Probability in Action

edited by

Tadeusz Banek, Edward Kozłowski

Politechnika Lubelska

Lublin 2014

Probability in Action

Monografie – Politechnika Lubelska

Politechnika Lubelska
Wydział Zarządzania
ul. Nadbystrzycka 38
20-618 Lublin

Probability in Action

edited by
Tadeusz Banek
Edward Kozłowski

Politechnika Lubelska
Lublin 2014

Reviewer:
prof. dr hab. Jurij Kozicki, Maria Curie Skłodowska University

Editorial staff and typesetting:
Edward Kozłowski
Przemysław Kowalik

Publication approved by the Rector of Lublin University of Technology

ISBN: 978-83-7947-036-5

Publisher: Lublin University of Technology
ul. Nadbystrzycka 38D, 20-618 Lublin, Poland
Realization: Lublin University of Technology Library
ul. Nadbystrzycka 36A, 20-618 Lublin, Poland
tel. (81) 538-46-59, email: wydawca@pollub.pl
www.biblioteka.pollub.pl
Printed by : TOP Agencja Reklamowa Agnieszka Łuczak
www.agencjatop.pl

The digital version is available at the Digital Library of Lublin University of Technology: www.bc.pollub.pl
Impression: 100 copies

Content

Preface

Probability in Action is a set of papers made by the members of Department of Quantitative Methods in Management in Lublin Technical University listed in alphabetical order: Tadeusz Banek, Przemyslaw Kowalik, Edward Kozlowski, Agnieszka Surowiec and Tomasz Warowny. The volume includes also some papers written by our colleagues from other departments, Department of Mathematics: Adam Gregosiewicz, Dariusz Majerek, Malgorzata Murat, Ernest Nieznaj, Wojciech Rosa and Witold Rzymowski and our co-workers from Department of Mechanics and Vibroacoustics AGH Cracow University of Science and Technology: Wojciech Batko and Olaf Knapik. This initiative is more or less a continuation of our previous book Process Conrol in Management, published in 2009.

There are many reasons to collect papers in the book form. In the case of potential reader's diverse scientific interests the presentation in a single book with many different papers has some advantages over a monothematic monograph - something interesting for everyone, one may say. This approach is successfully applied by many editors and publishers. In view of the number of books which already exist on the subject, one can ask, with a reason, whether another book about 'probability in action' is really needed. In particular, what is our excuse for writing this one? Our answer is this – on the desks of scientific journals editors there are large piles of excellent up-to-day papers and manuscripts waiting for publication resulting with delay measured in years. We want to bring the results of the recent research to the reader without such a delay. We will utilize mechanism that allows a rapid publication in style of arXiv.org of Cornell University. Last but not least, this book has to substantiate our activity in the last two years. These are the goals of this book.

As the title Probability in Action suggests we chose the applications of probability methods as Ariadna's thread making the papers selection for this volume. The term 'applications' we understand in the broad sense. Here, by applications we understand a type of problems at hand and the methods for solving them rather then what purists consider as real applications, i.e., the applications in engineering, economy, biology, or social sciences, when one starts from a real data, has to build a model which is consistent with data and allows to predict new data. A natural determination to understand the world around which sometimes takes a form of scientific curiosity is a motivation to pose intriguing questions, state problems and hypothesis. If the inspiration comes from the reality outside world of mathematics and the results can be verified in practice, then such work fits into our cathegory of broad applications.

Tadeusz Banek[1]

A causal construction of diffusion processes

Keywords: diffusion processes, translation of Wiener processes, Girsanov theorem, calculus of variations

Abstract

A simple nonlinear integral equation for Ito's map is obtained. Although it does not include stochastic integrals, it does give a causal construction of diffusion processes which can be easily implemented by iteration systems. Next, the result is applied for calculations of various type of variational derivatives. Applications in financial modelling are discussed.

1 Introduction

Diffusions are an important class of stochastic processes. They are Markov, and have continuous trajectories. There are extensive, competent historical surveys of the topic by D.W. Stroock (2003), D.W. Stroock and S.R.S. Varadhan (1979, 1987) and we recommend Stroock's discussion to the interested reader. Here, we shall point out only main stages. Historically, the first construction was given by A.N. Kolmogorov (1931) and, since then, a problem of constructing diffusions in $\mathbb{R}^n$ having differential operators as generators and no barriers, is known as the Kolmogorov problem. We will restrict our attention to the later class and call it (for short), K$-$ diffusions. The second construction of K$-$ diffusions was given by K. Ito (1951) (and it is called Ito diffusions, too). The theory of ordinary-stochastic equations by H. Sussman (1977) and H. Doss (1977) and some its modifications (see I. Karatzas and S.E. Shreve 1991) may be regarded as a deterministic variant of Ito's theory. The third is known as a solution of D.W. Stroock and S.R.S. Varadhan martingale problem (see D.W. Stroock and S.R.S. Varadhan 1979). The fourth is given by the Isobe-Sato formula (see E. Isobe and S. Sato 1983), which gives Wiener-Ito integrals for chaos decomposition of K$-$ diffusions.

[1]Technical University of Lublin, Faculty of Management, Department of Quantitative Methods in Management, Nadbystrzycka 38, 20-618 Lublin, *e-mail: t.banek@pollub.pl*

In this paper, we propose a new, pathwise variant of Ito's construction of K$-$ diffusions. Although the construction uses a Wiener process (Ito's idea), it does not involve Ito's integrals. It consists in:

(a) Solving a nonlinear, deterministic, Volterra type integral equation

$$c\left(w(t)-\int_o^t \varkappa(x(s))\,ds\right)=x(t), \tag{1}$$

where $w \in C_T \triangleq C([0,T];\mathbb{R})$, c and $\varkappa$ are ordinary scalar function to be specify later. Under mild assumptions (1) can be solved pathwise and nonanticipative, i.e., for any w, $v \in C_T$ one finds x_w, $x_v \in C_T$ such that restrictions x_w, x_v, on $[0,t]$ coincide if $w(s)=v(s)$, $s \in [0,t]$.

(b) Forming a map $X_t(w):[0,T]\times C_T \to \mathbb{R}$, such that $X_t(w)=x_w(t)$. Hence, $X(w)$ belongs to the space $\mathfrak{G}(C_T)$ of all nonanticipative mappings from C_T to C_T and it is a fix point of the operator $\mathfrak{L}:\mathfrak{G}(C_T)\to\mathfrak{G}(C_T)$, defined by

$$\mathfrak{L}(X(w))(t)=c\left(w(t)-\int_o^t \varkappa(X(w)(s))\,ds\right), \tag{2}$$

where we adopt the convention $X(w)(t)=X_t(w)$.

(c) Showing that $X_t(w)$ is a K$-$ diffusion assuming that $w(t)$, $t\in[0,T]$ is a Wiener process.

(d) Proving that it is true in the opposite direction as well, i.e. if $X_t(w)$ is a K$-$ diffusion, then it is a fixed point of $\mathfrak{L}$.

It is instructive to compare an intuitive picture behind Ito's theory (here, we again recommend D.W. Stroock 2003, D.W. Stroock and S.R.S. Varadhan 1979, 1987) to the picture of K$-$ diffusions as suggested by (1). In the first picture, infinitesimal increments of K$-$ diffusions are resulting from combined effects of two forces: a deterministic drift and random (Gaussian) fluctuations. Since combination means a sum here, the both forces (deterministic and random) have the same status in creation of K$-$ diffusions. However, (1) suggests other picture, or, looking from a cybernetic perspective, better to say a "behaviour". Namely, x_w follows w, what is easily visible on the diagram below

$$\begin{array}{ccccccc} w \longrightarrow^{(+)} & \otimes & \longrightarrow [c(\cdot)] \longrightarrow & \circ & \longrightarrow x_w \\ & {}^{(-)}\uparrow & & \downarrow & \\ & \longleftarrow [\int] & \longleftarrow & [\varkappa(\cdot)] & \end{array}$$

which explains the idea of simple iteration system which works according to (1). With $y(t)\triangleq\int_0^t \varkappa(x(s))\,ds$, this behaviour is even more explicit

$$\frac{d}{dt}y(t)=\varkappa\circ c(w(t)-y(t)).$$

Hence $y_w(t) \triangleq \int_0^t \varkappa(x_w(s))\,ds$ follows w with the speed equal to the image of the difference $w(t) - y_w(t)$ under $\varkappa \circ c$. Thus, in this picture we have pure **deterministic** mechanism, expressed in the terms of $\varkappa \circ c$ composition, which forces y_w to follow a **random** path w. Even more, a rule of producing actions according to the current errors is known in Automatic Control as a classical **feedback rule**, which in turn is the most transparent idea of **Cybernetics**. If there is any Variational Principle responsible for this rule, it is an open question.

The paper is organized as follows. In a preliminary section we state an auxiliary result on (1). In the next section we prove an equivalence theorem, which is the main result of this section. Several corollaries are also included. In the next section we apply the equivalence theorem for calculations of $w-$ derivatives of $X_t(w)$. Indication for financial mathematics is discussed next.

2 Preliminaries

We state here the following

Lemma 1 *Assume $c : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz and $\varkappa : \mathbb{R} \to \mathbb{R}$ measurable and bounded. Then, (a) for any $w \in C_T$, there exists a unique $x_w \in C_T$ satisfying (1), (b) for any $w \in C_T$ and any $\xi \in C_T$, a sequence of successive approximation*

$$\begin{aligned} x_0 &= \xi,\ x_{n+1} = \Phi_w(x_n), \\ \Phi_w(x)(t) &\triangleq c\left(w(t) - \int_o^t \varkappa(x(s))\,ds\right) \end{aligned}$$

is convergent in any norm $\|\cdot\|_\lambda$, $\lambda \geq 0$, to x_w, where
$\|x\|_\lambda = \max\left\{e^{-\lambda t}|x(t)|; 0 \leq t \leq T\right\}$,
(c) a mapping $C_T \ni w \mapsto X(w) \triangleq x_w \in C_T$ is locally Lipschitz (in any $\|\cdot\|_\lambda$, $\lambda \geq 0$), and nonanticipating.

Proof. The proof consists in two steps. In the first one, one can show that (a),(b),(c) hold when c is globally Lipschitz. In the second one, one can apply a method of continuation in the locally Lipschitz case. ■

3 Equivalence theorem

Let $g \in C^1(\Delta) > 0$, $\Delta \subset \mathbb{R}$, an open interval and $f : \mathbb{R} \to \mathbb{R}$ measurable. Define two functions:

$$c'(x) = g(c(x)) \tag{3}$$

and

$$\varkappa(x) = \frac{g'(x)}{2} - \frac{f(x)}{g(x)}. \tag{4}$$

Example 2 *(a) Let* $g(x) = \sqrt{1+x^2}$. *Then* $c(x) = \sinh(a+x)$. *For an arbitrary* $\phi \in C(\mathbb{R})$, *set* $f(x) = \frac{x}{2} - \phi(x)\sqrt{1+x^2}$, *then* $\varkappa(x) = \phi(x)$. *(b) Let* $g(x) = |x|^{\alpha}$, $1/2 < \alpha < 1$. *Then for* $a \in \mathbb{R}$, *we have* $c(x) = [sign(a+x)][(1-\alpha)|a+x|]^{1/1-\alpha}$. *For* $\phi \in C(\mathbb{R})$, *set* $f(x) = \frac{\alpha}{2} sign(x)|x|^{2\alpha-1} - |x|^{\alpha}\phi(x)$, *then* $\varkappa(x) = \phi(x)$.

Theorem 3 *Assume* $c : \mathbb{R} \to \mathbb{R}$, $\varkappa : \mathbb{R} \to \mathbb{R}$ *satisfy (3)(4) and* $\varkappa$ *is bounded. If* $w(t)$, $t \in [0,T]$ *is a Wiener process on* $(\Omega, \mathfrak{F}, \mathbb{P})$, *then (I) the mapping* $[0,T] \times C_T \ni (t,w) \mapsto X_t(w) \in \mathbb{R}$ *satisfies the equation*

$$c\left(w(t) - \int_0^t \varkappa(X_s(w))\,ds\right) = X_t(w) \tag{5}$$

$\mathbb{P}-$ *a.s., iff it solves strongly and uniquely Ito's differential equation*

$$\begin{aligned} dx(t) &= f(x(t))\,dt + g(x(t))\,dw(t) && (6)\\ x(0) &= c(0) && (7) \end{aligned}$$

$\mathbb{P}-$ *a.s., (II) if (5) holds, then*

$$X_t(w) = c(\widetilde{w}(t)), \tag{8}$$

where

$$\widetilde{w}(t) = -\int_0^t \varkappa \circ c(\widetilde{w}(s))\,ds + w(t).$$

Proof. Assume that $X_t(w)$ solves (5) and denote

$$\widetilde{w}(t) \triangleq w(t) - \int_o^t \varkappa(X_s(w)).$$

>From Ito's formula and (3), (4) we get

$$\begin{aligned} & dc(\widetilde{w}(t)) && (9)\\ =\; & \left[\frac{1}{2}c''(\widetilde{w}(t)) - c'(\widetilde{w}(t))\varkappa(X_t(w))\right]dt + c'(\widetilde{w}(t))\,dw(t) \\ =\; & \left[\frac{1}{2}g(c(\widetilde{w}(t)))g'(c(\widetilde{w}(t))) - g(c(\widetilde{w}(t)))\varkappa(X_t(w))\right]dt + g(c(\widetilde{w}(t)))\,dw(t) \\ =\; & g(c(\widetilde{w}(t)))\left\{\frac{1}{2}g'(c(\widetilde{w}(t))) - \left[\frac{g'(X_t(w))}{2} - \frac{f(X_t(w))}{g(X_t(w))}\right]\right\}dt \\ & + g(c(\widetilde{w}(t)))\,dw(t). \end{aligned}$$

Since (by the assumption)

$$X_t(w) = c(\widetilde{w}(t)),$$

thus the RHS of (9) equals

$$= f(X_t(w))\,dt + g(X_t(w))\,dw(t).$$

Hence $X_t(w)$ solves (6),(7) since $X_0(w) = c(0)$).

Now, it is the reverse direction. Let $X_t(w)$, $X_0(w) = c(0)$ solve strongly (6),(7). Then

$$\begin{aligned} dX_t(w) &= [f(X_t(w)) + g(X_t(w))\,\varkappa(X_t(w))]\,dt + g(X_t(w))\,[dw(t) - \varkappa(X_t(w))\,dt] \\ &= [f(X_t(w)) + g(X_t(w))\,\varkappa(X_t(w))]\,dt + g(X_t(w))\,d\widetilde{w}(t), \end{aligned}$$

where $\widetilde{w}(t)$ (from Girsanov theorem) is a Wiener process on a "new" space $\left(\Omega, \mathfrak{F}, \widetilde{\mathbb{P}}\right)$ with a measure

$$\begin{aligned} \widetilde{\mathbb{P}}(A) &= \int_A \Lambda d\mathbb{P},\ A \in \mathfrak{F}, \\ \Lambda &= \exp\left[\int_0^T \varkappa(X_t(w))\,dw(t) - \frac{1}{2}\int_0^T \varkappa^2(X_t(w))\,dt\right], \end{aligned}$$

($\mathbb{E}_{\mathbb{P}}[\Lambda] = 1$, because $\varkappa$ is bounded). It follows that $X_t(w)$ satisfies on $\left(\Omega, \mathfrak{F}, \widetilde{\mathbb{P}}\right)$ the equation

$$\begin{aligned} & dX_t(w) \\ = & \left[f(X_t(w)) + g(X_t(w))\left[\frac{g'(X_t(w))}{2} - \frac{f(X_t(w))}{g(X_t(w))}\right]\right] dt + g(X_t(w))\,d\widetilde{w}(t) \\ = & \; g(X_t(w))\left[\frac{g'(X_t(w))}{2}dt + d\widetilde{w}(t)\right]. \end{aligned} \tag{10}$$

It can be verified directly that $c(\widetilde{w}(t))$ solves (10), hence

$$X_t(w) = c(\widetilde{w}(t)) \tag{11}$$

by uniqueness of (6),(7). Hence, we get

$$X_t(w) = c(\widetilde{w}(t)) = c\left(w(t) - \int_0^t \varkappa(X_s(w))\,ds\right) \tag{12}$$

on the "old" space $(\Omega, \mathfrak{F}, \mathbb{P})$. The proof of the part (I) is completed. For the part (II), note that (5) implies

$$\begin{aligned} \widetilde{w}(t) &= w(t) - \int_0^t \varkappa(X_s(w))\,ds \\ &= w(t) - \int_0^t \varkappa \circ c\left(w(s) - \int_0^s \varkappa(X_u(w))\,du\right) ds \\ &= w(t) - \int_0^t \varkappa \circ c(\widetilde{w}(s))\,ds. \end{aligned}$$

■

Example 4 *(continued). With g and f as above, we have the integral equation, case (a)*

$$X_t(w) = \sinh\left(a + w(t) - \int_0^t \phi(X_s(w))\,ds\right)$$

and Ito's equation

$$dx(t) = \left[\frac{x(t)}{2} - \phi(x(t))\sqrt{1+x^2(t)}\right]dt + \sqrt{1+x^2(t)}\,dw(t),$$

case (b)

$$\begin{aligned} X_t(w) &= \left[sign\left(a + w(t) - \int_0^t \phi(X_s(w))\,ds\right)\right] \\ &\quad \times \left[(1-\alpha)\left|a + w(t) - \int_0^t \phi(X_s(w))\,ds\right|\right]^{1/1-\alpha} \end{aligned}$$

and Ito's equation

$$dx(t) = \left[\frac{\alpha}{2} sign(x(t))\,|x(t)|^{2\alpha-1} - |x(t)|^{\alpha}\phi(x(t))\right]dt + |x(t)|^{\alpha}\,dw(t).$$

Remark 5 *From (11) and (12) we have on* $\left(\Omega, \mathfrak{F}, \widetilde{\mathbb{P}}\right)$

$$X_t\left(\widetilde{w} + \int_0^{\cdot} \varkappa \circ c(\widetilde{w}(s))\,ds\right) = c(\widetilde{w}(t)).$$

Example 6 *(a) Since in our example* $\varkappa \circ c(x) = \phi(\sinh(a+x))$*, hence*

$$d\widetilde{w}(t) = -\phi(\sinh(a + \widetilde{w}(t))) + dw(t),$$

(b) here $\varkappa \circ c(x) = \phi\left([sign(a+x)]\left[(1-\alpha)|a+x|\right]^{1/1-\alpha}\right)$*, hence*

$$d\widetilde{w}(t) = -\phi\left([sign(a+\widetilde{w}(t))]\left[(1-\alpha)|a+\widetilde{w}(t)|\right]^{1/1-\alpha}\right) + dw(t).$$

Corollary 7 *(weak solutions) Let* $b(t)$*,* $t \in [0,T]$ *be a Brownian motion on some probability space* $(\Omega', \mathfrak{F}', \mathbb{P}')$*. Define*

$$\Lambda = \exp\left[-\int_0^T \varkappa \circ c(b(t))\,db(t) - \frac{1}{2}\int_0^T (\varkappa \circ c)^2(b(t))\,dt\right].$$

If

$$\mathbb{E}_{\mathbb{P}}\Lambda = 1,$$

then

$$x_b(t) \triangleq c(b(t))$$

is a (weak) solution of (5).

Proof. According to Girsanov theorem

$$\mathbb{P}(A) \triangleq \int_A \Lambda d\mathbb{P}',\ A \in \mathfrak{F}$$

is a probability measure, $(\Omega, \mathfrak{F}, \mathbb{P})$ is a probability space, and

$$w(t) \triangleq b(t) + \int_0^t \varkappa \circ c(b(s))\,ds$$

is a Wiener process on it. Hence

$$\begin{aligned} x_b(t) &\triangleq c(b(t)) \\ &= c\left(w(t) - \int_0^t \varkappa \circ c(b(s))\,ds\right) \\ &= c\left(w(t) - \int_0^t \varkappa(x_b(s))\,ds\right) \end{aligned}$$

on $(\Omega, \mathfrak{F}, \mathbb{P})$. ■

Remark 8 *$K-$ diffusions starting from random initial conditions can be easily obtained. Let ξ is a random variable on $(\Omega, \mathfrak{F}, \mathbb{P})$, and consider the following generalization of (5)*

$$c\left(\xi + w(t) - \int_0^t \varkappa(X_s(w))\,ds\right) = X_t(w). \tag{13}$$

If ξ is stochastically independent on $w(t)$, $t \in [0, T]$, then the solution of (13) is a $K-$ diffusions with $X_0(w) = c(\xi)$.

4 Applications

4.1 Identification of financial instruments

Consider two financial instruments. Denote their prices by X and Y. Moreover, assume that X and Y are driven by the same Wiener process and assume X is a K$-$ diffusion with c and $\varkappa$ known. How can one identify Y? There is a well known method of a "black box" identification by Norbert Wiener. However, his method is essentially restricted to systems of special kind: input and output must be observable. This is not the case in financial modelling. Here we have the black box $w \to (X_t(w), Y_t(w))$ and one may observe the output only. Hence, this method cannot be applied directly. To overcome this difficulty, observe that, if c^{-1} exists, then the mapping $w \to X_t(w)$ is invertible, and

$$w(t) = c^{-1}(X(t)) + \int_0^t \varkappa(X_s)\,ds$$

is a Wiener process, hence, the input and output of this black box

$$X \to Y_t(X) = Y_t\left(c^{-1}(X(t)) + \int_0^t \varkappa(X_s)\,ds\right)$$

is observable. Now, Wiener's method of nonlinear systems identification can be applied to the $Y-$ black box (see N. Wiener 1958, Lecture 10 and 11)

4.2 Smooth densities

Set $\widetilde{F}(x) = \mathbb{P}(\widetilde{w}(t) < x)$. Then

$$\begin{aligned} \mathbb{P}(X_t(w) < x) &= \mathbb{P}(c(\widetilde{w}(t)) < x) \\ &= \widetilde{F} \circ c^{-1}(x). \end{aligned}$$

Hence, the smoothness density problem for $X_t(w)$ is reduced to investigation of ordinary function $\widetilde{F} \circ c^{-1}$.

Example 9

$$\begin{aligned} (a)\ \widetilde{F} \circ c^{-1}(x) &= \widetilde{F}\left(\sinh^{-1}(a+x)\right), \\ (b)\ \widetilde{F} \circ c^{-1}(x) &= \widetilde{F}\left(sign(a+x)(1-\alpha)^{1-\alpha}|a+x|^{1-\alpha}\right). \end{aligned}$$

Acknowledgement 10 *We would like to thank Professor Moshe Zakai for his remarks and suggestions.*

References

Biagini F., Hu Y., Oksendal B., Zhang T., *Stochastic Calculus for Fractional Brownian Motion and Applications*, Springer, 2008.

Doss H., *Liens entre equations differentielles stochastiques et ordinaires*, "Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques" 1977 vol. 13 (2), pp. 99-125.

Isobe E., Sato S., *Wiener-Hermite expansion of a process generated by an Ito stochastic differential equation*, "Journal of Applied Probability" 1983 vol. 20 (4), pp. 754-765.

Ito K., *On stochastic differential equations*, "Memoirs American Mathematical Society" 1951 vol. 4, pp. 1- 51.

Ito K., Selected Papers, Ed. D.W. Stroock and S.R.S. Varadhan, Springer-Verlag, 1987.

Karatzas I., Shreve S.E., *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1991.

Kolmogorov A.N. , *Uber die analytischen methoden in der wahrscheinlichkeitsrechnung*, "Mathematatische Annalen", 1931 vol. 104, pp. 415-458.
Stroock D.W. , Varadhan S.R.S., *Multidimensional Diffusion Processes*, Springer-Verlag, 1979.
Stroock D.W., *Markov Processes from K. Ito's Perspective*, Princeton University Press, 2003.
Sussman H., *An interpretation of stochastic differential equations as orinary differential equations which depend on the sample point*, "Bulletin of the American Mathematical Society" 1977 vol. 83 (2), pp. 296-298.
Wiener N., *Nonlinear Problems in Random Theory*, Tech. Press of MIT, John Wiley&Sons, 1958.

Tadeusz Banek, Edward Kozłowski [1]

Coupling in dynamical systems and its probabilistic consequences

Keywords: bidirectional coupling, mutual information, Wiener-Ito homogeneous chaos expansion, Fredholm alternative

Abstract

In the literature connected with environmental sciences one can observe a growing interest in description and analysis of so-called bidirectional coupling between dynamics of distinct systems. This subject belongs to the System Science and its investigation requires advanced methods. Inspired by Information Theory and Wiener's ideas of homogeneous chaos, we propose a probabilistic approach which seems to be quite general and appropriate for this kind of problems. First, we apply the tools of Information Theory getting a quantitative characterization of coupling's strength. Secondly, we use the Wiener-Ito chaos theory and the Fredholm alternative getting coupling's characterization in the kernel expansion terms.

1 Introduction

Multidirectional coupling is one of the most important phenomenon arising in complex dynamic systems and is, in fact, a central issue of General Systems Theory. Large, complex systems consist of many smaller, simpler subsystems which operate independently, or partially independently. Sometimes, parts of the subsystems are linked rigidly and the effect can be visible by outside observers. If the evolution of one subsystem influences the behaviour of the another, then we call it a one-directional coupling. If the influence comes from the both sides, we call it bidirectional coupling. Finally, multidirectional coupling means that each subsystem is influenced by many other ones. Perhaps, the most famous example is the many-body problem in classical celestial mechanics where the influences manifest themselves in the form of gravitation forces appearing in the right side of Newton's equations

$$m_i \overset{\circ\circ}{q_i} = G \sum_{j \neq i} \frac{m_i m_j (q_j - q_i)}{\left| q_j - q_i \right|^3}, \quad i = 1, ..., n.$$

[1]Technical University of Lublin, Faculty of Management, Department of Quantitative Methods in Management, Nadbystrzycka 38, 20-618 Lublin, *e-mail: t.banek@pollub.pl , e.kozlovski@pollub.pl*

The many-body problem story is very illuminating and in particular it shows how hard an investigation of multidirectional coupling can be. Hence, there is no surprise that there are no general methods available for quantitative analysis of the coupling phenomenons.

Consider the system with two inputs w,u and two outputs x,y. In the general case each output depends on both inputs. Absence of bi-coupling is, by definition, the case when each output depends on one input only, say x depends on w and y depends on u but does not depend on w. In mathematical terms x is a $w-$ function and y is a $u-$ function. However, in the case of bidirectional coupling (bi-coupling in short), both x and y depend upon both variables w and u. A typical example are the systems described by ordinary differential equations

$$\overset{\circ}{x}(t) = f(x(t)) + \varphi(x(t),y(t)) \tag{1}$$

$$\overset{\circ}{y}(t) = g(y(t)) + \psi(x(t),y(t)) \tag{2}$$

together with the initial conditions

$$x(0) = x_0 \tag{3}$$

$$y(0) = y_0 \tag{4}$$

With the exception for the particular case $\varphi = \psi = 0$, this system exhibits a complex bidirectional coupling behaviour due to the dependence of φ and ψ on the both variables x,y. As was pointed out in M. Rutha, E. Kalnaya, N. Zenga, R. S. Franklinb, J. Rivasc and F. Miralles-Wilhelmd (2011) "neglecting bidirectional coupling renders models unable to exhibit important real-world phenomena and thus reduces usefulness as tools to explain and prepare for these phenomena". But what is the real difference between the model with bi-coupling and with the model without bi-coupling, i.e., the case $\varphi = \psi = 0$? On the phase space (x,y) we get from (1)-(2) the equation

$$\frac{dy}{dx} = \frac{g(y)) + \psi(x,y))}{f(x) + \varphi(x,y)}$$

which if it is solvable, gives the formula

$$y = \mathscr{R}(x) \tag{5}$$

But this means that

$$y(t) = \mathscr{R}(x(t)) \tag{6}$$

i.e., $y(t)$ can be expressed in terms of $x(t)$ independently of $\varphi = \psi = 0$ or $\varphi \neq \psi \neq 0$. In conclusion we see that bi-coupling is invisible on the phase space (x,y) and other methods for defining what the coupling mathematically means are needed.

