Инженерен подход и приложението му в пазарния и финансовия сектор

Автор: Александър Ефремов

Engineering approach with applications in retail and finance sectors

Author: Alexander Efremov

Резюме: В статията се анализира тенденцията знанията, придобивани в инженерните специалности да намират приложение в нетехнически области. Разглеждат се приложения, където по правило статистическият подход е водещ. Също така се дискутират предимствата в познаването на теорията на динамичните системи и идентификацията, както и общото между теория на управлението и вземането на решения. Основните направления в изследването са: идентификация и прогнозиране (в смисъла на Калмановата филтрация), а по-задълбочено засегнатите приложни области са пазарни системи и финанси.

Ключови думи: многомерни системи, идентификация, Калманова филтрация, оптимизация на стратегии, системен подход, финанси, маркетинг.

Abstract: In this paper is analyzed the tendency of expanding the field, where the engineering approach is successfully applied. The focus is on some nontechnical areas, where historically the approach is purely statistical. Some advantages of the usage of dynamic theory, system identification and even of some aspects of control theory during the strategy optimization processes are discussed. The main directions of the study from theoretical point of view are in the field of system identification and forecasting (in terms of Kalman Filtering). The applications, considered in more details are in the retail and finance sectors of industry.

Keywords: multivariable systems, identification, Kalman filtering, strategy optimization, economic engineering, finance, market.

1. INTRODUCTION

Today many companies from different industry sectors possess large databases with many giga, tera, and even petabytes of data. Moreover, the variety of variables, which are collected increases more and more. As a result the data generated by mankind increases exponentially literally every second. It is forecasted that in 2020 the data, generated in USA will reach 6,6 ZB (1 ZB = 1021 B). The presence of such a huge amount of this raw material – the data, is the main reason for the rapid development of system identification and more specifically the theory about multivariable systems (i.e. Multiple Input Multiple Output (MIMO) systems). There are applications of MIMO system identification where the developed models have thousands and more input and output variables. In addition to modelling, the data is also used in the optimal decisions making process where the problem dimension also increases extremely. For instance, in marketing, there are optimization problems in which the number of parameters under optimization (like the actions, which have an influence in the market system) can be tens of billions [13].

In many nontechnical areas increases the need of experts [9] that are able to design suitable data mining solutions for the needs of particular businesses, to develop algorithms, to construct auto-modelling methodologies [3, 20], etc. In addition the dynamic aspects of the investigated systems are still neglected or are not appropriately accounted for. This is very important premise for the successful usage of engineering approach in such areas like finances, marketing, sociology, medicine etc. Many examples can be found for the valuable results of applying the engineering knowledge in areas like stock exchange and financial markets [18, 19], retail sector [5, 10, 11, 12, 7, 14], credit risk [8], marketing strategy and optimal decisions making [22] etc. Furthermore, the automation of the activities, related to identification [4], forecasting, including the Kalman filter applications [17, 15], strategy optimization, etc., nowadays are often the only way of managing with the increased systems dimension, which is typical for many nontechnical systems.

2. ENGINEERING ECONOMICS AND OTHER NONTECHNICAL SYSTEMS

In technics when the system under investigation is defined in terms of a single unit the input-output variables reflect the physics of the processes in that unit. For example the behavior of an electric motor is connected with variables associated with electricity and mechanics. But with expanding the boundaries of the investigated system (for instance when consider a system of units like a production line or a whole factory) the technical aspects of the system give way to the economic once. Usually the model of such a complex system doesn't aim to represent all particular relationships between the control variables and the system outputs but rather the attention is focused on variables like, quantity of raw materials, materials features, prices and energy needed for the production process, on one hand, and the final products quality, prime cost, market price (the last depends on the market demand and competitors behavior). Exactly this transition from purely technical specifics to economical once is the reason to start using the term engineering economics system (EES).

Next considerations are focused not only on EES but also on real life systems from nontechnical areas such as market, finance, medicine, social systems, etc. An example for a system from the credit risk is "applicant for a credit" (schematically shown on fig. 1). Here the goal of modelling is to develop a model, which will be used to forecast the behavior of potential credit applicants in terms of risk rate. More often the dependent variable(s) is the probability for an applicant to be "good" and the independent variables are the available applicants' characteristics like age, income, etc. An important specific of the applicants' data is that it incorporates only static aspects of the applicants' behavior. This naturally leads to static models, which at first sight looks like a simplification of the system identification task compared with dynamic case, where the models describe both statics and dynamics. But just dropping out the time scale, this leads to the usage of data processing technics, which are not typical for the engineering approach developed within the years. In spite of this difference the principles of experimental modelling, studied in engineering specialties still take place in behavioural modelling of the credit applicants.

