria. Among all the families $\{C \cdot a(t)\}_C$, we want to choose the best one.

In mathematical optimization problems, numerical criteria are most frequently used, when to every alternative (in our case, to each family) we assign some value expressing its performance, and we choose an alternative (in our case, a family) for which this value is the largest. In our problem, as such a numerical criterion, we can select, e.g., the average grade on some standardized test A.

However, it is not necessary to restrict ourselves to such numerical criteria only. For example, if we have several different families that have the same average average grade A, we can choose between them the one that has the minimal level of uncomfortablness U. In this case, the actual criterion that we use to compare two families is not numerical, but more complicated: *a family F_1 is better than the family F_2 if and only if either $A(F_1) < A(F_2)$, or $A(F_1) = A(F_2)$ and $U(F_1) < U(F_2)$*. A criterion can be even more complicated. What a criterion *must* do is to allow us, for every pair of families, to tell whether the first family is better with respect to this criterion (we'll denote it by $F_1 \succ F_2$), or the second is better ($F_1 \prec F_2$), or these families have the same quality in the sense of this criterion (we'll denote it by $F_1 \sim F_2$). Of course, it is necessary to demand that these choices be consistent, e.g., if $F_1 \prec F_2$ and $F_2 \prec F_3$ then $F_1 \prec F_3$.

The optimality criterion must select a unique optimal family. Another natural demand is that this criterion must choose a *unique* optimal family (i.e., a family that is better with respect to this criterion than any other family). The reason for this demand is very simple.

If a criterion does not choose a family at all, then it is of no use.

If several different families are "the best" according to this criterion, then we still have a problem to choose among those "best". Therefore, we need some additional criterion for that choice. For example, if several families turn out to have the same average grade, we can choose among them a with the minimal uncomfortableness.

So what we actually do in this case is abandon that criterion for which there were several "best" families, and consider a new "composite" criterion instead: F_1 is better than F_2 according to this new criterion if either it was better according to the old criterion or according to the old criterion they had the same quality and F_1 is better than F_2 according to the additional criterion.

In other words, if a criterion does not allow us to choose a unique best family, it means that this criterion is not final. We have to modify it until we come to a final criterion that will have that property.

The optimality criterion must be scale-invariant. The next natural condition that the criterion must satisfy is connected with the fact that the numerical value of the time t depends on the choice of the unit for measuring time.

If we replace the original unit of time by a new unit which is λ times larger (i.e., replace minutes by hours), then numerical values change from t to $\widetilde{t} = t/\lambda$. The autocorrelation function that in the old units is described by a family $\{C \cdot a(t)\}$, in the new units, has a new form $\{C \cdot a(\lambda \cdot t)\}$,

Since this change is simply a change in a unit of time, it is reasonable to require that going from $a(t)$ from $a(\lambda \cdot t)$ should not change the *relative* quality of the autocorrelation functions, i.e., if a family $\{C \cdot a(t)\}_C$ is better that the family $\{C \cdot a'(t)\}_C$, then for every $\lambda > 0$, the family $\{C \cdot a(\lambda \cdot t)\}_C$ must be still better than the family $\{C \cdot a'(\lambda \cdot t)\}_C$.

Definitions and the main result.

Definition 1. – *By an* autocorrelation function *we mean a monotonically non-strictly decreasing function from non-negative real numbers to non-negative real numbers.*

– *By a* family *of functions we mean the family $\{C \cdot a(t)\}_C$, where $a(t)$ is a given autocorrelation function and C runs over arbitrary positive real numbers.*

– *A pair of relations $(\prec, \sim)$ is called* consistent [4] *if it satisfies the following conditions:*
  (1) *if $a \prec b$ and $b \prec c$ then $a \prec c$;*
  (2) *$a \sim a$;*
  (3) *if $a \sim b$ then $b \sim a$;*
  (4) *if $a \sim b$ and $b \sim c$ then $a \sim c$;*
  (5) *if $a \prec b$ and $b \sim c$ then $a \prec c$;*
  (6) *if $a \sim b$ and $b \prec c$ then $a \prec c$;*
  (7) *if $a \prec b$, then $b \prec a$ or $a \sim b$ are impossible.*

Definition 2. – *Assume a set A is given. Its elements will be called* alternatives. *By an* optimality criterion *we mean a consistent pair $(\prec, \sim)$ of relations on the set A of all alternatives. If $b \prec a$, we say that a is* better *than b; if $a \sim b$, we say that the alternatives a and b are* equivalent *with respect to this criterion.*

– *We say that an alternative a is* optimal *(or* best*) with respect to a criterion $(\prec, \sim)$ if for every other alternative b either $b \prec a$ or $a \sim b$.*

– *We say that a criterion is* final *if there exists an optimal alternative, and this optimal alternative is unique.*

– *Let $\lambda > 0$ be a real number. By the* λ-rescaling $R_\lambda(\rho)$ *of a function $a(t)$ we mean a function $(R_\lambda a)(t) \stackrel{\text{def}}{=} a(\lambda \cdot t)$.*

– *By the* λ-rescaling $R_\lambda(F)$ *of a family F, we mean the set of the functions that are obtained from $f \in F$ by λ-rescaling.*

In this paper, we consider optimality criteria on the set $\mathcal{F}$ of all families.

Definition 3. *We say that an optimality criterion on F is* scale-invariant *if for every two families F and G and for every number $\lambda > 0$, the following two conditions are true:*

i) *if F is better than G in the sense of this criterion (i.e., $G \prec F$), then*

$$R_\lambda(G) \prec R_\lambda(F);$$

ii) *if F is equivalent to G in the sense of this criterion (i.e., $F \sim G$), then*

$$R_\lambda(F) \sim R_\lambda(G).$$

As we have already remarked, the demands that the optimality criterion is final and scale-invariant are quite reasonable. The only problem with them is that at first glance they may seem rather weak. However, they are not, as the following theorem shows:

Theorem 1. *[4] If a family F is optimal in the sense of some optimality criterion that is final and scale-invariant, then every function $a(t)$ from this optimal family F which has the form $a(t) = A \cdot t^\alpha$ for some real numbers A and α.*

In other words, the optimal configuration is a fractal random process. When $\alpha = 2$, we have a straightforward trajectory, without any randomness. The value $\alpha = 0$ means that values of $x(t)$ and $x(s)$ for $t \neq s$ are completely uncorrelated, i.e., that we have a white noise. Intermediate values of α correspond to different levels of randomness.

Our experience showed that such fractal order indeed leads to improvement in learning. The exact value of the parameter α – corresponding to the fractal dimension of the corresponding trajectories – should be adjusted to the learning style of the students.

References

1. Aló, R., Aló, K., Kreinovich, V.: Towards Intelligent Virtual Environment for Training Medical Doctors in Surgical Pain Relief, *Proceedings of The Eighth International Fuzzy Systems Association World Congress IFSA'99*, Taipei, Taiwan, August 17–20, 1999, pp. 260–264.
2. Horsch, S., Clayes, L.: *Spinal Cord Stimulation II*, Steinkopff Verlag, Darmstadt, 1995.
3. Mandelbrot, B. B.: *The fractal geometry of Nature*, Freeman, San Francisco, 1982.
4. Nguyen, H. T., Kreinovich, V.: *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.
5. Zipf, G. K.: *Human behavior and the principle of least-effort*, Addison-Wesley, Cambridge, MA, 1949.

Pricing Derivatives for Some Levy Process with Fuzzy Pattern

Piotr Nowak[1] and Maciej Romaniuk[2]

[1] Systems Research Institute Polish Academy of Sciences, ul. Newelska 6, 01–447 Warszawa, Poland
pnowak@ibspan.waw.pl
[2] Systems Research Institute Polish Academy of Sciences, ul. Newelska 6, 01–447 Warszawa, Poland
mroman@ibspan.waw.pl

Abstract. In the following paper we propose stochastic model of underlying asset trajectory based on some subclass of Levy processes. Because of the fluctuation of financial market from time to time, some input parameters of the financial models cannot always be expected in the precise sense. Therefore we apply the fuzzy number theory for the proposed process. We also use Levy characteristics and discuss the possibility of application of Monte Carlo simulations in order to price the derivatives based on our model. In detail, we present pricing formula for classical example of European call option.

1 Introduction

The Black-Scholes formula is a very important and one of the best known results in financial mathematics (see [4,5]). It enables to calculate the price of a derivative under the assumption that there is not a possibility of arbitrage in the market. It is assumed in the Black-Scholes model that the stochastic process which describes the price movements of an underlying financial instrument S is the geometrical Brownian motion. The advantage of this assumption is a simplicity of the pricing formula (see [4,5]). However, it is commonly known that the Black-Scholes formula does not describe the real behaviour of derivatives very well. Such an example is non-symmetric distribution of the standardized returns of logarithms of S in the real market (see [1]). Our paper is devoted to improve the pricing model. As the domain of our explorations we choose the exponential function of a linear combinations of Brownian motion, drift and Poisson process. The set of exponents of this type is a natural subclass of Levy processes. It contains non-continuous processes which may model jumps of the underlying asset. In our paper we would like to join two approaches. The first one is the martingale method based on Levy characteristics of the underlying asset process (see [8]). In detail we explore the possibility of applying the minimal entropy martingale measure, which is closely related to exponential utility function approaches, known in e.g. economy and finance. The second way to deal with the main problem is to apply the fuzzy sets theory to described

process of the underlying asset (see [9]). Because of imprecise information and the fluctuation of financial market from time to time, some input parameters of our model cannot always be expected in the precise sense. Therefore it is natural to consider fuzzy parameters of the stochastic model. Such a treatment is particularly useful for financial analysts giving them the tool to pick option price with an acceptable belief degree for the latter use. Our considerations will be mainly focused on European options, however they may be generalized for other types of derivatives. We will consider the stochastic model with fuzzy interest rate.

The paper is organized as follows. Section 2 contains preliminaries. Section 3 is an exposition of basic definitions and facts concerning the minimal entropy martingale measure. In Section 4 we prove the pricing formula for European call option in the crisp case. Section 5 is devoted to pricing the derivative with a fuzzy interest rate. The last section contains concluding remarks.

2 Preliminaries

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P\right)$ be a probability space with filtration satisfying standard assumptions. Let $T < \infty$.

A stochastic process $H = (H_t)_{t\in[0,T]}$ is cadlag, if its trajectories are functions, which are right-continuous with left limits.

H is $(\mathcal{F}_t)$-adapted if H_t is $(\mathcal{F}_t)$-measurable for each $t \in [0,T]$.

A probability measure Q on $(\Omega, \mathcal{F})$ is absolutely continuous with respect to P ($Q << P$) if for all $A \in \mathcal{F}$

$$P(A) = 0 \Rightarrow Q(A) = 0$$

and it is equivalent to P if P and Q have the same sets with zero measure.

Let $\mathcal{S}_t$ be an $(\mathcal{F}_t)$-adapted cadlag stochastic process describing the underlying asset. Let r denote a constant risk–free interest rate and

$$\mathcal{Z}_t = e^{-rt}\mathcal{S}_t \tag{1}$$

be the discounted process of values of the underlying asset. We have to find the measure $\tilde{P}$ equivalent to P for which $\mathcal{Z}_t$ is a martingale. Throughout this paper the measure $\tilde{P}$ will be called the equivalent martingale measure.

The next step is to find the form of the process $\mathcal{S}_t$ with respect to this new probability measure $\tilde{P}$. The price of a derivative with a payment function f is given by formula:

$$C_t = \exp\left(-r(T-t)\right)\mathbb{E}^{\tilde{P}}\left(f(S)\,|\mathcal{F}_t\right),\ t \in [0,T]\,. \tag{2}$$

Next we remind some facts about fuzzy sets and numbers. Let X be a universal set and $\tilde{A}$ be a fuzzy subset of X. We denote by $\mu_{\tilde{A}}$ its membership function $\mu_{\tilde{A}} : X \to [0,1]$, and by $\tilde{A}_\alpha = \{x : \mu_{\tilde{A}} \geq \alpha\}$ the α-level set of $\tilde{A}$, where $\tilde{A}_0$ is the closure of the set $\{x : \mu_{\tilde{A}} \neq 0\}$. In our paper we assume that $X = \mathbb{R}$.

Let $\tilde{a}$ be a fuzzy number. Then, under our assumptions, the α-level set $\tilde{a}_\alpha$ is a closed interval, which can be denoted by $\tilde{a}_\alpha = [\tilde{a}_\alpha^L, \tilde{a}_\alpha^U]$ (see e.g. [9]).

We can now introduce the arithmetics of any two fuzzy numbers. Let $\odot$ be a binary operator $\oplus$, $\ominus$, $\otimes$ or $\oslash$ between fuzzy numbers $\tilde{a}$ and $\tilde{b}$, where the binary operators correspond to $\circ$: $+$, $-$, $\times$ or $/$, according to the "Extension Principle" in [9]. Then the membership function of $\tilde{a} \odot \tilde{b}$ is defined by

$$\mu_{\tilde{a}\odot\tilde{b}}(z) = \sup_{(x,y):x\circ y=z} \min\{\mu_{\tilde{a}}(x), \mu_{\tilde{b}}(y)\} \ . \tag{3}$$

Let $\odot_{\text{int}}$ be a binary operator $\oplus_{\text{int}}$, $\ominus_{\text{int}}$, $\otimes_{\text{int}}$ or $\oslash_{\text{int}}$ between two closed intervals $[a, b]$ and $[c, d]$. Then

$$[a, b] \odot_{\text{int}} [c, d] = \{z \in \mathbb{R} : z = x \circ y, \forall x \in [a, b], \forall y \in [c, d]\} \ , \tag{4}$$

where $\circ$ is an usual operation $+, -, \times$ and $/$, if the interval $[c, d]$ does not contain zero in the last case.

Therefore, if $\tilde{a}, \tilde{b}$ are fuzzy numbers, then $\tilde{a} \odot \tilde{b}$ is also the fuzzy number and its α-level set is given by

$$\begin{gathered}
(\tilde{a} \oplus \tilde{b})_\alpha = \tilde{a}_\alpha \oplus_{\text{int}} \tilde{b}_\alpha = [\tilde{a}_\alpha^L + \tilde{b}_\alpha^L, \tilde{a}_\alpha^U + \tilde{b}_\alpha^U] \ , \\
(\tilde{a} \ominus \tilde{b})_\alpha = \tilde{a}_\alpha \ominus_{\text{int}} \tilde{b}_\alpha = [\tilde{a}_\alpha^L - \tilde{b}_\alpha^U, \tilde{a}_\alpha^U - \tilde{b}_\alpha^L] \ , \\
(\tilde{a} \otimes \tilde{b})_\alpha = \tilde{a}_\alpha \otimes_{\text{int}} \tilde{b}_\alpha = \\
= [\min\{\tilde{a}_\alpha^L\tilde{b}_\alpha^L, \tilde{a}_\alpha^L\tilde{b}_\alpha^U, \tilde{a}_\alpha^U\tilde{b}_\alpha^L, \tilde{a}_\alpha^U\tilde{b}_\alpha^U\}, \max\{\tilde{a}_\alpha^L\tilde{b}_\alpha^L, \tilde{a}_\alpha^L\tilde{b}_\alpha^U, \tilde{a}_\alpha^U\tilde{b}_\alpha^L, \tilde{a}_\alpha^U\tilde{b}_\alpha^U\}] \ , \\
(\tilde{a} \oslash \tilde{b})_\alpha = \tilde{a}_\alpha \oslash_{\text{int}} \tilde{b}_\alpha = \\
= [\min\{\tilde{a}_\alpha^L/\tilde{b}_\alpha^L, \tilde{a}_\alpha^L/\tilde{b}_\alpha^U, \tilde{a}_\alpha^U/\tilde{b}_\alpha^L, \tilde{a}_\alpha^U/\tilde{b}_\alpha^U\}, \max\{\tilde{a}_\alpha^L/\tilde{b}_\alpha^L, \tilde{a}_\alpha^L/\tilde{b}_\alpha^U, \tilde{a}_\alpha^U/\tilde{b}_\alpha^L, \tilde{a}_\alpha^U/\tilde{b}_\alpha^U\}] \ ,
\end{gathered}$$

if α-level set $\tilde{b}_\alpha$ does not contain zero for all $\alpha \in [0, 1]$ in the case of $\oslash$.

Proposition 1 *Let $f : \mathbb{R} \to \mathbb{R}$ be a function, $\mathcal{F}(\mathbb{R})$ be a set of all fuzzy subsets of $\mathbb{R}$ and $\tilde{\Lambda} \in \mathcal{F}(\mathbb{R})$. We assume that the membership function $\mu_{\tilde{\Lambda}}$ of $\tilde{\Lambda}$ is upper semicontinuous and for all y the set $\{x : f(x) = y\}$ is compact. The function $f(x)$ can induce a fuzzy-valued function $\tilde{f} : \mathcal{F}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$ via the extension principle. Then the α-level set of $\tilde{f}(\tilde{\Lambda})$ is $\tilde{f}(\tilde{\Lambda})_\alpha = \{f(x) : x \in \tilde{\Lambda}_\alpha\}$.*

3 Minimal entropy martingale measure for Levy processes

The cadlag stochastically continuous process $Y = (Y_t)_{t\in[0,T]}$, $Y_0 = 0$ a.s., is called a Levy process if it satisfies the following conditions.

1. $Y_t - Y_s$ is independent of $\mathcal{F}_s$ for all $0 \le s \le t \le T$.
2. $Y_t - Y_s$ and Y_{t-s} have the same distributions for all $0 \le s \le t \le T$.

We assume that a truncation function φ is defined by the formula $\varphi(x) = xI_{|x|\leq 1}$. We denote by $\mathcal{M}(\mathbb{R})$ the space of non-negative measures on $\mathbb{R}$. For Levy processes local characteristics, called Levy characteristics are defined. They are functions of the following form

$$B_t : [0,T] \to \mathbb{R}, \quad B_t = bt,$$

$$C_t : [0,T] \to \mathbb{R}, \quad C_t = ct,$$

$$\nu_t : [0,T] \to \mathcal{M}(\mathbb{R}), \quad \nu_t(dx) = \nu(dx)\, t,$$

$$\nu(\{0\}) = 0, \ \int_{\mathbb{R}} \left(|x|^2 \wedge 1\right) \nu(dx)$$

where $b, c \in \mathbb{R}$ and $\nu \in \mathcal{M}(\mathbb{R})$. Moreover, only constant b depends on the form of φ.

A stochastic process $S = (S_t)_{t\in[0,T]}$ is called the geometric Levy process, if it can be written in the following form

$$S_t = S_0 \exp(Y_t), \quad t \in [0,T] \tag{5}$$

where Y_t is a Levy process.

Throughout this paper we assume that S is of the form (5), $\mathcal{F} = \mathcal{F}_T$ and for $t \in [0,T]$

$$\mathcal{F}_t = \sigma(S_s, s \in [0,t]) = \sigma(Y_s, s \in [0,t]).$$

The relative entropy $I(Q,P)$ of Q with respect to P is defined by

$$I(Q,P) = \begin{cases} E_P\left(\frac{dQ}{dP} \ln\left(\frac{dQ}{dP}\right)\right) & \text{if } Q << P \\ +\infty & \text{otherwise.} \end{cases}$$

If an equivalent martingale measure P^* satisfies the inequality

$$I(P^*, P) \leq I(Q,P)$$

for all equivalent martingale measures Q is called the minimal entropy martingale measure (MEMM).

Let

$$g^{(MEMM)}(\theta) = b + \left(\frac{1}{2} + \theta\right) c + \int_{\{|x|\leq 1\}} \left(\left(e^x - 1\right) e^{\theta(e^x-1)} - x\right) \nu(dx)$$
$$+ \int_{\{|x|>1\}} \left(e^x - 1\right) e^{\theta(e^x-1)} \nu(dx)$$

Let $\tilde{Y} = \left(\tilde{Y}_t\right)_{t\in[0,T]}$ be the Levy process corresponding to the original Levy process Y, such that the following equality holds

$$S_t = S_0 \exp(Y_t) = S_0 \mathcal{E}\left(\tilde{Y}_t\right), \quad t \in [0,T],$$

where $\mathcal{E}\left(\tilde{Y}_t\right)$ is Doleans-Dade exponential of $\tilde{Y}_t$.

Let $P^{(ESS)}_{\tilde{Y}_{[0,T]},\theta^*}$ and $P^{(ESS)}_{\tilde{Y}_T,\theta^*}$ be the Esscher transformed measures of P (see [6]).

The following theorem from [6] will be useful in our paper.

Theorem 1. *If the equation*

$$g^{(MEMM)}(\theta) = r \tag{6}$$

has a solution θ^*, then the MEMM of S, P^* exsists and

$$P^* = P^{(MEMM)} = P^{(ESS)}_{\tilde{Y}_{[0,T]},\theta^*} = P^{(ESS)}_{\tilde{Y}_T,\theta^*}.$$

The process Y is also a Levy process under P^* and the generating triplet of Y under P^*, (b^*, c^*, ν^*) is

$$b^* = b + \theta^* c + \int_{\{|x|\leq 1\}} x\left(e^{\theta^*(e^x-1)} - 1\right)\nu(dx),$$

$$c^* = c,$$

$$\nu^*(dx) = e^{\theta^*(e^x-1)}\nu(dx).$$

4 The model of the underlying asset and pricing formula

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P\right)$ be a probability space with filtration. Let $T < \infty$.

The price of the underlying asset S_t is the geometric Levy process given by (5), where

$$Y_t = \mu t + \sigma W_t + kN^\kappa_t, \tag{7}$$

W_t is Brownian motion, $\sigma > 0$, $\mu, k \in \mathbb{R}$ and N^κ_t is Poisson process with the intensity $\kappa > 0$. We assume that W_t and N^κ_t are independent.

Theorem 2. *The price of European call option with strike price K and payment function $f(x) = (x-K)_+$ at time 0 is given by formula*

$$C_0 = e^{-\kappa^* T}\sum_{n=0}^{\infty}\frac{(\kappa^* T)^n}{n!}\left(S_0 e^{(\mu^*-r)T+\frac{\sigma^2 T}{2}+kn}\Phi\left(d^n_+\right) - e^{-rT}K\Phi\left(d^n_-\right)\right),$$

where Φ is cumulative distribution function of standard normal distribution and θ^* is the solution of the equation

$$\mu + \left(\frac{1}{2}+\theta\right)\sigma^2 + \kappa\left(e^k - 1\right)e^{\theta\left(e^k-1\right)} = r, \tag{8}$$

$$\mu^* = \mu + \theta^*\sigma^2\ ,\ \kappa^* = \kappa e^{\theta^*\left(e^k-1\right)},$$

$$d^n_- = \frac{\ln\frac{S_0}{K} + \mu^* T + kn}{\sigma\sqrt{T}},\ \ d^n_+ = \frac{\ln\frac{S_0}{K} + \mu^* T + \sigma^2 T + kn}{\sigma\sqrt{T}}.$$

Proof. We apply Theorem 1 to price the option. Equation (6) has the form (8). Since $\lim_{\theta\to-\infty} g(\theta) = -\infty$ and $\lim_{\theta\to\infty} g(\theta) = \infty$, the above equation has a solution. We denote it by θ^*. According to Theorem 1, Y with respect to P^* has the form

$$Y_t = \left(\mu + \theta^*\sigma^2\right) t + \sigma W_t^* + kN_t^{\kappa e^{\theta^*\left(e^k-1\right)}}, \; t \in [0,T], \tag{9}$$

where W^* is a Brownian motion and $N_t^{\kappa e^{\theta^*\left(e^k-1\right)}}$ is a Poisson process with respect to P^* and the processes are independent. The price of the derivative is given by formula.

$$\begin{aligned}
C_0 &= e^{-rT}\mathbb{E}^{P^*}(S_T-K)^+ = e^{-rT}\mathbb{E}^{P^*}(S_T-K)\, I_{\{S_T>K\}} \\
&= e^{-rT}\mathbb{E}^{P^*}\left(S_0 e^{\mu^* T+\sigma W_T^*+kN_T^{\kappa^*}} - K\right) I_{\left\{e^{\mu^* T+\sigma W_T^*+kN_T^{\kappa^*}}>\frac{K}{S_0}\right\}} \\
&= e^{-rT}\mathbb{E}^{P^*}\sum_{n=0}^{\infty} I_{\left\{N_T^{\kappa^*}=n\right\}}\left(S_0 e^{\mu^* T+\sigma W_T^*+kn} - K\right) I_{\left\{\mu^* T+\sigma W_T^*+kn>\ln\frac{K}{S_0}\right\}} \\
&= e^{-(\kappa^*+r)T}\sum_{n=0}^{\infty}\frac{(\kappa^* T)^n}{n!}\mathbb{E}^{P^*}\left(S_0 e^{\mu^* T+\sigma W_T^*+kn} - K\right) I_{\left\{\mu^* T+\sigma W_T^*+kn>\ln\frac{K}{S_0}\right\}}
\end{aligned}$$

Since

$$(\mu^*-r)\,T + \sigma W_T^* + kn \sim N\left((\mu^*-r)\,T + kn, \sigma\sqrt{T}\right),$$

$$\begin{aligned}
I &= e^{-rT}\mathbb{E}^{P^*}\left(S_0 e^{\mu^* T+\sigma W_T^*+kn} - K\right) I_{\left\{\mu^* T+\sigma W_T^*+kn>\ln\frac{K}{S_0}\right\}} \\
&= \frac{1}{\sqrt{2\pi\sigma^2 T}}\int_{\ln\frac{K}{S_0}-rT}^{\infty}\left(S_0 e^x - e^{-rT}K\right) e^{-\frac{[x-((\mu^*-r)T+kn)]^2}{2\sigma^2 T}}\,dx.
\end{aligned}$$

We apply standard integral operations and we obtain

$$\begin{aligned}
I &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-^n}\left(S_0 e^{-y\sigma\sqrt{T}+(\mu^*-r)T+kn} - e^{-rT}K\right) e^{-\frac{y^2}{2}}\,dy \\
&= S_0\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-^n} e^{-y\sigma\sqrt{T}+(\mu^*-r)T+kn} e^{-\frac{y^2}{2}}\,dy - e^{-rT}K\Phi(d_-) \\
&= S_0\frac{1}{\sqrt{2\pi}} e^{(\mu^*-r)T+kn} e^{-\frac{y^2}{2}}\int_{-\infty}^{d_-^n} e^{-\frac{2y\sigma\sqrt{T}+y^2}{2}}\,dy - e^{-rT}K\Phi(d_-) \\
&= S_0 e^{(\mu^*-r)T+\frac{\sigma^2 T}{2}+kn}\Phi\left(d_+^n\right) - e^{-rT}K\Phi\left(d_-^n\right)
\end{aligned}$$

Finally,

$$C_0 = e^{-\kappa^* T}\sum_{n=0}^{\infty}\frac{(\kappa^* T)^n}{n!}\left(S_0 e^{(\mu^*-r)T+\frac{\sigma^2 T}{2}+kn}\Phi\left(d_+^n\right) - e^{-rT}K\Phi\left(d_-^n\right)\right).$$

It is easy to prove that the above series is convergent. From the Taylor expansion of the function exp, the form of its remainder term and from the equality

$$C_0 = S_0 e^{\left(\mu^* + \frac{\sigma^2}{2} - r - \kappa^*\right)T} \sum_{n=0}^{\infty} \frac{\left(\kappa^* e^k T\right)^n}{n!} \Phi\left(d_+^n\right) - K e^{-(\kappa^* - r)T} \sum_{n=0}^{\infty} \frac{\left(\kappa^* T\right)^n}{n!} \Phi\left(d_-^n\right)$$

the following condition holds.

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N}$$
$$e^{-\kappa^* T} \left| \sum_{n=N+1}^{\infty} \frac{\left(\kappa^* T\right)^n}{n!} \left(S_0 e^{(\mu^* - r)T + \frac{\sigma^2 T}{2} + kn} \Phi\left(d_+^n\right) - e^{-rT} K \Phi\left(d_-^n\right) \right) \right| < \varepsilon.$$

Therefore in our further considerations we will replace the pracing formula by

$$\sum_{n=1}^{N} \frac{\left(\kappa^* T\right)^n}{n!} \left(S_0 e^{(\mu^* - r)T + \frac{\sigma^2 T}{2} + kn} \Phi\left(d_+^n\right) - e^{-rT} K \Phi\left(d_-^n\right) \right), \tag{10}$$

for sufficiently large N.

5 Pricing formula for fuzzy interest rate

In this section we will use the symbol ˜ above model parameters to indicate they are fuzzy. Parameters without ˜ are crisp.

We assume that the constant interest rate r is not known precisely. Therefore we will treat the interest rate as a fuzzy number $\tilde{r}$. Then the pricing formula (10) has the form

$$\tilde{C}_0 = e^{-\tilde{\kappa}^* \otimes T} \otimes \bigoplus_{n=0}^{N} \left(\left(\tilde{\kappa}^* \otimes T\right)^n \oslash n! \right) \otimes \Big(S_0 \otimes e^{[((\tilde{\mu}^* \ominus \tilde{r}) \otimes T) \oplus (\sigma \otimes \sigma \otimes T \oslash 2) \oplus (k \otimes n)]} \otimes \tilde{\Phi}\left(\tilde{d}_+^n\right) \ominus \left(e^{-\tilde{r} \otimes T} \otimes K \otimes \tilde{\Phi}\left(\tilde{d}_-^n\right) \right) \Big), \tag{11}$$

where $\tilde{\theta}^*$ is the solution of the equation

$$\mu \oplus \left(\left(\frac{1}{2} \oplus \tilde{\theta} \right) \otimes \sigma \otimes \sigma \right) \oplus \left(\tilde{\kappa} \otimes \left(\left(e^k \ominus 1\right) \otimes e^{\tilde{\theta} \otimes \left(e^k \ominus 1\right)} \right) \right) = \tilde{r}, \tag{12}$$

for

$$\tilde{\mu}^* = \mu \oplus \left(\tilde{\theta}^* \otimes \sigma \otimes \sigma \right), \quad \kappa^* = \kappa \otimes e^{\tilde{\theta}^* \otimes \left(e^k \ominus 1\right)},$$

$$\tilde{d}_-^n = \left[\ln\left(S_0 \oslash K\right) \oplus \left(\tilde{\mu}^* \otimes T\right) \oplus \left(k \otimes n\right) \right] \oslash \left(\sigma \otimes \sqrt{T} \right) \text{ and}$$
$$\tilde{d}_+^n = \left[\ln\left(S_0 \oslash K\right) \oplus \left(\mu^* \otimes T\right) \oplus \left(\sigma \otimes \sigma \otimes T\right) \oplus \left(k \otimes n\right) \right] \oslash \left(\sigma \otimes \sqrt{T} \right).$$

We calculate the α-level set of $\tilde{C}_0$ using Proposition 1.
$\left(\tilde{C}_0\right)_\alpha = \left[\tilde{C}^L_{0\alpha}, \tilde{C}^U_{0\alpha}\right]$, where

$$\tilde{C}^L_{0\alpha} = \min\left(e^{-\tilde{\kappa}^{*U}_\alpha T}\tilde{V}^L_\alpha, e^{-\tilde{\kappa}^{*L}_\alpha T}\tilde{V}^L_\alpha\right), \;\; \tilde{C}^U_{0\alpha} = \max\left(e^{-\tilde{\kappa}^{*U}_\alpha T}\tilde{V}^U_\alpha, e^{-\tilde{\kappa}^{*L}_\alpha T}\tilde{V}^U_\alpha\right),$$

$$\tilde{V}^L_\alpha = \sum_{i=1}^{N}\left[S_0 e^{\left(\tilde{\mu}^{*L}_\alpha - \tilde{r}^{*U}_\alpha\right)T + \frac{\sigma^2 T}{2} + kn}\Phi\left(\tilde{d}^{nL}_{+\alpha}\right) - e^{-\tilde{r}^{*L}_\alpha T}K\Phi\left(\tilde{d}^{nU}_{-\alpha}\right)\right],$$
$$\tilde{V}^U_\alpha = \sum_{i=1}^{N}\left[S_0 e^{\left(\tilde{\mu}^{*U}_\alpha - \tilde{r}^{*L}_\alpha\right)T + \frac{\sigma^2 T}{2} + kn}\Phi\left(\tilde{d}^{nU}_{+\alpha}\right) - e^{-\tilde{r}^{*U}_\alpha T}K\Phi\left(\tilde{d}^{nL}_{-\alpha}\right)\right],$$

$$\tilde{\mu}^{*L}_\alpha = \mu + \sigma^2\tilde{\theta}^{*L}_\alpha, \quad \tilde{\mu}^{*U}_\alpha = \mu + \sigma^2\tilde{\theta}^{*U}_\alpha,$$

$$\tilde{d}^{nL}_{-\alpha} = \frac{\left(\ln\frac{S_0}{K} + \tilde{\mu}^{*L}_\alpha T + kn\right)}{\sigma\sqrt{T}}, \;\; \tilde{d}^{nU}_{-\alpha} = \frac{\left(\ln\frac{S_0}{K} + \tilde{\mu}^{*U}_\alpha T + kn\right)}{\sigma\sqrt{T}},$$
$$\tilde{d}^{nL}_{+\alpha} = \frac{\left(\ln\frac{S_0}{K} + \tilde{\mu}^{*L}_\alpha T + \sigma^2 T + kn\right)}{\sigma\sqrt{T}}, \;\; \tilde{d}^{nU}_{+\alpha} = \frac{\left(\ln\frac{S_0}{K} + \tilde{\mu}^{*U}_\alpha T + \sigma^2 T + kn\right)}{\sigma\sqrt{T}},$$

$$\tilde{\kappa}^{*L}_\alpha = \begin{cases} \kappa e^{\tilde{\theta}^{*L}_\alpha\left(e^k - 1\right)} \text{ for } k \geq 0 \\ \kappa e^{\tilde{\theta}^{*U}_\alpha\left(e^k - 1\right)} \text{ for } k < 0 \end{cases},$$

$$\tilde{\kappa}^{*U}_\alpha = \begin{cases} \kappa e^{\tilde{\theta}^{*U}_\alpha\left(e^k - 1\right)} \text{ for } k \geq 0 \\ \kappa e^{\tilde{\theta}^{*L}_\alpha\left(e^k - 1\right)} \text{ for } k < 0 \end{cases}$$

and $\tilde{\theta}^{*L}_\alpha$ and $\tilde{\theta}^{*U}_\alpha$ are solutions of (8) for $r = \tilde{r}^{*L}_\alpha$ and $\tilde{r}^{*U}_\alpha$ respectively. From the equality

$$\mu_{\tilde{C}_0}(c) = \sup_{0\leq\alpha\leq 1} I_{\left(\tilde{C}_0\right)_\alpha}(c)$$

we obtain the membership function of $\tilde{C}_0$. If c is the option price and $\mu_{\tilde{C}_0}(c) = \alpha$, then the value of the membership function may be treated by a financial analyst as the belief degree of c.

6 Concluding remarks

It is easily seen, that even the pricing fuzzy formula for classical European call option (11) looks complicated, but for lower values of N may be calculated. However, for more complex derivatives it may be impossible to obtain even closed form for their price. But the equation giving new martingale measure (12) derived from its crisp counterpart (8) still remains true. It gives us the possibility to calculate for the given α the appropriate α-level interval $[\tilde{\theta}^{*L}_\alpha, \tilde{\theta}^{*U}_\alpha]$, and to generate simulations of underlying asset trajectory for all $\theta^* \in [\tilde{\theta}^{*L}_\alpha, \tilde{\theta}^{*U}_\alpha]$. These simulations may be derived from the iterative version of process (9), which

analogous for Black-Scholes formula is known as Euler scheme (see [3,7]). In our case we have

$$S_{t_{i+1}} = S_{t_i} \exp\left(\left(\mu + \theta^* \sigma^2\right) \delta t + \sigma\sqrt{\delta t}\epsilon_i + k\nu_i\right) \ ,$$

where $i = 0, 1, \ldots, m$ for $t_{i_0} = 0$ and $t_{i_m} = T$, $\delta t = t_{i+1} - t_i = \text{const}$, $\epsilon_0, \ldots \epsilon_{m-1}$ are *iid* random variables from standard normal distribution and $\nu_0, \ldots \nu_{m-1}$ are *iid* realizations from Poisson process with intensity $\kappa\delta t e^{\theta^*\left(e^k - 1\right)}$ (compare with (9)). Hence, we can simulate steps $S_0, S_{t_1}, \ldots, S_{t_m}$ and then generate appropriate bunch of trajectories. These data could approximate the shape of $\tilde{C}_0$ for given kind of derivative. Such an approach could be also used instead of the expansion in (11) for the case of European call option.

It is worth of noting, that from a practical point of view, the α-level set for fuzzy price like in (11) may be seen as the interval of prices which has belief degree α for the financial analysts. Therefore, the analysts could find the interval of prices for their desired α, e.g. $\alpha = 1$ or $\alpha = 0.95$, which are comfortable for them.

In the paper we discussed some stochastic model from subclass of Levy processes. After application of the martingale method and fuzzy sets theory we obtained the pricing formula for classical example, i.e. European call option. We also discussed generalization of the obtained formulas for other kind of derivatives via simulations. Further possible studies include fuzzification of other parameters of the presented model and more detailed studies on more complex derivatives.

References

1. Davis, M.: Mathematics of financial markets. In: Engquist, B., Schmid, W.: Mathematics Unlimited — 2001 & Beyond. Springer, Berlin (2001)
2. Duffie, D., Glynn, P.: Efficient Monte Carlo Simulation for Security Prices. Annals of Applied Probability **5** (1997)
3. Glasserman, P.: Monte Carlo Methods in Financial Engineering. Springer-Verlag (2004)
4. Hull, J.C.: Options, Futures and Other Derivatives. Prentice Hall (1997)
5. Korn, R., Korn, E.: Option Pricing and Portfolio Optimization. American Mathematical Society (2001)
6. Miyahara, Y.: A Note on Esscher Transformed Martingale Measures for Geometric Levy Processes. Discussion Papers in Economics, Nagoya City University **379** (2004) 1-14
7. Nowak, P., Nycz, P., Romaniuk, M.: Dobór optymalnego modelu stochastycznego w wycenie opcji metodami Monte Carlo (in Polish). In: Kacprzyk, J., Weglarz, J. (eds.): Modelowanie i optymalizacja — Metody i zastosowania. Exit (2002)
8. Shiryaev, A.N., Kruzhilin, N.: Essential of Stochastic Finance. World Scientific Publishing Co. Pte. Ltd. (1999/2000)
9. Zadeh, L.A.: Fuzzy sets. Information and Control **8** (1965) 338 – 353

A fuzzy geometric programming approach to a fuzzy economic order quantity model

Shiang-Tai Liu

Graduate School of Business and Management, Vanung University
Chung-Li, Tao-Yuan 320, Taiwan, ROC
E-mail: stliu@vnu.edu.tw

Abstract. The changing economic conditions make the selling price and demand quantity more and more uncertain in the market. The conventional inventory models determine the selling price and order quantity for a retailer's maximal profit with exactly known parameters. This paper develops a solution method to derive the fuzzy profit of the inventory model when the demand quantity and unit cost are fuzzy numbers. Since the parameters contained in the inventory model are fuzzy, the profit value calculated from the model should be fuzzy as well. Based on the extension principle, the fuzzy inventory problem is transformed into a pair of two-level mathematical programs to derive the upper bound and lower bound of the fuzzy profit at possibility level α. According to the duality theorem of geometric programming, the pair of two-level mathematical programs is transformed into a pair of conventional geometric programs to solve. By enumerating different α values, the upper bound and lower bound of the fuzzy profit are collected to approximate the membership function. Since the profit of the inventory problem is expressed by the membership function rather than by a crisp value, more information is provided for making decisions.

1 Introduction

The integration of production and marketing functions has been recognized to be crucial in practice for increasing a firm's profit and decreasing their conflicts by reducing losses incurred from separate decision making. Pricing and lot sizing are two important strategies that concern simultaneous determination of an item's price and lot size or economic order quantity (EOQ) to maximize a firm's profit for constant but price-dependent demands over a planning horizon. Different from the classic lot sizing or EOQ problem, the demand is typically determined as a decreasing power function of the selling price with constant elasticity as in monopolistic pricing situations [8]. Geometric programming is an efficient and effective method to solve nonlinear programming problem with the terms in power functional form in the objective function and constraints [1,4]. The applications of geometric programming techniques on inventory management are well documented in the literature. Worrall and Hall [16] utilize geometric programming technique to solve an inventory model with multiple items subject to multiple constraints. Cheng [3] proposed an EOQ model with demand-independent unit cost and derived the optimal solution by em-

ploying geometric programming technique. Lee [9] utilized geometric programming techniques to determine the selling price and order quantity for a retailer. Jung and Klein [6] employed geometric programming technique to analyze two EOQ based inventory models under total cost minimization and profit maximization. Chen [2] proposed an inventory model under return-on-inventory -investment maximization for an intermediate firm to determine the selling quantity and purchase cost of a product via geometric programming technique.

The parameters in those studies are assumed to be precisely known. However, in real world applications, the parameters in the model might be inexact and imprecise in nature. For example, the component costs and selling price of personal computer assembly industry are decreasing at a sustained and significant rate [7]. Today the selling price, purchase cost, and demand quantity of a technology product become more and more uncertain in the market. To cope with quantitatively with uncertain information in making decision, Zadeh [18] introduces the notion of fuzziness. The associated problem becomes a fuzzy inventory model. Roy and Maiti [13] employ both fuzzy nonlinear and geometric programming techniques to solve a fuzzy EOQ model with demand-dependent unit cost under limited storage capacity. Nevertheless, only crisp solutions are provided. Intuitively, when the parameters are fuzzy, the objective value should be fuzzy as well. In this paper, we develop a solution procedure that is able to calculate the fuzzy profit of the fuzzy inventory model where the demand quantity and unit cost are represented as fuzzy numbers. The idea is to apply Zadeh's extension principle [17,18,19]. A pair of two-level mathematical programs is formulated to calculate the upper bound and lower bound of the profit at possibility level α. The membership function of the fuzzy profit is derived numerically by enumerating different values of α.

The rest of this paper is organized as follow. We first state the problem and model formulation with fuzzy parameters. Next, a pair of two-level mathematical programs for calculating the upper bound and lower bound of the profit is formulated; we transform the two-level mathematical programs into one-level conventional geometric programs to solve. Then we use an example to illustrate the idea of this paper. Finally, some conclusions are drawn from the discussion.

2 Mathematical formulation

We suppose that one novel version of technology product is being introduced to customers. The demand quantity ($\tilde{D}$) can be regarded as a function of price (p) with constant elasticity; that is, $\tilde{D} = \tilde{K}p^{-\beta}$, where $\tilde{K}$ is a convex fuzzy number denoting the scaling variable and β is the price elasticity. Since the price of technology product usually have a great impact on demand quantity, we assume $\beta > 1$; i.e., the price is elastic. The demand quantity is also a fuzzy number since the scaling variable $\tilde{K}$ is a fuzzy number. The unit cost of this product is changing within a range and the supplier offers quantity discount. We represent the unit cost $\tilde{C}$ as the decreasing func-

tion $\tilde{C} = \tilde{R}q^{-\delta}$, $0 < \delta < 1$, in which $\tilde{R}$ is a convex fuzzy number and q is the lot-size. The discount factor δ is a very small value. If δ =0, it means no quantity discount. Similar to demand quantity, the unit cost $\tilde{C}$ is a fuzzy number since $\tilde{R}$ is fuzzy. According to Lee [9], we have the following formulation for the profit per unit time:

$$\begin{aligned}\pi = f_0(p,q) &= \text{revenue} - \text{ordering cost} - \text{inventory holding cost} - \text{purchase cost},\\ &= p\tilde{D} - A\tilde{D}/q - I\tilde{C}q/2 - \tilde{C}\tilde{D} \qquad (1a)\\ &= \tilde{K}p^{-\beta+1} - A\tilde{K}p^{-\beta}q^{-1} - 0.5I\tilde{R}q^{-\delta+1} - \tilde{K}\tilde{R}p^{-\beta}q^{-\delta} \qquad (1b)\end{aligned}$$

where A is ordering cost and I is inventory carrying cost rate. In this model we have three assumptions, namely, no shortage cost, instantaneous replenishment, and batch order quantity, which are frequently found in the EOQ literature. The objective of this model is to maximize the profit with decision variables p and q.

Clearly, (1b) is an unconstrained signomial problem with fuzzy parameters $\tilde{K}$ and $\tilde{R}$. A signomial geometric programming problem is a class of nonconvex nonlinear programming problem possessing multiple locally optimal solutions. Although several methods [5,14,15] can tackle this type of problem, one simple approach is to transform the problem into a posynomial problem with one additional constraint and variable [1,9] as follows:

$$\begin{aligned}&\text{Max} \quad y \qquad (2)\\ &\quad \text{s.t.} \quad \tilde{K}p^{-\beta+1} - A\tilde{K}p^{-\beta}q^{-1} - 0.5I\tilde{R}q^{-\delta+1} - \tilde{K}\tilde{R}p^{-\beta}q^{-\delta} \ge y,\end{aligned}$$

which is equivalent to the following standard geometric program form:

$$\begin{aligned}&\text{Min} \quad y^{-1} \qquad (3)\\ &\text{s.t.} \quad y(\tilde{K})^{-1}p^{\beta-1} + A\,p^{-1}q^{-1} + .5\,I\tilde{R}(\tilde{K})^{-1}p^{\beta-1}q^{-\delta+1} + \tilde{R}p^{-1}q^{-\delta} \le 1.\end{aligned}$$

Clearly, Model (3) is a fuzzy posynomial geometric programming problem. It is guaranteed to have a global optimum solution if we properly choose the values of parameters β and δ. In the literature the solution techniques for geometric program may be categorized as either primal-based algorithms that directly solve the nonlinear primal problems, or dual-based algorithms that solve the equivalent linearly constrained dual [11]. In view of Rajgopal and Bricker [12], the dual problem has the desirable features of being linearly constrained and having an objective function with attractive structural properties, thus making it a natural candidate solution. According to Beightler and Phillips [1] and Duffin et al. [4], one can transform Model (3) into the corresponding dual geometric program as follows:

$$\begin{aligned}&\underset{\mathbf{w}}{\text{Max}} \quad (1/w_0)^{w_0}\,[(\tilde{K})^{-1}\lambda/w_1]^{w_1}\,[A\lambda/w_2]^{w_2}\,[0.5I\tilde{R}(\tilde{K})^{-1}\lambda/w_3]^{w_3}\,[\tilde{R}\lambda/w_4]^{w_4}\\ &\text{s.t.} \quad w_0 = 1,\\ &\qquad -w_0 + w_1 = 0,\\ &\qquad (\beta-1)w_1 - w_2 + (\beta-1)w_3 - w_4 = 0, \qquad (4)\end{aligned}$$

$$-w_2 + (1-\delta)w_3 - \delta w_4 = 0,$$
$$w_0, w_1, w_2, w_3, w_4 \geq 0.$$

where $\lambda = w_1 + w_2 + w_3 + w_4$. Note that the objective value derived from Model (4) is the reciprocal of the profit π.

Let $\mu_{\tilde{K}}$ and $\mu_{\tilde{R}}$ denote their membership functions. We have,

$$\tilde{K} = \{(k, \mu_{\tilde{K}}(k)) \mid k \in S(\tilde{K})\}$$
$$\tilde{R} = \{(r, \mu_{\tilde{R}}(r)) \mid r \in S(\tilde{R})\}$$

where $S(\tilde{K})$ and $S(\tilde{R})$ are the supports of $\tilde{K}$ and $\tilde{R}$, which represent the universe sets of the scaling variables of demand quantity and unit cost, respectively. Denote the α-cuts of $\tilde{K}$ and $\tilde{R}$ as:

$$(K)_\alpha = [(K)_\alpha^{\mathrm{L}}, (K)_\alpha^{\mathrm{U}}] \tag{5a}$$
$$= [\min_k \{(k, \mu_{\tilde{K}}(k)) \mid S(\tilde{K}) \geq \alpha\}, \max_k \{(k, \mu_{\tilde{K}}(k)) \mid S(\tilde{K}) \geq \alpha\}]$$
$$(R)_\alpha = [(R)_\alpha^{\mathrm{L}}, (R)_\alpha^{\mathrm{U}}] \tag{5b}$$
$$= [\min_r \{(r, \mu_{\tilde{R}}(r)) \mid S(\tilde{R}) \geq \alpha\}, \max_r \{(r, \mu_{\tilde{R}}(r)) \mid S(\tilde{R}) \geq \alpha\}]$$

Suppose we are interested in deriving the membership function of the profit. The major difficulty lies on how to deal with the varying ranges of the scaling variables k and r. One idea is to apply Zadeh's extension principle [17,18,19].

Based on the extension principle, the membership function $\mu_{\tilde{Z}}$ can be defined as:

$$\mu_{\tilde{Z}}(z) = \sup_{k,r} \min \{\mu_{\tilde{K}}(k), \mu_{\tilde{R}}(r) \mid z = Z(k,r)\} \tag{6}$$

where $Z(k,r)$ is defined in Model (4). The application of the extension principle to $\tilde{Z}$ may be viewed as the application of this extension principle to the α-cuts of $\tilde{Z}$ [10]. If the α-cuts of $\tilde{Z}$ at all α values degenerate to the same point, then this number is a crisp number. Otherwise, it is a fuzzy number. In Equation (6), two membership functions are involved. To derive $\mu_{\tilde{Z}}$ in closed form is hardly possible. According to (6), $\mu_{\tilde{Z}}$ is the minimum of $\mu_{\tilde{K}}$ and $\mu_{\tilde{R}}$. We need $\mu_{\tilde{K}}(k) \geq \alpha$ and $\mu_{\tilde{R}}(r) \geq \alpha$, and at least one $\mu_{\tilde{K}}(k)$ or $\mu_{\tilde{R}}(r)$ equal to α such that $z = Z(k,r)$ to satisfy $\mu_{\tilde{Z}}(z) = \alpha$. To find the membership function $\mu_{\tilde{Z}}$, it suffices to find the right shape function and left shape function of $\mu_{\tilde{Z}}$, which is equivalent to finding the upper bound Z_α^{U} and lower bound Z_α^{L} of the α-cuts of $\tilde{Z}$. Since Z_α^{U} is the maximum of $Z(k,r)$ and Z_α^{L} is the minimum of $Z(k,r)$, they can be expressed as:

$$Z_\alpha^{\mathrm{U}} = \max \{Z(k,r) \mid (K)_\alpha^{\mathrm{L}} \leq k \leq (K)_\alpha^{\mathrm{U}}, (R)_\alpha^{\mathrm{L}} \leq r \leq (R)_\alpha^{\mathrm{U}}\} \tag{7a}$$

$$Z_\alpha^{\mathrm{L}} = \min \{Z(k,r) |\ (K)_\alpha^{\mathrm{L}} \le k \le (K)_\alpha^{\mathrm{U}},\ (R)_\alpha^{\mathrm{L}} \le r \le (R)_\alpha^{\mathrm{U}} \tag{7b}$$

Since the inventory model contains fuzzy parameters, the objective value calculated from the model should be fuzzy as well. The values of k and r that attain the largest value of the objective value can be determined from the following two-level mathematical programming model:

$$Z_\alpha^{\mathrm{U}} = \underset{\substack{K_\alpha^{\mathrm{U}} \le k \le K_\alpha^{\mathrm{U}} \\ R_\alpha^{\mathrm{U}} \le r \le R_\alpha^{\mathrm{U}}}}{\mathrm{Max}} \left\{ \begin{array}{l} \underset{\mathbf{w}}{\mathrm{Max}}\ (1/w_0)^{w_0}[k^{-1}\lambda/w_1]^{w_1}\ [A\lambda/w_2]^{w_2}[0.5Ir\,k^{-1}\lambda/w_3]^{w_3}[r\lambda/w_4]^{w_4} \\ \quad \text{s.t.} \\ \quad\quad w_0 = 1, \\ \quad\quad -w_0 + w_1 = 0, \\ \quad\quad (\beta - 1)w_1 - w_2 + (\beta - 1)w_3 - w_4 = 0, \\ \quad\quad -w_2 + (1-\delta)w_3 - \delta w_4 = 0, \\ \quad\quad w_0, w_1, w_2, w_3, w_4 \ge 0. \end{array} \right. \tag{8}$$

where $\lambda = w_1 + w_2 + w_3 + w_4$. At least one k or r must hit the boundary of its α-cut to satisfy $\mu_{\tilde{Z}}(z)=\alpha$. In Model (8), the inner program calculates the objective value for both k and r specified by the outer program, while the outer program determines the values of k and r that generate the largest objective value. At α level, the objective value Z_α^{U} is the upper bound of Model (4).

By the same token, to find the value of k and r that produce the smallest objective value, a two-level mathematical program is formulated by simply replacing the outer program of Model (8) from "Max" to "Min":

$$Z_\alpha^{\mathrm{L}} = \underset{\substack{K_\alpha^{\mathrm{U}} \le k \le K_\alpha^{\mathrm{U}} \\ R_\alpha^{\mathrm{U}} \le r \le R_\alpha^{\mathrm{U}}}}{\mathrm{Min}} \left\{ \begin{array}{l} \underset{\mathbf{w}}{\mathrm{Max}}\ (1/w_0)^{w_0}[k^{-1}\lambda/w_1]^{w_1}\ [A\lambda/w_2]^{w_2}[0.5Ir\,k^{-1}\lambda/w_3]^{w_3}[r\lambda/w_4]^{w_4} \\ \quad \text{s.t.} \\ \quad\quad w_0 = 1, \\ \quad\quad -w_0 + w_1 = 0, \\ \quad\quad (\beta - 1)w_1 - w_2 + (\beta - 1)w_3 - w_4 = 0, \\ \quad\quad -w_2 + (1-\delta)w_3 - \delta w_4 = 0, \\ \quad\quad w_0, w_1, w_2, w_3, w_4 \ge 0. \end{array} \right. \tag{9}$$

where $\lambda = w_1 + w_2 + w_3 + w_4$. In this model the inner program calculates the objective value for each given set of (k, r), while the outer program determines the set of (k, r) that produces the smallest objective value. The objective value Z_α^{L} is the lower bound of the objective value of the α-level cut for Model (4).

Since the objective values derived from (8) and (9) are the reciprocal of the profit defined in (1), the values $[\left(Z_\alpha^{\mathrm{U}}\right)^{-1}, \left(Z_\alpha^{\mathrm{L}}\right)^{-1}] = [\pi_\alpha^{\mathrm{L}}, \pi_\alpha^{\mathrm{U}}]$ constitute the interval at the specific possibility level α.

3 One-level transformation

Models (8) and (9) are a pair of mathematical programs for calculating the upper and lower bounds of the fuzzy objective value at the specific α level. Since both the inner program and outer program of (8) have the same maximization operation, they can be combined into one level with the constraints of the two programs considered at the same time. That is:

$$
\begin{aligned}
Z_\alpha^{\mathrm{U}} = \underset{\mathbf{w}}{\mathrm{Max}} \;\; & (1/w_0)^{w_0} \; [k^{-1}\lambda/w_1]^{w_1} \; [A\lambda/w_2]^{w_2} \; [0.5Ir\,k^{-1}\lambda/w_3]^{w_3} \; [r\lambda/w_4]^{w_4} \\
\text{s.t.} \;\; & w_0 = 1, \\
& -w_0 + w_1 = 0, \\
& (\beta - 1)w_1 - w_2 + (\beta - 1)w_3 - w_4 = 0, \\
& -w_2 + (1-\delta)w_3 - \delta w_4 = 0, \\
& (K)_\alpha^{\mathrm{L}} \le k \le (K)_\alpha^{\mathrm{U}}, \; (R)_\alpha^{\mathrm{L}} \le r \le (R)_\alpha^{\mathrm{U}}, \\
& w_0, w_1, w_2, w_3, w_4 \ge 0.
\end{aligned}
\tag{10}
$$

Since variable k is in the denominator of the objective function, the smaller the value of k the larger the objective value is. To derive the upper bound of the objective value, one should set k to its lower bound K_α^{L}. On the other hand, variable r is in the numerator of the objective function. To obtain the largest objective value, we should set variable r to its upper bound $(R)_\alpha^{\mathrm{U}}$. Consequently, Model (10) can be reduced to the following conventional geometric program:

$$
\begin{aligned}
Z_\alpha^{\mathrm{U}} = \underset{\mathbf{w}}{\mathrm{Max}} \;\; & (1/w_0)^{w_0} \; [(K_\alpha^{\mathrm{L}})^{-1}\lambda/w_1]^{w_1} \; [A\lambda/w_2]^{w_2} \; [0.5IR_\alpha^{\mathrm{U}}\,(K_\alpha^{\mathrm{L}})^{-1}\lambda/w_3]^{w_3} \\
& [R_\alpha^{\mathrm{U}}\lambda/w_4]^{w_4} \\
\text{s.t.} \;\; & w_0 = 1, \\
& -w_0 + w_1 = 0, \\
& (\beta - 1)w_1 - w_2 + (\beta - 1)w_3 - w_4 = 0, \\
& -w_2 + (1-\delta)w_3 - \delta w_4 = 0, \\
& (K)_\alpha^{\mathrm{L}} \le k \le (K)_\alpha^{\mathrm{U}}, \; (R)_\alpha^{\mathrm{L}} \le r \le (R)_\alpha^{\mathrm{U}}, \\
& w_0, w_1, w_2, w_3, w_4 \ge 0.
\end{aligned}
\tag{11}
$$

where $\lambda = w_1 + w_2 + w_3 + w_4$. Model (11) is to find a stationary point of Largrangian function for a concave objective function subject to a set of convex constraints. Hence, Model (11) has a unique stationary point of Largrangian function- a global maximum [1,4] and we can derive the upper bound of the objective value Z_α^U by solving Model (11).