As was pointed out in T. Banek (2012) this phenomenon can be effectively investigated by applying random independent signals w, u as inputs. If outputs are independent as well, then there is no deterministic relation $y = \mathscr{R}(x)$ between outputs, hence there is no bi-coupling and the complex system consists of two disjoint autonomic subsystems. The aim of this paper is to justify this claim in a general setting. Moreover, we propose a quantitative characterization of this phenomenon, which we call a strength of bi-coupling. The paper consists of two parts. In the first we apply the tools of Information Theory getting a quantitative characterization of coupling's strength. In the second we use the Wiener-Ito chaos theory and the Fredholm alternative getting coupling's characterization in the kernel expansion terms.

2 Systems given by SDE's

In order to overcome the effect expressed in (5)-(6) one can introduce into the system (1)-(2) two stochastically independent "sources of randomness". This can be done in two different ways.

The first way is accomplished by making the initial conditions (3)-(4) the random independent variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Indeed, even for the case $\varphi = \psi = 0$, and $f \equiv g$, independence assumptions of x_0, y_0 imply that the solutions $x(t), y(t)$ of (1)-(2) are stochastically independent as well. Thus the function $\mathscr{R}$ appearing in (5)-(6) does not exist any more.

The second way consists in perturbing the right hand side of (1)-(2) by adding independent random "white noise" disturbances and obtaining Ito's stochastic differential equations

$$dx(t) = [f(x(t)) + \varphi(x(t), y(t))]dt + dw(t) \tag{7}$$

$$dy(t) = [g(y(t)) + \psi(x(t), y(t))]dt + du(t) \tag{8}$$

where $(w(t), u(t))$ is a pair of independent Wiener processes. Again, even for the case $\varphi = \psi = 0$ and $f \equiv g$, the independence assumption of $w(t), u(t)$ implies that the solutions $x(t), y(t)$ of (7)-(8) are stochastically independent as well.

On the other hand, if the two "sources of randomness" are stochastically independent and φ, ψ are indeed the functions of two variables, then the solutions $x(t), y(t)$ of (7)-(8) are stochastically dependent. In this way the stochastic dependence is connected with the phenomenon of bi-coupling in dynamical systems.

2.1 Mutual information as measure of bi-coupling

Mutual information is a concept from Information Theory. It is defined by the formula

$$I_T(x, y) = \mathbb{E} \ln \frac{d\mu_{x,y}}{d[\mu_x \times \mu_y]}(x, y),$$

where $\mathbb{E}$ stands for expectation with respect to the measure $\mathbb{P}$, $\mu_{x,y}$ is a measure on the space of trajectories $\{(x(t),y(t));0\leq t\leq T\}$ being the solutions of (7)(8), while $\mu_x\times\mu_y$ is a product of marginal measures of $\mu_{x,y}$.

Theorem 1 *Let f,g,φ,ψ are Lipschitz and satisfy a linear growth condition. Then*

$$I_T(x,y)=\frac{1}{2}\int_0^T\mathbb{E}\left\{[\varphi(x(t),y(t))-\overline{\varphi}(t,x)]^2+[\psi(x(t),y(t))-\overline{\psi}(t,y)]^2\right\}dt \quad (9)$$

where

$$\begin{aligned}\overline{\varphi}(t,x) &= \mathbb{E}[\varphi(x(t),y(t))\,|x(s)\;0\leq s\leq t]\\ \overline{\psi}(t,y) &= \mathbb{E}[\varphi(x(t),y(t))\,|y(s)\;0\leq s\leq t]\end{aligned}$$

Proof. Under theorem's conditions there exists a unique strong continuous solution $(x(t),y(t),\,0\leq t\leq T$. We begin by considering the case $f=g=0$ first. For this case we have

$$\mu_{x,y} \ll \mu_w\times\mu_y \quad (10)$$

$$\mu_x \ll \mu_w,\mu_y\ll\mu_u \quad (11)$$

$$\frac{d\mu_{x,y}}{d[\mu_x\times\mu_y]}(x,y) = \frac{d\mu_{x,y}}{d[\mu_w\times\mu_u]}(x,y)\Big/\frac{d\mu_x}{d\mu_w}(x)\frac{d\mu_y}{d\mu_u}(y) \quad (12)$$

where (10) means that the measure $\mu_{x,y}$ is absolutely continuous with respect to $\mu_w\times\mu_y$, (11) means that μ_x is absolutely continuous with respect to μ_w and μ_y is absolutely continuous with respect to μ_u. Finally, (12) gives the Radon-Nikodym derivative of $\mu_{x,y}$ with respect to $\mu_x\times\mu_y$. Moreover

$$\begin{aligned}&\frac{d\mu_{x,y}}{d[\mu_w\times\mu_u]}(x,y)=\exp\{\int_0^T\varphi(x(t),y(t))dx(t)+\int_0^T\psi(x(t),y(t))dy(t)\\ &-\frac{1}{2}\int_0^T\left[\varphi^2(x(t),y(t))+\psi^2(x(t),y(t))\right]dt\}\end{aligned} \quad (13)$$

and

$$\frac{d\mu_x}{d\mu_w}(x) = \exp\left\{\int_0^T\overline{\varphi}(t,x)dx(t)-\frac{1}{2}\int_0^T[\overline{\varphi}(t,x)]^2dt\right\} \quad (14)$$

$$\frac{d\mu_y}{d\mu_u}(y) = \exp\left\{\int_0^T\overline{\psi}(t,y)dy(t)-\frac{1}{2}\int_0^T[\overline{\psi}(t,y)]^2dt\right\} \quad (15)$$

All formulae (10)-(15) follow easily from the results included in R.S. Liptser and A.N. Shiryaev (1978) (see chapter VII and XVI), so we omit their proofs. Substitution of (13)-(15) into (12) and taking expectation gives (9). In the case $f\neq 0$, $g\neq 0$ note that

$$\begin{aligned}\overline{f}(t,x) &= \mathbb{E}[f(x(t))\,|x(s)\;0\leq s\leq t]=f(x(t))\\ \overline{g}(t,y) &= \mathbb{E}[g(y(t)\,|y(s))\;0\leq s\leq t]=g(y(t))\end{aligned}$$

hence

$$\begin{aligned} f(x(t))+\varphi(x(t),y(t))-\overline{f}(t,x)-\overline{\varphi}(t,x) &= \varphi(x(t),y(t))-\overline{\varphi}(t,x) \\ g(y(t))+\psi(x(t),y(t))-\overline{g}(t,y)-\overline{\psi}(t,y) &= \psi(x(t),y(t))-\overline{\psi}(t,y) \end{aligned}$$

what ends the proof. ■

Corollary 2 *If* $\varphi=\psi=0$*, then* $I_T(x,y)=0$*.*

Corollary 3 *If*

$$\begin{aligned} \varphi(x,y) &= ax+by \\ \psi(x,y) &= Ax+By, \end{aligned}$$

then

$$\begin{aligned} I_T(x,y) = & \frac{1}{2}\int_0^T \mathbb{E}\{b^2[y(t))-\mathbb{E}[y(t)\,|x(s)\;0\le s\le t]]^2 \\ & +a^2[x(t)-\mathbb{E}[x(t)\,|y(s)\;0\le s\le t]]^2\}dt \end{aligned}$$

or

$$\begin{aligned} I_T(x,y) = & \frac{b^2}{2}\int_0^T \mathbb{E}[var(y(t)\,|x(s)\;0\le s\le t)]\,dt \\ & +\frac{a^2}{2}\int_0^T \mathbb{E}[var(x(t)\,|y(s)\;0\le s\le t)]\,dt \end{aligned}$$

Corollary 4 *Denote*

$$\begin{aligned} p(t) &= \mathbb{E}[var(y(t)\,|x(s)\;0\le s\le t)], \\ q(t) &= \mathbb{E}[var(x(t)\,|y(s)\;0\le s\le t)]. \end{aligned}$$

From Kalman-Bucy filtering theory, $(p(t),q(t))$ *are given by the Riccati equations*

$$\begin{aligned} \overset{\circ}{p}(t) &= 1+2Bp(t)-b^2p^2(t) \\ \overset{\circ}{q}(t) &= 1+2aq(t)-A^2q^2(t) \end{aligned}$$

Conclusion 5 *From the above Corollaries we see that Mutual Information is:*
(1) a non-negative quantity $I_T(x,y)\ge 0$*,*
(2) vanishes $I_T(x,y)=0$ *if there is no bi-coupling,*
(3) it is almost explicitly described in the linear systems case.

3 Chaos expansion of random signals

According to the Wiener-Ito homogeneous chaos theory (N. Wiener, 1958; D.W. Stroock, 1987; P. Malliavin, 1997) any random variable F_T defined on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ having a finite second moment $\mathbb{E}F_T^2 < \infty$ and measurable with respect to some Wiener process $w(t)$, $t \in [0,T]$ can be uniquely expressed as a limit (in the mean square sense) of the series

$$F_T = \mathbb{E}(F_T) + \sum_{n \geq 1} I_n^W(h_n) \tag{16}$$

of iterated stochastic integrals

$$I_n^W(h_n) = \int_0^T \dots \left(\int_0^{t_2} h_n(t_1, \dots, t_n)\, dw(t_1) \right) \dots dw(t_n) \tag{17}$$

which are centred

$$\mathbb{E}\left[I_n^W(h_n) \right] = 0$$

and orthogonal

$$\mathbb{E}\left[I_n^W(h_n) I_m^W(h_m) \right] = \begin{cases} 0, & n \neq m \\ \|h_n\|_2^2, & n = m \end{cases} \tag{18}$$

where

$$\|h_n\|_2 = \left(\int_0^T \dots \left(\int_0^{t_2} h_n^2(t_1, \dots, t_n)\, dt_1 \right) \dots dt_n \right)^{1/2}$$

The functions h_n, $n = 1, 2, \dots$ are called kernels of chaos expansions and are uniquely defined and square integrable on the simplexes

$$S_n[0,T] = \{(t_1, \dots, t_n); 0 < t_1 < \dots < t_n < T\}$$

In the multidimensional extension of the theory, the Wiener process appearing in (16)-(18) is multidimensional. Now we shall apply the chaos expansion theory to the situation illustrated in the diagram above. This means that for any function $c : \mathbb{R}^2 \to \mathbb{R}$, such that $\mathbb{E}[c(x(T), y(T))]^2 < \infty$, there exists a unique sequence of square integrable functions

$$(t_1, \dots, t_n; \tau_1, \dots, \tau_m) \longmapsto h_{n,m}^c \begin{pmatrix} t_1, \dots, t_n \\ \tau_1, \dots, \tau_m \end{pmatrix} \in \mathbb{R}$$

on $S_n[0,T] \times S_m[0,T]$, such that, for

$$\begin{aligned} & I_{n,m}^{W,U}\left(h_{n,m}^c\right) \\ = & \int_{S_m[0,T]} \left(\int_{S_n[0,T]} h_{n,m}^c \begin{pmatrix} t_1, \dots, t_n \\ \tau_1, \dots, \tau_m \end{pmatrix} dw(t_1) \dots dw(t_n) \right) du(\tau_1) \dots du(\tau_m) \end{aligned} \tag{19}$$

we have

$$c(x(T),y(T)) = \mathbb{E}[c(x(T),y(T))] + \sum_{n+m\geq 1} I_{n,m}^{W,U}\left(h_{n,m}^{c}\right), \tag{20}$$

where $w(t)$, $u(t)$, $t\in[0,T]$ are independent Wiener processes on $(\Omega,\mathcal{F},\mathbb{P})$. Since the kernels of chaos expansion of $c(x(T),y(T))$ depend itself on the function $c(x,y)$, we have used the subscript c in the notation $h_{n,m}^{c}$.

3.1 Bi-coupling determined by kernels of expansion

Under the notation from previous section we have

Theorem 6 *Assume $c(x,y) = c_1(x)c_2(y)$. If $w(t)$ and $u(t)$ are stochastically independent Wiener processes, then $x(t)$ and $y(t)$ are stochastically independent if and only if*

$$h_0^{c_1c_2} = h_0^{c_1}h_0^{c_2}, \tag{21}$$

where

$$\begin{aligned} h_0^{c_1c_2} &= \mathbb{E}[c_1(x(T))c_2(y(T))] \\ h_0^{c_1} &= \mathbb{E}[c_1(x(T))],\; h_0^{c_2} = \mathbb{E}[c_2(y(T))] \end{aligned}$$

and for any $n,m\geq 0$

$$h_{n,m}^{c}\begin{pmatrix} t_1,\ldots,t_n \\ \tau_1,\ldots,\tau_m \end{pmatrix} = h_n^{c_1}(t_1,\ldots,t_n)\otimes h_m^{c_2}(\tau_1,\ldots,\tau_m). \tag{22}$$

In that case

$$\begin{aligned} I_{n,m}^{W,U}\left(h_{n,m}^{c}\right) &= I_n^W\left(h_n^{c_1}\right)\otimes I_m^U\left(h_m^{c_2}\right) \\ I_n^W\left(h_n^{c_1}\right) &= \int_{S_n[0,T]} h_n^{c_1}(t_1,\ldots,t_n)\,dw(t_1)\ldots dw(t_n) \\ I_m^U\left(h_m^{c_2}\right) &= \int_{S_m[0,T]} h_n^{c_1}(\tau_1,\ldots,\tau_m)\,du(\tau_1)\ldots du(\tau_m) \end{aligned} \tag{23}$$

and

$$c(x(T),y(T)) = \mathbb{E}[c_1(x(T))]\,\mathbb{E}[c_2(y(T))] + \sum_{n+m\geq 1} I_n^W\left(h_n^{c_1}\right)\otimes I_m^U\left(h_m^{c_2}\right) \tag{24}$$

Proof. Indeed, a lack of bi-coupling and independence of $x(t)$ and $y(t)$ implies

$$c_1(x(T)) = \mathbb{E}[c_1(x(T))] + \sum_{n\geq 1} I_n^W\left(h_n^{c_1}\right) \tag{25}$$

$$c_2(y(T)) = \mathbb{E}[c_2(y(T))] + \sum_{m\geq 1} I_m^U\left(h_m^{c_2}\right) \tag{26}$$

hence

$$\begin{aligned} & c_1(x(T))\,c_2(y(T)) \\ = {} & \mathbb{E}[c_1(x(T))]\,\mathbb{E}[c_2(y(T))] + \sum_{n+m\geq 1} I_n^W(h_n^{c_1}) \otimes I_m^U(h_m^{c_2}) \\ = {} & \mathbb{E}[c_1(x(T))]\,\mathbb{E}[c_2(y(T))] + \sum_{n+m\geq 1} I_{n,m}^{W,U}(h_n^{c_1} \otimes h_m^{c_2}) \end{aligned} \tag{27}$$

Now, assume (21),(22). Let $c_2(y) = 1$. Then we must have

$$h_0^{c_2} = 1, h_m^{c_2} \equiv 0, m \geq 1$$

and, in conclusion, from (27) we get (25). Let $c_1(x) = 1$. Then obviously we have

$$h_0^{c_1} = 1, h_n^{c_1} \equiv 0, n \geq 1$$

and, in conclusion, from (27), we get (26). Since W, U are independent stochastically, then $c_1(x(T)), c_2(y(T))$ are independent as well. ■

3.2 Strength of bi-coupling

The results of the last section clearly indicate that there exists a close relationship between intensity of bi-coupling and the kernels $h_{n,m}^{c_1c_2}$ of chaos expansion of $c_1(x(T))\,c_2(y(T))$. The aim of this section is to propose a formula which reflects this relation. We begin our analysis of bi-coupling on a subspace of order (n,m) of the chaos spaces $L^2\left((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^2\right)$.

Definition 7 *Strength of bi-coupling on a subspace of order (n,m) of the chaos space is defined by*

$$S_T^{n,m}(x,y) = \inf \left\| h_{n,m}^{c_1c_2} - h_n^{c_1} \otimes h_m^{c_2} \right\|_{L^2(S_n \times S_m)}^2, \tag{28}$$

where

$$\|h_{n,m}\|_{L^2(S_n\times S_m)} = \left(\int_{S_n[0,T]\times S_m[0,T]} h_{n,m}^2 \begin{pmatrix} t_1, \ldots, t_n \\ \tau_1, \ldots, \tau_m \end{pmatrix} dt_1 \ldots dt_n d\tau_1 \ldots d\tau_m \right)^{1/2}$$

and infimum is taken over all $h_n^{c_1} \in L^2(S_n[0,T]), h_m^{c_2} \in L^2(S_m[0,T])$ of the following form

$$\begin{aligned} h_n^{c_1}(t_1, \ldots, t_n) &= \int_{S_m[0,T]} h_{n,m}^{c_1c_2} \begin{pmatrix} t_1, \ldots, t_n \\ \tau_1, \ldots, \tau_m \end{pmatrix} \psi(\tau_1, \ldots, \tau_m)\, d\tau_1 \ldots d\tau_m \\ &= \left\langle h_{n,m}^{c_1c_2}(t_1, \ldots, t_n), \psi \right\rangle_{L^2(S_m)} \\ h_m^{c_2}(\tau_1, \ldots, \tau_m) &= \int_{S_n[0,T]} \varphi(t_1, \ldots, t_n)\, h_{n,m}^{c_1c_2} \begin{pmatrix} t_1, \ldots, t_n \\ \tau_1, \ldots, \tau_m \end{pmatrix} dt_1 \ldots dt_n \\ &= \left\langle h_{n,m}^{c_1c_2}(\tau_1, \ldots, \tau_m), \varphi \right\rangle_{L^2(S_n)}, \end{aligned}$$

where

$$\varphi \in L^2(S_n[0,T]), \;\; \psi \in L^2(S_m[0,T]).$$

Remark 8 *From Definition 7 we have $S_T^{n,m}(x,y) \geq 0$ for any signals x,y. In the absence of bi-coupling we have from Theorem 6 the conclusion $S_T^{n,m}(x,y) = 0$.*

Theorem 9 *Fix $n,m \geq 0$. Then*

$$S_T^{n,m}(x,y) = \left\| h_{n,m}^{c_1 c_2} \right\|^2_{L^2(S_n \times S_m)} - \|H^{n,m}\|^2 \tag{29}$$

where $H^{n,m} = \left(H_{i,j}^{n,m}\right)$ is an infinite matrix with elements

$$H_{0,0}^{n,m} \triangleq h_0^{c_1} h_0^{c_2}$$

and

$$H_{i,j}^{n,m} \triangleq$$

$$\int_{S_n \times S_m} h_{n,m}^{,j}(t_1,...,t_n)\, h_{n,m}^{c_1 c_2}\begin{pmatrix} t_1,...,t_n \\ \tau_1,...,\tau_m \end{pmatrix} h_{n,m}^{i,}(\tau_1,...,\tau_m)\, dt_1...dt_n d\tau_1...d\tau_m,$$

where (e_i), $\left(e^j\right)$ are an orthonormal basis in $L^2(S_n[0,T])$, $L^2(S_m[0,T])$ respectively. Moreover, the infimum in (28) is achieved on the elements
$\hbar_n^{c_1}, \hbar_m^{c_2}$, such that

$$\begin{aligned} (H^{n,m})(H^{n,m})^T \hbar_n^{c_1} &= \lambda_{c_1} \hbar_n^{c_1} && (30) \\ (H^{n,m})^T (H^{n,m}) \hbar_m^{c_2} &= \lambda_{c_2} \hbar_m^{c_2} && (31) \end{aligned}$$

i.e., $\hbar_n^{c_1}, \hbar_m^{c_2}$ are eigenvectors of the matrices $(H^{n,m})(H^{n,m})^T$ and $(H^{n,m})^T(H^{n,m})$ respectively and λ_{c_1}, λ_{c_2} are the greatest eigenvalues of these matrices.

Proof. We have the expansions

$$\varphi = \sum \alpha_i e_i, \;\; \psi = \sum \beta_j e^j$$

Denote

$$\begin{aligned} h_{n,m}^{,j}(t_1,...,t_n) &= \int_{S_m[0,T]} h_{n,m} e^j d\tau_1...d\tau_m, \\ h_{n,m}^{i,}(\tau_1,...,\tau_m) &= \int_{S_m[0,T]} e_i h_{n,m} dt_1...dt_n, \end{aligned}$$

for which we have

$$\begin{aligned} h_n^{c_1}(t_1,...,t_n) &= \int_{S_m[0,T]} h_{n,m} \psi d\tau_1...d\tau_m \\ &= \sum \beta_j \int_{S_m[0,T]} h_{n,m} e^j d\tau_1...d\tau_m = \sum \beta_j h_{n,m}^{,j}(t_1,...,t_n), \end{aligned}$$

$$\begin{aligned} h_n^{c_2}(\tau_1,...,\tau_m) &= \int_{S_n[0,T]} \varphi h_{n,m} dt_1...dt_n \\ &= \sum \alpha_i \int_{S_m[0,T]} e_i h_{n,m} dt_1...dt_n = \sum \alpha_i h_{n,m}^{i,}(\tau_1,...,\tau_m). \end{aligned}$$

Thus

$$\begin{aligned} \left\| h_{n,m}^{c_1c_2} - h_n^{c_1} \otimes h_m^{c_2} \right\|^2_{L^2(S_n\times S_m)} &= \left\| h_{n,m}^{c_1c_2} - \sum_{i,j} \alpha_i \beta_j h_{n,m}^{,j} \otimes h_{n,m}^{i,} \right\|^2_{L^2(S_n\times S_m)} \\ &= \left\| h_{n,m}^{c_1c_2} \right\|^2_{L^2(S_n\times S_m)} - 2\sum_{i,j} \alpha_i H_{i,j}^{n,m} \beta_j \\ &\quad + \sum_{i,j,p,q} \alpha_i \beta_j \alpha_p \beta_q \left\langle h_{n,m}^{,j} \otimes h_{n,m}^{i,}, h_{n,m}^{,q} \otimes h_{n,m}^{p,} \right\rangle \\ &= \left\| h_{n,m}^{c_1c_2} \right\|^2_{L^2(S_n\times S_m)} - 2\alpha^T H^{n,m} \beta \\ &\quad + \sum_{i,j,p,q} \left(\alpha_i \left\langle h_{n,m}^{i,}, h_{n,m}^{p,} \right\rangle \alpha_p \right) \left(\beta_j \left\langle h_{n,m}^{,j}, h_{n,m}^{,q} \right\rangle \beta_q \right) \\ &= \left\| h_{n,m}^{c_1c_2} \right\|^2_{L^2(S_n\times S_m)} - 2\alpha^T H^{n,m} \beta + \alpha^T \mathbb{H}^{n,m} \alpha \beta^T \mathscr{H}^{n,m} \beta, \end{aligned} \tag{32}$$

where

$$\begin{aligned} \left(\mathbb{H}_{i,p}^{n,m} \right) &= \left(\left\langle h_{n,m}^{i,}, h_{n,m}^{p,} \right\rangle \right) \\ \left(\mathscr{H}_{j,q}^{n,m} \right) &= \left(\left\langle h_{n,m}^{,j}, h_{n,m}^{,q} \right\rangle \right). \end{aligned}$$

Taking infimum in (32) with respect to the sequence $\alpha = (a_i)$, we get

$$\alpha = H^{n,m} \frac{\beta}{\|\beta\|^2} \tag{33}$$

where $\beta = (\beta_j)$. Substitution (33) into (32) gives

$$\left\| h_{n,m}^{c_1c_2} \right\|^2_{L^2(S_n\times S_m)} - \left\| H^{n,m} \frac{\beta}{\|\beta\|} \right\|^2 \tag{34}$$

hence to prove (29) it is enough to show that infimum with respect to β of (34) exists. From J.B. Conway (2007, see Proposition 4.7, pp. 43) we know, that the operator

$$\left(\mathbb{H}^{n,m} f \right)(t_1,...,t_n) \triangleq \int_{S_m} h_{n,m}^{c_1c_2} \begin{pmatrix} t_1,...,t_n \\ \tau_1,...,\tau_m \end{pmatrix} f(\tau_1,...,\tau_m)\, d\tau_1...d\tau_m$$

is compact and $\|\mathbb{H}^{n,m}\|^2 \le \left\| h_{n,m}^{c_1c_2} \right\|^2_{L^2(S_n\times S_m)}$. Note that the relationship between this operator $\mathbb{H}^{n,m}$ and the matrix $H^{n,m} = \left(H_{i,j}^{n,m} \right)$, namely

$$H_{i,j}^{n,m} = \int_{S_n} e_i(t_1,...,t_n) \left(\mathbb{H}^{n,m} e^j \right)(t_1,...,t_n)\, dt_1...dt_n = \left\langle e_i, \mathbb{H}^{n,m} e^j \right\rangle_{L^2(S_n)}.$$

Since $\left(e_i \otimes e^j\right)$ is a basis for $L^2\left(S_n \times S_m\right)$ we get

$$\begin{aligned}\sum_{i,j}\left|\left\langle e_i, \mathbb{H}^{n,m} e^j\right\rangle_{L^2(S_n)}\right|^2 &= \sum_{i,j}\left|\left\langle h_{n,m}^{c_1c_2}, e_i \otimes e^j\right\rangle_{L^2(S_n\times S_m)}\right|^2 \\ &= \sum_{i,j}\left|H_{i,j}^{n,m}\right|^2,\end{aligned}$$

what says that matrix $H^{n,m}$ is the representation of $\mathbb{H}^{n,m}$ in basis $\left(e_i \otimes e^j\right)$ and $\|\mathbb{H}^{n,m}\| = \|H^{n,m}\|$. Hence, taking infimum with respect to β of (34), we get (29). Now we are going to show that $\|\mathbb{H}^{n,m}\|$ is the greatest eigenvalue of $\mathbb{H}^{n,m}$. For $\mathbb{H}^{n,m} = 0$ it is obvious, so assume $\|\mathbb{H}^{n,m}\|^2 > 0$ is not the eigenvalue of $\left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)$. Since $\left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right) : L^2\left(S_m\right) \to L^2\left(S_m\right)$ is compact and self-adjoint hence by the Fredholm alternative an operator

$$\left(\|\mathbb{H}^{n,m}\|^2 I - \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)\right)^{-1} : L^2\left(S_m\right) \to L^2\left(S_m\right) \tag{35}$$

exists and it is continuous. Thus, there exists $\alpha > 0$, such that for $x \in L^2\left(S_m\right) \setminus \{0\}$ we have

$$\left\|\|\mathbb{H}^{n,m}\|^2 \frac{x}{\|x\|} - \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)\frac{x}{\|x\|}\right\| \geq \alpha$$

hence

$$\inf\left\{\left\|\|\mathbb{H}^{n,m}\|^2 e - \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)e\right\|; \|e\| = 1\right\} \geq \alpha$$

On the other hand, if $\|e\| = 1$, then

$$\begin{aligned}& 2\|\mathbb{H}^{n,m}\|^2\left(\|\mathbb{H}^{n,m}\|^2 - \|\mathbb{H}^{n,m}e\|^2\right) \\ =\ & 2\|\mathbb{H}^{n,m}\|^2\left(\|\mathbb{H}^{n,m}\|^2 - \left\langle e, \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)e\right\rangle\right) \\ =\ & \|\mathbb{H}^{n,m}\|^4 - 2\|\mathbb{H}^{n,m}\|^2\left\langle e, \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)e\right\rangle + \|\mathbb{H}^{n,m}\|^4 \\ \geq\ & \|\mathbb{H}^{n,m}\|^4 - 2\|\mathbb{H}^{n,m}\|^2\left\langle e, \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)e\right\rangle + \left\|\left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)e\right\|^2 \\ =\ & \left\|\|\mathbb{H}^{n,m}\|^2 e - \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)e\right\|^2 \geq \alpha^2.\end{aligned}$$

Thus

$$\begin{aligned}0 &= \|\mathbb{H}^{n,m}\|^2 - \|\mathbb{H}^{n,m}\|^2 = \|\mathbb{H}^{n,m}\|^2 - \sup\left\{\|\mathbb{H}^{n,m}e\|^2; \|e\| = 1\right\} \\ &= \inf\left\{\|\mathbb{H}^{n,m}\|^2 - \|\mathbb{H}^{n,m}e\|^2; \|e\| = 1\right\} \geq \frac{\alpha}{2\|\mathbb{H}^{n,m}\|^2} > 0,\end{aligned}$$

what is impossible. Hence $\|\mathbb{H}^{n,m}\|^2$ is the eigenvalue of $\left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)$. To see that it is the greatest, note that

$$\lambda = \frac{\left\langle x, \left(\mathbb{H}^{n,m}\right)^T\left(\mathbb{H}^{n,m}\right)x\right\rangle}{\|x\|^2} \leq \|\mathbb{H}^{n,m}\|^2$$

for any eigenvalue. We denote this eigenvalue by λ_{c_2} and the corresponding eigenvector by $\hbar_m^{c_2}$. This finish the proof of (31). For (30) the reasoning is similar. ■

Having defined what we mean by the strength of bi-coupling on a subspace of order (n,m) of the chaos space $L^2\left((\Omega,\mathcal{F},\mathbb{P}),\mathbb{R}^2\right)$, we now give the general definition.