Another example from market systems is a store (fig. 2) or a retail chain. The models of such systems are again used for forecasting, but here variables of interest are the products demand. Input variables are the retailers' actions and output variables are product sales. The difference from the above mentioned example is that here the data explains the dynamics in the products' demand. It is typical for these systems the forecasted processes to have components like trend and seasonalities. Also very often the demand contains oscillating components with different periods (e.g. monthly, weekly, daily period, etc.). Then further actions are needed like modelling and temporarily removing of these oscillating components and the trend from the initial data. After this signals decomposition, the approaches for model building are in mostly the same as those used in the technical area.

3. FREQUENTLY OBSERVED FEATURES OF NONTECHNICAL SYSTEMS

Although the engineering methods are developed for technical systems, they may also be applied successfully for economic systems. In many cases the methods are not applied directly for objective considerations like: small number of values across the time scale due to the large sampling time; too low level of the signal to noise ratio, which is not typical for the engineering practice; requirements for the usage of the not typical in technics categorical variables, etc. For these and other reasons the application of engineering methods is connected with an additional activity, which plays the role of an interface between them and the particular nontechnical problem. This interface is based on statistical methods, machine learning [2], information extraction from audio signals [9], text mining, etc. – theoretical fields developed in different branches of the applied mathematics, which are not directly related to engineering.

Some features of the nontechnical systems are discussed below. Here the aim is not a comprehensive outlining of their specifics, but rather, to pay an attention on the link ensuring the correct usage of engineering techniques mostly among the above examples.

A small number discrete time instants

The main issue that arises when trying to apply dynamic systems theory to an economic system is the very limited number of time instants for which are available the input-output observations. For example, in the above mentioned market systems, the practice is the data to be accumulated on a weekly base. The reason is that usually in the retail market chains, the supply and also activities like marketing campaigns, promotions, prices' discounts, etc. are updated weekly and hence the demand forecast should be synchronized with these events. So, working with a sampling time of one week, 52 observations would be available per year. This is a major difference with the technical systems, where the sampling time could be seconds, milliseconds or even smaller. When talk about retail chains a significant model improvement can be obtained, if the data from all stores is appropriately gathered. Then, even in this case – with a small number of observations in time, the resulting reduction of the data uncertainty would increase the reliability of the final model.

More extreme is the case when represent the behavior of accepted credit applicants where the sampling time is one month (usually for instalment credits the payments are made once a month). Here the largest dimension of data set is not connected with the time scale but is across the applicants. The large number of observed individuals gives the possibility to extract a deterministic component of applicants' behavior (the common behavioural aspects) and in this way to reduce the data uncertainty.

Although applying the just explained uncertainty reduction, obtained by accounting for a large number of realizations, the problem of the small number of discrete time instants remains. In engineering it is normal to work with datasets containing hundreds or thousands of discrete observations, but not tens. To overcome this issue it is appropriate to apply machine learning methods like decision trees, clustering, etc. With their help it is possible the problem to be reformulated and then resolved by the well-known engineering approaches like Kalman filtering and interacting multiple models methods for target tracking.

Time-varying behavior

As already mentioned, nowadays is available a huge amount of data. But often it appears that large amount of values don't represent the current systems performance. For example in the finance there are big data sets, collected from many years and looks like there is a possibility to provide a large number of time instants for the modelling data sets. But due to macro-economic changes; drifts in the population segments, which use particular products; also social, political and other phenomena; historical data become unsuitable for the modelling or optimization purposes. Hence only recent samples should be used, as they contain valuable information for the current systems behavior. This leads to the above-mentioned issue (small number discrete time instants). Some well-known engineering techniques for time-varying systems can be used, like introducing of forgetting factor, may give an opportunity to extend the observation interval and to bring the problem formulation closer to the typical engineering problems.