In Model (9) we search for a set of (k , r) that derives the smallest objective value. Similar to Model (8), both variables k and r are in the objective function. To obtain the lower bound of the objective value, we should set k to its upper bound K_α^U; conversely, we should let r to its lower bound R_α^L. Hence, Model (9) becomes the following conventional geometric program

$$
\begin{aligned}
Z_\alpha^L = \underset{\mathbf{w}}{\text{Max}} \quad & (1/w_0)^{w_0}\,[(K_\alpha^U)^{-1}\lambda/w_1]^{w_1}\,[A\lambda/w_2]^{w_2}\,[0.5IR_\alpha^L\,(K_\alpha^U)^{-1}\lambda/w_3]^{w_3} \\
& [R_\alpha^L\lambda/w_4]^{w_4} \\
\text{s.t.} \quad & w_0 = 1, \\
& -w_0 + w_1 = 0, \\
& (\beta-1)w_1 - w_2 + (\beta-1)w_3 - w_4 = 0, \\
& -w_2 + (1-\delta)w_3 - \delta w_4 = 0, \\
& (K)_\alpha^L \le k \le (K)_\alpha^U,\ (R)_\alpha^L \le r \le (R)_\alpha^U, \\
& w_0, w_1, w_2, w_3, w_4 \ge 0.
\end{aligned}
\tag{12}
$$

where $\lambda = w_1 + w_2 + w_3 + w_4$.

To find the optimal solution for Model (11), one simple way is to deal with the reduced problem because the model has only one degree of difficulty. From the constraints of Model (11), we can express w_0, w_1, w_3, w_4, and λ in terms of w_2:

$$w_0 = 1, \tag{13a}$$

$$w_1 = 1, \tag{13b}$$

$$w_3 = [\delta(1-\beta) + (\delta-1)w_2]/(\beta\delta-1), \tag{13c}$$

$$w_4 = [(\beta-1)(\delta-1) + (2-\beta-\delta)w_2]/(\beta\delta-1), \tag{13d}$$

$$\lambda = [\beta(\delta-1)(w_2+1)]/(\beta\delta-1). \tag{13e}$$

Since $\lambda = [\beta(\delta-1)(w_2+1)]/(\beta\delta-1)$ should be greater than 0 and the original assumptions are $\beta > 1$ and $0 < \delta < 1$, we derive $\beta\delta < 1$. It implies that Model (11) has feasible solutions only if $\beta > 1$ and $\beta\delta < 1$, where $0 < \delta < 1$. Lee [9] has also shown this feasibility condition.

The substituted program is formed by replacing variables in the objective function of Model (11) with (13a) through (13e):

$$Z_\alpha^{\mathrm{U}} = \underset{w_2}{\mathrm{Max}} \;\; (K_\alpha^{\mathrm{L}})^{-1} \left(\frac{A}{w_2}\right)^{w_2} \left[\frac{0.5IR_\alpha^{\mathrm{U}}(K_\alpha^{\mathrm{L}})^{-1}}{[\delta(1-\beta)+(\delta-1)w_2]/(\beta\delta-1)}\right]^{[\delta(1-\beta)+(\delta-1)w_2]/(\beta\delta-1)}$$

$$\left[\frac{R_\alpha^{\mathrm{U}}}{[((\beta-1)(\delta-1)+(2-\beta-\delta)w_2)/(\beta\delta-1)]}\right]^{[(\beta-1)(\delta-1)+(2-\beta-\delta)w_2]/(\beta\delta-1)}$$

$$\left[\frac{\beta(\delta-1)(w_2+1)}{(\beta\delta-1)}\right]^{\beta(\delta-1)(w_2+1)/(\beta\delta-1)} \qquad (14)$$

We can take the logarithm of the objective function to derive a concave function [1] as follows:

$$\log Z_\alpha^{\mathrm{U}} = \underset{w_2}{\mathrm{Max}} \;\; -\log K_\alpha^{\mathrm{L}} + w_2 \log\left(\frac{A}{w_2}\right) + + \frac{\delta(1-\beta)+(\delta-1)w_2}{(\beta\delta-1)} +$$

$$\log\left[\frac{0.5IR_\alpha^{\mathrm{U}}(K_\alpha^{\mathrm{L}})^{-1}}{[\delta(1-\beta)+(\delta-1)w_2]/(\beta\delta-1)}\right] + \frac{(\beta-1)(\delta-1)+(2-\beta-\delta)w_2}{(\beta\delta-1)} +$$

$$\log\left[\frac{R_\alpha^{\mathrm{U}}}{[(\beta-1)(\delta-1)+(2-\beta-\delta)w_2]/(\beta\delta-1)}\right] +$$

$$\frac{\beta(\delta-1)(w_2+1)}{(\beta\delta-1)} \log[\beta(\delta-1)(w_2+1)/(\beta\delta-1)] \qquad (15)$$

Since there is only one variable w_2 in this logarithmic function, one can set the first derivative to zero to derive the optimal solution w_2^*

$$\frac{\partial \log Z_\alpha^{\mathrm{U}}}{\partial w_2} = \log\left(\frac{A}{w_2}\right) - 1 + \frac{(\delta-1)}{(\beta\delta-1)} \log\left[\frac{0.5IR_\alpha^{\mathrm{U}}(K_\alpha^{\mathrm{L}})^{-1}}{[\delta(1-\beta)+(\delta-1)w_2]/(\beta\delta-1)}\right] - \frac{(\delta-1)}{(\beta\delta-1)}$$

$$+ \frac{(2-\beta-\delta)}{(\beta\delta-1)} + \frac{\beta(\delta-1)}{(\beta\delta-1)} \log\left[\frac{\beta(\delta-1)(w_2+1)}{\beta\delta-1}\right] + \frac{\beta(\delta-1)}{(\beta\delta-1)} = 0. \qquad (16)$$

Now, we can solve (16) easily by any commercial nonlinear software. Substituting w_2^* into (13c), (13d), and (13e), respectively, we obtain w_3^*, w_4^*, and λ^*. With these values of $\mathbf{w}^*$, the objective value Z_α^{U} can be calculated from the objective function of Model (11) or Eq. (14).

In the constrained geometric programming problem, the dual feasible solutions $\mathbf{w}^*$ and λ^* provide weights of the terms in the constraints of transformed primal problem. Let $\theta_i = w_i^*/\lambda^*$, i =1,2,3,4, where $\sum_{i=1}^{4}\theta_i$ =1. These weights represent

proportions of the profit (θ_1), the ordering cost (θ_2), the inventory holding cost (θ_3), and the purchase cost (θ_4) to the total revenue as follow:

$$\theta_1 = y(K_\alpha^{\mathrm{L}})^{-1} p^{\beta-1}, \tag{17a}$$

$$\theta_2 = A\, p^{-1} q^{-1}, \tag{17b}$$

$$\theta_3 = 0.5\, IR_\alpha^{\mathrm{U}}\, (K_\alpha^{\mathrm{L}})^{-1} p^{\beta-1} q^{-\delta+1}, \tag{17c}$$

$$\theta_4 = R_\alpha^{\mathrm{U}} p^{-1} q^{-\delta}. \tag{17d}$$

Using these equations, we can obtain the corresponding primal solution:

$$\begin{aligned} p^* &= [0.5IA\theta_4 / (K_\alpha^{\mathrm{L}} \theta_2 \theta_3)]^{1/(1-\beta)} \\ &= [R_\alpha^{\mathrm{U}} A^{-\delta} (\theta_2)^{\delta} (\theta_4)^{-1}]^{1/(1-\delta)}, \end{aligned} \tag{18}$$

$$\begin{aligned} q^* &= [K_\alpha^{\mathrm{L}} (\theta_2)^{\beta} \theta_3 / (0.5IA^{\beta}\theta_4)]^{1/(1-\beta)} \\ &= [(R_\alpha^{\mathrm{U}})^{-1} A (\theta_2)^{-1} \theta_4]^{1/(1-\delta)}. \end{aligned} \tag{19}$$

Similar to the discussion of Model (11), Model (12) can be transformed into the following mathematical equation:

$$\begin{aligned} Z_\alpha^{\mathrm{L}} = \max_{w_2} \; & (K_\alpha^{\mathrm{U}})^{-1} \left[\frac{A}{w_2}\right]^{w_2} \left[\frac{0.5IR_\alpha^{\mathrm{L}} (K_\alpha^{\mathrm{U}})^{-1}}{[\delta(1-\beta)+(\delta-1)w_2]/(\beta\delta-1)}\right]^{[\delta(1-\beta)+(\delta-1)w_2]/(\beta\delta-1)} \\ & \left[\frac{R_\alpha^{\mathrm{L}}}{[((\beta-1)(\delta-1)+(2-\beta-\delta)w_2)/(\beta\delta-1)]}\right]^{\frac{[(\beta-1)(\delta-1)+(2-\beta-\delta)w_2]}{(\beta\delta-1)}} \\ & \left[\frac{\beta(\delta-1)(w_2+1)}{(\beta\delta-1)}\right]^{\beta(\delta-1)(w_2+1)/(\beta\delta-1)} \end{aligned} \tag{20}$$

By the same token, we can derive the optimal solution w_2^* by solving Eq. (20). The corresponding weights w_3^*, w_4^*, λ^*, and the objective value Z_α^{L} are calculated from (13c), (13d), (13e), and (20), respectively. Since λ^* and w_i^*, i =1,2,3,4, are obtained, we can easily derive θ_i, $\forall i$, by previous definition

$$\theta_1 = y(K_\alpha^{\mathrm{U}})^{-1} p^{\beta-1}, \tag{21a}$$

$$\theta_2 = A\, p^{-1} q^{-1}, \tag{21b}$$

$$\theta_3 = 0.5\, IR_\alpha^{\mathrm{L}}\, (K_\alpha^{\mathrm{U}})^{-1} p^{\beta-1} q^{-\delta+1}, \tag{21c}$$

$$\theta_4 = R_\alpha^{\mathrm{L}} p^{-1} q^{-\delta}. \tag{21d}$$

Based on the values of θ_i, the solutions of p^* and q^* can be calculated

$$p^* = [0.5IA\theta_4 / (K_\alpha^{\mathrm{U}} \theta_2 \theta_3)]^{1/(1-\beta)} \tag{22}$$
$$= [R_\alpha^{\mathrm{L}} A^{-\delta} (\theta_2)^{\delta} (\theta_4)^{-1}]^{1/(1-\delta)},$$
$$q^* = [K_\alpha^{\mathrm{U}} (\theta_2)^{\beta} \theta_3 / (0.5IA^{\beta} \theta_4)]^{1/(1-\beta)} \tag{23}$$
$$= [(R_\alpha^{\mathrm{L}})^{-1} A (\theta_2)^{-1} \theta_4]^{1/(1-\delta)}.$$

At the upper bound Z_α^{U} and the lower bound Z_α^{L}, we can derive the associated selling price p^* and lot-size quantity q^*, respectively. In (1) the original objective value $\pi = y = 1/Z$. Since $Z_\alpha^{\mathrm{U}} \geq Z_\alpha^{\mathrm{L}}$, the profit will be the interval $[\left(Z_\alpha^{\mathrm{U}}\right)^{-1}, \left(Z_\alpha^{\mathrm{L}}\right)^{-1}]$ at α-level. The numerical solutions for Z_α^{U} and Z_α^{L} at different possibility level α can be collected to approximate the shape of the membership function of the profit.

4 Conclusion

The integration of production and marketing functions is a critical issue in practice for increasing a firm's profit. In previous studies, the demand and the unit cost are regarded as constants and are represented as a decreasing function of the selling price and a decreasing function of lot-size quantity, respectively. The obtained profit and its associated selling price and lot-size quantity are point values. This paper develops a method to find the membership function of the fuzzy profit when the demand quantity and unit cost are represented as fuzzy numbers. The idea is based on Zadeh's extension principle to transform the fuzzy inventory problem to a pair of mathematical programs to derive the upper bound and lower bound of the fuzzy profit. According to the duality theorem of geometric programming technique, the two-level mathematical program is transformed into the one-level conventional geometric program. We can easily find the global optimum solution for the profit via the well-developed theories of geometric programming technique. By enumerating different α values, the upper bound and lower bound of the fuzzy profit are collected to approximate the membership function. Since the profit of the inventory problem is expressed by the membership function rather than by a crisp value, more information is provided for making decisions.

Acknowledgments

Research was supported by the National Science Council of Republic of China under Contract NSC94-2416-H-238-001.

References

1. Beightler,C.S., Philips, D.T.: Applied Geometric Programming. John Wiley & Sons, New York (1976).
2. Chen, C.K.: A return-on-inventory-investment maximization model for intermediate firms. International Journal of Systems Science 32 (2001) 855-861.
3. Cheng, T.C.E.: An economic order quantity model with demand-dependent unit production cost and imperfect production process. IIE Transaction 23 (1991) 23-28.
4. Duffin, R.J., Peterson, E.L., Zener, C.: Geometric Programming- Theory and Applications. John Wiley & Sons, New York (1967).
5. Gao,Y., Xu, C., Wang, Y., Zhang, L.:. A new two-level linear relaxed bound method for geometric programming problems. Applied Mathematics and Computation 164 (2005) 117-131.
6. Jung, H., Klein, C.M.: Optimal inventory policies under decreasing cost functions via geometric programming. European Journal of Operational Research 132 (2001) 628-642.
7. Khouja, M., Park, S., Saydam, C.: Joint replenishment problem under continuous unit cost change. International Journal of Production Research 43 (2005) 311-326.
8. Kim, S., Lee, W.J.: Optimal joint pricing and lot sizing with fixed and variable capacity. European Journal of Operational Research 109 (1998) 212-227.
9. Lee, W.J.: Determining order quantity and selling price by geometric programming: Optimal solution, bounds, and sensitivity. Decision Sciences 24 (1993) 76-87.
10. Nguyen, H.T.: A note on the extension principle for fuzzy sets. Journal of Mathematical Analysis and Applications 64 (1978) 369-380.
11. Peterson, E.L.: The fundamental relations between geometric programming duality, parametric programming duality, and ordinary Lagrangian duality. Annals of Operations Research 105 (2001) 109-153.
12. Rajgopal, J., Bricker, D.L.: Solving posynomial geometric programming problems via generalized linear programming. Computational Optimization and Applications 21 (2002) 95-109.
13. Roy, T.K., Maiti, M.: A fuzzy EOQ model with demand-dependent unit cost under limited storage capacity. European Journal of Operational Research 99 (1997) 425-432.
14. Shen, P.: Linearization method of global optimization for generalized geometric programming. Applied Mathematics and Computation 162 (2005) 353-370.
15. Sherali, H.D., Tuncbilek, C.H.: A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique. Journal of Global Optimization 2 (1992) 101-112.
16. Worrall, B.M., Hall, M.A.: The analysis of an inventory control model using posynomial geometric programming. International Journal of Production Research 20 (1982) 657-667.
17. Yager, R.R.: A characterization of the extension principle. Fuzzy Sets and Systems 18 (1986) 205-217.
18. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1 (1978) 3-28.
19. Zimmermann, H.J.: Fuzzy Set Theory and Its Applications, 3rd edn. Kluwer-Nijhoff, Boston (1996).

LU-fuzzy call option prices*

Maria Letizia Guerra[1,2], Laerte Sorini[2], Luciano Stefanini[2]

[1]Department of Mathematics for Social and Economic Sciences,
University of Bologna, Viale Filopanti, 5, 40126 Bologna, Italy
marialetizia.guerra@rimini.unibo.it
[2]Faculty of Economics, University of Urbino "Carlo Bo",
Via Saffi, 42 61029 Italy
laerte@uniurb.it, lucste@uniurb.it

Abstract

In this paper we show that the vagueness induced by the fuzzy mathematics (see [5], [7] and [14] for more details) can be relevant in modelling objects in finance (see also [2], [8] and [11]), especially when a flexible parametrization is adopted to represent the fuzzy numbers.

Fuzzy calculus for financial applications requires a big amount of computations and the LU-fuzzy representation produces good results due to the fact that it is computationally fast and it reproduces the essential quality of the shape of fuzzy numbers involved in computations. The paper considers the Black and Scholes option pricing formula, as long as many other have done in the last few years (see [3], [13], [15]).

We suggest the use of the LU-fuzzy parametric representation for fuzzy numbers, introduced in Guerra and Stefanini (see [10]) and improved in Stefanini, Sorini and Guerra (see [12]), in the framework of the Black and Scholes model for option pricing, everywhere recognized as a benchmark; the details of the computations and an illustrative example are also incuded.

1 Introduction

Recent literature on fuzzy numbers is rich of several approaches to approximate operations between fuzzy numbers. The desirable feature is to preserve the real shape of the fuzzy numbers resulting from the operations, without loosing in simplicity and applicability and in goodness of the approximations.

*This research was partially supported by the National Project FIRB (financed by the Italian Ministery of University), N.RBAU01KZ7Z-002.

In [10] we introduced a representation of the fuzzy numbers, based on the use of parametrized monotonic functions to model the $\alpha - cuts$ (or the membership functions) of the fuzzy numbers. We call it the LU representation, as it models directly the Lower and the Upper branches of the fuzzy numbers and it uses the parametrization to perform the arithmetic operations and more generally for the fuzzy calculus. The LU-fuzzy numbers can also be viewed as a parametrized extension of the standard LR-fuzzy numbers and are related to this extension by a one-to-one (inverse) correspondence. In [12] we show the advantages of the use of LU-fuzzy numbers in the principal applications of fuzzy calculus: they generalize the LR-fuzzy setting in the direction of the shape preservation but also they allow easy error-controlled approximations in fuzzy calculus.

It is well known in economics and in finance that the application of the option theory, both for the real and for the financial markets, is strongly dependent on the precision of the input data and that in many cases the quality of the information becomes critical to the validity of the results. A suitable methodology to approach this problems is based of the fuzzy calculus as it allows the description of uncertain or imprecise interest rates, volatility, prices, etc in combination with the stochastic (risky) characters of the real world.

On the other hand, the inclusion of vagueness and stochasticity into the models frequently implies massive computations to simulate or perform the calculations.

In section 2 we recall briefly some fundamental properties of fuzzy calculus and in section 3 we apply the LU model to the Black and Scholes fuzzy option pricing formula, in a way similar to the one proposed by Wu in [13] (see also [15] and [3] for recent related aspects).

2 LU parametric fuzzy numbers

The relevant aspects of Fuzzy Set Theory have started by its invention due to Zadeh [14] in 1965. In what follows, according to the representation theorem for fuzzy numbers or intervals (see [9]), we use the so called $a - cut$ setting to define a fuzzy number or interval; a fuzzy number (or interval) u is completely determined by any pair $u = (u^-, u^+)$ of functions $u^{\pm} : [0,1] \longrightarrow R$, defining the end-points of the $a - cuts$, satisfying the three conditions:
(i) $u^- : \alpha \longrightarrow u_\alpha^- \in R$ is a bounded monotonic increasing (non decreasing) left-continuous function $\forall \alpha \in]0,1]$ and right-continuous for $\alpha = 0$;
(ii) $u^+ : \alpha \longrightarrow u_\alpha^+ \in R$ is a bounded monotonic decreasing (non increasing) left-continuous function $\forall \alpha \in]0,1]$ and right continuous for $\alpha = 0$;
(iii) $u_\alpha^- \leq u_\alpha^+ \ \forall \alpha \in [0,1]$.

If $u_1^- < u_1^+$ we have a fuzzy interval and if $u_1^- = u_1^+$ we have a fuzzy number; for simplicity we refer to fuzzy numbers or intervals without nominal distinction.

The notation

$$[u]_\alpha = \left[u_\alpha^-, u_\alpha^+\right], \ \alpha \in [0,1] \tag{1}$$

denotes explicitly the $\alpha - cuts$ of u. We can refer to u^- and u^+ as the lower and the upper branches of u, respectively.

The arithmetic operations for two fuzzy numbers $u = (u^-, u^+)$ and $v = (v^-, v^+)$ are all defined in terms of the $\alpha - cuts$.

The LU-representation of the fuzzy numbers is obtained by writing the lower and the upper branches $u^-, u^+ : [0,1] \rightarrow R$ as monotonic functions. As mentioned in [10] and [12], a family of monotonic rational or mixed cubic-exponential splines can be used as flexible and easy-to-implement models for the parametrization of the LU-fuzzy numbers. The parameters included in the models (number and position of the nodes, values and slopes) allow a wide range of shapes to be taken into account; further, any given fuzzy number can be approximated to a given prescribed precision at the cost of increasing the number of the nodes. The computational experimentations reported in [10] and the extended applications in [12] suggest that the LU-fuzzy numbers are useful tools in the fuzzy calculus.

In the current application we use two monotonic models: a (2,2)-rational spline given by the ratio of two second degree polynomials) and a mixed cubic-exponential spline.

In the simpler forms, the two (continuous) monotonic branches u^- and u^+ are parametrized by 8 numbers $(u_0^-, \delta_0^-, u_0^+, \delta_0^+; u_1^-, \delta_1^-, u_1^+, \delta_1^+)$ giving the values u_0, u_1 and the slopes δ_0, δ_1 (we omit the superscripts $+$ and $-$ for simplicity) at the extremal points of the $[0,1]$ interval; the two interpolating functions for u^- and u^+ for $\alpha \in [0,1]$ are obtained as functions $s(\alpha)$ satisfying $s(0) = u_0$, $s'(0) = \delta_0$, $s(1) = u_1$, $s'(1) = \delta_1$ (if $u_1 = u_0$ we assume $\delta_0 = \delta_1 = 0$ and a constant $s(\alpha)$):

$$\begin{cases} s(\alpha) = \frac{p(\alpha)}{q(\alpha)} \\ \text{where} \\ p(\alpha) = (u_1 - u_0)u_1\alpha^2 + (u_1\delta_0 + u_0\delta_1)\alpha(1-\alpha) + (u_1 - u_0)u_0(1-\alpha)^2 \\ q(\alpha) = (u_1 - u_0)\alpha^2 + (\delta_0 + \delta_1)\alpha(1-\alpha) + (u_1 - u_0)(1-\alpha)^2 \end{cases} \tag{2}$$

for the (2,2)-rational model, and

$$\begin{cases} s(\alpha) = u_0 + (u_1 - u_0 - \frac{\delta_0+\delta_1}{v})\alpha^2(3-2\alpha) + \frac{\delta_0}{v} - \frac{\delta_0}{v}(1-\alpha)^v + \frac{\delta_1}{v}\alpha^v, \\ \text{where} \\ v = 1 + \frac{\delta_0+\delta_1}{u_1-u_0} \end{cases} \tag{3}$$

for the mixed model.

If the data are increasing (i.e. $u_1 \geq u_0$ and $\delta_0 \geq 0, \delta_1 \geq 0$) then $s(\alpha)$ is increasing for all the values of $\alpha \in [0,1]$; if the data are decreasing (i.e. $u_1 \leq u_0$ and $\delta_0 \leq 0, \delta_1 \leq 0$) then $s(\alpha)$ is decreasing for all the values of $\alpha \in [0,1]$. So, simply by choosing the data in the appropriate way, we can model the lower (increasing) and the upper (decreasing) branches of the fuzzy numbers.

An interesting property of (2) is that its inverse can be computed analytically, while (3) is such that if the data are linear ($\delta_0 = \delta_1 = u_1 - u_0$) or quadratic ($\delta_0 + \delta_1 = 2(u_1 - u_0)$) or polynomial ($\delta_0 + \delta_1 = n(u_1 - u_0)$, $n > 2$) then $s(\alpha)$ is a linear, a quadratic or an $(n+1)$-polynomial (a cubic shape is obtained for $\delta_0 = \delta_1 = \frac{3}{2}(u_1 - u_0)$ or $\delta_0 = \delta_1 = 0$ but $u_1 > u_0$). Note also that in models (2)

or (3) any non negative values are admitted for the slopes δ_0 and δ_1 to reproduce non decreasing functions of a very large family (depending on u_0 and $u_1 > u_0$ for the values and on two arbitrary non negative parameters, ranging from zero to infinity, for the slopes); we will use une of the following notations to indicate the spline function s:

$$s(\alpha) = s(\alpha; u, \delta) = s(\alpha; u_0, \delta_0, u_1, \delta_1). \tag{4}$$

The models allows a representation of the $\alpha - cuts$ of the LU-fuzzy number by four pairs of values $u = (u_0^-, \delta_0^-, u_0^+, \delta_0^+; u_1^-, \delta_1^-, u_1^+, \delta_1^+)$: each pair (u_i^*, δ_i^*), with $i \in \{0, 1\}$ and $* \in \{-, +\}$ refers to one of the four interpolating points corresponding to $\alpha_0 = 0$ and to $\alpha_1 = 1$ for the lower and the upper $\alpha - cut$ branches.

To obtain the general (continuous) LU-representation, we introduce a decomposition of the α's into $N + 1$ nodes (N subintervals) $0 = \alpha_0 < \alpha_1 < ... < \alpha_N = 1$ and we model each branch u^- and u^+ by a piecewise function of the forms above. Let $(u_i^-, \delta_i^-)_{i=0,1,...,N}$ and $(u_i^+, \delta_i^+)_{i=0,1,...,N}$ be given values of the functions u^- and u^+ and of their first derivatives at the $N + 1$ nodes of the decomposition; valid values must satisfy the following conditions:

$$\begin{array}{rcl} u_0^- & \leq & u_1^- \leq ... \leq u_N^- \leq u_N^+ \leq u_{N-1}^+ \leq ... \leq u_0^+ \\ \delta_i^- & \geq & 0 \ \text{ and } \ \delta_i^+ \leq 0. \end{array}$$

For all the values of $\alpha \in [0, 1]$ the functions u^- and u^+ are computed on each of the N subintervals of the decomposition with a different version of the mixed spline s: after the transformation $\alpha \in [\alpha_{i-1}, \alpha_i] \longleftrightarrow t_\alpha \in [0, 1]$, i.e. $t_\alpha = \frac{\alpha - \alpha_{i-1}}{\alpha_i - \alpha_{i-1}}$ and $\widetilde{\delta}_i^- = \delta_i^-(\alpha_i - \alpha_{i-1})$, $\widetilde{\delta}_i^+ = \delta_i^+(\alpha_i - \alpha_{i-1})$, we get

$$\begin{array}{rcl} u & = & (u_i^-, \delta_i^-, u_i^+, \delta_i^+)_{i=0,1,...,N} \\ & \Updownarrow & \\ u_\alpha & = & [s(t_\alpha; u_{i-1}^-, \widetilde{\delta}_{i-1}^-, u_i^-, \widetilde{\delta}_i^-), s(t_\alpha; u_{i-1}^+, \widetilde{\delta}_{i-1}^+, u_i^+, \widetilde{\delta}_i^+)]_{i=1,2,...,N}. \end{array} \tag{5}$$

For $N \geq 1$, an array of $4\,(N + 1)$ parameters is available for the lower branch u_α^- (monotonic increasing) and the upper branch u_α^+ (monotonic decreasing); the simple conditions $\delta_i^- \geq 0$, $\delta_i^+ \leq 0$ and $u_0^- \leq u_1^- \leq ... \leq u_N^- \leq u_N^+ \leq u_{N-1}^+ \leq ... \leq u_0^+$ are required.

The set of LU-fuzzy numbers (for a fixed monotonic-shaped model) is denoted by

$$\mathbb{F}_N = \left\{ (u_i^-, \delta_i^-, u_i^+, \delta_i^+)_{i=0,1,...,N} \mid u_i^- \nearrow, \ \ u_i^+ \searrow, \ \ \delta_i^- \geq 0, \ \ \delta_i^+ \leq 0 \right\}.$$

In particular, corresponding to the nodes of the $\alpha-$decomposition, the membership function of u is given by the relations

$$\mu(u_i^-) = \mu(u_i^+) = \alpha_i \ \text{ for } \ i = 0, 1, ..., N$$

and, for the differentiable case, it holds:

$$\mu'(u_i^-) = \frac{1}{\delta_i^-}, \quad \mu'(u_i^+) = \frac{1}{\delta_i^+} \quad \text{for} \quad i = 0, 1, ..., N.$$

For the particular (2,2) rational spline (2) it is easy to compute the analytic inverse of the spline $s(\alpha)$ by solving the following equation of second degree, to obtain $\alpha = s^{-1}(x) \in [0,1]$ for $x \in [u_0, u_1]$ or for $x \in [u_1, u_0]$:

$$p(\alpha) - q(\alpha)x = 0$$

i.e.

$$(u_1-u_0)(u_1-x)\alpha^2+[(u_1-x)\delta_0+(u_0-x)\delta_1]\alpha(1-\alpha)+(u_1-u_0)(u_0-x)(1-\alpha)^2 = 0.$$

If we define $a_x = (u_1 - u_0)(u_1 - x) \geq 0$, $b_x = (u_1 - x)\delta_0 + (u_0 - x)\delta_1$ and $c_x = (u_1 - u_0)(u_0 - x) \leq 0$, then the inverse spline is given by:

$$s^{-1}(x) = \frac{2c_x - b_x \pm \sqrt{b_x^2 - 4a_x c_x}}{2(a_x + c_x - b_x)} \tag{6}$$

where we select the + or the − in the numerator for which the value is in $[0,1]$ (note that it is unique and that $b_x^2 - 4a_x c_x \geq 0 \ \forall x$).

The arithmetic operations, the Zadeh's fuzzy extensions, the fuzzy integral and derivative and other elements of fuzzy calculus can be defined in F_N.

As an example, the fuzzy multiplication is obtained by a relatively simple algorithm: denote $uv = w = \left(w_i^-, f_i^-, w_i^+, f_i^+\right)_{i=0,1,...,N}$, and

$$(uv)_i^- = \min\{u_i^- v_i^-, u_i^- v_i^+, u_i^+ v_i^-, u_i^+ v_i^+\} \tag{7}$$

$$(uv)_i^+ = \max\{u_i^- v_i^-, u_i^- v_i^+, u_i^+ v_i^-, u_i^+ v_i^+\}; \tag{8}$$

let (p_i^-, q_i^-) be the pair associated to the combination of + and - of $u_i^{\pm} v_i^{\pm}$ giving the minimum for $(uv)_i^-$ in (7), and similarly let (p_i^+, q_i^+) be the pair associated to the combination of + and - of $u_i^{\pm} v_i^{\pm}$ giving the maximum for $(uv)_i^+$ in (8), then (for the values and the slopes of $w = uv$)

$$\begin{cases} w_i^- = u_i^{p_i^-} v_i^{q_i^-} \ , & w_i^+ = u_i^{p_i^+} v_i^{q_i^+} \\ f_i^- = \delta_i^{p_i^-} v_i^{q_i^-} + u_i^{p_i^-} e_i^{q_i^-} \ , & f_i^+ = \delta_i^{p_i^+} v_i^{q_i^+} + u_i^{p_i^+} e_i^{q_i^+} . \end{cases}$$

To obtain the fuzzy extension of a real function $f : \mathbb{R} \longrightarrow \mathbb{R}$ (we consider here for simplicity the simple case where f is monotonic) let's introduce the following notation:

$$p_\alpha^-, p_\alpha^+ \in \{-, +\} \text{ and}$$

$$p_\alpha^- = \begin{cases} - & \text{if} \quad \min\left\{f\left(u_\alpha^-\right), f\left(u_\alpha^+\right)\right\} = f\left(u_\alpha^-\right) \\ + & \text{if} \quad \min\left\{f\left(u_\alpha^-\right), f\left(u_\alpha^+\right)\right\} = f\left(u_\alpha^+\right) \end{cases}$$

$$p_\alpha^+ = \begin{cases} - & \text{if} \quad \max\left\{f\left(u_\alpha^-\right), f\left(u_\alpha^+\right)\right\} = f\left(u_\alpha^-\right) \\ + & \text{if} \quad \max\left\{f\left(u_\alpha^-\right), f\left(u_\alpha^+\right)\right\} = f\left(u_\alpha^+\right) \end{cases}$$

(where we simplify $p^{\pm}_{\alpha_i} \equiv p^{\pm}_i$, $i = 0, 1, ..., N$, in the points of the $\alpha-$decomposition).

So, we have $f(u)^-_\alpha = f\left(u^{p^-_\alpha}_\alpha\right)$ and $f(u)^+_\alpha = f\left(u^{p^+_\alpha}_\alpha\right)$.

If X is the LU-fuzzy number $X = \left(x^-_i, \delta^-_i, x^+_i, \delta^+_i\right)_{i=0,1,...,N}$ then its image $f(X)$ is

$$f(X) = \left(f(x_i^{p_i^-}), f'(x_i^{p_i^-})\delta_i^{p_i^-}, f(x_i^{p_i^+}), f'(x_i^{p_i^+})\delta_i^{p_i^+}\right)_{i=0,1,...N}.$$

An LU-fuzzy calculator, implementing the basic fuzzy arithmetic operations and calculus by an interactive desk-top easy-to-use software is available by contacting the authors[1].

3 LU-fuzzy Black-Scholes model

Many empirical studies have shown that some hypothesis of the Black-Scholes (B-S) model for european options (introduced in [1]) do not reflect in a satisfactory way the behavior of financial markets. In particular the risk-free rate r is considered constant but in the real world it varies in an imprecise way. Moreover, a big amount of recent financial literature is devoted to the volatility modeling and many different approaches have been studied; we believe that the fuzzy modeling of volatility can contribute in the research of a benchmark volatility model, especially when an easy model to implement is adopted.

The B-S formulation consists in determining the current value C_t of an european call option of a derivative security having known current stock price and exercise price at maturity time T. The volatility and the so called risk-neutral interest rate are the other parameters by which the B-S formula is composed. All these parameters, with the possible exception of the strike price and of the time to maturity, can be modelled as fuzzy numbers. In our calculations, the stock price, the strike price, the volatility and the interest rate are handled as fuzzy numbers due to their imprecise character. The strike price and the time to maturity are supposed to be real numbers because they have this characteristics in the real world.

The general procedure is very similar to that described by Wu (see [13]); we focus on the alternative way of representing the fuzzy numbers, the LU representation illustrated in the preceeding section, and on its advantages in producing quickly the final solution.

Denote, as usual, by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

the cumulative standard normal function.

[1] L. Sorini - L. Stefanini, An LU-fuzzy calculator for the basic fuzzy calculus, EMS working paper series, 98, 2005, University of Urbino "Carlo Bo", Italy (contact laerte@uniurb.it or lucste@uniurb.it)

It is an increasing function with derivative

$$\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

To introduce the standard Black Scholes formula for the European put/call options (without dividends), denote

$$D_1(S, K, r, \sigma, t) = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right) t}{\sigma\sqrt{t}}$$

and

$$D_2(S, K, r, \sigma, t) = D_1(S, K, r, \sigma, t) - \sigma\sqrt{t}.$$

We also write

$$D_1 = \frac{\ln\left(\frac{S}{K}\right) + rt}{\sigma\sqrt{t}} + \frac{\sigma}{2}\sqrt{t}$$

$$D_2 = \frac{\ln\left(\frac{S}{K}\right) + rt}{\sigma\sqrt{t}} - \frac{\sigma}{2}\sqrt{t}.$$

If S_t is the (current) stock price at time $t \in [0, T]$ (where T is the time to maturity), K is the exercise price (strike price), r is the interest rate (continuously compounded), σ is the (standard deviation) volatility, then the price C_t of the corresponding European call option at time t is given by:

$$C_t = S_t \Phi(D_1(S_t, K, r, \sigma, T - t)) - K e^{-r(T-t)} \Phi(D_2(S_t, K, r, \sigma, T - t)) \quad (9)$$

and the price P_t of the corresponding European put option at time t (with the same expiry date T and strike price K) is given by:

$$P_t = C_t + K e^{-r(T-t)} - S_t. \quad (10)$$

It is also important to compute a probability density function for the option value C_t, given by (log-normal type):

$$D(C_t) = \frac{a}{2\pi} e^{-\frac{1}{2}b^2} \quad (11)$$

where

$$a = \frac{e^{r(T-t)}}{\sigma(T-t)\left(C_t e^{r(T-t)} + S_t\right)}$$

and

$$b = \frac{\ln\left(\frac{C_t e^{r(T-t)} + k}{S}\right) - \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sqrt{\sigma(T-t)}}.$$

3.1 The LU-fuzzy representation

The fuzzyfication of (9) and (10) assumes that K, T are given crisp numbers and $r, \sigma, \{S_t; t \in [0,T]\}$ are given fuzzy numbers; in the LU-fuzzy setting, denote:

$$r = \left(r_i^-, \delta r_i^-, r_i^+, \delta r_i^+\right)_{i=0,1,...,N}$$

$$\sigma = \left(\sigma_i^-, \delta\sigma_i^-, \sigma_i^+, \delta\sigma_i^+\right)_{i=0,1,...,N}$$

$$S_t = \left(S_{t,i}^-, \delta S_{t,i}^-, S_{t,i}^+, \delta S_{t,i}^+\right)_{i=0,1,...,N}$$

and $0 = \alpha_0 < \alpha_1 < ... < \alpha_N = 1$.

To compute the fuzzy extensions of (9), (10) and (11), observe first that all the involved quantities are non negative crisp or fuzzy numbers so that all the arithmetic operations are simplified; the procedure works as follows:

Step 1. Compute the fuzzy extension of the function

$$S_t \quad \rightarrow \quad \ln\left(\frac{S_t}{K}\right) \tag{12}$$

(S_t fuzzy, K crisp, t given).

The strike price K is a crisp positive quantity and we use the scalar multiplication

$$\frac{S_t}{K} = \left(\frac{S_{t,i}^-}{K}, \frac{\delta S_{t,i}^-}{K}, \frac{S_{t,i}^+}{K}, \frac{\delta S_{t,i}^+}{K}\right)_{i=0,1,...,N}$$

as the function $x \rightarrow \ln(x)$ is increasing, the fuzzy extension of (12) is the following

$$\ln\left(\frac{S_t}{K}\right) = \left(\ln\left(\frac{S_{t,i}^-}{K}\right), \frac{\delta S_{t,i}^-}{S_{t,i}^-}, \ln\left(\frac{S_{t,i}^+}{K}\right), \frac{\delta S_{t,i}^+}{S_{t,i}^+}\right)_{i=0,1,...,N}$$

Denote this fuzzy number by:

$$L_t = \left(L_{t,i}^-, \delta L_{t,i}^-, L_{t,i}^+, \delta L_{t,i}^+\right)_{i=0,1,...,N}$$

Step 2. Compute $D_{1,t}$ and $D_{2,t}$ by extending the fuzzy functions

$$D_{1,t} : (S_t, r, \sigma) \longrightarrow \frac{\ln\left(\frac{S_t}{K}\right) + r\,(T-t)}{\sigma\sqrt{T-t}} + \frac{\sigma\sqrt{T-t}}{2} \tag{13}$$

and

$$D_{2,t} : (S_t, r, \sigma) \longrightarrow \frac{\ln\left(\frac{S_t}{K}\right) + r\,(T-t)}{\sigma\sqrt{T-t}} - \frac{\sigma\sqrt{T-t}}{2} \tag{14}$$

as functions of S_t, r and σ for given (crisp) K, T, t.

If

$$A_t = \ln\left(\frac{S_t}{K}\right) + r\,(T-t)$$

we denote

$$A_t = \left(A^-_{t,i}, \delta A^-_{t,i}, A^+_{t,i}, \delta A^+_{t,i}\right)_{i=0,1,\ldots,N}$$

in order to obtain:

$$A^-_{t,i} = \left(L^-_{t,i} + r^-_i\,(T-t)\,, \delta L^-_{t,i} + \delta r^-_i\,(T-t)\,, L^+_{t,i} + r^+_i\,(T-t)\,, \delta L^+_{t,i} + \delta r^+_i\,(T-t)\right).$$

The fuzzy extensions of $D_{1,t}$ and $D_{2,t}$ in the LU-representation can be written as:

$$D_{1,t} = \left(D^-_{1,t,i}, \delta D^-_{1,t,i}, D^+_{1,t,i}, \delta D^+_{1,t,i}\right)_{i=0,1,\ldots,N}$$

$$D_{2,t} = \left(D^-_{2,t,i}, \delta D^-_{2,t,i}, D^+_{2,t,i}, \delta D^+_{2,t,i}\right)_{i=0,1,\ldots,N}$$

and the values of the parameters are obtained by applying the extension principle to the functions (13) and (14). We omit here the tedious details for brevity; observe only that, with respect to $x = \sigma\sqrt{T-t}$, $D_{1,t}$ and $D_{2,t}$ have the following form:

$$D_{1,t} : x \longrightarrow \frac{A_t}{x} + \frac{x}{2}, \qquad D_{2,t} : x \longrightarrow \frac{A_t}{x} - \frac{x}{2}$$

and their fuzzy extensions require to distinguish the two cases $A_t \geq 0$ and $A_t \leq 0$; in fact, $D_{1,t}$ is increasing if $A_t \leq 0$ but has a global minimum point at $x = (A_t)^{\frac{2}{3}}$ if $A_t > 0$; analogously, $D_{2,t}$ is decreasing if $A_t \geq 0$ but has a global maximum point at $x = (-A_t)^{\frac{2}{3}}$ if $A_t < 0$.

Step 3. Compute $\Phi(D_{1,t})$ and $\Phi(D_{2,t})$ as the fuzzy extension of the cumulative standard normal function.

As the function $x \to \Phi(x)$ is differentiable and monotonic increasing, it is immediate to obtain its fuzzy extension $\Phi(D_{k,t})$ $k = 1, 2$:

$$\Phi_{k,t} = \left(\Phi\left(D^-_{k,t,i}\right), \delta D^-_{k,t,i}\Phi'\left(D^-_{k,t,i}\right), \Phi\left(D^+_{k,t,i}\right), \delta D^+_{k,t,i}\Phi'\left(D^+_{k,t,i}\right)\right) \quad (15)$$

To compute the (integral) cumulative function $\Phi(x)$ we use a well known approximation[2].

Step 4 Compute the fuzzy extension of

$$\begin{array}{rcl} r & \to & e^{-r(T-t)} \\ & & (T \text{ and } t \text{ crisp, } r \text{ fuzzy}) \end{array} \quad (16)$$

The fuzzy extension of this differentiable decreasing function has the following LU-representation

$$e^{-r(T-t)} = \left(\exp\left(-r^+_i\,(T-t)\right), -(T-t)\,\delta r^+_i \exp\left(-r^+_i\,(T-t)\right),\right.$$

$$\left.\exp\left(-r^-_i\,(T-t)\right), -(T-t)\,\delta r^-_i \exp\left(-r^-_i\,(T-t)\right)\right)_{i=0,1,\ldots,N}$$

denote it by

$$E_t = \left(E^-_{t,i}, \delta E^-_{t,i}, E^+_{t,i}, \delta E^+_{t,i}\right)_{i=0,1,\ldots,N}$$

[2] For more details see W.H. Press *et al.* Numerical Recipes in C: the Art of Scientific Computing, Cambridge University Press, second edition, 1992.

Step 5 Compute C_t by the arithmetic operations in (9).

If the LU-representations of the fuzzy extension of C_t, P_t and $D(C_t)$ are denoted by

$$C_t = \left(C^-_{t,i}, \delta C^-_{t,i}, C^+_{t,i}, \delta C^+_{t,i}\right)_{i=0,1,\ldots,N}$$

$$P_t = \left(P^-_{t,i}, \delta P^-_{t,i}, P^+_{t,i}, \delta P^+_{t,i}\right)_{i=0,1,\ldots,N}$$

$$D_t = \left(D^-_{t,i}, \delta D^-_{t,i}, D^+_{t,i}, \delta D^+_{t,i}\right)_{i=0,1,\ldots,N}$$

we easily obtain (due to the fact that all quantities are positive) the following

$$C^-_{t,i} = S^-_{t,i}\Phi^-_{1,t,i} - KE^+_{t,i}\Phi^+_{2,t,i}$$

$$\delta C^-_{t,i} = \delta S^-_{t,i}\Phi^-_{1,t,i} + S^-_{t,i}\delta\Phi^-_{1,t,i} - K\left(\delta E^+_{t,i}\Phi^+_{2,t,i} + E^+_{t,i}\delta\Phi^+_{2,t,i}\right)$$

$$C^+_{t,i} = S^+_{t,i}\Phi^+_{1,t,i} - KE^-_{t,i}\Phi^-_{2,t,i}$$

$$\delta C^+_{t,i} = \delta S^+_{t,i}\Phi^+_{1,t,i} + S^+_{t,i}\delta\Phi^+_{1,t,i} - K\left(\delta E^-_{t,i}\Phi^-_{2,t,i} + E^-_{t,i}\delta\Phi^-_{2,t,i}\right)$$

The computation of the fuzzy price P_t in (10) and of the probability D_t in (11) can be made in a similar way as for C_t. We only note that equation (10) defines P_t as the Hukuhara difference:

$$P_t = C_t + Ke^{-r(T-t)} \overset{h}{-} S_t$$

where $\overset{h}{-}$ is defined by the equality (if the fuzzy Z exists)

$$Z = X \overset{h}{-} Y \Longleftrightarrow Z + Y = X.$$

Introducing the term $E_t = Ke^{-r(T-t)}$ the following equalities hold $\forall\alpha \in [0,1]$:

$$\begin{cases} P^-_{t,\alpha} + S^-_{t,\alpha} &= C^-_{t,\alpha} + E^-_{t,\alpha} \\ P^+_{t,\alpha} + S^+_{t,\alpha} &= C^+_{t,\alpha} + E^+_{t,\alpha} \end{cases}$$

that are equivalent to the following (after (9) and (15)):

$$\begin{cases} P^-_{t,\alpha} &= S^-_{t,\alpha}\left(\Phi^-_{1,t,\alpha} - 1\right) + E^-_{t,\alpha} - E^+_{t,\alpha}\Phi^+_{2,t,\alpha} \\ P^+_{t,\alpha} &= S^+_{t,\alpha}\left(\Phi^+_{1,t,\alpha} - 1\right) + E^+_{t,\alpha} - E^-_{t,\alpha}\Phi^-_{2,t,\alpha}. \end{cases}$$

The final step of our computation is the construction of the membership function by inverting the LU representation. As illustrated in the previous section, depending on the adopted model for the spline, we can use the analytic inverses as in (6) and (**??**) for the (2,2)-rational spline, or one of the two methods (**??**) or (**??**) for the mixed spline model.

3.2 Computational results

As an example of the illustrated procedures, we consider the same evaluation reported by Wu ([13]). In our case, the complexity of the computations is extremely reduced: as we will see, only 5 points in the $\alpha-$decomposition ($N = 4$) give an error less than 0.004% (with a precision between the fourth and the fifth decimal place) for all the values of the membership greater than $\alpha = 0.5$ (note that, in the financial applications, usual values of interest for α are between 0.9 and 1.0). In the framework of the LU-fuzzy calculus, the results are obtained in a much simpler way and the calculation of the membership function of the involved quantities makes also possible their use in further fuzzy calculations. In particular, the problem to find the value of α such that $\mu(C) = \alpha$ (where μ is the membership function) can be solved in closed form and does not require additional computational effort.

In our computations, the results are obtained by the LU-fuzzy decomposition with $N = 1, 2, 4$ and 10 and by an exact procedure in 101 uniform α points. The approximation with $N = 10$ is exact up to 8 decimal places and is not reported further.

The same data as in ([13]) are used: $K = 30$, $T = 0.25$, $t = 0$, $S_0 = (32, 33, 34)$, $r = (0.048, 0.05, 0.052)$ and $\sigma = (0.08, 0.1, 0.12)$ (triangular and symmetric).

The following tables give the LU-representation for the data S_0, r and σ for the first $\alpha-$ decomposition ($N = 1$):

$u = S_0$	α_i	u_i^-	du_i^-	u_i^+	du_i^+
	0	32	1	34	-1
	1	33	1	33	-1

$u = r$	α_i	u_i^-	du_i^-	u_i^+	du_i^+
	0	0.048	0.002	0.052	-0.002
	1	0.05	0.002	0.05	-0.002

$u = \sigma$	α_i	u_i^-	du_i^-	u_i^+	du_i^+
	0	0.08	0.02	0.12	-0.02
	1	0.1	0.02	0.1	-0.02

Figure 1 illustrates the calculated exact and approximated membership functions for the α values of interest. Note that the approximations with $N = 2$ and $N = 4$ are graphically coincident with the exact solution.

The obtained LU-representation of C_0 is reported for the two $\alpha-$ decompo-

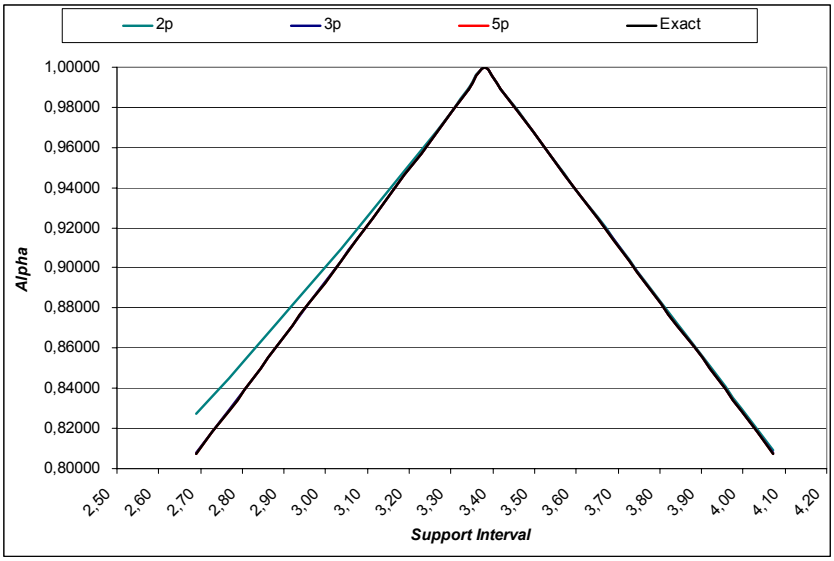


Figure 1: Computed exactly and approximated membership functions

sitions (N = 2, 4) in the following table:

$u = C_0$	α_i	u_i^-	du_i^-	u_i^+	du_i^+
N=2	0	-0.7549	5.0881	7.5312	-5.1247
	0.5	1.5173	4.0371	5.2478	-4.0486
	1	3.3813	3.5673	3.3813	-3.5673
N=4	0	-0.7549	5.0881	7.5312	-5.1247
	0.25	0.4479	4.5355	6.3212	-4.5570
	0.5	1.5173	4.0371	5.2478	-4.0486
	0.75	2.4792	3.6903	4.2839	-3.6951
	1	3.3813	3.5673	3.3813	-3.5673

The percentage error of the 5-point approximation is reported in figure 2 (here a zero error means exact to the fifth decimal).

The remaining three figures illustrate the membership functions of the fuzzy option value C_t, the fuzzy price value P_t and the fuzzy probability density D_t obtained by the mixed spline model.

It appears that the LU parametric representation and the corresponding calculus maintain the shape of the fuzzy numbers and are precise also when few nodes are considered. This means a relevant saving in the computations and the advantage of obtaining the membership functions which can be used for further fuzzy operations.

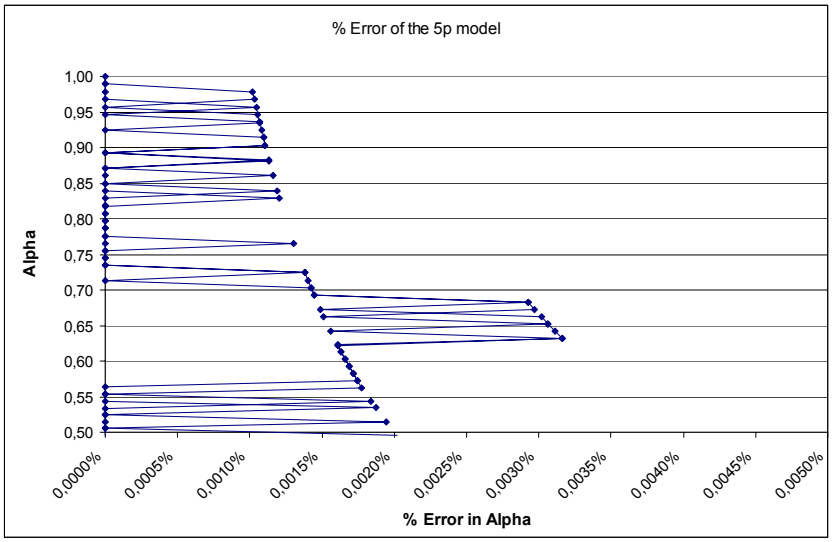


Figure 2: Error expressed in percentage

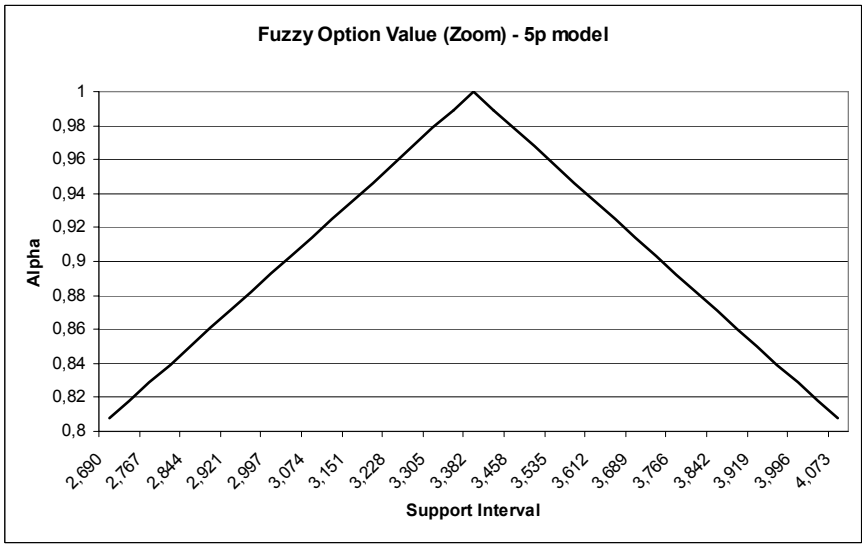


Figure 3: The membership function of the fuzzy option value C_t

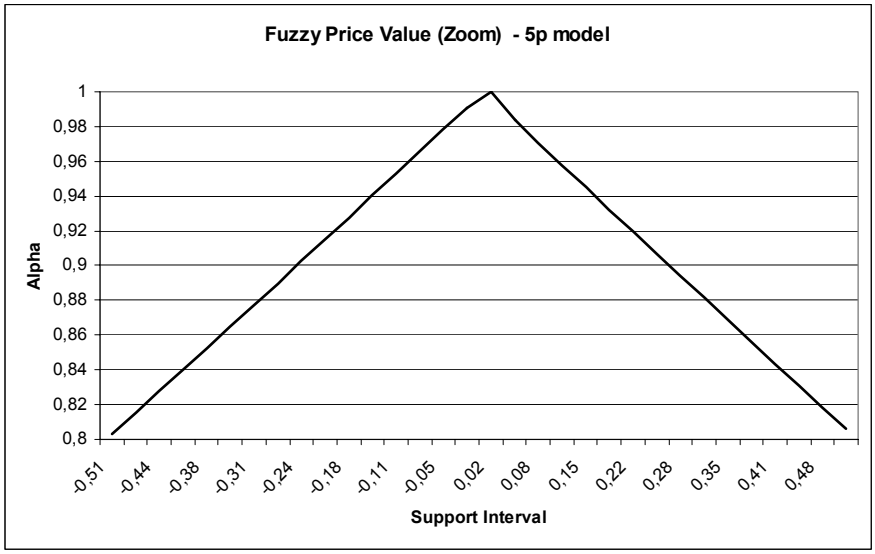


Figure 4: The membership function of the fuzzy price value P_t.

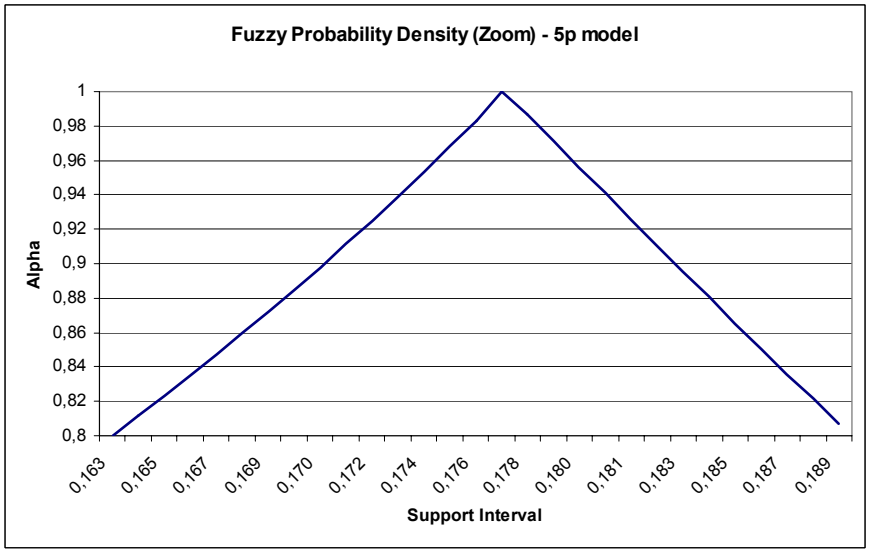


Figure 5: The memebership function of the fuzzy probability density D_t.

References

[1] Black F., Scholes M., The Pricing of Options and Corporate Liabilities, Journal of Political Economics, 87, (1973), 637-659.

[2] J.J. Buckley, The Fuzzy Mathematics of Finance, Fuzzy Sets and Systems, 21, (1987), 57-73.

[3] C. Carlsson, R. Fullér, Real Option Evaluation in Fuzzy Environment, Proceedings of the International Symposium of Hungarian Researchers on Computational Intelligence, Budapest, (2000), 69-77.

[4] C. Carlsson, R. Fullér, On Possibilistic Mean Value and Variance of Fuzzy Numbers, Fuzzy Sets and Systems, 122, (2000) 315-326.

[5] D. Dubois, H. Prade, Fuzzy Sets and Systems: Theory and Applications, Academic Press, New York, (1980).

