ON THE INFORMATION-BASED APPROACHES IN ECONOMICS AND FINANCE

Jan Kalina

Abstract

The amount of information relevant for decision making in economic and financial applications keeps dramatically increasing. This will allow a radical transformation of economic and financial practices towards the ideals that may be denoted as information-based. These information-based procedures will exploit modern statistical tools for extracting useful knowledge from the available information in the form of data. This paper is devoted to a discussion of information-based approaches to investing, which will be based on portfolio diversification and optimization, or to transfer learning, where the last is based on information flow from given learning tasks (especially in finance) to a new related task. In addition, a novel robust procedure for estimating the parameters of the Gompertz curve is proposed here. On the whole, the information-based approaches in economics and finance will definitely go beyond the currently discussed evidence-based approaches, which are inspired by the currently popular concept of the evidence-based medicine.

Key words: information sciences, information flow, transfer learning, nonlinear regression, robust estimation

JEL Code: D81, C14, D83

Introduction

In economic and financial sciences and applications, it is crucial to see the fundamental difference between information and knowledge. The information has the form of available data and does not contribute to any decision making without being transformed to useful knowledge. Information as a capital or opportunity with a potential to be transformed to wealth or power may be regarded as an accelerator of changes of the economy (Assaf et al., 2022). Information sciences (statistics, information theory, and also information economics) are able to extract the relevant knowledge from available data under the presence of uncertainty. Still, only the knowledge obtained from the data may be applied to particular decisions or actions related to finance and economy.

The current digitalization processes going hand in hand with an increasing availability of big data represents a key tool for transforming the decision making processing towards the ideals that may be described as information-based. The increasing availability of artificial intelligence (AI) tools in finance represents also a revolution e.g. in automatic trading, decision support related to investing, or constructing government budgets (Jeong and Kim, 2019). Big data are nowadays commonly used e.g. for credit risk, allowing to construct targeted (i.e. tailor-made) decisions about the likelihood of a debtor to repay the loan. As another application, we may mention decision support systems as promising tools of artificial intelligence. The information-based approaches based on exploiting massive data (Big Data), may be viewed as an extension of evidence-based approaches, which have been discussed in the literature.

From the statistical point of view, the information-based tools will exploit linear or nonlinear regression analysis, the analysis of (high-frequency) time series, extensions of quantiles (value at risk), neural networks for smoothing nonlinear and nonstationary time series (Kalina, 2013), or dimensionality reduction e.g. for finding the most relevant assets. The unstable nature of the complex data in the currently turbulent economy will require the use of robust statistics for highly contaminated data by noise or outlying values (Kalina, 2015). Also robust versions of regression quantiles (particularly for predictions) or (possibly Bayesian) estimation techniques based on maximizing the entropy, which search for the conditions with the smallest available information, will be needed. Theoretical frameworks connecting information with entropy for financial time series have to be developed e.g. in Benedetto et al. (2021). The information-based approaches also require (often conceptually new) reliable tools for obtaining knowledge from data and the principal directions of research include robust statistical estimators, robust feature extraction, robust training of deep neural networks, their reliability (e.g. hypothesis tests for verifying their probabilistic assumptions), and explainability. The modern advanced statistical methods have not been taught at economic or business universities yet.

This paper formulates a novel insight into the perspectives of the information-based approaches in finance and economics. Section 1 is devoted to information-based investing. The transfer of information and transfer learning are discussed in Section 2. In Section 3, a novel robust procedure for estimating the parameters of the Gompertz curve is proposed.

1 Information-based investing

Investors have been using subjective approaches for their decision making to a large extent. Psychology still plays a prominent role in current decisions related to investing. This is true in spite of the availability of theoretical results for the construction of the optimal portfolio. It is namely known in the field of portfolio management that the optimal portfolio should be diverse, avoiding strongly correlated stocks. A diverse portfolio has a smaller volatility compared to any individual asset. This section recalls available concepts of active investment and evidence-based investing, and proceeds to the vision of information-based investing and its theoretical foundations.

A success in investing is ensured for the investors that have better information or more reliable predictions compared to their rivals. For example, stock prices are heavily influenced by announced macroeconomic facts or predictions. The so-called active investment is an investment strategy characterized by an active search for perspective stocks that are likely to outperform available combinations of stocks (stock market indexes). The aim to overcome the performance of the market is difficult and requires to predict future performance under risk (uncertainty) in a better way compared to the majority of other investors. The useful information for active investors may include newly available information about a given company or about the development of the whole economy. The ability to combine multi-modal information coming from various sources represents an advantage; at the same time, automatic information extraction from free (narrative) texts, which represents a task for natural language processing, is beneficial here. A special case of active investing is the evidence-based investing (factor investing), which requires to take various macroeconomic and microeconomic factors into consideration.