Definition 10 *Strength of bi-coupling between x,y is defined by*

$$S_T(x,y) = \left|h_0^{c_1c_2} - h_0^{c_1}h_0^{c_2}\right| + \sum_{n+m\geq 1} S_T^{n,m}(x,y). \tag{36}$$

Conclusion 11 *>From (28)(29) follows that for all $n,m \geq 0$*

$$\left\|H^{n,m}\right\|^2 \leq \left\|h_{n,m}^{c_1c_2}\right\|^2_{L^2(S_n\times S_m)}$$

hence

$$0 \leq S_T(x,y) \leq \sum_{n,m\geq 0} \left\|h_{n,m}^{c_1c_2}\right\|^2_{L^2(S_n\times S_m)} = \mathbb{E}\left[c_1(x(T))\,c_2(y(T))\right]^2 < \infty. \tag{37}$$

Conclusion 12 *According to the famous results of Norbert Wiener included in his book (1958, see Chapter 10 and 11), the kernels $h_{n,m}^c$ of the chaos expansion can be find due to the orthogonality relations (18). The identification process is described in this book in details. Having the kernels identified one may apply the results included in Theorems 6 and 9 to settle the bi-coupling question. Further results can be used for numerical evaluation of the phenomena.*

References

Banek T., Coupling in dynamical systems and its probabilistic consequences, "24th Int. Conf. on Sys. Res. Infrmatics & Cybernetics", July 30-August 3, Baden Baden, Germany, 2012.

Conway J.B., A Course in Functional Analysis, Springer, Berlin, 2007.

Lax P.D., Functional Analysis, Wiley-Interscience, New York, 2002.

Liptser R.Sh., Shiryaev A.N., Statistics of Stochastic Processes, Springer-Verlag, New York, 1978.

Malliavin P., Stochastic Analysis, Springer-Verlag, New York, 1997.

Stroock D.W., Homogeneous chaos revisited, Seminaire de Probabilites XXI. Lecture Notes in Mathematics 1247, Springer, Berlin 1987, pp. 1-8.

Rutha M., Kalnaya E., N. Zenga N., Franklinb R. S., Rivasc J., Miralles-Wilhelmd F., Sustainable prosperity and societal transitions: Long-term modeling for anticipatory management, "Environmental Innovation and Societal Transitions" 2011 vol.1 (1), pp. 160-165.

Wiener N., Nonlinear Problems in Random Theory, Technology Press Research Monographs MIT and J. Wiley & Sons, N.Y, Chapman & Hall, London, 1958.

Przemysław Kowalik[1]

On an implementation of the method of capacity of information bearers (the Hellwig method) in spreadsheets

Keywords: econometric model, independent variables, dependent variables, information bearers, covariance, spreadsheets.

Abstract

A method of selecting independent variables in an econometric model known as the method of capacity of information bearers (the Hellwig method) is a very simple concept. Unfortunately, its computational complexity grows exponentially along with the number of potential independent variables what can make the method practically unusable unless appropriate software is used. However, formulas used in the method, at least in the form most often presented in literature are not "uniform" and that is why their literal implementation into computer software may encounter remarkable difficulties. In particular, spreadsheets, despite a large number of built-in statistical features, are apparently not a good tool for some calculations in the field, including those required by the Hellwig method. This paper shows a simple transformation of the formulas used in the Hellwig method which makes an implementation of the method in spreadsheets very easy. It is compatible with most popular spreadsheet programs (both commercial and free) and appropriate for both educational purposes and analysis of real-world econometric models. The implementation was illustrated with a simple example with real-world data.

1 Introduction – key ideas of the Hellwig method

The method of capacity of information bearers (also called the method of optimal choice of predictors or, after its author, the Hellwig method) is of one methods of selecting independent variables for an econometric model. As many others, the method consists in selecting such variables which are strongly correlated with the dependent variables and, simultaneously, weakly correlated with other independent variables (Hellwig 1968, 1969). The selection of variables requires "testing" all of $L = 2^k - 1$ combinations of k potential independent variables ("zero" combination i.e. rejecting all the variables is not considered). The following notation (based on Nowak 1994, p.23) will be used:

---

[1] Lublin University of Technology, Faculty of Management, Department of Quantitative Methods in Management, Nadbystrzycka 38, 20-618 Lublin, *e-mail: p.kowalik@pollub.pl*

- l – number of a combination ($l = 1,2,...,L$);
- k_l – number of variables in the l^{th} combination;
- j – number of a variable in the l^{th} combination ($j = 1,2,...,k_l$);
- r_j – correlation of the j^{th} independent variable with the dependent variable;
- r_{ij} – correlation of the j^{th} independent variable with other independent variables included in the l^{th} combination $i = 1,2,...,k_l$, $i \neq j$;

The *individual capacity of information bearer* (later referred to as *individual capacity of information*) h_{lj} for the j^{th} independent variable ($j = 1,2,...,k_l$) in the l^{th} combination $l = 1,2,...,L$ is defined as

$$h_{lj} = \frac{r_j^2}{1 + \sum_{i=1, i \neq j}^{k_l} \left| r_{ij} \right|}.$$

The *integral capacity of information bearer* (later referred to as *integral capacity of information*) for the l^{th} combination is the sum of the abovementioned individual capacities of information bearers for the l^{th} combination:

$$h_l = \sum_{j=1}^{k_l} h_{lj}.$$

The combination of independent variables for which the maximum of integral capacity is attained is chosen to the econometric model.

Both individual and integral capacities of information are normalized i.e. they are included in the [0;1] interval. They increase as the independent variables are strongly correlated with the dependent variable and the independent variables are weakly correlated one to another.

2 Transforming the formulas into a "spreadsheet-friendly" form

Any algorithm or formula which is claimed to be useful in practical application must be considered in connection with some fundamental questions. How to perform calculations: with what software and hardware, at what cost, how long will those calculations last? Without answering those questions a theoretically very useful method may turn out to be of no use to a practitioner.

When creating econometric models, spreadsheet can be used as easily available, well-known and rich of built-in statistical features. However, some calculations in the field including the Hellwig method are rather hard to perform in spreadsheets if appropriate formulas are implemented literally i.e. as exact equivalents of the mathematical notation. Due to rules of using spreadsheets, all the necessary formulas, whose numbers grows exponentially along with k – the number of potential independent variables, must be placed simultaneously in spreadsheet cells. Moreover, formulas for individual capacities of information are not "uniform" because of combination-dependent indexes of summation. It means that the spreadsheet equivalents of the mathematical formulas presented earlier cannot be placed into cells by copying of some "initial formula" or by using array formulas. In other words, if one follows the definition literally, the formula for each individual capacity of information should be placed "manually" in a cell what is time-consuming and may result in many mistakes. However, it is possible to transform formulas for individual capacities of information into a more "spreadsheet-friendly" form. The following additional notations are needed.

- R_Y^2 – vector (or matrix $[1\times k]$) of squared correlations of independent variables with the dependent variable (elements are denoted by r_j^2);
- R_A – matrix of absolute values of correlations of independent variables
- C_k – matrix ($[L\times k]$) of zeroes and ones which is a numerical description of all the possible combinations of choice $1,2\ldots,k$ independent variables out of k potential independent variables.
- c_{li} – entries of C_k where l (the number of a row) stands also for the number of a combination (1 on position i stands for selecting the variable X_i and 0 for rejecting the variable);
- H_I - matrix ($[L\times k]$) of individual capacities of information;
- H - 2^k-1-element vector of integral capacities of information h_l.

The entries of H_I - individual capacities of information for the l^{th} combination can be expressed by "uniform" formulas

$$h_{lj}=\frac{c_{lj}r_j^2}{\sum_{i=1}^{k}c_{li}\left|r_{ij}\right|},\qquad j=1,2,\ldots,k$$

The number „1" from the definition was replaced with $|r_{ii}|$ which are all obviously equal to 1. The main difference to compare with the definition is that selection of correlations which "occur" in an individual capacity of information is done by multiplication by 0-1 coefficients of the combination, not by the combination-dependent index of summation.

The above transformation is almost identical with that already presented in Kowalik 2010 and 2012. The only difference is that in this paper c_{lj} coefficients are placed in formulas for h_{lj}, not for h_l.

3 Entering the formulas to a spreadsheet

After introducing the transformations defined in the previous chapter, now it is possible to present consecutive steps of implementation of the Hellwig method in a spreadsheet. The steps will be illustrated with a simple real-world example concerning influence of five factors on production volumes of construction aggregates (Bajorek, Kiernia-Hnat, Skrzypczak 2012). It is worth notifying that the authors of the quoted paper mention explicitly "non-uniform" formulas as one of the disadvantages of practical usage of the Hellwig method. This paper is just an attempt to get rid of this disadvantage.

The steps of implementation are the following.

1) Calculating R_Y^2 - the vector of squared correlations of the dependent variable with independent variables.
2) Calculating R_A - the matrix of absolute values of correlations of independent variables.
3) Generating C_k - the matrix of 0-1 combinations.
4) Calculating H_I - the matrix of individual capacities of information.
5) Calculating H - the matrix of integral capacities of information.
6) Finding the combination matching the maximal integral capacity of information.

	A	B	C	D	E	F
1	**Y**	**X1**	**X2**	**X3**	**X4**	**X5**
2	4,696	13885	151,1	10485,5	216,5	0,46
3	4,800	14456	160,3	10635,2	216,3	0,45
4	5,349	15680	490,5	12335,8	220,2	0,45
5	5,167	16776	193,7	11607,5	226,9	0,44
6	6,972	13359	193,7	11968,8	239,1	0,40

Fig. 1. Values of variable placed in an Excel 2010 spreadsheet

Source: Bajorek, Kiernia-Hnat, Skrzypczak 2012, p. 15

The values of the dependent variable are located in A2:A6 and the values of the independent variables in B2:F8 (see Fig.1). Formulas in cells are grouped according to the steps mentioned earlier. English names of spreadsheet functions are used. Comma is used as the decimal separator and semicolon as the separator of arguments of spreadsheet functions. Screenshots were made in Microsoft Excel 2010. The compatibility was tested in the following spreadsheet software: Microsoft Excel 2002 to 2010, OpenOffice.org 3.2.0, LibreOffice 4.01, Gnumeric 1.10.16 and IBM Lotus Symphony 3.0.1. No array formulas were used.

In descriptions of formulas the phrase *the first cell of the range* stands for the top-left cell of the range under consideration (in case of single-row ranges it is the left-most cell and in case of single-column ranges it is the top-most cell).

Ad. 1. R_Y^2 - the vector of squared correlations of the dependent variable with independent variables.

Data from the example

Enter =CORREL($A2:$A8;B2:B8)^2 into the cell B10. Next, copy B10 to C10:D10.

General rule

Content of the first cell of the range (entered by the user)

=CORREL (range for the dependent variable; range for the independent variable X_1)^2

The first argument of CORREL must have the $ signs at letters and the second argument must not have the $ signs at letters. Entering the $ signs at numbers (in any argument) does not matter. Behind the closing bracket there must be symbols for the power operation i.e. "^2".

Copying

Copy the cell with the above formula right to $k-1$ adjacent cells.

Ad 2. R_A - the matrix of absolute values of correlations of independent variables

Data from the example

Enter =ABS(CORREL(B$2:B$6;OFFSET($B$2:$B$6;0;$A12-1))) into the cell B12. Next copy B12 to B12:F16.

Auxiliary data from the example

Numbers 1,2,3,4,5 were entered in A12:A16 (the column left to B12:F16 – the range for the correlation matrix) as well as in B11:F11 (the row above B12:F16). The numbers in A12:A16 are used to create references necessary to „extract" particular columns by using the OFFSET function. The numbers in B11:F11 are not actually used in calculations – they are just added for better readability.

General rule

It is necessary to prepare some auxiliary data in the spreadsheet. The first one is a vertical range of numbers $1,2,...,k$ increasing downward (preferably located just left to the range of the correlation matrix). The second one is a horizontal range of numbers $1,2,...,k$ increasing rightward (preferably located just over the range of the correlation matrix). The horizontal range is not actually used in calculations – it is just added for better readability.

Content of the first cell of the range (entered by the user)

=ABS(CORREL(range for the independent variable X_1 ;OFFSET(range for the independent variable X_1 ;0; first cell of the vertical auxiliary range -1)))

The first argument of CORREL must have the $ signs at numbers only. The first argument of OFFSET must have the $ signs at both letters and numbers. Finally, the third argument of OFFSET i.e. the address of the first cell of the vertical auxiliary range must have the $ sign at the letter only. References to all the pairs of vectors of the values of independent variables are created by changes of cell addresses while copying. The second argument of CORREL is created by using the OFFSET function. The OFFSET function in the considered context returns the reference to the range whose size (numbers of rows and colums) is identical to that of the first argument of the function. The relative location with respect to the first cell of the first argument ("offset") is defined by the second argument ("vertical offset", constantly equal to 0 here) and the third argument ("horizontal offset", which is calculated by values of the vertical auxiliary range minus one and equal to 0,1,…, $k-1$). Absolute values of correlations are calculated by using the ABS function

Copying

Copy the cell with the above formula right to $k-1$ adjacent cells. The resulting *k*-element row range copy down to $k-1$ adjacent rows. Finally, the resulting range must have *k* rows and *k* columns.

	A	B	C	D	E	F
8				squared correlations (Y, Xi)		
9		0,141301263	0,004608817	0,400957265	0,886992099	0,935953235
10				absolute values of correlations (Xi, Xj)		
11		1	2	3	4	5
12	1	1	0,385105025	0,333908334	0,185048016	0,343302868
13	2	0,385105025	1	0,732522539	0,092935914	0,133269615
14	3	0,333908334	0,732522539	1	0,588821307	0,513833684
15	4	0,185048016	0,092935914	0,588821307	1	0,968164749
16	5	0,343302868	0,133269615	0,513833684	0,968164749	1

Fig. 2. Values of correlations calculated by formulas (defined in items 1 and 2) in an Excel 2010 spreadsheet

Source: own calculations

Ad. 3. C_k - the matrix of 0-1 combinations.

Data from the example

Enter =INT(MOD($I19/(2^(J$18-1));2)) into the cell J19. Next, copy J19 to J19:N49.

Auxiliary data from the example

The numbers 1,2,3,4,5 were entered in J18:N18 (the row above J19:N49 – the range for the 0-1 matrix). The numbers 1,2,…31 were entered in I19:I49 (the column left to J19:N49)

General rule

It is necessary to prepare some auxiliary data first. The first one is a vertical range of numbers $1,2,...,2^k-1$ (i.e. numbers of combinations) increasing downward (preferably located just left to the range of the 0-1 matrix). The second one is a horizontal range of numbers $1,2,...,k$ (numbers of variables) increasing rightward.

"Generating" all the 0-1 combinations representing all the possible combinations of k independent variables is equivalent to binary notation of numbers $l=1,2,...,2^k-1$ (k digits with leading zeroes e.g. for $k=5$ $9_{10}=01001_2$). The binary representation of a number is calculated by dividing remainders of division of $l/2^{k-1},l/2^{k-2},...,l/2^1,l/2^0$ by 2 and rounded the results down to integers. Remainders are calculated with the MOD function and rounded with the INT function.

The results (binary representations) are returned in the reversed order e.g. 9 is 10010 instead of 01001. The reason is that using the reversed order of divisions ($l/2^0,l/2^1,...,l/2^{k-2},l/2^{k-1}$) results in simplification of spreadsheets formulas. The order of digits in binary representations does not really matter, however, as long as all the representations and, what follows, all the combinations are calculated.

Content of the first cell of the range (enter by the user).

=INT(MOD(first cell of the auxiliary vertical range/(2^(first cell of the auxiliary vertical range-1));2))

The first cell address in the first argument of MOD must have the $ sign at the letter and the second cell address must have the $ sign at the number.

Copying

Copy the cell with the above formula right to $k-1$ adjacent cells. The resulting k-element row range copy down to 2^k-2 adjacent rows. Finally, the resulting range must have 2^k-1 rows and k columns.

Ad. 4. H_I - the matrix of individual capacities of information.

Data from the example

Enter =J19*B$9/MMULT($J19:$N19;B$12:B$16) into the cell B19. Next, copy B19 to B19:F49.

General rule

The matrix is an array of quotients. Dividends are the elements of vector R_Y^2 multiplied by corresponding entries of a row of the 0-1 combination matrix C_k. Divisors are the entries of the matrix product C_kR_A, i.e. sums of products of rows of C_k and columns of R_A. They are calculated by using the MMULT function (formally each of those sums of products is the matrix product of a row matrix and a column matrix).

Content of the first cell of the range (entered by the user)

=first cell of the 0-1 combination matrix=1*first cell of the range of squared correlations of the dependent variable with independent variables/MMULT(first row of the 0-1 combination matrix; first column of the matrix of absolute values of correlations);0)

The address of the first cell of the 0-1 combination matrix must not have the $ signs at all. The address of the first cell of the range of squared correlations must have the $ signs at numbers only. The first argument of MMULT must have the $ signs at the letters only. The second argument of MMULT must have the $ signs at the numbers only.

Ad 5. *H* - the matrix of integral capacities of information

Data from the example

Enter =SUM(B19:F19) into the cell G19. Next, copy G19 to G20:G49.

General rule

Integral capacities of information are calculated as sums of individual capacities of information for each combination.

Content of the first cell of the range (entered by the user)

=SUM(first row of the matrix of individual capacities of information)

No $ signs required.

Copying

The cell with the above formula should be copied down to $2^k - 2$ adjacent cells.

Ad. 6. Finding the combination matching the maximal integral capacity of information

Data from the example

Enter =MAX(G19:G49) into the cell G50.

Enter =MATCH(G50; G19:G49;0) into the cell G51. No copying required.

General rule

The maximum of integral capacities of information is found by using the MAX function, whose argument is the range of all the integral capacities of information calculated in Item 5. The MATCH function finds the number of the best combination by looking for the position of the maximal integral capacity of information in the array of all the integral capacities of information.

Content of cells (entered by the user)

=MAX(range of all the integral capacities of information)

=MATCH(address of the cell with MAX described above; range of all the integral capacities of information;0). Zero as the third argument stands for exact matching when looking for the value from the first argument).

The final result of calculations is shown in Fig. 3.

	A	B	C	D	E	F	G
17							
18			individual capacities of information				integr. cap. of inform.
19	1	0,141301263	0	0	0	0	0,141301263
20	2	0	0,004608817	0	0	0	0,004608817
21	3	0,102014837	0,003327413	0	0	0	0,105342251
22	4	0	0	0,400957265	0	0	0,400957265
23	5	0,105930265	0	0,30058832	0	0	0,406518585
24	6	0	0,002660177	0,231429754	0	0	0,234089931
25	7	0,082199049	0,002176406	0,194033718	0	0	0,278409173
26	8	0	0	0	0,886992099	0	0,886992099
27	9	0,119236741	0	0	0,748486211	0	0,867722952
28	10	0	0,004216914	0	0,811568261	0	0,815785175
29	11	0,089992032	0,003118193	0	0,694055753	0	0,787165979
30	12	0	0	0,25236146	0,558270521	0	0,810631981
31	13	0,09302523	0	0,208535437	0,50003238	0	0,801593046
32	14	0	0,002524745	0,172726357	0,527419825	0	0,702670927
33	15	0,074210456	0,002084906	0,151005343	0,475139067	0	0,702439772
34	16	0	0	0	0	0,935953235	0,935953235
35	17	0,10518943	0	0	0	0,696755183	0,801944613
36	18	0	0,004066832	0	0	0,825887523	0,829954355
37	19	0,081752267	0,003035362	0	0	0,633868805	0,718656434
38	20	0	0	0,264862164	0	0,618266885	0,883129048
39	21	0,084247746	0	0,21699851	0	0,50397653	0,805222787
40	22	0	0,002470166	0,178492289	0	0,568241977	0,749204433
41	23	0,068515808	0,002047547	0,155393858	0	0,470232283	0,696189496
42	24	0	0	0	0,45066964	0,475546184	0,926215823
43	25	0,092453418	0	0	0,4119389	0,404917304	0,909309622
44	26	0	0,003758601	0	0,430348752	0,445387803	0,879495156
45	27	0,073846104	0,002860291	0	0,394894651	0,382844104	0,85444515
46	28	0	0	0,190690944	0,346889689	0,377096626	0,91467726
47	29	0,075876259	0	0,164558524	0,323479605	0,331275548	0,895189936
48	30	0	0,002352964	0,141422278	0,334723856	0,357880423	0,836379521
49	31	0,062874215	0,001966359	0,126521427	0,312875305	0,316353152	0,820590458
50						max	0,935953235
51						combin. with max	16

Fig. 3. Values of capacities of information (calculated by formulas defined in items 4 and 5) and the final result – the maximal integral capacity of information and the number of the corresponding combination (calculated by formulas defined in item 6) in an Excel 2010 spreadsheet

Source: own calculations

The maximal value of integral capacity of information is equal to 0.935953235. This value is attained for combination 16 (00001) what means selecting X_5. The result of calculations in Bajorek, Kiernia-Hnat, Skrzypczak 2012 – 0.913 (what, according to the authors, means selecting X_1, X_4) - is incorrect. It might have been caused by mistakes in calculations of results of complicated “original” formulas.

4 Summary

A method of implementation of the Hellwig method in spreadsheets described in this paper can be efficiently used for educational purposes but also for small sized real world problems. Obviously, spreadsheet software has many restrictions connected with the maximal sizes of single portions of data which can be processed in one file. Nevertheless, the presented approach may be interesting for many researchers because spreadsheets are widely available (also as free software) and relatively easy to use. The cited example shows how using a proper computational tool for can help to avoid mistakes which may happen when processing even relatively small-size input data. Additionally, the computational methods presented in the paper are of much wider importance than the Hellwig method itself. Namely, spreadsheet are missing the easy way of calculating correlation or covariance matrices and the concept presented here is an answer to this disadvantage. Easy method of generating 0-1 combinations may also be useful for many users.

Bibliography

Bajorek G., Kiernia-Hnat M., Skrzypczak I., *Aspekty środowiskowe w technologii produkcji kruszyw budowlanych*, Prace Instytutu Ceramiki i Materiałów Budowlanych, Rok V, Nr 9, 2012, pp.7-19.

Hellwig Z., *On the Optimal Choice of Predictors*, [in:] Study VI, Toward a System of Quantitative Indicators of Components of Human Resources Development, UNESCO, Paris 1968.

Hellwig Z., *Problem optymalnego wyboru predykant*, Przegląd Statystyczny, R.XVI, 3-4, 1969.

Kowalik P, *Implementacja metody wskaźników pojemności informacyjnej (metody Hellwiga) w arkuszach kalkulacyjnych*, [in:] Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne, Tom 2, Wydawnictwo Wyższej Szkoły Handlowej, Kielce 2011, pp. 186-194.

Kowalik P., *Wykorzystanie arkuszy kalkulacyjnych do wyboru zmiennych objaśniających przy pomocy metody wskaźników pojemności informacyjnej (metody Hellwiga)*, [in:] Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne, Tom 2/2012, Wydawnictwo Wyższej Szkoły Handlowej, Kielce 2012, pp. 168-178.

Nowak E., *Zarys metod ekonometrii. Zbiór zadań*, Wydawnictwo Naukowe PWN, Warszawa 1994.

Edward Kozłowski [1]

Optimal route and control for the LQG problem

Keywords: optimal route, linear quadratic control, navigation, landmark

Abstract

This paper presents a routing problem of linear system with Gaussian disturbances. The linear quadratic control problem was reduced to determining the optimal trajectory , which must be tracked by linear system.The general aim of optimal route determining consist in minimization of composite cost function. Moreover, it is compared the optimal route for linear system with trajectory, on which the system was controlled optimally. To illustrate those paths a numerical example is included.

1 Introduction

Finding the shortest (or the cheapest/fastest) route is a very important class of optimization problems because of their practical applicability in many fields of human activity. Among the most commonly known examples there are the following: the shortest path problem (SSP), the travelling salesman problem (TSP), the vehicle routing problem (VRP) and their extensions. They consist in visiting some or all nodes in a network when minimizing the total weight of arcs connecting the visited nodes. All those problems are linear and deterministic, nevertheless even with such - often very restrictive - assumptions, they may be hard to solve. The deterministic features are not only fixed weights assigned to arcs (distances, unit costs, travel times) but also fixed locations of nodes and availability of arcs. The stochastic shortest path problem with recourse (SSPPR) is an example of a problem with randomly available arcs. Going further, if an object is affected by random disturbances, we consider optimal routing without any specific nodes and arcs. Optimal routing is performed then by finding some route, called a trajectory, which is represented by its "geometrical" shape rather than by a "logical" sequence of nodes.

The control, navigation, stabilization, learning, identyfication, cost minimization etc. problems for various systems are widely considered in literature. (see

[1]Technical University of Lublin, Faculty of Management, Department of Quantitative Methods in Management, Nadbystrzycka 38, 20-618 Lublin, *e-mail: e.kozlovski@pollub.pl*

e.g. M. Aoki, 1967; S. Azuma, M.S. Sakar and G.J. Papas, 2012; T. Banek and E. Kozłowski, 2005, 2006; Z. Bubnicki, 2000; Y. Chena, T. Edgarb and V. Manousiouthakisa, 2004; F. Kozin, 1972; E. Kozłowski, 2011; G.N. Saridis, 1995; J. Zabczyk, 1996). Each of these tasks is characterized by necessity of controlling the object in order to achieve some aim. These tasks are connected with some optimization tasks. By solving optimization tasks the control law of the system can be determined in the explicit or-non explicit form. As a result, we can control the object to perform the control aim. Sometimes, in order to achieve the aim the system should be moved after a certain path (trajectory). (see e.g. S. Azuma, M.S. Sakar and G.J. Pappas, 2012; B. Yuan, M. Orlowska and S. Sadig, 2007; M.S. Machmoud, 2011; M.K. Mainali, K. Shimada, S. Mabu and K. Hirasawa, 2008). Thus, the problem arises when there are many guide paths. Which of these trajectories is optimal? In this way, we have the problem of system navigating, where first we must determine the landmarks and next we must lead the system in such a way to mimic these marks (points). The tasks of optimal control and route are dual ones. Namely, if the control law is known, then we can specify the path after where you want the object to move, and if it is know the optimal trajectory then we can control so that the object imitated it.