[6] D. Dubois, H. Prade, Possibility Theory. An approach to Computerized Processing of Uncertainty, Plenum Press, New York, (1988).

[7] D. Dubois, E. Kerre, R.Mesiar, H. Prade, Fuzzy Interval Analysis, in: Fundamentals of Fuzzy Sets, D. Dubois, H. Prade, Rds. Kluwer, Boston, The Handbooks of Fuzzy Sets Series, (2000) 483-581.

[8] M.Fedrizzi, M.Fedrizzi, W.Ostasiewicz, Towards Fuzzy Modelling in Economics, Fuzzy Sets and Systems, 54, (1993), 259-268.

[9] R. Goetschel, W. Voxman, Elementary fuzzy calculus, Fuzzy Sets and Systems, 18, 1986, 31-43.

[10] M.L. Guerra, L. Stefanini, Approximate Fuzzy Arithmetic Operations Using Monotonic Interpolations, Fuzzy Sets and Systems,150/1, 5-33, 2005.

[11] M. Li Calzi, Towards a General Setting for the Fuzzy Mathematics of Finance, Fuzzy Sets and Systems, 35, (1990), 265-280.

[12] L. Stefanini, L.Sorini, M.L.Guerra, Parametric Representations of Fuzzy Numbers and Application to Fuzzy Calculus, Fuzzy Sets and Systems, forthcoming, 2006.

[13] H-C. Wu, Pricing European options based on the fuzzy pattern of Black Scholes formula, Computers and Operations Research, 31, 1069-1081, 2004.

[14] L.A. Zadeh, Fuzzy Sets, Information and Control, 8, (1965), 338-353.

[15] Z. Zmeskal, Application of the fuzzy-stochastic methodology to appraising the firm value as a European call option, European Journal of Operational Research, 135, (2001), 303-310.

Fuzzy strategies in modelling effects of technological implementations

Henryk Piech, Aleksandra Ptak, Marcin Machura

Technical University of Czestochowa, Institute of Mathematics and Informatics
Dabrowskiego Street, 73, 42-200 Czestochowa, Poland,
tel.:00480343250331
hpiech@adm.pcz.czest.pl

Abstract. The description of the technological process one finds in von Neumann's model which looks as follows:

$$A = \begin{Bmatrix} a(1,1) & a(2,1) & \dots & a(k,1) \\ a(1,2) & a(2,2) & \dots & a(k,2) \\ \dots & \dots & \dots & \dots \\ a(1,j) & a(2,j) & \dots & a(k,j) \\ \dots & \dots & \dots & \dots \\ a(1,m) & a(2,m) & \dots & a(k,m) \end{Bmatrix} \quad B = \begin{Bmatrix} b(1,1) & b(2,1) & \dots & b(k,1) \\ b(1,2) & b(2,2) & \dots & b(k,2) \\ \dots & \dots & \dots & \dots \\ b(1,j) & b(2,j) & \dots & b(k,j) \\ \dots & \dots & \dots & \dots \\ b(1,m) & b(2,m) & \dots & b(k,m) \end{Bmatrix} \qquad (1)$$

A simple method of describing technology can be presented in the following way: consumption of goods - a(1,j) units of number 1 goods, a(2,j) units of number 2 goods, etc., a(k,j) units of number k goods, one obtains the result which includes: b(1,j) units of different number 1 goods, b(2,j) units of different number 2 goods, etc., and finally b(k,j) units of different number k goods. The simplified equation corresponds with such a technology:

$$(a(1,j)a(2,j) \dots a(k,j)) \longrightarrow (b(1,j)\ b(2,j) \dots b(k,j)) \qquad (2)$$

Assume that the technologies were realized with the following intensities: x(1),x(2),...,x(m), which means that using technology 1 one supplies multiple x(1) outlays (a(1,1),a(2,1), ..., a(k,1), i.e. (x(1)*a(1,1),x(1)*a(2,1),...,x(1)*a(k,1)) and expects similar multiples of the result, i.e. (x(1)*b(1,1), x(1)*b(2,1), ... ,x(1)*b(k,1)). Using technology *j* one generates outlays (x(j)*a(1,j), x(j)*a(2,j), ... ,x(j)*a(k,j)) and obtains results: (x(j)*b(1,j), x(j)*b(2,j), ... ,x(j)*b(k,j)). Total outlays of the first goods are (x(1)*a(1,1),x(2)*a(1,2),...,x(m)*a(1,m)), and *r-th* goods - (x(1)*a(1,1), x(2)*a(1,2), ... ,x(r)*a(r,m)). Thus, the vector of outlays **x*A** arises. The vector of the obtained results looks as follows:

(x(1)*b(1,1) + x(2)*b(1,2) + ,..., + x(m)*b(1,m);
(x(1)*b(2,1) + x(2)*b(2,2) + ,..., + x(m)*b(2,m);
.....
(x(1)*b(k,1) + x(2)*b(k,2) + ,..., +x(m)*b(k,m)) = **x*B**

One is only interested in those vectors which have at least one positive coordinate. Technological growth is defined as:

$$\text{techn\_growth} = \min_{1 \le i \le k} \{ \sum_{j=1}^{m} (x(j)*b(i,j)) / \sum_{j=1}^{m} (x(j)*a(i,j)) \} \qquad (3)$$

1 Introduction

Prices of individual goods can be presented by means of the prices' vector *(p(1), p(2),...,p(k)), p(i) ≠ 0, i =1,...,k*. Financial control of the studied undertaking can be ascribed to "the second" player in the strategy game [1]. Vector ***Ap*** describes the value of unit outlays of processing results. Thus, if one divides each coordinate of vector ***Bp*** (values of technological effects) by the appropriate coordinate of vector ***Ap***, then the smallest such value (for one of the vectors' coordinates) will be accepted as the technological profitability coefficient:

$$techn\_prof\_coeff = \min_{1 \le i \le m} \{ \sum_{i=1}^{k} (b(i,j)*p(i)) / \sum_{i=1}^{k} (a(i,j)*p(i)) \} \quad (4)$$

The rate of economic growth is the largest value of the quotient by the assumption that at least one of the coordinates *p(i)* is larger than zero:

$$econ\_growth = \max_{1 \le i \le m} \{ \sum_{i=1}^{k} (b(i,j)*p(i)) / \sum_{i=1}^{k} (a(i,j)*p(i)) \} \quad (5)$$

The introduction of fuzziness with reference to both independent and dependent parameters takes the model closer to the real situation. One certainly senses that an increase in the outlays as well as the intensity degree lead to an increase of production and technological results. One also is aware of the fact that through some technological solutions it is possible to incur a loss. Also, one does not always have an influence on the intensity level. A global problem related to many technologies, various outlays, various intensity levels and various final products requires a wider analysis and can lead not only to surprising results but can also provide suggestions for the control of production and technological processes.

2 Introduction rules of fuzzy strategies

Fuzziness with reference to outlays ***A*** represents a quantitative range of use of individual products for the realizations of selected technologies. This is statistic information or experts' knowledge. This fuzziness can be connected with the quality of products, rejects and technological requirements. Fuzziness brought into intensity ***X*** can determine a controlled parameter; thanks to it one can control technological and economic growth as well as profitability indexes. Fuzziness of technology implementation effects, i.e. parameters of matrix ***B***, one can determine on the basis of statistical data or can calculate using the interval analysis on the basis of fuzziness of matrixes ***A*** and ***X:***

$$[Xp,Xk]*[\mathbf{A}p,\mathbf{A}k] = [B'p,B'k] \quad (6)$$

where: *p* denotes values of minimal limitations, *k* – values of maximal limitations, *B'* – row vectors of matrix ***B***.

If, on the basis of intensity vector ***X*** and matrix ***A***, one wants to estimate lower and upper limits of the results, then it is possible to use the interval mathematics [7]. The mathematical apparatus will apply to addition and multiplication operations. The interval operations for closed intervals are realized in the following way:

$$[xp(i),xk(i)] + [xp(j),xk(j)] = [xp(i) + xp(j), xk(i) + xk(j)] \quad (7)$$

If one wants to determine the intensity on the basis of results and outlays it is possible to apply the interval division:

$$[xp(i),xk(i)] * [xp(j),xk(j)] = [\min\{xp(i)* xp(j), xp(i)* xk(j), xk(i)* xp(j), xk(i) * xk(j)\},\max\{xp(i)* xp(j), xp(i)* xk(j), xk(i)* xp(j), xk(i) * xk(j)\}] \quad (8)$$

$$[xp(i),xk(i)] / [xp(j),xk(j)] = [xp(i),xk(i)]*1/[xp(j),xk(j)] \quad (9)$$

while

$$1/[\mathrm{xp(j)},\mathrm{xk(j)}] = \begin{cases} 0 & for \quad [\mathrm{xp(j)},\mathrm{xk(j)}] = [0,0] \\ [1/\mathrm{xk(j)}, 1/\mathrm{xp(j)}] & for \quad 0 \notin [\mathrm{xp(j)},\mathrm{xk(j)}] \\ [1/\mathrm{xk(j)}, \infty] & for \quad \mathrm{xp(j)} = 0 \ \mathrm{i} \ \mathrm{xk(j)} > 0 \\ [-\infty, 1/\mathrm{xk(j)}] & for \quad \mathrm{xp(j)} < 0 \ \mathrm{i} \ \mathrm{xk(j)} = 1 \\ [-\infty, \infty] & for \quad \mathrm{xp(j)} < 0 \ \mathrm{i} \ \mathrm{xk(j)} > 0 \end{cases} \quad (10)$$

Multiplying intensity vector ***X*** by productivity matrix ***A*** one obtains the result vector which can estimate fuzziness as follows:

$$[\sum_{i=1}^{m} \min\{xp(i)*ap(i,j), xp(i)*ak(i,j), xk(i)*ap(i,j), xk(j)*ak(i,j)\}, \sum_{i=1}^{m} \max\{xp(i)*ap(i,j), xp*ak(i,j), xk(i)*ap(i,j), xk(j)*ak(i,j)\}] \quad l \le j \le k \quad (11)$$

In the situation when the basis information is result ***B*** and borne outlays ***A*** and one wants to estimate intensity ranges ***X***, it is possible to use intervalization methods relating to solving the system of equations: $\boldsymbol{A}^{\mathrm{T}}*X=\boldsymbol{B}$, $k=m$. Gauss's elimination method leads to the intensity intervalization according to the following description [3] (so-called contractor):

$$[x(j)] = ([b(r,j)] - \sum_{i=j+1}^{k} [a(j,i)]*[x(i)])/[a(j,j)], \ 1 \le j \le k, \ 1 \le r \le m \quad (12)$$

Similar contractors can be presented for the Gauss-Seidel method, Krawczyk's method, forward and backward propagation, linear programming ($k>m$) [10], etc.

Obtained ranges are either equal or edged in relation to real intensity ranges. On the other hand, one obtains the real values of range limitations using accurate solution methods for the system of linear equations with reference to the full set of boundary values (minimal and maximal) of matrix ***A*** parameters and a selected column of matrix ***B***. The creation strategy of setting up parameters ***A*** and ***B*** is analogous to filling the binary register as in Table 1.

Table 1. Creation strategy of parameter sets for technological intensity analysis, 0 - denotes range lower limit and 1- upper limit

a(1, 1)	a(1, 2)	...	a(1, k0	a(2, 1)	a(2, 2)	…	a(2, k)	…	….	a(m ,1)	a(m ,2)	…	a(m .k)	b(r, 1)	b(r, 2)	...	b(r, k)
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

3 Realization algorithm of task research of growth index and technological profitability

The goal of interference in the intensity control process is the maximization of technological growth and profitability by simultaneous minimization of outlays and intensity of implementations. Several of the above mentioned criteria induce one to choose the most important - or what is convenient for deterministic analysis - validity weights. Finally, the multicriterion notation of the goal function can be presented as follows:

$$z=\boldsymbol{w\_techn\_growth} * techn\_growth + \boldsymbol{w\_techn\_prof\_coeff} * techn\_prof\_coeff \quad (\mathbf{13})$$

$$\boldsymbol{w\_int} * \sum_{i=1}^{k} x(i) - \boldsymbol{w\_out} * \sum_{i=1}^{k}\sum_{j=1}^{m} a(i,j) + \boldsymbol{w\_result} * \sum_{i=1}^{k}\sum_{j=1}^{m} b(i,j) \rightarrow max$$

where:
w_techn_growth – weight of technological growth in optimization criterion,
w_techn_prof_coeff – weight of technological profitability coefficient,
w_int – weight of intensity level,
w_out – weight of outlays' level,
w_result – weight of technological result.

The fuzzy strategies certainly oblige one to consider the variability intervals of the individual parameters mentioned in the previous section. In this case one adds the fuzzy intervals to the price vector *p(1), p(2), ...,p(k):*

$$[pp(1),pk(1)], [pp(2),pk(2)], \ldots , [pp(k),pk(k)] \quad (\mathbf{14})$$

Example 1. Estimation of growth and implementation profitability of technology

Table 2. Technological outlays

matrix A											
0.15	0.45	0.66	0.24	0.91	0.38	0.28	0.16	0.39	0.26	0.4	0.94
0.19	0.05	0.78	0.55	0.21	0.65	0.74	0.62	0.53	0.27	0.76	0.07
0.77	0.35	0.81	0.64	0.39	0.69	0.03	0.27	0.56	0.18	0.68	0.25
0.97	0.13	0.79	0.59	0.91	0.61	0.36	0.12	0.02	0.33	0.53	0.09
0.87	0.27	0.96	0.79	0.97	0.1	0.87	0.74	0.52	0.65	0.45	0.45
0.98	0.49	0.83	0.58	0.7	0.43	0.37	0.48	0.35	0.19	0.3	0.63
0.01	0.59	0.54	0.04	0.45	0.73	0.49	0.49	0.35	0.04	0.23	0.15

Table 3. Results of technology implementation

matrix B											
0.36	0.55	0.74	0.58	0.92	0.9	0.54	0.68	0.57	0.9	0.91	0.26
0.59	0.06	0.21	0.08	0.86	0.98	0.56	0.36	0.07	0.24	0.35	0.9
0.7	0.62	0.09	0.52	0.51	0.83	0.62	0.98	0.26	0.09	0.64	0.64
0.28	0.64	0.49	0.62	0.07	0.37	0.66	0.29	0.25	0.19	0.18	0.01
0.69	0.81	0.46	0.82	0.95	0.71	0.44	0.57	0.69	0.12	0.04	0.43
0.69	0.17	0.54	0.34	0.9	0.33	0.76	0.49	0.8	0.33	0.13	0.14
0.07	0.77	0.55	0.45	0.86	0.46	0.36	0.06	0.6	0.12	0.52	0.33

Table 4. Technology implementation intensity

0.06	0.31	0.99	0.09	0.08	0.42	0.44	matrix X

Table 5. Table 6. Vectors for research of technological growth

vector X*A											
1.41	0.89	1.81	1.19	1.16	1.47	0.76	0.96	1.09	0.46	1.24	0.71

vector X*B											
1.29	1.19	0.75	1.04	1.66	1.62	1.39	1.43	0.98	0.44	1.1	1.16

Table 7. Table 8. Analysis results of technological and economic growth

	1	2	3	4	5	6	7	8	9	10	11	12
XB/ XA	0.92	1.35	0.41	0.87	1.44	1.1	1.84	1.49	0.9	0.95	0.88	1.64

0.41	min techn. growth
1.84	max economic growth

Table 9. Table 10. Table 11. Price vector and vectors for research of technological implementation profitability

price vector P	0.51	0.46	0.17	0.76	0.13	0.42	0.81	0.17	0.63	0.97	0.99	0.1

A*P	2.1	3.03	2.83	2.69	3.78	2.74	1.75

B*P	4.2	2.18	3.08	2.25	2.93	2.7	2.47

Table 12. Table 13. Technology implementation profitability indexes

BP / AP
2
0.72
1.09
0.84
0.78
0.99
1.41

0.72	min	profit.
2	max	profit.

Table 14. Weight values

weights				
growth	profitability	intensity	outlays	results
0.49	0.46	0.34	0.68	0.61

Table 15. Optimization coefficient of technology implementation *(13)*

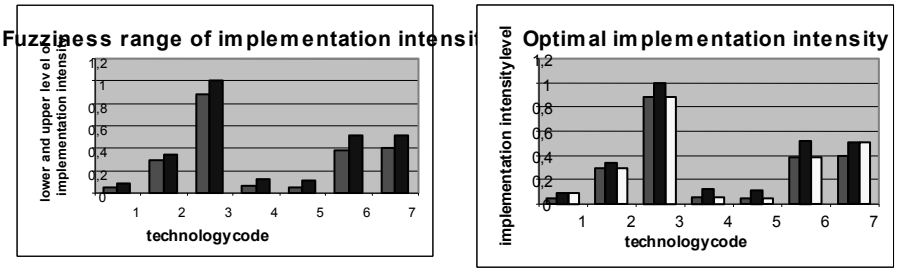
optim. coeff.
0.99

The row vectors in matrixes ***A*** and ***B*** (Tables 1 and 2) correspond with successive implemented technologies and the columns correspond with raw material outlays or products obtained as a result of the implementation of individual technologies. Table 3 includes the intensity levels ***X*** of the implementation of individual technologies.
Example 2

Table 16. Fuzziness of intensity level and its optimal level

Xb	Xe	Xopt
0.05	0.09	0.09
0.29	0.34	0.29
0.88	1	0.88
0.06	0.12	0.06
0.05	0.11	0.05
0.38	0.52	0.39
0.4	0.51	0.51

In this example one assumes fuzziness of the intensity level of technological implementations. The non-linear character of the model of technological outlays and effects allows one to conclude (the analysis results confirm it) that optimal intensity levels can be inconsistent with the lower or upper limit of its fuzziness (searching for the optimal solution was done by means of Solver's method). However, it is possible



for the sake of the occurrence of local monotonicity.

Fig. 1. Fuzziness range of implementation intensity and optimal levels of this parameter

Example 3. Optimization of technological outlays

In this example the intensity vector agrees with the optimal solution from Example 2 and is constant. Elements of the technological outlay matrix are subjected to optimization.

Table 17. Constant level of intensity (*X*)

0.09	0.29	0.88	0.06	0.05	0.39	0.51

Table 18. Optimal values of technology implementation outlays (*Aopt*)

-0.1	0.22	0.56	-0.1	0.84	0.17	-0.1	0.07	0.08	-0.2	-0.1	0.88
0.19	0.05	0.74	0.54	0.21	0.64	0.74	0.62	0.52	0.26	0.75	0.06
0.77	0.34	0.7	0.63	0.38	0.68	0.02	0.27	0.56	0.17	0.68	0.25
0.97	0.12	0.77	0.58	0.9	0.6	0.35	0.11	0.02	0.32	0.53	0.09
0.86	0.26	0.95	0.78	0.96	0.09	0.86	0.73	0.51	0.64	0.44	0.45
0.98	0.48	0.78	0.57	0.7	0.42	0.37	0.47	0.34	0.19	0.29	0.63
0.01	0.59	0.48	0.04	0.44	0.72	0.49	0.48	0.34	0.04	0.22	0.15

As the experiments show, by limiting of the technology outlay range *[Ap,Ak]*, it is enough to select the limit value which is the closest to the optimal value (if the optimal value lies outside the fuzziness interval) or optimal value (if it lies inside the technology outlay fuzziness):

$$afopt(i,j) \in [ap(i,j),ak(i,j)]; afopt(i,j);\ if\ aopt(i,j) \notin [ap(i,j),ak(i,j)] = \quad (15)$$

$$\begin{cases} ap(i,j)\ dla\ |ap(i,j)-aopt(i,j)| \le |ak(i,j)-aopt(i,j)| \\ ak(i,j)\ dla\ |ap(i,j)-aopt(i,j)| > |ak(i,j)-aopt(i,j)| \end{cases}$$

$$afopt(i,j) \in [ap(i,j),ak(i,j)]; \quad afopt(i,j) = aopt(i,j) \text{ if } aopt(i,j) \in [ap(i,j),ak(i,j)] \qquad (16)$$

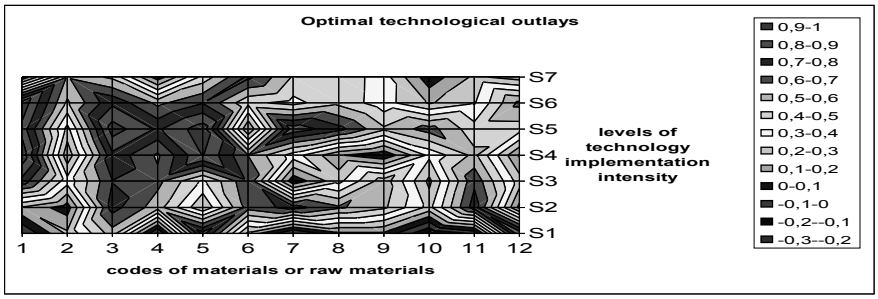


Fig. 2. Diagram of optimal levels of technological outlays for constant intensity values (*X*) and results of technology implementation (*B*)

4 Conclusions

1. If one has the possibility of controlling the fuzziness parameters of technology implementation outlays, it is possible to construct a criterion function similar to *(13)* and, on the basis of the optimization results, to determine directions of outlay quantities in order to increase their use efficiency *(16).*
2. Another way to influence technology implementation effects is to control the intensity vector elements (Example 2). Its influence on the growth and technology profitability is definitely (on average 8.5 – 10.5 times) smaller than the influence of the outlays (for the data from Examples 1, 2 and 3).
3. The control of technology implementation results is possible thanks to technical procedures and evidently influences the effects of growth and profitability of technology.

References

[1] Bamos A..: On Pseudo-Games. Annals of Mathematical Statistics, No. 39, 1968
[2] Binmore K., Kirman P., Tani P.: Frontiers and Game Theory. MIT Press, 1993
[3] Conway J.H.: On number and Games. Academic Press, 1976
[4] Ellison G.: Learning with One Rational Player. MIT Press, 1994
[5] Fudenberg D., Tirole J.: Game Theory. MIT Press, 1991
[6] Fundenberg D., Kreps D.: Lectures on Learning and Equilibrium in Strategic Form Games. Core Lecture Series, 1990
[7] Harsanyj J., Selten R.: A General Theory of Equilibrium Selection in Games. MIT Press, 1988
[8] Harsanyj J.: Games with Randomly Disturbed Payoffs. International Journal of Game Theory, No. 2, 1973
[9] Hofbauer J.: Stability for Best Response Dynamic. University of Viena, 1995
[10] Isaacs R.: Differential games. Wiley, 1965

Fuzzy Approach to Modelling of the Socio-Economic System

Q.C.Imanov[1], R.R.Rzayev[2]

Neuro-fuzzy economics laboratory of the Cybernetic Institute of the National Academy of Azerbaijan Republic, Az1141, F.Agaev str. 9, Azerbaijan Republic, Baku.
[1] korkmazi2000@yahoo.com; [2] raminrza@yahoo.com.

Abstract. The socio-economic system as an aggregate of economic system and social environment is investigated. A mathematical tool of the theory of fuzzy logic is proposed as a tool for analysis and forecasting of socio-economic system under influence of examined factors of uncertainty. Some fuzzy models for different level problems of economic system and social environment have been developed.

1 Introduction.

Socio-economic system (SES) as an aggregate of economic system and the social environment is considered. While the economic system covers production, distribution, an exchange and consumption of an gross product, the social environment covers sosaity, material and spiritual conditions of human existence. SES represents the hierarchical information system and majority of parameters have a non-numerical nature [1]. The final product of SES as an information system is the integrated population's life quality parameter of the state [2] formed on the basis of the following five integrated characteristics: population quality; well-being of the population; quality of social sphere; quality of ecological niche and natural-climatic conditions. As it is not difficult to notice, parameters of these characteristics can be the qualitative indexes formed, for example, by gradational scaling.

The information environment forming the population's life quality parameter is actually the economic system and the social environment of the state which determines economic behavior in production, distribution, an exchange and consumption. The given information environment differs by numerous factors of uncertainty, while majority of its parameters being qualitative, as a rule, are immeasurable. Therefore, for overcoming the complexities related to the record and processing of immeasurable data, in the present work we propose to use the mathematical tools of the fuzzy sets theory. In particular, for modeling of the level problems of SES we propose to represent qualitative indicators by linguistic variables accepting values as fuzzy term-sets [3]. In this case for definition of the population's life quality integrated parameter fully reflecting the level of SES development, it is possible to take advantage of fuzzy logic inference, or particularly the fuzzy model

based on linguistic rules. For achievement of isomorphism of this model the number of preconditions and consequences of fuzzy linguistic rules are being selected based on the requirements of the fuzzy environment, while its parameters such as those of membership functions of linguistic variables should be appropriately optimized.

Models at all levels of the economic system (micro, mezo, macro and meta) and model of economic behavior corresponding to its social environment are used as a source of information for such modeling. For studying of the social environment and consequences of its influence on global parameters of SES it is proposed to use the fuzzy model of economic behavior in production, distribution, an exchange and consumption of the gross product. Here fuzzy curves of consumer indifference and fuzzy utility functions from consumption sets are used as a tool for modeling.

Below the existing factors of uncertainty in SES are investigated and proceeding from this alternative fuzzy models for solution of the level problems of economic system are presented.

2 The Factors of Uncertainty in SES.

SES is an open, dynamic and ill-structured system [4] functioning in conditions of complete uncertainty and, secondly, there are no effective methods to overcome this uncertainty. Moreover, from 50% to 90% of the data in SES are non-numerical [1], i.e. they have uncertain or fuzzy nature. Therefore, it is quite justified that the further researches of SES underlie in the area of studying the nature of the present uncertainty. Uncertainty is a concept reflecting absence of unambiguity. There are two types of uncertainty: true uncertainty – stipulated by internal properties of investigated objects, and uncertainty related to incompleteness of information about these objects. As many researches have shown, there are both types of uncertainty in SES [5].

The first researcher of a phenomenon of uncertainty in SES was the American economist F.Knight [6], who particularly noted that for definition of a reasonable way of behaviour the economic subject should define causal relations between the decisions and their consequences. However by virtue of increased uncertainty of economic environment he can resort only to defining probabilities of event occurrence. However, the more we approach to innovative field, the more we face unique events, the mathematical whose calculation of probabilities is simply impossible. In this respect, F.Knight has proposed to name uncertainty of the first type (when there is some information about processes) as *a risk*, while for the second type he has kept the definition *an uncertainty* itself.

It is known that the basic sources of uncertainty in economy, as well as in other areas, are incompleteness or inadequacy of human knowledge about processes that take place and about the environment, and also contingency, which in similar conditions occurs unequally and which beforehand cannot be foreseen. In this case the research of uncertainty in its all appearances on SES scale is being conducted trough application of the alternative approaches.

Each SES is the complex self-organizing object developing under influence of many changing certain and uncertain factors – both internal and external. From the

factors forming a principle of uncertainty in the SES, it is possible to distinguish *economical*, *political*, *socio-psychological*, *technological* and *natural* components of uncertainty. Departing from general structure of functioning environment of the SES, it is possible to consider the appropriate factors of uncertainty. Among *political factors* of uncertainty we distinguish internal and external. Internal political factors include 1) legislative system of the state; 2) economic policy of the government; 3) social policy of the government; 4) activity of opposition forces; 5) ethno-political situation. The external political factors of uncertainty include set of conditions incorporated in general diplomatic, military-political and politico-economic groups.

Social-psychological factors of uncertainty are considered in aspects of economic and human behaviour. The classical types of economic behaviour are realized on various phases of the reproduction cycle. In general, economic behaviour is not the independent factor of development of economic life. It depends on a number of deeper factors of formation of economic culture and economic thinking, features of the existing systems of the economic and social relations. In the field of the human relations one might distinguish the whole group of uncertainty factors defined by value-normative space, with cultural features and state of the society. These factors reflect human, public and cultural measurements of the economic phenomena and, therefore, are uncertain.

The technological factors of uncertainty in SES are caused by functioning of existing and appearance of new engineering, and also with development of new technologies. The general direction of the development of a science and technology, especially on the near future, can be predicted with known accuracy. However, complete definition of specific consequences of those or other scientific findings and technical inventions is almost impossible. The technical progress is not reachable without risk and uncertainty.

Uncertainty *of natural environments* take place owing to spontaneity of the natural phenomena and, first of all, to those which result in various natural disasters, which can have strong negative influence on results of economic processes, and may become a source of the unforeseen costs. Furthermore, the special role in the modern world is played by ecological problems being the consequence of human's unsystematic activity which is destructive for the nature.

Thus, we see that under influence of the political, socio-psychological, technological and natural factors of uncertainty at the various levels SES (micro, mezo, macro and meta) there are various categories of uncertainty.

3 Some Fuzzy Models of SES.

At micro level of SES the task of positioning of the consumer market is considered, where in the specific market segment the rational choice of a competitive food commodity is carried out by a method of fuzzy logic inference. To realize of this task qualitative criteria of estimations of considered interchangeable food products are used. Fuzzification of their, the subsequent formation of relations of preference and a subset undominated alternatives are carried out on the basis of the mathematical tool of the fuzzy sets theory. Furthermore, on the example of the Azerbaijan consumer

market the model of consumer behavior is developed on the basis of application of the fuzzy logic inference mechanism. As curves of indifference of this model fuzzy sets are chosen, and utility from each chosen set of the goods and services (value of utility function) is determined by means of dot estimations of these fuzzy sets. In particular, suppose that we deal with a segment of food commodities of the consumer market, where space of possible sets of the goods is closed convex and continuous m-dimensional hyperspace $C = \left\{ X = (x_1, x_2, ..., x_m)^T \middle| x_j \geq 0, j = \overline{1,m} \right\}$, in which x_j denotes quantity of j-th commodity acquired by the consumer. Let's break all consumers in q linguistic categories by levels of interval incomes r_k $(k = \overline{1,q})$, part of which they are ready or capable to spend for the purchase of food commodities from the given segment. Then, taking into account "degraded" values of the prices p_j of the goods from a set X, the proposed fuzzy model in a general can look as:

$$\text{If } I = \tilde{r}_k \text{ and } P_1 = \tilde{p}_1 \text{ and } P_2 = \tilde{p}_2 \text{ and } \ldots \text{ and } P_m = \tilde{p}_m \text{ , then } IC = \tilde{l}_i \qquad \textbf{(1)}$$

where $\tilde{r}_k$ $(k = \overline{1,q})$ is the k-th fuzzy level of consumers' incomes; $\tilde{p}_j$ $(j = \overline{1,m})$ is the fuzzy level of the price of j-th commodity from a set X; $\tilde{l}_i$ $(i = \overline{1,s})$ is the i-th fuzzy level of indifference curve IC. Each of submitted in (1) levels has a non-numerical nature and can be expressed by terms such as, for example, "low", "average", "high", "lower than average", etc.

The choice of this approach is proved by the following reasons. During optimization of his choice each consumer can use a so-called map of indifference curves, each of which in classical interpretation represents a geometrical place of points (sets of commodities) in space whose dimension is defined by a number of the consumed goods. It is obvious that by virtue of limitation of the incomes among these indifference curves only one curve has a point (set) in which the consumer reaches the maximal utility from consumption of the appropriate set of goods. In a multivariate case it is a point of touch of the appropriate indifference curve with a hyper plane of the consumer incomes in space of the commodities prices. As it is known, each of indifference curves corresponds to the certain utility level, estimated in conventional units - utiles. However, we recognize that concept "utility" is rather a qualitative category than a quantitative one. Therefore, as criterion of utility it is proposed to use a linguistic variable whose values would be fuzzy term-sets. This in turn allows considering each indifference curve as fuzzy set. Among components of a basic vector of this set only for one, in which the curve of indifference touches the appropriate hyper plane of the consumer income, membership function will be equal to 1. In the proposed fuzzy model of consumption it is possible to regard linguistic variable "utility" as endogenous value. As exogenous values of considered model the consumer income and retail prices for the goods or services are chosen. These values are never strictly fixed. Each of them varies within the limits of the appropriate interval and consequently one has to use its average value. Finally this results in errors which at times do not provide required adequacy of model to consumer behavior. Therefore, we shall interpret these values also as linguistic variables with "degraded" values from the appropriate intervals. As a result with the help of the

acquired model it was possible to construct family of fuzzy indifference curves, whose defuzzyfied values of levels became conditional alternatives to utiles.

Key parameter describing mezo economic level of SES is the level of regional development (*RD*). It is determined trough indexes of *RD*, being, in essence, endogenous values of mezo-economic model. In this case, values such as levels of supply of the regional population with crops areas, production capacities, social objects, workplaces, natural resources, and also level of production volume per capita are exogenous. It is obvious, that it is not the full list of values of which finally level of *RD* develops. There are many factors of aesthetic, psychological and social character which can not be put on balance and can not be measured. In general they do not have expression in any units. The relative importance of such factors on regional space is determined by instinctive sympathies and antipathies which have developed as a result of nurture, education and a social environment of representatives of local population. Therefore, exact definition of level of *RD* generally represents the most complex problem of measurement and comparison of many diverse (incommensurable) variables of which, actually, parameters of mezo-economic model develop. As a result the estimation data of *RD* obtained in practice can not be considered absolute since the desirable result can not be determined without taking into account the designated factors above which it is impossible to carry out usual arithmetic operations. Moreover, even the listed above exogenous values of mezo models are not absolute – they always change in the certain limits. Therefore, for achievement of the greater adequacy it is more expedient to estimate parameters and variables of mezo models by trough intervals with their subsequent fuzzification.

For mezo level of SES the fuzzy productive model for definition of level of regional development *RD* can look as following:

$$\text{If } P_1 = \tilde{p}_1^{k_1} \text{ and } P_2 = \tilde{p}_2^{k_2} \text{ and } \ldots \text{ and } P_{13} = \tilde{p}_{13}^{k_{13}}, \text{ then } RD = \tilde{l}_j \qquad (2)$$

where P_i $(i = \overline{1,13})$ are linguistic variables describing the basic social and economic indices of regions; $\tilde{p}_i^{k_i}$ is the fuzzy term-set corresponding to k_i-th value of linguistic variable P_i; $\tilde{l}_j$ $(j = 1,2,...)$ is the fuzzy j-th level of regional development. Number of rules and amount of term-sets for each linguistic variable P_i are chosen on the basis of accessible heuristic knowledge. On the basis of this model and the expert socio-economic information on eleven regions of Azerbaijan the ranging of the appropriate regional levels of development is carried out.

On macro level we propose the fuzzy analogue for econometric model of definition of gross demastic product (GDP) whose level in the generalized form depends on investment volume, inflation rate and the oil price in the world market. In classical interpretation the parameter of GDP and the designated factors influencing its value are.averaged. Actually these values vary in the certain limits which in essence establish corresponding levels to them. Therefore, to increase degree of adequacy of model for definition of GDP it is proposed to use the mathematical tool of fuzzy sets trough fuzzification of elements and parameters of the model. In this case the fuzzy analogue of model of definition of GDP can be presented as the following linear dependence:

$$Q\tilde{D}P = \tilde{A}_0 + \tilde{A}_1 \cdot I\tilde{N}V + \tilde{A}_2 \cdot I\tilde{N}F + \tilde{A}_3 \cdot \tilde{P}_{\text{oil}} \quad (3)$$

where $Q\tilde{D}P$ is a required fuzzy parameter of GDP; $I\tilde{N}V$ is a fuzzy volume of investment contributions; $I\tilde{N}F$ is a fuzzy inflation rate; $\tilde{P}_{\text{oil}}$ is a fuzzy price level of one barrel of petroleum in the world market; $\tilde{A}_k$ $(k = \overline{0,3})$ are fuzzy parameters.

For definition of fuzzy level of investment contributions $I\tilde{N}V$ it is possible to take advantage of a fuzzy recurrent parity:

$$I\tilde{N}V = \tilde{B}_0 + \tilde{B}_1 \cdot \tilde{S}(t-1) + \tilde{B}_2 \cdot \tilde{S}(t-2) + \tilde{B}_3 \cdot \tilde{S}(t-3) \quad (4)$$

where $\tilde{S}(t-1), \tilde{S}(t-2), \tilde{S}(t-3)$ are fuzzy volumes of savings relative to the current t-th year relative to years $(t-1)$, $(t-2)$ and $(t-3)$; $\tilde{B}_k$ $(k = \overline{0,3})$ are fuzzy parameters.

To determine fuzzy analogue of inflation rate it is possible to apply the following fuzzy linear model:

$$I\tilde{N}F = \tilde{C}_0 + \tilde{C}_1 \cdot \tilde{M} \quad (5)$$

where $\tilde{M}$ is a fuzzy volume of money available in circulation; $\tilde{C}_{\text{o}}$ and $\tilde{C}_1$ are unknown fuzzy parameters.

Finally, for definition of the current fuzzy price level of one barrel of oil in the world market it is possible to take advantage of a fuzzy recurrent parity:

$$\tilde{P}_{\text{oil}}(t) = F\left(\tilde{P}_{\text{oil}}(t-1), \tilde{P}_{\text{oil}}(t-2), \tilde{P}_{\text{oil}}(t-3)\right) \quad (6)$$

where $\tilde{P}_{\text{oil}}(t-1), \tilde{P}_{\text{oil}}(t-2), \tilde{P}_{\text{oil}}(t-3)$ are the fuzzy price levels of one oil barrel relative to the current t-th year accordingly in years $(t-1)$, $(t-2)$ and $(t-3)$.

The next important problem of macro level of SES is development of the balance reflecting interbranch proportions. The exact formulation of the purpose this problem of input-output balance (IOB) and criteria of its achievement (of criterion function) generally represents a most complex problem of measurement and comparison of many diverse (incommensurable) variables of which parameters of its model develop. By virtue of these circumstances it is possible to replace classical model of IOB by fuzzy analogue, i.e. system of the algebraic equations with fuzzy variable and fuzzy parameters:

$$\tilde{X} = \tilde{A} \cdot \tilde{X} + \tilde{Y} \quad (7)$$

where $\tilde{Y} = (\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_n)$ is a vector of the final products expressed as fuzzy levels of volumes; $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$ is a set of fuzzy levels of volumes of production possible at the present resources; $\tilde{A} = (\tilde{a}_{ij})_{n \times n}$ is a matrix of fuzzy factors of direct expenses $\tilde{a}_{ij}$, which show fuzzy levels of production volumes of *i*-th branch necessary for production of one unit of production of *j*-th branch. Fuzzy problem of IOB (7) can be simulated and solved by means of a fuzzy neural network [7]. For this purpose we shall present this system in well caused, equivalent aspect:

$$\tilde{B} \cdot \tilde{X}^* = \tilde{Y} \qquad \textbf{(7')}$$

where $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$ is the required vector with fuzzy components; $\tilde{B} = [\tilde{b}_{ik}]_{i,k=1}^{n}$ is a fuzzy matrix with fuzzy elements $\tilde{b}_{ik} = \begin{cases} \tilde{a}_{ik}, & i \neq k \\ 1 - \tilde{a}_{ik}, & i = k \end{cases}$; $\tilde{Y}^T = (\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_n)^T$ is a fuzzy vector-column. As a fuzzy neural network we shall choose a two-layer network, whose weights of internal connections will be required fuzzy numbers $\tilde{x}_k \, (k = \overline{1,n})$. For their finding, for example, it is possible to choose "error backpropagation" learning algorithm and to realize it under the circuit of the neuro-identifier (Fig. 1). In its output the fuzzy neural network induces a signal $\tilde{z}_i = \sum_{k=1}^{n} \tilde{a}_{ik} \tilde{x}_k \, (i = \overline{1,n})$. Herewith, as criterion of its comparison with a desirable target $\tilde{y}_i$ it is possible to use a fuzzy measure of comparison $E(\tilde{z}_i, \tilde{y}_i)$ [7]. Assuming that the optimum fuzzy output of a neural network finally should satisfy the condition $E(\tilde{z}_i, \tilde{y}_i) = 1$, trough "error backpropagation" learning algorithm it is possible to adjust fuzzy parameters $\tilde{x}_k \; (k = \overline{1,n})$ so that this condition on the certain steps of training is accomplished. For realization of the arithmetic operations stipulated by these algorithm it is possible to use, for example [8] in which the techniques of operations on fuzzy sets is considered. The proposed fuzzy model of IOB was confirmed on data of four branches of Azerbaijan economy.

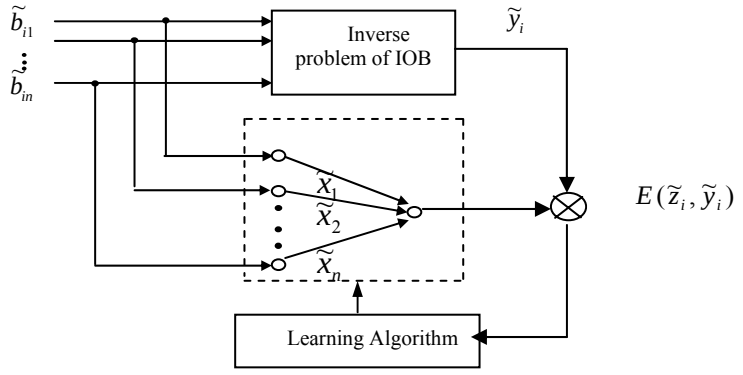


Fig. 1. Neural identification of fuzzy solutions of IOB problems

Other approach to the decision of the fuzzy problem of IOB is based on finding of an inverse matrix. In this case decision of the fuzzy problem (7) relative of $\tilde{X}$ in the matrix form is expressed as:

$$\tilde{X} = (E - \tilde{A})^{-1} \cdot \tilde{Y} \qquad (8)$$

or in the extended expression as:

$$x_{ik} = \sum_{j=1}^{n} \beta_{ij} y_{jk}\ ;\ i,k = \overline{1,n} \qquad (8')$$

where β_{ij} are the fuzzy numbers of the inverse matrix $(E - \tilde{A})^{-1}$. As is known, the scheme of construction of the inverse matrix including calculation of the matrix determinant $|E - \tilde{A}|$ and algebraic adjuncts of the transposed elements $\tilde{B}_{ji}$, is realized by formula:

$$\beta_{ij} = \frac{\tilde{B}_{ji}}{det(E - \tilde{A})}; i,j = \overline{1,n} \qquad (9)$$

To find β_{ij} it is possible to take advantage of the construction of "linearized history" of the fuzzy number and its corresponding algebra [10].

For meta level of SES the alternative approach for estimation of the global level of state development based on fuzzy logic inferences is considered. On the basis of the proposed fuzzy model and on the basis of researches which have been done within the framework of the United Nations Development Program, the estimation of global level of development of Azerbaijan is carried out. Endogenous value of the given model is the quality indicator "global level of development" (GD) of SES and exogenous values are linguistic variables: "index of human development" (HDI) and "index of technological achievements" (TAI), accepting values in the fuzzy term-sets determined by gradational scales. On their basis and based on the chosen dominant judgments fuzzy linguistic rules for the estimation of GD are formed as:

$$\text{If } HDI = \tilde{A}_i \text{ and } TAI = \tilde{B}_j \text{, then } GD = \tilde{C}_k \qquad (10)$$

where $\tilde{A}_i$ ($i = 1,2,...$) is the fuzzy term-set corresponding to i-th value of linguistic variable HDI; $\tilde{B}_j$ ($j = 1,2,...$) is the fuzzy term-set corresponding to j-th value of linguistic variable TAI; $\tilde{C}_k$ ($k = 1,2,...$) is the fuzzy term-set corresponding to k-th value of linguistic variable GD.

Conclusion.

On the basis of the performed research of uncertainty factors under whose influence the socio-economic system functions, we propose the fuzzy models for solution of some problems of economic system on micro, mezo, macro and meta levels and of socio-economic behavior. The carried out researches do not apply for the completed

character. For full research it is necessary to develop the integrated system of models of economic system and the social environment.

References

1. A.I. Orlov: Razmitiye Tseni. Nechislovaya ekonomika i upravleniye investisionnim processom. «Rossiyskoye Predprinimatelstvo», Journal 2001, #12 pp.103-108, (in Russian).
2. S.A.Ayvazian: K Metodologiyi izmereniya sinteticheskih kategoriy kachestva jizni naseleniya. Ekonomika i matematicheskiye metodi. 2003, b. 39, #2, pp. 33-53, (in Russian).
3. L.A.Zadeh: The concept of a linguistic variable and its application to approximate reasoning. American Elsevier Publishing Company, New York, 1973.
4. H.Simon: The Structure of ill-structured Problems. Artifical Intelligence. 1973. Vol. 4.
5. Matematika i kibernetika v ekonomike. Slovar-spravochnik. M., Ekonomika, 1975 (in Russian).
6. F.Knight: Risk, Uncertainty and Profit. Boston, 1921, pp. 210-232.
7. Fuzzy Sets, Neural Networks, and Soft Computing. Edited by Yager R.R., L.A.Zadeh. Van Nostrand Reinhold, 1994, pp.440.
8. R.A.Aliev, B.Fazlollahi, R.R.Aliev: Soft Computing and its Applications in Business and Economics, Springer-Verlag Berlin Heidelberg, 2004.
9. A.V. Yevdokimov: Metod nechetkoy linearizasii dla chislennogo resheniya algebrayicheskih i differencialnih uravneniy. Electronic journal «Issledovano v Rossiyi» http://zhurnal.ape.relarn.ru. (in Russian).
10. Q.C.Imanov: Soft Computing in Models of Socioeconomic Systems Management. ICCI 2004, Nicosia, North Cyprus, Near East University, May 27-29 2004, pp.71-75.

Value at Risk Implication Before and After the Euro Introduction

Juan Pineiro Chousa[1], Artur Tamazian[1], and Davit N. Melikyan[2]

[1] University of Santiago de Compostela,
Avda. Juan XXIII, s/n. 15782 Santiago de Compostela,
efjpch@usc.es, oartur@usc.es
[2] AEPLAC
28, Charents Str., 375025, Yerevan,
dmelikyan@aeplac.org

Abstract. To evaluate the goodness of a VaR model, banks as well as regulators use backtesting to confirm their judgments. We applied variance techniques of daily VaR estimation to evaluate changes in financial risk caused by Euro introduction. The analysis covers 10 European indices and neutral DJIA as a mirror for common world developments. Estimations are made on 1000 ex-ante and 1000 ex-post data points and backtested on the next 250 for each index by Kupiec's (1995) methodology. At this stage of the ongoing research it is already clear that in general VaR has grown significantly after introducing Euro, which in turn claims for new commercial bank capital requirements according to Basle Accord.

1 INTRODUCTION

In the last years we have experienced a "perfect storm" in the confluence of a stock market bubble that favoured stock-based incentive compensation, increased focus on short-term stock market movements, research reports written by conflicted analysts, audits performed by accountants focused on growing their consulting businesses. These circumstances have grown the interest in the study of the financial market risk (Jorion, 1995).

Financial risk, caused by movements in financial markets, is one of the three types of risk distinguished in the financial literature. In turn financial risk is broken down further into different categories. Among them the market risk brought on by changes in the prices of financial assets and liabilities is our target. Uniform methodology of risk measurement called Value-at-Risk (VaR) has received a great attention both from regulatory and academic fronts. It was made popular by US investment bank J.P. Morgan (1996), who incorporated it in their risk management model RiskMetricsTM. During a short span of time, a serious number of papers have studied various aspects of VaR methodology (Hull and White, 1997), (Goorberg and Vlaar, 1999; Lee and Saltoglu, 2002). The availability of information from financial markets allows us to empirically examine this type of risk better than any other kind. VaR is defined as the maximum potential change in value of a portfolio of financial instruments with a given probability over a certain time horizon, with assumption that the composition of the portfolio remains the same. We will return to this indicator's definition and to its detailed characteristics a little later in our next chapters.

VaR measures can have many applications, such as in risk management, to evaluate the performance of risk takers and for regulatory requirements. In particular, the Basel Committee on Banking Supervision (1996) at the Bank for International Settlements imposes to financial institutions such as banks and investment firms to meet capital requirements based on VaR estimates. Providing accurate estimates is of crucial importance. If the underlying risk is not properly estimated, this may lead to a sub-optimal capital allocation with consequences on the profitability or the financial stability of the institutions.

We did not try to find the best possible model for VaR estimation as many other authors do before, making the process into Holy Grail search. We used variance techniques of VaR estimation previously used by the other authors (Goorberg and Vlaar (1999)) and applied them on the exante and expost periods of Euro introduction. We evaluated VaR using Maximum Log-likelihood method for normally and t-student distributed returns, along with RiskMetrics and GARCH(1.1) models with Gaussian innovations. The analysis cover 10 European indexes corresponding to Netherlands, Austria, Belgium, Germany, France, UK, Spain, Denmark, Italy and Swiss (AEX, ATX, BFX, DAX, FCHI, FTSE, IBEX, KFX, MIB30, SSMI) and DJIA for USA, covering larger geographical area than

most other related studies. Kupiec's (1995) back-testing procedure is run for the estimation results. We compare ex-ante and ex-post VaR estimations to evaluate the "Euro effect". Then several ideas are discussed to characterize changes in world financial markets caused by Euro.

The outline of the paper is as follows: Section 2 revisits the general view about Value at Risk. In the 3rd section VaR regulation is described. Section 4 examines different techniques for VaR computing. Section 5 summarizes the empirical results of VaR estimates. Kupiec backtest is the issue of the 6th section and finally in the 7th section present our conclusions.

2 VALUE AT RISK: DEFINITION

Three aspects need to be kept in mind when judging the Value-at-Risk of a portfolio. In the first place, we need to know the initial value of the portfolio. For analytical purposes, the initial portfolio value is usually normalized to 100 currency units, but it could be any other amount, of course. A second ingredient is the holding period to which the VaR pertains. And finally, the confidence level is of importance. Evidently, the higher the confidence level the larger the Value-at-Risk of the portfolio. By varying the confidence level, one is able to explore a whole risk profile, i.e. the entire distribution of results is revealed.

Defined as the maximum potential change in value of a portfolio of financial instruments with a given probability over a certain time horizon VaR is a little difficult to estimate. However, VaR indicator is suitable to be defined analytically. It just runs into the idea of probability distribution.

$$\Pr[r_t < -VaR\,(h)] = \alpha \tag{1}$$

where r_t is the return at time t, $\mathtt{V}_0$ and $\mathtt{V}_t$ are initial and time t values of the asset (portfolio, derivative, etc.) respectively, is the left tail probability and VaR(h) is the VaR for time horizon h. We shall evaluate and compare daily VaR, so in our observation $h = 1$

Analytically, the VaR is defined by the top limit of integral of the function of expected returns r(s):

$$\alpha = \int_{-\infty}^{E(r)-VaR} r\,(s)\,ds \tag{2}$$

Usually it is assumed that the expected value of the returns is zero so we can transform (2) into

$$\alpha = \int_{-\infty}^{-VaR} r\,(s)\,ds \tag{3}$$

An alternative representation consists of considering the VaR through the following expression:

$$VAR(h) = \alpha\sqrt{\sigma^2 h} \tag{4}$$

where is the factor that defines the area of returns, σ^2 is returns variance and h is the time horizon for which the factor of risk will be calculated.

Value at Risk (VaR) concept, or valuation of the risk, comes of the need from quantifying with certain level of significance or uncertainty the amount or percentage of loss that a portfolio will face in a predefined period of time (Jorion 2000, Penza and Bansal 2001). Its measurement has statistical foundations and the standard of the industry is to calculate the VaR with a significance level of 5%. This means that only 5% of times, or 1 of 20 times we can consider too that once a month with daily information, or once every five months with weekly information the return of the portfolio will fall down of what indicates the VaR, in relation with the expected return.

If we consider a series of historical returns of a portfolio that has N number of assets, it is feasible to visualize the distribution of density of those returns across the analysis of the histogram. It is common to find fluctuations of returns around an average slightly different value of zero, in other words, to find mean reversion process and which distribution comes closer one normally. Skewness is sometimes perceived in the returns and from a practical point of view it is not so realistic to assume symmetry in the distribution. That is why the assumption about the distribution, in order to compute the Value at Risk, is a very important issue.

3 VALUE AT RISK REGULATION: MARKET AND ASSET LIQUIDITY RISKS

Market risk refers to the potential losses arising from the changes in the value or price of an asset, such as those resulting from fluctuations in interest rates, currency exchange rates, stock prices and commodity prices. Asset liquidity risk is clearly allied with market risk and represents the risk that an entity will be unable to unwind a position in a particular financial instrument at or near its market value because of a lack of depth or disruption in the market for that instrument.

Market risk together with liquidity risk, are the most important risks for securities firms, which typically operate on a fully mark-to-market basis. Securities firms, which engage in the business of

underwriting, trading, and dealing in securities, must necessarily maintain proprietary positions in a wide range of financial instruments. Therefore, the aim of such firms is not to eliminate all market risk, but rather to manage it to a level at which acceptable returns, net of market losses, can be generated.

Market risk is also important for banks and their affiliates that hold significant positions that are marked to market. Banks typically manage market and liquidity risks associated with such positions in the same manner and with the same kinds of tools as their securities firm counterparts. The situation is somewhat different in regard to assets that the firms intend to hold to maturity and may be illiquid. Insurance companies are also subject to market risks. Here, such risks are generally classified as asset or investment risks in insurance activities. The investment of premiums must generate income and have a realizable liquidation value sufficient to meet the firms' liabilities. Shifts in market prices could affect achievement of this objective.

Most securities firms and banks, together with insurance companies running significant trading positions, use statistical models to calculate how the prices and values of assets are potentially impacted by the various market risk factors. These models generate a Value-at Risk estimate of the largest potential loss the firm could incur, given its current portfolio of financial instruments. More precisely, the VaR number is an estimate of maximum potential loss to be expected over a given period a certain percentage of the time.

A number of vast VaR models depend on statistical analysis of past price movements that determine returns on the assets. The VaR approach evaluates how prices and price volatility behaved in the past to determine the range of price movements or risks that might occur in the future. These models are commonly back-tested to evaluate the accuracy of the assumptions by comparing predictions with actual trading results. In practice, while VaR models provide a convenient methodology for quantifying market risks and are helpful in monitoring and limiting market risk.

The Basel Committee has developed a so-called "internal models" approach to the calculation of a market risk capital charge. For those banks that meet a series of qualifying criteria, this approach effectively relies on their own value-at-risk calculations of market risk. Banks and securities firms choosing to use an internally developed VaR model to calculate market risk capital charges must demonstrate to their supervisor that their model meets minimum qualitative and quantitative standards, including incorporation of VaR into the firm's daily risk management process, back-testing to determine the precision of the model and continuous adjustment of the latter.

For back-testing proposal we used Kupiec (1995) methodology performed according to the Basel Internal Model approach regulations.

4 MODELING VAR: EMPLOYED VARIANCE METHODS

To measure VaR we used so-called variance methods based on some assumption concerning distribution of returns. Applied methods include both static Maximum Log-likelihood method and non-static methods [Risk Metrics and GARCH (1.1)], which take into account volatility clustering phenomenon.

4.1 Normality Assumption

One of the static models that we use in our paper is that assumed that high frequency financial return data have fatter tails than can be explained by the normal distribution, this artefact seems odd. However, the normal distribution entails some very convenient characteristics that do not carry over to other distributions.

First of all the parameters of normal distribution are easier to estimate as there is often an analytical solution for them. One of the other advantages is the additivity of the normal distribution and this characteristic is especially important for the calculation of the multi-day VaR based on one-day VaR.

Hence, if we assume independence of normally distributed returns and a mean return of zero (Figlewski,1994) it can be show that:

$$\mathtt{VAR}^{(T)} = \sqrt{T}\mathtt{VAR} \tag{5}$$

Also assuming i.i.d log-returns we can express the Value at Risk as:

$$VAR = -\mathtt{V}(e^{\mu+\sigma\phi^{-1}(\alpha)} - 1) \tag{6}$$

where V represents the initial value of our portfolio and $\phi(\cdot)$ is the cumulative distribution function of the standard normal probability distribution. For this model we can estimate the parameters of μ and σ by means of normal distribution's log-likelihood function maximization.

4.2 t-Student Distribution Assumption

As we can see the results related on the application of the first model based on the assumption that our portfolio returns follow normal distribution are underestimates the portfolio risk. So it seems obvious to try the estimation with the consumption that the log returns of portfolio are Student-t distributed, as it is known that Student-t distribution is the most suitable for VaR estimation because of its fatter tails (Goorberg and Vlaar,1999).

The Student-t probability distribution has three main characteristics which are the scale ($\gamma > 0$) degrees of freedom ($\upsilon > 0$) and location parameter (μ). At this stage we can say that any variable distributed by means of Student-t distribution has a variance $\frac{\upsilon\gamma^2}{\upsilon-2}$ for $\upsilon > 2$[3], mean (μ) and provided $(\upsilon > 1)$[4].

Hence, assuming that our portfolio log returns follow t-distribution the log likelihood function is given by:

$$ML_{t-dist} = T\left[\log\Gamma(\tfrac{\upsilon+1}{2}) - \log\Gamma(\tfrac{\upsilon}{2}) - \tfrac{1}{2}\log\pi\upsilon - \log\gamma\right] - \tfrac{\upsilon+1}{2}\sum_{t=1}^{T}\log[1 + (\tfrac{r_t-\mu}{\gamma\sqrt{\upsilon}})^2] \quad (7)$$

Gamma function is defined as $\Gamma(\xi) = \int_0^\infty e^{-x}\mathtt{x}^{\xi-1}\,dx$ and for this case the maximum likelihood optimization has to be done numerically. There is no analytical expression for γ, υ and μ. The Value at Risk for Student-t distribution assumption can be expressed as:

$$VAR = -V(e^{\mu+\gamma F_\nu^{-1}(\alpha)} - 1) \quad (8)$$

where $F(\cdot)$ is the cumulative distribution function of a standardized t-distributed random variable.