Within the vision of a prospective approach exceeding the tools of evidence-based investing, which may be denoted as information-based investing, investors will base their decision making entirely on the objective analysis of data and computations. Information-based approach will exploit the potential to obtain useful knowledge from available big data; this will be enabled by logical and analytical thinking, probabilistic reasoning, and Bayesian statistical inference. The risk in portfolio investing is namely in the uncertainty of future results or in the fractal nature of volatility (Bessler et al., 2021).

The aim of investing is to obtain a large (maximal) gain at the price of a small (minimal) risk. Statistically speaking, the investor aims at high mean but low standard variance (or standard deviation) of the random prices. Often, this couple of contradictory criteria is denoted as a risk-return trade-off. The situation is in fact more complex and it is needed to consider the information triangle, where the gain, risk (evaluated as variance), and liquidity play their roles. The theoretical basis for information-based investing is the portfolio optimization approach developed by Harry Markowitz (*1927). This mean-variance model represents one of central topics of modern finance and is formulated as a mathematically elegant solution in the form

$$max \frac{\alpha^T x}{\sqrt{x^T \Sigma x}} \quad \text{subject to } 1_n^T x = 1, \quad \text{where } 1_n = (1, \dots, 1)^T. \tag{1}$$

A sparse (regularized) approach to financial portfolio selection was later appraised by Steland and von Sachs (2017). The covariance matrix plays the key role in the Markowitz model. Because it is highly vulnerable with respect to outlying values in the data, it is natural to consider robust alternatives of the empirical covariance matrix. Another problem is the formulation of the covariance matrix in situations with the number of variables (stocks) exceeding the number of observations. The solution in the form of a regularized covariance matrix was proposed by Ledoit and Wolf (2022), two economists, who applied their approach to stocks in their paper.

2 Transfer of information and transfer learning

Dominant markets have been well known to have influence on less important ones. Also, a time series of stock prices may have influence on the time series of cryptocurrency returns. In both situations we speak about **information flow**, i.e. transfer of information (about the performance of the stock market) with causal effects on financial returns in another market. Available attempts to evaluate (quantify) the effect of one stock market on another by means of measures of information theory focused on the use of transfer entropy as a measure allowing to evaluate the transfer of information. Transfer entropy has a better interpretation compared to correlation coefficients, where the latter do not imply causal relationships. Information flow has been subject of theoretical studies in the context of information theory and/or cybernetics with relevant applications also in biology, sociology, or neurosciences. One of interesting applications was presented by Assaf et al. (2022), who used information theoretical measures (entropy, approximate entropy) to evaluate the information sharing in the cryptocurrency markets during the COVID-19 pandemic. A regression model for the information flow related to tourism was modeled by means of fractals and correlation dimensions in Chen et al. (2017).

Smith (2015) applied the concept of economic equilibrium to a model for information transfer. He explained several macroeconomic phenomena (laws) to be in accordance with a law of **information equilibrium** between information supply and information demand. Smith formulated the information equilibrium as a general economic principle and denoted economic reasoning based on statistical methodology to represent so-called statistical economics.

Transfer learning, which is a popular technique in current machine learning, has already found promising applications in the field of finance. It has the aim to improve learning in a new task using the knowledge acquired in previous learning tasks and is based on exploiting a transfer of information from given learning tasks to a new (but related) task. It aims at learning common features across similar problems. As transfer learning starts by analyzing previous tasks and extracting knowledge from them, it is used as a prior knowledge for the new dataset. Thus, transfer learning is often understood as a special case of **metalearning**, which can be characterized as an approach of automated machine learning, often successful in the task of automatic selection (recommendation, prediction) of the most appropriate method or algorithm from several alternatives for a particular dataset. The problem of automatic recommendation of the most suitable method has been denoted as the algorithm selection problem (ASP), optimal algorithm selection, or automatic method selection (AMS).

To describe at least a couple of recent financial applications of transfer learning, let us mention the work of Jeong and Kim (2019), who proposed an automatic system for financial trade decisions based on reinforcement learning (particularly on the approach commonly denoted as Q-learning). To prevent from overfitting, the system contains a transfer learning component. This performs a pretraining of the models on a subset of the stocks under considerations. Kraus and Feuerriegel (2017) analyzed financial disclosures (documents related to financial statements of individual companies) in various forms, including context of articles in newspapers or internet announcements. These narrative texts may be valuable for investors. In their paper, Kraus and Feuerriegel proposed and implemented a decision support system allowing to assist investors with decision making. This exploits deep learning for the supervised learning task, while transfer learning is applied to pretrain the system itself on a different (and much larger) corpus. This transfer of information helps to improve the decision making process significantly.