In the case considered in the paper, a problem of movement and control of linear systems is presented whose main optimality criterion is to minimize the overall cost. Obviously, for a fixed horizon the energy (control) and landmarks are evenly distributed out over time.The general aim of paper is to present the problem of determining the optimal trajectory (path/route) for controlled system (the object, e.g a robot). To solve above problem this paper exploits an idea of dynamic programming. The solution of the presented task gives the optimal trajectory as a set of landmarks (statemarks). Additionally the optimal control for a LQC problem is presented. The numerical simulation for a simple linear system shows that the differences between the optimal route and the simulated path for optimal control are negligible.

The paper is organized as follows. In section II the linear quadratic routing problem is formulated and the idea of conversion from control to navigation is outlined. Next, the solution of routing problem is provided in section III. The section IV presents an example, which illustrates the controls and track for a dual tasks.

2 Linear quadratic routing problem

Sometimes for dynamical systems it is better to determine the optimal route (path) instead controls. Next the system must be controlled so as to follow a designated path. Therefore, the task consists in determining the optimal trajectory, after which we want to move the system with minimal cost. In this case the routing means determining the set of points (marks, landmarks), which must be tracked by system to satisfy the aim. The objective function represents the total costs, which are the sum of control costs and costs associated with not to hit the point (target). Let $(\Omega, \mathscr{F}, P)$ be a complete probability space. Suppose that $w_1, w_2, ...$ are independent n-dimensional random vectors on this space, with normal $N(0, I_n)$ distribution. We assume that all the above mentioned objects are stochastically independent and an initial state $\|y_0\| < \infty$.

Let the stochastic linear system is described by a state equation

$$y_{i+1} = y_i - Bu_i + C\xi + \sigma w_{i+1}, \tag{1}$$

where $i = 0, ..., N-1$, $y_i \in \mathbb{R}^n$, $B \in \mathbb{R}^{n\times k}$, $C \in \mathbb{R}^{n\times l}$ and $\sigma \in \mathbb{R}^{n\times n}$. Below we assume that the parameters of linear system $\xi \in \mathbb{R}^k$ are unknown and has a normal distribution $N(m, Q)$. On probability space $(\Omega, \mathscr{F}, P)$ we define a family of sub-σ-fields $Y_j = \sigma\{y_i : i = 0, 1, ..., j\}$.

The classical aim of control consists in optimization of performance criterion. Let for the time i the value $u_i^T R u_i$ presetnts a cost of control and the value $y_\tau^T Q y_\tau$ prezents a heredity function as losses (add costs) associated with not to hit to target. For the linear quadratic control problem the aim of control is to minimize the total cost, which is a sum of costs and losses. Then, the task is to find

$$\inf_{u\in U} E\left\{\sum_{i=0}^{N-1} u_i^T R u_i + y_N^T Q y_N\right\}, \tag{2}$$

where Y_j-measurable vector $u_j \in \mathbb{R}^l$ is called a control action, and $u = (u_0, u_1, ...)$ an admissible control and the class of admissible controls is denoted by U.

The main aim is to move the system from state y_0 to state origin coordinates $col\,(0, 0, ..., 0)$. The system should be carried out (controlled) at the cheapest cost. On the other hand, when we need to determine an optimal route, then the task (2) should be formulated in a slightly different form. Let $det\left(B^T B\right) \neq 0$. When we want to move the system (1) from state y_i to y_{i+1}, $i = 0, 1, ..., N-1$ then the control has a form

$$u_i = -\left(B^T B\right)^{-1} B^T\left(y_{i+1} - y_i - C\xi - \sigma w_{i+1}\right). \tag{3}$$

Thus, the task (2) may be replaced by

$$\inf_{y\in Y} E\sum_{i=0}^{N-1}\left[\left(y_{i+1} - y_i - C\xi - \sigma w_{i+1}\right)^T K\left(y_{i+1} - y_i - C\xi - \sigma w_{i+1}\right) + y_N^T Q y_N\right], \tag{4}$$

where

$$K = B\left(B^T B\right)^{-1} R\left(B^T B\right)^{-1} B^T.$$

We see, that the objective function of task (4) represents the total cost, which is composed of costs of transformation the system (1) along the trajectory $y_0, y_1, ..., y_N$

and cost of losses associated with not to hit to target. Thus the task (4) consist in determinig the optimal path on wich the total cost is least.

3 Optimal route determining

By solving the task (4) we obtain a set of admissible points (marks) $y=(y_0,...,y_{N-1})$ for which the infimum is attained. The sequence $\{y_i\}_{0\le i\le N}$ presents a route (optimal path, trajectory), after which the system (1) should move. The theorem below presents the method of determining of optimal route, which must be tracked by a system.

Theorem 1 *Let*

$$\begin{aligned} A_j &= K-K^T\left(K+A_{j+1}\right)^{-1}K, &(5)\\ L_j &= KC-K^T\left(K+A_{j+1}\right)^{-1}\left(KC+L_{j+1}\right), &(6)\\ M_j &= \left(M_{j+1}+C^TKC\right)+\left(KC+L_{j+1}\right)^T\left(K+A_{j+1}\right)^{-1}\left(KC+L_{j+1}\right), &(7)\\ Z_j &= Z_{j+1}+tr\left(A_{j+1}H_j\right)-tr\left(C^TKC\Sigma_j\right)-2tr\left(L_{j+1}\left(\Sigma_j-\Sigma_{j+1}\right)C^T\right)\\ &\quad -tr\left(\left(KC+L_{j+1}\right)^T\left(K+A_{j+1}\right)^{-1}\left(KC+L_{j+1}\right)\Sigma_j\right). &(8)\end{aligned}$$

where $A_N=Q$, L_N, M_N *are matrix zero,* $Z_N=0$ *and*

$$\begin{aligned}\Sigma_j &= E\left(\left(\xi-E\left(\xi|F_j\right)\right)\left(\xi-E\left(\xi|F_j\right)\right)^T\Big|Y_j\right),\\ H_j &= C\Sigma_jC^T+\sigma\sigma^T.\end{aligned}$$

If $\det\left(K+A_{i+1}\right)\neq 0$ *for* $i=0,1,...,N-1$ *then the optimal state (mark, position) for the time* $j+1$ *based on information available to time* j *is*

$$E\left(y_{j+1}\middle|Y_j\right)=\left(K+A_{j+1}\right)^{-1}\left(Ky_j+\left(KC+L_{j+1}\right)E\left(\xi\middle|Y_j\right)\right) \tag{9}$$

and the total cost is

$$\inf_{y\in Y}E\left[\sum_{i=0}^{N-1}\left(y_{i+1}-y_i-C\xi-\sigma w_{i+1}\right)^TK\left(y_{i+1}-y_i-C\xi-\sigma w_{i+1}\right)\right.$$
$$\left.+y_N^TQy_N\right]=W_0\left(y_0\right), \tag{10}$$

where

$$\begin{aligned} W_N\left(y_N\right) &= y_N^TQy_N, &(11)\\ W_i\left(y_i\right) &= y_j^TA_jy_j+2y_j^TL_jE\left(\xi\middle|Y_j\right)+E\left\{\xi^TM_j\xi\middle|Y_j\right\}+Z_j. &(12)\end{aligned}$$

Proof. First we define the Bellmann's function as

$$W_N(y_N) = y_N^T Q y_N$$

and

$$\begin{aligned} W_i(y_i) &= \min_{y_{i+1}} E\left\{ (y_{i+1} - y_i - C\xi - \sigma w_{i+1})^T K (y_{i+1} - y_i - C\xi - \sigma w_{i+1}) \right. \\ &\quad \left. + W_{i+1}(y_{i+1}) \middle| Y_i \right\} \end{aligned} \tag{13}$$

for $j = 0, 1, ..., N-1$. At time $N-1$ from (13) we have

$$\begin{gathered} W_{N-1}(y_{N-1}) = \min_{y_N} E\left\{ y_N^T (K+Q) y_N - 2y_N^T K (y_{N-1} + C\xi + \sigma w_N) \right. \\ \left. + 2y_{N-1} KC\xi + \xi^T C^T KC\xi \middle| Y_{N-1} \right\} + y_{N-1}^T K y_{N-1} + tr\left(\sigma^T K \sigma\right) \\ = \min_{y_N} \left\{ E\left(y_N^T \middle| Y_{N-1} \right) (K+Q) E\left(y_N | Y_{N-1} \right) \right. \\ \left. -2E\left(y_N^T \middle| Y_{N-1} \right) K \left(y_{N-1} + CE\left(\xi | Y_{N-1}\right)\right) \right\} + tr\left((K+Q) H_{N-1}\right) \\ +2y_{N-1} KCE\left(\xi | Y_{N-1}\right) + tr\left(\sigma^T K \sigma\right) - 2tr\left(C^T KC\Sigma_{N-1} + \sigma^T K \sigma\right) \\ + E\left\{ \xi^T C^T KC\xi \middle| Y_{N-1} \right\} + y_{N-1}^T K y_{N-1}. \end{gathered}$$

The expected optimal state (position, mark) at time N based on information available to time $N-1$ *is*

$$E\left(y_N | Y_{N-1}\right) = (K+Q)^{-1} K \left(y_{N-1} + CE\left(\xi | Y_{N-1}\right)\right)$$

and

$$\begin{gathered} W_{N-1}(y_{N-1}) = y_{N-1}^T \left(K - K^T (K+Q)^{-1} K \right) y_{N-1} \\ +2y_{N-1}^T \left(I - K^T (K+Q)^{-1} \right) KCE\left(\xi | Y_{N-1}\right) \\ +E\left\{ \xi^T C^T \left(K + K^T (K+Q)^{-1} K \right) C\xi \middle| Y_{N-1} \right\} + tr\left((K+Q) H_{N-1}\right) \\ -tr\left(C^T K^T (K+Q)^{-1} KC\Sigma_{N-1} \right) - 2tr\left(C^T KC\Sigma_{N-1}\right) - tr\left(\sigma^T K \sigma\right) \\ = y_{N-1}^T A_{N-1} y_{N-1} + 2y_{N-1}^T L_{N-1} E\left(\xi | Y_{N-1}\right) + E\left\{ \xi^T M_{N-1} \xi \middle| Y_{N-1} \right\} + Z_{N-1}. \end{gathered}$$

We assume, that equation (12) is true for $i+1$. From (12) - (13) and the properties of condition expectation we have

$$
\begin{gathered}
W_j(y_j) = \min_{y_{j+1}} E\left\{ (y_{j+1} - y_j - C\xi - \sigma w_{j+1})^T K (y_{j+1} - y_j - C\xi - \sigma w_{j+1}) \right. \\
\left. + y_{j+1}^T A_{j+1} y_{j+1} + 2 y_{j+1}^T L_{j+1} E(\xi | Y_{j+1}) + E\{\xi^T M_{j+1} \xi | Y_{j+1}\} + Z_{j+1} | Y_j \right\} \\
= \min_{y_{j+1}} \left\{ E\left(y_{j+1}^T \middle| Y_j\right) (K + A_{j+1}) E\left(y_{j+1} \middle| Y_j\right) \right. \\
\left. -2E\left(y_{j+1}^T \middle| Y_j\right) \left(K y_j + (KC + L_{j+1}) E(\xi | Y_j)\right) \right\} \\
+ E\left\{ \xi^T \left(M_{j+1} + C^T K C\right) \xi \middle| Y_j \right\} + tr((K + A_{j+1}) H_j) + y_j^T K y_j \\
+ 2 y_j K C E(\xi | Y_j) + Z_{j+1} + tr\left(\sigma^T K \sigma\right) \\
-2tr\left(C^T K C \Sigma_j + \sigma^T K \sigma\right) - 2tr\left(L_{j+1} (\Sigma_j - \Sigma_{j+1}) C^T\right).
\end{gathered}
$$

Thus, the expected optimal state (position) at time $j+1$ is

$$
E\left(y_{j+1} \middle| Y_j\right) = (K + A_{j+1})^{-1} \left(K y_j + (KC + L_{j+1}) E(\xi | Y_j)\right)
$$

and finally

$$
\begin{gathered}
W_j(y_j) = y_j^T K y_j + E\left\{ \xi^T \left(M_{j+1} + C^T K C\right) \xi \middle| Y_j \right\} + 2 y_j K C E(\xi | Y_j) \\
- \left(K y_j + (KC + L_{j+1}) E(\xi | Y_j)\right)^T (K + A_{j+1})^{-1} \left(K y_j + (KC + L_{j+1}) E(\xi | Y_j)\right) \\
+ Z_{j+1} + tr((K + A_{j+1}) H_j) - 2tr\left(C^T K C \Sigma_j\right) - 2tr\left(L_{j+1} (\Sigma_j - \Sigma_{j+1}) C^T\right) \\
- tr\left(\sigma^T K \sigma\right) = y_j^T \left(K - K^T (K + A_{j+1})^{-1} K\right) y_j \\
+ 2 y_j^T \left(KC + K^T (K + A_{j+1})^{-1} (KC + L_{j+1})\right) E(\xi | Y_j) \\
+ E\left\{ \xi^T \left(\left(M_{j+1} + C^T K C\right) + (KC + L_{j+1})^T (K + A_{j+1})^{-1} (KC + L_{j+1}) \right) \xi \middle| Y_j \right\} \\
+ Z_{j+1} - tr\left(\sigma^T K \sigma\right) + tr((K + A_{j+1}) H_j) - 2tr\left(C^T K C \Sigma_j\right) \\
-2tr\left(L_{j+1} (\Sigma_j - \Sigma_{j+1}) C^T\right) - tr\left((KC + L_{j+1})^T (K + A_{j+1})^{-1} (KC + L_{j+1}) \Sigma_j \right) \\
= y_j^T A_j y_j + 2 y_j^T L_j E(\xi | F_j) + E\left\{ \xi^T M_j \xi \middle| Y_j \right\} + Z_j,
\end{gathered}
$$

what finish the proof. ■

Remark 2 *The equation (9) gives formula (recipe, rule) how to determine the optimal route (state- or land-marks) for time* $j+1$ *if the system (1) to time j traveled the way (path, track)* $y_0, ..., y_j$.

Remark 3 *The random vector* ξ *in system (1) has a normal* $N(m, \Sigma)$. *From the theory of filtering conditionally normal sequences (see R.Sh. Liptser and*

A.N. Shiryaev, 1978), the conditional distribution $P\left(d\xi\left|Y_j\right.\right)$ *has a normal distribution* $N\left(m_j,\Sigma_j\right)$, *where the conditional expectation* $m_j = E\left(\xi\left|Y_j\right.\right)$ *and the conditional covariance matrix* $\Sigma_j = E\left([\xi - m_j][\xi - m_j]^T\left|Y_j\right.\right)$ *are given by formulas*

$$\begin{aligned} m_j &= \left(I + \Sigma\sum_{i=0}^{j-1} C^T\left(\sigma\sigma^T\right)^{-1} C\right)^{-1} \\ &\quad\times\left(m + \Sigma\sum_{i=0}^{j-1} C^T\left(\sigma\sigma^T\right)^{-1}[y_{j+1} - y_j + Bu_j]\right) \end{aligned} \tag{14}$$

and

$$\Sigma_j = \left(I + \Sigma\sum_{i=0}^{j-1} C^T\left(\sigma\sigma^T\right)^{-1} C\right)^{-1}\Sigma. \tag{15}$$

4 Linear quadratic control problem

Let us consider a linear system, which is described by the state equation (1). The optimal control of linear system (1) for the task (2) contains in follow

Theorem 4 *If the matrices R and Q are positive-definite,*
$\det\left(R + B^T Q_{j+1} B\right) \neq 0$ *and* Q_j *is defined as*

$$Q_j = Q_{j+1} - Q_{j+1}B\left[R + B^T Q_{j+1} B\right]^{-1} B^T Q_{j+1},$$

for any $j = 0, \ldots, N-1$, *where* $Q_N = Q$, *then*
a). the optimal control of stochastic system (1) is

$$u_j^* = \left[R + B^T Q_{j+1} B\right]^{-1} B^T Q_{j+1}\left[y_j + (N-j)BE\left(\xi\left|Y_j\right.\right)\right], \tag{16}$$

b). the value of performance criterion is

$$\begin{aligned} \inf_{u\in U} E\left\{\sum_{i=0}^{N-1} u_i^T R u_i + y_N^T Q y_N\right\} &= \sum_{j=0}^{N-1} tr\left(C^T Q_{j+1} C\Sigma_j\right) \\ + \left(y_0 + NBE\left(\xi\left|Y_0\right.\right)\right)^T Q_0\left(y_0 + NBE\left(\xi\left|Y_0\right.\right)\right) &+ \sum_{j=1}^{N} tr\left(\sigma^T Q_j \sigma\right) \\ + \sum_{j=0}^{N-1}(N-j+1)(N-j-1)\, tr\left(C^T Q_{j+1} C\left(\Sigma_j - \Sigma_{j+1}\right)\right), \end{aligned} \tag{17}$$

where for any $j = 0, \ldots, N-1$ *matrices* Σ_j *is defined by (15).*

Proof of this theorem we can find in E. Kozłowski, 2010.

5 Numerical example

Let us determine the optimal route and controls for a linear system with state equation

$$y_{i+1} = y_i - Bu_i + \sigma w_{i+1}, \tag{18}$$

where the initial state y_0 is $(53;32)$ and the fixed horizon $N = 10$. This system must be moved to origin coordinates. Let us assume $Q = I$ and

$$R = \begin{bmatrix} 2 & 0.4 \\ 0.3 & 1.5 \end{bmatrix}, \; B = \begin{bmatrix} 7 & 0.53 \\ 0.5 & 9 \end{bmatrix}, \; \sigma = \begin{bmatrix} 0.4 & 0.02 \\ 0.02 & 0.6 \end{bmatrix}.$$

Remark 5 *The optminal route (trajectory, set of ladmarks) for the system (18) is*

$$E\left(y_{j+1} \middle| Y_j\right) = \left(K + A_{j+1}\right)^{-1} K y_j \tag{19}$$

$j = 0,1,...N-1$, *where* A_j *is defined as (5)*

Remark 6 *When we want to plane a trajectory (route, path) at time* $t = 0$ *then we must determine optimal route conditioned on* σ*−field* Y_0

$$E\left(y_{j+1} \middle| Y_0\right) = \left(K + A_{j+1}\right)^{-1} K E\left(y_j \middle| Y_0\right)$$

or in dynamical form

$$E\left(y_j \middle| Y_0\right) = \left(\left(K + A_{j+1}\right)^{-1} K\right)^j y_0.$$

Remark 7 *When the optimal route for the linear system is known and calculated as (19) then from (18) the expected control conditioned on* σ*−field* Y_j *is*

$$\begin{aligned} \tilde{u}_j &= E\left(u_j \middle| Y_j\right) = -\left(B^T B\right)^{-1} B^T \left(E\left(y_{j+1} \middle| Y_j\right) - y_j\right) \\ &= \left(B^T B\right)^{-1} B^T \left(I - \left(K + A_{j+1}\right)^{-1} K\right) y_j. \end{aligned} \tag{20}$$

For this case the route planning $E\left(y_j \middle| F_0\right)$, simulated states y_j and landmarks $E\left(y_{j+1} \middle| Y_j\right)$ (expected optimal states conditioned on information to time j), optimal controls u_j are given in the table 1. We see that the route planning, simulated states, landmarks are very close to each other, the differences between them are negligible.

The table 2 presents the values of optimal controls u_j^* and the expected controls $\tilde{u}_j$ for $j = 0,1,...,9$. The optimal controls are calculated from (16), but the expected controls are calculated from (20). Also, we see that the optimal and expected controls are similar.

Table 1. The similation of states y_j, trace planning $E\left(y_j \middle| Y_0\right)$ and landmarks $E\left(y_j \middle| Y_{j-1}\right)$ for linear system

j	$E\left(y_j \middle| Y_0\right)$	y_j	$E\left(y_j \middle| Y_{j-1}\right)$
0	$(53,32)$	$(53,32)$	——
1	$(42.46,36.00)$	$(46.56,32.44)$	$(42.46,36.00)$
2	$(33.58,36.96)$	$(39.98,33.12)$	$(37.65,34.66)$
3	$(26.26,35.52)$	$(34.37,32.26)$	$(32.56,33.15)$
4	$(20.32,32.22)$	$(28.11,30.85)$	$(28.32,30.19)$
5	$(15.57,27.47)$	$(22.48,27.59)$	$(23.15,26.54)$
6	$(11.74,21.65)$	$(17.10,23.43)$	$(18.35,21.38)$
7	$(8.552,15.03)$	$(12.67,18.94)$	$(13.43,15.61)$
8	$(5.711,7.893)$	$(8.756,13.47)$	$(8.93,9.399)$
9	$(2.907,0.499)$	$(4.788,6.386)$	$(4.155,1.405)$
10	$(0.114,0.011)$	$(0.367,-0.24)$	$(0.202,0.118)$

Table 2. The optimal controls u_j^* and expected controls $\tilde{u}_j$ for linear system

j	u_j^*	$\tilde{u}_j$
0	$(0.920,-0.172)$	$(1.545,-0.531)$
1	$(0.894,-0.044)$	$(1.299,-0.319)$
2	$(0.855,0.095)$	$(1.064,-0.063)$
3	$(0.819,0.219)$	$(0.851,0.183)$
4	$(0.757,0.351)$	$(0.675,0.442)$
5	$(0.692,0.455)$	$(0.540,0.659)$
6	$(0.619,0.547)$	$(0.461,0.843)$
7	$(0.571,0.639)$	$(0.457,1.035)$
8	$(0.558,0.710)$	$(0.558,1.309)$
9	$(0.605,0.663)$	$(0.605,0.663)$

6 Conclusion

In this article, the optimal track problem of stochastic discrete-time linear system with quadratic objective function for fixed horizon was presented. The described problem is an the idea of conversion from control to navigation of linear system. The tasks of optimal control and optimal route are dual. To determine optimal trajectory the algorythm dynamic programming was used. As a result we have a set of landmarks. To perform aim the system (robot, object) must track the optimal path (trajectory).

The extension of described problem can be used, for example, to the source seeking problem, route determinig with reach information, plannig navigation, perfect tracking etc.

References

Aoki, M., Optimization of Stochastic Systems. Academic Press, 1967.

Azuma S., Sakar M.S., Pappas G.J., Stochastic Source Seeking by Mobile Robots. "IEEE Transactions on Automatic Control", 2012 vol. 57 (9), pp. 2308-2321.

Banek T., Kozłowski E., Adaptive control of system entropy. "Control and Cybernetics", 2006 vol. 35 (2), pp. 279-289.

Banek T., Kozłowski E., Active and passive learning in control processes application of the entropy concept, "Systems Sciences", 2005 vol. 31 (2), pp. 29-44.

Bellman R., Adaptive Control Processes, Princeton, 1961.

Bubnicki Z., General approach to stability and stabilization for a class of uncertain discrete non-linear systems, "International Journal of Control", 2000 vol. 73 (14), pp. 1298-1306.

Chena Y., Edgarb T., Manousiouthakisa V., On infinite-time nonlinear quadratic optimal control, "Systems and Control Letters", 2004 vol.51 (3-4), pp. 259 - 268.

Yuan B., Orlowska M., Sadig S., On the Optimal Robot Routing Problem in Wireless Sensor Networks, "IEEE Transactions on knowledge and data engineering", 2007 vol. 19 (9), pp. 1252-1261.

Fleming W. H., Rishel R., Deterministic and Stochatic Optimal Control, Springer-Verlag, Berlin, 1975.

Kozin F., Stability of stochastic dynamical systems, "Lecture Notes in Mathematics", 1972 vol. 294, pp. 186-229.

Kozłowski E., The linear quadratic stochastic optimal control problem with random horizon at finite number of events intependent of state system, "Systems Science", 2010 vol. 36 (3), pp. 5-11.

Kozłowski E., Identyfication of linear system in random time, "International Journal of Computer and Information Technology", 2011 vol. 1 (2), pp. 103-108.

Liptser R.Sh., Shiryaev A.N., Statistics of Stochastic Processes, Springer-Verlag, New York, 1978.

Machmoud M.S. Design of control strategies for robust dynamic routing in traffic networks, "IET Control Theory Applications", 2011 vol. 5 (15), pp. 1716 – 1728.

Mainali M.K., Shimada K., Mabu S., Hirasawa K., Optimal route of road networks by dynamic programming, "IEEE International Joint Conference on Neural Networks", Hong Kong 2008, pp. 3416 - 3420.

Saridis G.N., Stochastic processes, estimation and control: the entropy approach, John Wiley and Sons, 1995.

Zabczyk J., Chance and decision, Scuola Normale Superiore, Pisa, 1996.

Witold Rzymowski[1], *Agnieszka Surowiec*[2]

Analysis of average exchange rates

Keywords: exchange rates, linear autoregression, absolute relative error, periodic components, normal distribution, gamma distribution.

Abstract

The quantitative analysis of the average exchange rates is discussed in this article. The study was undertaken in order to establish the existence of possible quantitative regularities in the observed variables. The distributions of the average exchange rates of the chosen currency as well as the distributions of the random disturbances of first order autoregression models are analyzed. The results for these models are compared to the results for the model in which the regression coefficient equals 1. The existence of the periodic components in the random disturbances of each model is also analyzed.

1 Introduction

Studying of financial markets has a long tradition and continues to arouse much excitement. In this paper the average exchange rates in years 2003-2011 (2281 observations) of ten chosen currencies published by the National Bank of Poland such as: the US dollar (USD), the euro (EUR), the Danish krone (DKK), the Hong Kong dollar (HKD), the Japanese yen (JPY), the rand (ZAR), the Russian ruble (RUB), the Swiss franc (CHF), the pound sterling (GBP), the International Monetary Fund (XDR) in relation to the Polish zloty (PLN) are analyzed (NBP Archive 2013). The 1st order autoregressive model:

$$b_t = \beta b_{t-1} + \varepsilon_t,\ t = 1,2,...,2280 \tag{1}$$

where b_t is t^{th} observation of the average exchange rate of chosen currency in relation to PLN, β is the regression coefficient and ε_t is a random disturbance for the t^{th} observation, is used in this analysis.

[1] Lublin University of Technology, Faculty of Fundamentals of Technology, Department of Applied Mathematics, Nadbystrzycka 38, 20-618 Lublin, *e-mail: w.rzymowski@pollub.pl*
[2] Lublin University of Technology, Faculty of Management, Department of Quantitative Methods in Management, Nadbystrzycka 38, 20-618 Lublin, *e-mail: a.surowiec@pollub.pl*

Three versions of model (1) for ten chosen currencies will be discussed: M1, M2, M3. In the first version M1, the parameter β is estimated with the most important method of the Least Squares (LSM), in the second one, M2 – with the method of the Least Absolute Relative Error Estimation (LAREEM), whereas in the third one, M3, we accept arbitrarily $\beta = 1$. The rests of the latter model correspond to the daily quotation rises of the average exchange rates for chosen currencies.