4.3 Non static models

The static models are that they do not take the volatility clustering into account. By far the most popular model to model this phenomenon is the so called Generalized Autoregressive Conditional Heteroskedasticity, or GARCH, model introduced by Bollerslev (1986). It is an extension of the Autoregressive Conditional Heteroskedasticity, or ARCH, model by Engle (1982). In the GARCH model we start by defining an innovation η_{t+1}, i.e., some random variable with mean zero conditional on time $\mathtt{t}$ information, I_t. This time t information is a set including the innovation at time t, $\eta_t \in I_t$, and all previous innovations, but any other variable available at time t as well. In finance theory, η_{t+1} might be the innovation in a portfolio return. In order to capture serial correlation of volatility, or volatility clustering, the GARCH model assumes that the conditional variance of the innovations depends on the latest past squared innovations as is the assumption in the less general ARCH model, possibly augmented by the previous conditional variances. In its most general form, the model is called GARCH(p,q), and it can be written as:

$$\sigma_t^2 = \omega + \sum_{j=1}^{p}\beta_j\sigma_{t-1}^2 + \sum_{i=1}^{q}\alpha_i\eta_{t-i+1}^2 \quad (9)$$

p lags are included in the conditional variance, and q lags are included in the squared innovations. In this section, we shall regard these innovations as deviations from some constant mean portfolio return:

$$\mathtt{r}_{t+1} = \mu + \eta_{t+1} \quad (10)$$

so that σ_t^2 is also the conditional variance of the portfolio returns. We can write the innovation η_{t+1} as $\sigma_t\varepsilon_{t+1}$, where ε_{t+1} is assumed to follow some probability distribution with zero mean and unit variance, such as the standard normal distribution.

A great many empirical studies have proved it unnecessary to include more than one lag in the conditional variance, and one lag in the squared innovations. This is why our point of departure will be the GARCH(1,1) model:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha\eta_t^2 \quad (11)$$

with $\omega > 0$, $\beta \geq 0$ and $\alpha \geq 0$ to ensure positive variances. If the market was volatile in the current period, next period's variance will be high, which is intensified or offset in accordance with the magnitude of the return deviation this period. If, on the other hand, today's volatility was relatively low, tomorrow's volatility will be low as well, unless today's portfolio return deviates from its mean considerably. Naturally, the impact of these effects hinges on the parameter values. Note that for $\alpha + \beta < 1$, the conditional variance exhibits mean reversion, i.e., after a shock it will

[3] For $\upsilon \to \infty$ the t-distribution is reduced to the normal one with mean (μ) and variance (γ^2)

[4] The smaller υ gets, the fatter the tails are

eventually return to its unconditional mean $\frac{\omega}{(1-\alpha-\beta)}$. If $\alpha+\beta=1$ this is not the case, and we have persistence.

In order to estimate these parameters by means of likelihood maximization, one has to make assumptions about the probability distribution of the portfolio return innovations η_{t+1}. We shall consider Gaussian innovations.

$$\varepsilon_t \sim \mathtt{N}(0,1), \qquad \eta_{t+1}|\mathtt{I}_t \sim \mathtt{N}(0,\sigma_t^2) \tag{12}$$

leading to a conditional log likelihood of η_{t+1} equal to

$$l_t(\eta_{t+1}) = -\log\sqrt{2\pi} - \tfrac{1}{2}\log\sigma_t^2 - \frac{\eta_{t+1}^2}{2\sigma_t^2} \tag{13}$$

The log likelihood of the whole series $\eta_1;\eta_2;:::;\eta_T$ is

$$\mathtt{L} = \sum_{t=1}^{T} l_t(\eta_{t+1}) \tag{14}$$

As an alternative to the GARCH(1,1) model, or as a special case of this model, US investment bank J.P. Morgan introduced $\mathtt{RiskMetrics}^{TM}$, a VaR assessment method that basically restricts both μ and ω to 0, and α to $1-\beta$ in formula (11). The parameter β, called the decay factor and renamed λ, is set at 0.94 for daily data. This makes estimation elementary, since there are no parameters to estimate left. The portfolio return variance conditional on time t information is just:

$$\sigma_t^2 = \lambda\,\sigma_{t-1}^2 + (\mathtt{1}-\lambda)\,r_t^2 \tag{15}$$

or

$$\sigma_t^2 = \lambda^t\,\sigma_0^2 + (\mathtt{1}-\lambda)\sum_{k=0}^{t-1}\lambda^k\,r_{t-k}^2 \tag{16}$$

From the formulae above, it is clear that the conditional variances are modeled using an exponentially weighted moving average: the forecast for time t is a weighted average of the previous forecast, using weight λ, and of the latest squared innovation, using weight $\mathtt{1}-\lambda$. The $\mathtt{RiskMetrics}^{TM}$ approach essentially boils down to keeping track of the return data and using these along with the decay factor to update the conditional volatility estimates.

The imposition of the restriction that α and β should sum to unity implies persistence in the conditional variance, i.e., a shock moving the conditional variance to a higher level does not die out over time but 'lasts forever'. If there are no shocks offsetting this volatility increase, the conditional variance will remain high, and will not display mean-reverting behavior.

As you will make sure the non-static models can better predict VaR than the comparably naïve static ones. All these methods are applied to those 11 indexes under observation.

5 ESTIMATING THE DAILY VAR

Before estimating VaR it is suitable to verify the existence of autocorrelation in the daily returns of our time series for every eleven indexes. The correlograms of the returns suggest hardly any evidence of autocorrelation in the first six lags (the results can be provided upon request).

Therefore, following the techniques described in the past section of the paper we compute VaR estimation for 11 countries. The analysis cover 10 European indexes and 1 USA index: the time series of every country comprises data from the end of January of 1994 till December of 2003 or 2500 daily data. Dividing our data in two subsets, before and after Euro introduction in the EU markets, obtain 1250 daily data for every period. Hence, the estimation sample consists in 1000 daily data and 250 for evaluation sample, for both periods. For the purpose of not overloading the paper we attach VaR estimates results only for left tail probability of 1% (see table 1).

Table 1. Ex-ante and Ex-post VaRs for 1% Confidence Level[a]

VaR ESTIMATES	*AEX*	*ATX*	*BFX*	*DJIA*	*FCHI*	*FTSE*	*IBEX*	*KFX*	*MIB30*	*SSMI*	*DAX*
Normal (ex-ante)	-2,39	-2,29	-1,87	-2,03	-2,68	-1,85	-2,67	-1,93	-3,21	-2,33	-2,65
Normal (ex-post)	-3,94	-2,27	-3,08	-3,15	-3,95	-3,19	-3,82	-3,06	-3,76	-3,23	-4,37
Student-t (ex-ante)	-2,74	-2,62	-2,12	-2,32	-2,86	-1,98	-2,95	-2,12	-3,43	-2,62	-3,02
Student-t (ex-post)	-4,53	-2,53	-3,55	-3,47	-4,38	-3,52	-4,10	-3,40	-3,43	-3,71	-4,88
RM (ex-ante)	-3,69	-3,50	-2,91	-2,67	-3,39	-2,83	-3,33	-2,82	-3,80	-2,71	-3,88
RM (ex-post)	-5,42	-2,17	-3,66	-3,22	-5,11	-3,49	-3,91	-2,60	-4,20	-3,80	-6,30
GARCH(1,1) (ex-ante)	-3,59	-2,40	-2,83	-2,40	-3,69	-2,87	-2,76	-2,74	-4,37	-2,65	-3,33
GARCH(1,1) (ex-post)	-4,68	-2,42	-3,66	-3,07	-4,55	-3,46	-3,77	-2,57	-4,20	-3,62	-6,06

[a] The results for the rest confidence levels are excluded in order to not overloading the paper but may be provided upon request.

6 KUPIEC BACK-TESTING OF THE "INTERNAL MODEL"

Once we have estimated VaR our goal is focused on Kupiec back-test which is one of the most popular tests applied in backtesting. The likelihood ratio developed by, will be used to conclude if the VaR model has to be rejected or otherwise has to be accepted. For this proposal we split our indexes returns into an estimation sample and an evaluation sample.

The estimation sample were used to estimate the model and to predict the VaR, were upon a statistically back-test is applied by means of the evaluation sample in order to find the adequacy of the model.

Be N the number of failures (it is the number of cases in which loss exceeds the one forecasted by VaR model) in a sample of size T, then the number of VaR violations follows a binomial distribution and the failure rate N/T should be equal to the left tail probability p.

The likelihood ratio statistic is:

$LR = -2Ln[(1-p^*)^{T-N}(p^*)^N] + 2Ln[(1-\frac{N}{T})^{T-N}(\frac{N}{T})^N$ (17)

with the null hypotheses: H_0: N/T=p, $H_1 : N/T \neq p$.

The likelihood ratio is asymptotically χ^2 distributed under the null hypotheses that is the true probability the VaR is exceeded.

Thus, we can construct rejection/acceptance intervals with certain confidence level that advice us whether a model has to be rejected or not. Conventionally the confidence level is set to 0.05. The table 2 shows the rejection/acceptance regions according to the likelihood ratio statistic.

Table 2. Acceptance/Rejection regions[a]

Left Tail Prob.	EVALUATION SAMPLE SIZE (T)				
	250	500	750	1000	1250
0.05	7<N<19	17<N<35	27<N<49	38<N<64	60<N<92
0.01	1<N<6	2<N<9	3<N<13	5<N<16	9<N<23
0.005	0<N<4	1<N<6	1<N<8	2<N<9	3<N<13
0.001	0<N<1	0<N<2	0<N<3	0<N<3	0<N<4

[a] The test size is 5%

Also, in order to not overloading the paper, as in the past section, and according to the rejection/acceptance regions we present the numerical results for the failure rates for 0.05, 0.01, 0.005, 0.001 and 0.0001 left tail probabilities which are tabulated at the end of the paper (see table 11).

According to the results, the pre-Euro period presents violations in case of Normal, t-Student, Riskmetrics and Garch(1,1) . The assumption of normality is rejected at almost all left tail probability levels, except for DJIA in which one is accepted at 5% of left tail probability. In case of t-Student assumption we improve the failure rates time due to the fatter tails that generates more realizations in the tails than is to be expected on the basis of a normal distribution. The accepted regions for t-Student distribution mostly corresponds to 0,1%. The last two assumption, Riskmetrics and Garch(1,1), fit better. Kupiec back-test does not reject them mostly for 5%, 1% and 0,1% for 11 countries (see table 11).

As can be observed not the same occurs for post Euro period. Both, Normal and t-Student assumption improve VaR estimates compared with the previous period but still exists regions in which ones VaR estimates are rejected according to the Kupiec back-test. On the contrary, the back-test results for Riskmetrics and Garch(1,1) models are very successful. They are accepted at all levels and for entire group of countries (see table 11).

Other of our remarks is the comparison of the financial risk before and after the Euro. We used average VaR arrived by different methods for 1% confidence level. As we one can see from the figure 1 only in Greece the average VaR was reduced. In all remaining countries VaR trends to grow after Euro Zone establishment. The largest jump in VaR was registered in Germany, Netherlands and France.

Therefore, after Euro introduction the world financial markets have a quite different model with higher but more predictable Value at Risk.

7 CONCLUSION

After introducing the Euro Europe has pretended to deepen the economic integration. The most important component of the economy is the financial market. Thus we tried to evaluate the effect of

Euro on the financial markets stability through VaR estimations. We conclude that after Euro VaR has grown in European and world financial markets as well. Only in Greece the average VaR was reduced. In all remaining observed countries VaR trends to grow after Euro Zone establishment. The largest jump in VaR was registered in Germany, Netherlands and France. Though it is a bit difficult to be sure that the Euro zone establishment is the very source for those changes, several hypotheses can be suggested however.

E.g. a common currency for EU member states, the Euro/Dollar exchange rate now concerns entire competitiveness of the both industrial giants of the world - EU and USA. The same is true also for relative competitiveness of EU products compared with any other trade partners of its own. The often changes in competitiveness may by the events preconditioning changes in prices of company shares all over the world. So slight changes in exchange rates now are more significant for the indices under observation. The often fluctuations in exchange rates after establishment of Euro can explain the higher VaR indicated after creation of Euro zone. In fact, in spite of higher financial risk, now it is more predictable, as it is obvious from Kupiec back-testing of our VaR measures. This fact makes other explanations of VaR levels jump even less persuasive.

Higher VaR does not mean that we must loose in market regulation. Only accurate evaluation of potential risk makes it harmless for the business. Concluding remarks we have to state that non static models of VaR estimation (GARCH(1.1) and RiskMetrics) are more suitable and non rejectable on all the left tail probability confidence levels. However after the Euro the static models also became more efficient, characterizing with lower failure rates.

References

[] Basle Committee on Banking Supervision, 1996, Supplement to the Capital Accord to Incorporate Market Risks, Basle: Basle Committee on Banking Supervision.

Bollerslev, T. (1986), Generalized Autogreressive Conditional Heteroskedasticity, Journal of Econometrics, 31, 307-327.

Engle, R. (1982), Autogregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation, Econometrica, 50, 987-1008.

Figlewski, S. (1994), Forecasting Volatility Using Historical Data, New York University, Working Paper 13.

Jorion, P. (1997), Value at Risk, Irvine, Chicago.

Jorion, P. (2000), Value-at-Risk: The New Benchmark for Managing Financial Risk, McGraw-Hill.

J.P. Morgan Bank (1996), RiskMetricsTM Technical Document (4 ed.), New York: J.P. Morgan Bank.

Hull, John and Alan White, (1998), Value at Risk When Daily Changes in Market Variables are not Normally Distributed" Journal of Derivatives, 5, 9-19.

Hull, J. and A. White (1998), Incorporating Volatility updating into Historical Simulation Method for Value at Risk, Journal of Risk, 1, 5-19.

Kupiec, P. (1995), Techniques for Verifying the Accuracy of Risk Measurement Models, Journal of Derivatives, 2, 73-84.

Lee and Saltoglu (2002), Assessing the Risk Forecasts for Japan, Japan and the World Economy, 14, 1, 63-87.

McNeil, A.J. and R. Frey (2000), Estimation of Tail-related Risk Measures for Heteroskedastic Financial Time Series: An Extreme Value Approach", Journal of Empirical Finance, 7, 271-300.

Penza and Bansal (2001), "Measuring Market Risk with Value at Risk", John Wiley and Sons.

Rockinger, M. and E. Jondeau (2002), Entropy Densities with an Application to Autoregressive Conditional Skewness and Kurtosis", Journal of Econometrics, 106(1), 119-142.

Taylor, J.W. (1999), Evaluating Volatility and Interval Forecasts", Journal of Forecasting, 18, 111-128.

Van den Goorbergh R. and Vlaar P. (1999), Value at Risk Analysis of Stock Returns: Historical Simulation, Variance Techniques or Tail Index Estimation? manuscript, Tilburg University.

Table 3. Failure Times Number for 250 evaluation sample size

α	0.05	0.01	0.005	0.001	0.0001	0.05	0.01	0.005	0.001	0.0001
			AEX exante					AEX expost		
Normal	32	18	13	9	7	18	7	5	2	0
Student-t	35	13	9	2	0	10	1	1	0	0
RM	14	7	5	4	0	13	4	2	1	0
GARCH(1,1)	15	0	0	0	0	21	10	7	4	1
			ATX exante					ATX expost		
Normal	28	16	14	11	7	1	1	1	0	0
Student-t	30	13	9	4	0	1	1	0	0	0
RM	14	7	7	4	1	1	1	1	0	0
GARCH(1,1)	27	8	13	9	7	1	1	1	0	0
			BFX exante					BFX expost		
Normal	32	13	10	8	4	16	2	2	1	0
Student-t	33	10	7	1	0	23	1	0	0	0
RM	10	4	1	1	0	7	1	1	0	0
GARCH(1,1)	10	4	3	1	0	7	1	1	0	0
			DJIA exante					DDJIA expost		
Normal	19	12	10	6	3	3	1	1	0	0
Student-t	21	10	6	1	0	4	1	0	0	0
RM	13	8	4	2	1	2	1	1	0	0
GARCH(1,1)	16	8	8	4	2	3	1	1	0	0
			FCHI exante					FCHI expost		
Normal	25	12	10	7	4	9	3	1	1	0
Student-t	20	9	6	0	0	13	1	1	0	0
RM	16	9	7	3	1	5	1	0	0	0
GARCH(1,1)	13	7	6	2	0	6	1	1	0	0
			FTSE exante					FTSE expost		
Normal	32	20	18	11	7	7	2	1	1	0
Student-t	32	18	13	7	0	9	1	1	0	0
RM	18	8	5	0	0	7	1	1	1	0
GARCH(1,1)	18	8	5	0	0	7	1	1	1	0
			IBEX exante					IBEX expost		
Normal	31	15	13	9	7	4	2	0	0	0
Student-t	33	13	9	6	0	5	0	0	0	0
RM	20	10	8	7	5	3	2	0	0	0
GARCH(1,1)	29	15	13	8	7	4	2	0	0	0
			KFX exante					KFX expost		
Normal	33	22	21	12	7	4	1	0	0	0
Student-t	33	21	15	5	1	4	0	0	0	0
RM	21	12	7	3	1	11	2	2	0	0
GARCH(1,1)	22	12	8	5	1	11	2	2	0	0
			MIB30 exante					MIB30 expost		
Normal	31	18	14	7	4	7	1	0	0	0
Student-t	31	18	11	4	0	11	1	0	0	0
RM	24	11	8	4	0	5	0	0	0	0
GARCH(1,1)	19	7	4	2	0	5	0	0	0	0
			SSMI exante					SSMI expost		
Normal	32	22	17	13	9	11	3	2	1	0
Student-t	32	17	12	4	0	13	2	1	0	0
RM	27	17	14	9	5	5	2	1	0	0
GARCH(1,1)	27	17	14	10	7	6	2	1	1	0
			DAX exante					DAX expost		
Normal	30	15	14	10	6	31	19	16	11	7
Student-t	32	14	8	4	0	32	15	11	3	0
RM	14	7	6	4	0	6	0	0	0	0
GARCH(1,1)	19	10	8	6	4	6	1	0	0	0

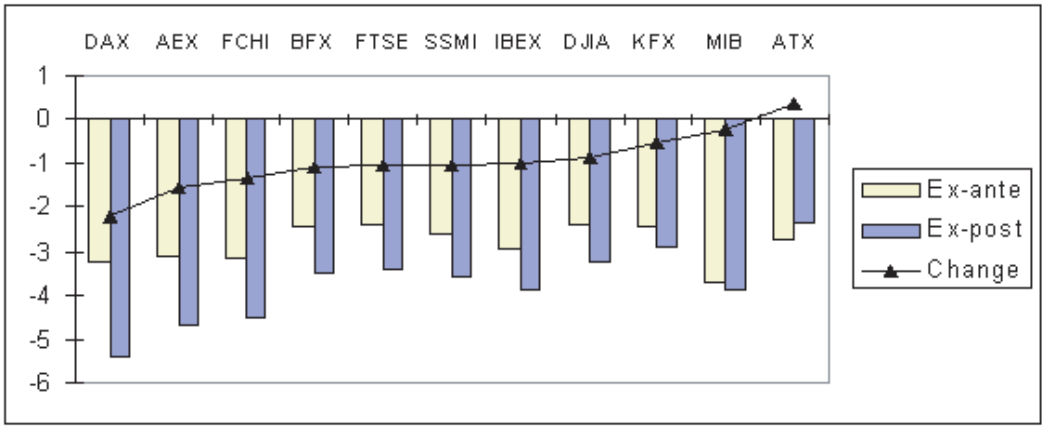


Fig. 1. Ex-ante and Ex-post Average VaR Estimates for 1% Confidence Level

Asset portfolio optimization. The case of portfolio consisted of both underlying assets and options

Alexey Nedosekin[1],

[1] IFEL Rus, Vice-President, Doctor of Economics, an@ifel.ru

Abstract. A fast algorithm for a portfolio efficient frontier reconstruction is represented. Portfolio consists of both real assets and derivatives. No more Monte-Carlo simulation, no probabilistic framework, fuzzy number analytics only.

1 Introduction

In papers **[1-3]** new methods of portfolio optimization were proposed, considering the case of portfolio consisted of a combination "asset + option of one type" as well as of portfolio consisted of only options. Now there is every reason to optimize portfolio of a common type.

It is clear from the content of the cited papers how the future price of the combination "asset + 2 options" looks like. On the one hand, due to the minimum price of the asset the minimum of the price is limited by the fixed price, i.e. by the strike of the put option and the extent of hedging $0 \le \gamma$. In the maximum-field the price of the combination is forced and defined as the relation of the maximum price of the asset, the strike of the call option and the extent of forcing $0 \le \delta$. The approximate future price is presented on Fig. 1 (for the case when the underlying asset price is a triangle fuzzy number)

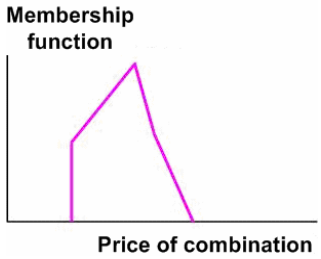


Fig. 1. The future combination price "asset + 2 options"

In this paper it is assumed that underlying assets and options on them can have any proportion. That means that options in this concept can play role of an independent

investment object, regardless whether the investment is made into underlying asset or not. Thus, we consider the case for the competitive investing into assets of different nature.

2 Quantitative analysis of investment into the combination "asset + 2 options"

2.1 Return on an asset is a triangle number

Consider the simplest case for the description of the return on underlying asset as a fuzzy number.

Let us examine one asset in time. Its today's price is Po. Our market analysis indicates that for the moment T the price for asset P can fluctuate in the following intervals:

- the worst price $P=P_{min;}$
- the average expected price – $P=P_{av}$;
- the maximum price – $P=P_{max.}$

If the return on such asset is designed as a triangle number, it can be written down as:

$$r = \left(\frac{P_{min} - P_0}{T \times P_0}, \frac{P_{av} - P_0}{T \times P_0}, \frac{P_{max} - P_0}{T \times P_0} \right), \quad (1)$$

The analytical type of the triangle fuzzy number of the future price, represented as an interval-segmented type, can be written as:

$P = [P_1(\alpha), P_2(\alpha)]$, $\alpha = [0,1]$, where

$$\begin{aligned} P_1(\alpha) &= P_{min} + \alpha(P_{av} - P_{min}) \\ P_2(\alpha) &= P_{max} - \alpha(P_{max} - P_{av}) \end{aligned} . \quad (2)$$

The "triangle" representing of the future price for the asset is shown on Fig. 2.

Now it is assumed that the investment is undertaken into two types of options on the examined undertaking asset – into put- and call-option. The depth of hedging for call options is δ, and for put-option is correspondingly γ. If a virtual lot of 100 shares, which make up the underlying asset, is considered (it is necessary to note that shares on the Stock Exchange are traded in round lots for 100 shares), it would mean that divisible lots of call-options in the amount of $100*\delta$ shares and divisible lots of put-options in the amount of $100*\gamma$ shares are purchased.

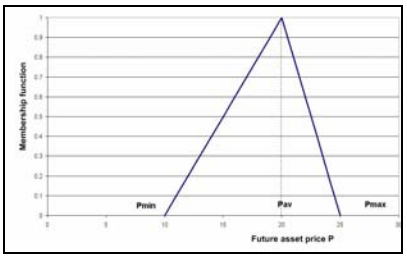


Fig. 2. Future asset price as a triangle number

Let the price of the option lots bought account for z_c and z_p, and the strike price (the option's exercise price) of the underlying asset, listed in the option, account for y_c and y_p correspondingly for call- and put-options. Naturally, $y_c < P_{max}$ and $y_p > P_{min}$ - otherwise, there is no point investing in the option. The condition $y_c > P_{av} > y_p$ is true for the rational market (i.e. nobody, who buys an option, will not act at one's loss, regardless the price expectations). Here and further the combination of "the underlying asset + δ% call-option + γ% put-option" will be considered. This combination is called "strangle"; if strike prices are the same then the combination is called "straddle".

This gives us a relationship for the future return on our combination. In accordance with **[3]**, the future price of the combination «δ% call-option + γ% put-option» is:

$Z_{PC} = [\gamma Z_{P1}(\alpha)+\delta Z_{C1}(\alpha),\gamma Z_{P2}(\alpha)+\delta Z_{C2}(\alpha)] = [Z_1(\alpha), Z_2(\alpha)]$, $\alpha = [0,1]$, where

$$Z_1(\alpha) = \begin{cases} \gamma(P_{av} - P_{min})(\alpha_p - \alpha), \alpha \le \alpha_p \\ 0, \alpha > \alpha_p \end{cases}$$

$$Z_2(\alpha) = \begin{cases} \delta(P_{max} - P_{av})(\alpha_c - \alpha), \alpha \le \alpha_c \\ 0, \alpha > \alpha_c \end{cases} . \quad (3)$$

$$\alpha_p = \frac{y_p - P_{min}}{P_{av} - P_{min}} > 0, \alpha_c = \frac{P_{max} - y_c}{P_{max} - P_{av}} > 0$$

Correspondingly, the resultant future price of the combination "the underlying asset + δ% call-option + γ% put-option" is:

$Z_{APC} = [P_1(\alpha)+\gamma Z_{P1}(\alpha)+\delta Z_{C1}(\alpha), P_2(\alpha)+\gamma Z_{P2}(\alpha)+\delta Z_{C2}(\alpha)] = [Z_1(\alpha), Z_2(\alpha)]$, $\alpha = [0,1]$, where

$$Z_1(\alpha)=\begin{cases} P_{min}+\alpha(P_{av}-P_{min})+\gamma(P_{av}-P_{min})(\alpha_p-\alpha), \alpha\le\alpha_p \\ P_{min}+\alpha(P_{av}-P_{min}), \alpha>\alpha_p \end{cases}$$

$$Z_2(\alpha)=\begin{cases} P_{max}-\alpha(P_{max}-P_{av})+\delta(P_{max}-P_{av})(\alpha_c-\alpha), \alpha\le\alpha_c \\ P_{max}-\alpha(P_{max}-P_{av}), \alpha>\alpha_c \end{cases}. \quad (4)$$

$$\alpha_p=\frac{y_p-P_{min}}{P_{av}-P_{min}}>0, \alpha_c=\frac{P_{max}-y_c}{P_{max}-P_{av}}>0$$

It follows from (4) that the calculated price and return are fuzzy generalized broken line numbers. For these numbers a comparison of segmented intervals is possible **[3]**. Thus, the scheme of types' succession represented in a series of our papers can be shown on Fig. 3.

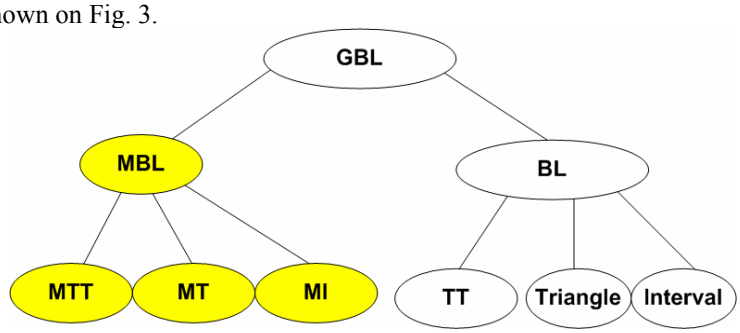


Fig. 3. The scheme of succession of fuzzy number types.

Abbreviations in the scheme: BL – classical numbers of broken line type, MBL – mirror broken line numbers (it was proved in [3] that the future price of the put-option is a number of this type), TT – truncated triangle numbers [1], T – triangle numbers, I – rectangle or interval numbers, MTT – mirror TT, MT – mirror T, MI – mirror intervals.

2.2 Return on asset as an interval

If the future price of the asset is an interval $[P_{min}, P_{max}]$, than the price for the call-option on such asset is also an ordinary interval with bounds $[0, P_{max} - y_c]$. At the same time the put-option price is a mirror (modal) interval with bounds $[y_p - P_{min},0]$. Thus, the price of the combination "the underlying asset + δ% call-option + γ% put-option" is $[P_{min} + \gamma*(y_p - P_{min}), P_{max} + \delta* (P_{max} - y_c)]$. It can be either a classical or a mirror interval depending on the relationship among the underlying share, the strike option prices and the depth of hedging.

Correspondingly, the return of such combination is I or MI type:

$$r = \left[\frac{P_{min} + \gamma(y_p - P_{min}) - (P_0 + \gamma z_p + \delta z_c)}{(P_0 + \gamma z_p + \delta z_c)T}, \frac{P_{max} + \delta(P_{max} - y_c) - (P_0 + \gamma z_p + \delta z_c)}{(P_0 + \gamma z_p + \delta z_c)T} \right]. \quad (5)$$

3 Problem setting for the generalized portfolio optimization

A generalized portfolio of 3 sub-portfolios is considered:

A. A sub-portfolio of N underlying assets.

B. A sub-portfolio of put-options. It consists of N lots of call-options; each of lots implements the hedging of the relevant underlying portfolio with the depth γ_i from 0 to 1. If $\gamma_i = 0$, then there are no options, and if $\delta_i = 1$, then the hedging is 100%, i.e. on each dollar of the underlying asset.

C. A sub-portfolio of call-options. It consists of N lots of call-options. Each of them implements the forcing of the relevant underlying asset with the depth δ_i from 0 to 1. If $\delta_i = 0$, there are no options, and if $\delta_i = 1$, the forcing is 100%, i.e. on each dollar of the underlying asset.

This, there are 3N constitutes in the generalized portfolio. The weight of each constitute is estimated on the basis of the prices of the underlying assets and option premiums. As a result, each constitute has a weight of x_i, and

$$\sum_{i=1}^{3N} x_i = 1, \; 0 \le x_i \le 1. \quad (6)$$

Here and further it is assumed that the depths of hedging (forcing) are infinitely divisible values. It is also assumed that the initial sum of investments remains constant, i.e. if weight of one constitute decreases, the increase of other weights of other constitutes in the portfolio follows.

The portfolio holder, while making an investment decision, set a criterion for a lower bound of the return on the portfolio at the moment T. It can be represented as a fuzzy number $[r_{Pmin}(\alpha), r_{Pmax}(\alpha)]$. In singular case it is a common numerical criterion r_P, for example, 15% per annum. However, in most cases the portfolio holder has hardly idea about these criteria. For example, tomorrow inflation rate is not precisely known; it can fluctuate in some bounds. If there is no convex, the most expected rate of inflation, around which other opportunities are grouped, then it is a common interval. Therefore, the portfolio holder wants the portfolio's rate of return to be ahead of the inflation rate (a rather natural desire). But the investor has rather fuzzy expectations about the portfolio as well as about the rate of inflation. If the expectations about the portfolio and inflation are crossed on some membership intervals, there is an obvious risk that the portfolio will not "fight" inflation, i.e. investments turn out to be inefficient (such situation is presented in Fig. 4). As it was shown, the return on the portfolio is a GBL number.

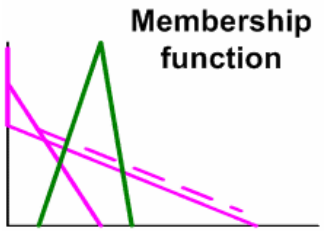


Fig.4. A comparison between the return on the portfolio and the rate of inflation

Observations. A part of a portfolio's fuzzy number is marked on Figure 4. It is formed by mirror (model) segmented intervals of membership.

The maximum of the expected return is required, because we consider forcing as well as hedging instruments, which have different impact. Any maximization leads to the crowding out effect of certain classes of assets from the portfolios, which form an efficient frontier of the portfolio set.

4 Solution of the optimization problem

It is carried out on the same principle as it is described in **[1-3]**. The risk degree of each portfolio is defined and the effective frontier of the portfolio set is reestablished according to the gradient method.

Consider a numerical example.

5 A numerical example of the portfolio with options, 3N=12 assets

The initial data on assets and options are borrowed from **[3]** and form Table 1. The total amount of investment in portfolio is constant and equals 1000 conventional units. The limit on the lower bound of return on the asset is r_P = 9.5%..11% per annum.

The preliminary analysis indicates that if there are no limits on the shares of the assets, all optimal assets, which lie on the efficient frontier, consist of only options (solution from **[3]**). To derive not an "extreme" solution, the following restrictions are to implied: a). there is at least one underlying asset in the portfolio; b). the depth of hedging of the asset lies in between from 0 to 1. In this case we have a classical strategy of portfolio investment into underlying assets, which is "diluted" by the application of derived securities which relate to these assets as the hedging.

N	Initial pricing parameters						
	P_0	P_{min}	P_{max}	y_c	z_c	y_p	z_p
1	100	109	114	110	0.2	110	1
2	100	108	118	110	0.6	110	1.2
3	100	106	125	110	1	110	1.6
4	100	105	130	110	3	110	2.1

Table 1. Initial data on assets and options.

The solution of the problem turns out to be singular. Right point of the efficient frontier is the asset number 4 forced at 100%. Left point is represented by the combination of the asset number 4, the put-option with 100% hedging and the call-option with 70% hedging. Thus, the efficient frontier is formed through the reallocation of the capital from the underlying asset and its hedging call-option in the direction of the put-option, i.e. from the forcing to the hedging as it must be. Other assets and options are not able to compete with the 4^{th} asset in our statement of the problem, and, therefore, they do not hit the efficient frontier in the portfolio.

Conclusion

The theory of portfolio optimization developed here and in [1-3] establishes good assumptions for the development of a specialized portfolio calculator. Having connected our assumptions represented in interval or triangle form with objective pricing data of the market, we can immediately define, which options are efficient in use, which is not and which form an efficient frontier considering the criterion of optimization, which will be actually proposed by us. Such calculator can be a very supportive tool for investors-professionals.

References

1. Nedosekin A.O. Optimization of a portfolio with put options // Banks and risks, №1, 2005.
2. Nedosekin A.O. Optimization of a portfolio with call options // Banks and risks, №1, 2005.
3. Nedosekin A.O. Optimization of a portfolio that consists of options only// Banks and risks, №2, 2005.

Fuzzy Pareto Set (FPS) as an Investment Planning Tool

Alexey Nedosekin[1],

[1] IFEL Rus, Vice-President, Doctor of Economics, an@ifel.ru

Abstract. Fuzzy Pareto Set (FPS) is an excellent tool for investment planning that allows excluding extreme investment scenarios if even they are formally optimal. Exclusion of extreme scenarios gives more relevant estimations for costs & prices in structure of investment budgets.

1 Introduction

The classic Pareto set is a set of non-dominated alternatives, either countable or non-countable. It is a mathematical formalism which is broadly applied in economic theory. Consider some examples to gain better understanding of the topic.

Countable Pareto Set. Consider a set of N investment projects, where each of them can be described by two parameters: X – risk of the project's failing, and Y – internal rate of return. Hence, there are N points in XY coordinates, from which we need to find the subset, where risk is minimal OR (not simultaneously) IRR is maximal (in singular case logical connector "OR" can be substituted by "AND"). The Pareto subset of the actual set, which satisfies these conditions, is represented on Fig. 1 (there are 4 elements in the subset).

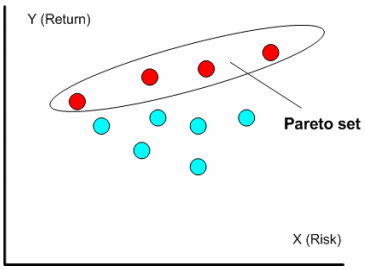


Fig. 1. Countable Pareto set

Thus, the actual (basic) set can be divided into two subsets: the Pareto subset and the subset of inefficient alternatives. Moreover, any element from the Pareto subset has characteristics that are not worse on one of the criteria than any other element from this subset (for the example on Fig. 1 – criteria are return and risk). If the set of inefficient alternatives is not empty, then there is such element in the Pareto set, which dominates some element from the "inefficient" set on two criteria at the same

time (in case of complete domination, risk is lower and return is higher at the same time).

Non-countable Pareto Set. One can think of the classical example of Markowitz's equation for stock portfolio. If assets introduced into a portfolio are infinitely divisible, then the set of investment alternatives is non-countable. Therefore, in "volatility - return" plane a so-called "portfolio cloud" can be shown. This means a complete set of possible portfolios, which can be combined from N infinitively divisible assets.

Consequently, a line without gaps, where return on a portfolio reaches its maximum considering its fixed volatility – a so-called efficient frontier of the portfolio set – is also the Pareto set. The example of such frontier for the portfolio set consisted of 3 assets is represented on Fig. 2. Dots on Fig. 2 indicate portfolios, which consist of only one asset.

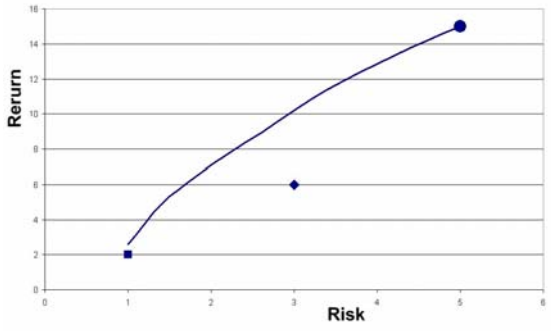


Fig. 2. Non-countable Pareto set – efficient frontier

In applied economic problems it is often the case that the Pareto subset, which is worked by the means of classical optimization, can not be proved for optimality if we want to consider other issues defined by the means of natural language. The mechanism of "fuzziness" of the actual Pareto set is turned on, its conversion begins.

2 Assumptions for the transformation of the classical Pareto set

Let us look at Figure 3, where the actual countable set of investment alternatives is graphed.

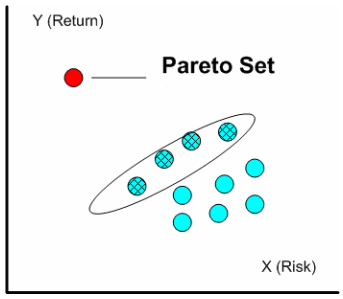


Fig. 3. The Pareto set consisted of only one element

A formal optimization shows that the Pareto subset for the basic set represented on Figure 3 consists of only one element. It is also seen that this element is situated far away from the others elements, which constitute a densely group. As a result, an analyst would regard this pointed Pareto set with distrust. He would see in this "star" element some suspicious anomaly, which is not typical of the actual set.

Variants of the further analysis are the following:

- An analyst would ignore his anxiety and consider an anomalous element as the Pareto set. From this point of view he remains being a mathematician, however, stops being an analyst.
- An analyst would exclude an anomalous element from the analysis consideration, assuming those elements to be the Pareto set, which are the cross-hatched and oval-shaped area. Besides, the analyst's measurement of trust to the formed Pareto set is 100%, even though this set from the oval-shaped figure is not the Pareto set according to the formal characteristics. Here one points out a quasi-Pareto set (as if it the Pareto set was considered not absolutely, but under condition of exclusion of anomalous element from the actual set). The analyst does his best as an analyst and no other analyst would condemn him. However, the mathematician would be unsatisfied and would require further research. And, by the way, would be absolutely right.
- An analyst would make a deal with himself as if he were a mathematician and put the anomalous element back into the actual set and the Pareto subset, however, not to a full extent. He would have some limited degree of trust to the fact that this anomalous element can really be optimal, that there is no computational mistake, no garbled facts or something like this. Therefore, the anomalous element and elements from the quasi-Pareto subset are included in the new set. Moreover, they are included not conditionally independent, but belong to this broadened subset with their certain membership degree.

Thus, there are assumptions for determining the quasi-Pareto set (in that the anomalous element is also included) as a fuzzy set. This will be defined later, and now consider one typical example.

Assume that alternatives have the following allocation (Fig. 4):

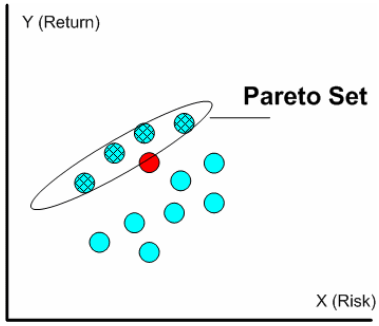


Fig. 4. The Pareto set with closely located elements

In fact, all the balls are closely-grouped (we will allow us such free "geometrical" explanation), and there is no anomaly. At the same time, there is one ball, which is situated rather closely to the oval-shaped figure, which covers the Pareto-set (exactly on the oval's frontier). On the one hand, this "closely situated" alternative is dominated by other elements from the Pareto set, and from this point of view it is formally inefficient. On the other hand, however, if we remember that the is an analyst in us, then this alternative is considered to be so close to the Pareto set, that it is a pity just to ignore it. And what if we had calculated something wrong for this alternative or took bad data (without considering white noise) for the analysis of the alternative?

As a result, we turn back again to the idea of the quasi-Pareto set, where the elements of the actual Pareto set (with membership 1) will be included as well as other elements of the basic set which have some closeness to the Pareto set. Moreover, the measurement of this closeness is, actually, a degree of elements' membership to the set.

3 The formal definition of the Fuzzy Pareto Set (FPS)

Thus, the Fuzzy Pareto Set (the quasi-Pareto set, FPS) – is a fuzzy subset of the actual set (in traditional definition – of the carrier), which is formed on the carrier by introducing the membership functions of the elements of the actual subset in accordance with the fuzzy set.

Therefore, the membership function (discrete and continuous in its domain) – is a method for an analyst, which allows him to adjust initial formal solution of the optimization problem by expanding the classical Pareto-set up to the actual set. The analyst gives the grade of certainty to each element of the carrier that all these elements are considered as conditionally optimal or conditionally non-optimal (as you like – to see the glass as half full or half empty).

There, where we are not confident about something, or confident that it is vise versa, the membership function is close to or equals zero. On the contrary, the closer to the formal Pareto-set we are, the higher the analyst's confidence is. In this case it useful to imply formalism of the Hamming distance between two elements of the set in the phase space of these criteria (calibration of the set is implemented by linear

transformation of the criteria scale, considering its changes for specific scales with membership area as a unit interval).

Thus, we turn back again to our discrete examples (Fig. 5). On these figures there is a membership degree near the each element of the FPS, indicating its membership to the set. If there is no membership, we do not put zero values down.

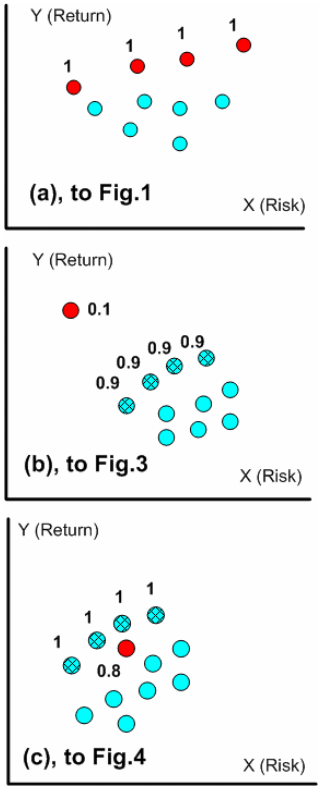


Fig. 5. Different cases of FPS

The case (a) on Fig. 5 is when the classic Pareto set coincides with the fuzzy set. The case (b) is when the analyst actually excludes the anomaly from the fuzzy Pareto set (giving it a zero degree of membership) and adds elements to the FPS, which are considered to be the Pareto set to the actual set excluding the anomalous element. The case (c) is when the FPS consists of the classic set (full membership) and a border with a rather high membership degree.

A typical example of a fuzzy Pareto set in case of non-countable carrier is an efficient fuzzy Markowitz' frontier (Fig. 6). Consider the classic problem of portfolio optimization to be fuzzy concerning its initial inputs: return and risks of assets, the correlation matrix – LR fuzzy numbers (Dubua-Prad case). Then, the efficient frontier of the portfolio set is also fuzzy and turns to be a curvilinear line, where each level of risk corresponds with a fuzzy number of maximal return. Therefore, the FPS for Markowitz' problem is a segment of the membership function.

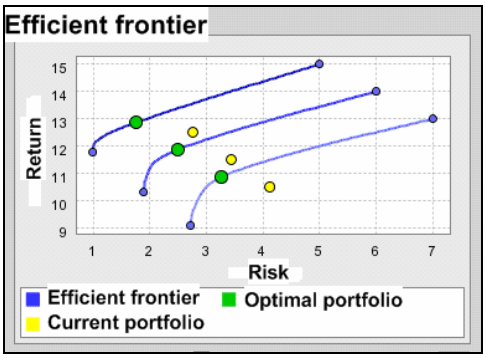


Fig. 6. FPS as a segment of a fuzzy function

4 Example of definition of FPS "taken from life"

The best is, if we give here an example how the FPS has helped us in the investment projecting.

Consider a building project of a business center (BC) in Saint-Petersburg. We know land parameters, its location, purpose of BC, its class and other input parameters of the project. There is also an architect project available. From this point of view the problem of determining of initial investment costs and current operating costs can be solved rather precisely.

It's quite another matter when we need to estimate return on the project, which includes rental payments of leases by tenants in the BC. Firstly, the volume of the demand for the BC is unclear, in other words, the number of leases is unknown. Secondly, and this is mostly important, the rental rate of the effective business size, dollars per square meter per annum. To define it, it is necessary that marketing research is carried out, gathering data on all BC of the same category in the region. Moreover, it is more important to get concept of not only rates, but also utility, which is gained by a tenants form the BC. Thus, in the analysis we must include qualitative estimate of the BC, i.e. a complex estimate of the consuming characteristics of the BC (further – a quality of the BC, Fig. 7).

The simplest model of the BC's quality is based on the qualitative and quantitative estimate of the parameters and further aggregation of the estimates, which were found by the methods of fuzzy mathematics (by analogy with how it was done on **[1]**).

The factors to be estimated are:

- F1 - region, location
- F2 – state and quality of repair and maintenance
- F3 – sum of required advance (1 – 6 months)
- F4 – type, area of parking and price of parking for tenants
- F5 – infrastructure of the BC (café, restaurants, shops and so on)

Expert estimation of chosen criteria has shown that they have the following preference relations (preference criteria on estimation) among each other:

$$F1 \succ F2 \approx F5 \succ F3 \approx F4, \quad (1)$$

Where «$\succ$» - a weak preference relation, «≈» - an indifference relation. Application of Fishburn's rule to (1) gives us weights of criteria in matrix form.

$$F1 = 0.333,\ F2 = F5 = 0.222,\ F3 = F4 = 0.111. \quad (2)$$

Business Center (BC)					
#	Weight P	Factor	Factor's quality		
			low	*blend*	*high*
1	0.333	Location		1	
2	0.222	Repairing quality			1
3	0.111	Advance			1
4	0.111	Parking quality			1
5	0.222	Infrastructure		1	
Weight Y			0.3	0.5	0.7
Quality			**0.589**		

Fig. 7. Example of quality matrix

Having applied the matrix scheme of quality evaluation for all business centers of B and B+ class, situated in Saint-Petersburg (for us, they are objects analogies), we face with two-dimensional evaluation of each business center by "quality – rental rate" criteria (Fig. 8). Data are valid for September, 2005.

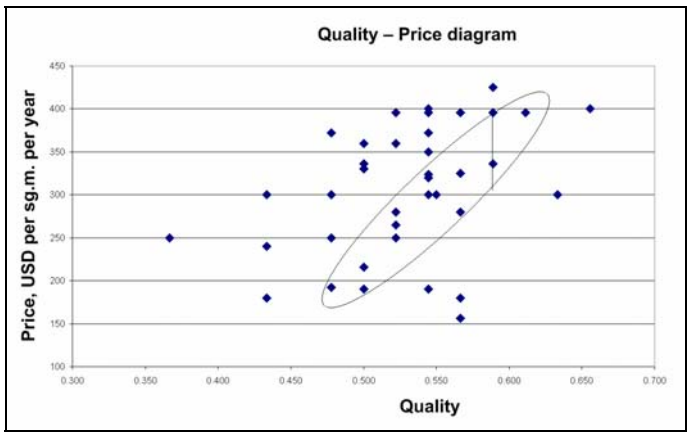


Fig. 8. Characteristics of BC's of B and B+ class in SPb

Taking into account the variation of points on Fig. 8, we come to conclusion that the formal Pareto set here (3 points) is not exactly what we need. Firstly, there are few points, i.e. they are not densely located. Centers represented by these points are either of too high quality or too cheap, which makes analyst regard these alternatives

as unsuitable for the analysis. On the contrary, our BC (the basic object) has such level of quality that there is a tightly plotted group of points near the vertical line passing through abscissa of 0.589.

Therefore, it necessary to search the group of relevant objects-analogous for defining an equilibrium rental rate directly in the dense set of points, and not on the "anomalous" formal Pareto frontier.

Finally, to define computable rental rate for the basic object, a so-called method of "net" is applied. This method includes a set of techniques for identifying (they are usually all of a heuristic type). The "net" itself is represented by ellipse, whose semi-major axe has angle value of 45 degrees to abscissa. Such orientation of the net means that we are going to net alternatives located in the right down corner of the coordinate space. Length of the axes is defined the way that all (or a majority) of the alternatives, which possess a "good" membership to FPS, can be covered by the net. The whole question rests here on the expert estimations and value judgments about what anomalies exist, which of them to neglect and which of them are closely located to the optimal set (in the sense of Figure 5c) and if there is point in applying these alternatives.

In the end, as the "net" is located on the surface, cut it vertically by the abscissa responsible for the quality of the basic object. The coordinates for the points of intersection of the "net" and the "vertical" gives us an interval, where the rental rate for the basic object of investment is to be found (in our example it is 300-400 USD per sq.m. of useful size of the BC per annum)

Considering this estimation in the budget, we apply a prudence method. By that, it is assumed that we are able to offer the most favorable rental rates for the consumers, staying at the same quality of rental areas offered by other business centers in the city.

Reference

1. Nedosekin A.O., Kuznetsov D.D. The development of fuzzy qualitative model for comparative cost analysis of the objects // Banks and risks, №2, 2005.

An Approach to Composing Optimum Investment Program in Conditions of Uncertainty

Larissa Rozanova, Zoya Maximenco

Department of Computer Science and Robotics,
State Aviation Technical University of Ufa,
12 K. Marx St., Ufa, Russia, 450000
rozanova_lara@mail.ru, zoya@mail.rb.ru

Abstract. The paper regards the problem of optimum investment program composition in conditions of imprecise initial information about the investment projects. The problem of composing an optimum investment program is formulated as a problem of fuzzy mathematical programming. The approach to solution of the stated problem that we propose is based on the use of heuristic algorithms and allows composing an optimum investment program which is characterized by the given efficiency criterion and also by probability of not exceeding the available financial budget. The formalization of imprecise initial data and mathematical fundamentals of the suggested approach are considered.

1 Introduction

Making decisions about realization of real investment projects is inevitably attended by uncertainty. In these conditions it's important to prove expediency and efficiency of supposed investments. One of the problems of evaluating investment projects' efficiency is composing an optimum investment program on the basis of given investment portfolio with given limitations on projects' selection. It's known that often the initial parameters of the investment projects may be given imprecisely, only as a certain intervals, thus one should not precisely limit the amount of monetary resources for financing the investment program.

In this paper the problem of optimum investment program composition in conditions of imprecise initial information about the investment projects is considered. The approach to solution of the stated problem with the use of the formalism of fuzzy mathematics is suggested.

2 Formalization of Imprecise Initial Data

As the precise values of parameters are unknown according to concepts founded in [2, 3] we suggest formalizing the initial data using the following fuzzy quantities:

– the parameters of projects and investment program are described by fuzzy numbers $\underline{A} = [a_{\min}, \overline{a}, a_{\max}]$ with triangular membership function:

$$\mu_A(x) = \begin{cases} 0, & x \le a_{\min} \\ \dfrac{x - a_{\min}}{\overline{a} - a_{\min}}, & a_{\min} \le x \le \overline{a} \\ \dfrac{a_{\max} - x}{a_{\max} - \overline{a}}, & \overline{a} \le x \le a_{\max} \\ 0, & x \ge a_{\max} \end{cases}. \quad (1)$$

The parameters with indexes *min*, *max* show the boundaries of the interval of parameter values, and the parameter $\overline{a}$ shows the most expected parameter value.

– the available budget of monetary resources for financing the investment program is described by fuzzy set $\underline{A} = [\overline{a}, a_{\max}]$ with *z*-similar membership function:

$$\mu_A(x) = \begin{cases} 1, & x \le \overline{a} \\ \dfrac{a_{\max} - x}{a_{\max} - \overline{a}}, & \overline{a} \le x \le a_{\max} \\ 0, & x \ge a_{\max} \end{cases}. \quad (2)$$

This function describes a statement: "the budget for financing the investment program is equal to $\overline{a}$, but if it's necessary it may be increased up to $a_{\max}$".

If any of parameters $\underline{A}$ is known quite precisely or is uniquely assigned, the fuzzy number $\underline{A}$ degenerates to a real number *A* with fulfillment of the condition $a_{\min} = \overline{a} = a_{\max}$.

3 Mathematical Fundamentals of Composing an Optimum Investment Program

Maximum value of net present value, maximum profit on each unit of investment capital, minimum payback period, maximum value of investments' profitability index etc. may serve as a criterion of selection projects into investment program [5, 6]. In this paper we suggest to use the maximum of net present value (*NPV*) and profitability index (*PI*) of the investment program.

Composition of an optimum investment program is carried out on the basis of the following premises [1]:

1. Considered projects are equivalent for person, who makes decisions.
2. Financial resources are limited.
3. Investment program is composed on the beginning of planned period, initial expenses might not exceed the capital budget.

4. All projects have passed the evaluation of efficiency, which proved their right for implementation.
5. All projects are indivisible.

Let's consider the following designations:

1. $A = \{a_j\}, j = \overline{1,n}$ – a set of projects included into investment portfolio.
2. $\underline{IC_j} = (IC_{j\min}, \overline{IC}_j, IC_{j\max})$ – the volume of initial investments required for implementation of project *j*.
3. $\underline{I} = (\overline{I}, I_{\max})$ – available to investor budget of monetary resources for financing the investment program.
4. $\underline{NPV_j} = (NPV_{j\min}, \overline{NPV}_j, NPV_{j\max})$ – the value of *NPV* of project *j*.
5. $x_j \in \{0,1\}, \quad j = \overline{1,n}$ – Boolean variable, which value defines if the project will be included into the optimum investment program:
   $x_j = 1$, if project *j* will be included into the optimum investment program,
   $x_j = 0$, otherwise.
6. *P* – probability that the volume of monetary resources required for financing the investment program won't exceed the prescribed budget.
7. P_{req} – given by an expert number from an interval [0, 1], which presents the required value of *P*.

Then the optimum investment program will consist of such set of parameters x_j, which will be the solution of the following problem of fuzzy mathematical programming:

It's required to maximize *NPV* of the investment program

$$\left(\sum_{j=1}^{n} \right) \underline{NPV_j} x_j \to \max \qquad (3)$$

with following restrictions:

1. on volume of initial investment resources:

$$\left(\sum_{j=1}^{n} \right) \underline{IC_j} x_j \subseteq \underline{I} \ , \qquad (4)$$

2. on sufficiency of monetary resources:

$$P \geq P_{req} \ , \qquad (5)$$

3.

$$x_j \in \{0,1\}, \quad j = \overline{1,n} \ . \qquad (6)$$

Thus the problem of composing an optimum investment program down to the problem of integer programming with Boolean variables in conditions of fuzzy input data.

The following algorithms are used for solving the stated problem:

− Algorithm "first suitable with sequencing by *NPV*" – FSS_{NPV} (see fig. 1):

1. Sort projects by descending of *NPV*.
2. Select a project.
3. Check the sufficiency of investment resources for implementation of the selected project.
4. If there are sufficient resources($P \geq P_{req}$) then include the project into the program. If all projects have been reviewed then – end, otherwise go to p. 2.

− Algorithm "first suitable with sequencing by *PI*" – FSS_{PI} :

1. Calculate the profitability of each project using the following formula:

$$\underline{PI_j} = \frac{NPV_j}{\underline{IC_j}} \cdot \qquad (7)$$

2. Sort projects by descending of *PI*.
3. Select a project.
4. Check the sufficiency of investment resources for implementation of the selected project.
5. If there are sufficient resources ($P \geq P_{req}$) then include the project into the program. If all projects have been reviewed then – end, otherwise go to p. 3.

The probability of sufficiency of monetary resources is calculated by means of comparison of membership functions of fuzzy numbers $\underline{I} = (\overline{I}, I_{\max})$ and $\underline{ICS} = \left(\sum_{j=1}^{n} \right) \underline{IC_j}, \text{для } \forall j \neq 0$.

There may be several situations, for example see fig. 2.

In the situation shown in fig.2 the probability of sufficiency of resources for financing the investment program *P* is calculated as a ratio of area of zone restricted by lines $(ICS_{\min}, \overline{ICS})$, $\underline{I} = (\overline{I}, I_{\max})$ to area of triangle $(ICS_{\min}, \overline{ICS}, ICS_{\max})$.

As the suggested algorithms are heuristic we can't guarantee their complete efficiency. It's suggested to solve the stated problem with the use of both algorithms and take the maximum result as an optimal decision [1].

4 Analysis of Efficiency of Suggested Approach

A number of experiments were carried out to check the efficiency of suggested algorithms. The results of calculations with the use of suggested algorithms were compared with the results of calculations with the use of the exhaustive search algorithm.

The results of experiments are presented in table 1.