Early applications of transfer learning included mainly smaller and very similar (homogeneous) tasks. Current applications are mainly in image analysis, robotics or reinforcement learning. Another important area of transfer learning applications in the proposal of Bayesian networks, which can be (according to a particular task) easily modified. Most commonly, transfer learning is performed by means of neural networks, while the features common across individual learning tasks are encoded in their weights.

1 Sequential transfer, based on a sequential training of the given model by individual tasks.
2 Parallel transfer: the models of neural networks are trained in a parallel way so that individual networks share some of their parts.
3 Multitask learning: the last layer of the neural network typically serves as a switch among individual tasks. The joint information about all tasks is included in the weights of the neural network.

3 The Gompertz curve as a nonlinear regression model

The information-based approaches will have to exploit reliable (resistant) statistical methods and this is true also for the Gompertz curve, which is a popular model in the actuarial science for modeling mortality in the population. In this section, its history sand importance are briefly recalled and a novel robust approach for estimating the parameter of the Gompertz curve is then proposed in Section 4.1.

Let us start by recalling the predecessors of the approach of Gompertz for mortality modeling. The year 1654 is considered to represent the founding year of the probability theory, when Pierre de Fermat (1607-1665) and Blaise Pascal (1623-1662) started their correspondence about the problem of division of stakes. Jacob Bernoulli (1655-1705) founded probability theory as a self-standing scientific discipline. Gottfried Wilhelm Leibniz (1646-1716) formulated the concepts and aims of the modern probability theory. As the first philosopher of probability, he did not consider probability only mathematically, but with a potential for a broader employment in practice. Leibniz as a legal scholar earned his licence to practice law; he aimed at finding the principal logical concepts (principal components) in the legal system. Leibniz had an influence on Abraham de Moivre (1667-1754), who started to use probability theory to construct mortality tables for life insurance applications. His law of mortality relied on the assumption that the occurence of deaths is uniformly distributed over the ages until very high age values.

The Gompertz curve was proposed by Benjamin Gompertz (1779-1865) for the context of mortality investigations. Although Gompertz was working at the London Stock Exchange, he is now renowned especially for his mortality models. Gompertz was direclty influenced by de Moivre's law of mortality, but did not assume the mortality rate to be a constant, but rather used a nonlinear trend for its modeling and prediction. The Gompertz model is very flexible, the curve is a nonlinear curve suitable for time series analysis not only for mortality data and may look quite similarly to the logistic curve (Li et al., 2021). Gompertz was able to perform numerical estimation (i.e. computed hismself the estimates) of parameters based on given data. The methods proposed by Gompertz have been used even nowadays, e.g. for modeling mortality data from the popular Human mortality database and for applications to life insurance (Tai and Noymer, 2018).

The literature on curve fitting for time series commonly describes the Gompertz curve without explaining how to estimate the parameters, so the readers are left to trust the available software procedures. For practitioners, it is however important to describe not only the model, but also to explain how the parameters are estimated. Available ways for estimating the parameters include the maximum likelihood, least squares, method of moments, or others. All these approaches are parametric estimation techniques so that the estimates suffer under the presence of outliers in the data. A novel robust estimation will be now proposed for the Gompertz curve.

3.1 Robust estimation of the Gompertz curve

Let us consider the nonlinear regression model

$$Y_i = f(X_i) + e_i, \quad i = 1, \dots, n, \tag{2}$$

where $Y =$ is a continuous response, $X_i =$ for $i = 1, \dots, n$ is the vector of independent variables observed for the i-th measurement, f is a given nonlinear function, and is the vector of random regression errors (disturbances). Time series in financial or actuarial applications are typically considered as equidistant, thus the notation

$$Y_t = f(X_t) + e_t, \quad t = 1, \dots T, \tag{3}$$

is often used instead of (2), where $t = 1, \dots T$corresponds to time.

The Gompertz curve (Gompertz curve) is defined as

$$Y_t = exp(\gamma + \alpha\beta^t) + e_t, \quad t = 1, \dots T, \tag{4}$$

where $\beta > 0$. In such a model, the response depends on time only. The formulation of the model (4) allows also different parametrizations, but the key issue is the estimation of the parameters. Replacing the unknown function f in (3) by (4) actually replaces the nonparametric estimation task to a parametric one, i.e. to a model, which is completely specified up to a finite number of parameters. Also, there exists a nonlinear regression model (not only for time series) inspired by (4); the model is known as the Gompertz regression model.