2 Least Absolute Relative Error Estimation Method (LAREEM)

The Least Absolute Relative Error Estimation Method has been used since relatively recently (Ashar and Wallace 1963, Hyb and Kaleta 2004, Chen et al. 2010). A short description of the theoretical basis of this method for linear model of the form:

$$Y_t = \alpha_1 X_{1,t} + \alpha_2 X_{2,t} + \ldots + \alpha_p X_{p,t} + \varepsilon_t,$$

for $t = 1,2,\ldots N$ where $Y_t > 0$, we published in Rzymowski and Surowiec 2011. In the case of model (1) in the form

$$Y_t = \beta X_t + \varepsilon_t \qquad (2)$$

with the dependent variable Y_t and with one explanatory variable X_t it is possible to create explicit formulae for the regression coefficient β and the relative error δ_t.

For the variables X_t, Y_t taking positive values only, for $t = 1,2,\ldots N$, we are looking for such a parameter $\beta \in R$, which minimizes the size of the relative error

$$\delta_t \overset{def}{=} \left| \frac{Y_t - \beta X_t}{Y_t} \right| = \left| 1 - \frac{X_t}{Y_t} \beta \right|.$$

Speaking a little more, we are looking for such a parameter $\beta^* \in R$ which minimizes the size of

$$\max_{t=1,2,\ldots N} \left| 1 - \frac{X_t}{Y_t} \beta \right|.$$

Therefore we are looking for such a parameter $\beta^* \in R$, for which

$$\max_{t=1,2,\ldots N} \left| 1 - \frac{X_t}{Y_t} \beta^* \right| = \min_{\beta \in R} \max_{t=1,2,\ldots N} \left| 1 - \frac{X_t}{Y_t} \beta \right|.$$

It is possible to prove, that the unknown parameter $\beta^* \in R$ of model (2) is given by the formula

$$\beta^* = \frac{2}{\chi^+ + \chi^-}, \tag{3}$$

where

$$\chi^- = \min_{t=1,2,\ldots N} \frac{X_t}{Y_t}, \ \chi^+ = \max_{t=1,2,\ldots N} \frac{X_t}{Y_t}, \tag{4}$$

and the relative percentage error of model (2) is equal to

$$\delta^* = \frac{\chi^+ - \chi^-}{\chi^+ + \chi^-} 100\%. \tag{5}$$

In addition to this, for all $t = 1,2,\ldots,N$, we have

$$-\delta^* \le \frac{Y_t - \beta X_t}{Y_t} 100 \le \delta^*,$$

$$\min_{t=1,2,\ldots N} \frac{Y_t - \beta^* X_t}{Y_t} 100 = -\delta^* \text{ and } \max_{t=1,2,\ldots N} \frac{Y_t - \beta^* X_t}{Y_t} 100 = \delta^*.$$

3 Results

The results of our research concern the analysis of models of the average exchange rates of the chosen currencies in versions M1, M2, M3, the analysis of distributions of the average exchange rates as well as the distributions of the random disturbances of first order autoregression models of the average exchange rates in versions M1, M2, M3. We also present the results concerning the harmonic analysis of the random disturbances of first order autoregression models of the average exchange rates.

3.1 Models

Table 1 presents the appropriate regression coefficients for M1 and M2 models and relative error percentage for M1, M2 and M3 models for the chosen currencies in years 2003-2011.

The regression coefficients and relative percentage errors for models M2 for the chosen currencies in years 2003-2011 were obtained with the LAREEM, using patterns (3), (4) and (5). On the basis of the data from Table 1 one can conclude that the relative percentage errors for the uncomplicated M3 model are comparable to M1 model. The relative errors for M2 models for all analyzed currencies are not more than 9%, and the relative errors in the models M1 and M3 are larger on average by 0.5% from M2 models errors. In addition one can observe that the relative errors in the models M1 and M3 are close to each other with an accuracy of 0.005% on average.

Table 1. Regression coefficients for models M1 and M2 and relative percentage errors for models M1, M2 and M3 in years 2003-2011

Currencies	β_{M1}	β_{M2}	δ_{M1}	δ_{M2}	δ_{M3}
1 USD	0.9999	0.9940	6.97%	6.33%	6.98%
1 EUR	1.0000	0.9960	4.70%	4.28%	4.70%
1 DKK	1.0000	0.9965	4.68%	4.32%	4.68%
1 HKD	0.9999	0.9934	7.01%	6.32%	7.03%
100 JPY	1.0001	1.0042	9.32%	8.95%	9.33%
1 ZAR	0.9999	0.9936	6.18%	5.51%	6.19%
1 RUB	0.9999	0.9995	4.94%	4.90%	4.95%
1 CHF	1.0001	0.9872	8.13%	6.73%	8.11%
1 GBP	0.9999	0.9971	4.81%	4.51%	4.82%
1 XDR	1.0000	0.9986	4.58%	4.44%	4.59%

Source: own elaboration

Figure 1a presents the nature of the changes over time of the regression coefficients and Figure 1b presents nature of the changes over time of the relative percentage errors for M1, M2, M3 models for USD. One can see that the changes of the values of regression coefficients and the changes of the values of the relative percentage errors for M1 and M3 models are comparable in the entire analyzed period 2003-2011 (see Table 1) as well as in subsequent years of this period (see Figure 1).

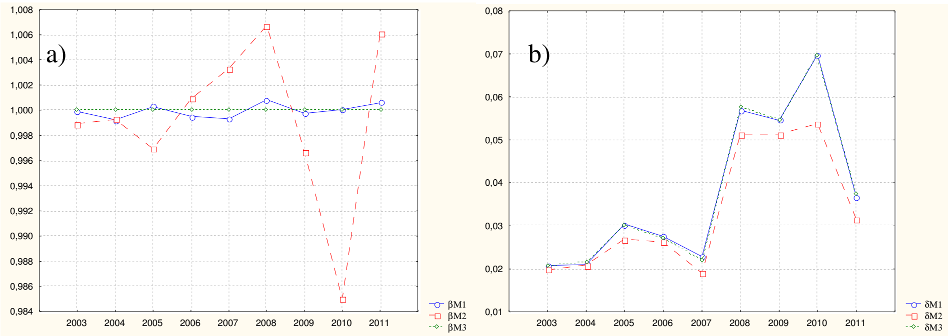

Figure 1. Nature of the changes over time in years 2003-2011 of a) regression coefficients b) relative errors for M1, M2, M3 models for USD.

Source: own elaboration

Looking at Figure 1 a) and 1 b) one can see the coincidence in changes of the regression coefficients of M2 model and the relative percentage errors. This coincidence occurs for all the analyzed currencies. The biggest change of values of the regression coefficients of M2 model and the relative percentage errors has occurred in 2008 for the analyzed currencies with the exception of the JPY for which this change has occurred earlier, namely in 2007. One can see it in Figure 2 a), 2 b). The noticeable change of value of the regression coefficient of the M2 model confirms the collapse of the value of exchange rates for all analyzed currencies in 2008.

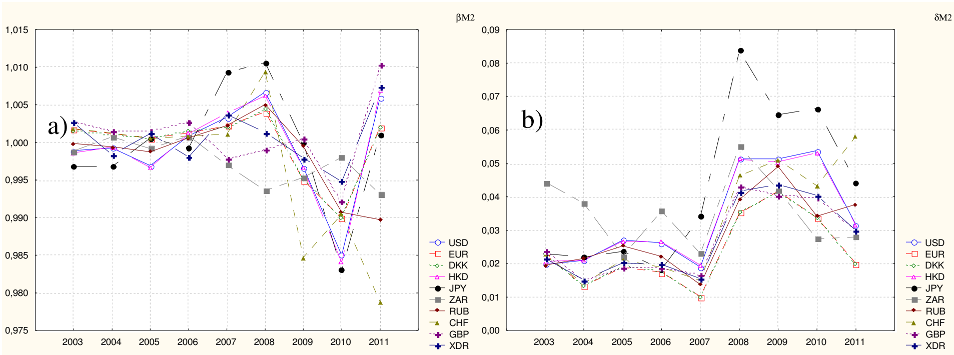

Figure 2. Value of a) regression coefficients b) relative errors for M2 models for ten currencies in subsequent years of the period 2003-2011

Source: own elaboration

3.2 Distributions

In this section we present the results concerning the distributions of the average exchange rates of the ten chosen currencies as well as the results concerning the distributions of the random disturbances of first order autoregression models in versions M1, M2, M3 of the average exchange rates.

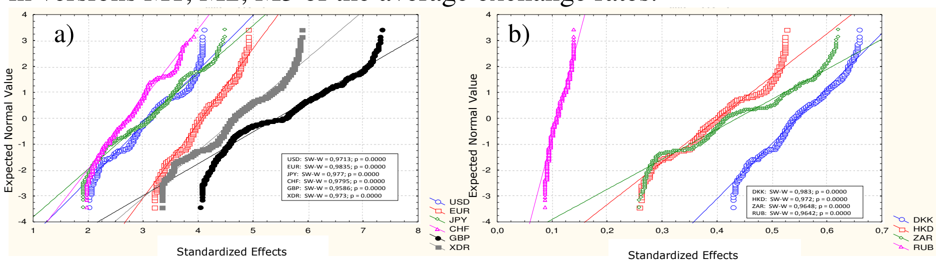

Fig. 3. Probability Plot for normal distribution for average exchange rates in years 2003-2011 for a) USD,EUR,JPY,CHF GBP,XDR b) DKK,HKD,ZAR,RUB with Shapiro-Wilk Test

Source: Own elaboration

On the basis of the Probability Plot for the normal distribution for the average exchange rates in years 2003-2011 presented in Figure 3 one can conclude that none of the distributions of the average exchange rate is a normal distribution during that period.

The analysis of distributions of random disturbances of the M1, M2 and M3 models for USD in years 2003-2011 is presented in Figure 4 a) and 4 b). One can conclude that the distributions of random disturbances of M1, M2 and M3 models are very similar and they are not the normal distributions.

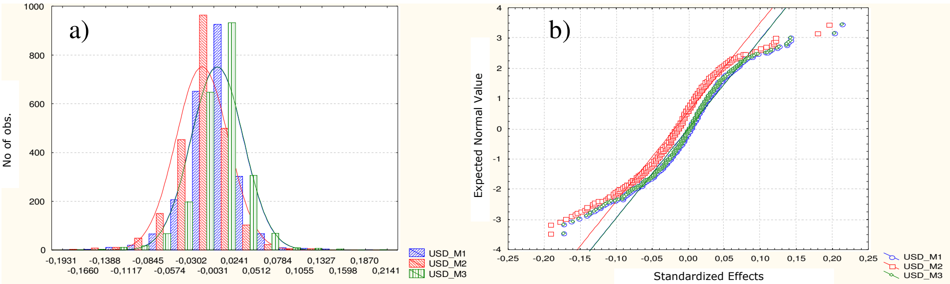

Fig. 4. A) Frequency Distribution with normal distribution b) Probability – Probability Plot for normal distribution for M1, M2 and M3 model for average exchange rates for USD in years 2003-2011

Source: own elaboration

The Figures 5, 6 and 7 present the Probability-Probability Plots. In these figures the x-axis represents Theoretical Cumulative Distribution Function, the y-axis represents Observed Cumulative Distribution Function.

The results presented in Figure 5 and 6 relate to the chosen average exchange rates and the positive values of sequences of residuals for M2 models only. We also analyzed the sequences of residuals for M1 and M3 models.

One can see in Figure 5 and 6 that the distributions of the average exchange rates of the chosen currencies as well as the distributions of positive values of the random disturbances of first order autoregression models of the average exchange rates in versions M2 in the entire period 2003-2011 (see Figure 5) and in subsequent years in period 2003-2011 considered separately for USD only (see Figure 6) have the distribution close to Gamma distribution.

Furthermore, one can assume that the Gamma distribution for positive values of random disturbances of M2 model persists in time, but Gamma distribution for average exchange rate for USD has not this property. We observed these properties not only for average exchange rate for ten chosen currencies and for positive values of the random disturbances of M2 models of the average exchange rates of ten chosen currencies but also for negative values of the random disturbances of first order autoregression models of the average exchange rates taken with a minus sign and for the absolute values for these residuals without zero value.

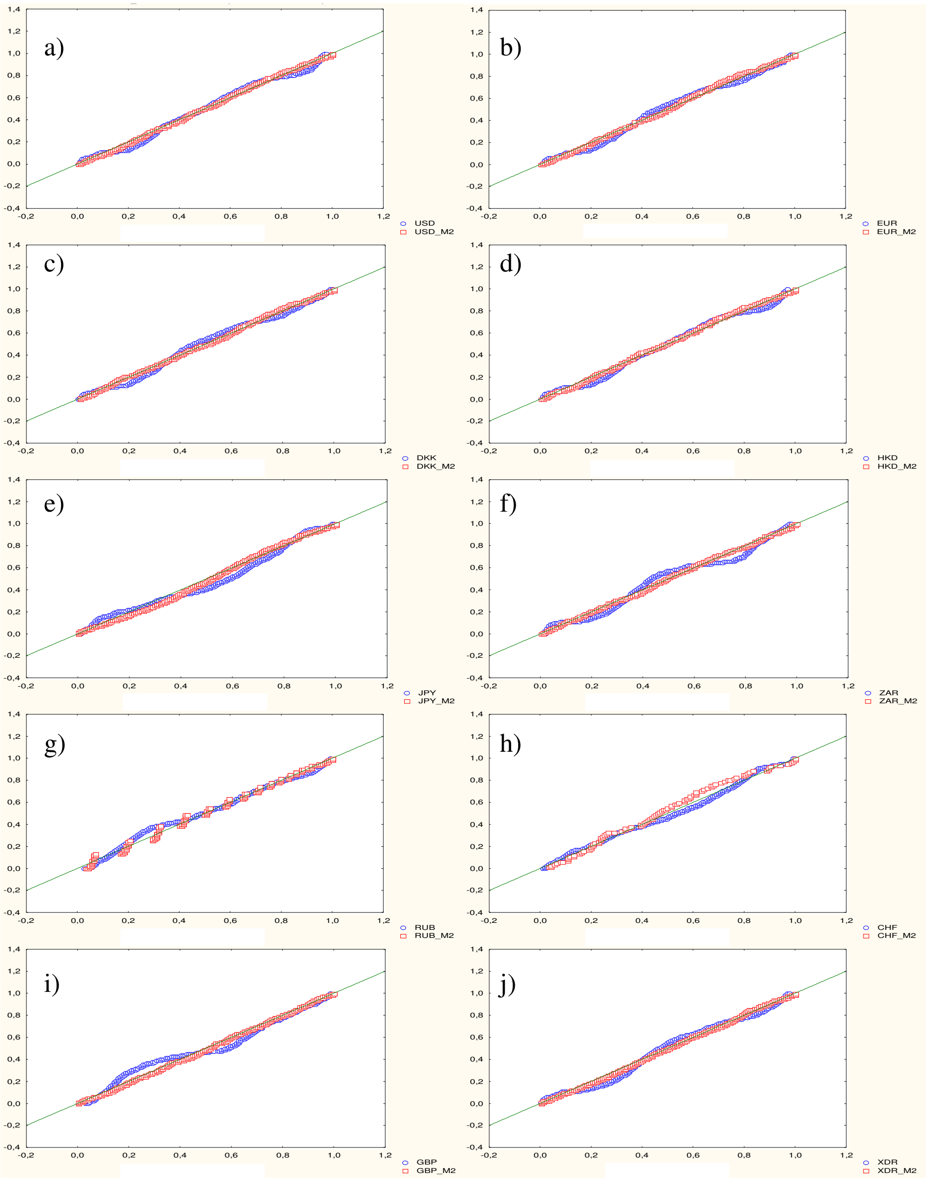

Fig. 5. Probability-Probability Plots for Gamma distribution for average exchange rates and for positive random disturbances of M2 models for a) USD, b) EUR, c) DKK, d) HKD e) JPY, f) ZAR, g) RUB, h) CHF, i) GBP j) XDR in years 2003-2011

Source: own elaboration

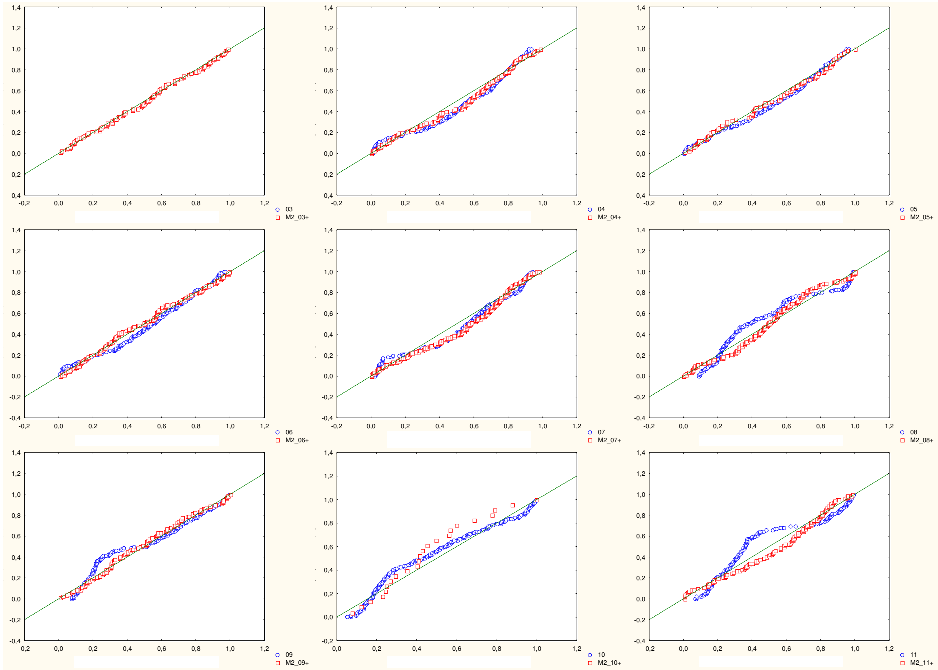

Fig. 6. Probability-Probability Plots for Gamma distribution for average exchange rates and for positive random disturbances of M2 models for USD in the following years from the period 2003-2011

Source: own elaboration

In Figure 6, which presents the exchange rates for USD, one cannot observe the Gamma distribution in 2003.

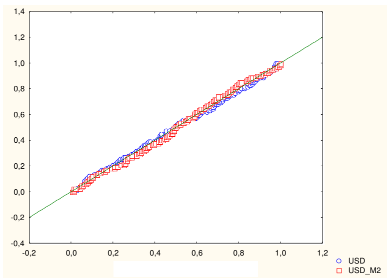

Fig. 7. Probability-Probability Plot for normal distribution for average exchange rates and for positive random disturbances of M2 model for USD in 2003

Source: own elaboration

A better distribution for exchange rates of this currency and for positive values of the random disturbances of M2 model of this currency in 2003 is a normal distribution. One can observe it in Figure 7.

Unfortunately, the normal distribution for the average exchange rate for USD does not persist over time.

3.3 Harmonic analysis

In this section, the existence of periodic components in the random disturbance of M1, M2 and M3 models is analyzed. We do not consider the existence of periodic components in the quotations of average exchange rates of chosen currencies due to the fact it is very probable that those periodic components can have long periods. In order to search the periodicities of harmonic components the standard methods presented in many econometrics textbooks (Chow 1995) are used.

For random disturbances of model (2), for every $k = 1,2,..., k \le N/2$ one defines:

$$\kappa_k(C) = \sum_{t=1}^{N} \varepsilon_t \cos\left(k \cdot \frac{2\pi}{N} \cdot t\right), \ \kappa_k(S) = \sum_{t=1}^{N} \varepsilon_t \cos\left(k \cdot \frac{2\pi}{N} \cdot t\right)$$

and critical values

$$\kappa_k^* = \begin{cases} 3 \cdot \sqrt{N} \cdot s_{N-1}(\varepsilon_t), & \text{gdy } k = \frac{N}{2} \\ 3 \cdot \sqrt{\frac{N}{2}} \cdot s_{N-1}(\varepsilon_t), & \text{gdy } k < \frac{N}{2}, \end{cases}$$

where $s_{N-1}(\varepsilon_t)$ is the root mean square deviation of random disturbances of model (2).

Table 2 presents the values of periods for random disturbances for 1^{st} order autoregressive M2 model for all analyzed currencies in years 2003-2011. A shaded cell indicates that 1^{st} order autoregressive M2 model for analyzed currency (column) can be described by the given period (row).

To summarize the results included in Table 2, one can observe some regularities in the analyzed currencies.

- We can divide all analyzed currencies into two groups. The first group consists of the currencies which all have the periods of 2.5 weeks (13.73 days) and 1.5 weeks (7.73 days). The second group consists of the currencies with some quite unique periods, i.e. different from periods of other currencies. USD, EUR, DKK, HKD, XDR belong to the first group, JPY, ZAR, RUB belong to the second one. GBP has the period of 5.15 days which does not occur in other currencies and it has also the period equal to 10 days, similarly as USD, HKD and XDR.

In addition:

- USD and HKD have the same periodic components; the same phenomenon occurs for EUR and DKK.
- RUB, CHF, GBP and XDR are the currencies which also have long periods (over 75 days).

All analyzed currencies have a period of 2.02 days.

Table 2. Regression coefficients for models M1 and M2 and relative percentage errors for models M1, M2 and M3 in years 2003-2011

Period	USD	EUR	DKK	HKD	JPY	ZAR	RUB	CHF	GBP	XDR
2.02	■	■	■	■	■	■	■	■	■	■
2.38		■	■							■
3.23					■					
5.15									■	
5.39						■				
6.83					■					
7.73	■	■	■	■		■	■	■		■
8.29							■			
10.00	■			■					■	■
13.73	■	■	■	■	■					■
15.20							■			

Source: own elaboration

4 Conclusions

The obtained results are the following:

1. The results presented in Table 1 confirm the popular sentence: the best forecast for the next day for exchange rate is today's exchange rate.
2. The distributions of the exchange rate for all currencies are unstable in time. The distributions of the positive random disturbances of the first order autoregression models of type M1, M2 and M3 are much more stable and they are similar to the gamma distribution. However, the standard assumption in econometrics textbooks is that the sequences of residuals for econometric models have a normal distribution. This also applies to the absolute values of these sequences of residuals.
3. The maximum values of relative percentage errors in all models, which are estimated by the method of Least Absolute Relative Errors Estimation, are less than 9%. For half of the models the maximum values of relative percentage errors, which are estimated by this method, are less than 5% (see Table 1).
4. The results obtained for M3 model are similar to the results obtained for M1 model.

5. In the majority of cases analyzed currencies (see Table 2) one can state occurring the same periods. This contradicts the independence of successive increases in the average daily exchange rates

In conclusion we can state that some obtained results suggest certain limitations of the applicability of the standard models of financial mathematics in relation to currency market.

Bibliography

Ashar V. G., Wallace T. D., *A sampling Study of Minimum Absolute Deviations Estimators*, “Operations Research”, 1963 vol. 11, pp. 747-758.

Chen K., Guo S., Lin Y., Ying Z., *Least Absolute Relative Error Estimation*, Journal of American Statistical Association, 2010 vol. 105(491), pp. 1104-1112.

Chow G. C., *Ekonometria*, Warszawa, PWN, 1995.

Hyb W., Kaleta J., *Porównanie metod wyznaczania współczynników modelu matematycznego na przykładzie prognozy liczby ludności świata*, Przegląd naukowy Inżynieria i Kształtowanie Środowiska,, 2004, vol. 2(29), pp. 94-99.

Rzymowski W., Surowiec A., *Autoregression models for unemployment rate*, [in:] Rola Informatyki w naukach ekonomicznych i społecznych, Innowacje i implikacje interdyscyplinarne, ed. Z. E. Zieliński, Kielce, WSH, 2011, vol. 2, pp. 262-273.

NBP Archive of average exchange rates, http://www.nbp.pl/home.aspx?c=/ascx/archa.ascx, accessed 4th June 2013.

Witold Rzymowski[1], Tomasz Warowny[2]

Bluff in games

Keywords: econometric model, independent variables, dependent variables, information bearers, covariance, spreadsheets.

Abstract

A special kind of two-person games with incomplete information is considered in the paper. The result of the game can be win or loss and the result is uncertain for both players. Prior to the game, one player, let's call him player A (attacker), assesses the strength, capabilities, etc. of his opponent B (defender). Probability of A's accession to the game depends on this evaluation. If A joins the game, B has to take part in it. Probability of A's winning depends then upon the real power of B. Before A joins the game, the player B can allocate part of his forces on the additional activities, which we call a bluff, reducing their own strength, but at the same time reducing the probability of A's accession to the game. Player B's aim is to minimize the probability of A's winning. We construct a mathematical model of this type of games and investigate the desirability of using a bluff.

1 Introduction

In many areas of human activities the result of actions of a specific entity, say player B, depends on his actions and the actions of others. Furthermore, assessment of ability (strength, capacity, etc.) of a given player B, made by other players, may have a significant impact on their performance, and thus also on the outcome of player B. Player B, allocating a part of his resources to cause a false assessment of the situation by other participants in the game, can sometimes get a better result than the one generated by his real possibilities. These additional B's activities we call a bluff.

Actions resulting in an erroneous assessment of the opponent are a common element of warfare. In past centuries Tatars escaped often from the battlefield in order to pull an opponent in a prepared trap. Obviously the result of the battle with a well-prepared opponent is uncertain and the losses can be very large. Simulated battle with simulated escape does not require a great loss and a chance

[1] Lublin University of Technology, Faculty of Fundamentals of Technology, Department of Applied Mathematics, Nadbystrzycka 38, 20-618 Lublin, *e-mail: w.rzymowski@pollub.pl*
[2] Lublin University of Technology, Faculty of Management, Department of Quantitative Methods in Management, Nadbystrzycka 38, 20-618 Lublin, *e-mail: t.warowny@pollub.pl*

of winning the battle on a prepared area with disorganized pursuing enemy is much better. The only exception was sometimes the Polish hussars prepared especially for the battles with the Tatars.

During the Second World War, before the invasion in Normandy, the Allies conducted wide-ranging measures to confuse the Germans about the place of a planned attack. Such action was undertaken to increase the initial advantage for the Allies. But, on the other hand, it probably cost a lot of resources and human efforts, which could reduce the effectiveness of attack if the Germans had discovered its real place.

In some cases, while doing business in a competitive environment, a company does not want to reveal its actual position in the market. To do this, it carries out some additional measures such as lowering prices of traded goods to increase sales ("limit pricing", see: Malawski et al 1997, p. 82), increasing expenditures on advertising or presenting a fake image of the company in the media. The purpose of it is to deceive the competing companies in such a way that the position of the deceiving company in the eyes of its competitors becomes greater than it is really.

In the animal world we can also encounter some kind of bluff. Some species of tropical frogs, in the moment of danger, inflate and take the vivid coloration to present the aggressor the illusion that they are more dangerous than they really are. Other animals, such as the antelope for example, are doing short runs with high jumps to discourage predators to initiate the chase. In doing so, they lose some of their energy which is necessary to a possible escape in case the predator has not abandoned the attack.

A good example of the use of bluff is poker game. Player, despite his weak position, can confuse opponents by aggressive manoeuvrings suggesting his much better position. This behaviour leads sometimes to a victory, but it also weakens the ability to continue the game if the opponents do not fear the bluff and decide to continue.

A simplified model of the use of bluffing in poker was presented in the book Von Neumann, Morgenstern 1967 (already recognized as a classic item) dealing with the theory of games and economic behaviour. Currently, there are many items concerning the strategies of bluffing in card games, see e.g. Caro 2003, Hansen 2003, Sklansky 1999. An interesting item concerning provocations is Karwat 2007. However, it is difficult to find an item with mathematical approach to the problem of the use of bluff or provocation in games.

This paper presents the model of two-person games with incomplete information. We consider a game with player A (attacker) and player B (defender). Roughly speaking, player B is bluffing in order to avoid playing. However, if player A joins the game, player B will also have to adhere to it. The aim of the player B is to minimize the probability of winning of his opponent.