Uncertainty

An important point when work with data is to account for the presence of uncertainty in the available observations. Usually in technics, at the design stage of the plants (units, production lines, etc.) are set the input and output variables, also a low level of noise in the measurements is ensured and the influence of environmental disturbances is suppressed. On the other hand, in economics the systems under investigation are not designed by engineers and such uncertainty reduction cannot be provided. For example, in credit industry the system, which is modelled, is a credit applicant. Here the variables used in the identification process are indirect and as a rule they are not enough for a precise customers' description. Sometimes, because of low restrictions in some countries are banned for use even accessible for observation data, such as "sex". In the market sector some significant factors, which affects the demand, are also hard to observe (like the effects caused by competitors). Other type of uncertainty is the discrepancy of the real system and its approximation – the model. It is named system uncertainty and usually it could not be neglected as typically the systems are too complicated and cannot be entirely represented, also are nonlinear, time-varying, etc.

Fighting down the high level of uncertainty is to use the available, often insufficient a-prior information about the system, which helps to suppress the effect of not deterministic variations in data. In the credit industry, the a-priori information usually is reduced to expected trends or with other words how the dependent variable(s) are correlated with factors' variables. For example, it is expected that when the factor 'Age' grows the probability of an applicant to be "good" accordingly to increase and hence the estimation of the corresponding model parameter should be positive.

The incorporation of the a-priori information in the modelling process leads to complication of the (more or less) standard identification methodologies and also makes the eventual automation of model development much more difficult. Nevertheless, the usage of this information is the most appropriate way to reduce wrong (from business perspective) choices of factors and not to relay only on statistics, when build models.

Data pre-processing

Getting back in credit industry where applicant is a MIMO system and especially where data uncertainty is huge, the transformation of the rough data into an appropriate data set for modelling is a key role in identification. For this reason here the data pre-processing may take about 80% and even more from the total time needed to build the model. Unlike in technics where data is highly informative, in finance is necessary to reformulate the initial variables. The aims are: to save the meaningful information, to reduce uncertainty, to "clean" the data set from missing values and outliers, to standardize, to transform variables (if the relations between them are nonlinear), to encode (if the variables are nominal or ordinal), to categorize variables, etc. At the stage of data preparation exactly this activities play the role of bridge to the engineering approach. Once prepared, the dataset for modelling can be considered as the well-known data matrix or matrices [3].

The situation in the market sector is very similar. Here, if a product is not supplied in the store for a long time interval, it is normal to assume the price, sales, etc. of this product as missings. For this reason it is very important how to accept the lack of data. Sometimes, the presence of outliers can strongly distort the model. These values often are named "leverage points" [16] because of their significant influence on the model.

Number of potential factors

Frequently, when expand the system limits, e.g. by including more units from a production line or adding more aspects under consideration, the number of potential factors grows. For instance, in credit industry, the eligible factors for entering in the model at the data pre-processing stage can grow up to several thousands. Furthermore a hypermarket may have tens of millions inputs and hundreds of thousands outputs. Moreover, if dynamic models are used this leads to an additional growth of the factors number as it is necessary to include historical data as well.

The above-mentioned data size is the main reason to focus the attention on auto-modelling solutions. Here again some methods, developed in nontechnical sectors, can be used to reduce the data size.

In addition, typical for the MIMO systems is the presence of multicollinearity within some factors, which imposes the usage of numerically stable realizations of the applied methods where matrix inversion is necessary. For this purpose are suitable: Tikhonov regularization [23], LU, QR, SVD, EVD, Choleskey and other matrix decompositions. [12, 13, 15, 21, 25, 24].

Large size of data samples

As mentioned earlier, one way for a significant reduction of data uncertainty is to increasing the size of data samples. In practice, there is a natural limit of the data size after which a further concatenation of data does not improve sensibly the model accuracy. Also, in marketing strategy optimization the optimized scheme (the solution) should be rollout to all potential customers and they could be millions and more. Here again machine learning and numerical methods [6, 16] may help to make both modelling and optimization tasks practically solvable.

Another approach for managing with large data sets is the usage of distributed systems and parallel computations. As a result tasks which would take days can be performed in minutes. This topic is not in the scope of the paper.

4. CONCLUSION

There are many aspects of the nontechnical systems, which prevent the direct application of engineering approach. But even in these cases, after appropriate transformation of the original task, it is possible to apply the engineering principles and methods and so to achieve qualitative improvement of the solutions. For example, the correct application of the dynamic systems theory, including Kalman filtering, etc., can increase the solutions quality in terms of demand forecast, prediction of the borrowers' behavior and risk assessment on the base of their historical data, market strategy optimization etc.