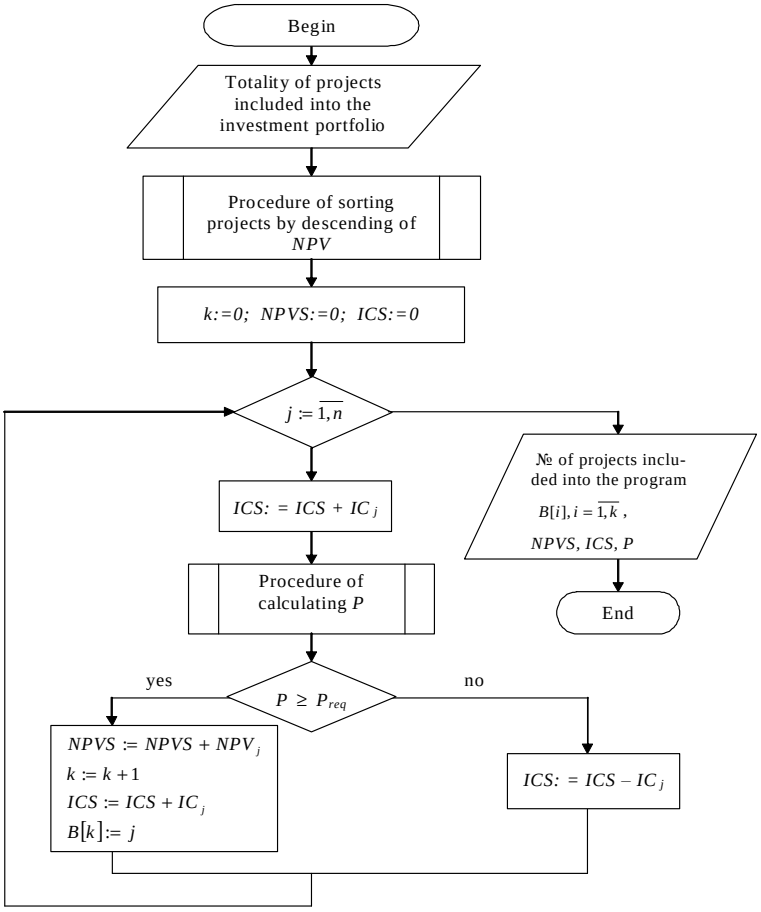


Fig. 1. Block-scheme of the algorithm “first suitable with sequencing by *NPV*”

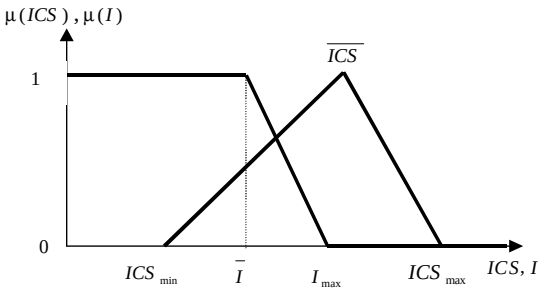


Fig. 2. Membership functions of *I* and *ICS*

Table 1. The results of checking the efficiency of algorithms FSS_{NPV} and FSS_{PI}

Number of projects	Number of experiments	Coincidence of results with the results received with the use of the exhaustive search algorithm			Relative error when algorithms FSS_{NPV} and FSS_{PI} were used jointly, $\frac{\sum_{i=1}^{N} \frac{NPV_{EXSEAR} - NPV_{FSS}}{NPV_{EXSEAR}} * 100}{\text{number of not coincided experiments}}$
		FSS_{NPV}	FSS_{PI}	FSS_{NPV} and FSS_{PI}	
10	1000	45,6%	54,7%	84,7%	1,87% – 3,12%
15	1000	49,8%	52,6%	82,1%	3,11% – 3,56%
18	600	53%	60,83%	85,33%	1,012% – 1,05%
20	500	48,2%	53,6%	73,8%	1,125% – 1,99%
22	100	42%	50%	74%	2,1% – 2,34%

The average difference of NPV calculated with the use of suggested algorithms from NPV calculated with the use of the exhaustive search algorithm was taken as a relative error. So when the algorithms FSS_{NPV} and FSS_{PI} are used jointly the average error makes up 1,84% – 2,41% of *NPV*.

Thus the joint implementation of algorithms FSS_{NPV} and FSS_{PI} just insignificantly reduces the efficiency (precision) of decision. Besides the suggested algorithms significantly reduces the computational difficulty of solution [4].

5 Conclusion

In this paper the approach to composing optimum investment program in conditions of imprecise initial information about the investment projects is suggested. The formalization of imprecise initial data and mathematical fundamentals of the suggested approach are considered.

The problem of composing an optimum investment program in conditions of limited budget is formulated as a problem of fuzzy mathematical programming. Fuzzy value of *NPV* of the investment program is considered as the efficiency criterion. The budget limitation is set as *z*-similar function and represents the investor's preferences.

The suggested solution of the stated problem is based on the use of heuristic algorithms FSS_{NPV} and FSS_{PI} and allows to compose an optimum investment program which is characterized not only by given efficiency criterion but also by probability of not exceeding the available financial budget.

The analysis of efficiency of algorithms FSS_{NPV} and FSS_{PI} showed that the results of joint implementation of these algorithms coincide with optimum in 73%-85% of experiments and in the rest of experiments the average error doesn't exceed 4% of *NPV*.

References

1. Maximenko Z.V.: Composition of Optimum Investment Program in the ISIDMS. In: Proceedings of CSIT'2004 6th International Workshop on Computer Science and Information Technologies, Vol. 2. Budapest, Hungary (2004) 261−265
2. Nedosekin A.O., Voronov K.I.: Investment Risk Analysis with Use of Fuzzy Sets. www.cfin.ru
3. Nedosekin A.O.: Fuzzy Analysis of Portfolio Investments Risk. Sezam, Saint Petersburg (2002)
4. Rozanova L.F., Maximenko Z.V., Osipova V.P., Oleynik J.A.: Information System for Investment Decision Making Support. In: Proceedings of CSIT'2002 4th International Workshop on Computer Science and Information Technologies. Patras, Greece (2002) paper on CD
5. Shelobaev S.I.: Mathematical Methods and Models in Economics, Finances, Commerce. UNITI, Moscow (2000)
6. Vilenskii P.L., Livshits V.K., Orlova E.R., Smoljak S.A.: Evaluation of Investment Projects' Efficiency: Theory and practice. Delo, Moscow (2002)

The influence of fuzziness of payments on the stability level of imputation sets in strategic games

Henryk Piech, Aleksandra Ptak, Marcin Machura

Technical University of Czestochowa, Institute of Mathematics and Informatics
Dabrowskiego Street, 73, 42-200 Czestochowa, Poland,
tel.: 00480343250331
hpiech@adm.pcz.czest.pl

Abstract. In an *n*-person game one wants to determine which coalitions can form and how the winnings of a coalition will be divided among its members. The division plan of winnings by a potential coalition essentially influences the formation of the coalition by individual players [4]. The payment vector should meet the following conditions:
- payment for all players within confines of cooperation should be greater than payment without this cooperation: $xi>v(i)$, i – player's number, $v(i)$ – characteristic function for *i*-th player (individual rationality),
- sum of payments for players by formed cooperations should be equal to sum of payments in non-cooperation play

$$\sum_{i=1}^{n} xi = v(N) \text{ (collective rationality)}$$

The *N*-dimensional vector of payments meeting the conditions of individual and collective rationality is called imputation.
Note that in real conditions the following are subjected to fuzziness:
- amount of payments,
- percentage division of payments within confines of coalition,
- value of pool for division.

The idea of stabilization consists in this, that external stability "attracts" imputation to the stability zone and internal stability protects against "leaving" this zone.
Controlled fuzziness can influence an increase in the possibilities of the stabilization of imputation sets and uncontrolled fuzziness can lead to a loss of stability of the game solution. This is the main thesis which will be verified here and the conditions of its confirmation will be analyzed.

1 Selection of tools for research of stability of imputation sets

As basic models describing the distribution of payments in coalitions one can use a normalized triangle, with the sum of distances to all sides equal to *1* (rectangle, polygon), in which points describing the selected imputation will be placed (Figure 1).

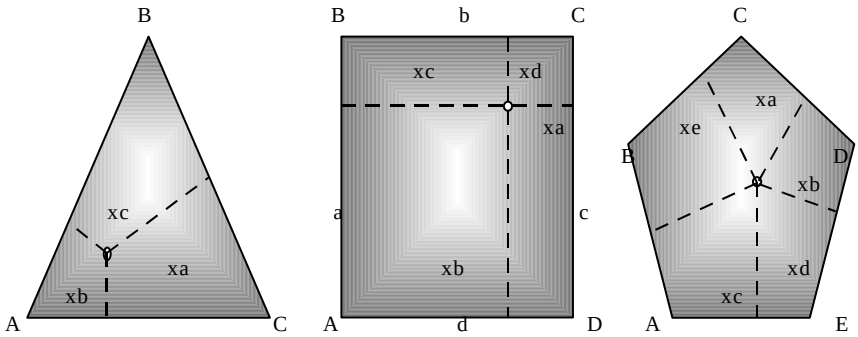


Fig. 1. Research models of cooperation in 3, 4 and 5-person games

In the internal stability of an imputation set one component is not dominated by another imputation belonging to this set. The domination of imputation *X* or over imputation *X* (Figure 2 and 3) consists in the existence of a coalition for which payments for all players in this coalition are greater than the payments in imputation *Y* which is dominated in the coalition and the sum of payments for the players of the coalition is not greater than the pool of payments for this coalition *S*:

$$xi > yi \text{ for players of coalition } S,$$

$$\sum_{i \in S} xi \le v(S).$$

External stability is the domination of imputations belonging to a given set over all the imputations which do not belong to the set.

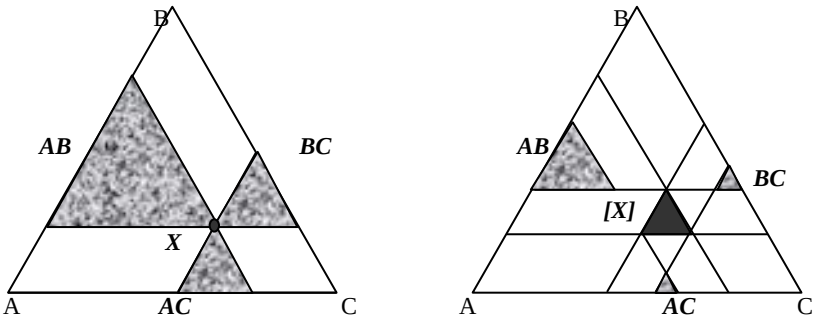


Fig. 2. Domination of *AB, BC, AC* coalitions over imputation *X* in situation of determined weights and over fuzzy set *[X]* in situation of fuzzy weights.

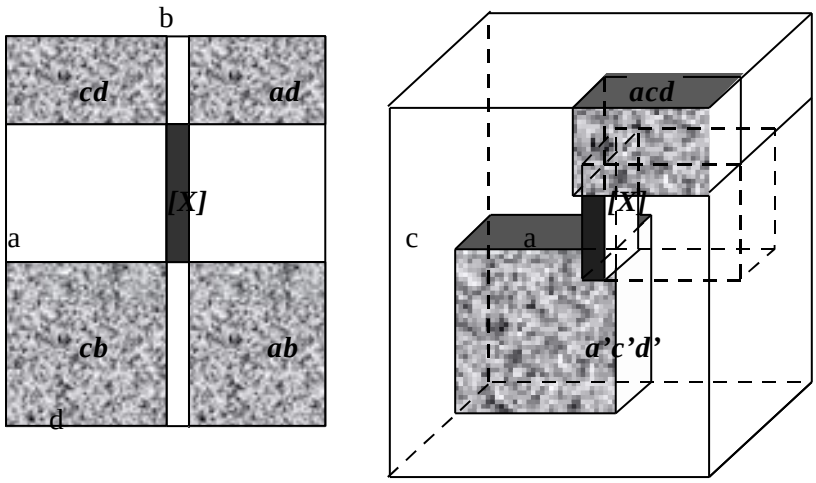


Fig. 3. . Examples of domination presentation: coalitions *cd*, *ad*, *cb*, *ab* dominating fuzzy imputation ***[X]*** (on the left), coalition *acd* dominating fuzzy imputations ***[X]*** (on the right), coalition *a'c'd'* is dominated by ***[X]*** (on the right)

The use of rectangle does not allow the presentation all the coalitions as dominant, because the rectangle's sides making the basis of the payment calculation lie on opposite sides.

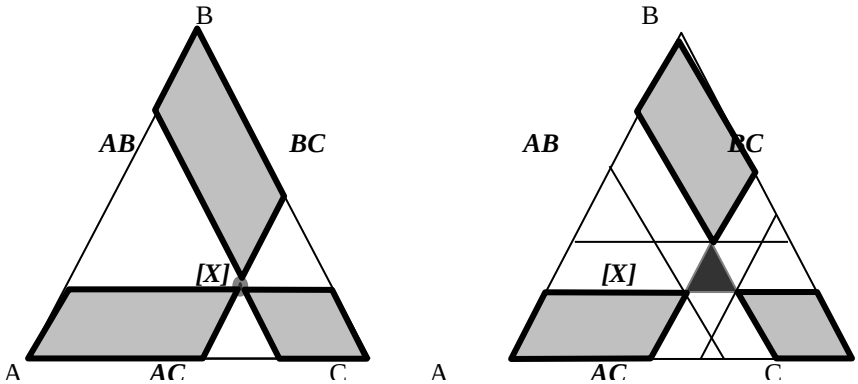


Fig. 4. Coalitions ***AB, BC, AC*** dominated by imputation ***X*** in situation of determined weights and by imputation fuzzy set ***[X]*** in situation of fuzzy weights

If one determines a set in which none of the imputations dominates any of the remaining imputations, then the set is internally stable (Figure 5). If in each imputation not belonging to the given set players of the coalition obtain less than within the confines of the imputation from the given set, then the set is externally stable (Figure 5).

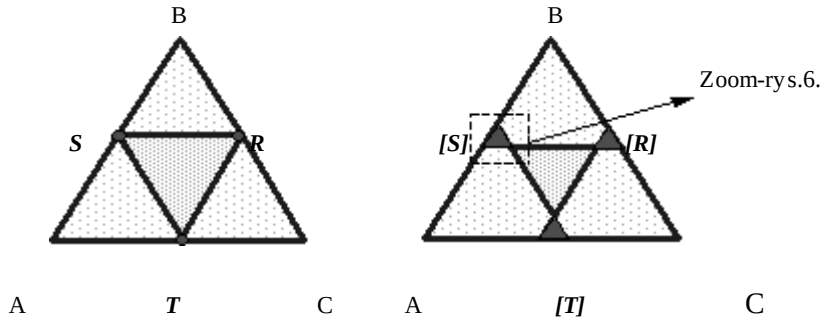


Fig. 5. Set of imputations ***{S,R,T}*** (on the left) is not internally dominated, i.e. internally stable; sets ***[S],[R],[T]*** are not internally stable (on the right) (Figure 6); each imputation from external set is externally dominated by internal imputations from ***{S,R,T}*** set (external stability); fuzzy sets of ***[S],[R],[T]*** imputation are also externally stable

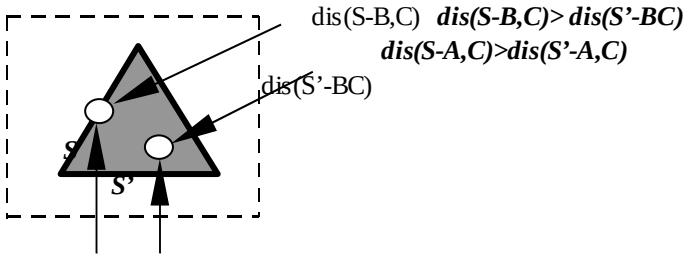


Fig. 6. Interpretation of internal domination inside of fuzzy field ***[S]*** (direct domination)

Figure 6 shows the graphical way of demonstrating internal domination in the case if one allows more than one imputation point from the fuzzy field. In this case the other point ***S'*** is located inside the same fuzzy subfield ***[S].*** One deals here with direct internal domination. It is possible to obtain the same result by allowing the existence of an imputation inside both remaining subsets ***[R]*** and ***[T].*** One can define this situation as indirect internal domination (Figure 7).

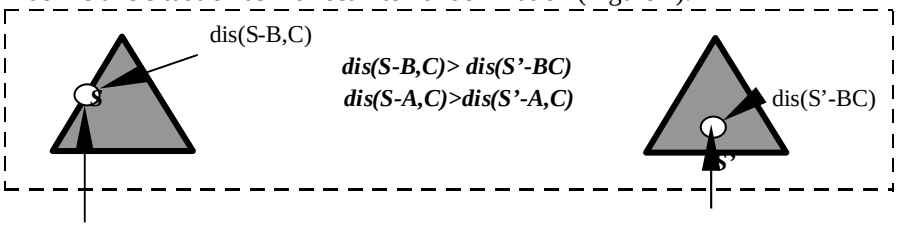


Fig. 7. Interpretation of internal domination outside of internal field (indirect domination)

An analysis of the external domination shows that the fuzzy subfield ***[S]*** dominates over the ***f*** and ***g*** fields**.** These fields are external fields in relation to ***[S]*** (the condition of external stability) (Figure 8).

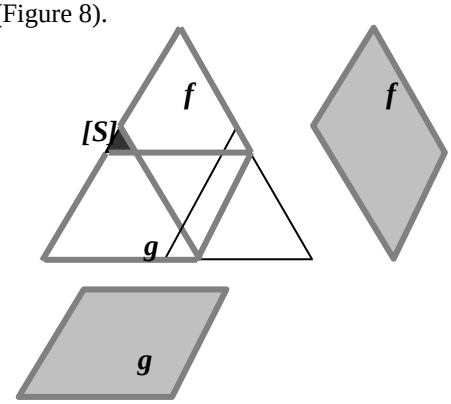


Fig. 8. Example of domination of fuzzy set ***[S]*** over ***f*** and ***g*** sets

The assurance of external domination and the lack of internal domination is a situation which induces the estimation of the stability degree.

2 Fuzzy instability measure

It is obvious that a lack of stability results from fuzziness. The analysis of instability can be carried out by means of fuzziness polygons (here - ***[S],[R],[T]*** triangles). One can intuitively suggest that the greater fuzziness of payments, the greater instability. However, it is excessively simplified, because the position of the imputation polygon is an additional factor influencing destabilization (Figure 9).

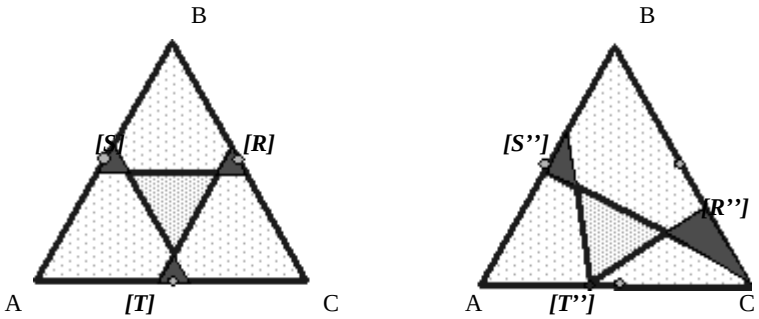


Fig. 9. Examples of instability resulting from imputation fuzziness. Points are stable places

The estimation of the measurements of instability as the sum of distances from stable places can be carried out in the way illustrated in Figure 10.

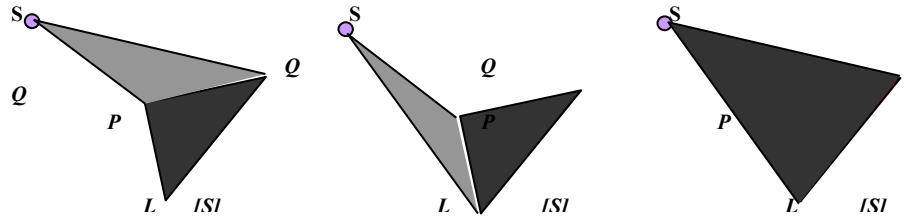


Fig. 10. Measurement of instability as sum of distances from stability place ***S*** realized as sum of surfaces of triangles ***SLP, SPQ*** and ***SLQ*** (instability triangles)

In general, instability can be estimated as follows:

$$\sum_{i=1}^{n} surf\Delta\{S \longleftrightarrow side(i)\} \text{ if side (i) and point S do not lie on the same straight line}$$

$$\sum_{i=1}^{m} dist\_vert\{S \longleftrightarrow vert(i)\} \text{ if points S and interval limits of imputation fuzziness lie on the same straight line} \quad (1)$$

where:
surfΔ {S↔ side(i)} – surface of triangle created from vertex *S* and side of triangle with number *i*,
dist_vert {S↔ vert(i)} – distance between stability point *S* and vertex with code *i*,
n – number of players,
m – number of instable imputations.

The presented kind of instability has an external character, because it does not take the surface of imputation polygon into consideration. This surface increases the danger of domination and internal instability as well. There are conceivable situations when an increase in the imputation polygon surface causes a simultaneous increase in internal instability and a decrease in external instability (Figure 11).

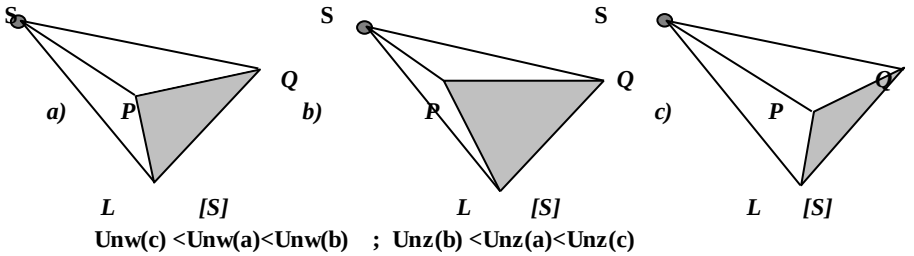


Fig. 11. Example of change influence of external instability on reverse reaction of internal instability as confirmation of thesis on possibility of expression of total instability in the form of contour surface created from combination of stable point and imputation polygon.

Thus, one can write down:

$$Uii = \Delta PLQ$$

$$Uei = \Delta SLP + \Delta SPQ + \Delta SLQ$$

$$Ui = Uii + Uei = \Delta PLQ + \Delta SLP + \Delta SPQ + \Delta SLQ = 2*\Delta SLQ = 2*\text{contour} \quad (\mathbf{2})$$

If the stability point lies on a side or inside the imputation polygon, then the instability contour obviously "reduces" the surface of the polygon (Figure 12).

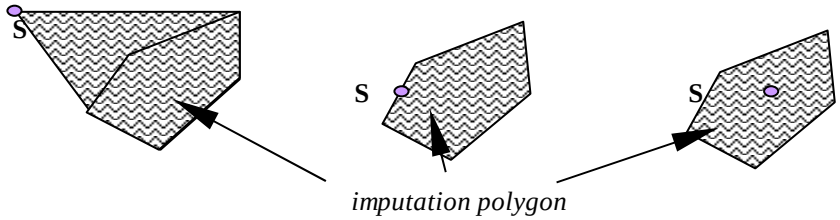


Fig. 12. Rules of construction of instability contour

3 Conclusions

1. The imputation fuzziness of payments influences the solution instability increase of cooperation rules within the confines of the players' coalition. It causes a loss of internal stability and the possibility of a decrease in external stability.
2. The tool which one can use to research the stability level is the instability contour constructed on the fuzzy imputation polygon and the stability point.
3. The control of the imputation fuzziness level takes into consideration its direct, and frequently simultaneous, influence on the increase or decrease of the external and internal domination of the set of payments. A decrease in the internal domination level can cause an increase in the external domination scale (Figure 11).

References

[1] Aumann R., Survey of repeated games. In Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Institute Mannheim, Vein, 1981.
[2] Bamos A., On Pseudo-Games. Annals of Mathematical Statistics, No. 39, 1968.
[3] Binmore K., Kirman P., Tani P., Frontiers and Game Theory, MIT Press, 1993
[4] Conway J.H., On number and Games, Academic Press, 1976.

[5] Czogala E., Perdycz W., Elements and methods in fuzzy sets theory, PWN, Warsaw, 1985 (in Polish).
[6] Ellison G., Learning with One Rational Playe, MIT Press, 1994.
[7] Fudenberg D., Kreps D., Lectures on Learning and Equilibrium in Strategic Form Games, Core Lecture Series, 1990.
[8] Fudenberg D., Tirole J., Game Theory, MIT Press, 1991.
[9] Harsanyj J., Selten R., A General Theory of Equilibrium Selection in Games, MIT Press, 1988.
[10] Harsanyj J., Games with Randomly Disturbed Payoffs, International Journal of Game Theory, No. 2, 1973.
[11] Hofbauer J., Stability for Best Response Dynamic, University of Viena, 1995.
[12] Isaacs R., Differential games, Wiley, 1965.
[13] Lachwa A., Fuzzy world of files, numbers, relations, facts, rules and decisions, Akademicka Oficyna Wydawnicza Exit, Warsaw 2001 (in Polish).
[14] Loeve M., Probability Theory, Berlin Springer Verlag, 1978.
[15] Luce D., Raiffa H., Games and Decisions, PWN, Warsaw, 1996 (in Polish).
[16] Nachbar J., Evolutionary Selection Dynamic in Games, International Journal of Games Theory, No. 19, 1990.
[17] Nachbar J., Prediction, Optimization and Learning in Repeated Games, Econometrica, 1995.
[18] Nash J.F., Non cooperative games, Ann. Math., 2, pp. 295-296, 1951.
[19] Nash J., Equilibrium points in N-persons games, National Academy of Sciences, 36, pp. 48-49, 1950.
[20] Owen G., Theory of Games, PWN, Warsaw, 1982 (in Polish).
[21] Ozyildirim S., A discrete dynamic game approach, Computers Math. Applic., 32 (5), pp. 43-56, 1996.
[22] Papadimitriou C. H., Games against nature, J. Comp. System Sci., 31, pp. 288-301, 1985.
[23] Piegat A., Fuzzy modelling and controlling, Akademicka Oficyna Wydawnicza Exit, Warsaw 2003 (in Polish).
[24] Shapley L., Some Topics in Two-Person Games, In Advances in Game Theory, Princeton University Press, 1964.
[25] Straffin P. D., Game theory, Wydawnictwo Naukowe Scholar, Warsaw 2001 (in Polish).
[26] Syslo M., Deo N., Kowalik J., Discrete optimization algorithms, PWN, Warsaw 1995 (in Polish).
[27] Vavock V., Aggregating Strategies, Conference on Computational Learning Theory, 1990.
[28] Wetzeel A., Evaluation of the effectiveness of genetic algorithms in combinatorial optimization, University of Pittsburgh, 1983.
[29] Zadeh L. A., Fuzzy limitations calculus; design and systems; methodological problems, Ossolineum 1980 (in Polish).

Implementation of Genetic Algorithm for Agents in E-learning in Finance

Nonika Bajaj[1] and Rakesh Agarwal[2]

Accenture Services Pvt. Ltd, Bangalore, India

E mail: {nonika.bajaj}{rakesh.agarwal}@accenture.com

Abstract. E-learning has become demanding and competitive in the recent times. Internet has now made financing dynamic by introducing the concept of financing through internet. E-learning in finance refers to learning about financial systems using Internet and e-commerce systems. The worldwide interest into financial markets has gained the interest in learning about the financial market. E-learning provides an alternative for classroom learning. In e-learning system actual problem is getting the best knowledge with in short period of time. Intelligent agents act as a catalyst in improving the quality of the resource for the learner in finance. Genetic Algorithm will create an appropriate sequence of the resources from the set of all possible solutions. The proposed knowledge based genetic algorithm incorporates agent to make it capable of finding optimal resources and provides the best alternative for effective learning in finances. This paper extends the idea to use intelligent agents using Genetic algorithm for making e- learning in finance more effective.

1 Introduction

The information age overwhelms us with loads of information every second yet we do not have the appropriate tools to convert the data into usable information that a user needs at a particular moment. Managing overloaded information is a challenge. The approach that can be used to manage overload information is through agents. Agents are primarily computer programs that can work for the users. They are responsible for doing tasks on behalf of the users. The agents could act as search engines and retrieve- user beneficial information from the internet based on which the e-learning module of the user could be made more appropriate and interactive. A database could be used to store the knowledge of such retrievals. The use of genetic algorithm for retrieving the information from the database can be very beneficial. Genetic algorithms are adaptive search techniques that can learn high performance knowledge structure. The genetic algorithms can improve the quality of business decision making [2]. Objective of this paper is to apply intelligent agent and genetic algorithm to make e-learning in finance more interactive and effective for the users.

1.1 Why e-learning in finances?

Derek Stockley [9] defines e-learning a "The delivery of a learning or education program by electronic means." e- learning involves the use of a computer or electronic device in some way to provide educational or learning material'. As per Lance Dublin[4] e-learning is defined as "The use of technologies internet, intranet, WANs, Satellite, CD-ROM, wireless – to provide and manage data, information, knowledge and learning to improve the performance of Organization and employees."

E-learning in finances describes an alternative to one-to-one classroom sessions. It provides enterprise-wide solutions through the usage of a set of tools and techniques. It reduces the learners training time which saves the cost related to training. It also maintains training consistency. It is the best way to deliver training to a number of people across different locations.

But the challenge is keeping the requirements of the end-user in mind - How would the user learn? How would he acquire the skill? And most important, how would a user be able to access the information which would be beneficial to him? The success of the 'e' in e-learning lies in the proper addressing of these questions and in ensuring proper design and implementation of all the concerned factors.

1.2 What are Agents?

Intelligent agents are called e-assistants or helper programs that can sit inside a computer and make the learning in e-learning happen dynamically to suit the need of the users. They could basically help in extracting knowledge relevant to the users from the available information. Intelligent agents can customize the content of e-learning according to the user's knowledge and skill. It can trap the user's likes and dislikes in various areas, the level of knowledge and the learning style and accordingly provide the information dynamically [11]. A dedicated e-learning portal can be built with multi-agent technology which would be responsible for finding the best possible match for the given requirement.

The difference [2] between software programs and agents are:

Autonomous: The agent goes beyond a simple software program where functionality is considered. It has a fair degree of control over its actions and it does not have to always wait for commands. It can take decisions of its own and function independently.

Persistent: Agents have the ability to run continuously. They persist over time i.e. the output of one stage affects the next stage.

Reactive: Agents can perceive changes in the environment and adapt its behavior in response to the changing environment.

Proactive: Agents are goal-oriented and take proactive initiatives towards fulfilling the goals set for them.

Personalized: Agents learn over time and can also be taught what to do in a particular

situation.

Social Behavior: Agents can interact and collaborate with other users so as to help them to achieve their goals.

2 Genetic Algorithms

Genetic Algorithms (GAs) were proposed by Holland [3], inspired by the concepts of evolution and natural selection [3]. In nature, the population of a species evolves according to principles of natural selection, of the "Survival of the fittest". Accordingly, the characteristics of an individual which is fit (good) are most likely to be passed to the next generation than the characteristics of an unfit individual. Genetic Algorithm [5] attempts to simulate the natures' theory. The algorithm supports a population of individuals, each one representing a possible solution to the problem of interest. Genetic algorithm is the most widely used random search technique. A genetic algorithm [3][5] is a domain – independent global search technique where elements in a given set of population are randomly combined until the termination criteria is satisfied. Population is evolved iteratively in order to improve the quality of the population. Initially the population is generated and with each generation a new set of solutions is generated by applying the three basic operators. The process is done by three important operators' i.e selection, crossover and mutation. Crossover is one of the most important operators of GA as it explores new search space by recombining existing elements of the space [5]. Mutation introduces new search elements that were not found in existing solutions.

3 Implementation of Intelligent agent and genetic algorithm to e-Learning

Intelligent agents – the so called e-assistants or helper programs - can sit inside a computer and make the learning in e-learning in finances happen dynamically to suit the need of the user. They would basically help in extracting knowledge relevant to the user from the available information. The supply and demand equation of the information market has changed. Today the information market is dominated by the demand for information rather than the supply of information as in the earlier days [10]. Similarly, the e-learning modules which were supplier-based earlier can give the end-user a more competitive edge in the global market by becoming demand-based. Intelligent agents can play a major role in making e-learning demand-based.

Intelligent agents can customize the content of e-learning according to user's knowledge and skill. It can trap the user's likes and dislikes in various areas, the level of knowledge and the learning style and accordingly provide the information dynamically [11].

A dedicated e-learning portal can be built with multi-agent technology which would be responsible for finding the best possible match for the given demand of

requirements. This portal would result in efficient search in a user-friendly environment. The agent based e-learning system [12] would be responsible for:

- Suitable matching of supply to demand of information.
- Re-matching the previous match dynamically as and when the supply/demand scenario changes
- Continuing the matching process until the best match is achieved or the time specified for the search runs out.

The agents could act as search engines and retrieve user-beneficial information from the Internet based on which the e-learning module of the user could be made more interesting and interactive. The content would thus change according to the different user-learning styles. The progress at each level would be judged and accordingly more and more information would be provided. A database could be used to store the knowledge. Figure-1 shows a proposed model below which shows how Intelligent agents can be used for effective e-learning.

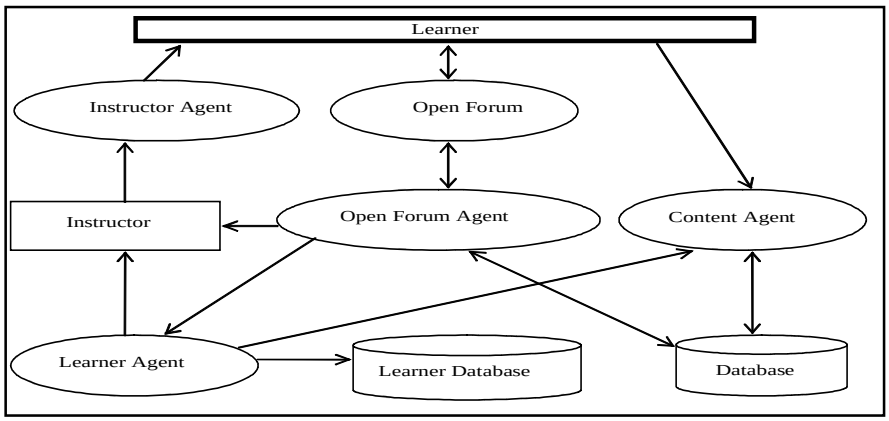


Figure 1: Intelligent agents in e-learning

The different actors in this model are:

Learner- The learner gets the contents related to the subject with the help of Content agent. If there is any query, the actor (Learner) can post the query in the "Open Forum".

Open Forum Agent- Open Forum agent will parse the query and analyze it. If it is a proper query then it searches the database based on the tokens. If the answer is found then it sends it back to learner. At the same time it interacts with the Learner agent to store the query and the result of the query in the Learner Database. If the result based on the tokens is not found, the query is sent back to the "Instructor".

Instructor-The Instructor has to send the information with the help of the Instructor Agent to the Learner regarding the query. The Instructor gets the details about the learner with the help of Learner agent.

Learner Agent- Learner agent stores the activities of a Learner as well as the query and the result in the Learner Database.

Content Agent- The Content agent takes the request from the Learner and also gets the details of the Learner from the Learner agent and then gives the content of the specific subject to the Learner based on his level of knowledge [13].

In this approach, the content agent search for the solution in the database. To optimize the search genetic algorithms are used. The life cycle of the knowledge based genetic algorithm for database as shown in Figure 2. An important fact with respect to e- learning is the ability of modifying the content in such a manner that the outcome can fit the learner's conditions.

The criteria for effectiveness of the solution are the time and the quality of the solution. The quality of the solution is improved by learner agent. The search is optimized by applying GA.

This will be an intelligent question answering system. The user or the learner will be able to get the required information asked for and the agents will reply the information according to the ability or understandability of the user or learner.

3.1 Why Genetic Algorithms?

The Genetic Algorithm will surpass their traditional cousins for robustness because GA is different in fundamental ways.

- GA searches from population and not from single point.
- Use of Objective functions
- GA uses probabilistic transition rules
- The use of Genetic Algorithm will optimize the search and the required information will be retrieved in shorter period of time.

3.2 Parameters of Genetic Algorithm

The two basic operators used are Cross over and Mutation. The cross over has a probability of sticking in the local minima so the mutation is performed so as to change the population.

Cross over

Cross over select the genes from the parent chromosome and create new offsprings. The easiest way to perform crossover is to choose randomly some crossover point and everything before this point copy from first parent and then everything after a cross over point copy from the second parent. The rate of crossover should be as high as 85%.

Mutation

The mutation is performed after the crossover and it avoids the possibility of solution getting stuck in the local optimum. Mutation is performed by changing randomly bits from 1 to 0 or 0 to 1. Mutation rate should be as low as 1%.

The life cycle of the genetic algorithm for the initial population chromosomes P using the fitness function and selection operator it chooses the subset of the chromosomes that is optimal for the moment. Then the process follows a iterative process by applying crossover and mutation operators unless the termination

criteria is satisfied. Let R= {R1, R2……..Rn} is a set of contents or resources. We will relate a collection of X=(KR) set of key words i.e KR is a set of keywords in R1. The Learner agent will update the database with the learning ability of the learner. To search for the query given by the learner the database needs to be searched for the matching keyword. The components of the GA, presenting for optimal search from database [7].

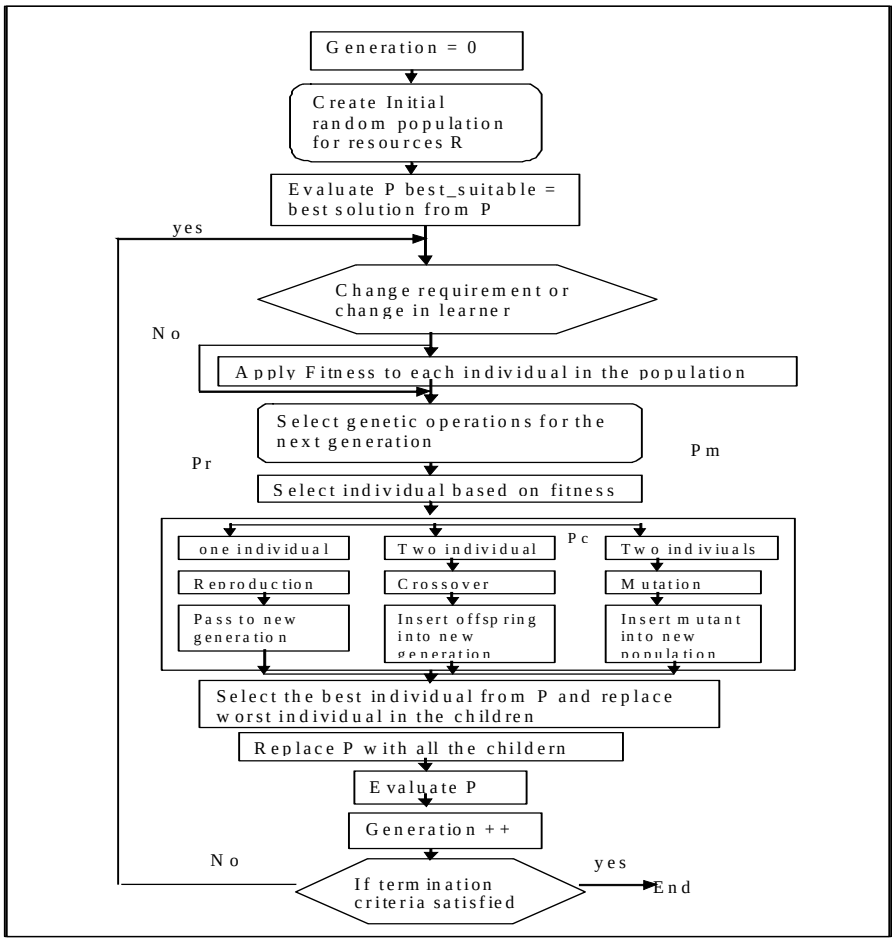


Figure 2: Lifecycle of the Knowledge based genetic algorithm for e- learning

The initial population is created by GA. The elements from the population (chromosomes) encode teaching resources. The important operator of GA is crossover operator. This operation creates new chromosomes via the crossing over operation. Select two chromosomes based on the fitness of the chromosome. Let we get two chromosomes C1 and C2, crossing breaks them in some points and get C11,C111 and C21C211 parts of those chromosomes. Then two new chromosomes are created Cs= C11C211and Ck = C21C111. The Pseudo code [10] for genetic algorithm for knowledge base is shown in Figure 3.

The process can be repeated until the best match resources or most effective chromosomes are obtained.

```
Begin /* genetic Algorithn*/
Create initial random population of
resources R1,R2…
compute fitness of each individual

WHILE  Not finished Do
Begin/*produce new population*/

FOR  (population size / 2) DO
Begin

IF Set of keywords K1 = No match THEN
selection of the two chromosome
(resourse) from the
population is made via roulette

operator selection for the
recombintion of the
existing chromosome to create new
generation offsprings

offsprings are inserted in the new
generation
End

If  Match = Found or population =
converges then
Finished
End
End
```

Figure 3 Pseudo code for a genetic algorithm for knowledge Base

4 Case Study

The experiments shows that if there were 15 set of keywords and the initial population generated is 300 then the best result is obtained at 198th generation. The suggested approach can be used for e-learning. This approach helps the learner to get the resources or content according to his/her ability and the results are found in possible short time. Learner agent stores the queries of the learner in a database and the contents that are searched using genetic algorithm from the database helps the learner to get the quality resources.

5 Conclusions

Implementing agents and GA in e-learning will make the learning process more dynamic. It will lead to a revolution in the education world. The issue that we presented in the paper is to construct the resources in such a way that solutions obtained are useful for the learner. The GA will work better and the learners will get knowledge of demand within short period of time [7]. However, by our experience we feel that this will act like a catalyst to the learning world in the industry.

6 Disclaimer

The author(s) of this paper gratefully acknowledge Accenture for their encouragement in the development of this paper. The information contained in this document represents the view of the author(s) and the company is not liable to any party for any direct/indirect consequential damage.

7 References

1. Quesenbery and Whitney: The Logic Underlying the Intelligent Technologies Used in Performance Support. Technical Communication. Nov2002. Vol. 49 Issue 4. p449. 9p. 1 chart. 1 diagram; (AN 7685560).
2. Deshar-nais P and Radhakrishnan J: Exploring agent support at the user interface in e-commerce applications. International Journal on Digital Libraries. 2002. Vol. 3 Issue 4. p284, 7p; (AN 6688344)
3. Xu Y and Vukovich G: A Fuzzy genetic Algorithm with Effective search and Optimization. International joint conference on neural networks 1993. p. 2967-2970.
4. VIRTUALLY HELPFUL. Telephony, 2/11/2002, Vol. 242 Issue 6, p56, 1p, 2bw; (AN 6128484)
5. Yi-wen Z and Jian-gang Y: A hybrid genetic Algorithm with Lamarckian Indiviual learning for tasks scheduling. Man and Cybernetics. 2004 IEEE International Conference on Volume 4. 10-13 Oct. 2004 Page(s):3343 - 3348 vol.4 2004.
6. Hu Y and Yang S.X: A Knowledge based genetic algorithm for path planning of a mobile robot. IEEE International conference on Robotics & automation New Orleans. LA – April 2004.
7. Culver E.R.S. and Hackos J.T: Perry's Model of Intellectual Development. Education, 72, 221-226 (1982).
8. Hovakimyan A, Sargsyan S and Barkhoudaryan S: Genetic Algorithm and the problem of getting knowledge in e-learning systems. IEEE International conference on Advanced learning Technologies (ICAL T'04) IEEE Computer Society. 2004 0-7695-2181-9.
9. E- learning in the new economy : http://www.linezine.com/elearning.htm
10. White S: Enhancing Knowledge Acquisition with constraint technology. Thesis degree of doctor of philosophy by. Department of Computing Science at university of Aberdeen.
11. Intelligent Software Agents on the Internet : http://www.firstmonday.dk/issues/issue2_3/ch_123/
12. Zhang T..Z, Fu Y.G and Shen N.M: Improve questions& Answers system by applying genetic algorithm. Third international conference on machine learning and cybernetics. 26-29. 2004
13. Virtual learning http://www.brunel.ac.uk/depts/cs/research/madira/collab.htm
14. Genetic algorithm : http://www.icce2001.org
15. Agarwal R, Deo A and Das S: From Education to Learning: A paradigm shift. Journal on Excellence in College Teaching. 2003.

Possibilities of fuzzy neuron networks in soft expert system

Nadezhda G. Yarushkina
Ulyanovsk State Technical University
Severny Venets str. 32, 432027 Ulyanovsk, Russia
jng@ulstu.ru

Abstract. Relied upon Zadeh's soft computing methodology, an integration of the expert system and the fuzzy neuron networks is presented. It is considered as a new approach to development the architecture of Soft Expert System. This paper is devoted to analysis of Fuzzy Neural Networks (FNN) with Genetic Tuning of parameters. The possibilities of FNN are considered for solving of basic and applied problems. That example is Data Miner for Relational Data Base in Soft Expert System for Performance Assessment

Keywords. Fuzzy sets and systems, Fuzzy Neural Network, fuzzy data base, complex system, expert system, multi-criterion choice, soft computing, Data Mining

1. Introduction

eBusiness is becoming an ever-increasingly complex activity involving numerous software tools, communication/transformation of data and collaboration among enterprises within a corporation. Among the new tasks that have appeared, assessing the performance of an enterprise involved in a large-scale collaboration in all phases of activity seems to hold the lead. The past decade has witnessed an obvious proliferation of software tools available for business managers and assessors such as knowledge bases, expert systems, rule-based dynamics analysis tools, decision making and support systems, cost estimation tools and others [1-8]. While this has resulted in greater accuracy and finer detail in the level of estimation, it also added to the complexity of the process overall. The mere existence of and local improvements in these individual analysis tools does not mean that they increase the productivity and efficiency of the business; in fact, the opposite may very well be true at a stage. The reason for this is that different tools are based on different principles. Moreover, they use different types of data: crisp or fuzzy, qualitative or quantitative. The multiplicity of ill-matched tools, heterogeneous data and polytypic approaches puts obstacles in the way of placing eBusiness on a broad footing. This reflects the well-known principle of uncertainty by Lotfi Zadeh. According to this principle, accuracy and meaning start to conflict since some point of analysis [9]. The problem becomes even more overbearing if part of the assessing process involves contributions from

partner teams in geographically disparate locations. Under conditions of geographical remoteness, lack of coordination is observed, firewalls create communication barriers, security is seen as a matter of utmost concern, and data transfer among tools becomes cumbersome and error-prone.

Considering all of these obstacles, eBusiness leadership and partnership, as well as isolated managers, auditors, performance assessors and consulting experts, are now keenly expecting to be provided with software tools that are intuitive, are easy to use, and that can interact with other tools involved in the process fairly seamlessly. The solution is not more or even better assessment tools – the solution is a framework that provides a solid foundation for overcoming these problems by combining different approaches, tools and technologies within a unified web based assessment environment.

The theoretical grounds for such solution have been prepared by many works in the field of fuzzy sets and systems [10], [11], decision analysis [12] and analytical planning [13]. On the other hand, recent technological advancements enable noticeable improvements in distributed computing and have made it feasible to construct new web-enabled environments such as FIPER [14] and the Globus Toolkit [15] providing a standard way for companies to publish their services for common use.

Recognizing the need for the aforementioned framework, Ulyanovsk State Technical University has teamed with Ulyanovsk State University, Russian Association of Artificial Intelligence, and a group of Volga Region enterprises in a three-year collaborative effort to develop a Component-based Integration Environment (COMBINE), a project sponsored by the Administration of Ulyanovsk City and partly by Russian Foundation for Basic Research (RFBR). The COMBINE joint venture team is well into its second year of work, with the various team members and additional sub-contractors, investigating different aspects of the environment and components to be provided in it. The ultimate outcome of the project is a commercially developed and supported Component-based Integration Environment for Performance Assessment of a Complex Enterprise (PACE), COMBINE-for-PACE. The requirements for this project continue to be established as the problems and obstacles become more clearly understood [16]. The remainder of this paper provides a high-level overview of this environment including objectives, methodology and technology used up to this stage of work.

2. Objectives

For a more penetrating insight into our COMBINE-for-PACE solutions, we start from the very notion of and need for performance assessing.

When several enterprises pooled to form a partnership, every partner, wishing to succeed in one's business (and in their common business, as well), should (and wish to) be clear in one's mind about the main issues:

How much is one's activity efficient?

How much is a partner's activity efficient?

What further effort should be mounted to increase the outcome of one's management activities and/or financial activities?
How to ensure confidentiality while getting answers on the above questions?
Solutions to the first and second questions mean that every partner should have a possibility to obtain the results of self-assessment and cross-assessment, correspondingly. The COMBINE tools must take into account all the data available from partners concerning their activities, first of all, – management activities and financial activities. Solutions to the third issue imply that the tool can present some positive recommendations to any partner who asks for them. Lastly, solution to the fourth issue must fully exclude any chance for unauthorized access to the data used within the COMBINE.
This interactive web-enabled consulting system has been developed to serve as a plug-and play framework for
1. top managers to
    - obtain the integrated express-estimates of their management activity, as well as financial activity of their enterprise, at a reasonable price and
    - compare their enterprise performance indices with the mean level of the same indices existing over the set of all partner or similar enterprises

in order to continually improve their managing decisions and increase the company income;
2. auditing firms to
    - obtain a new sort of service that they may sell,
    - attract new clients who can wish to order a complete set of auditing and consulting services after having an express-estimate from the COMBINE,
    - accumulate knowledge about businesses' state within a particular market segment and
    - verbalize their cumulative experience and formulate, on this basis, their own techniques and procedures for the enhanced performance analysis and assessment

in order to increase their client network; and for
3. any individual consulting expert to
    - automatically collect information about a market segment or business type,
    - express his/her expertise in consulting business in the form of his/her own inference rule base and
    - develop new components and publish them to the COMBINE Library (thus enlarging its capabilities)
    - to raise the level of his/her skill and his/her value as a professional within the consulting community.

3. Technology description

By the statement that the COMBINE is a web-based consulting service intended to assess the performance of a complex enterprise, we mean that the following technology instruments exist at an end user's disposal:
The IP-net supporting the base set of web services for data transfer with the unified numbering and routing policy under DNS.
The allotted Web-server providing availability of hypertext documents through the IP-net in response on Web-client's request.
Fuzzy relational data server providing storage and retrieval of large scale data bases.
The intellectual processing component viz. the above considered Soft Expert System.

Thus, the COMBINE is based on Client-Server technology and is implemented as a combination of the FRDS (viz. MS/SQL) server and Web-server. The main operation – assessing – is the direct responsibility of the fourth component, SES, organized as a toolbox for MATLAB. It uses the data from SQL server, forms the assessment and transfers it to the web server. The web server transforms this result into an HTML code and returns it to the web-client who requested for assessment.
The soft expert system is a fuzzy expert system possessing the following features:
For data mining, the SES uses statistical data interpreting them as the learning samples.
The SES represents knowledge in the form of linguistic variables (or membership functions, MF), fuzzy productions and the learned neural networks. Reduction of the set of fuzzy productions and tuning MF and the rule base are performed with a Genetic Algorithm (GA).
The SES operates as a system of interacting components.
The expert diagnosis procedure is realized by the SES, the intellectual heart of the system, while the others are realized as traditional data server and WWW-server. It is caused by the fact that usually the actual data available for an expert are crisp, precise values presented by the time series. This input information is transformed by the expert into the fuzzy (qualitative) conclusions, forecasts and recommendations. Analysis of the real-world expert conclusions shows that up to 80 % of them are either qualitative judgments or estimates of data trends, and only 20 % are recommendations. Consequently, the workflow of expert diagnosis can be determined as the soft expert diagnosis procedure shown in Table 1.

Table 1. Steps of the soft expert diagnosis procedure

The name of step	The name question to be answered at the step	Mathematical approach used
Step 1. Transformation of the crisp time series (CTS) into the fuzzy time series (FTS), – Fuzzification	What estimates should be assigned to the values of variables chosen to describe the object state?	Theory of fuzzy intervals [10]
Step 2. Estimation of the FTS tendencies	What are tendencies in the change in variables describing the object state?	Algebra of fuzzy intervals [11], neural networks theory
Step 3. Forming the expert summary for a state of the object under analysis	What are general terms that characterize the object as a whole?	Fuzzy inference system [9], decision making support system [14]
Step 4. Forecasting trend of developments	What is the prognosis of developments in the object state?	Neural network theory

Let us characterize the main components and methodology
Soft Expert System (SES) [12]containing:
Fuzzy Neural Network (FNN).
Fuzzy Relational Data Server (FRDS).
Fuzzy Logic Inference System (FLIS).
Genetic Algorithm (GA) for tuning parameters of FLIS and FNN [6].
Hierarchy Analysis Processor (HAP).
Decision Making Support System (DMSS).

4. Fuzzy neural network (FNN) as data miner in SES

Fuzzy time series is generated by Fuzzificator (**Fig. 1**) as a series of linguistic terms (fuzzy labels of quality) such as: “excellent”, “good”, “adequate”, “alarming” and so

on. The FNN transforms the FTS into the wider concepts expressing tendencies such as: “rise”, “decay”, “stabilization”, “fluctuation” or “chaos”.

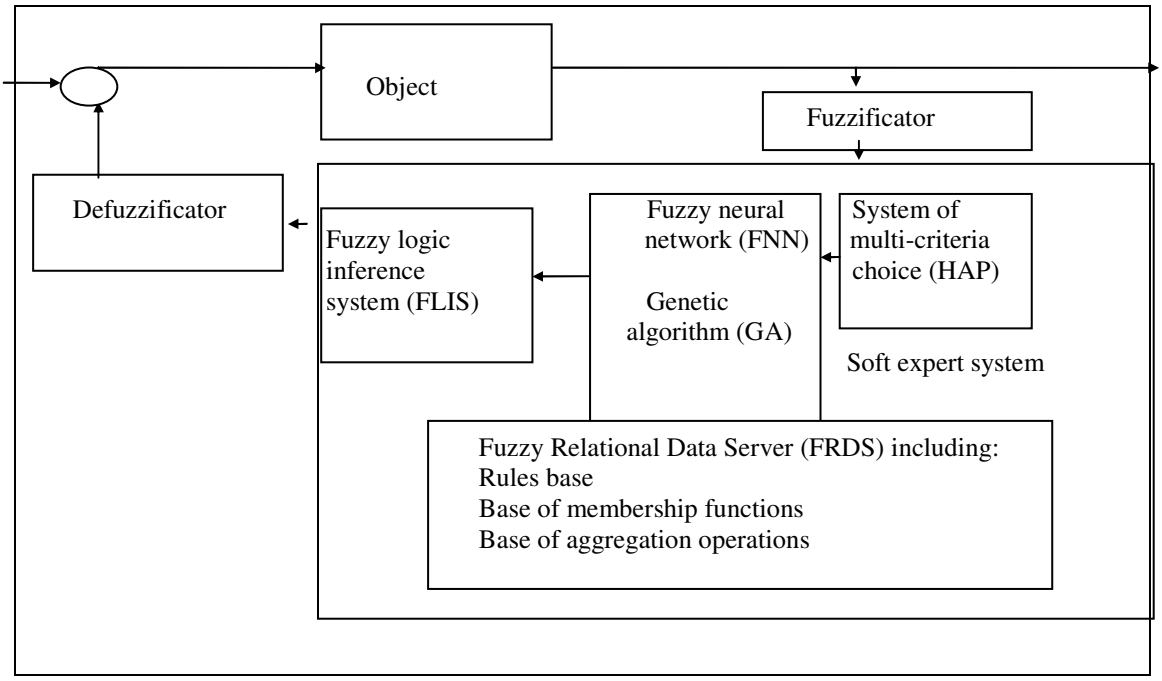


Fig.1. Integration of components in the SES

5. Fuzzy relational data server (FRDS)

The FRDS has been developed as a set of triggers and procedures for data server Oracle 8i to save membership functions, linguistic estimates and the data base containing fuzzy rules, crisp and fuzzy time series, and intermediate assessment results together with estimated states and tendencies. Because Oracle 8i does not have a means to save fuzzy data, a special repository has been developed.

6. Fuzzy logic inference system (FLIS)

The FLIS is implemented as a neural network able to learn to predict behaviour of a FTS and by virtue of this ability, to forecast the future state of object. Expert systems combining the traditional production inference mechanism and the learned neural networks, form the new class of expert systems called here neural expert systems [7,8,9].

7. Developments

In the expert systems traditionally used for economic analysis, visualization of results is performed by such means as verbal expert conclusions, line diagrams, histograms, circle diagrams, and/or three-dimensional diagrams. These traditional means of representing results require much time for understanding the information by the enterprise leader or expert economist. For this reason, it seems advisable to develop the Chernoff faces as an integral estimate expressed in a graphic image.
The objects (enterprises under assessment) are usually of large dimension, i.e., their economic state is determined by a large number of primary economic indices recorded in the account books. Presently, the account balance plan includes about 90 account types. Taking into consideration that some of the accounts are of regulation and control types, we can assume that there are 10-20 basic indices for enterprises with various forms of property and various activities. The constructive elements of the Chernoff faces are designed by using the images suggested earlier, which are adapted to special features of the enterprise economic state.

8. Results

At this point, the following functions and corresponding tools have been developed:

1. Visualization of the analytical parameter diagrams.
2. On-line formation of characteristic functions for qualitative estimates of economic data (the knowledge base of estimate scales).
3. On-line formation of the knowledge base of linguistic terms (using the scale within the interval [0, 1]).
4. Generating linguistic estimates of the basic economic parameter values using the scaling.
5. Evaluation of linguistic estimates of the analytic parameter values based on fuzzy interval algebra.
6. Generating a linguistic description of data trend diagrams and their analysis.
7. Evaluating the state of the enterprise under performance assessment.
8. Inference of a summary for the expert conclusion.
9. Forecasting the values of economic variables.

We used the COMBINE for assessing two Stock Corporations: “Volga-Dnepr” and “Nomatex”. The results of analysis and recommendations have proven to be in demand and of great business benefit.

9. Business benefits

Consulting services are very expensive. Of particularly high value are prices asked for full-scale assessment. Thus, the main benefit from the COMBINE is reduction of consulting costs. The benefit exists for consulting firms, too, because in one way or

another the data server and web server exist and the COMBINE intelligent component is not very costly.

The SES relies upon the knowledge base containing the rules that are based on the interrelations between economical indices and able to link typical business situations and advisable management solutions. The SES analyzes the enterprise initial data, establishes the critical values of indices, makes a conclusion about the effect of particular factors on the economic and financial result, recognizes the characteristic economical situations and, finally, forms recommendations for enterprise management.

Fig. 2. The example of Chernoff face

As a result, the user obtains the WWW-response to his Internet-request displayed in the form of a conclusion-resume. To illustrate the benefits from such a service, let us give a fragment of the conclusion-resume formed for the public stock corporation "Volga-Dnepr" (Ulyanovsk).