2 Bluff

Two players can take part in the game: attacker A and defender B. The game can be won by either player A or player B, but A does not need to take part in the game. However, if A decides to take part in the game then B will have to take part as well. Prior to the game, player A assesses the strength (power, possible efficiency, etc.) of the player B on a scale from 0 to 1. If $z \in [0,1]$ stands for such an evaluation then probability of the accession of A to the game is equal to $q(z)$, where $q(0)=1$ and $q:[0,1] \to [0,1]$ is a decreasing function. Probability of winning for the player A will then depend on the real power $r \in [0,1]$ of the player B, which can be different from its evaluation z. We shall denote this probability by $p(r)$, where $p(0)=1$ and $p:[0,1] \to [0,1]$ is a decreasing function. The probability that the player A joins the game and win is therefore equal to

$$P(z,r) = q(z) \cdot p(r).$$

Let $s \in [0,1]$ denote the initial, real the strength (power, possible efficiency, etc.) of the player B. Allocating part of their capabilities for additional activities (bluff), player B can lead to a reassessment of its forces to the level of $z(s) \geq s$. On the other hand, additional activities can reduce initial strength of the player B to a level $r(s) \leq s$. As a result, the probability that the player A joins the game and win will equal

$$P(z(s), r(s)) = q(z(s)) \cdot p(r(s)).$$

Player B's aim is to minimize the probability $P(z(s), r(s))$, which leads to the problem

$$\min_{s \in [0,1]} q(z(s)) \cdot p(r(s)). \tag{1}$$

We will consider functions p, q, r, z of a special form and introduce an additional auxiliary parameter $\sigma \in [0,1]$, which can be interpreted as the intensity of bluffing. Given $a, b \in (0,1)$. Let us define, for arbitrary $x, y \in [0,1]$,

$$q(y) = 1 - by$$

and

$$p(x) = 1 - ax.$$

Clearly $q(0) = p(0) = 1$ and both p and q are decreasing functions. Next, for arbitrary $s, \sigma \in [0,1]$, we set

$$z(s,\sigma) = s + (1-s)\sigma \text{ and } r(s,\sigma) = (1-\sigma)s \quad .$$

If $\sigma = 0$, which means that B does not bluff, we obtain

$$z(s,\sigma) = r(s,\sigma) = s,$$

so that the strength of B and its assessment by player A coincide with initial value s.

If $\sigma = 1$, which means that B consumes over all of his strength to bluffing, we obtain

$$z(s,\sigma) = 1 \text{ and } r(s,\sigma) = 0.$$

In this case

$$q(1) = 1 - b < 1 = p(0).$$

Therefore A will take part in the game with probability $q(1) < 1$ and will then this game with probability $p(0) = 1$.

If $0 < \sigma < 1$, then

$$z(s,\sigma) = s + (1-s)\sigma > s, \text{ when } s < 1$$

$$r(s,\sigma) = (1-\sigma)s < s, \quad \text{when } s > 0,$$

which is consistent with the above description of the properties of the functions z and r.

Setting now

$$P(s,\sigma) = (1 - b(s + (1-s)\sigma))(1 - a(1-\sigma)s), \; s,\sigma \in [0,1],$$

we change general problem (1) to the following problem

$$\min_{s\in[0,1]} \min_{\sigma\in[0,1]} P(s,\sigma). \tag{2}$$

The solution to the problem (2) will be given in next section. This section will be finished with

Example 1

Suppose that, for all $z, r \in [0,1]$,

$$q(z) = 1 - 0{,}7z \text{ and } p(r) = 1 - 0{,}6r.$$

Let $s = 0{,}8$ be the initial strength of the player B.

(a) If B does not bluff ($\sigma = 0$) then player A's probability of accession to the game is equal to

$$q(z(s,\sigma)) = q(s) = 1 - 0{,}7s = 1 - 0{,}7 \cdot 0{,}8 = 0{,}44$$

and A will then win possible game with the probability

$$p(r(s,\sigma)) = p(s) = 1 - 0{,}6s = 1 - 0{,}6 \cdot 0{,}8 = 0{,}52 .$$

Therefore the probability that the player B will lose the game is equal to

$$q(z(s,\sigma))p(z(s,\sigma)) = 0{,}44 \cdot 0{,}52 = 0{,}229 .$$

(b) Suppose now that the player B will bluff with intensity $\sigma = 0{,}4$. This will increase its ranking to

$$z(s,\sigma) = z(0{,}8;0{,}4) = 0{,}8 + 0{,}4 \cdot (1 - 0{,}8) = 0{,}88 ,$$

which in turn will decrease A's probability of accession to the game to

$$q(z(s,\sigma)) = 1 - 0{,}7 \cdot 0{,}88 = 0{,}384 .$$

Simultaneously, intensity $\sigma = 0{,}4$ decreases B's real power to

$$r(s,\sigma) = r(0{,}8;0{,}4) = (1 - 0{,}4)0{,}8 = 0{,}48$$

which in turn increases probability of winning possible game by the player A to

$$p(r(s,\sigma)) = 1 - 0{,}6 \cdot 0{,}48 = 0{,}712 .$$

Consequently, the probability that the player B will lose the game is equal to

$$q(z(s,\sigma))p(r(s,\sigma)) = 0{,}384 \cdot 0{,}712 = 0{,}273 .$$

Comparing cases (a) and (b) we see that bluffing does not need to be a good strategy in the game

3 Solution

We are now going to solve the problem (2). If $s \in (0,1)$ then

$$P(s,\sigma) = -abs(1-s)\sigma^2 + (as(1-bs) - b(1-s)(1-as))\sigma + (1-as)(1-as)$$

is a square trinomial with respect to the variable σ with negative factor at σ^2. If $s = 0$ or $s = 1$ then $P(s,\sigma)$ is a linear function with respect to the variable σ. Therefore, for each $s \in [0,1]$, the function $P(s,\cdot):[0,1] \to R$ is concave. Since each concave function attains its minimum at the boundary of its domain we have

$$\min_{s\in[0,1]}\min_{\sigma\in[0,1]} P(s,\sigma)=$$
$$=\min_{s\in[0,1]}\min\{P(s,0),P(s,1)\}=$$
$$=\min_{s\in[0,1]}\min\{(1-as)(1-bs),1-b\}.$$

Note that

$$f(s)\overset{def}{=}P(s,0)=(1-as)(1-bs)=abs^2-(a+b)s+1,\ \ s\in R,$$

is a square trinomial with respect to the variable s with positive factor at s^2 and

$$f(0)=1>1-b,$$
$$f(1)=(1-a)(1-b)<1-b.$$

Moreover, both roots $\frac{1}{a}$ and $\frac{1}{b}$ of f are greater than 1, so that the equation

$$f(s)=1-b$$

has a unique solution $s_0=s_0(a,b)\in(0,1)$. It is easy to check that

$$s_0=s_0(a,b)=\frac{a+b-\sqrt{(a+b)^2-4ab^2}}{2ab}$$

is such a solution (see Fig. 1).

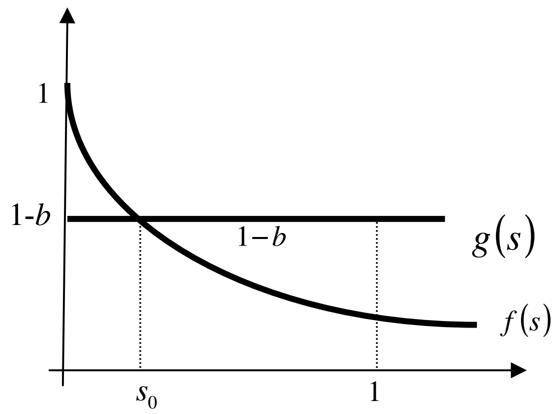

Fig. 1. Plot of functions f **and** g

Source: own elaboration

Consequently

$$\min_{s\in[0,1]}\min_{\sigma\in[0,1]} P(s,\sigma)=\begin{cases} 1-b, & \text{when } 0\le s\le s_0,\\ (1-as)(1-bs), & \text{when } s_0\le s\le 0.\end{cases}$$

4 Remarks

1. The player B should bluff with the maximal intensity in the case of $s<s_0$, because of

$$1-b=P(s,1)<(1-as)(1-bs)=P(s,0),\ s\in[0,s_0).$$

2. Bluffing is not necessary (and even can be harmful) in the case of $s>s_0$, because of

$$1-b=P(s,1)>(1-as)(1-bs)=P(s,0),\ s\in(s_0,1].$$

3. We have

$$s_0(a,b)=\frac{a+b-\sqrt{(a+b)^2-4ab^2}}{2ab}=$$

$$=\frac{\left(a+b-\sqrt{(a+b)^2-4ab^2}\right)\left(a+b+\sqrt{(a+b)^2-4ab^2}\right)}{2ab\left(a+b+\sqrt{(a+b)^2-4ab^2}\right)}=$$

$$=\frac{(\mathrm{a}+\mathrm{b})^2-\left(\sqrt{(a+b)^2-4ab^2}\right)^2}{2ab\left(a+b+\sqrt{(a+b)^2-4ab^2}\right)}=$$

$$=\frac{2b}{a+b+\sqrt{(a+b)^2-4ab^2}}$$

for each $a\in(0,1)$, so

$$\lim_{b\to 0} s_0(a,b)=0.$$

When b approaches 0, then probability $q(z)$ approaches 1 in the whole interval $[0,1]$. Thus (see Fig. 1) bluffing becomes effective only for small s, which is consistent with intuition.

4. Similarly as it was done in the cited book (Von Neumann, Morgenstern 1967) we considered here a very special and simple model of bluffing. Modelling the use of bluff in real conflicts is interesting but rather difficult problem. Our model can be viewed as a first step in this direction.

Example 2

In the case of p,q from Example 1 we have

$$s_0 = \frac{0{,}6 + 0{,}7 - \sqrt{(0{,}6+0{,}7)^2 - 4 \cdot 0{,}6 \cdot 0{,}7^2}}{2 \cdot 0{,}6 \cdot 0{,}7} = 0{,}6941$$

and

$$\min_{s\in[0,1]} \min_{\sigma\in[0,1]} P(s,\sigma) = \begin{cases} 0{,}3; & \text{when } 0 \le s \le 0{,}6941, \\ 0{,}42 \cdot s^2 - 1{,}3 \cdot s + 1; & \text{when } 0{,}6941 \le s \le 1. \end{cases}$$

By Remark 1 bluffing can be effective only for $s < s_0 = 0{,}6941$, which explains results of Example 1.

Example 3

Army Command B expects an attack by an army of A. Analysts of B have estimated the value of military capabilities sufficient to repel any attack, but also found that the real potential of a defensive army B is 70 percent ($s = 0{,}7$) of this sufficient value. For the lack of time, the military potential can no longer be increased. Army Command B must make one of two decisions:

- wait for an attack without further action;
- perform the action the enemy disinformation, which consists of additional motion of troops, sending false messages, creating dummy weapons, etc. These activities may increase the defence capability assessment (z) carried out by an army of A, but also undermine the effectiveness of any self-defence (r).

Analysts have estimated the probability of B army attack as

$$q(z) = 1 - 0{,}8z\,.$$

A probability of winning by the army of A in case of attack is estimated as

$$p(r) = 1 - 0{,}9r\,.$$

We have $a = 0{,}9$, $b = 0{,}8$.
Calculating the value of s_0, we get:

$$s_0 = \frac{a+b-\sqrt{(a+b)^2-4ab^2}}{2ab} =$$
$$= \frac{0{,}9+0{,}8-\sqrt{(0{,}9+0{,}8)^2-4\cdot 0{,}9\cdot 0{,}8^2}}{2\cdot 0{,}9\cdot 0{,}8} = 0{,}65.$$

Because $s > s_0$, the command of the army should decide not to carry out additional activities misinforming the opponent. The minimum value of the probability that the opponent will proceed to attack and win is then equal

$$\min_{\sigma\in[0,1]} P(s,\sigma) = P(s,0) = (1-as)(1-bs) = (1-0{,}9\cdot 0{,}7)(1-0{,}8\cdot 0{,}7) = 0{,}16\,.$$

If the command of the army B, however, decides to bluff, then the probability of failure would be the value of

$$\min_{\sigma\in[0,1]} P(s,\sigma) = P(s,1) = 1-b = 0{,}2\,.$$

Bibliography

Caro M., *Caro's Book of Poker Tell's*, Cardoza Publishing, 2003.
Hansen G., *Every Hand Revealed*, Kensington Publishing, 2003.
Karwat M., *Teoria prowokacji*, Wydawnictwo Naukowe PWN, 2007.
Malawski M., Wieczorek A., Sosnowska H., *Konkurencja i kooperacja. Teoria gier w ekonomii i naukach społecznych*, Wydawnictwo Naukowe PWN, 1997.
Sklansky D., *The Theory of Poker*, Two Plus Two, 1999.
Von Neumann J., Morgenstern O., *Theory of Games and Economic Behavior*, Wiley, 1967.

Wojciech Batko[1], *Oskar Knapik*[2]

Modelling the variability of the controlled environmental noise hazard levels by the ARMA processes

Keywords: time series, autoregressive moving average model, noise monitoring

Abstract

A new method of processing the results of noise measurements for the needs of estimation of long-term noise indicators is proposed in the paper. It is based on using the ARMA processes – autoregressive-moving average processes – in modelling the results. It enables an estimation of the expected value of sound level and variance, which can be used in uncertainty assessments of the realized evaluates. The presented approach is illustrated by calculation examples related to the data of the continuous noise monitoring in one of the urban areas.

1 Introduction

Problems of assessing acoustic hazards of the environment are subjects to mandatory obligations (Box, Jenkins, Reinsel 2008, Brockwell, Davis 2002, Brown, Mayer 1961) which require performing statistic control inspections, connected with the selection of the proper forecasting method of acoustic hazards in the analysed areas. Realization of these processes should take into account several measuring and calculating conditions determining the way of estimation the expected values of the controlled noise indicators as well as methods of their uncertainty assessments. Currently binding estimation procedures of noise indicators (Brown, Mayer 1961) are derived from the classic solutions of statistic conclusions. Their estimating rules (Guide to the Expression of Uncertainty Measurement 1995) are related to assumptions of the equivalence of the results of the random control test and the normal probability distribution of occurrence of these results, and also to the condition that successive meas-

[1] AGH University of Science and Technology, Faculty of Mechanical Engineering and Robotics, Department of Mechanics and Vibroacoustics, *e-mail: batko@agh.edu.pl*
[2] Cracow University of Economics, Faculty of Management, Department of Statistics, *e-mail: knapiko@uek.krakow.pl*

urements $x_i, i = 1,...,n$ are statistically independent one of another, what means they are uncorrelated. When the environment hazard state control is realized on the bases of random tests, these assumptions are generally accepted without wider analyses and discussions of their reliability. This concerns mainly the correctness of the assumptions that the results of sound level measurements are of the normal distribution and that they are uncorrelated. The first one, being the result of physical properties of the measured value and the observed - in practice - asymmetry of the density of probability distribution of the results is generally difficult to be accepted. The second one is of a small likelihood due to often occurring measuring disturbances of relatively high levels, which influence successive results of random tests as well as conditions of performing these tests for the needs of the long-term noise indicators estimation.

Abovementioned limitations of the current estimation methods of long-term noise indicators generated the need of looking for new modelling attempts. In some papers (Regulation of the Minister of Environment 2007, Bal 2010, Batko, Bal 2008) the attention was paid to the possibility of estimation of noise indicators and their uncertainty assessments by modelling the control tests results by means of time series models, especially by adaptation trend models from the R.G. Brown's method perspective (Batko, Knapik 2013).

Positive results obtained in this model approach induced the authors to continue this way of looking for new estimation solutions.

The aim of the hereby paper is the presentation of this model concept with taking into account the formalism of the *ARMA* processes, which describes a broader class of possible modelling solutions, which have already found numerous applications (Batko, Bal 2008). Its presentation, in relation to modelling the variability of assessing measurements of the controlled noise indicators together with the analysis of its suitability for the probabilistic characteristics of estimation, constitutes the subject of the hereby paper.

2 Description of the applied ARMA method in modelling the measurement results of the controlled noise indicators

The assessment of acoustic hazards in areas requiring protection and related to them selection of environmental acoustic protection solutions is conditioned by knowledge of long-term average sound levels $L_{LT}^{(j)}$ at times of the day: $j = 1$, i.e. in hours [6.00 -18.00], in the evening: $j = 2$, hours: [18.00-22.00] and at night: $j = 3$, hours: [22.00-6.00] - during the whole calendar year.

The $L_{LT}^{(j)}$ values for individual times: $j = 1, 2, 3$, are determined as:

$$L_{LT}^{(j)} = 10\log\left[\frac{1}{365}\sum_{k=1}^{365}10^{0.1L_{A,eq,LT,k}^{(j)}}\right] \tag{0.1}$$

$$L_{LT}^{(j)} = 10\log\left[\frac{1}{365}\sum_{k=1}^{365}0.1^{L_{A,eq,LT,k}^{(j)}}\right] \tag{1}$$

from equivalent sound A levels ($L_{A,eq,T}$) [dB/A] in the k^{th} day of the calendar year, in the considered reference time period $T^{(j)}$ proper for the examined time: $j = 1, 2, 3$.

The long-term day-evening-night level L_{DEN}, being the basic indicator for the selection the realization strategy of the environment acoustic protection, is calculated from their values

$$\left\{L_{A,eq,LT}^{(1)} = L_D,\ L_{A,eq,LT}^{(2)} = L_E,\ L_{A,eq,LT}^{(3)} = L_N\right\}:$$
$$L_{DEN} = 10\log\left[\frac{1}{24}\left(12\cdot 10^{0.1L_D} + 4\cdot 10^{0.1(L_E+5)} + 8\cdot 10^{0.1(L_N+10)}\right)\right] \tag{2}$$

The knowledge of the expected value and variance related to the results of the equivalent sound level $L_{A,eq,t}^{(j)}$ at the determined time: $j = 1, 2, 3$ of the controlled day $t = 1, 2, ..., n$ during the calendar year is necessary in such an estimation process.

Their values $L_{A,eq,t}^{(j)}$, $t = 1, 2, ..., n$ can be connected with the time series $\{X_t, t \in Z\}$ representing them (Brown, Mayer 1961), i.e. a sequence of random variables. They are determined by results of the successive '*momentary*' control observations $\{x_1, ..., x_n\}$ describing the environment acoustic state. It can be assumed that the probabilistic structure of the control results changes can be shaped by the additive process. Such a mechanism can be expressed as:

$$X_t = \mu_t + \varphi_t + \xi_t. \tag{0.2}$$

Thus, its structure can be shaped by the factor influencing the average level, i.e. μ_t related to the constant tendency forcing the level of the analysed noise indicators in the given area, the cyclic component φ_t - representing periodical changes related to recurrent characteristic forcing (influencing changes of the controlled noise indicators) and the residual component representing random disturbances (or imperfections of model descriptions) ζ_t of the expected value being zero and variance: σ_ζ^2. The most often, it is assumed that the distribution of disturbances related to the model error ζ_t is $N\left(0,\sigma_\zeta^2\right)$.

In contrast to the classic model of the statistic random control test (generally recommended in the estimation of the controlled parameter and its uncertainty, Guide to the Expression of Uncertainty Measurement, 1995) assuming that successive control results are random variables with a normal distribution, in the approach proposed in Regulation of the Minister of Environment 2007, Bal 2010, Batko, Bal – Pyrcz 2006 the presence of a certain mechanism forcing changes of control results is assumed. This mechanism is subjected to random disturbances of expected values being zero and variances σ_ζ^2.

The problem of estimating the expected value and variance of the analysed noise indicators is reduced (*in such an approach*) to the identification and estimation of the structure of time series. This requires the proper estimates: $\hat{\mu}_t$, $\hat{\varphi}_t$ for μ_t and φ_t components, which should ensure the correct variability description of successive control tests.

In order to increase the effectiveness of the obtained evaluations, in relation to the proposition given in papers (Regulation of the Minister of Environment 2007, Bal 2010, Batko, Bal – Pyrcz 2006), an introduction of a sufficiently broad class of models enabling a proper description of properties of process realizations is appropriate.

We assume that the component φ_t responsible for cyclic fluctuations can be neglected (Batko, Bal 2008) and we limit ourselves to the class of stationary models. Those models can be approximated by means of the $ARMA(p,q)$ models, definitions of which are given below.

A stochastic process $\{X_t, t\in Z\}$ is called the $ARMA(p,q)$ process if:

$$X_t = \mu_t + \xi_t \tag{4}$$

where:

$\mu_t = \phi_0 + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p}$, $\xi_t = \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q} + Z_t, \{Z_t\} \sim N(0,\sigma^2)$.

Two groups of solutions are utilised in the identification and estimation processes of the $ARMA$ model parameters. The first group determines solutions based on the sample autocorrelation function (ACF) and on the partial autocorrelation function (PACF). Comparisons of individual autocorrelation functions with the theoretical representatives – resulting from the $ARMA$ processes properties – are being done on their bases. The second group constitutes methods based on comparisons of the estimated models with using information criteria (among others: Akaike, Schwarz). However, they require additional assumptions concerning the nature of the mechanism generating random disturbances. It is assumed the most often that $\{X_t\}$ is the zero-mean Gaussian time series. Technical details concerning the classic estimation of the $ARMA$ process parameters with the application of *the Maximum Likelihood Method* can be found e.g. in Batko, Bal 2008 or Batko, Knapik 2013. The authors of the hereby paper (see Box, Jenkins, Reinsel 2008) propose the application of the robust statistical procedures for identification of the orders and estimation of unknown parameters of the $ARMA(p,q)$ model.

The modelling process, understood in the above way, can be applied for the description of variability of the monitored noise indicators. In this approach the $ARMA(p,q)$ model allows the evaluation of the expected sound level, since it is represented by $\hat{\varphi}_0$ - the ARMA model estimate. This solution is justified from the formal point of view.

In the empirical part we will show the example of the average of the equivalent sound level estimation.

3 Results of the modelling experiments

The proposed modelling formalism was assigned to the results of the continuous noise monitoring carried out in the town of Kielce. Their data are contained in the database of the acoustic monitoring system developed in the Department of Mechanics and Vibroacoustics (www.monitoringakustyczny.pl). The bases of the realised analyses constituted three time series. Each represented the original time-history of the results of the sound level measurements $L_{A,i}$ taken in 1-second intervals in three discussed periods *j*. Each series was formed by joining the corresponding to each other series from the 7-days period. Their graphical presentation is given in Figures 1 - 3.

Next, the identification and estimation procedure of the $ARMA$ process parameters, described in Bal 2010, was applied on the bases of the considered time series. The best model from the $ARMA$ class for the description of time series of the day time noise indicators (6:00-18:00) is the $ARMA(5,2)$ model. In case of series of the evening noise indicators (18:00-22:00), it is the $ARMA(2,5)$ model. For the series corresponding the night (22:00-6:00) the best model is, similarly as for the evening, the $ARMA(2,5)$ model.

Examples of simulations of the modelling description for variability of the sound level results for the day, evening and night - are presented in Figures 1-3 and the corresponding theoretical sound level obtained on the basis of $ARMA$ models in Figures 4-6.

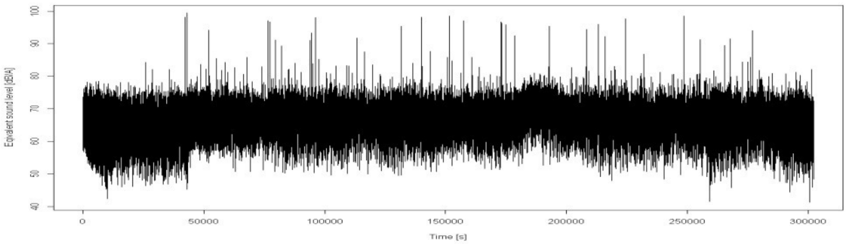

Fig. 1. Variability of the sound level (in dB/A) for series *j* = 1

Source: Own research

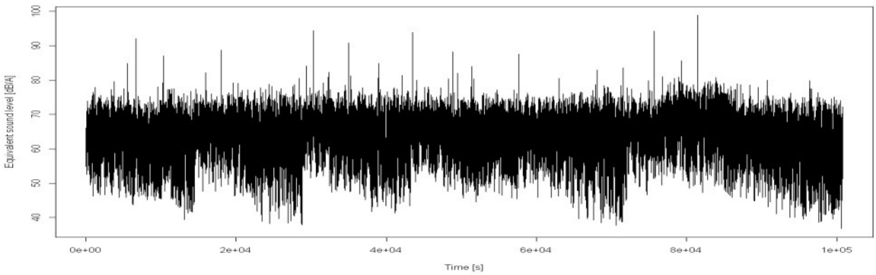

Fig. 2. Variability of the sound level (in dB/A) for series *j* = 2

Source: Own research

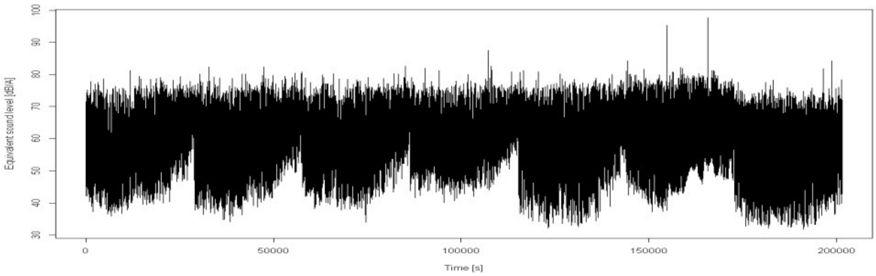

Fig. 3. Variability of the sound level (in dB/A) for series j=3.

Source: Own research

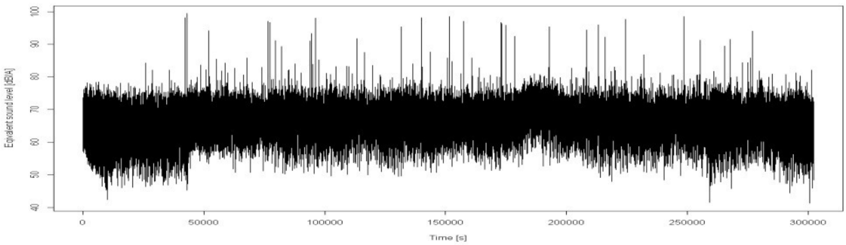

Fig. 4. Theoretical variability of the sound level \\ (in dB/A) for series j=1

Source: Own research

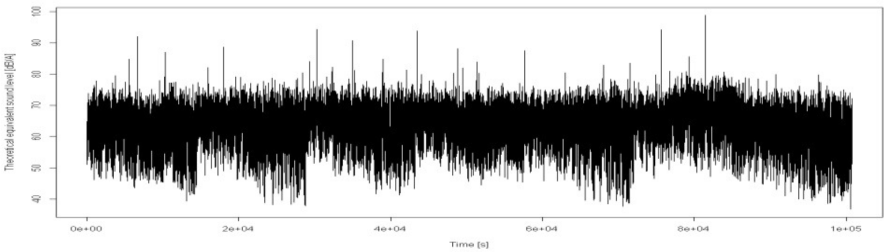

Fig. 5. Variability of the sound level (in dB/A) for series j=2

Source: Own research

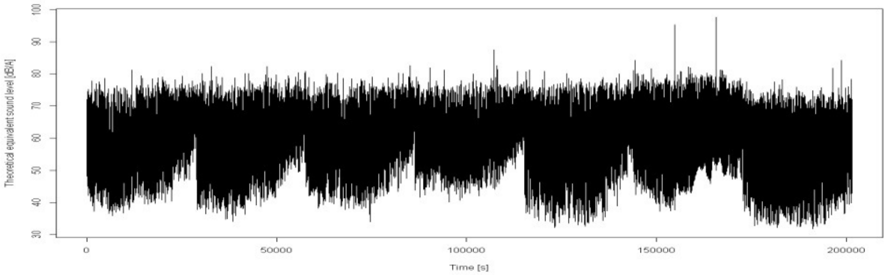

Fig. 6. Variability of the sound level (in dB/A) for series j=3

Source: Own research

The expected sound level representing the control process is determined by the estimator of the intercept of the best fitted to the data class model. The results for the considered time series are given in the table below.