For the trend model (4), one can exploit the NLWS (nonlinear least weighted squares) method for estimating the parameters. Such approach does not represent a new model, but rather a novel estimation method for existing regression models. The approach is a nonlinear extension of the least weighted squares estimator for linear regression (Kalina, 2012). To describe the NLWS procedure formally, let the vector parameters in the given model (5) be denoted by θ. In order to give the formal definition of the NLWS estimator, we will use the notation $u_i(\theta)$ for the residual corresponding to the i-th observation for a given estimator of β. We consider the residuals arranged in ascending order in the form

$$u_{(1)}^2(\theta) \leq \cdots \leq u_{(n)}^2(\theta). \tag{5}$$

Let us define the least weighted squares estimator of the parameters in the model (5) as

$$argmin \sum_{i=1}^{n} w_i u_{(i)}^2(\theta), \tag{6}$$

where the argument of the minimum is computed over all possible values of θ and where $w_1, \dots, w_n$ are magnitudes of weights determined by the user. The NLWS estimation principle is known to assess a high robustness in terms of a high breakdown point and also other appealing properties such as efficiency for non-contaminated data.

Conclusions

The dramatically increasing availability of information can be perceived not only as a potential, but also as a threat: the current revolution in economics and finance may, just like the industrial revolution in the 19th century, lead to an impoverishing of large layers of the population. As the amount of available relevant information has been increasing so dramatically, there is a threat that the decision making in an abundance of information will be even worse compared to decision making completely without any information. The transformation of economic and financial process towards the information-based ideals will require lifelong education of experts as well as adjusting the curricula taught at economics and business universities or within Master of Business Administration (MBA) courses. This new content of curricula will have to reflect the recent advances in quantitative approaches or in financial statistics and financial informatics, which are fields centered around the crucial concept of information. Also the financial literacy of the whole population requires to be redefined. In this paper, we omit various other aspects of information, e.g. information asymmetry, the role of information in the macroeconomic equilibrium (Navrátil et al., 2022), text mining applications in finance, or encoder-decoder frameworks for learning stock trading rules.

Acknowledgements

Aleš Neoral was helpful with the section on transfer learning. The work was supported by the project 24-11146S of the Czech Science Foundation.

References

Assaf, A., Charif, H., Demir, E. (2022). Information sharing among cryptocurrencies: Evidence from mutual information and approximate entropy during COVID-19. *Finance Research Letters,* 47A, 102556.

Benedetto, F., Mastroeni, L., Vellucci, P. (2021). Modeling the flow of information between financial time-series by an entropy-based approach. *Annals of Operations Research*, 299, 1235–1252.

Bessler, W., Taushanov, G., Wolff, D. (2021). Factor investing and asset allocation strategies: A comparison of factor versus sector optimization. *Journal of Asset Management*, 22, 488–506.

Chen, Y., Wang, J., Feng, J. (2017). Understanding the fractal dimensions of urban forms through spatial entropy. *Entropy*, 19, 600.

Jeong, G., Kim, H.Y. (2019). Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning. *Expert Systems with Applications*, 117, 125–138.

Kalina, J. (2012). Highly robust statistical methods in medical image analysis. *Biocybernetics and Biomedical Engineering*, 32(2), 3–16.

Kalina, J. (2013). Highly robust methods in data mining. *Serbian Journal of Management*, 8, 9–24.

Kalina, J. (2015). Three contributions to robust regression diagnostics. *Journal of Applied Mathematics, Statistics and Informatics*, 11(2), 69–78.

Kraus, M., Feuerriegel, S. (2017). Decision support from financial disclosures with deep neural networks and transfer learning. *Decision Support Systems,* 104, 38–48.

Ledoit, O., Wolf, M. (2022). The power of (non-)linear shrinking: A review and guide to covariance matrix estimation. *Journal of Finance Econometrics,* 20, 187–218.

Li, H., Tan, K.S., Tuljapurkar, S., Zhu, W. (2021). Gompertz law revisited: Forecasting mortality with a multi-factor exponential model. *Insurance: Mathematics and Economics,* 99, 268–281.

Navrátil, R., Taylor, S., Večeř, J. (2022). On the utility maximization of the discrepancy between a perceived and market implied risk neutral distribution. *European Journal of Operational Research,* 302, 1215–1229.

Smith, J. (2015): Information equilibrium as an economic principle. ArXiv:1510.02435v1.

Steland, A., von Sachs, R. (2017). Large-sample approximations for variance-covariance matrices of high-dimensional time series. *Bernoulli*, 23, 2299–2329.

Tai, H.T., Noymer, A. (2018). Models for estimating empirical Gompertz mortality with an application to evolution of the Gompertzian slope. *Population Ecology,* 60, 171–184.

Contact

Jan Kalina
The Czech Academy of Sciences, Institute of Computer Science
Pod Vodárenskou věží 2, 182 07, Prague 8, Czech Republic
& Charles University, Faculty of Mathematics and Physics
Sokolovská 83, 186 75 Prague 8, Czech Republic
kalina@cs.cas.cz