"Volume of business increases at a rapid pace thus testifying that the turnover rising trend is stable. Average annual rate of growth of airlift tonnage was 16 % by total flying time, considerably lower than corresponding turnover rate of growth. Volume of business outside of the NIS was increasing lower than within the NIS."

The fragment shows that the SES separates first of all the most significant economical indices and describes their dynamics. Next is selection of the typical situations for which the management actions may be recommended:

"The Stock Corporation "Volga-Dnepr" by the end of 1992 year is a company with high rate of growth. Skilfully and successfully, the enterprise made the use of the market factors and opportunities: drew additional income from the forming rate of exchange and borrowed funds. However at the same time, the enterprise has made some errors:

Profitability of main economic activity was weakly controlled that was a result of euphoria from the galloping fall in the ruble exchange.

The inflation factors remained out of control that allowed inflation "to eat up" the gain obtained from the exchange rate in the second half of the year.

Lack of attention was given to correlation between interest rate and borrowed funds that resulted in “eating away” up to one third of income.
Most likely there was not posed the problem of forming the own liquid assets that led to the quickened crises of solvency.
In this situation, the following steps should be taken:
Abandon the “common stock” practice of cost-flow assumptions.
Determine beforehand each contract payments profitability, was this a contract on air transport service, credit contract or otherwise.”
In this work, the COMBINE-for-PACE team develops an advanced environment and middleware object-oriented components in response to the primary obstacles in existing individual analysis tools as described in the introduction of this paper. It enables eBusiness partnership to obtain access to a novel web-based service for assessing the performance of a complex enterprise in order to continually improve its performance and increase thereby the company income. On its merits, this rather complex intelligent system is placed at end-user-client’s disposal in a relatively simple, easy to use, form and at a reasonable price. With the COMBINE, a new segment of consulting market emerges in the WWW-space as an essential factor of further eBusiness development.
The ultimate outcome of the project is a commercially developed and supported Component-based Integration Environment for Performance Assessment of a Complex Enterprise (PACE), COMBINE-for-PACE.

9. Conclusion

Thus, the COMBINE is based on Client-Server technology and is implemented as a combination of the FRDS (viz. MS/SQL) server and Web-server. The main operation – assessing – is the direct responsibility of the main component, Soft Expert System.
The soft expert system is a fuzzy expert system possessing the following features:

1. For data mining, the SES uses statistical data interpreting them as the learning samples.
2. The SES represents knowledge in the form of linguistic variables (or membership functions, MF), fuzzy productions and the learned neural networks. Reduction of the set of fuzzy productions and tuning MF and the rule base are performed with a Genetic Algorithm (GA).
3. The SES operates as a system of interacting components.

References

1. R. Askers, I. Bica, E. Kant, C. Randall and R. Young. SciFinance: A Program Synthesis Tool for Financial Modeling, AI Magazine, Vol. 22, No. 2, 2001, pp. 27-43.
2. J. Nolan. An Intelligent System for Case Review and Risk Assessment in Social Services. AI Magazine, Vol. 19, No. 1, 1998, pp. 39-47.

3. N. Matsatsinis, M. Doumpos and C. Zopounidis. Knowledge Acquisition and Representation for Expert Systems in the Field of Financial Analysis. Expert Systems with Applications, Vol. 12, No. 2, 1997, pp. 247-263.
4. Sh. Chen and Sh. Chang. Fuzzy Economic Production Quantity Model with Quality Improvement Investment and Fuzzy Costs. In: Proceedings of the 10th IFSA World Congress, 2003, Turkey, Istambul, pp. 523-527.
5. C. Kahraman and Z. Ulukan. Fuzzy Cash Flows under Inflation. In: Proceedings of the 7th IFSA World Congress, Czech Republic, Praque, Vol. 4, 1997, pp. 104-109.
6. D. Lakov. Fuzzy Neural Network Structures in Credit Risk Assessment. In; Proceedings of the 7th IFSA World Congress, Czech Republic, Praque, Vol. 4, 1997, pp. 109-116.
7. L.A. Zadeh. A Theory of Approximate Reasoning (AR). Machine Intelligence, Vol. 9, 1979.
10. D. Dubois and H. Prade. Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York, 1980.
8. D. Dubois and H. Prade. Criteria Aggregation and Ranking of Alternatives in the Framework of Fuzzy Set Theory. In: Fuzzy Sets and Decision Analysis, H. J. Zimmermann, L.A. Zadeh and B.R. Gaones (Eds.). TIMS Studies in Management Sciences, North-Holland, 1984, pp. 209-240,
9. G. J. Klir. Architecture of Systems Problem Solving. Plenum Press, New York, 1985.
10. T.L. Saaty and K.P. Kearns. Analytical Planning. The Organization of Systems. Pergamon Press, Oxford, 1985.
11. N. Yarushkina. Data Flow Distribution in Concurrent Engineering. In: R. Goncalves, R. Roy and A. Steigler-Garcao (Eds.). Advances in Concurrent Engineering, Proc. Intern. Conf. CE2002, Balkema: Lisse, Abington, Exton (PA), Tokyo, 2002, pp. 735-743.
12. N. Yarushkina, I. Semoushin COMBINE-for-PACE: Performance Assessment of a Complex Enterprise, In: P.
13. N. Yarushkina. Application of Fuzzy Relational Data Server for Concurrent Engineering. In: J. Cha, R. Jardim-Goncalves and A. Steiger-Garcao (Eds.). Concurrent Engineering - The Vision for the Future Generation in Research and Applications; Proc. Intern. Conf. CE2003, Book 2, Balkema: Lisse, Abington, Exton(PA), Tokyo, 2003, pp. 87-93.
14. N. Yarushkina. Soft Computing and Complex System Analysis. Int. J. General Systems, Vol. 30. No. 1, 2000, pp. 71-88.
15. A.N. Gorban, D.A. and Rossiev. Neural Networks on Personal Computers, Nauka, Novosibirsk, 1996.
16. H. Chernoff. Using Faces to Represent Points in K-Dimensional Space Graphically. J. ASA, 1973, p. 68.
17. J.S.R. Jang ANFIS: Adaptive Network-based Fuzzy Inference System. IEEE Trans. Systems, Man&Cybernetics, 1991Cunningham, M. Cunningham(Eds) eAdoption and Knowledge Economy. Issues, Applications, Case Studies, Part 2, IOS Press, 2004, pp. 1472-1479

The fuzzy neuron network for data store management system (ISMS)

Nadezhda G. Yarushkina
Ulyanovsk State Technical University
Severny Venets str. 32, 432027 Ulyanovsk, Russia
jng@ulstu.ru

Abstract. It is considered as a new approach to development the architecture of fuzzy data store management system. This paper is devoted to analysis of Fuzzy Neural Networks (FNN) [5]. An effective approach to study of FNN is to create the intellectual Data Store system with neuro-fuzzy management system [2]. The performance of the proposed scheme is tested through practical example: Neural-Fuzzy Management System for Text Warehouse.

Keywords. Fuzzy sets and systems, Fuzzy Neural Network, fuzzy data base, complex system, soft computing, Data Mining

1. Introduction

Fast increase of information volumes in a concrete organization promotes development of the special class of program complexes — information stores (repositories). In such complexes the unit of processing and storage is an information resource. The information resource is the file or a set of files united by common semantics and having a text summary. The text of the summary or the text of the resource unambiguously reflects the semantics of this resource. Automated identification of semantics will allow publishers not to think of which categories to publish the resource in. And users do not become hostages of experts and publishers subjectivity. Clusterisation automation generates the index allowing to realize various kinds of search, including browsing by categories. This accelerates search from the side of store users.

The basic functions of the information store are:

1. storage of information resources;
2. ensuring of access to resources from workplaces;
3. ensuring the possibility of publication of new resources;
4. the possibility of search of necessary resources by certain criteria.

2. Basic functions of ISMS

Use of a corporate file-server as the store of various electronic documentation can be the most simple example. In this case, storage of a resource is carried out on the file system of the server. Access from workplaces is organized through access by ftp-, http-, or smb-protocols in a local area network. The publication is done by loading file/files to the server under the same protocols. Search of the necessary resource is carried out as a usual travel round catalogues of the file-server. At first sight, this is a very cheap, convenient and simple decision. But it is convenient and cheap only till the certain sizes. The above mentioned set of functions of an information store does not take into account several facts which are not plain and always remain in the background.

Let us consider the problem more in detail. The structure of catalogues reflects the structure of categories of the information being stored. So, the problem of placing a resource in catalogues is the problem of a resource classification. The following questions must be solved during classification of a resource. Firstly, the tree of categories must be made. In our case, this catalog tree on the file-server is made by the group of publishers. Secondly, it is required to determine to which of these categories a resource corresponds most of all. The result of classification and, therefore, of subsequent search of the resource directly depends on the decision of these questions. But the result of the decision of these questions strongly depends on subjective representations of an expert about an object field. In case of independent publication of resources of one object field by different people, they can appear in different categories. Besides, these representations, even of the same person, can vary eventually. This will unavoidably affect quality and speed of search of the necessary resource by an end user. Thus, we cannot rely upon subjectivity of a person when we are talking about a large information store with great volumes of incoming resources. Some unified representation about the conformity of resources with the tree of categories is necessary.

3. Fuzzy clusterization on base Kohonen NN

The being offered decision of the problem of the tree of categories construction is the machine clusterisation of resources [1,3,4]. When clustering we rely upon the hypothesis that the semantic content of a text can be taken from the statistical distribution of words. That is, we can determine its semantics by frequency distribution of the words making the text of a resource (or a summary). According to

this supposition clusterisation is based on the data of the frequency analysis of the text. The process of the frequency analysis is connected with the morphological analysis of the words making the text. In this process all word forms of one word should be taken into account as one word to lower noise on inputs of a network. The mechanism of stemming is applied for this. This mechanism is the formal isolation of the stem of a word — the stable, graphically constant by declination and conjugation part of a word. Further, being based on results of the frequency analysis, the neural network of Kohonen is constructed, which, actually, carries out resources clusterisation . Relative frequencies of occurrence of key words on each resource are applied to inputs of the network. The network is trained on algorithm "the winner receives all". Outputs of the network correspond to the list of categories. The network can be considered to be trained after the termination of processing of all resources. An expert has to name clusters formed by the neural network for completion of the tree of categories construction. By this he can base on sets of key words and resources related by the network to the given category.

After clusterisation the problem of resources publication is solved by the same network. It is switched to the classifier mode and the function of activation most appropriate for the problem is chosen. Incoming resources pass stemming and the frequency analysis. The results of the frequency analysis are applied to inputs of the network. It is possible to make the decision on belonging of the resource to categories by the state of outputs of the network.

Thus, we relieve publishers and experts of the work on creation of the tree of categories and making the decision on the publication of a resource. The question "Where to publish a resource" is solved.

Increase of volumes of the designing information stored in an electronic form demands development of means of necessary resources search. The following kinds of search are usually used:

1. usual text search by sub-line with some patterns;
2. the search taking into account language morphology;
3. fuzzy search;
4. browsing by categories-catalogues.

Having great volumes and semantic heterogeneity of the stored information the search by sub-line with patterns and the morphological search are not enough. Precise search with set criteria is replaced by so-called "fuzzy search" allowing to specify inexact values of search criteria. But frequently there is no possibility to formulate precisely search criteria. It is only known that a resource may belong to certain (general enough) categories. In that case, the search by categories as the simple choice of the category from a tree with browsing all resources filling it happens to be much more convenient. Though this kind of search is historically the oldest, so far it has no developed standards for automated means for index construction.

The automated indexing of information resources for the subsequent search as browsing by categories demands creation of a special kind of an index. The index should reflect the structure of categories and their interconnection with information resources. The special format of the index satisfying these requirements has been developed by us. In it the structure of categories and their interconnection with information resources will be implicitly contained in the weights matrix of the neural network. Relative frequencies (F_R) are computed by the results of stemming and the

frequency analysis of the text making a resource. The neural network in the index is the neural network of Kohonen trained in the process of resources clusterisation. That is, the process of resources clusterisation also gives us the unique by opportunities index. On the basis of this index we can carry out all above described kinds of search, including browsing, as well as combine them. For example, it is possible to enable to carry out text-through fuzzy search only in certain categories.

By subsequent increase of volumes of the information incoming to the store, the situation when the current tree of categories will not cover all incoming resources any more may occur. The symptoms of such situation will be appearing of resources which do not belong to any category, or, on the contrary, which belong to a large number of categories. In this case reconstruction of the tree of categories taking into account new resources have come in is required. That is, repeated clusterisation of all the volume of resources with subsequent repeated naming of received clusters.

1. The information repository can construct the tree of categories for resources stored in it itself and carry resources along the tree. Now experts' time is only required to name obtained clusters on the base of lists of key words and resources related to the given category.
2. The repository decides itself which category this or that resource to relate to. This allows publishers not to think of which categories to publish a resource in. And users do not become hostages of experts and publishers subjectivity.
3. Clusterisation automation generates the index allowing to realize various kinds of search, including browsing by categories. It accelerates search on the side of store users.
4. The store looks after urgency of the tree of categories itself. By nonconformity of the tree of categories with incoming resources, that is by impossibility to classify a resource correctly, the store must make the decision about reclusterisation.

Thus, the basic result of the project will be the intelligent information store, including except type functions a number of extended functions, namely:

1. import of resources already available;
2. available resources clusterisation with the purpose of making the tree of categories;
3. incoming resources processing, that is performance of classification of incoming resources and their publication in corresponding categories;
4. the possibility of text-through and fuzzy search of resources in texts, the possibility of search by the tree of categories;
5. looking after the tree of categories so that it always covers incoming resources (so that there were no resources not corresponding to some category, or, on the contrary, corresponding to all categories).

Efficiency of realization of the project can be easily measured in such a way. By construction of documents classification systems, the basic quality criterion is the relative quantities of documents correctly related to the category. It is supposed that realization of the project will allow to reach more than 90 per cent correct classifications.

4. Conclusion

The basic result of the FNN-based Data Store project is the intelligent information store, including except type functions a number of extended functions, namely: import of resources already available; available resources clusterisation with the purpose of making the tree of categories; incoming resources processing, that is performance of classification of incoming resources and their publication in corresponding categories; the possibility of text-through and fuzzy search of resources in texts, the possibility of search by the tree of categories; looking after the tree of categories so that it always covers incoming resources (so that there were no resources not corresponding to some category, or, on the contrary, corresponding to all categories). The neural network in the index is the neural network of Kohonen trained in the process of resources clusterisation [6, 7, 8, 9].

References

1. L.A. Zadeh. A Theory of Approximate Reasoning (AR). Machine Intelligence, Vol. 9, 1979.
2. D. Dubois and H. Prade. Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York, 1980.
3. G. J. Klir. Architecture of Systems Problem Solving. Plenum Press, New York, 1985.
4. T.L. Saaty and K.P. Kearns. Analytical Planning. The Organization of Systems. Pergamon Press, Oxford, 1985.
5. J.S.R. Jang ANFIS: Adaptive Network-based Fuzzy Inference System. IEEE Trans. Systems, Man&Cybernetics, 1991
6. N. Yarushkina. Data Flow Distribution in Concurrent Engineering. In: R. Goncalves, R. Roy and A. Steigler-Garcao (Eds.). Advances in Concurrent Engineering, Proc. Intern. Conf. CE2002, Balkema: Lisse, Abington, Exton (PA), Tokyo, 2002, pp. 735-743.
7. N. Yarushkina. Application of Fuzzy Relational Data Server for Concurrent Engineering. In: J. Cha, R. Jardim-Goncalves and A. Steiger-Garcao (Eds.). Concurrent Engineering - The Vision for the Future Generation in Research and Applications; Proc. Intern. Conf. CE2003, Book 2, Balkema: Lisse, Abington, Exton(PA), Tokyo, 2003, pp. 87-93.
8. N. Yarushkina. Soft Computing and Complex System Analysis. Int. J. General Systems, Vol. 30. No. 1, 2000, pp. 71-88.
9. N. Yarushkina Application of a fuzzy relational server to concurrent engineering. Proceedings of Int. conf. On Fuzzy Sets and Soft Computing in Economics and Finance, Saint-Petersburg, Russia, 2004, pp. 530 –537

A NEURAL NETWORK APPROACH TO SOLVING A KIND OF NONLINEAR OPTIMIZATION PROBLEMS

R.Z. Salakhutdinov[1], A.V. Rubtsov[2]

1 - Institute of Problems of Informatics, Academy of Sciences, Tatarstan, Russia. & Institute of Economics, Management & Law (Kazan), Russia.
2 - Saratov State Socio-Economic University, Saratov, Russia
ravsal@mail.ru

Abstract. This paper offers a neural network that performs nonlinear constrained optimization. It can also be used in data compression and in extraction of principal components. As an economic application of the neural network, the task of portfolio optimization was chosen.

1. Architecture of Neural Network

Here we consider two neural networks. The first network performs unconstrained optimization. The second one is employed when constrained optimization is required.

Architecture of AKM Neural Network. Let's consider a network [4], [5] that consists of n neurons with weight vectors $\overline{W_i} = (w_{i1},...,w_{ik})^T, i = 1,...,n$. In fig.1. $n = 3$.

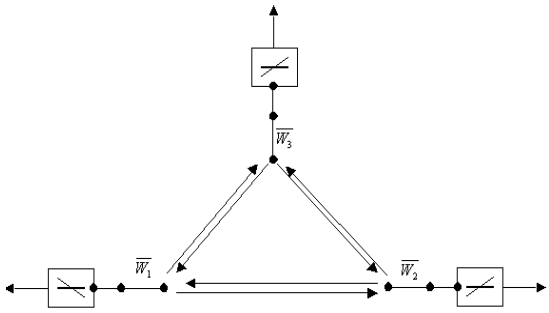


Fig. 1. Architecture of AKM Neural Network

The network computes the distances between the weight vectors of its neurons. The number of output signals of each neuron equals to the number of neurons in the network.

Algorithm of the network's functioning.

1. Each neuron computes the distances between its weight vector and the weight vectors of other neurons in the network.

Outputs of the 1st neuron	$\rho'_{11} = \left\|\overline{W_1} - \overline{W_1}\right\|$	$\rho'_{12} = \left\|\overline{W_1} - \overline{W_2}\right\|$	$\rho'_{1n} = \left\|\overline{W_1} - \overline{W_n}\right\|$
Outputs of the 2nd neuron	$\rho'_{21} = \left\|\overline{W_2} - \overline{W_1}\right\|$	$\rho'_{22} = \left\|\overline{W_2} - \overline{W_2}\right\|$	$\rho'_{2n} = \left\|\overline{W_2} - \overline{W_n}\right\|$
Outputs of the n-th neuron	$\rho'_{n1} = \left\|\overline{W_n} - \overline{W_1}\right\|$	$\rho'_{n2} = \left\|\overline{W_n} - \overline{W_2}\right\|$	$\rho_{nn} = \left\|\overline{W_n} - \overline{W_n}\right\|$

2. Computed outputs $R' = \begin{pmatrix} \rho'_{11} & \rho'_{12} & \rho'_{1n} \\ \rho'_{21} & \rho'_{22} & \rho'_{2n} \\ \rho'_{n1} & \rho'_{n2} & \rho'_{nn} \end{pmatrix}$ are compared with desirable outputs

$R = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{1n} \\ \rho_{21} & \rho_{22} & \rho_{2n} \\ \rho_{n1} & \rho_{n2} & \rho_{nn} \end{pmatrix}$ in objective function $E(\rho) = \frac{1}{n} e^2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} (\rho_{ij} - \rho'_{ij})^2$.

3. Weight vectors are updated according to the following learning rule

$$\overline{W_l}(t+1) = \overline{W_l}(t) - \alpha \frac{\partial E}{\partial \overline{W_l}(t)}, \quad \frac{\partial E}{\partial \overline{W_l}(t)} = -\sum_{j=1}^{n} \left(\rho_{lj} - \rho'_{lj}\right)\left(\overline{W_l}(t) - \overline{W_j}(t)\right).$$

As a result, the closest to R matrix R' is constructed.

Architecture of Neural Network for Constrained Optimization (NNCO). To construct NNCO, the outward signals $\overline{X_p} = \left(x_{p1}, ..., x_{pk}\right)^T, p = \overline{1, N}$ are attached to the inputs of AKM neural network (fig.2.)

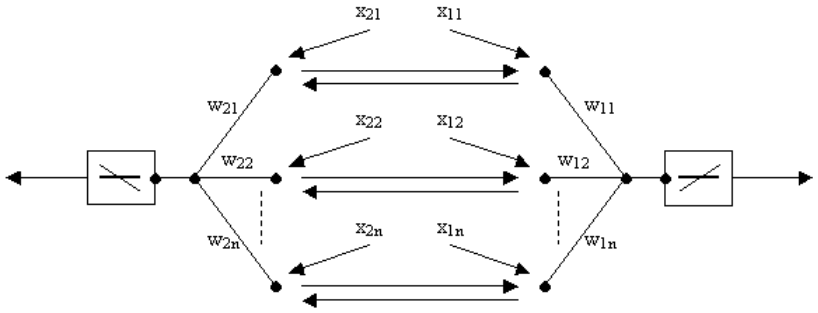


Fig. 2. Architecture of NNCO

Algorithm of the network's functioning.

1. The outward signals $\overline{X_p}, p = \overline{1, N}$ are processed: $out_j^p = \overline{X_p} \circ \overline{W_j}$.

2. The signals $\overline{W_i}, i = \overline{1, n}$ from other neurons are processed: $\rho'_{ij} = \overline{W_i} \circ \overline{W_j}$.

Vector operation ($\circ$) depends on the task to be performed. Below we consider some

tasks that employ NNCO.

2. The training algorithm for NNCO

To train AKM neural network a gradient descent technique is used. In case of training of NNCO we deal with constrained optimization. Here we consider some mathematical issues related to NNCO training process.

Let's there is a decreasing sequence $\{\beta_i\}$ ($\sum_{n=1}^{\infty}\beta_n=\infty$, $\sum_{n=1}^{\infty}\beta_n^2<\infty$). Then the minimum of the function $L_n(\overline{W})=\beta_n F(\overline{W})+(1-\beta_n)G(\overline{W})$, $n\to\infty$ is

$$\min F(\overline{W}) \quad (1)$$

$$G(\overline{W})=0 \quad (2)$$

under the following conditions:

$G(\overline{W})$, $F(\overline{W})$ are convex, continuous and bounded below; their first derivatives are continuous functions;

The first derivative of $F(\overline{W})$ is bounded.

To find a solution to (1)-(2) the following formula is used $\overline{W}(t+1)=\overline{W}(t)-\alpha\frac{\partial L_n}{\partial\overline{W}}$, $\alpha\in(0;1)$.

Let's there is a decreasing sequence $\{\beta_i\}$ ($\sum_{n=1}^{\infty}\beta_n=\infty$, $\sum_{n=1}^{\infty}\beta_n^2<\infty$). Then the maximum of the function $L_n(\overline{W})=\beta_n F(\overline{W})-(1-\beta_n)G(\overline{W})$, $n\to\infty$ is

$$\max F(\overline{W}) \quad (3)$$

$$G(\overline{W})=0 \quad (4)$$

under the following conditions:

$G(\overline{W})$, $F(\overline{W})$ are convex, continuous and bounded below; also, their first derivatives are continuous functions;

The first derivative of $F(\overline{W})$ is bounded.

To find a solution of (3)-(4) the following formula is used $\overline{W}(t+1)=\overline{W}(t)+\alpha\frac{\partial L_n}{\partial\overline{W}}$, $\alpha\in(0;1)$

3. Applications

Principal Component Analysis

A neural network approach to extraction of the principal component was first suggested by Oja [2]. Afterwards Sanger [1] developed a learning rule to extract all components from data distribution. According to Sanger's learning rule the weight vector of the $l-th$ neuron ($\overline{W_j}$) is updated as follows $\overline{W_l}(t+1) = \overline{W_l}(t) + \Delta\overline{W_l^p}(t)$,

$\Delta\overline{W_l^p}(t) = \alpha out_l^p(t)\left(\overline{X_p} - \sum_{\substack{i=1 \\ i\neq l}}^{n} out_i^p(t)\overline{W_i^p}(t)\right)$, where $out_l^p(t) = \overline{X_p}^T\overline{W_l}(t)$ the output of the $l-th$ neuron.

The rule performs the extraction of components sequentially, i.e. the first neuron is trained (the first (principal) component is extracted), the second neuron is trained (the second component is extracted) etc.

NNCO can also be used in feature extraction. According to derived earlier learning rules the objective function and constraints must be constructed. According to Hebbian model 0, the objective function is as follows

$$F(\overline{W_l}) = \sum_{i=1}^{N}\left(\overline{X_i}^T\overline{W_l}\right)^2 \to \max$$

It is known that neuron trained by Hebbian learning rule maximizes the scalar product of input vectors and weight vector of the neuron. As a result, the weight vector tends to principal component.

Normalization of weight vectors and their mutual orthogonality are constraints $\overline{W_i}\overline{W_i} = 1, \forall i$, $\overline{W_i}\overline{W_j} = 0, i \neq j$.

The function to be minimized by the $l-th$ neuron (extraction of the $l-th$ component) can be written as

$$L(\overline{W_l}) = \beta\sum_{i=1}^{N}\left(\overline{X_i}^T\overline{W_l}\right)^2 - (1-\beta)\left[\left(\overline{W_l}^T\overline{W_l} - 1\right)^2 + \sum_{\substack{j=1 \\ j\neq l}}^{n}\left(\overline{W_l}^T\overline{W_j}\right)^2\right].$$

The algorithm of the network's functioning during the extraction of the $l-th$ component.

1. The network processes outward signal $\overline{X_p}, p = \overline{1,N}$. The scalar product of outward signal and weight vector is calculated: $out_l^p = \overline{X_p}^T\overline{W_l}$.

2. The signals $\overline{W_i}, i = \overline{1,n}$ from other neurons are processed. The scalar products of weight vectors are calculated: $\rho_{ij}^{/} = \overline{W_i}^T \overline{W_j}, i, j = \overline{1,n}$.

3. Computed at step 2 matrix $R^{/}$ is compared with the matrix of desirable outputs R (unity matrix).

4. The weight vector of the $l-th$ neuron is updated according to the following formula $\overline{W_l}(t+1) = \overline{W_l}(t) + \alpha \frac{\partial L}{\partial \overline{W_l}}$,

$$\frac{\partial L}{\partial \overline{W_l}} = 2\beta \overline{X_i}^T \overline{W_l} \overline{X_i} - 4(1-\beta)\left[\overline{W_l}\left(\overline{W_l}^T \overline{W_l} - 1\right) + \sum_{\substack{j=1 \\ j \neq l}}^{n} \overline{W_j}^T \overline{W_l} \overline{W_j}\right].$$

The network is trained on-line, i.e. the weight vectors are changed after each sample $\overline{X_p}, p = \overline{1,N}$ is presented to the network.

Objects' Classification

Here we consider objects' classification performed by NNCO [5]. Let's assume that there are samples $\overline{X^i}, i = \overline{1,N}$ and an object W that requires to be classified (object) to one of the samples. Figure 4 illustrates a particular case of this classification.

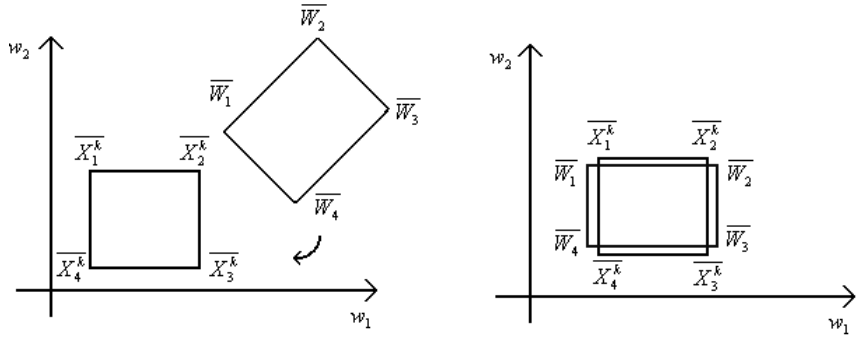


Fig. 3. Illustration of classification

According to the task the objective function to be minimized by NNCO can be written as $L(W) = \beta \sum_{i=1}^{n} \left\| \overline{X_i^k} - \overline{W_i} \right\|^2 + (1-\beta) \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\rho_{ij}^{/}(0) - \rho_{ij}^{/}(t)\right)^2$.

The algorithm of the network's functioning.

The network processes outward signals $\overline{X_i}, i = \overline{1,n}$ (signal $\overline{X_i}, i = \overline{1,n}$ is presented to the neuron with weight vector $\overline{W_i}, i = \overline{1,n}$).

The signals $\overline{W_i}, i = \overline{1,n}$ from other neurons are processed. The network computes the square distances between weight vectors of its neurons: $\rho_{ij}^{/} = \left\| \overline{W_i} - \overline{W_j} \right\|^2, i, j = \overline{1,n}$.

Computed at step 2 matrix $R^{/}$ is compared with the matrix of desirable outputs R (the matrix of initial distances between object's vectors).

The weight vector of the $l-th$ neuron is updated according to the following formula

$$\overline{W_l}(t+1) = \overline{W_l}(t) - \beta \frac{\partial L}{\partial \overline{W_l}},$$

$$\frac{\partial L}{\partial \overline{W_l}(t)} = -\beta\left(\overline{X_l^k} - \overline{W_l}(t)\right) - 2(1-\beta)\sum_{j=1}^{n}\left(\rho_{ij} - \rho_{ij}^{/}\right)\left(\overline{W_i}(t) - \overline{W_j}(t)\right).$$

The network is trained off-line, i.e. the weight vectors are changed after all samples $\overline{X_p}, p = \overline{1,N}$ are presented to the network.

It's worth noting that NNCO can be employed in many other economic tasks that require constrained optimization. That's why, it seems possible to use it in portfolio optimization.

Conclusion

Suggested neural network performs the tasks that require constrained optimization. Presented learning algorithm is used in extraction of principal components and in objects' classification.

References

1. Sanger T.D., «Optimal unsupervised learning in a single-layer linear feedforward neural network» IEEE Trans. Neural Networks, vol. 2, pp. 459-473, 1989
2. Oja E., «A simplified neuron model as a principal component analyzer» Journal of Mathematical Biology, vol. 15, pp. 267-273, 1982
3. Ezov A.A., Shumskii S.A. Neurokomputing i ego primenenie v ekonomike i biznese. M.:1998 g.-222s. (in Russian)
4. Salakhutdinov R.Z., Ismagilov I., Rubtsov A.V., "A Neural-Fuzzy Approach to Economic Data Classification", Proc. of Int. Conf. on Fuzzy Sets & Soft Computing in Economics and Finance, June 17-20, 2004, St.-Petersburg, Russia, pp.394-400.

5. Salakhutdinov R.Z., Rubtsov A.V., “A Neural Network Approach to Classification Based on Samples”, Proc. of Third International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, September 1-2, 2005, Antalya, Turkey, pp.143-148.

AHP, fuzzy AHP and its application for information security risks assessment

Igor Anikin

Kazan State Technical University, CS and Information Security Department,
K. Marx str. 10, Kazan, Russia
Institute of Problems of Informatics, Academy of Science Republic of Tatarstan
email: anikinigor@evm.kstu-kai.ru

Abstract. The paper considers the problem of information security (IS) risks assessment and analysis in the case of both qualitative and quantitative risk factors are presence based on analytic hierarchy process (AHP) and fuzzy AHP. We constructed in paper 3-level hierarchies for assess impact levels from realized the threats. We also constructed 6-levels hierarchies for assess likelihood of realizing some threats and impact levels of its realizing.

Information security risk is a function of the likelihood of a given threat-sources's exercising a particular potential vulnerability, and the resulting impact of that adverse event on the organization [4].

There are two basic approaches to the information security risks assessment – two-factor and three-factor. According to the two-factor approach, the information security risk assess by the formula (1), according to the three-factor approach – by the formula (2).

$$RISK=P_{incident}*IMPACT \quad (1)$$

$$RISK=P_{threat}*P_{vulnerability}* IMPACT \quad (2)$$

where $P_{incident}$ – possibility of realizing some information security incident, P_{threat} – possibility of realizing the threat, $P_{vulnerability}$ – possibility of exercising the vulnerability, making possible realizing the threat, IMPACT– resulting impact value of adverse event on the organization.

Methodological aspects of information security risks managment is consider in such manuals, guides and standards as CRAMM, NIST SP 800-30, OCTAVE, COBIT and etc. But these documents don't consider as a rule mathematic models of risks managment and its relevant assessment. Investigation such kind of models have considerable actuality.

The main difficulty on the stage of information security risks assessment is that necessity adequate assessment of risks factors in (1) and (2) – $P_{incident}$, P_{threat}, $P_{vulnerability}$, impact level. In practice these factors as a rule assess by experts. These factors assessment needs to provide the following requirements.

1. It's necessary to take into consideration oblique qualitative factors, not having elementary measurable features.

2. It's necessary to take into attention the model of intruder into procedures of risks factors assessment. We must consider its aims, motives, restrained factors.

3. It's necessary to provide relevance and consistency of risk factors assessments.

4. It's desire to form the end risks assessments in quantitative form for possibility to use well known analytical procedures to proceed risk values.

The necessity of simultaneous satisfaction above enumerated requirements don't let effectively use well known quantitative approaches to assess risk factors (like Delphi method) and qualitative methods based on introduction of qualitative scales and risk matrixes. We may to decide this problem with using Analytic Hierarchies Process (AHP) and fuzzy AHP for the risks factors assessment.

Analytic Hierarchies Process

The AHP was developed by T. Saaty [1] and has been identified as an important approach to multi-criteria decision-making problems of choice and prioritization. The main idea of AHP lie in the building multilevel hierarchy by decomposing complex problem or goal target into the more simple composite elements, factors, subtargets, that impacts on the goal, its allocation on certain levels of hierarchy, and identification the influence level between them. Next, the central question is being examined in AHP: how much influence the particular elements of the bottommost level of hierarchy on its top (goal target). For getting the answer on this question on each level of hierarchy (except topmost) the expert make all possible elements pair comparisons by its preferences to targets more high level of hierarchy on the qualitative scale from 9 point, offered by T.Saaty.

Let's consider the task of prioritization information security (IS) threats on the resulting impact value by their realization. We will use the following IS threats from IT Baseline Protection Manual catalogue (Table 1).

Table 1. The set of threats (from threats catalogue IT BPM)

T.1.4. Fire	T. 5.9. Unauthorized use of IT systems
T.4.1. Disruption of power supply	T 3.2. Negligent destruction of equipment or data
T.5.23. Computer viruses	T.3.36. Misinterpretation of events
T.5.4. Theft	T.4.22. Software vulnerabilities or errors
T.2.62. Inappropriate handling of security incidents	T.5.69. Higher risk of theft from a working place at home
T.3.6. Hazards posed by cleaning staff or outside staff	

We will use the following factors of resulting impact of these threats on some information asset.

1. K1 = «Confidentially violation»
2. K2 = «Integrity violation».
3. K3 = «Availability violation».

4. K4 = «Destruction of information asset».

Hierarchy corresponding to this task represented on the figure 1.

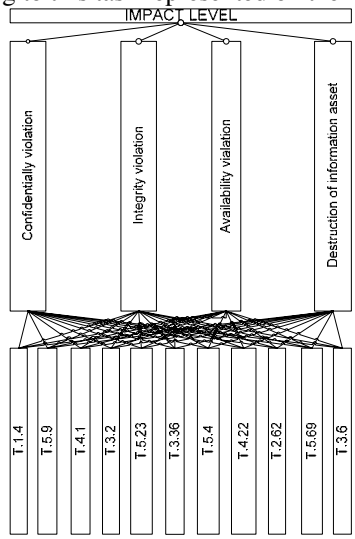


Fig. 1. Hierarchy of impact level of realizing the threat

On the results of pair comparsions the second level criterions of hierarchy were evaluated its priorities. They represented in table 2.

Table 2. Criterions priorities

Criterions	Weight
Confidentially violation	0,0645694
Integrity violation	0,5478897
Availability violation	0,2403545
Destruction of information asset	0,1471864

Results of calculation the total priorities of treats and priority index represented in table 3.

Building and analysis hierarchy of many levels for risk factors assessment is possible. These hierarchies let include to themselves models of intruder, motives of his beheiviour, his politics, restrainig factors. It's their great advantage. We have built the example of sixlevel hierarchy for assessment the possibility of realizing 8 threats, that represented on the figure 2. On the figure 3 hierarchy for assessment the impact level from realizing threats have represented. Have calculated the priorities of alternatives, which determined possibilities from realizing threats on one of the hierarchies, impact level on the another hierarchy and find their composition, we'll get the quantitative assessment of IS risks value.

Table 3. Results of calculation the total priorities of threats

Threat ID	Threats priorities for the «Confidentially violation»	Threats priorities for the «Integrity violation»	Threats priorities for the «Availability violation»	Threats priorities for the «Destruction of information asset»	Total threats priorities	Priority index
T.1.4.	0,0213105	0,3287222	0,2864647	0,2708314	0,304994	1
T.5.9.	0,2504948	0,0908009	0,1130992	0,0551396	0,096552	4
T.4.1.	0,0213105	0,0366213	0,0217385	0,035262	0,032612	8
T.3.2.	0,0213105	0,0559718	0,078117	0,0897835	0,067701	7
T.5.23	0,0761002	0,0908009	0,0502552	0,0897835	0,087619	5
T.3.36	0,0330271	0,0239192	0,0330406	0,035262	0,027613	9
T.5.4	0,1947716	0,0908009	0,2104532	0,1982236	0,134422	2
T.4.22	0,0761002	0,0908009	0,0502552	0,0897835	0,078226	6
T.2.62	0,0493015	0,1437237	0,1130992	0,0897835	0,124057	3
T.5.69	0,1781563	0,0239192	0,0217385	0,0230737	0,023102	10
T.3.6	0,0781167	0,0239192	0,0217385	0,0230737	0,023102	10

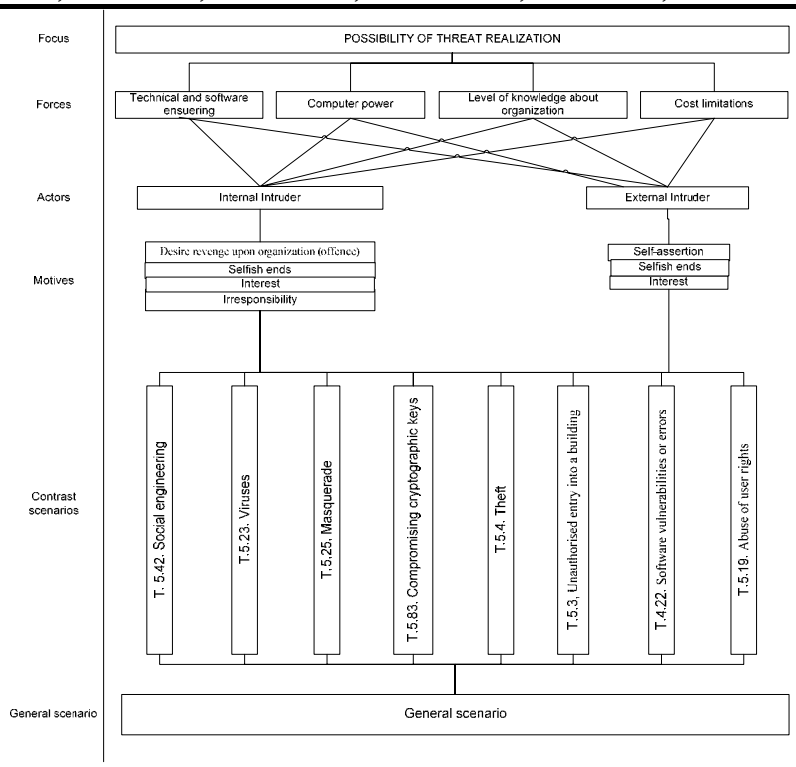


Fig. 2. Hierarchy of possibilities from realizing threats

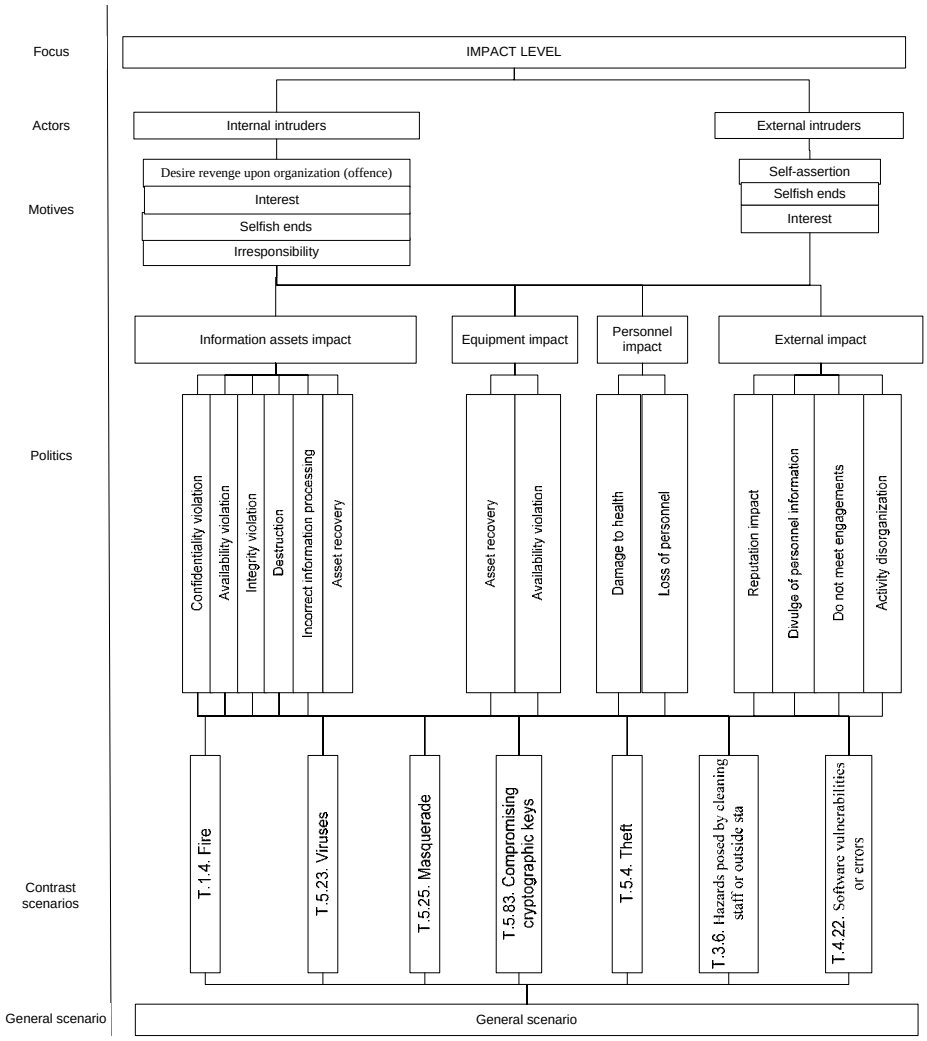


Fig. 3. Impact level hierarchy from realizing the threat

AHP in the task of building questionnaire for assessment possibilities of realizing threats and vulnerabilities

We solve the task of building questionnairies in the paper (analogous to well-known one from psychology) for assessment possibility of realizing threats. Questionnairies like this showed its high effect in CRAMM standard [5] on the stage of risks assessment. These questionnairies contain list of questions with variants of answers for them. Choosing some answer get some points to overall sum. Analysis the overall

number of points (sum), after answering on all questions of the test, let determine the possibility of realizing the threat, taking on the qualitative scale (for example, low, middle, high, very high).

One of the problems in building that kind of questionnaires is the necessity of definite proof of points belonging on each answering question. Often these points define subjective by experts, but this approach couldn't be proved and defined points might be call a question.

For define proof of points belonging on each answering question, we used AHP. We suggested three-level hierarchy shown on the figure 4, which analysis let solve this task.

This hierarchy include following levels:

1. Focus of the problem.

2. Names of questions (Factors, influense on possibility of realizing threat (vulnerability)).

3. Answers on the questions.

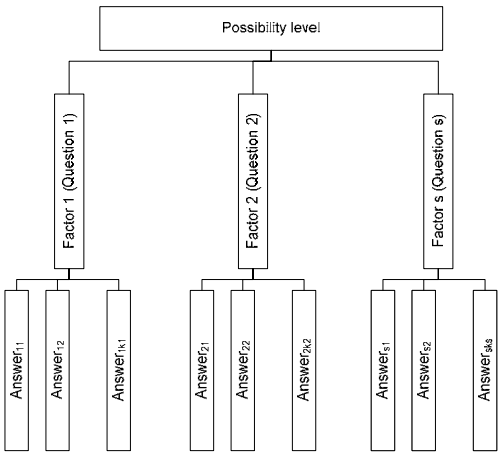


Fig. 4. Hierarchy the questionnaire

Weights of connections between levels 1-2 of hierarchy determine overall volume of points that belonging to the questions. Weights of connections between levels 2-3 determine the distribution of points between the answers of the question.

On the figure 5 shown the example of hierarchy for the assessment possibility of realizing threat «Masquerade» by the factors, proposed for it in the CRAMM method. Analysis this hierarchy and comparison getting points with points, represented in the CRAMM method, let make conclusions about more high exactness of our approach.

Analogous to this approach may be used for building questionnaires for assessment impact value, caused by realizing the threat. This kind of approach may be used for building methodic of information assets categorization into security levels.

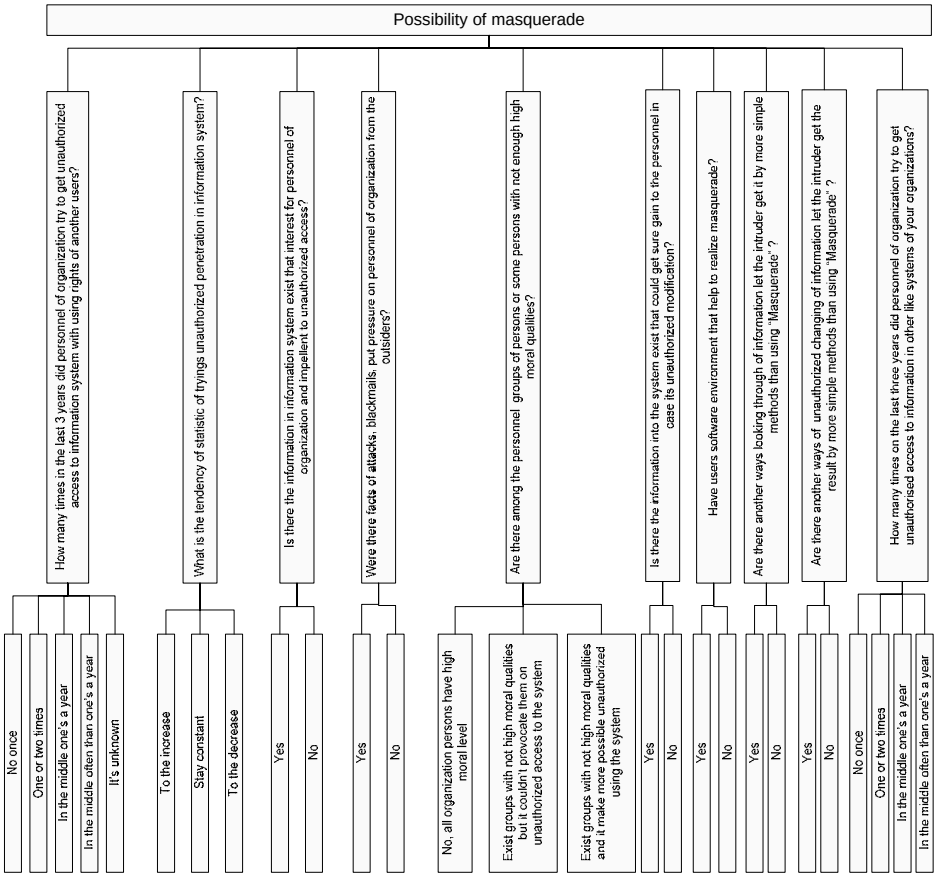


Fig. 5. Hierarchy the questionnaire for the assessment possibility for realizing the threat «Masquerade»

Fuzzy AHP and its application for information security risks assessment

The main deficiency of investigated approaches for information security risks assessment appear that the AHP usage is difficult, when expert have uncertainty in pair comparison or variations individual pair comparison by making group decisions.

In this case the result of pair comparison of elements C_i and C_j more prefer formed as fuzzy numbers $\tilde{a}_{ij}$, with membership function in triangle form, for example [2,3].

In this case we'll get fuzzy matrix of pair comparisons.

$$A = \begin{pmatrix} \tilde{a}_{11} & \dots & \tilde{a}_{1n} \\ \dots & \dots & \dots \\ \tilde{a}_{n1} & \dots & \tilde{a}_{nn} \end{pmatrix}$$

where $\tilde{a}_{ji} = \tilde{a}_{ij}^{-1}$.

Further we culculate fuzzy priorities $\tilde{w}_i$ of investigate elements in a following way:

$$\tilde{Z}_i = (\tilde{a}_{i1} \otimes \tilde{a}_{i2} \otimes \dots \otimes \tilde{a}_{in})^{1/n}$$
$$\tilde{w}_i = \tilde{Z}_i / (\tilde{Z}_1 \oplus \tilde{Z}_1 \oplus \dots . \oplus \tilde{Z}_n)$$

So all considered methods for information security risks assessment might be generalized on the fuzzy level. This approach is more prefer when expert have uncertainty in pair comparison or variations individual pair comparison by making group decisions.

Conclusion

Preference of usage AHP for information security risks assessment conclude in possibility of include models of intruder into the process of risk assessment, get assessments of qualitative criterions in quantitative form. AHP is effective for decision the problems, where it's difficult to gather objective data and leading motives for making decisions appear subjective preferences of peoples. The main difficulties of using AHP appear in the big excess leading comparisons and also absence formal methods of decompositions the system for hierarchy building. In general case AHP is difficult to use for analysis of threats and vulnerabilities when set of them consist more than 15 elements.

Fuzzy AHP is more prefer when expert have uncertainty in pair comparison or variations individual pair comparison by making group decisions.

References

1. Saaty T.L. The analytic hierarchy process. McGraw-Hill, New York, 1980.
2. van Laarhoven P.J.M., Pedrycz W. A fuzzy extension of Saaty's priority theory. Fuzzy Sets and Systems, V. 11, 1983, pp. 229-241.
3. Salo A.A. A fuzzy ratio comparisons in hierarchical decision models. Fuzzy Sets and Systems, V 84, 1996, pp. 21-32.
4. Petrenko S.A. Simonov S.V. Information risks management. Economically defensible security. Moscow: IT Press, 2004.
5. www.insight.co.uk/cramm/index.htm

Benefit/cost ratios estimation for information security countermeasures based on AHP and fuzzy AHP

Igor Anikin

Kazan State Technical University, CS and Information Security Department,
K. Marx str. 10, Kazan, Russia
email: anikinigor@evm.kstu-kai.ru

Abstract. The paper considers the problem of benefit/cost ratios estimation for information security countermeasures when both qualitative and quantitative benefit and cost factors are presence. To decide this problem we use Analytic Hierarchy Process (AHP) and its fuzzy extension. We suggested some benefit and cost hierarchies for the problem of choosing the best Intrusion Detection System for the one of the national enterprises in Kazan city.

Information security is the field of information assets and its supposing infrastructure secure from accidental or premeditated threats natural or artificial character, pregnating with making damage for owners or users of information assets and its supposing infrastructure [1].

Existing information assets has hundreds of possible threats and vulnerabilities be able to break its security states. Also there are many countermeasures exist would be pretty expencive and difference by their efficiency.

So management of information security of information systems expect not purposeless implementing and using security systems. Approach, based on unaimness covering threats and vulnerabilities in a way of choosing “any” countermeasure, doomed to failure. In this conditions information security management will be ineffective and organization will pay too much money for the ensuring information security never returning it.

So in present days information security transformed into the science about realizing adequate and effective on quality and cost approach to ensuring protection all elements of information system.

One of the tasks need to be decided within such kind of opinions on information security is the task of reasoning choice some variant of implementation countermeasure; proceed from its possible benefits and expenses (costs). This task may be considering like choosing one from several alternatives on the benefit-cost ratio.

Let we have set C that including n alternatives. $C = \{ C_i \}_{i=\overline{1,n}} = \{ C_1,...,C_i,...,C_n \}$

Choosing each of alternatives $C_i \in C$ implies its possible costs $P_i \geq 0$ and possible benefits $B_i \geq 0$. So benefit-cost ratio CE_i for the alternative C_i determined as (1).

$$CE_i = \frac{B_i}{P_i} \quad (1)$$

So rule of choosing the best effectiveness alternative on the benefit-cost ratio is (2)

$$C_{i^*}, \text{ where } i^* = \arg\max_i CE_i = \arg\max_i \frac{B_i}{P_i} \quad (2)$$

where C_{i^*} - the best alternative.

The main imperfection of formula (2) is the necessity of assess qualitative factors, determined possible benefits and costs of choosing the alternative, such as «Simplicity of product support», «Flexibility of licensing policy», «Operation complexity» and etc. The elementary, measurable on some quantitative scale factors are absent for its. So that factors often evaluate on some qualitative scales with 3-7 qualitative gradations. But such kind of qualitative values couldn't be used straight in formulas (1) and (2).

So deciding the task of quantitative evaluate benefit/cost ratio when qualitative cost and benefit factors are present is very actuality. This task may be decided with the analytic hierarchies process (AHP) and fuzzy analytic hierarchies process [2,3,4]. Fuzzy AHP is prefer for using when uncertainty in knowledge about investigated object exist.

Analytic Hierarchies Process (AHP) and Fuzzy Analytic Hierarchies Process (fuzzy AHP)

The AHP was developed by T. Saaty and has been identified as an important approach to multi-criteria decision-making problems of choice and prioritization. The main idea of AHP lie in the building multilevel hierarchy by decomposing complex problem or goal target into the more simple composite elements, factors, subtargets, that impacts on the goal, its allocation on certain levels of hierarchy, and identification the influence level between them. Next, the central question is being examined in AHP: *how much influence the particular elements of the bottommost level of hierarchy on its top (goal target).*

For getting the answer on this question on each level of hierarchy (except topmost) the expert make all possible elements pair comparisons by its preferences to targets more high level of hierarchy on the qualitative scale from 9 point, offered by T.Saaty (table 1) [2]. So, several pair-comparison matrixes are formed.

For pair-comparison matrix $A = (a_{ij})$, where a_{ij} - value representing the ratio for the relative importance of the *i*th subelement over the *j*th subelement, we have $a_{ij} = 1/a_{ji}$. All of the pair-comparison matrix elements $a_{ij} \in \{1,2,...,9\} \cup \{1/2, 1/3, ..., 1/9\}$.

Table 1. 9-Point Saaty scale

Scale value	Definition	Interpretation
1	Equal significance	Each of the elements compared is equally significant to the target immediately higher
3	Somewhat more significant	Experience and estimations indicate that one element is somewhat more significant than the other
5	Considerably more significant	Experience and estimations indicate that one element is considerably more significant than the other
7	Extremely more significant	One element has shown itself in the past to be extremely more significant than the other
9	Absolutely dominant significant	The two elements differ as much as it possible in terms of their significance
2,4,6,8	Intermediate values	A compromise must be found between two neighboring assessments

Let we have pair-comparison matrix $A = (a_{ij})$. We could to obtain priority vector *w* of alternatives being compared with respect to element more high level of hierarchy by evaluating the eigenvector of the matrix *A* belonging to maximum eigenvalue λ_{max} of *A*.

To determine how much inconsistency is in the *A* matrix, Saaty defined a measure of deviation from consistency, called a consistency index, as: C.I. = ($\lambda_{max} - N$)/(*N*-1), where *N* is the dimension of the matrix. Then, Saaty calculated a consistency ratio (C.R.) as the ratio of the C.I. to a random index (R.I.) which is the average C.I. of sets of judgments for randomly generated reciprocal matrices (table 2).

Table 2. Average Random Consistency Index (R.I)

N	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
R.I.	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

When we assess weights of all relations between hierarchy elements, we could assess intensity influence all the elements on bottommost level of hierarchy to the top of hierarchy.

To obtain the result coincides with the eigenvector solution for the pair-comparison matrix *A*, one can evaluate vector *w* with normalized geometric means of the rows of the matrix *A*.

To obtain the result coincides with the maximum eigenvalue λ_{max} of the pair-comparison matrix *A*, one can add the columns of *A* and multiply the resulting vector by the priority vector *w*.

The classic AHP is imperfection when we have uncertainty in the judgments or fuzziness of individual judgments in group decision making. In this case the ratio of pair comparison of ith and jth hierarchy elements most prefer formed as fuzzy number $\widetilde{a}_{ij}$, for example as fuzzy triangle $\widetilde{a}_{ij} = (4,5,6)$. So we'll get fuzzy pair-comparison matrix.

$$\ddot{\widetilde{A}} = \begin{pmatrix} \widetilde{a}_{11} & \dots & \widetilde{a}_{1n} \\ \dots & \dots & \dots \\ \widetilde{a}_{n1} & \dots & \widetilde{s}_{nn} \end{pmatrix} \text{ where } \widetilde{a}_{ji} = \widetilde{a}_{ij}^{-1}$$

For the fuzzy pair-comparison matrix $\widetilde{A}$ we may calculate fuzzy priority vector $\widetilde{w}_i$.

$$\widetilde{Z}_i = \left(\widetilde{a}_{i1} \otimes \widetilde{a}_{i2} \otimes \dots \otimes \widetilde{a}_{in}\right)^{1/n}$$
$$\widetilde{w}_i = \widetilde{Z}_i / \left(\widetilde{Z}_1 \oplus \widetilde{Z}_1 \oplus \dots \oplus \widetilde{Z}_n\right)$$

Fuzzy AHP allow obtaining more relevant assessments of elements priority when we have uncertainty in the judgments or fuzziness of individual judgments in group decision making.