Table 1. Evaluation of the average sound level for the considered series together with the corresponding standard deviation

Series	$j=1$	$j=2$	$j=3$
Evaluation of the average sound level (Standard error)	66.17 [dB(A)] (0.042)	62.64 [dB(A)] (0.12)	57.06 [dB(A)] (0.13)

Source: own research

It is worth mentioning that the evaluations, given in the Table above, are in agreement with the results obtained on the widerly controlled data. The highest value i.e. 66.17 [dB(A)] is obtained for the sound level corresponding with hours: 6:00-18:00. The standard deviation related to the evaluation of the expected sound level equals 0.04 [dB(A)], which means that the expected sound level can differ by 0.04 [dB(A)] from the estimate. The second value, 62.64 [dB(A)], corresponds the evening time: 18:00-22:00. In this case, the expected sound level can differ, on average, by 0.12 [dB(A)] from the estimated value. The lowest value corresponds to the night (57.06 [dB(A)]). The standard deviation equals in this case 0.13 [dB(A)]. It is also worth noting that standard deviations, related to the estimation of the expected sound level, are in each case very small.

4 Conclusions

The proposed and shown in the paper procedure of the modelling the variability of the environment monitored sound levels by the ARMA processes provides the formally adequate mathematical tool, ensuring the estimation of the expected sound level variability and the accuracy of such estimations. The results of the simulation investigation indicate its usefulness for obtaining diagnostics information on the environment hazard. The essential improvement of the estimation accuracy can be obtained by using longer time series in the described algorithm. These can be obtained by combining the results of noise monitoring from more measuring days. Although the proposed estimating approach to long-term noise indicators can be considered more complex than the one currently applied it is characterized by a higher universalism. It is not limited by the restrictions pointed out in the introduction. However, its broader implementations in the environment acoustic control practice will require developing the software dedicated solely to this solution.

Bibliography

Bal R.: *Analysis of estimation conditions of the long-term noise indicators*, Ph.D. Thesis, AGH, 2010 (in Polish).

Batko W., Bal – Pyrcz R. *Uncertainty analysis in the assessment of long – term noise indicators*, Archives of Acoustics vol.31, No 4, pp. 253 – 260, 2006.

Batko W., Bal R. : *Choosing model approximation for the purpose of the estimation of the long-term noise indicators LDEN and LN*, Archives of Acoustics, Vol. 32, 2008.

Batko W., Knapik O., *Robust statistical procedures in statistical analysis of controllable noise indicators*, Acta Physica Polonica A, 2013 .

Box G.E.P., Jenkins G.M., Reinsel G. C., *Time Series Analysis: Forecasting and Control* (Wiley Series in Probability and Statistics). Wiley, 4th Edition, 2008.

Brockwell P.J., Davis R. A., *Introduction to Time Series and Forecasting*, Springer, 2nd edition, 2002.

Brown R.G., Mayer R.F., *The Fundamental Theorem of Exponential Smoothing*, Operations Research, Vol.9, 1961.

Directive 2002/49/WE of the European Parliament and of the Council of 25 June 2002, relating to the assessment and management of environmental noise, Official Journal of the European Communities 18 July 2002. The Act of 27th April 2001.

Environment Protection Law (the Journal of Law 2008, No. 25, item 150 with later amendments) / Ustawa z dnia 27 kwietnia 2001 r. Prawo ochrony środowiska (Dz. U. 2008 r., nr 25, poz. 150 z późn. zm.) .

Guide to the Expression of Uncertainty Measurement. International Organization for Standardization, ISBN 92-67-10188-9, 1995.

Regulation of the Minister of Environment of 2nd October 2007, concerning requirements of performing in the environment the measurements of substances or energy levels for managers of roads, railways, tramways, airports and harbours (the Journal of Law 2007, No. 192, item 1392).

Adam Gregosiewicz [1]

Asymptotic Behaviour of Diffusions on Graphs

Keywords: fast diffusions, Markov processes, convergence of semigroups

Abstract

We investigate fast diffusions on finite directed graphs. We prove results in a way dual to presented in A. Bobrowski (2012), A. Bobrowski and K. Morawska (2012) and obtain an asymptotic behaviour of a diffusion semigroup on a graph in L^1 as the diffusion speed increases and the probability of a particle passing through a vertex decreases.

1 Introduction

Assume that $\mathscr{G}$ is a directed graph in $\mathbb{R}^3$ without loops, and there is a Markov process on $\mathscr{G}$ which on each edge behaves like a Brownian motion with given variance. Moreover, assume that each vertex is a semipermeable membrane with given permeability coefficients, i.e. for each vertex there are non-negative numbers p_{ij} describing the probability of a particle passing through the membrane from the ith to the jth edge.

In A. Bobrowski (2012), A. Bobrowski and K. Morawska (2012), the authors prove that if the diffusion speed increases to infinity at the same rate as the permeability coefficient decreases to zero, then we obtain a limit process which is a Markov chain on the vertices of the line graph of $\mathscr{G}$.

The aim of this paper is to prove a similar asymptotic result but in a different space. In A. Bobrowski (2012), A. Bobrowski and K. Morawska (2012), the authors deal with the space of continuous functions on a graph $\mathscr{G}$. Here we investigate diffusions on the space of Lebesgue integrable functions on a graph $\mathscr{G}$.

One may wish to mimic the original proof of the continuous case but this is not fully possible. In particular, in the space of continuous functions on a graph there exists an isomorphism transforming boundary conditions associated with the original process to much simpler Neumann boundary conditions. Because of that, one can obtain the limit for the isomorphic semigroups which leads to required

[1] Department of Mathematics, Technical University of Lublin, Nadbystrzycka 38, 20-618 Lublin, *e-mail: a.gregosiewicz@pollub.pl*

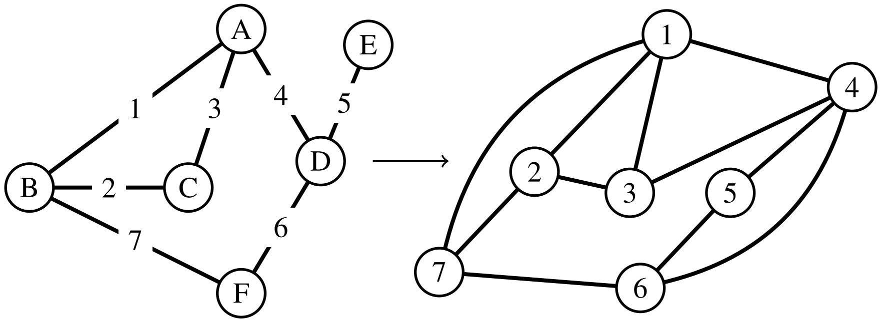

Figure 1: Diffussion on a graph becomes a Markov chain on the vertices of the line graph.

asymptotics. What is crucial, in the space of integrable functions such isomorphism does not exist. However, there is an isomorphism of the Sobolev space $W^{2,1}$ in a way similar to the isomorphism in the space of continuous functions, which leads to different approach via the Kurtz theorem.

2 Continuous case

As in A. Bobrowski (2012), let $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ be a finite geometric graph (D. Mugnolo (2007)) without loops, where $\mathscr{V} \subset \mathbb{R}^3$ is the set of vertices and $\mathscr{E}$ is the set of edges of finite lengths. The number of edges is N and the edges are seen as C^1 curves connecting vertices. Let $\mathscr{N} = \{1, \dots, N\}$. For $i \in \mathscr{N}$, by convention, we call the initial and terminal vertices of the ith edge E_i its "left" and "right" endpoints. We denote them by L_i and R_i, respectively. Moreover, let V_i, $i \in \mathscr{N}$ denote the vertex $V \in \mathscr{V}$ as an endpoint of the ith edge. If V is not an endpoint of this edge, we leave V_i undefined.

Let $S = \bigcup_{i \in \mathscr{N}} E_i$ be a disjoint union of the edges. What is in particular important, since by convention $V_i \neq V_j$ for $i, j \in \mathscr{N}$, $i \neq j$, in S there can be many "copies" of the same vertex, treated as an endpoint of different edges. Then S is a disconnected compact topological space, and we denote by $C(S)$ the space of continuous functions on S with the standard supremum norm. If $f \in C(S)$, then we may identify $f = (f_i)_{i \in \mathscr{N}}$ where f_i is a member of $C(E_i)$. The latter space is isometrically isomorphic to the space $C[0, d_i]$ where d_i is the length of the ith edge.

Let $\sigma \in C(S)$ be defined by $\sigma(p) = \sigma_i$ for $p \in E_i$ where σ_i are given positive numbers. Define the operator A in $C(S)$ by

$$Af = \sigma f'', \tag{1}$$

on the domain composed of twice continuously differentiable functions, satisfying the transmission conditions described below.

For each $i \in \mathcal{N}$, let l_i and r_i be non-negative numbers describing the possibility of passing through the membrane from the ith edge to the edges incident in the left and right endpoints, respectively. Also, let l_{ij} and r_{ij}, $j \neq i$ be non-negative numbers satisfying $\sum_{i\neq j} l_{ij} \leqslant l_i$ and $\sum_{i\neq j} r_{ij} \leqslant r_i$. These numbers determine the probability that after filtering through the membrane from the ith edge a particle will enter the jth edge.

By default, if E_j is not incident with L_i, we put $l_{ij} = 0$. In particular, by convention $l_{ij} f(V_j) = 0$ for $f \in C(S)$ if V_j is not defined. The same remark concerns r_{ij}. In these notations, the transmission conditions mentioned above are as follows: if $L_i = V$, then

$$f'(V_i) = l_i f(V_i) - \sum_{j\neq i} l_{ij} f(V_j), \tag{2}$$

where $f'(V_i)$ is the right-hand derivative of f at V_i and if $R_i = V$, then

$$-f'(V_i) = r_i f(V_i) - \sum_{j\neq i} r_{ij} f(V_j), \tag{3}$$

where $f'(V_i)$ is the left-hand derivative of f at V_i.

It is showed in A. Bobrowski (2012) that the operator A generates a Feller semigroup $\{\mathrm{e}^{tA}, t \geqslant 0\}$ in $C(S)$. Moreover, the semigroup is conservative if and only if $\sum_{j\neq i} l_{ij} = l_i$ and $\sum_{j\neq i} r_{ij} = r_i$, for $i \in \mathcal{N}$. What is more, the following theorem is proved.

Theorem 2.1. *Let $(\kappa_n)_{n\geqslant 1}$ be a sequence of positive numbers converging to infinity and let operators A_n be defined by* (1) *with σ replaced with $\kappa_n\sigma$ and with the domain composed of $C^2(S)$ functions satisfying the transmission conditions* (2) *and* (3) *with permeability coefficients (i.e., all l_i, r_i, l_{ij} and r_{ij}'s) divided by κ_n. Then,*

$$\lim_{n\to\infty} \mathrm{e}^{tA_n} f = \mathrm{e}^{tQ} P f, \quad f \in C(S),\ t > 0,$$

where P is the projection of $C(S)$ on the space $C_0(S)$ of functions that are constant on each edge, given by $Pf = \left(d_i^{-1}\int_{E_i} f\right)_{i\in\mathcal{N}}$, while Q is the operator in $C_0(S)$ which may be identified with the matrix $(q_{ij})_{i,j\in\mathcal{N}}$ with $q_{ij} = \sigma_i d_i^{-1}(l_{ij} + r_{ij})$ for $i \neq j$ and $q_{ii} = -\sigma_i d_i^{-1}(l_i + r_i)$. The limit here is strong and almost uniform in $t \in (0,\infty)$; for $f \in C_0(S)$, the formula holds for $t = 0$ as well and the limit is almost uniform in $t \in [0,\infty)$.

3 Adjoint of the operator A

Using the same identification as in the previous section, we will now consider the space $L^1(S) = \{\varphi\colon \varphi = (\varphi_i)_{i\in\mathcal{N}},\ \varphi_i \in L^1(E_i)\}$. Here $L^1(E_i)$ is the space of Lebesgue integrable functions on E_i. Let $W^{2,1}(S)$ be the Sobolev space, i.e. the space composed of functions $\varphi \in L^1(S)$ such that φ, φ' are weakly differentiable and $\varphi', \varphi'' \in L^1(S)$. We define the operator A^* in $L^1(S)$ by

$$A^*\varphi = \sigma\varphi'' \tag{4}$$

with domain composed of members of $W^{2,1}(S)$ satisfying the transmission conditions

$$\begin{aligned}\sigma_j\varphi'(L_j) &= \sigma_j l_j \varphi(L_j) - \sum_{i\in I_j^L}^{*}\big[\sigma_i l_{ij}\varphi(L_i) + \sigma_i r_{ij}\varphi(R_i)\big],\\ \sigma_j\varphi'(R_j) &= \sum_{i\in I_j^R}^{*}\big[\sigma_i l_{ij}\varphi(L_i) + \sigma_i r_{ij}\varphi(R_i)\big] - \sigma_j r_j\varphi(R_j),\end{aligned} \tag{5}$$

for $j \in \mathscr{N}$. Here, I_j^L and I_j^R are the sets of indexes $i \neq j$ of edges incident in L_j and R_j, respectively. The asterisk in the sums denotes the fact that, since there are no loops, at most one of the terms $\sigma_i l_{ij}\varphi(L_i)$ and $\sigma_i r_{ij}\varphi(R_i)$ is taken into account.

Lemma 3.1. *If $f \in D(A)$ and $\varphi \in D(A^*)$, then*

$$\int_S \varphi A f = \int_S (A^*\varphi) f. \tag{6}$$

Proof. Integrating by parts, we obtain

$$\begin{aligned}\int_{E_i} \varphi f'' = \varphi(R_i)f'(R_i) - \varphi'(R_i)f(R_i) + \varphi'(L_i)f(L_i) - \varphi(L_i)f'(L_i)\\ + \int_{E_i}\varphi'' f\end{aligned}$$

for $i \in \mathscr{N}$. Hence, equality (6) holds if and only if

$$\sum_{i\in\mathscr{N}}\big[\varphi(R_i)f'(R_i) - \varphi'(R_i)f(R_i) + \varphi'(L_i)f(L_i) - \varphi(L_i)f'(L_i)\big] = 0, \quad i \in \mathscr{N}. \tag{7}$$

Since f belongs to $D(A)$, the transmission conditions (2) and (3) are satisfied. Thus condition (7) holds if and only if

$$\begin{aligned}\sum_{i\in\mathscr{N}} \sigma_i\varphi(R_i)\Big[\sum_{j\neq i} r_{ij}f(R_{ij}) - r_i f(R_i)\Big] - \sum_{i\in\mathscr{N}}\sigma_i\varphi'(R_i)f(R_i)\\ + \sum_{i\in\mathscr{N}}\sigma_i\varphi(L_i)\Big[\sum_{j\neq i} l_{ij}f(L_{ij}) - l_i f(L_i)\Big] + \sum_{i\in\mathscr{N}}\sigma_i\varphi'(L_i)f(L_i) = 0,\end{aligned}$$

where L_{ij} and R_{ij} are, by definition, respectively left and right ends of E_i, seen as members of E_j. Changing the order of summation, the last equality becomes

$$\begin{aligned}\sum_{j\in\mathscr{N}}\sum_{i\neq j}\sigma_i r_{ij}\varphi(R_i)f(R_{ij}) - \sum_{j\in\mathscr{N}}\sigma_j f(R_j)\big[r_j\varphi(R_j) + \varphi'(R_j)\big]\\ + \sum_{j\in\mathscr{N}}\sum_{i\neq j}\sigma_i l_{ij}\varphi(L_i)f(L_{ij}) - \sum_{j\in\mathscr{N}}\sigma_j f(L_j)\big[l_j\varphi(L_j) - \varphi'(L_j)\big] = 0.\end{aligned}$$

Notice that L_{ij} is either L_j or R_j or is left undefined, and the same holds for R_{ij}. Thus we can rewrite the last condition in the form

$$\sum_{j\in\mathcal{N}} f(R_j)\left\{\sum_{i\in I_j^R}^{*}\left[\sigma_i r_{ij}\varphi(R_i)+\sigma_i l_{ij}\varphi(L_i)\right]-\sigma_i r_i\varphi(R_i)-\sigma_i\varphi'(R_i)\right\}$$
$$+\sum_{j\in\mathcal{N}} f(L_j)\left\{\sum_{i\in I_j^L}^{*}\left[\sigma_i r_{ij}\varphi(R_i)+\sigma_i l_{ij}\varphi(L_i)\right]-\sigma_j l_j\varphi(L_j)+\sigma_j\varphi'(L_j)\right\}=0,$$

which is true, since φ satisfies the transition conditions (5). □

Keeping in mind the Riesz representation theorem, the above lemma shows that the operator A^* is in a way "adjoint" to A. More precisely, A^* is the part of the adjoint operator of A in $L^1(S)$.

4 Operator A^* generates a sub-Markov semigroup

As we said before, we know from A. Bobrowski (2012) that A generates a Feller semigroup in $C(S)$. We will prove that operator A^* generates a sub-Markov semigroup $\{e^{tA^*}, t \geqslant 0\}$ in $L^1(S)$, i.e. a semigroup of operators such that for a non-negative $\varphi \in L^1(S)$ we have $e^{tA^*}\varphi \geqslant 0$ and $\int_S e^{tA^*}\varphi \leqslant \int_S \varphi$ for $t \geqslant 0$. Since the domain of A^* is dense in $L^1(S)$ (see the remarks after Theorem 5.1) this is equivalent to all the resolvents being sub-Markov (see A. Lasota, M.C. Mackey (1985), Corollary 7.8.1), which by definition means that the operators $\lambda(\lambda - A^*)^{-1}$ are sub-Markov for $\lambda > 0$.

Lemma 4.1. *For all $\lambda > 0$ there exists the resolvent $(\lambda - A^*)^{-1}$ and the following equality*

$$\int_S \varphi(\lambda - A)^{-1} f = \int_S f(\lambda - A^*)^{-1}\varphi, \qquad f \in C(S),\ \varphi \in L^1(S), \tag{8}$$

holds.

Observe that if the resolvent of A^* exists, then equality (8) becomes obvious since A^* is the part of the adjoint of A in $L^1(S)$. However, the existence of $(\lambda - A^*)^{-1}$ is not obvious. We will prove Lemma 4.1 in a moment.

It is well known, (see e.g. A. Lunardi (1985), Theorem 3.1.3) that the operator G in $L^1(a,b)$ given by $G\varphi = \varphi''$ with the domain composed of functions $\varphi \in W^{2,1}(a,b)$ satisfying the Neumann boundary conditions, i.e. $\varphi'(a) = \varphi'(b) = 0$, is sectorial with angle $\pi/2$. It is also easy to see that the domain of G is dense in $L^1(a,b)$. Indeed, assume that $\varphi \in C^2[a,b]$, and let $\varphi_n \in C^2[a,b]$, $n \geqslant 1$, be given by $\varphi_n(x) = \mathbb{1}_{(a,b)}(x)\exp\{\frac{-1}{n(x-a)(b-x)}\}\varphi(x)$. Then $\varphi_n'(a) = \varphi_n'(b) = 0$ and $\|\varphi - \varphi_n\|_{L^1(a,b)} \to 0$ as $n \to \infty$ by the Lebesgue dominated convergence theorem.

Since $C^2[a,b] \subset W^{2,1}(a,b)$ is dense in $L^1(a,b)$, this proves that $\overline{D(G)} = L^1(a,b)$. Hence, by K.J. Engel, R. Nagiel (2000), Theorem II.4.6, the operator G generates a equi-bounded analytic semigroup in $L^1(a,b)$.

Lemma 4.2. *For G defined above,*

$$\lim_{t\to\infty} \mathrm{e}^{tG}\varphi = \lim_{\lambda\to 0^+} \lambda(\lambda - G)^{-1}\varphi = \frac{1}{b-a}\int_a^b \varphi \tag{9}$$

for $\psi \in L^1(a,b)$, $\lambda > 0$, *where* $\int_a^b \varphi$ *is identified with the constant function on* (a,b).

Proof. For simplicity, we will assume that $a = 0$ and $b = 1$.

Since G is sectorial, it suffices to show the second equality. For $\nu \in \mathbb{R}$ let $\mathrm{e}_\nu \in C(\mathbb{R})$ be defined by $\mathrm{e}_\nu(x) = \mathrm{e}^{-\nu x}$. Given $\lambda > 0$, $\psi \in L^1(0,1)$, the unique solution $\psi \in D(G)$ of $\lambda\psi - \psi'' = \varphi$ is of the form

$$\psi(x) = \frac{1}{2\mu}\int_0^1 \mathrm{e}_\mu(|x-y|)\varphi(y)\,\mathrm{d}y + c_1\mathrm{e}_{-\mu}(x) + c_2\mathrm{e}_\mu(x), \tag{10}$$

where $\mu = \sqrt{\lambda}$. Moreover, constants c_1, c_2 are chosen so that $\psi'(0) = \psi'(1) = 0$. Precisely, we have

$$c_1 = \frac{\xi_1\mathrm{e}^{-\mu} + \xi_2}{\mu(\mathrm{e}^\mu - \mathrm{e}^{-\mu})}, \qquad c_2 = \frac{\xi_1\mathrm{e}^{\mu} + \xi_2}{\mu(\mathrm{e}^\mu - \mathrm{e}^{-\mu})}, \tag{11}$$

where

$$\xi_1 = \frac{1}{2}\int_0^1 \mathrm{e}_\mu(y)\varphi(y)\,\mathrm{d}y, \qquad \xi_2 = \frac{1}{2}\int_0^1 \mathrm{e}_\mu(1-y)\varphi(y)\,\mathrm{d}y.$$

Thus, by the Lebesgue dominated convergence theorem, we have

$$\lim_{\mu\to 0^+} \mu^2 c_1 = \lim_{\mu\to 0^+} \mu^2 c_2 = \frac{1}{2}\lim_{\mu\to 0^+}(\xi_1 + \xi_2).$$

Finally, since

$$\lim_{\mu\to 0^+} \mu\int_0^1 \mathrm{e}_\mu(|x-y|)\varphi(y)\,\mathrm{d}y = 0,$$

we have

$$\lim_{\lambda\to 0^+} \lambda\psi = \lim_{\mu\to 0^+}(\xi_1 + \xi_2) = \int_0^1 \varphi,$$

which completes the proof. □

Lemma 4.3. *The part G_0 of G in $W^{2,1}(a,b)$ is a sectorial operator in $W^{2,1}(a,b)$ with angle $\frac{\pi}{2}$, i.e. for each $\delta \in (0,\pi/2]$ the sector*

$$\Sigma_{\frac{\pi}{2}+\delta} = \{\lambda \in \mathbb{C}\setminus\{0\} \colon |\arg\lambda| < \pi/2 + \delta\}$$

is contained in the resolvent set $\rho(G_0)$ of G_0 and there exists $M > 0$ such that

$$\|(\lambda - G_0)^{-1}\|_{\mathscr{L}(W^{2,1})} \leqslant \frac{M}{|\lambda|}, \qquad \lambda \in \Sigma_{\frac{\pi}{2}+\delta}.$$

Proof. Without loss of generality we may assume that $a = 0$ and $b = 1$. For $\delta \in (0, \pi/2)$ let $\lambda = |\lambda| e^{i\theta}$, $\theta \in (-\pi/2 - \delta, \pi/2 + \delta)$, and set $\mu = \sqrt{\lambda}$ such that $m := \operatorname{Re}\mu > 0$. Given $\varphi \in W^{2,1}(0,1)$, the unique solution $\psi \in W^{2,1}(0,1)$ of $\lambda\psi - \psi'' = \varphi$ satisfying the Neumann boundary conditions is defined by (10). Rewrite (11) in the form

$$c_1 = \frac{1}{2\mu} \frac{1}{1 - e^{-2\mu}} \int_0^1 (e_\mu(2+y) + e_\mu(2-y)) \varphi(y) \, dy,$$

and

$$c_2 = \frac{1}{2\mu} \frac{1}{1 - e^{-2\mu}} \int_0^1 (e_\mu(y) + e_\mu(2-y)) \varphi(y) \, dy.$$

Then (10) takes the form

$$\begin{aligned} \psi(x) = \frac{1}{2\mu} \frac{1}{1 - e^{-2\mu}} \int_0^1 \Big[& e_\mu(|x-y|) - e_\mu(2 + |x-y|) + e_\mu(2 - x + y) \\ & + e_\mu(2 - x - y) + e_\mu(x+y) + e_\mu(2 + x - y) \Big] \varphi(y) \, dy. \end{aligned}$$

Expanding $\frac{1}{1-e^{-2\mu}}$ into the geometric series, we obtain

$$\psi(x) = \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_0^1 \big(e_\mu(|2n + x - y|) + e_\mu(|2n + x + y|) \big) \varphi(y) \, dy. \tag{12}$$

By changing variables in the integral, we get

$$\begin{aligned} \psi(x) = & \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_{x-1}^{x} e_\mu(|2n + t|) \varphi(x - t) \, dt \\ & + \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_{x}^{x+1} e_\mu(|2n + t|) \varphi(t - x) \, dt. \end{aligned}$$

Differentiating the above under the integral sign leads to

$$\begin{aligned} \psi'(x) = & \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_{x-1}^{x} e_\mu(|2n + t|) \varphi'(x - t) \, dt \\ & - \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_{x}^{x+1} e_\mu(|2n + t|) \varphi'(t - x) \, dt, \end{aligned}$$

since all the "non-integral" terms cancel out. Similarly,

$$\begin{aligned} \psi''(x) = & \frac{1}{\mu} \sum_{n \in \mathbb{Z}} e_\mu(|2n + x|) \varphi'(0) - \frac{1}{\mu} \sum_{n \in \mathbb{Z}} e_\mu(|2n - 1 + x|) \varphi'(1) \\ & + \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_{x-1}^{x} e_\mu(|2n + t|) \varphi''(x - t) \, dt \\ & + \frac{1}{2\mu} \sum_{n \in \mathbb{Z}} \int_{x}^{x+1} e_\mu(|2n + t|) \varphi''(t - x) \, dt. \end{aligned}$$

Now, returning to the original variables, we can write

$$\psi'(x) = \frac{1}{2\mu}\sum_{n\in\mathbb{Z}}\int_0^1 \big(\mathrm{e}_\mu(|2n+x-y|) - \mathrm{e}_\mu(|2n+x+y|)\big)\varphi'(y)\,\mathrm{d}y, \tag{13}$$

and

$$\begin{aligned}\psi''(x) &= \frac{1}{\mu}\sum_{n\in\mathbb{Z}}\mathrm{e}_\mu(|2n+x|)\varphi'(0) - \frac{1}{\mu}\sum_{n\in\mathbb{Z}}\mathrm{e}_\mu(|2n-1+x|)\varphi'(1)\\ &\quad + \frac{1}{2\mu}\sum_{n\in\mathbb{Z}}\int_0^1 \big(\mathrm{e}_\mu(|2n+x-y|) + \mathrm{e}_\mu(|2n+x+y|)\big)\varphi''(y)\,\mathrm{d}y.\end{aligned} \tag{14}$$