From the above analysis can be concluded that the combination of engineering methods with statistics, machine learning, numerical methods, etc. is the key to manage with the enormous amount of available data in the nontechnical areas. But these instruments should be used accounting for the a-priori information (known aspects of the particular systems). In fact the a-priori information gives the possibility to design models, to formulate optimal solutions, etc. that lead to adequately solutions. This is especially true when large number possible solutions exist and the probability of making wrong conclusions, laying only on the statistics and the available data, is significant.

References

1. Agre, G., 2011. Lectures on data mining regularities. Sofia University "St Kliment Ohridski", Faculty of Mathematics and Informatics.
2. Agre, G., 2011. Lectures on machine learning. Sofia University "St Kliment Ohridski", Faculty of Mathematics and Informatics
3. Efremov, A., 2014. Multivariable System Identification, ISBN 978-954-9489-42-2.
4. Efremov, A., 2008. Multivariate Time-Varying System Identification at Incomplete Information. Ph.D. thesis.
5. Tikhonov, A. N. and Arsenin, V. Y., 1974. Solutions of ill posed problems, Nauka, Moscow.
6. Todorov, A., Kostov, K., Georgiev, B., Enev, S., Sergieva, V., Efremov, A. and Mehmed, A., 2011. Automation of Technological Processes – manual for laboratory exercises. Technical University – Sofia.
7. Baltagi, B. H. and Levin, D., 1992. Cigarette taxation: raising revenues and reducing consumption. – In: Structural Change and Economic Dynamics, volume 3, № 2, pp. 321 – 335.
8. Bartels, R. H. and Golub, G. H., 1969. The simplex method of linear programing using LU decomposition, Communications of the ACM, pp. 266–268.
9. Bernard, J. T., Idoudi, N., Khalaf, L. and Yeloud, C., 2007. Finite sample multivariate structural change tests with application to energy demand models. – In: Journal of Econometrics.
10. Brandimarte, P., 2002. Numerical Methods in Finance. A MATLAB-Based Introduction. John Wiley & Sons, Inc., New York.
11. Data Science Society: http:// www.datasciencesociety.net/.
12. Deschamps, P. J., 1997. Full maximum likelihood estimation of dynamic demand models. – In: Journal of Econometrics, volume 82, pp. 335 – 359.
13. Experian: http://www.experian.com/.
14. Friesen, J., Capalbo, S. and Denny, M., 1992. Dynamic factor demand equations in U.S. and Canadian agriculture. – In: Agricultural Economics, volume 6, pp. 25l – 266.
15. Gill, J. and King, G., 2004. What to do when your hessian is not invertible – alternatives to model respecification in nonlinear estimation. – In: Sociological Methods & Research, volume 32, – 4, pp. 1 – 34.
16. Golub, G. H. and Van Loan, Ch. F., 1989. Matrix Computations, Second edition, Johns Hopkins University Press, Baltimore, MD.

17. Gonzalo, J. and Martinez, O., 2004. Threshold integrated moving average models (Does size matter? Maybe so). – In: European Economic Association, Econometric Society, North American Winter Meetings 145 – EEA-ESEM. Econometric Society.

18. Grewal, M. S. and Andrews, A. P., 2001. Kalman Filtering: Theory and Practice Using Matlab. John Wiley & Sons, Inc., New York, second edition.

19. Hoffmann, J. P., 2004. Generalized Linear Models – an Applied Approach. Pearson Education, Inc., Boston.

20. Kalman, R. E., 1960. A new approach to linear ltering and prediction problems. – In: Transactions of the ASME – Journal of Basic Engineering, Vol. 82, Series D, pp. 35 – 45.

21. Marchev, A., Jr and Marchev, A., 2012. Selecting and Simulating Models for Management of Investment Portfolios Using Cybernetic Approach. Economic Alternatives, Issue 2.

22. Marchev, A., Jr., 2010. Model for sequential dynamic competition between random investment portfolios and portfolios selected by collective expert opinions. – In: Proceedings of the Challenges for Analysis of the Economy, the Businesses, and Social Progress, pp 579 – 598, At Szeget, Hungary.

23. Nelles, O., 2001. Nonlinear System Identification. From Classical Approaches to Neural Networks and Fuzzy Models. Springer-Verlag, Berlin Heidelberg.

24. Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., 1992. Numerical Recipes in C – The Art of Scientific Computing. Cambridge University Press, 40 West 20th Street, New York, NY 10011-4211, USA.

25. Verhaegen, M. and Verdult, V., 2007. Filtering and System Identification. A least Squares Approach. Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK.

26. Watkins, D. S., 2007. The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods. Siam, ISBN-13: 978-0898716412.