Analytic Hierarchy Process in benefit/cost ratios estimation

The basic concept of benefit/cost analysis comes form economics. The analyst investigates a project or alternative course of actions and attempts to quantify all possible positive (benefit) and negative (cost) aspects into a common currency as dollars. With positives and negative aspects in common currency, their ratio can be calculated to determine whether the benefits outweigh the costs according (1). The project with the highest benefit/cost ratio would give the highest return for monies expended (2).

When using AHP and fuzzy AHP for benefit/cost analysis, the same approach is taken expect that AHP priorities rather than dollars are used as the common currency of comparison. One may use such qualitative priorities as «Simplicity of product support», «Flexibility of licensing policy», «Operation complexity» and etc.

The decision-maker proceeds with two AHP hierarchies – one measuring benefit and the other measuring cost for the same set of projects [4]. The benefit priorities are then compared to the cost priorities to see which option has the highest ratio of benefit/cost.

Simple benefit and cost hierarchies represented on the figures 1, 2. Top level of these hierarchies is goal - benefits or costs assessment process, level two – benefit or cost categories (for example capital investments costs, training costs, information security risk decrease benefits, security management benefits and etc).

Benefit/cost ratios estimation for information security countermeasures

Let's consider the task of choosing the one of the four Intrusion Detection System (IDS) for one of the national enterprises in Kazan city:

1. Secure INTRANET with IDS Snort+IDS Center.
2. Secure INTRANET with IDS Real Secure Net Sensor.
3. Secure each of the host with Outpost Personal Firewall.
4. Secure each of the host with Sygate Personal Firewall.

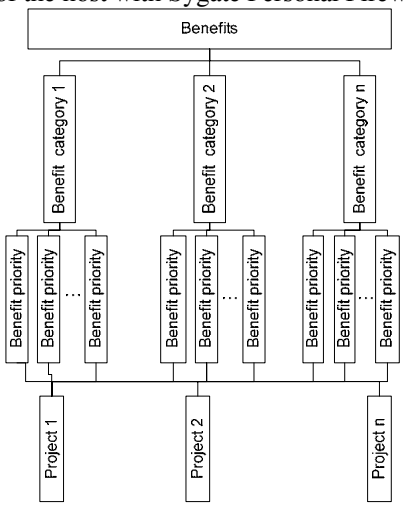


Fig. 1. Benefits hierarchy

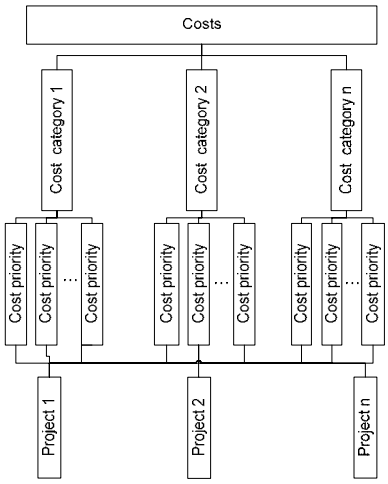


Fig. 2. Costs hierarchy

The following factors were used in benefit and cost hierarchies.

BENEFIT FACTORS

Economic benefits

1. Reduction the service staff, that supporting information security management.
2. Number of attacks identificated by IDS.
3. Frequency of realizes updates with new attack signatures.
4. Information security risk decrease value.

Security management benefits

1. Possibility for the centralized management of the security policy.
2. Monitoring of users activity.
3. Attacks references system completeness.
4. Intelligent processing sensors' information.
5. Log files contents.
6. Possibility for filter data on application level.

Counteraction to hacker

1. Effectiveness of the decoy system.
2. Incidence responding policy effectiveness.

COST FACTORS

Economic costs

1. Capital investment.
2. Implementation and configuration complexity.

Technical costs

1. Management difficulty.
2. Demands to technical environment.

Working costs

1. Inconvenience and difficulty of using of application interface.
2. Number of false alarms.
3. Unsteadying to the attacks.
4. Possibility of IDS deception.

The benefits and costs hierarchies for choosing the IDS represented on the figures 3 and 4.

We obtain benefit priority vector b and cost priority vector c using AHP. These priorities are shown in the table 3. The benefit-cost ratios are shown in the table 4.

Table 3. Benefits and costs of choosing IDS

Benefits		**Costs**	
$b = \begin{pmatrix} 0{,}261 \\ 0{,}596 \\ 0{,}077 \\ 0{,}066 \end{pmatrix}$	*Snort* *REAL Secure* *Outpost PF* *Sygate PF*	$c = \begin{pmatrix} 0{,}222 \\ 0{,}519 \\ 0{,}126 \\ 0{,}138 \end{pmatrix}$	*Snort* *REAL Secure* *Outpost PF* *Sygate PF*

Table 4. Benefit-cost ratios

Snort	**Real Secure**	**Outpost PF**	**Sygate PF**
$\frac{0.261}{0.222} = 1{,}18$	$\frac{0.596}{0.519} = 1{,}15$	$\frac{0.077}{0.125} = 0{,}6$	$\frac{0.066}{0.138} = 0{,}48$

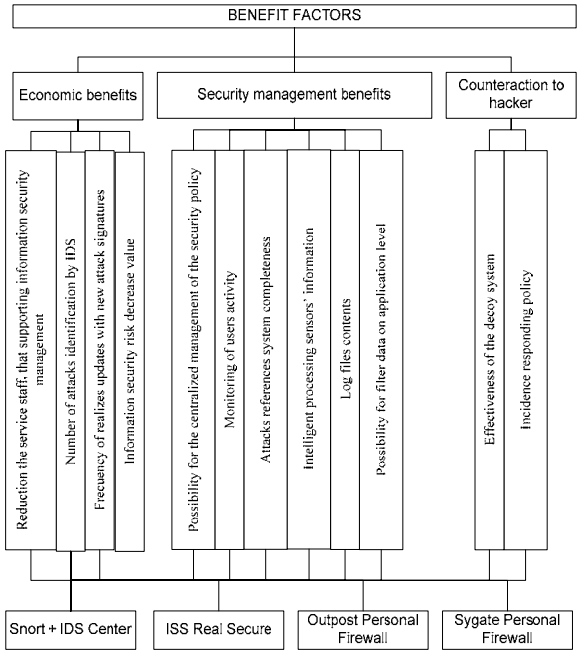


Fig. 3. Hierarchy of benefits from taking root IDS

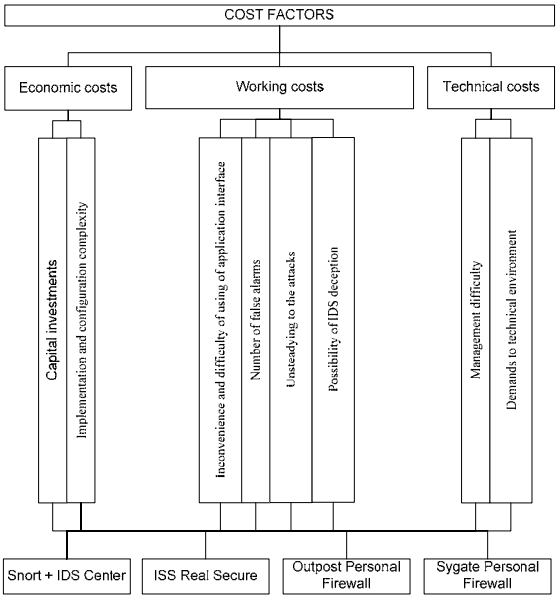


Fig. 4. Hierarchy of expenses from taking root IDS

So it would be better for organization take IDS Snort. But closes the benefit/cost ratio values for Real Secure and Snort meaning that it is possible would be needed take Real Secure in the future.

We decide also this task with using Fuzzy AHP. We obtain fuzzy priority vector $\widetilde{b}$, fuzzy cost priority vector $\widetilde{c}$ with using Fuzzy AHP. Elements of vectors $\widetilde{b}$ and $\widetilde{c}$ are fuzzy triangles. Next we calculate fuzzy benefit-cost ratio according (3)

$$\widetilde{CE}_i = \frac{\widetilde{b}_i}{\widetilde{c}_i} \quad (\mathbf{3})$$

where $\widetilde{b}_i$ - fuzzy benefits for ith IDS (element of fuzzy priority vector $\widetilde{b}$), $\widetilde{c}_i$ - fuzzy costs for ith IDS (element of fuzzy priority vector $\widetilde{c}$).

By prioritization fuzzy ratio as fuzzy numbers, we receive, that better for organization is IDS Snort too, but the worst for organization is the IDS Outpost PF.

References

1 Devyanin P.N., Michalskiy .., Pravikov D.I., Tcherbakov A.Y. Theoretical bases of computer security. ..: Radio and communication, 2000.

2. Saaty T.L. The analytic hierarchy process. McGraw-Hill, New York, 1980.

3. van Laarhoven P.J.M., Pedrycz W. A fuzzy extension of Saaty's priority theory. Fuzzy Sets and Systems, V. 11, 1983, pp. 229-241.

4. Salo A.A. A fuzzy ratio comparisons in hierarchical decision models. Fuzzy Sets and Systems, V 84, 1996, pp. 21-32.

5. Lai V.S., Wong B.K., Cheung W. Group decision making in a multiple criteria environment. A case using the AHP in software selection. European Journal of Operational Research, V. 137, 2002, 134-144.

Creation of the Rating of Stock Market Analytical Systems on the Base of Expert Qualitative Estimations

Olga A. Siniavskaya, Boris A. Zhelezko

Belorussian State Economic University 220672 Republic of Belarus, Minsk, Partisan Ave. 26
o.siniavskaya@sam-solutions.net, BorisZh@yandex.ru

Abstract. In this article the algorithms of creation of the rating of stock market analytical automated systems are proposed. They are based on Zahedi's analytic hierarchy process modification in which triangular fuzzy numbers are used to transform qualitative expert estimations into quantitative ones. An example of practical using of these algorithms is considered.

1 Introduction

Analytic hierarchy process (AHP) devised by Thomas Saaty [3] became one of the wide-spread multi-criteria methods of the best alternative choice. There are many modifications of Saaty's AHP method. One of them was made by Fatemah Zahedi in her work [4] and adopted for quantitative evaluation of expert systems. However, the principle of superiority evaluation, proposed in that work, entails the essential difficulties for expression of the expert opinions if estimation criteria have a qualitative nature. The objective of this article is working out of the algorithms of creation of the rating of stock market analytical systems on the base of Zahedi's method modification, in which qualitative expert opinions about alternatives would be employed during the calculation of generalized quantitative scores of alternatives. This work is the continuation of authors' investigation concerning the AHP improvement [1, 5] and the estimation of decision support systems quality [6].

2 AHP method suggested by F. Zahedi

F. Zahedi proposes the following simplified algorithm for comparative evaluation of expert systems (ES) [4].

Stage 1. The functional structure of expert systems is defined in the form of evaluation criteria hierarchical tree.

Stage 2. Relative significances (weights) of criteria are defined on each hierarchy level. The sum of all criteria weights must be equal one in each criteria group on hierarchy levels. Then global significances of the lowest level criteria (hierarchical tree leaves) are calculated by the following formula:

$$w_k^G = \prod_{v=1}^{H_k - 1} w_v^L \cdot w_k^L \qquad (1)$$

where w_k^G – global significance of k-th criterion which is a leaf of the hierarchical criteria tree; v – numbers of hierarchy levels; H_k – amount of hierarchy levels in the branch, which includes k-th criterion; w_v^L – local significances of parent criteria of branch, which includes k-th criterion on H_k-th hierarchy level; w_k^L – local significance (weight) of k-th criterion which is a leaf of the hierarchical criteria tree.

Stage 3. Comparison of alternatives superiority is conducted on criteria, which are the leaves of the hierarchical criteria tree. Relative superiority of one alternative by another one is expressed within the [0, 100] interval. We shall mark superiority of i-th ES above j-th ES on k-th criterion as e_{ijk}. If there is no superiority, then e_{ijk}=e_{jik}=0, otherwise e_{jik}=-e_{ijk}. Comparison matrices have the following form (table 1).

Table 1. Matrix for alternatives comparison according to Zahedi's AHP-method

Criterion k	**ES 1**	**ES 2**	**…**	**ES N**
ES 1	*0*	e_{12k},c_{12k}	…	e_{1Nk},c_{1Nk}
ES 2	-e_{12k},c_{12k}	*0*	…	e_{2Nk},c_{2Nk}
…	…	…	*0*	…
ES N	-e_{1Nk},c_{1Nk}	-e_{2Nk},c_{2Nk}	…	*0*

In order to reduce uncertainty in expert answers Zahedi proposes alongside with the significance estimation e_{ijk} to indicate a coefficient of confidence in this estimation $c_{ijk} \in (0, 1]$. If c_{ijk}=1, it means that expert confidence in the estimation e_{ijk} is equal to 100%. Definitive superiority estimation on each criterion is calculated as $e_{ijk} \cdot c_{ijk}$.

Stage 4. Aggregate score of i-th alternative is calculated as follows:

$$Z_i = \sum_{k=1}^{m} \left(w_k^G \sum_{j=1}^{n} c_{ijk} \cdot e_{ijk} \right) \qquad (2)$$

where m – amount of evaluation criteria, which are the leaves of the hierarchical criteria tree; n – amount of alternatives.

Stage 5. Rating of alternatives is created by sorting them on decreasing of their aggregate scores Z_i. Alternative with the maximal Z_i score is recognized as the best alternative.

The main shortcoming of Zahedi's method is necessity of expression of expert opinion about alternatives superiority through the quantitative scale from 0 to 100. For instance, intermediate estimation of superiority may be equal to 30, 35, 37, 40 points, and it is difficult for expert to discern these values, especially if the evaluating software is complex and estimation criteria are qualitative (possibilities of graphical interpretation of some financial indices, quality of programming editor, etc). But such difference in quantitative estimation may lead to incorrect result.

3 Using of qualitative estimations and fuzzy numbers for representation of experts' opinions

The most convenience form of representation of experts' opinions about alternatives superiority could be the verbal scale, such as «very high superiority, high superiority, average superiority, etc». However, verbal assessment of superiority has two shortcomings. Firstly, probability of unconformity of the matrix with quantitative superiority estimations generated from the matrix with qualitative superiority estimation is not excluded. Secondly, if a number of alternatives is large then estimation of their relative superiority in qualitative form would be difficult and lengthy.

In order to eliminate these shortcomings we propose to estimate not relative superiority of alternatives, but alternatives quality (x_{ik}) on defined criteria by means of the verbal scale. In table 2 and on figure 1 an example of possible scale of verbal estimations correspondence to quantitative estimations within [0, 100] interval is represented.

Table 2. An example of qualitative and quantitative estimations correspondence

Qualitative (verbal) estimation	Quantitative (fuzzy interval) estimation in $[a_1, a_2, a_3]$ form
Absence (ABS)	[0, 0, 10]
Low (L)	[0, 10, 30]
Below Average (BA)	[10, 30, 50]
Average (AVG)	[30, 50, 70]
Above Average (AA)	[50, 70, 90]
High (H)	[70, 90, 100]
Very High (VH)	[90, 100, 100]

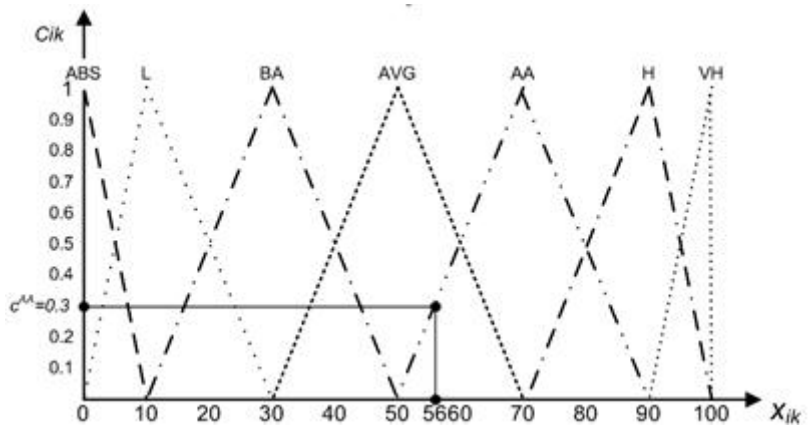


Fig. 1. An example of possible fuzzy interval estimating scale

Quantitative fuzzy interval estimation, which corresponds to verbal formulation, is represented as symmetric or asymmetric fuzzy triangular number $[a_1, a_2, a_3]$. That

means that quality of alternative could be estimated as a_2 with highest confidence, but if the expert's confidence in his (her) estimation is less than 100%, then quantitative value of estimation may vary within $[a_1, a_3]$ interval.

There is also an example of transformation of ambiguous verbal estimation «AVG, c^{AVG}=0.7; AA, c^{AA}=0.3» into quantitative estimation on figure 1.

Production rules of verbal estimations transformation into quantitative form are represented in table 3.

Table 3. Rules of verbal estimations transformation into quantitative form

Rules	Antecedent (qualitative estimation, x_{ik}')	Consequent (quantitative estimation, x_{ik})
Transformation rule for unambiguous estimations	IF expert's confidence in some qualitative (verbal) estimation is equal 100%, i.e. c_{ik}=1	THEN $x_{ik}=a_2$ for corresponding fuzzy number from table 2
	Example	
	IF x_{ik}'=«H, c^H=1»	THEN $x_{ik}=90$
Transformation rule for ambiguous estimations	IF two coefficients of confidence c_{ik}^1 and c_{ik}^2 correspond to two nearby qualitative (verbal) estimations	THEN $x_{ik}=c\cdot(a_2-a_1)+a_1$, where c – coefficient of confidence in higher verbal estimation; a_1, a_2 – elements of fuzzy triangular number, which corresponds to higher verbal estimation
	Example	
	IF x_{ik}'=«AVG, c^{AVG}=0.7; AA, c^{AA}=0.3»	THEN $x_{ik}=0.3\cdot(70-50)+50=56$

Superiority of i-th alternative above j-th on k-th criterion is defined by the following formula [5]:

$$e_{ijk}=x_{ik}-x_{jk} \quad (3)$$

Generalized scores of alternatives are calculated as follows:

$$Z_i = \sum_{k=1}^{m}\left(w_k^G \sum_{j=1}^{n} e_{ijk} \right) \quad (4)$$

Suggesting algorithms of creation of alternatives rating, based on qualitative expert estimations (*Verbal AHP* and *Verbal$^+$AHP*), are compared with Zahedi's algorithm in table 4. Stages 1, 2 and 4 of all this algorithms are executed by experts. The rest stages can be automated. Hereupon AHP algorithm based on qualitative estimations of alternatives quality on criteria, (*Verbal$^+$AHP)*, which doesn't requires the relative comparison of the alternatives, more simply then two another AHP algorithms in spite of the fact that it contains the greater number of stages.

Table 4. Comparison of Zahedi's AHP algorithm with the algorithms based on qualitative experts estimations (*Verbal AHP* and *Verbal$^+$AHP*)

Zahedi's AHP algorithm	*Algorithm based on qualitative estimations of relative superiorities of alternatives (Verbal AHP)*	*Algorithm based on qualitative estimations of alternatives quality on criteria (Verbal$^+$AHP)*
1. Forming of evaluation criteria hierarchy tree		
2. Assessment of local criteria weights by means of Saaty's AHP	2. Assessment of local criteria weights by means of expert methods in which there are no relative comparison of criteria pairs (for instance, *Fishburn* method)	
3. Calculation of global criteria weights by the formula (1)		
4. Evaluation of alternatives relative superiority in quantitative form within [0, 100] interval with taking into account of confidence in estimations	4. Evaluation of alternatives superiority in qualitative form by means of scale from table 2 with taking into account of confidence in estimations	4. Evaluation of quality of alternatives by means of scale from table 2 with taking into account of confidence in estimations
	5. Transformation of verbal estimations into quantitative form by means of rules from table 3 (with use of e_{ijk}' and e_{ijk} instead of x_{ijk}' and x_{ijk}).	5. Transformation of verbal estimations into quantitative form by means of rules from table 3
		6. Calculation of relative superiorities of alternatives as differences (distances) between their absolute estimation by the formula (3).
5. Calculation of aggregate scores of alternatives Z_i by the formula (2).	6. Calculation of aggregate scores of alternatives Z_i by the formula (4).	7. Calculation of aggregate scores of alternatives Z_i by the formula (4).
6. Creation of the rating of alternatives on decreasing of Z_i value	7. Creation of the rating of alternatives on decreasing of Z_i value	8. Creation of the rating of alternatives on decreasing of Z_i value

4 Creation of the rating of stock market analytical systems on the base of suggested AHP modifications

Suggested AHP algorithms may be efficiently used for the evaluation of stock market analytical systems and creation of their rating. We shall consider an example of the practical using of the given AHP algorithms. Evaluating systems are described in table 5, created on the base of the article [2].

Table 5. Description of stock market analytical automated systems

Stock market analytical automated systems	Supported methods, functions and features
OMEGA Research Trade Station (TS)	Includes 150 indicators, 15 trade systems, 30 graphical tools, additional tools *ShowMe* and *PaintBar*, built-in programming language *EasyLanguage*, programming editor *PowerEditor*. Has a possibility to analyze different length trends.
OMEGA Research Super Charts (SC)	Includes more than 80 indicators, 15 trade systems, 15 graphical tools, additional tools *ShowMe, PaintBar* and *Expert Indicator* (latter tool for indicators explanation). Has a possibility to analyze only day, week and month trends. Has no programming editor.
Reuters Graphics Professional (RGP)	Includes 65 indicators, 6 types of trend lines (up to 200 lines on the graph), possibilities of new indicators and return curves creation. Allows showing up to 16 graphs simultaneously. Has an ability of data exchange with Microsoft Excel. Supports fundamental analysis.
DanaLyzer Euro Charts (DEC)	Includes 40 indicators. Has a possibility of data representation in the form of linear graphs, histograms, «Japanese candles», etc. Uses different length trends.
EQUIS MetaStock Professional (MSP)	Includes more than 120 indicators and linear dependences, a possibility of new indicators creation. Supports 9 graphical styles. Allows creating, testing and optimizing trade systems.

Qualitative estimations of stock market analytical automated systems on 8 criteria are represented in table 6. In this table global criteria weights w_k^G are also adduced. In table 7 stock market analytical systems ratings are built based on the using of suggested *Verbal AHP* and *Verbal*$^+$*AHP* algorithms.

High score of *Reuters Graphics Professional* is explained by that only this system among considered software supports the fundamental analysis, which has high significance. *Super Charts* has the worst score because this system is intended mainly for training in technical analysis and has limitations in functionality.

The sequence of the alternatives coincides in both ratings. The difference in absolute values of rating indexes is explained by using the same scales in the first case – for estimation quality of stock market analytical systems on criteria, and in the second case – for estimation their relative superiorities. The given ratings characterize func-

tionality of considered analytical systems and could be used hereinafter for the analysis of acceptability of their prices, as that was made in the works [4] for expert systems and in the book [5] for office software.

Table 6. Qualitative estimations of stock market analytical automated systems

Criteria	w_k^G	TS	SC	RGP	DEC	MSP
Fundamental analysis	0.3	ABS, c^{ABS}=1	ABS, c^{ABS}=1	AVG, c^{AVG}=1	ABS, c^{ABS}=1	ABS, c^{ABS}=1
Possibility of using different length trends	0.15	H, c^{H}=1	BA, c^{BA}=0.6 AVG, c^{AVG}=0.4	AVG, c^{AVG}=1	H, c^{H}=1	AA, c^{AA}=1
Possibility of different graphical interpretation of stock price trends	0.1	H, c^{H}=1	BA, c^{BA}=0.6 AVG, c^{AVG}=0.4	BA, c^{BA}=1	AVG,c^{AVG}=0.3 AA, c^{AA}=0.7	BA, c^{BA}=0.8 AVG,c^{AVG}=0.2
Possibility of new indicators creation	0.1375	ABS,c^{ABS}=0.5 L, c^{L}=0.5	ABS, c^{ABS}=1	H, c^{H}=1	ABS,c^{ABS}=0.5 L, c^{L} =0.5	H, c^{H}=1
Amount of indicators	0.1125	AA, c^{AA}=0.1 H, c^{H}=0.9	AVG,c^{AVG}=0.2 AA, c^{AA}=0.8	AVG,c^{AVG}=0.6 AA, c^{AA}=0.4	BA, c^{BA}=0.8 AVG,c^{AVG}=0.2	AA, c^{AA}=0.3 H, c^{H}=0.7
Analysis of stock returns	0.1	ABS, c^{ABS}=1	ABS, c^{ABS}=1	AA, c^{AA}=1	ABS, c^{ABS}=1	ABS, c^{ABS}=1
Additional tools	0.06	AA, c^{AA}=0.6 H, c^{H}=0.4	H, c^{H}=0.8 VH, c^{VH}=0.2	AA, c^{AA}=0.1 H, c^{H}=0.9	H, c^{H}=1	AA, c^{AA}=0.9 H, c^{H}=0.1
Built-in programming language and programming editor	0.04	H, c^{H}=1	BA, c^{BA}=0.9 AVG,c^{AVG}=0.1	ABS, c^{ABS}=1	ABS, c^{ABS}=1	ABS, c^{ABS}=1

Table 7. Ratings of stock market analytical automated systems

Number	**Stock market analytical system**	**Generalized score Z_i** with using *Verbal*$^{+}$*AHP*	with using *Verbal AHP*
1	Reuters Graphics Professional	91.8	170.5
2	Trade Station	15.2	5.2
3	MetaStock Professional	8.6	2.3
4	Euro Charts	-42.6	-87.3
5	Super Charts	-73.0	-90.7

5 Conclusion

As a result of given research two algorithms for evaluation of stock market analytical automated systems were worked out, based on Zahedi's AHP-method modifications. The main distinctive peculiarity of these algorithms consists in using of qualitative (verbal) expert estimations in the generalized quantitative score obtaining. Verbal estimations applied for the expression of alternative relative superiority on criteria in the first algorithm (*Verbal AHP*), and for the alternatives evaluation on criteria – in the second algorithm (*Verbal$^+$AHP)*. Such a principle of alternatives and their superiorities evaluation is more convenient that quantitative assessment during special software estimation on qualitative criteria (for example, on quality of graphical features, programming editor, etc).

Production rules of transformation of the qualitative estimations into quantitative ones by means of the fuzzy interval scale were worked out in this research and used in AHP for the first time. Research results were applied to creation of stock market analytical software rating.

References

1. Ahrameiko, A.A., Zhelezko, B.A., Ksenevitch, D.V., Ksenevitch, S.V. Generalization of Saaty's analytic hierarchy process for using of fuzzy interval expert data. New Information Technologies (NITE-V): Proceedings of International scientific conference, Minsk, BSEU. Vol. 1 (2002) 217–222 (in Russian).
2. Logai, P. Main software for technical analysis of commodity and financial markets. "Companion", (1998) #49 [Electronic resource]. – Mode of access: http://www.forex.ua/soft/soft_osn.shtml (in Russian).
3. Saaty, T. Scaling Method for Priorities in Hierarchical Structures. J. of Mathematical Psychology 3, Vol. 15 (1977) 234–281.
4. Zahedi, F.M. A method for quantitative evaluation of expert systems. European Journal of Operational Research 48 (1990) 136–147.
5. Zhelezko, B.A., Morozevitch, A.N. Information-analytical decision support systems. Minsk, NIU (1999) (in Russian).
6. Zhelezko, B.A., Siniavskaya, O.A., Ahrameiko, A.A., Berbasova, N.Y. Methodology of the estimation of quality of objects with complex structure under conditions of non-stochastic uncertainty. International Conference on Fuzzy Sets and Soft Computing in Economics and Finance (FSSCEF 2004): Proceedings, Saint-Petersburg. Vol. 2 (2004) 360 – 367.

Spyware: Business Impact Beyond of a Technical View

Luis Villa[1], Manuel Romero-Salcedo[1], Mario Ramos[1], Felipe Davalos[1], Jair García-Lamont[2]

[1] Mexican Petroleum Institute, Program of Research in Applied Mathematics and Computing, Eje Central Lázaro Cárdenas 152, Col. San Bartolo Atepehuacan, 07730 Mexico City, Mexico
{lvilla, mromeros, mario, fdavalos}@imp.mx
http://www.imp.mx

[2] National Polytechnic Institute, Computer Research Center, Av. Juan de Dios Batíz s/n, 07738 Mexico City, Mexico
{jgarcial}@cic.ipn.mx

Abstract. Security have become increasingly important issues for organizations. An increasing variety of malware, such as worms, spyware and adware, threatens both personal and business computing. Problems related to security threats normally are analyzed around the CIA (Confidentiality, Integrity and Availability) concept. However, from the financial perspective, is very import to know the relation between these security threats and their cost for the business. In this paper, we explore the area for a technical estimation of financial losses due to breaches of security, particularity due to spyware traffic. Our preliminary results shows that most of 12% traffic over the network infrastructure are used by process that are not related with services and neither related with the business.

1 Introduction

Due to the confluence of advances in processors performance, network bandwidth and accessibility prices, the internet users have seen explosive growth in both personal and business use. At the same time increasing variety of malware, such as worms, spyware and adware threatens, appeared in both personal and business computing [1, 2]. The FBI estimated that only in 2005 around $67 billion was lost to cybercriminals [3]. This report shows that viruses and spyware were the most common problems, and the most costly were viruses and worms. However, despite the negative impact only a small fraction of affected organizations report the attacks.

Previous works has focused on studying the risk of the malicious traffic for the organization analyzing the impact over the information's Confidentiality, Integrity and Availability, leaving aside a study about how this extra-traffic impact in the organization.

The availability of the network throughput depend highly in the number of users and the type of resource requested by the services (internal and external). From the

IT best practices framework (ITIL – IT Infrastructure Library) perspective [9], the Capacity Management ensures that appropriate capacity is available at all times to meet the requirements of the business, and it is directly related to the business services (internal and externals). At the same time this process is responsible for planning future business requirements. Normally this plan is constructed by using the existing data on the current resource utilization, in order to predict future requirements, which might are the result of planning improve, growth or new services. This information is used for the Financial Management process in order to planning the business budget.

A very adjusted planning, could cause that the increment of the traffic produced by spyware, and in this way could cause a reduction in the quality of the services offered through the network infrastructure. In this way it is important to know precisely the traffic that it can be supported.

In this paper, we study the characteristics of network traffic in order to quantify the extra traffic produced by malicious threats. We develop quantitative estimates of extra-traffic which can be used to estimate a financial losses due to breaches of security, particularity due to spyware traffic. Our main focus is on organizations and institutions with high number of Internet users. Our preliminary results shows that most of the 30% of the infrastructure related to the internet is used and wasted by process that are not related with services which are not related with the business.

The rest of this paper is organized as follows: Section 2 reviews the related work. The experimental framework and the analysis of results is presented in Section 3 and finally Section 4 summarizes the main conclusions and future work.

1.1 Understanding what is Spyware

This is a relatively new malware related with security [6]. Among their qualities we can highlight that these are programs that without authorization causing severe security exposure and risk. Such software can gain complete control of the user's system, starting whenever the user turns on the system. These programs can sniff the system for any desired data and can transmit anything to outside source [7]. Problems related with the information confidentiality and integrity and also a traffic growth are highlighted due to spyware malware.

Any system exposed to the Internet, using well-known services (a browser or e-mail), can become a spyware victim. Symptoms might appear, depending on which form of spyware you have encountered [8].

2 Related Work

Several researchers have previously analyzed the traffic generated by malicious threats [5]. Most of them are focused in DDoS like attack detection and response of worm traffic propagation. In this section we provide a brief overview of these research works. In [12] the effect of malicious traffic on the background traffic is analyzed. The result shows that an increase in the average DNS latency by 230% and an

increase in the average web latency by 30%. An analysis of popular on-line games traffic was developed. Reported results show that the traffic behavior of these servers is highly predictable. In our study we can predict that it is possible to filter the information related to games from traffic produced by spyware.

Most of work achieved about traffic analysis is related with the risk analysis in order of prevent a contingence. That is, the effort is concentrated to ensure the systems operation. Different techniques have been tackled using fuzzy like techniques and association rules to detect anomalous pattern [4], [10], [11].

3 Experimental Framawork and Results

We have implemented an ETL (Extract, Translate, Load) system. The complete network traffic was captured and analyzed. We collect traces due to well-known services and applications (see Table 1).

We analyzed a complete day and a sampling frequency of one hour. On average there were 3000 users using the network infrastructure.

Table 1. Logs analyzed by de ETL System. Each log is processed and then is integrated into a matrix in order of analyze the flow of information in the net

Service	Description
ftp	File transfer protocol
Apache	Apache web server
IIS	Internet information server
Procmail	E-mail filter
Sendmail	Internet e-mail server

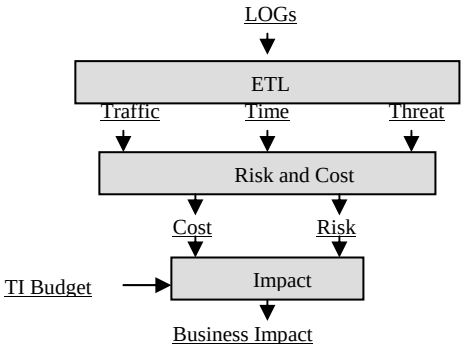


Fig. 1. The three main phases of the traffic analysis system proposed to detecting risk and quantify the business cost impact.

Figure 1 shows the main architecture of the system used in our analysis. This system is organized in three main modules. The ETL to extract de information from the logs, the next one where the risk and cost are quantified and finally the module where

the business impact is estimated. The design of the ETL system is completed and the results are reported in this paper. The other two modules are still under construction.

3.1 Experiments results

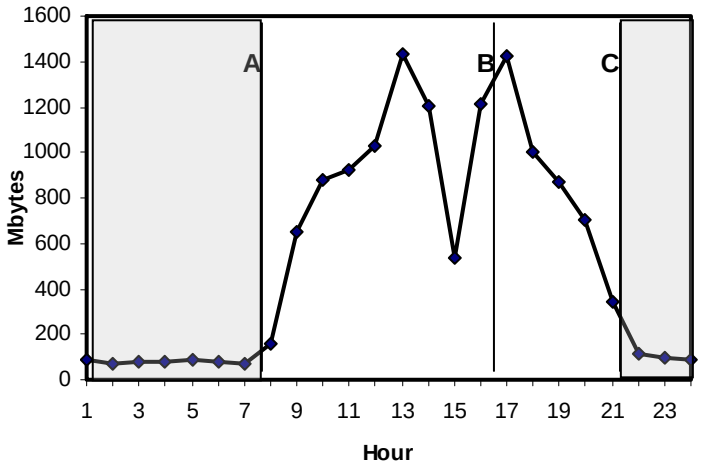


Fig. 2. Traffic distribution (*solid line*) along complete day, from 00:00hrs to 23:59 hrs. This clearly shows network usage in: office hour (from A to B) and with gray color the network use when we expect very low traffic

The figure 2 shows the different activity periods in the organization. From 8:00hrs. to 17:00hrs. the office hour is plenty identified (with a lunch break at 15:00 hrs). The important point in our study is analyze the traffic in the no-office hour window (depicted with gray).

We can notice that although traffic has lessened in the no-office hour, represent more than 12% of the day analyzed. This traffic is the results of the contribution of different factors: users, backups and anomalous traffic. A deeper analysis reveals that 80% correspond to a traffic produced by spyware threats. These results also show that most of the flow is output traffic, that means that the probability to loose information is high and at the same time the confidentiality is low.

The results are promising approach and confirm that our hypothesis is in force to explore the area for a technical estimation of financial losses due to breaches of security.

4 Conclusions and Future Work

In this paper we have explored the idea for a technical estimation of financial losses due to breaches of security, particularity due to spyware traffic. The results presented in this paper shows that most of the 12% of the infrastructure related to the network is used and wasted by process that are not related with services which are not related with the business, and 80% of this anomalous flow of information correspond to a traffic produced by spyware threats.

The results confirm our hypothesis about the importance to explore the area for a technical estimation of financial losses due to breaches of security. This work motivates the need to study more closely about the infrastructure usage for malicious threats.

At the present time we are simulating related schemes described in this paper to provide comparative performance information between the analysed proposals.

References

1. S. Staniford, V. Paxson, and N. Weaver: How to own the internet in your spare time. In: Proceedings of the 11th Usenix Security Symposium, (2002)
2. Symantec: Symantec Internet Security Threat Report. http://enterprisesecurity.symantec.com, (2004)
3. Prachi Patel Predo: Cybercrime at a glance Spectrum, Vol 43. IEEE, (2006) 17
4. Dong-Mei Zhao; Jing-Hong Wang; Jing Wu; Jian-Feng Ma. : Using fuzzy logic and entropy theory to risk assessment of the information security. In: Machine Learning and Cybernetics, Proceedings of 2005 International Conference on. Vol 4. (2005) 2448 - 2453
5. Zhao Dongmei: Comprehensive Risk Assessment of the Network Security. Computer Science, Vol 31. IEEE (2004) 66-69
6. Ames, W. : Understanding spyware: risk and response. IT Professional, Vol. 6. (2004) 25 - 29
7. E.W.; Halderman, J.A. : Digital Rights Management, Spyware, and Security Felten. Security & Privacy Magazine, Vol. 4. IEEE (2006) 18 - 23
8. McFedries, P.: Technically Speaking: The Spyware Nightmare. Spectrum. IEEE Vol. 42. (2005) 72 - 72
9. OGC: "ITIL-Service Delivery". Office of Government. The Stationary Office: London. (2002)
10. Verykios, V.S., Elmagarmid, A.K., Bertino, E., Saygin, Y., Dasseni, E.: Association rule hiding. Transactions on Knowledge and Data Engineering, Vol. 16. IEEE (2004) 434 – 447
11. S Harmantzis, F.; Malek, M.: Security risk analysis and evaluation. International Conference Communications, Vol. 4. IEEE (2004) 1897 - 1901
12. Wu-chang Feng, Chang, F., Wu-chi Feng, Walpole, J.: A traffic characterization of popular on-line games. Transactions on Networking, Vol. 13. IEEE/ACM (2005) 488 - 500

A Groupware System for Document Collaboration in Business Management

Manuel Romero-Salcedo[1], Luis Villa[1], Jair García-Lamont[2]

[1] Mexican Petroleum Institute, Program on Applied Mathematics and Computing, Eje Central Lázaro Cárdenas 152, 07730 San Bartolo Atepehuacan, Álvaro Obregón, Mexico, D.F.
{mromeros, lvilla}@imp.mx
http://www.imp.mx
[2] Instituto Politécnico Nacional, Computer Research Center, Av. Juan de Dios Bátiz s/n, 07738, Mexico, D.F.
jgarcial@cic.ipn.mx
http://www.cic.ipn.mx

Abstract. A lot of the work that takes place in complex production environments consists of the development, revision, exchange and sharing of contracts, proposals, projects, technical papers and reports, legal procedures and other business documents. In fact, document collaboration is probably one of the most difficult aspects of business management. Being able to share information in a way that is effective, precise and understandable is a key issue. Groupware technology is part of the answer. Effective use of groupware systems for document collaboration would greatly help business' management process and it would provide massive benefits in complex industries. In this paper, we present ELXI, a groupware tool to allow people work together, even when co-authors are in different locations geographically dispersed, sharing common business documents. ELXI was tested on a real complex production environment. We report on the results of a users' satisfaction study.

1 Introduction

In the last decades, many industries have transformed their work processes and their business practices. They have passed of being monolithic organizations and centralized into much more structured organizations, globally distributed and with complex production environments. Resource distribution entails new challenges in the way how people work together, generating knowledge and breaking the boundaries of time and space. In fact, most of the knowledge that is generated collectively in an organization is stored in documents. In this way, documents constitute one of the most valuable assets. Business Management in organizations and complex production environments require managing contents of business documents which, depending on the kind of organization processes and their information, could be voluminous and complex.

On the other hand, there exist some guides, formats and regulations that need to be applied and/or to be followed when business documents are developed. Moreover, it is important and necessary to take into account the time constrains.

Document writing is a long and complex task [1], [3], [4], [5]. Many authors decide to join their efforts in order to reduce time during document production; to improve quality of documents; to obtain points of view from their co-authors; to increase productivity; etc. Until now, several research works and formal studies have been conducted into the way groups write together [3], [4], [5]. These research works have been carried out taking into account studies of cases, experiments, interviews and observations of real situations of collaborative writing. Surveys on writing have shown that a large number of documents are produced through a collaborative effort [6].

The proliferation of interdisciplinary studies, international projects and distributed work groups within large and complex industries, has led to pressure on writers to work in collaboration. Writing groups may consist of people who rarely meet face-to-face, yet they are expected to work closely together and to tight schedules. In complex industries, for developing common documents [5] is necessary to put effort into the way how groups of specialists and experts can be organized for accomplishing joint writing tasks. In a business documents, each one of the participants plays a specific role (reviewer, publisher, reader, writer, manager, editor, etc.).

An important problem in business management as well as in several scientific and technical areas is to know the best way to develop, produce and share common documents. For tackling this problem, we developed a groupware system designed to support collaborative writing on business documents.

The rest of this paper is organized as follows. Section 2 describes business documents in complex production environments. Section 3 describes a groupware system for document collaboration in business management. Evaluation results are reported in Section 4 and finally, Section 5 presents conclusions and future work.

2 Business Documents in Complex Industries

In complex industries, teams of experts and specialists are dispersed throughout the organization and they require working together on common business documents in order to accomplish their productive activities. As an example, in PEMEX, the Mexican National Oil Company, the Standardization Committee and Subsidiaries Organisms, as well as industries and technical consultants, join their efforts to work together to develop technical documents. Currently, productivity of different experts and specialists, working on a Technical Standardization, is limited because of problems caused by difficulty of integrating different revisions; necessity of attending meetings in order to take decisions; difficulty of following the evolution of a document; tracking and execution of delivery dates; etc.

Most of the business documents have to be elaborated with the joint participation of different experts. Examples of these documents are: administrative and technique standardizations, alignments, procedures, projects' protocols, projects' logs, business and strategic planning, technical and research reports, scientific articles, products' specifications, legal and administrative regulations, etc. Experts in PEMEX, as well as in many other complex production environments, develop their business documents, using standard word processor systems, which are designed for individual

editing. Nowadays, these word processors include some tools for making change control; adding comments and revisions; and producing different versions. When different versions of the same document are produced, each version usually contains comments, revisions and/or corrections from different experts and specialists. Consolidation of all these versions, with the purpose of producing a single document, is a complex task, time consuming and error prone. Then, quality of the document is always in detriment. Thus, these tools are helpful, but they do not completely solve the problem of the collaborative writing process.

3 Collaboration on Business Documents with ELXI

ELXI (**E**ditor co**L**aborativo de documentos **X**ml en **I**nternet - *Collaborative Editor of Xml Documents over the Internet)* is a groupware system designed to support multiple authors who jointly create, edit, annotate and revise shared business documents. With ELXI, authors can improve the efficiency and quality of group writing and support the collaborative writing process among distributed authors. A document is defined as a strongly structured entity, which can be divided it into several fragments. On each fragment, several co-authors can apply different editing actions, according to their editing roles. Authors can also track evolution of the same document, playing different roles on different fragments. Thus, a document is a collaboration space through which co-authors can join their efforts, communicate, capture and organize their ideas and produce their contributions (see Fig. 1).

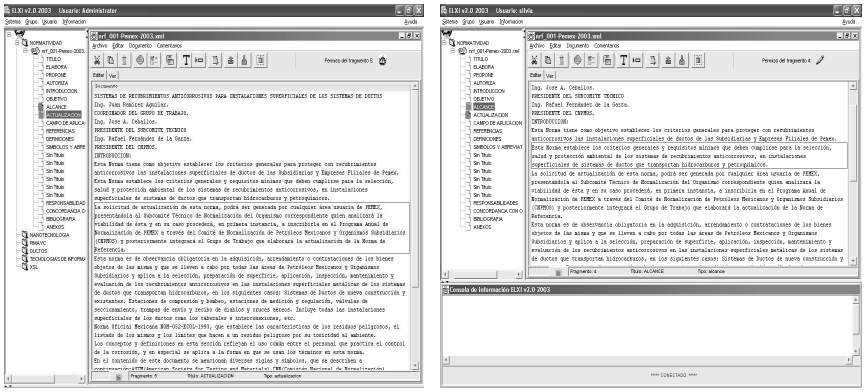


Fig. 1. Two different authors collaborating on the same business document

ELXI has a component-based client-server architecture (see Fig. 2) in which its components implement mechanisms for *document management* (replication, consistency, concurrency control), *shared workspace management*, *group awareness*, and *communication and interaction* through a facilitator. Business documents are produced in XML (an industrial standard and universal document format for electronic documentation) and are managed as objects, which allows to access and store these objects efficiently. When documents are created, they are firstly stored in the server file system. Then, a copy is also stored in the co-authors' file systems.

The document management component guarantees, in any moment, document consistency. The facilitator component allows communication and interaction among co-authors. The shared workspace provides a common space (in this case is the document itself) where co-authors are able to make their contributions. The group awareness component allows to any author to perceive and use everything that is in the shared workspace, without contacting other users or leaving the system. This is perhaps one of the most important components because all information that is transmitted includes knowledge about other co-authors' locations, states, actions and abilities as well as the shared document state.

Finally, the *ELXI Server* is in charge of controlling all document editing actions performed by co-authors. These editing actions are remotely accessed by all co-authors (*ELXI clients*). ELXI allows creating business documents through the Internet, so co-authors can be spread out over different geographical locations. ELXI guarantees that different co-authors can make their contributions on the same document, at the same time or differed times and from different places without consistency problems.

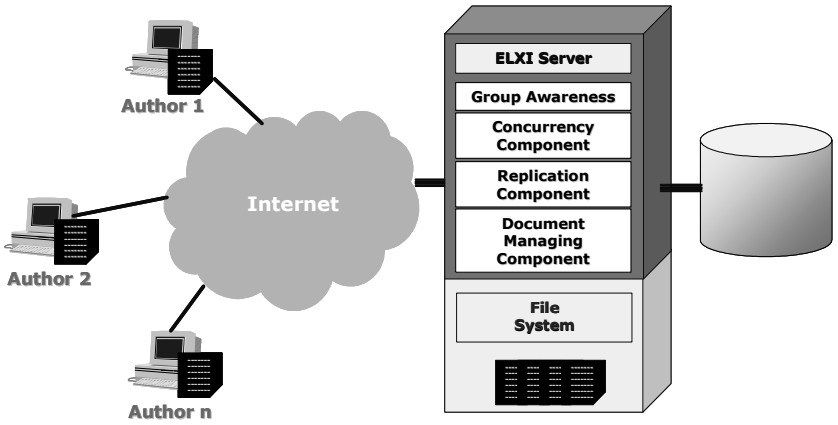


Fig. 2. Component-based Client-Server Architecture for ELXI

In ELXI, the structure of any business document is defined in XML[1] (*eXtensible Markup Language*). Then, business documents in ELXI can easily be exploited, processed and exchanged using different standard and commercial applications and tools (Web Browsers, Content Management Systems, Document Management Systems, Office applications, Search Engines, etc). In any moment, co-authors can take a business document and convert it to another format (such as HTML, RFT or PDF). ELXI implements technologies associated to XML, such as: XSL (*eXtensible Stylesheet Language*), XSLT (*eXtensible Stylesheet Language Tranformations*) and XSL-FO (*XSL Formatting Objects*). Fig. 3 shows a business document which resulted from a collaborative writing process using ELXI. It was exported to PDF and RTF formats and then visualized with Adobe Acrobat Reader and MS Word, respectively.

[1] http://www.w3.org/XML

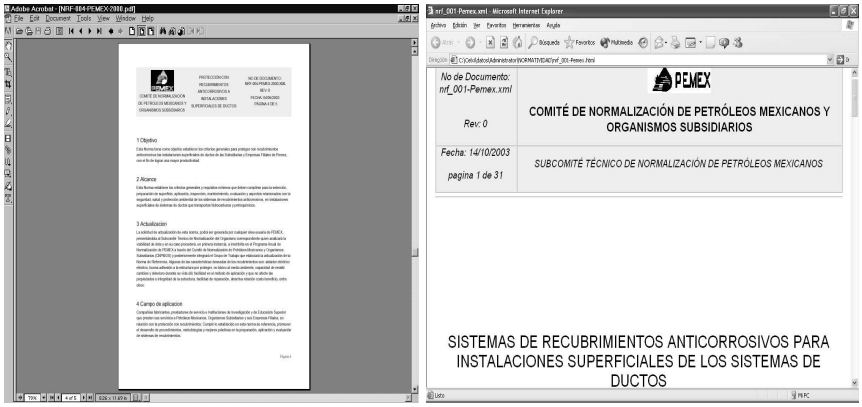


Fig. 3. A business document exported to PDF format and RTF formats and visualized with Adobe Acrobat Reader and MS Word

ELXI provides the following features:

- *Collaborating strategy*. Multiple dispersed authors can access (edit, consult) the same shared document.
- *Basic text editing*. Each author can create documents using the editing functionalities available (text, lists, tables, images, footnotes, headers, etc.). Moreover, authors are able to make searches and replacements within any fragment or the whole document.
- *Managing authors and groups of authors*. A manager defines a list of authors and he/she can define work groups. Then, collaboration can be carried out among the work group members. When an author creates a document, he/she becomes a manager of that document. Then, he/she is able to invite new authors to collaborate on the document.
- *Access control to the collaborative writing environment.* Only authorized authors can participate in the collaborative writing process. Authors log to the system using a login and a password.
- *Document division into different fragments*. Following the inherent structure of a document, it can be divided into different fragments. A fragment is the minimum unit of collaboration. It can be any element inside the structure of a document (a section, a subsection, a paragraph, a list, etc.).
- *Access control to the document*. An author playing the manager role is in charge of assigning editing roles to each author who is going to participate in the collaborative writing process. The editing roles available are: reader, writer and manager. The first two roles allow an author to read and write, respectively, on a fragment. The manager is the author who orchestrates and tracks document evolution. A manager is always able to modify the editing roles of other authors. He/she is also authorized to add new fragments to the document or delete them. If necessary, a manager can delegate his/her role to another author. Taking into account that collaborative writing is a dynamic process, changes of editing roles can take place in a dynamic way.

- *Asynchronous and synchronous document interaction.* Different co-authors, spread out over different locations, can work together on the same document, at the same time (real time) or they can work at different times (in differed times).
- *Messages among authors*. Any author can send(receive) messages to(from) another author (suggestions, invitations to join collaborative writing, comments, problems, etc.). An author can also send a message to all the work group members.
- *Production (creation and viewing) of business documents*. The editing core allows generating structured documents in the standard and industrial XML format. It includes all necessary edition functionalities to allow producing business documents as well as technical and/or scientific documentation. Stylesheets are used for viewing and formatting documents. Then, authors do not be worried about formatting styles and they can only be concentrated in document production.
- *Document exportation*. Business documents can be exported, in any moment, to standard formats, such as PDF, RTF and HTML. Then, authors can use corresponding commercial applications to apply final or professional formatting styles to the document.
- *Group awareness support.* Thanks to different group awareness elements, an author can be aware of any editing action (modifications to the document or fragments, roles changes, new documents, new fragments, new work groups, etc.) made by any co-author. An author can also know who is connected to the system; who is working on a given fragment; which editing role is being played; etc.
- *Document consistency support.* Any editing action (writing, update, modification, etc.) on a document must be consistent with all editing actions made from all other co-authors working on the same document. A document consistency mechanism guarantees that only one author at a time can edit or modify a fragment.
- *Annotations support*. Authors can attach electronic notes to document fragments or to the whole document. A record of notes can be done by attaching a note to a note. Using annotations on documents or fragments, authors can express comments, ideas, revisions, etc. Co-authors can decide to make public or private their electronic notes. Each author is able to manage his/her own annotations (make chances; add new notes or delete notes).
- *Distributed component-based architecture*. ELXI implements a client-server model. Four main components constitute the core of the whole system: *document management*; *shared workspace management*; *group awareness control*; and *communication and interaction.*
- *Run on Windows and/or UNIX (Linux, Solaris) platforms*. *ELXI Clients* and the *ELXI Server* can be executed in any available computing platform.
- *Collaboration through Internet and/or Intranet*. All authors are able to keep communicated and collaborate through a common document. They can be working on their own workstations connected to Internet and/or Intranet. Co-authors can be close one from another and/or they can be geographically dispersed over the world.

4 Evaluation

ELXI was tested by a group of users in order to evaluate its functionalities and impact. Evaluation was carried out by 14 users. All users received training on the use of ELXI. All users worked on a specific business document. When document was finished, all users were asked to fill out an evaluation questionnaire. This questionnaire consisted of ten questions arranged in three chapters: *document collaboration* (questions 1.1, 1.2, 1.3 and 1.4); *document viewing* (questions 2.1, 2.2 and 2.3); and *user interface* (questions 3.1 and 3.2). Questions were scored according to the following scales of values: *Very well*; *Well*; *Sufficient*; *Bad*. For each one of the questions required, Fig. 3 shows the results obtained in a grouped bar graph and in table form.

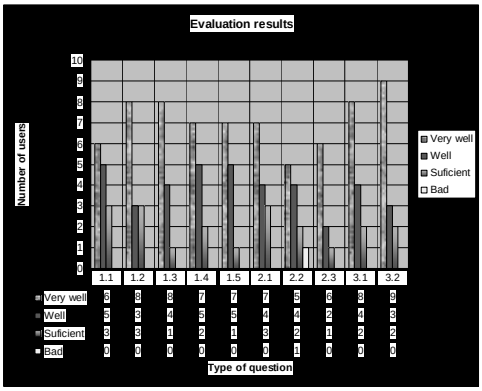


Fig. 3. Users' satisfaction evaluation results

From Fig. 3, it can be seen that, questions were scored in the highest ranges. This fact shows that ELXI had very good acceptance from users. For the cases where the score was *Sufficient*, we carefully read all answers and comments in order to analyze them. In almost all the cases, this score was given because they considered that functionalities for editing documents were the minimum required. They do not make any comments about problems with other functionalities of the system. Thus, ELXI was satisfactorily accepted by all users. Concerning user interface, users proposed to improve some interfaces for editing basic elements such as: tables, lists and images. Concerning document collaboration, users proposed to add new tools for helping users to be aware of all editing changes. Finally, concerning document viewing, users proposed to enhance document quality, adding other formatting features (centering, type fonts, underlined text, text colors, etc.).

5 Conclusion and future work

Until now, several groupware systems for collaborative document have been developed. However, many of these systems provide inadequate support for the collaborative writing process. They mainly focus on implemented-oriented problems (such as: concurrency control, network failures, data sharing, conflict management, access control) while overlooking very important social issues (such as: interaction protocols, cognitive models of collaborative writing, human factors, conflict management). ELXI takes into account both technological and social issues, proposing an efficient solution to the problem of collaboration on business documents. Authors using ELXI will find the following benefits: promotes dynamic participation of work groups developing the same business document; improves quality of documents, making more efficient their production; reduces costs while increasing productivity; reduces time allowing consolidation of different revisions (authors are not worried about multiple versions), opinions, comments and/or co-authors' arguments; compatible with multiple current applications, producing XML documents. Thus, using ELXI, experts and specialists will be able to share and collaborate on business documents in a coordinated and consistent way, avoiding separated and individual writing phases, which they will pay in important efforts of communication and agreement. With our users' satisfaction evaluation study, we could collected several comments and suggestions. As part of our future work, we plan to evaluate all these users' suggestions. Then, we will convert suggestions in user requirements and we will develop firstly those which will contribute with a valuable impact in users.

References

1. Baecker, R..: A Survey of Experiences of Collaborative Writing, Computer Supported Collaborative Writing. M. Sharples (ed.), Springer-Verlag (1993) 87-112
2. Baecker, R.., Belloti, V.,: Informed Oportunism as Strategy: Supported Coordination in Distributed Collaborative Writing, Proceedings of the Third European Conference on Computer Supported Collaborative Work, Kluwer Academic Publishers, Londres, (1993) 233-248
3. Ede, L., Lunsford, A.: Singular Texts/Plural Authors, Carbondale Southern University Press, Southern Illinois University Press, Estados Unidos (1990)
4. Flower, L., Hayes, J. R..: A Cognitive Process Theory of Writing, College Composition and Communication, Vol. 32, (1978) 365-387
5. Lay, M. M., Karis, W. M..: Collaborative Writing in Industry: Investigations in Theory and Practice, Baywood Publishing Company, Amityville (1991)
6. Mendoza Chapa, S., Romero Salcedo, M.., Oktaba, H.: Group Awareness Support in Collaborative Writing Systems, Sixth International Workshop on Groupware (CRIGW'2000), IEEE Computer Society Press, Portugal (2000), 112-118
Schlichter, J., Koch, M., Bürger, M.: Workspace Awareness for Distributed Teams, Workshop Coordination Technology for Collaborative Applications, W. Conen (ed.), Lecture Notes on Computer Science, Springer-Verlag (1997)

Modeling Volatility of Time Series Using Fuzzy GARCH Models

Alexander Popov[1], Konstantin Bykhanov[2]

[1] Novosibirsk State Technical University, Department of Applied Mathematics and Informatics, Karl Marx str. 30, 630092 Novosibirsk, Russia
[2] Novosibirsk State Technical University, Department of Applied Mathematics and Informatics, Karl Marx str. 30, 630092 Novosibirsk, Russia
kvib@ngs.ru

Abstract. Fuzzy Modeling is an effective way of construction models for complex dynamic systems. Here we present a new application of fuzzy rule-based models to analysis of discrete time series. Fuzzy generalization of autoregressive conditional heteroscedasticity (ARCH/GARCH) models is proposed and technology of Fuzzy GARCH modeling is basically overviewed. A new method for parametrical identification of fuzzy rule-based models is presented. A comparison with usual GARCH models is made both for modeled and real time series.

1 Introduction

Fast development of risk and uncertainty concepts in economics and finances necessitated a creation of such new mathematical models that could describe a dynamic behavior for both first- and higher-order conditional moments. Thereupon a "volatility" term is often mentioned by researchers. Volatility is a relative rate at which the price or other financial asset moves up and down, i.e. a characteristic of its changeability and unsteadiness. Usually volatility is found as a conditional variance of historical observations on price or other financial asset.

Volatility models appeared in early 80s of last century [3] and have been improving for all last years. Different types of models were introduced, both data- and parameter-driven (so called state-space models). In the class of data-driven models the biggest attention was expectedly paid to linear models. Therefore, in the beginning of 90s wants for new nonlinear volatility models were noticed [4]. Many real financial processes have a complex behavior of their conditional variance along with non-trivial dynamics of conditional expectation. For example a leverage effect is often take place, i.e. negative correlation between past asset returns and current volatility values. Some of such effects can be captured by linear models modification, but others demand nonlinear approach. Unfortunately, due to complexity, nonlinear models are still of a very limited use.

In this paper we propose another class of autoregressive heteroskedasticity models based on constructing a system of Takagi-Sugeno fuzzy rules – a class of Fuzzy GARCH models. Methods of fuzzy modeling have already showed themselves as a good technique for description of complex dynamic systems [2]. Combining easiness and convenience of linear models with ability to capture complex system correlations (which is commonly typical for nonlinear models), fuzzy models therefore could be a promising approach to analyzing dynamic processes with time-varying variance.

This paper is organized as follows:

Section 2 contains basic information about conventional ARCH/GARCH models.

Section 3 presents a class of Fuzzy GARCH models as a fuzzy generalization of conventional GARCH. Main aspects of construction of such models are also affected.

In section 4 we give a comparison between conventional and fuzzy GARCH models on several tests. An example of option-pricing models for real-based financial time series is presented to show the advantages of proposed class of models. Finally we make some conclusions about the work.

The preparation of manuscripts which are to be reproduced by photo-offset requires special care. Papers submitted in a technically unsuitable form will be returned for retyping, or canceled if the volume cannot otherwise be finished on time.

2 ARCH/GARCH Models

Before describing the new class of Fuzzy GARCH models it is necessary to give some basic definitions and concepts of conventional ARCH/GARCH models theory.

Definition 1. Let x(t), t∈Z – is a discrete random process and

$$E\left[x(t)\middle|I_{t-1}\right]=0, \tag{1}$$

$$E\left[x^2(t)\middle|I_{t-1}\right]=g(I_{t-1}), \tag{2}$$

where $I_{t-1}=\{x(t-1),x(t-2),...\}$ – previous process values, E denotes an expectation operator and $g(I_{t-1})$ is a positive function of process prehistory. Then x(t) is a process with autoregressive conditional heteroskedasticity.

Definition 2. Let x(t), t∈Z – is a discrete random process with autoregressive conditional heteroskedasticity. Then x(t) is said to be an ARCH(p) process if

$$x(t)\middle|I_{t-1}\in Norm[0,d^2(t)], \tag{3}$$

$$d^2(t)=E\left[x^2(t)\middle|I_{t-1}\right]=a_0+\sum_{i=1}^{p}a_i x^2(t-i). \tag{4}$$

In right part of equation (4) all coefficients are non-negative: $a_0>0$, $a_1\geq 0,\ldots, a_p\geq 0$.

ARCH models were introduced by R. Engle in 1982 [3]. Analysis of different financial time series (for example, inflation rate in Great Britain in 60s-70s) let him to find out an interesting regularity: big and small values were often grouped into alternate series (so-called volatility clusters). As the stationarity and homoskedasticity

hypotheses still could not be discarded for the whole amount of data, this phenomenon was unexplainable, so new class of ARCH models was brought in.

Certain ARCH(p) process has following essential properties:

1. It is not autocorrelated, so it is weakly stationary if its unconditional variance is finite, i.e. $a_1+a_2+\ldots+a_p<1$.
2. Its unconditional distribution has a higher kurtosis in comparison with normal one. This fact meets well with observations of real financial time series, which are often characterized by fat-tailed distributions.
3. Squared values of ARCH(p) process follows a non-Gaussian autoregressive process AR(p).

A generalization of ARCH(p) models are GARCH(p,q) models [4]. For GARCH(p,q) models modified type of $g(I_{t-1})$ function is used.

Definition 3. Let $x(t)$, $t\in Z$ – is a discrete random process such that equations (1),(3) are valid and

$$d^2(t)=a_0+\sum_{i=1}^{p}a_i x^2(t-i)+\sum_{j=1}^{q}b_i d^2(t-j)\cdot \qquad (5)$$

In right part of equation (5) all coefficients are non-negative: $a_0>0$, $a_1\geq 0,\ldots, a_p\geq 0$, $b_1\geq 0,\ldots, b_q\geq 0$. Then $x(t)$ is called a process with generalized autoregressive conditional heteroskedasticity (or GARCH(p,q) process).

Properties of certain GARCH(p,q) process are basically similar to properties of ARCH(p). Squared values of GARCH(p,q) processes follow an ARMA($max(p,q),q$) process. Existence requirement for unconditional variance of GARCH(p,q) process looks like: $a_1+a_2+\ldots+a_p+b_1+\ldots+b_q<1$. Kurtosis rate for GARCH(p,q) processes is also higher in comparison with normal cases.

Using GARCH models allowed to obtain relatively exact prognoses for different financial data, therefore conventional GARCH models were the most popular instrument for analyzing time series with time-varying conditional variance [4]. However, in some cases an adequate match between real data and certain GARCH model could not be reached. Today two main drawbacks of GARCH(p,q) models are mentioned:

1. Economist's researches proved [5] that assert returns often have negative (asymmetric) correlation with changes of volatility (leverage effect). Conventional GARCH models cannot capture such correlation, because its conditional variance depends only on absolute values of former observations but not on their sign.
2. Recent researches showed [6] that GARCH models often exaggerate the length of volatility clusters as compared to actual data of financial markets. This is especially noticeable when the market comes out of a recession. The reason of such inadequacy of GARCH models is believed to be the presence of latent non-linearity in studied processes. Ignoring of them can often lead to bad prognostic characteristics of used models.

This work is dedicated to Fuzzy GARCH models that have already showed themselves as a good technique for description of complex dynamic systems. Being a fuzzy generalization of conventional GARCH(p,q) models described above they can combine easiness and convenience of linear models with an ability to capture complex system correlations. Fuzzy models therefore seem to be a promising approach to analyzing dynamic processes with time-varying variance.

3 Fuzzy GARCH Models

Let {*X(1), X(2), … X(N)*} is a time series, produced by a random process *x(t)* with autoregressive conditional heteroskedasticity with variance function $g(I_{t-1})$ defined as Takagi-Sugeno rule system [1]:

$$
\begin{aligned}
&R_1 : \textit{if } q_1(I_{t-1}) \in A_1^1, q_2(I_{t-1}) \in A_2^1, ..., q_r(I_{t-1}) \in A_r^1 \\
&\quad \textit{then } d^2(t) = g^1 = g^1(t, I_{t-1}, s^1) \\
&R_2 : \textit{if } q_1(I_{t-1}) \in A_1^2, q_2(I_{t-1}) \in A_2^2, ..., q_r(I_{t-1}) \in A_r^2 \\
&\quad \textit{then } d^2(t) = g^2 = g^2(t, I_{t-1}, s^2) \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&R_K : \textit{if } q_1(I_{t-1}) \in A_1^K, q_2(I_{t-1}) \in A_2^K, ..., q_r(I_{t-1}) \in A_r^K \\
&\quad \textit{then } d^2(t) = g^K = g^K(t, I_{t-1}, s^K)
\end{aligned}
\tag{6}
$$

In (6) $A_j^i, i = 1,2,...,K$ are fuzzy sets, $q_2(I_{t-1}),...,q_r(I_{t-1})$, $j = 1,2,...,r$ – explanatory variables, $g^i(t, I_{t-1}, s^i)$, i = 1,2,…K – local rule outputs, i.e. positive functions and s^1,s^2, … ,s^K – are unknown parameters of model.

Definition 4. Fuzzy model defined by (1), (3) and fuzzy rule system (6) is called fuzzy model with autoregressive conditional heteroskedasticity or Fuzzy GARCH model.

The identification problem for the class of introduced Fuzzy GARCH models can be split into several steps.

At the first step the expediency of using GARCH models is studied. The data to be analyzed are tested for presence of any effects of conditional heteroskedasticity. There are many methods of such testing, for example Lagrangian coefficients method or Box-Ljung statistic test.

Second step is dedicated to structural identification of model. In other words, the structure of rule antecedents and consequents is determined. This step can also be divided into five separate problems:

1. Definition of explanatory variables set

It is hard to give any universal recommendations for choosing the set of explanatory variables for fuzzy rule system (6). In our work we used natural definition for them, i.e. previous values of studied time series:

$$\{q_i(I_{t-1}) = x(t-u_i)\},\ i=1, 2,\ldots,r. \tag{7}$$

Explanatory variables (7) let the fuzzy model capture the leverage effect (if one exists) because model output now depends on the sign of previous time series values. Our further researches showed that it is sometimes more effective to define explanatory variables as multiplication of previous values:

$$\{q_i(I_{t-1}) = x(t-u_i)\cdot x(t-v_i)\},\ i=1, 2,\ldots,r. \tag{8}$$

or use either of definitions (7) and (8). In (7) and (8) values u_1, u_2, …, u_r, v_1, v_2, …,v_r are unknown and have to be identified.

2. Definition of rule output functions

When choosing the structure of $g^i(t, I_{t-1}, s^i)$ which are in charge of local fuzzy rule outputs, we followed the natural principle of generalization of conventional class of GARCH(*p,q*) models, so the structure for it is proposed to be similar to (5):

$$g^i(t, I_{t-1}, s^i) = s_0 + \sum_{i=1}^{p} s_i x^2(t-i) + \sum_{j=1}^{q} s_{p+j}\, g(t-j, I_{t-1}, s) \tag{9}$$

3. Identification of antecedent structure

To obtain the certain structure of fuzzy rule antecedents (i.e. identify the values of $u_1, u_2, \ldots, u_r, v_1, v_2, \ldots, v_r$) we used the Takagi-Sugeno treelike algorithm as an essential. Its detailed description can be found in [2]. This algorithm belongs to class of inclusion algorithms and forms an optimal (with relation to chosen criterion) antecedent structure after the certain number of steps. Applying Takagi-Sugeno procedure, we used the
determination coefficient R2 as the performance criterion:

$$R^2 = [corr(x^2(t), G(t))]^2 \tag{10}$$

In (10) G(t) denotes the whole output of Fuzzy GARCH model. The determination coefficient (10) represents the portion of variance dynamic that could be explained by model factors. To avoid the tuning of model on specific features of current dataset (which are casual but not structural) the performance criterion must be calculated for test sample. Usually we used 10-15% of last time series values as the test sample to calculate the determination coefficient.

4. Obtaining fuzzy sets

To determine fuzzy sets $A_j^i, i = 1,2,...,K$ (that correspond to explanatory variables $q_j(I_{t-1})$, $j = 1,2,...,r$) one of fuzzy clustering methods can be used. In our researches we have used a well-known FCM ("Fuzzy c-Means") clustering procedure. Its detailed formulation can be found in [2]. Number of clusters c and fuzzy partition matrix UN×c are initialized as input parameters. Elements u_{zk} of matrix U denote the membership degree of *z*-th sample to the *k*-th cluster. They can be initialized randomly. Running the algorithm represents the optimization of distance functional. When finished, FCM algorithm returns set of cluster prototypes and its fuzzy partition matrix U. Fuzzy c-Means procedure automatically provides us with membership functions of exponential type (see Fig. 1):

$$m_{i,j}(x) = \begin{cases} \dfrac{\left[d(v_{i,j} - x)\right]^{-\frac{2}{m-1}}}{\sum_{k=1}^{c_j}\left[d(v_{i,k} - x)\right]^{-\frac{2}{m-1}}}, & if\ d(v_{i,j,}, x) > 0; \\ 1, \quad otherwise \end{cases} \tag{11}$$

In (11) $v_{i,j}$ – is the *j*-th element of *i*-th cluster prototype, and $d(v,x)$ – is an Euclidean distance function. Parameter m represents the fuzziness of cluster. It is also possible to use triangular or trapezoidal membership functions with obtained cluster prototypes as centers.

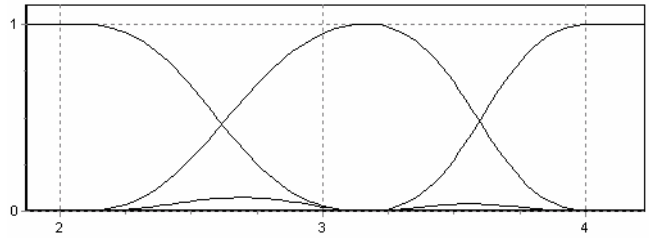


Fig.1. Example of exponential membership functions

5. Generating fuzzy rules

Thus, FCM algorithm allows us to form a collection of fuzzy sets $\{A_i^1, A_i^2, \ldots, A_i^{c(i)}\}$, covering the definitional domain for the *i*-th explanatory variable $q_j(I_{t-1})$, $i = 1,2,\ldots,r$. Fuzzy rules are generated by full enumeration of all possible memberships of i-th explanatory variable to j-th fuzzy set A_i^j. The whole number of rules is $K = c(1) \times c(2) \times \ldots \times c(r)$. The possibility value of *k*-th rule evaluates as:

$$\begin{aligned} m_k &= m_k(q_1(I_{t-1}), q_2(I_{t-1}), ...q_r(I_{t-1})) = \\ &= m_{k,1}(q_1) \times m_{k,1}(q_2) \times ... \times m_{k,r}(q_r), \end{aligned} \tag{12}$$

where $m_{k,j}(q_j)$ is a membership degree of q_j to fuzzy set $A_j^{l(i)}$ (function l(i) depends on the order of fuzzy sets enumeration).

The third step of common identification problem is dedicated to parameterization of fuzzy rules system. Since the structure of rule antecedents is presumed to be defined in the previous step, the consequent coefficients are to be evaluated now. The structure of local rule output is determined by (9).

Usually fuzzy rule systems are based on ARX-models (see Ref. [1]), therefore linear least-squares method is mainly used for estimation of fuzzy models. Unfortunately, local output (9) is not strictly linear, because it depends on past outputs of the rule, $g^i(t-i, I_{t-1}, s^i)$, i = 1,2,…,q that are unknown. Although is possible to use non-linear least-squares method, we propose a new method for estimation of unknown consequent parameters – *weighted* maximum-likelihood method.

Before using maximum-likelihood method it is necessary to make an assumption about the distribution law of studied process $x(t)$. As we assume that statement (3) is valid, the log-likelihood function looks like:

$$\ln[f(x(1), x(2),...x(N))] =$$
$$\ln[f(x(1),x(2),..x(m))\cdot f(x(p+1)\,|\,I_{m-1})\cdot ...\cdot f(x(N)\,|\,I_{N-1})], \qquad (13)$$

where $m = \max(p,q)$. Although distribution law for $f(x(1),...x(m))$ is unknown, this factor has minor contribution to the log-likelihood function and can be neglected. Therefore the L' likelihood function is usually considered:

$$L' = \ln[\, f(x(1),\, x(2),... x(m)) \cdot$$
$$\cdot f(x(m+1)\,|\,I_m)\cdot ...\cdot f(x(N)\,|\,I_{N-1})] =$$
$$= \ln\left[\prod_{t=1}^{N} f(x(t), d(t)\big|I_{t-1})\right] = \ln\left[\prod_{t=1}^{N} \frac{1}{\sqrt{2\cdot\pi}\cdot d(t)} \exp(-\frac{x^2(t)}{2d^2(t)})\right] =$$
$$= \sum_{t=1}^{N} \ln\left[\frac{1}{\sqrt{2\pi}d(t)} \exp(-\frac{x^2(t)}{2d^2(t)})\right] =$$
$$= -\frac{n}{2}\cdot \ln\pi + \sum_{t=1}^{N}\left[-\frac{1}{2}\cdot \ln d^2(t) - \frac{1}{2}\cdot\frac{x^2(t)}{d^2(t)}\right]. \qquad (14)$$

Function L' is to be maximized. Constant component $-\frac{n}{2}\cdot\ln\pi$ can also be ignored. Denoting

$$\tilde{L} = \sum_{t=1}^{N}\left[-\frac{1}{2}\cdot \ln d^2(t) - \frac{1}{2}\cdot\frac{x^2(t)}{d^2(t)}\right],$$

we have now the problem of optimization for the function $\tilde{L}$.

The contribution of t-th observation into likelihood function $\tilde{L}$ is equal to:

$$\tilde{L}_t = -\frac{1}{2}\cdot \ln d^2(t) - \frac{1}{2}\cdot\frac{x^2(t)}{d^2(t)} \qquad (15)$$

For estimation of local model parameters (9) it is proposed to maximize the weighted likelihood function, $\tilde{L}(W_i)$, that takes into account also membership degree of t-th observation to the i-th fuzzy rule:

$$\tilde{L}(W_i) = \sum_{t=1}^{N} m_i(t)\tilde{L}_t \ , \tag{16}$$

where $m_i(t)$ is calculated as in (12) and $W_i = diag[m_i(1),...,m_i(N)]$.

There are several methods of optimizing likelihood functions. As recent researches proved that direct and first-order methods were not always able to localize the extremum before stopping the algorithm, we used second-order methods to maximize $\tilde{L}(W_i)$. Approximation of Hesse matrix was made by using of conditional expectations as in [6].

To guarantee that local constrains of partial GARCH model are obeyed, i.e. that

$$s_0^i > 0,\ s_j^i \geq 0\ \forall\ i \in \overline{1,K},\ j \in \overline{1,p+q} \text{ and}$$

$$\sum_{j=1}^{p+q} s_j^i \leq 1\ \forall\ i \in \overline{1,K}$$

we projected a directing vector to the accessible region on every step of optimizing procedure. We also provided our procedure for the possible dimension reduction to avoid premature algorithm completion.

After finishing the parameterization of all local models it is necessary to make a defuzzification of evaluated Takagi-Sugeno rule system. We used standard (centroid) method of deffuzification that looks like

$$G = \frac{\sum_{i=1}^{K} m_i \cdot g^i(t, I_{t-1}, s^i)}{\sum_{i=1}^{K} m_i} =$$

$$= \frac{\sum_{i=1}^{K}\left[m_i \cdot (s_0^i + \sum_{j=1}^{p} s_j^i x^2(k-j) + \sum_{j}^{q} s_{p+j}^i g^i) \right]}{\sum_{i=1}^{K} m_i} \tag{17}$$

for Fuzzy GARCH(*p*,*q*) models.

Last step of any identification procedure covers the testing of the estimated model to find out whether it corresponds with the time series data or not. If results are inadequate, the researcher goes back to the second step. Otherwise, the identified model is accepted and can be used for modeling or prediction.

4 Examples of fuzzy GARCH modeling

In this section we present two examples of using Fuzzy GARCH models for analysis of time series with conditional heteroskedasticity and make a comparison between Fuzzy and conventional GARCH models.

4.1 Modeling volatility for certain nonlinear GARCH process

Let $x(t)$ is a discrete random process with a conditional heteroskedasticity and

$$E[x^2(t) \mid I_{t-1}] = d^2(t) = w + a \cdot x^2(t-1) + \exp(c - b \cdot d^2(t-1)) \qquad (18)$$

Statement (18) signifies that $x(t)$ is a non-linear GARCH process. Figure 2 shows a time series of 1000 observations generated by $x(t)$ with $w = 0.05$, $a = 0.15$, $b = 0.7$, $c = 4$.

There are many volatility clusters of different length that depends on exponential component of $x(t)$.

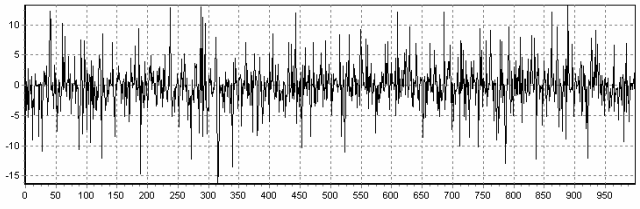


Fig.2. Generated time series for nonlinear GARCH process

Using of common GARCH models didn't give any adequate results: the value of performance criterion R^2 (see (10)) never exceeded the level of *0.02* on the test sample, formed by the last 149 observations.

Table 1. Results of modeling volatility for nonlinear GARCH process (18)

Model order		**Conventional GARCH(*p*,*q*)**	**Fuzzy GARCH(*p*,*q*)**	
p	q	R^2 value on the test sample	Number of fuzzy rules	R^2 value on the test sample
1	1	0.00002	4	0.023
2	2	0.005	24	0.068
3	3	0.018	24	0.147
5	0	0.018	24	0.196

Modeling volatility with Fuzzy GARCH models allowed to obtain much better results. Table 1 represents the results of modeling (values of determination coefficient on the test sample) for conventional and Fuzzy GARCH(*p*,*q*) models of different orders. For Fuzzy models both statements (7) and (8) were used to form the explanatory variables set.

It is necessary to note that Fuzzy GARCH models need much more time to be evaluated than usual GARCH(p,q) models. In this work we didn't make any timing and were concentrated only on model performances. Figure 3 demonstrates the results of one-step prediction for both classes of models. Dots represent a confidence interval, calculated by Fuzzy GARCH(5,0) model, whereas dashed line denotes a confidence interval for conventional GARCH(5,0) model. The confidence level is equal to 0.95. It is visible that fuzzy rule-based model commonly gives more exact prediction of *volatility splashes* than usual one.

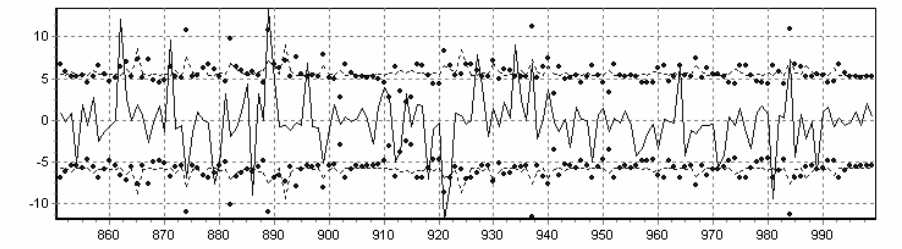


Fig.3. One-step prediction on the test sample of last 149 observations

4.2 Analysis of real-life time series

To demonstrate an application of Fuzzy GARCH models to analysis of real financial time series, we used Dow Jones Industrial Index variations as basic data. Figure 4 shows a time series formed by daily DJII values for the period from 10th February, 2004 till 25th February, 2005 (264 observations).

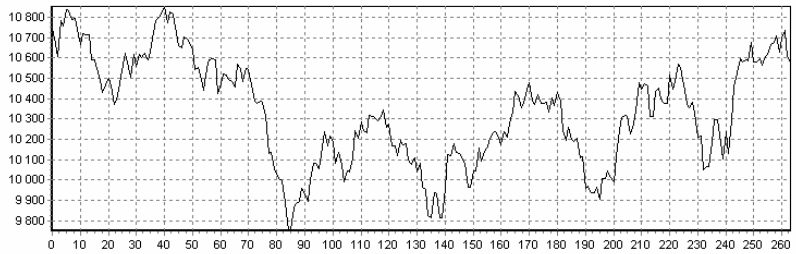


Fig.4. Time series for Dow Jones Industrial daily variations, DJ_t

To perform volatility analysis, it is necessary to pass on to the time series, made of initial by taking the logarithm of quotient of last two values (see Fig. 5):

$$\ln\left(\frac{DJ_{t+1}}{DJ_t}\right). \tag{19}$$

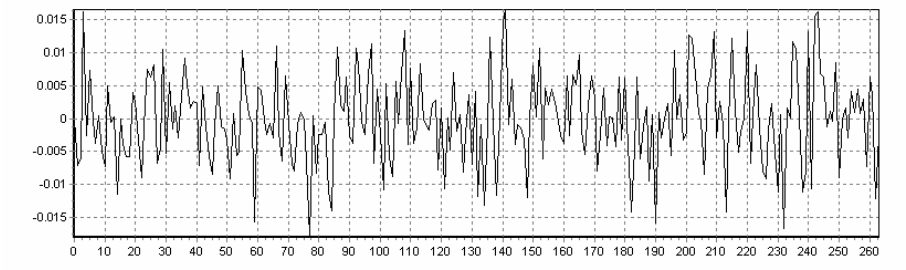


Fig.5. Time series for $\ln\left(\frac{DJ_{t+1}}{DJ_t}\right)$

Conventional GARCH(p,q) models could not model volatility for this time-series in a proper way. Although Box-Ljung statistics showed the significant presence of GARCH effects, application of this class of models gave no more information about the process behavior. The value of determination coefficient on the test sample made of last 39 observations is very close to null.

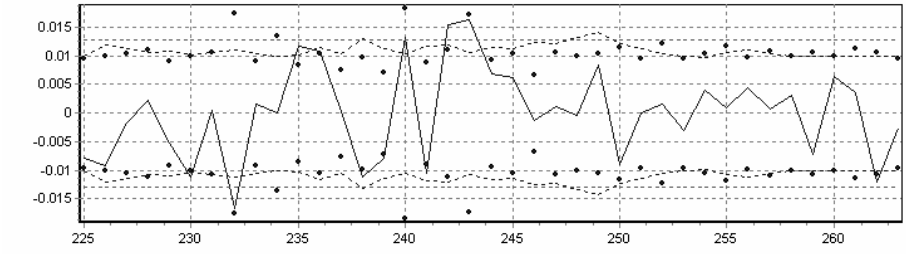


Fig.6. One-step prediction of volatility rate for $\ln\left(\frac{DJ_{t+1}}{DJ_t}\right)$ time series on the test sample made of last 49 observations (using GARCH(5,5) and Fuzzy GARCH(5,*) models)

Application of Fuzzy GARCH models to volatility analysis of $\ln\left(\frac{DJ_{t+1}}{DJ_t}\right)$ time series allowed to get quite adequate results. Figure 6 demonstrates the comparison between conventional GARCH(5,5) and Fuzzy GARCH(5,*) models, related to one-step prediction of volatility on the test sample. Asterisk instead of q order parameter means that local outputs of fuzzy models are *singletons*, i.e.

$$g^i = const,\ i = 1,2,...K\ ,$$

for the statement (6). First parameter, $p=5$, defines the cardinality of explanatory variables set. In the figure 6 dots represent a confidence intervals, calculated by Fuzzy GARCH(5,*) model, whereas dashed line denotes a confidence interval for usual GARCH(5,5) model. The confidence level is equal to 0.95.

Dow Jones Industrial Index is a good playground to demonstrate the application of fuzzy volatility models for calculation of fair option prices. Here we make an estimation of fair price for the *European option*. Basic asset is assumed to be similar to DJII. Other option parameters are supposed to be as follows: date of transaction is 10th February, 2005 (i.e. 253th observation of DJ time series), option period T = 10 workdays, risk-free interest rate r = 0.1. Since "at-the-money" option is considered, the performance price K is set to DJ_{253}. The key parameter is the volatility level d of the basic asset. We used two methods for pricing the option: Black-Scholes formula [4] and binomial Monte-Carlo scheme [7]. The Black-Scholes formula is often [4] accepted as the standard fair option price estimation. It describes the dependence between option price and option parameters K, T, r and d, where volatility level d is supposed to be constant:

$$BS = S(t_0) \cdot N(c) - K \cdot \exp(-rT) \cdot N(c - d\sqrt{T}), \tag{20}$$

where

$$c = \frac{\ln\left[\frac{S(t_0)}{K}\right] + (r + \frac{d^2}{2}) \cdot T}{d\sqrt{T}},$$

$$N(c) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{c} \exp\left(-\frac{z^2}{2}\right) dz \cdot$$

Applying (20) we used sample variance of $\ln\left(\frac{DJ_{t+1}}{DJ_t}\right)$ time series calculated for the first 253 observations.

Table 2. Results of European option pricing for Dow Jones Index

Option pricing method	Fair option price
Black-Scholes formula	101.22
Monte-Carlo with constant volatility	99.38
Monte-Carlo with GARCH(5,5) volatility	100.47
Monte-Carlo with Fuzzy GARCH(5,*) volatility	97.22

Monte-Carlo method lies in making a numerous simulations of basic asset behavior during the option period and averaging them to obtain the final estimation. Monte-Carlo scheme for option pricing models is expounded in [7]. Monte-Carlo pricing models can include different volatility models. We present here usage of three of them: constant volatility level (the same as in the Balck-Scholes foemula), linear (GARCH) model and fuzzy model described above. All volatility models are identified with the first 253 observations of the time series (19). We conducted 1000 simulations for each Monte-Carlo estimations. Table 2 shows the fair option prices obtained by all four methods.

The results in table allow us to conclude that first three methods presume the higher level of volatility during the option period. However the actual volatility level

during the option period (i.e. from 10th till 25th February, 2005) was quite small. To conduct the comparison of different pricing methods we can use the fact that heteroskedasticity hypothesis is declined on the 20 tailing $\ln\left(\frac{DJ_{t+1}}{DJ_t}\right)$ observations, though the normality hypothesis for the whole time series is not declined. It means that the whole option period lies within the same volatility cluster. That is why we can evaluate the sample variance d^* for the last 10 DJ observations and count the appropriate Black-Scholes estimation as the most accurate option price estimation. In practical situations we can't obtain this estimation because at the date of the transaction we have no information about the real behavior of the asset during the option period. Black-Scholes formula with d^* volatility level gave 95.71 as the fair price for the option. Direct comparison of different option pricing methods shows that Monte Carlo scheme with Fuzzy GARCH(5,*) volatility model provides us with the most realistic estimation, quite close to the "historical" Black-Scholes price. This fact comes out of the ability of fuzzy volatility models to predict levels and duration of volatility clusters even for complex time series.

5 Work conclusions

The results of researches that were carried out allow us to make several conclusions:

1. The technique of fuzzy modeling is applicable to analysis of time series with time-varying volatility (i.e. time-varying conditional variance).
2. Using Fuzzy GARCH models often allows to obtain better results of modeling and prediction in comparison with classical GARCH(p,q) models.
3. It is recommended to use Fuzzy GARCH models for estimating volatility and for further application to different financial and other problems (for example to construction of option-pricing models etc.).

References

1. T. Takagi, M. Sugeno. Fuzzy identification of systems and its application to modeling and control. IEEE. Trans.systems, Man and Cybernetics, 15(1):116-132, 1985
2. Wolkenhauer, O. *Data Engineering*, 2001.
3. Engle, R. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", Econometrica, 50 (1982), 987-1007.
4. Shpehard N. "Statistical aspects of ARCH and stochastic volatility", Time Series Models in Econometrics, Finance and Other Fields, p.1-55. Publisher: Chapman & Hall, 1996.
5. D. Nelson. Conditional heteroscedasticity in asset pricing: a new approach // Econometrica, 1991, vol. 59, p. 347-370.
6. Lanne M., Saikkonen P. "Nonlinear GARCH Models for Highly Persistent Volatility" Econometrics Journal , vol.. 8, No. 2, p. 251-276, July 2005.
7. Shvedov A. S. "On the mathematical methods used for work with options" (*in Russian*) Econometrics Journal of HSE, 1998, No. 3, p. 385-409 .

Combining Local Trend Association Network and Clustering in Visualization of Relationships in Time Series Data Bases

Batyrshin I.Z., Klimova A.S., Sheremetov L.B., Velasco-Hernandez J.X.

Mexican Petroleum Institute, Mexico
Kazan State Technological University, Russia
batyr1@hotbox.ru

Abstract. The methods of visualization of relationships between time series in economics, finance, petroleum industry, geophysics etc based on moving approximation transform and local trend associations are considered. These methods are combined with relational invariant clustering based on transitive transformation of similarity relation obtained from the measure of local trend association. The proposed approach to visualization of relationships between time series is demonstrated on real time series from statistical data bases in economics.

1 Introduction

Terabytes of statistical information is collected in all areas of human activity. An important part of such information is represented by time series data bases. Huge time series data bases exist in economics, finance, meteorology, geophysics, petroleum industry etc. For example, all countries regularly publish statistical data in economics, demography etc. [1-3], in petroleum industry there are regularly collected results of geophysical measurements, oil, gas and water production in wells etc. How to use this information for analysis relationships existing between different elements of economical or geophysical systems described by time series is not clear. For example, state statistical centers publish regularly time series of economic indicators of regions, branches of economics etc but what should do with these data a potential user? Is it supposed that he should read huge tables and visually compare the content of strings or columns? Of course he can apply some sophisticated statistical methods to process these data and use it for his goals. But what should do with these data an inexperienced user, for example, sitting in some business or country office? How he can compare these statistical data?

In this paper we propose a method of visualization of relationships between time series by graph based on aggregation of local trend association network [4,5] with relational invariant clustering [6, 7]. It would be nice to make such analysis in statistical centers and to publish time series together with results of visualization of relationships between them to help users in interpretation of statistical data.

Novel methods of analysis of relationships between time series in economics and finance were proposed in [4,5]. These methods are based on moving approximation (MAP) transform, measure of local trend associations and local trend association network. MAP transform calculates slopes of linear approximations of time series in sliding windows and replace time series by the sequence of these slopes. Local trend association measure calculates associations between time series based on distance or similarity measure between their MAP transforms. The important property of these measures is their invariance under linear transformations of time series, and hence, under normalizations of them. Also these measures are sufficiently insensitive to errors in data due to smoothing of data by moving approximations. Association network shows associations between time series as a graph where nodes represent time series and two nodes are connected by edge if an absolute value of association between corresponding time series is greater that given threshold. An application of this technique to analysis of associations between currencies, economic indicators etc was discussed in [4,5].

Really, it is not a new idea to show high associations or similarity between the objects by graph. The set of maximal connected subgraphs (components) of the graphs obtained from association measure for different thresholds correspond to clusters of single linkage (nearest neighbor) clustering algorithm [8, 9]. A single linkage is a popular algorithm of cluster analysis satisfying two important properties: invariance under numeration of objects and invariance under monotone transformation of similarity values. A specific property of this algorithm is that it builds clusters as chains of connected objects [10]. In some applications it may be desirable to construct such clusters. But a drawback of such clusters is that the objects located on the ends of such chains and hence joined in one cluster can have small mutual similarity.

In this paper we propose to join local trend association network of time series, which shows high associations between time series, with clustering of time series based on a scheme of relational clustering procedures satisfying both invariant properties fulfilled for single linkage [6, 7]. The proposed clustering scheme contains clustering procedures based on analysis of similarity between objects with respect to group of neighbors. This scheme includes single linkage as a partial case [11, 12] but generally it can build more compact clusters with high similarity between objects joined in one cluster.

In the following section we give basics of MAP transform and local trend association measures. In Section 3 we describe relational clustering scheme. In Section 4 we discuss the proposed method of aggregation of association networks with invariant clustering and demonstrate it on example of visualization of relationships between real time series. Finally, we make some conclusions and discuss future directions of research.

2 Basics of MAP Transform and Local Trend Associations

A time series (y,t) is a sequence $\{(y_i,t_i)\}$, $i \in I = (1,\ldots, n)$, such that $t_i < t_{i+1}$ for all $i = 1,.., n\text{-}1$, where y_i and t_i are real numbers called time series values and time points, correspondingly,. A time series (y,t) will be denoted also as y. A window W_i of a length $k > 1$ is a sequence of indexes $W_i = (i, i+1,\ldots, i+k\text{-}1)$, $i \in \{1,\ldots, n\text{-}k+1\}$. The sequence $y_{W_i} = (y_i, y_{i+1},..., y_{i+k-1})$ of the corresponding values of time series y is called a partial time series induced by window W_i. A sequence $J = (W_1, W_2,\ldots, W_{n\text{-}k+1})$ of all windows of size k, $(1 < k \le n)$, is called a moving (or sliding) window.

Suppose J is a moving window of size k and $y_{W_i} = (y_i, y_{i+1},..., y_{i+k-1})$, $i \in (1,2,\ldots, n\text{-}k+1)$, are corresponding partial time series in time points $(t_i, t_{i+1},\ldots, t_{i+k-1})$. A linear function $f_i = a_i t + b_i$ with parameters $\{a_i,b_i\}$ minimizing the criterion

$$Q(f_i, y_{W_i}) = \sum_{j=i}^{i+k-1} \left(f_i(t_j) - y_j\right)^2 = \sum_{j=i}^{i+k-1} \left(a_i t_j + b_i - y_j\right)^2, \quad (1)$$

is called a moving (least squares) approximation of y_{Wi}. The solution of (1) is well known and optimal values of parameters a_i, b_i can de calculated as follows:

$$a_i = \frac{\sum_{j=i}^{i+k-1} (t_j - \overline{t_i})(y_j - \overline{y_i})}{\sum_{j=i}^{i+k-1} (t_j - \overline{t_i})^2}, \; b_i = \overline{y_i} - a_i \overline{t_i}, \quad (2)$$

where $\overline{t_i} = \frac{1}{k} \sum_{j=i}^{i+k-1} t_j$, $\overline{y_i} = \frac{1}{k} \sum_{j=i}^{i+k-1} y_j$.

Denote $a = (a_1, \ldots, a_{n\text{-}k+1})$ a sequence of slope values obtained as a result of moving approximations of time series (y,t) in moving window of size k. A transformation $MAP_k(y,t) = a$ is called a moving approximation (MAP) transform of time series y. The slope values $a_1, \ldots, a_{n\text{-}k+1}$ are called local trends.

In many applications time points $t_1, \ldots, t_n$ are increasing with a constant step h such that $t_{i+1} - t_i = h$ for all $i = 1,\ldots, n\text{-}1$. In such cases in MAP transform the set of time points $t = (t_1, \ldots, t_n)$ can be replaced by the set of indexes $I = (1,\ldots, n)$ as follows: $MAP_k(y,t) = (1/h) MAP_k(y,I)$ and the formula (2) for local trends can be simplified as follows [5].

$$a_i = \frac{6 \sum_{j=0}^{k-1} (2j - k + 1) y_{i+j}}{hk(k^2 - 1)}, \; i \in (1,2,\ldots, n\text{-}k+1). \quad (3)$$

Further, for time series with a fixed step we will replace time points by indexes $I = (1,\ldots, n)$ and in (3) the value $h = 1$ will be used. We will denote time series also as $y = (y_1,\ldots,y_n)$ and will use a notation $MAP_k(y)$ for $k \in \{2,\ldots,n\text{-}1\}$.

As a measure of similarity between time series one can use measures of similarity between their MAP transforms. Some of these measures satisfy very nice properties

of invariance to linear transformations of time and time series values. Suppose $y = (y_1,\ldots,y_n)$, $x = (x_1,\ldots, x_n)$ are two time series and $MAP_k(y)= (a_{y1}, \ldots, a_{ym})$, $MAP_k(x)= (a_{x1}, \ldots, a_{xm})$, $(k\in\{2,\ldots,n-1\}, m= n- k+1)$, are their MAP transforms. The following function is called a measure of local trend associations:

$$coss_k(y,x)= \frac{\sum_{i=1}^{m} a_{yi}\cdot a_{xi}}{\sqrt{\sum_{i=1}^{m} a_{yi}^2 \cdot \sum_{j=1}^{m} a_{xj}^2}}. \qquad (4)$$

Suppose p,q,r,s, $(p,r\neq 0)$ are real values and (y,t) is a time series. Denote $py+q = (py_1+q, \ldots, py_n+q)$ and $rt+s = (rt_1+s, \ldots, rt_n+s)$. A transformation $L(y,t)= (py+q, rt+s)$ is called a linear transformation of time series (y,t). It can be shown that if L_1 and L_2 are two linear transformations of time series (y,t) and (x,t) given by the sets of parameters (p_1,q_1,r_1,s_1) and (p_2,q_2,r_2,s_2), respectively, where $p_1, p_2, r_1, r_2 \neq 0$, then

$$coss_k(L_1(y,t), L_2(x,t))= sign(p_1)\cdot sign(r_1)\cdot sign(p_2)\cdot sign(r_2)\cdot coss_k((y,t), (x,t)). \qquad (5)$$

It is a very nice invariance property of local trend association measure under various types of normalization of time series. Additionally to *coss* one can use other distance of similarity measures between MAP of time series.

Analysis of associations between time series is based on the analysis of associations between them for different window size. The sequence of association values $AF(y,x)= (coss_2(y,x),\ldots, coss_n(y,x))$ for all sizes of window is called an association function [5]. Generally some subset $K\subseteq \{2,\ldots,n\}$ of all possible windows may be considered, in this case we will talk about function $AF_K(y,x)$ defined on the set of windows K. The average value of this function may be used as a measure of association between time series:

$$AM(y,x) = \frac{1}{|K|}\sum_{k\in K} \cos s_k(y,x) \qquad (6)$$

Another, "max association" measure, can be used, for example, when K contains a short subsequence of $\{2,\ldots,n\}$, for example $K=\{2,3,4\}$:

$$AM(y,x) = \max_{k\in K}(abs(\cos s_k(y,x)))\cdot sign(\max_{k\in K}(\cos s_k(y,x)) + \min_{k\in K}(\cos s_k(y,x))).$$

This measure calculates maximal positive or minimal negative association between time series y and x for window sizes given in K.

Association measure takes values in interval [-1,1]. Association values greater than 0.5 and less than -0.5 are considered, respectively, as highly positive and highly negative associations. Association network is as a graph such that nodes represent time series and two nodes are connected by edge if an absolute value of association between corresponding time series is greater that given threshold. Examples of application of association measures and association networks in economics and finance are considered in [4,5].

3 A Scheme of Hierarchical Invariant Clustering Procedures

Association measure can be transformed into a similarity measure S or dissimilarity measure D as follows:

$$S(y,x) = (AM(y,x)+1)/2, \qquad D(y,x) = (1-AM(y,x))/2. \tag{7}$$

These measures satisfy the properties: $S(y,x)$, $D(y,x) \in [0,1]$,

$$S(y,x) = S(x,y), \qquad S(y,y) = 1, \tag{8}$$

$$D(y,x) = D(x,y), \qquad D(y,y) = 0.$$

These measures can be used by some clustering procedure for clustering time series and the result of such clustering will depend on applied procedure. Often, the result will depend also on numeration of time series. For this reason it is important to apply clustering procedures invariant under numeration of classified objects. Such scheme of clustering procedures was studied in [6, 7].

A function $S: X \times X \rightarrow [0,1]$ satisfying on X properties (8) is called a valued (fuzzy) similarity relation. This function is called a fuzzy equivalence relation if it satisfies on X $(\vee,\wedge)$ – transitivity condition:

$$S(x,y) \geq min\{(S(x,z),S(z,y)\}.$$

The properties of equivalence relations were studied in [11-13]. The property of $(\vee,\wedge)$ - transitivity is dual to the ultrametric inequality:

$$D(x,y) \leq max\{(D(x,z),D(z,y)\}.$$

Fuzzy equivalence relations and ultrametric distances define hierarchy of nested partitions of the set X. For this reason a clustering procedure can be considered as a transformation $E = Q(S)$ of similarity relation S into a fuzzy equivalence relation E or as a transformation of distance measure in ultrametrics [9, 14, 15]. A general scheme of clustering procedures based on these considerations was proposed and studied in [6,7]:

$$E = Q(S) = TC(F(S)),$$

where F is some "correction" of given similarity relation S and TC is a procedure of $(\vee,\wedge)$-transitive closure of valued similarity relation. The procedure of transitive closure is studied in the theory of fuzzy relations, in graph theory and in cluster analysis and may be realized by single linkage clustering method or by special algorithms [11-13]. This procedure possesses two types of invariance: invariance under numeration of objects and invariance under monotone transformation of similarity values in S. The second invariance property is also important for clustering procedures because it insures insensitivity of these procedures to some choice of similarity measure . When a correction procedure F also exhibits both types of invariance, then the clustering procedure Q will also satisfy both invariance properties.

A clustering procedure consisting of these two procedures F and TC is called a relational clustering procedure. In [6] it was shown also that a reasonable correction

procedure F should satisfy the following constraint: $F(S) \subseteq S$, where $\subseteq$ is a partial ordering of similarity relations. Below is a description of parameterized correction procedure from this scheme [7].

Suppose $f_1, f_2, f_3: R \to R$ are monotone functions. A correction procedure depends on the following sets and functions:

$$V_y(x) = \{z \in X \setminus \{x,y\} \mid S(x,z) \geq f_1(S(x,y))\}, \quad V_x(y) = \{z \in X \setminus \{x,y\} \mid S(y,z) \geq f_1(S(x,y))\}.$$

The sets $V_y(x)$, $V_x(y)$ denote, respectively, the sets of objects "similar" to x and to y when the value $f_1(S(x,y))$ serves as a criterion of this similarity. The set

$$V(x,y) = \{z \in X \setminus \{x,y\} \mid \max\{S(x,z), S(y,z)\} \geq f_2(S(x,y))\},$$

contains the objects from X which are "similar" at least to one of the objects x and y. When $f_1 \equiv f_2$ we have $V(x,y) = V_y(x) \cup V_x(y)$. This set will be considered as the set of "neighbors" of x and y. The objects in $V(x,y)$ will be taken into account when decision about correction of the value $S(x,y)$ will be made. The set

$$W(x,y) = \{z \in V(x,y) \mid \min\{S(x,z), S(y,z)\} \geq f_3(S(x,y))\},$$

denotes the set of "strong" or "common" neighbors, i.e. objects which are "similar" to both objects x and y. The objects from $W(x,y)$ will "support" the value $S(x,y)$. When $f_1 \equiv f_3$ we have $W(x,y) = V_y(x) \cap V_x(y)$. The functions f_1, f_2, f_3 used in clustering procedure are called neighborhood functions.

The decision about correction of the value $S(x,y)$ will depend on the relative part of objects "supporting" the similarity value $S(x,y)$. It can be used the following method to calculate for each pair of objects x and y this relative part:

$$h = \frac{|W(x,y)|}{|V(x,y)|},$$

where, by definition, $h = 1$ if it denominator is equal to 0. The correction procedure $F(S)$ in the clustering procedure Q may be defined as follows:

$$F(S(x,y)) = \begin{cases} S(x,y) & \text{if } h \geq p \\ F_j(x,y) & \text{otherwise} \end{cases},$$

where $p \in [0,1]$, j are parameters and $F_j(x,y)$ is a corrected value such that $F_j(x,y) \leq S(x,y)$. We will suppose that $F_j(x,y)$ depends on the values $S(x,z)$, $S(y,z)$ for all objects z belonging to the sets of neighbors $V_y(x)$, $V_x(y)$ and $V(x,y)$. We require also that $F_j(x,y) \geq \min_{z \in V}\{S(x,z), S(y,z)\}$, where $V = V_y(x) \cup V_x(y) \cup V(x,y)$. The specific definition of $F_j(x,y)$ will be discussed later. When $p = 0$, from $h_j \geq 0$ it follows that $F(S(x,y)) = S(x,y)$, i.e. for all x,y from X the values $S(x,y)$ will be unchanged and $Q(S) = TC(F(S)) = TC(S)$, i.e. clustering procedure will coincide with single linkage method.

Let $L_V(x,y)$ denotes the list of all values $S(x,z), S(y,z)$, $(z \in V)$, which are less than $S(x,y)$, ordered in descending order. Denote the number of elements in $L_V(x,y)$ as $m = |L_V(x,y)|$ and the elements of $L_V(x,y)$ as l_k (k=1, m). If $m > 1$ then $l_k \geq l_{k+1}$ for all k=1, m-1. When m>1, possible corrections will be defined by parameter j:

j=1: $F_j(x,y) = l_m$, i.e. l_m is the minimal value of $L_V(x,y)$;

j=2: $F_j(x,y) = l_1$, i.e. l_1 is the maximal value of $L_V(x,y)$;
j=3: $F_j(x,y) = (\sum l_k)/m$, i.e. $F_j(x,y)$ is the mean of all values from $L_V(x,y)$;
j=4: $F_j(x,y) = l_k$, where $k \in \{1, \ldots, m\}$ - parameter, F_2 is a special case of F_4;
j=5: $F_j(x,y) = median(L_V(x,y))$.

All correction procedures $F_j(x,y)$ for j=1,…,5 are invariant under numeration of objects and correction procedures $F_j(x,y)$ for j=1,2,4,5 are invariant under monotone transformations of similarity values. As result, clustering procedures using these corrections will also satisfy these invariance properties.

4 Example of Visualization of Relationships Between Time Series

In this paper the following procedure for visualization of relationships between time series is proposed. First, a measure of local trend association between time series is calculated. This measure is transformed into similarity function which is used further for clustering time series by relational clustering algorithm. From hierarchical clustering obtained by clustering procedure for different parameter values the most often presented clustering is selected. To show the associations between time series the threshold is selected such that the number of "strong" associations, with absolute value higher than this threshold, maximize the number of strong associations between time series joined in the same cluster and minimize the number of strong associations between time series from different clusters. Such association network will be called an optimal one.

The method is demonstrated on the example of time series of investments in fixed capital in 14 regions of Volga Federal District, Russia [16]. These regions have the following abbreviations and names: (BA) Bashkortostan, (MA) Mariy El, (MO) Mordoviya, (TA) Tatarstan, (UD) Udmurtiya, (CH) Chuvashiya, (KI) Kirovskaya oblast, (NI) Nigegorodskaya oblast, (OR) Orenburgskaya oblast, (PN) Pensenskaya oblast, (PR) Permskaya oblast, (SM) Samarskaya oblast, (SR) Saratovskaya oblast, (UL) Ulyanovskaya oblast.

For considered example the set of window sizes K= {2,3} was used for calculation of "max association" measure. This association measure was converted to similarity relation by (7). After application of relational clustering procedure based on this similarity relation four different clusters were obtained: C1= {BA, TA, OR, PR, UD}, C2= {KI, UL, MA, NI, CH}, C3= {PN, SM, MO}, C4={SR}. Fig. 1 and Fig. 2 show obtained optimal association networks (for threshold values H1= 0.95868 and H2= 0.94686). Solid lines and dash-dot lines show associations between time series from the same clusters and from different clusters, respectively. For both networks the number of associations which not confirm the obtained clustering is equal to 6.

Additional criteria can be used for selection one from these two optimal visualizations of associations between time series. For the higher value of threshold (Fig. 1) the number of associations between time series from different clusters is minimal. For the minimal value of threshold (Fig. 2) the number of intra-cluster associations is maximal. One of these criteria can be used as additional criteria for final selection of visualization.

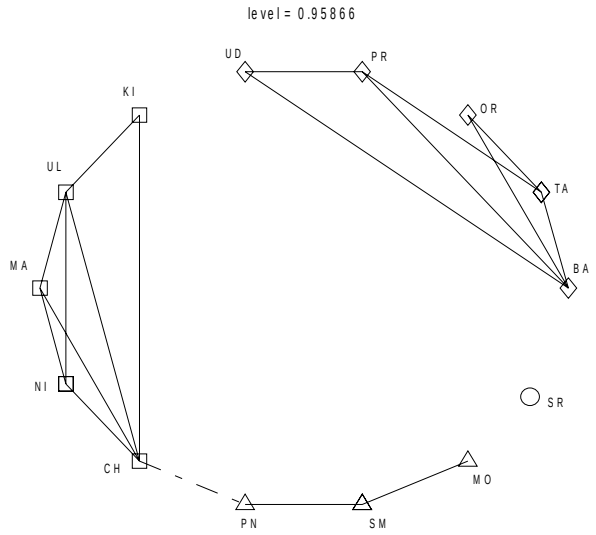


Fig. 1. Optimal association network containing minimal number of associations between time series from different clusters (shown by dash-dot lines)

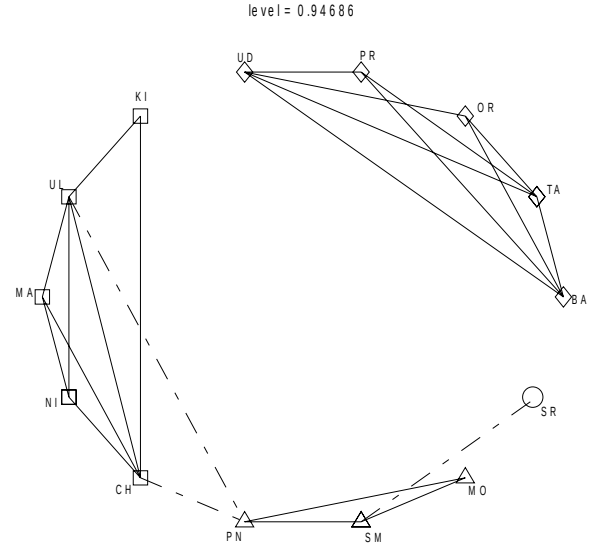


Fig. 2. Optimal association network containing maximal number of associations between time series from the same clusters (shown by solid lines)

Fig. 3 shows time series of investments in fixed capital in 14 regions of Volga Federal District, Russia, used in this exmaple. These time series are ordered such that time series from the same cluster are located sequentially. After such ordering the similarity between time series from the same clusters can be seen visually more or less clear.

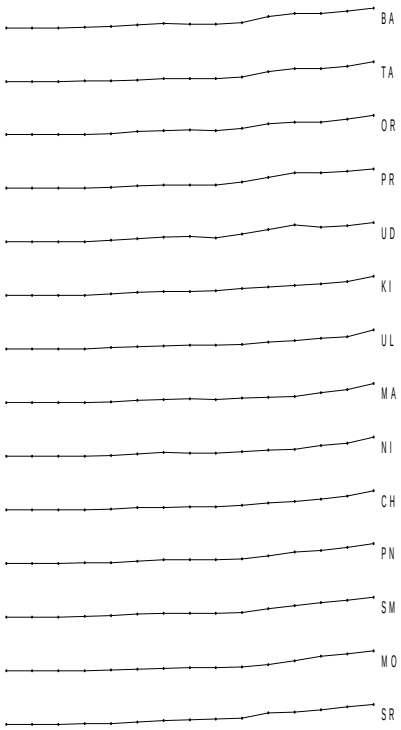


Fig. 3. Time series of investments in fixed capital in 14 regions of Volga Federal District, Russia

Conclusions

The paper introduces a new technique for visualization of relationships between time series combining local trend association measure and relational clustering procedures. This technique can be used by centers of statistics to facilitate a perception of statistical information published regularly in economics and finance. This technique can be useful also for visualization of relationships between geophysical, industrial and other time series, e.g. in petroleum industry. Generally, the proposed approach

can be used for visualization and clustering of any objects. Another possibility to visualize relationships between time series is to combine local trend associations with two- or three-dimensional representation [17, 18].

References

1. Russia's Federal State Statistics Service, http://www.gks.ru/wps/portal/english
2. Australian Bureau of Statistics, http://www.abs.gov.au/
3. Instituto Nacional de Estadística, Geografía e Informática (INEGI), http://www.inegi.gob.mx/inegi/default.asp
4. Batyrshin, I., Herrera-Avelar, R., Sheremetov, L., Suarez, R. Moving approximations in time series data minig. In: Proc. Intern. Conf. on Fuzzy Sets and Soft Computing in Economics and Finance, FSSCEF 2004, St. Petersburg, Russia, vol. I (2004) 62-72
5. Batyrshin, I., Herrera-Avelar, R., Sheremetov, L., Panova, A. Moving approximation transform and local trend associations in time series data bases. In: Batyrshin, I., Kacprzyk, J., Sheremetov, L., Zadeh, L.A. (Eds.). Perception-based Data Mining and Decision Making in Economics and Finance. Springer-Verlag, Series: Studies in Computational Intelligence, vol. 36 (2006) (to be published)
6. Batyrshin, I.Z., Shuster, V.A. The structure of semantic spaces of verbal estimates of actions. Acta et Commentationes Universitas Tartuensis, Transactions on Artificial Intelligence, Principle Questions of Knowledge Theory, Tartu, 688 (1984) 20-38 (in Russian)
7. Batyrshin, I.Z., Rudas, T., Klimova, A. On general scheme of invariant clustering procedures based on fuzzy similarity relation. In: Proc. Int. Conf. Fuzzy Sets and Soft Computing in Economics and Finance, FSSCEF 2004, St. Petersburg, Russia, vol. I (2004) 122-129
8. Jardine, N., Sibson, R. Mathematical taxonomy. London: John Wiley & Sons (1971)
9. Johnson, S.C. Hierarchical clustering schemes. Psychometrika, 32 (1967) 241-254
10. Duda, R.O., Hart, P.E. Pattern Classification and Scene Analysis. Wiley, New York (1973)
11. Tamura, S., Higuchi, S., Tanaka, K. Pattern classification based on fuzzy relations. IEEE Trans. Systems, Man, Cybernetics, SMC-1 (1971) 61-66
12. Dunn, J.C. A graph-theoretic analysis of pattern classification via Tamura's fuzzy relation. IEEE Trans. on Systems, Man and Cybernetics, SMC-4 (1974) 310-313
13. Zadeh, L.A. Similarity relations and fuzzy orderings. Information Sciences, 3 (1973) 177-200
14. Jardine, C.J., Jardine, N., Sibson, R. The structure and construction of taxonomic hierarchies. Math. Biosci., 1 (1967) 173-179
15. Hartigan, J.A. Representation of similarity matrices by trees. J. Amer. Statist. Assoc., 62 (1967) 1140-1158
16. Russia in figures. Federal Service of state statistics http://www.gks.ru/scripts/db_inet/dbinet.cgi?pl=2702005
17. Kruskal, J.B., The relationship between multidimensional scaling and clustering, In: Classification and Clustering, John Van Ryzin (ed.), Academic Press, New York (1977) 17-44
18. Klimova, A. Evolutionary procedures of visualization of multidimensional data. Proc. Int. Conf. Fuzzy Sets and Soft Computing in Economics and Finance, St. Petersburg, Russia, vol. I (2004) 130–139