In order to estimate the $W^{2,1}$ norm of ψ, we at first estimate $C_1 = C_1(y) := \sum_{n\in\mathbb{Z}}\int_0^1 \big|\mathrm{e}_\mu(|2n+x-y|)\big|\,\mathrm{d}x$ for $y\in[0,1]$. Observe that

$$\int_0^1 \big|\mathrm{e}_\mu(|2n+x-y|)\big|\,\mathrm{d}x = \begin{cases} \frac{1}{m}\big(2-\mathrm{e}^{-my}-\mathrm{e}^{-m(1-y)}\big), & n=0,\\ \frac{1}{m}\mathrm{e}^{-2mn}\mathrm{e}^{my}(1-\mathrm{e}^{-m}), & n\geqslant 1,\\ \frac{1}{m}\mathrm{e}^{2mn}\mathrm{e}^{-my}(\mathrm{e}^{m}-1), & n\leqslant -1.\end{cases}$$

Hence

$$C_1 \leqslant \frac{1}{m}\left[2+\frac{\mathrm{e}^{-m(2-y)}}{1+\mathrm{e}^{-m}}+\frac{\mathrm{e}^{-my}}{1+\mathrm{e}^{m}}\right] \leqslant \frac{4}{m}.$$

Very similar calculations show that

$$C_2 := \sum_{n\in\mathbb{Z}}\int_0^1 \big|\mathrm{e}_\mu(|2n+x+y|)\big|\,\mathrm{d}x \leqslant \frac{1}{m}\left[\frac{\mathrm{e}^{-my}}{1+\mathrm{e}^{-m}}+\frac{\mathrm{e}^{my}}{1+\mathrm{e}^{m}}\right] \leqslant \frac{2}{m}.$$

Since $m = \sqrt{|\lambda|}\cos(\theta/2) \geqslant \sqrt{|\lambda|}\cos(\frac{\pi}{4}+\frac{\delta}{2})$, by the Fubini theorem applied to (12) we have

$$\|\psi\|_{L^1(0,1)} \leqslant \frac{3}{|\lambda|\cos(\frac{\pi}{4}+\frac{\delta}{2})}\|\varphi\|_{L^1(0,1)}.$$

In the same way, by (13) we obtain

$$\|\psi'\|_{L^1(0,1)} \leqslant \frac{3}{|\lambda|\cos(\frac{\pi}{4}+\frac{\delta}{2})}\|\varphi'\|_{L^1(0,1)}.$$

Moreover, by the fact that $\varphi'\in W^{1,1}(0,1)$ and the Sobolev embedding theorem, $\varphi'\in C[0,1]$ and $\|\varphi'\|_{C[0,1]}\leqslant M\|\varphi'\|_{W^{1,1}(0,1)}$ for some constant $M>0$. Thus, by (14),

$$\|\psi''\|_{L^1(0,1)} \leqslant \frac{6M}{|\lambda|\cos(\frac{\pi}{4}+\frac{\delta}{2})}(\|\varphi'\|_{W^{1,1}(0,1)}+\|\varphi''\|_{L^1(0,1)}),$$

since we can assume that $M\geqslant\frac{1}{2}$. Finally, for $\lambda\in\Sigma_{\frac{\pi}{2}+\delta}$,

$$\|\psi\|_{W^{2,1}(0,1)} \leqslant \frac{12M}{|\lambda|\cos(\frac{\pi}{4}+\frac{\delta}{2})}\|\varphi\|_{W^{2,1}(0,1)},$$

which completes the proof. □

Let G_j be a version of G in $L^1(E_j)$, $j \in \mathscr{N}$ and let B be the operator in $L^1(S)$ defined by $B\varphi = \sigma\varphi''$ with the domain composed of functions $\varphi = (\varphi_j)_{j\in\mathscr{N}}$ where $\varphi_j \in D(G_j)$. Since B is equal to G_j on each E_j, operator B generates a strongly continuous analytic semigroup $\{\mathrm{e}^{tB}, t \geqslant 0\}$ and the same is true for operators $\kappa_n B$, $n \geqslant 1$. Moreover, by (9), the following limit

$$\lim_{n\to\infty} \lambda(\lambda - \kappa_n B)^{-1}\varphi =: P\varphi, \qquad \varphi \in L^1(S),$$

exists, and $P\varphi = P(\varphi_j)_{j\in\mathscr{N}} = \left(d_j^{-1}\int_{E_j}\varphi_j\right)_{j\in\mathscr{N}}$. Evidently, range P is the space of functions that are constant on each E_j. Let us denote this space by $L^1_0(S)$.

Rewrite conditions (5) in the form

$$\varphi'(L_j) = F_{L,j}\varphi,$$
$$\varphi'(R_j) = F_{R,j}\varphi,$$

$j \in \mathscr{N}$, where the functionals $F_{L,j}, F_{R,j} \in (W^{2,1}(S))^*$ are given by

$$F_{L,j}\varphi = l_j\varphi(L_j) - \sum_{i\in I_j^L}^{*}\left[\frac{\sigma_i}{\sigma_j}l_{ij}\varphi(L_i) + \frac{\sigma_i}{\sigma_j}r_{ij}\varphi(R_i)\right],$$

$$F_{R,j}\varphi = \sum_{i\in I_j^R}^{*}\left[\frac{\sigma_i}{\sigma_j}l_{ij}\varphi(L_i) + \frac{\sigma_i}{\sigma_j}r_{ij}\varphi(R_i)\right] - r_j\varphi(R_j).$$

Using functions similar to $[\alpha,\beta] \ni x \mapsto \frac{1}{\gamma}(x-\alpha)^2\sin\gamma(x-\beta)$, we can pick $h_{L,j}, h_{R,j} \in W^{2,1}(E_j)$, $j \in \mathscr{N}$ with an arbitrarily small norm such that $h_{L,j}(L_j) = h_{L,j}(R_j) = h_{R,j}(L_j) = h_{R,j}(R_j) = 0$ and $h'_{L,j}(L_j) = h'_{R,j}(R_j) = 1$, $h'_{L,j}(R_j) = h'_{R,j}(L_j) = 0$. Let J be the bounded linear operator $W^{2,1}(S) \to W^{2,1}(S)$ given by

$$J\varphi = \left((F_{L,j}\varphi)h_{L,j} + (F_{R,j}\varphi)h_{R,j}\right)_{j\in\mathscr{N}},$$

and for $\kappa \neq 0$ define $I_\kappa\colon W^{2,1}(S) \to W^{2,1}(S)$ by

$$I_\kappa = I_{W^{2,1}(S)} - \frac{1}{\kappa}J.$$

Here, $I_{W^{2,1}(S)}$ is the identity operator in $W^{2,1}(S)$. Then, it is easy to observe (see A. Bobrowski (2012), Lemma 3.1) that I_κ is an isomorphism of $W^{2,1}(S)$ with the inverse

$$I_\kappa^{-1} = I_{W^{2,1}(S)} + \frac{1}{\kappa}J. \tag{15}$$

Let $K\colon W^{2,1}(S) \to W^{2,1}(S)$ be defined as J with h's replaced by their second derivatives, i.e.

$$K\varphi = \left((F_{L,j}\varphi)h''_{L,j} + (F_{R,j}\varphi)h''_{R,j}\right)_{j\in\mathscr{N}}.$$

Finally, let $\widetilde{A} = \sigma K$.

We are now able to show that the resolvent of A^* exists, as claimed in Lemma 4.1.

Proof of Lemma 4.1. We proceed exactly as in A. Bobrowski (2012) pp. 1507-1508. Let $\widetilde{B} = I_1 A_0^* I_1^{-1}$, where A_0^* is the part of A^* in $W^{2,1}(S)$, be the operator in $W^{2,1}(S)$. We have

$$A_0^* I_1^{-1} = B_0 + \sigma K, \tag{16}$$

where B_0 is the part of B in $W^{2,1}(S)$. Moreover, $D(\widetilde{B}) = D(B_0)$, and by (16), for $\varphi \in D(B)$,

$$\widetilde{B}\varphi = B\varphi + C\varphi + D\varphi,$$

where $C\varphi = -J(\sigma\varphi'') = -JB_0\varphi$ and $D\varphi = \sigma K\varphi - J(\sigma K\varphi)$. Observe that if $\varphi \in D(B_0)$,

$$\|C\varphi\|_{W^{2,1}(S)} \leqslant M\|B_0\varphi\|_{W^{2,1}(S)},$$

for $M := \|J\|$, i.e. the operator C is B_0-bounded. Moreover,

$$M \leqslant \max_{j\in\mathscr{N}}\{\max(\|h_{L,j}\|_{W^{2,1}(S)}, \|h_{R,j}\|_{W^{2,1}(S)})\} \cdot \max_{j\in\mathscr{N}}\{\|F_{L,j}\| + \|F_{R,j}\|\}.$$

We will now prove that the operator $B_0 + C$ is sectorial. By Lemma 4.3, B_0 is sectorial with angle $\pi/2$, i.e. there exists $\widetilde{M} > 0$ such that $\|(\lambda - B_0)^{-1}\| \leqslant \widetilde{M}/|\lambda|$ for $\lambda \in \Sigma_{\pi/2+\delta}$, $\delta \in (0, \pi/2]$. Hence, using the fact that
$B_0(\lambda - B_0)^{-1} = \lambda(\lambda - B_0)^{-1} - I_{W^{2,1}(S)}$, we have

$$\|C(\lambda - B_0)^{-1}\| \leqslant M\|\lambda(\lambda - B_0)^{-1} - I_{W^{2,1}(S)}\| \leqslant M(\widetilde{M} + 1).$$

Thus, if we pick $h_{L,j}, h_{R,j}$, $j \in \mathscr{N}$ in such a way that $\|J\| \leqslant \frac{1}{\widetilde{M}+1}$, then
$q := \|C(\lambda - B_0)^{-1}\| < 1$ which implies $\lambda \in \rho(B_0 + C)$. Therefore we have the Neumann series expansion

$$(\lambda - B_0 - C)^{-1} = (\lambda - B_0)^{-1}\sum_{i=0}^{\infty}[C(\lambda - B_0)^{-1}]^i,$$

which implies

$$\|(\lambda - B_0 - C)^{-1}\| \leqslant \frac{\widetilde{M}}{1-q}\frac{1}{|\lambda|}.$$

This means that $B_0 + C$ is sectorial and, being densely defined, generates an analytic semigroup in $W^{2,1}(S)$. What is more, the operator D is bounded since J and K are. Hence, by K.J. Engel, R. Nagiel (2000, Proposition III.1.12) the operator $\widetilde{B}$ generates an analytic semigroup in $W^{2,1}(S)$ and so does A_0^*. In particular, $(\lambda - A_0^*)D(A_0) = W^{2,1}$.

Finally, observe that the operator A^* is dissipative. Let $A^\#$ be the adjoint operator of A. Since $\|(\lambda - A)^{-1}\| \leqslant \frac{1}{\lambda}$, we also have $\left\|\left[(\lambda - A)^{-1}\right]^*\right\| = \|(\lambda - A^\#)^{-1}\| \leqslant \frac{1}{\lambda}$ in the dual space $C[0,1]^* = M_b([0,1])$ of regular Borel measures. Thus $\|(\lambda - A^\#)\mu\| \geqslant \lambda\|\mu\|_{M_b}$ for $\mu \in D(A^\#) \subset M_b([0,1])$. Assume that μ is an absolutely continuous measure with respect to the Lebesgue measure, i.e. there exists

$\varphi_\mu \in L^1(0,1)$ such that $\mu(E) = \int_E \varphi_\mu$ for any Borel measurable set $E \subset [0,1]$. Then $\|\mu\|_{M_b} = \|\varphi_\mu\|_{L^1(0,1)}$. By Lemma 3.1 we know that $\int_S \varphi(\lambda - A)f = \int_S f(\lambda - A^*)\varphi$ for $f \in D(A)$, $\varphi \in D(A^*)$ and hence $(\lambda - A^\#)\mu = (\lambda - A^*)\varphi_\mu$ for absolutely continuous $\mu \in M_b([0,1])$. This implies $\|(\lambda - A^*)\varphi\| \geqslant \lambda\|\varphi\|_{L^1(0,1)}$, $\varphi \in D(A^*)$ and proves that A^* is dissipative. Furthermore, A^* is also a closed operator. Thus, since $W^{2,1}(S) = \mathrm{range}(\lambda - A_0^*) \subset \mathrm{range}(\lambda - A^*)$ is dense in $L^1(S)$, we have $\mathrm{range}(\lambda - A) = L^1(S)$ by K.J. Engel, R. Nagiel (2000, Proposition II.3.14(iii)), which completes the proof. □

Theorem 4.4. *The operator A^* generates a sub-Markov semigroup. Moreover, if the semigroup generated by A is conservative, then A^* generates a Markov semigroup, i.e. we have $\int_S \mathrm{e}^{tA^*}\varphi = \int_S \varphi$ for non-negative $\varphi \in L^1(S)$.*

Proof. By the remarks preceding the theorem, we are left with proving that $\lambda(\lambda - A^*)^{-1}$ are sub-Markov for $\lambda > 0$.

Let us now prove that if $\psi \geqslant 0$, $\psi \in L^1(S)$, then $(\lambda - A^*)^{-1}\psi \geqslant 0$, for $\lambda > 0$. Suppose, contrary to our claim, that there exist a function $\psi \geqslant 0$, a set $\Gamma \subset S$ with positive Lebesgue measure and a real number $\delta > 0$ such that $(\lambda - A^*)^{-1}\psi < -\delta$ a.e. on Γ for some $\lambda > 0$. Without loss of generality, we may assume that Γ is a subset of some edge E_i. Then, for a given $\varepsilon > 0$ we choose an open set G and a closed set Γ' such that $\Gamma' \subset \Gamma \subset G \subset E_i$ and $\mu(G \setminus \Gamma') < \varepsilon$. By the Urysohn lemma, there exists a continuous real function $0 \leqslant f \leqslant 1$ with $f \equiv 1$ on Γ' and $f \equiv 0$ outside G. Since A generates a Feller semigroup, we have $(\lambda - A)^{-1}f \geqslant 0$, hence

$$\begin{aligned} 0 &\leqslant \int_S (\lambda - A)^{-1}f \cdot \psi = \int_S f \cdot (\lambda - A^*)^{-1}\psi \\ &= \int_{\Gamma'} f \cdot (\lambda - A^*)^{-1}\psi + \int_{G\setminus\Gamma'} f \cdot (\lambda - A^*)^{-1}\psi \\ &\leqslant -\delta\mu(\Gamma') + \int_{G\setminus\Gamma'} (\lambda - A^*)^{-1}\psi. \end{aligned}$$

By the Lebesgue dominated convergence theorem and the fact that $(\lambda - A^*)^{-1}\varphi \in L^1(S)$, letting $\varepsilon \to 0$ leads to a contradiction.

Finally, let $\psi \in L^1(S)$. Since A generates a Feller semigroup we have

$$(\lambda - A)^{-1}\mathbb{1}_S = \int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{tA}\mathbb{1}_S\,\mathrm{d}t \leqslant \mathbb{1}_S \int_0^\infty \mathrm{e}^{-\lambda t}\,\mathrm{d}t = \tfrac{1}{\lambda}\mathbb{1}_S. \tag{17}$$

Thus

$$\int_S \lambda(\lambda - A^*)^{-1}\psi = \int_S \lambda(\lambda - A)^{-1}\mathbb{1}_S \cdot \psi \leqslant \int_S \psi,$$

which completes the first part of the proof.

If we assume that the semigroup generated by A is conservative, then inequality (17) becomes equality, and $\int_S \lambda(\lambda - A^*)^{-1}\psi = \int_S \psi$. □

5 Convergence

To prove similar convergence result as in Theorem 2.1, we begin with a theorem due to Kurtz (cf. S.N. Etheir, T.G. Kurtz (1986), Theorem 7.6). Let $\mathscr{A}_n$, $n \geqslant 1$ be the generators of equi-bounded semigroups $\{\mathrm{e}^{t\mathscr{A}_n}, t \geqslant 0\}$ in a Banach space $\mathbb{X}$ and suppose that an operator $\mathscr{Q}$ generates a strongly continuous semigroup $\{\mathrm{e}^{t\mathscr{Q}}, t \geqslant 0\}$ such that for $x \in \mathbb{X}$ the limit

$$\lim_{\lambda \to 0^+} \lambda(\lambda - \mathscr{Q})^{-1} x =: \mathscr{P} x,$$

exists. Moreover, assume that $(\varepsilon_n)_{n \geqslant 1}$ is a sequence of positive real numbers converging to 0.

Theorem 5.1. *Let $\mathscr{A}$ be an operator in a Banach space $\mathbb{X}$, D be a subset of its domain and assume that*

(i) *if $x \in D$, then $(x, \mathscr{A} x) \in \mathscr{A}_{ex}$ where $\mathscr{A}_{ex}$ is the extended limit of $\mathscr{A}_n$, $n \geqslant 1$,*

(ii) *if y is in a core D' of $\mathscr{Q}$, then $(y, \mathscr{Q} y) \in \mathscr{B}_{ex}$ where $\mathscr{B}_{ex}$ is the extended limit of $\varepsilon_n \mathscr{A}_n$, $n \geqslant 1$,*

(iii) *for $\mathbb{X}' := \operatorname{range} \mathscr{P}$ the operator $\mathscr{P}\mathscr{A}$ with domain $D \cap \mathbb{X}'$ is closable and its closure $\overline{\mathscr{P}\mathscr{A}}$ generates a strongly continuous semigroup in $\mathbb{X}'$.*

Then for $x \in \mathbb{X}$ and $t > 0$

$$\lim_{n \to \infty} \mathrm{e}^{t\mathscr{A}_n} x = \mathrm{e}^{t\overline{\mathscr{P}\mathscr{A}}} \mathscr{P} x,$$

almost uniformly in $(0, \infty)$.

As in Theorem 2.1 let $(\kappa_n)_{n \geqslant 1}$ be a sequence of positive real numbers such that $\lim_{n \to \infty} \kappa_n = \infty$, and let A_n^* be the operator defined by (4) with σ replaced by $\kappa_n \sigma$ and the domain composed of $W^{2,1}(S)$ functions satisfying the transmission conditions (5) with permeability coefficients divided by κ_n.

In order to verify conditions (i)- (iii) of the Kurtz theorem we need the following two lemmas.

Lemma 5.2. *We have $L_0^1(S) \subset D(A_{ex})$, and if $u \in L_0^1(S)$, then*

$$A_{ex} u = \widetilde{A} u,$$

where A_{ex} is the extended limit of A_n^*, $n \geqslant 1$.

Proof. Given $u \in L_0^1(S)$, set $\varphi_n = I_{\kappa_n}^{-1} u$. Then $\psi_n \to u$ as $n \to \infty$ by (15), for κ_n^{-1} converges to 0 and J is bounded. What is more,

$$A_n^* \varphi_n = \sigma \kappa_n \varphi_n'' + \sigma (J \varphi_n)'' = \sigma \kappa_n \varphi_n'' + \sigma K \varphi_n.$$

Hence $\lim_{n \to \infty} A_n^* \varphi_n = \sigma K u$ which completes the proof. □

Lemma 5.3. *If $\varphi \in D(B)$, then*

$$B_{ex}\varphi = B\varphi,$$

where B_{ex} is the extended limit of $\kappa_n^{-1}A_n^*$

Proof. Observe that if $\varphi \in D(B)$, then $\varphi_n := I_{\kappa_n}^{-1}\varphi \in D(A_n^*)$, $n \geqslant 1$. Evidently, by the same argument as in the previous lemma, we have $\varphi_n \to \varphi$, $n \to \infty$, and hence

$$\lim_{n\to\infty} \kappa_n^{-1}A_n^*\varphi_n = B\varphi,$$

as claimed. □

We are now ready to apply Theorem 5.1. First, let us observe that for $\varphi \in W^{2,1}(S)$ and $j \in \mathscr{N}$, we have

$$\int_{E_j}\big((F_{L,j}\varphi)h''_{L,j} + (F_{R,j}\varphi)h''_{R,j}\big) = d_j^{-1}F_{R,j}\varphi - d_j^{-1}F_{L,j}\varphi.$$

Thus, when $\varphi \in L_0^1(S)$, we obtain

$$\begin{aligned}
\sigma PK\varphi &= \Big(\sigma_j d_j^{-1}(F_{R,j}\varphi - F_{L,j}\varphi)\Big)_{j\in\mathscr{N}} \\
&= \Bigg(d_j^{-1}\sum_{i\in I_j^E}^{*}\big[\sigma_i l_{ij}\varphi(L_i) + \sigma_i r_{ij}\varphi(R_i)\big] \\
&\quad - \sigma_j d_j^{-1} l_j\varphi(L_j) - \sigma_j d_j^{-1} r_j\varphi(R_j)\Bigg)_{j\in\mathscr{N}} \\
&= \Bigg(d_j^{-1}\sum_{i\in I_j^E}^{*}\sigma_i(l_{ij}+r_{ij})\varphi(V_i) - \sigma_j d_j^{-1}(l_j + r_j)\varphi(V_j)\Bigg)_{j\in\mathscr{N}},
\end{aligned}$$

where I_j^E is the set of indexes $i \neq j$ of edges incident to E_j. Hence, we can identify the operator $Q := \sigma(PK)_{|L_0^1(S)}$ with the matrix $(q_{ji})_{j,i\in\mathscr{N}}$, such that $q_{ji} = d_j^{-1}\sigma_i(l_{ij}+r_{ij})$ for $j \neq i$, and $q_{ji} = -\sigma_j d_j^{-1}(l_j + r_j)$ for $j = i$. Since this matrix is finite, the operator Q generates strongly continuous semigroup $\{e^{tQ}, t > 0\}$. Therefore the operator $P\widetilde{A} = \sigma PK = Q$ is closed.

Finally, setting $\mathbb{X} = L^1(S)$, $D = L_0^1(S)$, we check that, by Lemma 5.2, Lemma 5.3 and the above remarks, all the conditions (i), (ii) and (iii) are satisfied with $\mathscr{A} := \widetilde{A}$, $\mathscr{A}_n := A_n^*$, $\mathscr{Q} := Q$, $\mathscr{P} = P$, and $\varepsilon_n := \kappa_n^{-1}$. Hence we obtain the following "dual" result to Theorem 2.1.

Theorem 5.4. *Let* $(\kappa_n)_{n\geqslant 1}$, *and* A_n^* *be the defined as in the beginning of this section. Then,*

$$\lim_{n\to\infty} \mathrm{e}^{tA_n^*} f = \mathrm{e}^{tQ} P f, \quad f \in L^1(S),\ t > 0.$$

Moreover, from the proof of the Kurtz theorem, the limit is almost uniform in $t \in (0,\infty)$.

References

Bobrowski A., *From diffusions on graphs to Markov chains via asymptotic state lumping*, "Annales Henri Poincaré", 2012 vol. 13(6), pp. 1501-1510.

Bobrowski A., Morawska K., *From a PDE model to an ODE model of dynamics of synaptic depression*, "Disc. and Cont. Dyn. Systems B", 2012 vol. 17(6), pp. 2313-2327.

Engel K.J., Nagel R., *One-parameter semigroups for linear evolution equations*, Springer-Verlag, New York, 2000.

Ethier S.N., Kurtz T.G., *Markov processes. Characterization and convergence*, Wiley, New York, 1986.

Lasota A., Mackey, M.C., *Probabilistic properties of deterministic systems*, Cambridge University Press, Cambridge, 1985.

Lunardi A., *Analytic semigroups and optimal regularity in parabolic problems*, Springer Basel AG, Basel, 1995.

Mugnolo D., *Semigroup methods for evolution equations on networks*, preprint, pp. 1-75, 2007.

Małgorzata Murat [1]

Incomplete moments of non-zero inflated modified power series distribution

Keywords: inflated distribution, lower partial (incomplete) moments, upper partial (incomplete) moments, factorial partial (incomplete) moments, recurrence relations, modified power series distributions, generalized Poisson distribution, generalized negative binomial distribution, generalized logarithmic series, lost game distributions

Abstract

The paper contains recurrence formulae for lower and upper partial moments and lower and upper partial descending and ascending factorial moments of non-zero inflated modified power series distributions.

1 Introduction

Modified power series distributions, introduced by Gupta (1974), form a useful subclass of one-parameter discrete exponential families suitable for modeling count data. An inflated modified power series distribution is a mixture of a modified power series distribution and a degenerate distribution at one of the support points s, with a mixing probability α for the degenerate distribution. This distribution is useful for modeling count data that may have extra observations. Count data with many zeros are common in a wide variety of disciplines. Application areas of inflated distributions are diverse and have included manufacturing defects (Lambert, 1992), patent applications (Crepon and Duguet, 1997), road safety (Miaou, 1994), species abundance (Welsh et al., 1996; Faddy, 1998), medical consultations (Gurmu, 1997), use of recreational facilities (Gurmu and Trivedi, 1996; Shonkwiler and Shaw, 1996), number of roots produced by a certain species of apple tree (Ridout, Demétrio, and Hinde, 1998). In probability and statistical literature one can also find many real and interesting examples of non-zero inflated distributions (Pandey, 1964-65; Murat and Szynal, 1998). A fair amount of probabilistic and statistical methodology has been developed to deal with such data.

[1] Department of Mathematics, Technical University of Lublin, Nadbystrzycka 38, 20-618 Lublin, *e-mail: m.murat@pollub.pl*

Gupta, Gupta, Thripati (1995) gave recurrence relations for ordinary, factorial and central moments of zero inflated modified power series distributions, while Murat and Szynal (1998) obtained recurrence relations for moments of non- zero inflated modified power series distributions. In the current paper those formulas are completed with relations for lower and upper partial (incomplete) ordinary moments and lower and upper partial (incomplete) factorial moments of non-zero inflated modified power series distribution. Partial moments can be use to define a risk measure (Stone, 1973; Nawrocki, 1991), to characterize and estimate asymmetric effects of inputs on output distributions (Antile, 2010). They are also applied to such problems as insurance purchasing (Hamburg and Matlack, 1968), bayesian point estimation (Britney and Winkler, 1968), inventory theory (Arrow, K. J., Karlin, S. and Scarf, H., 1958), theory of the firm (Horowitz 1970), stopping rules (Hayes, 1969; DeGroot, 1968).

Section 2 establishes basic notations and definitions. In Section 3 recurrence relations for lower partial ordinary and factorial moments about a point c are derived. Some illustrative examples are given in Section 4. Upper partial ordinary and factorial moments about a point c are given in Section 4.

2 Definitions and notations

We begin with definition of an inflated modified power series distribution. Let $\mathbb{N}$ be a set of nonnegative integers and $s \in \mathbb{N}$ be fixed.

Definition 1 *A discrete random variable X is said to have **an inflated modified power series distribution** (IMPSD) if its probability mass function is given by*

$$\mathbf{P}[X=x] = \begin{cases} \beta + \alpha \frac{[g(\theta)]^x a(x)}{f(\theta)}, & x = s, \\ \alpha \frac{a(x)[g(\theta)]^x}{f(\theta)}, & x \neq s,\ x \in \mathbb{N};\ 0 < \alpha \leq 1, \end{cases} \tag{1}$$

where $\beta = 1-\alpha$, $f(\theta) = \sum\limits_{x=0}^{\infty} a(x)[g(\theta)]^x$, $g(\theta)$ is positive, finite and differentiable and coefficients $a(x)$ are free of θ.

Next we establish notations and definitions of partial moments. Let $c, r \in \mathbb{N}$.

Definition 2 ***The** r**-th lower partial moment about** c (r-th incomplete moment on the left about c) is defined by*

$$\mu_r(t) = \sum_{x=0}^{t} (x-c)^r \mathbf{P}[X=x],\ t \in \mathbb{N}. \tag{2}$$

Definition 3 ***The r-th upper partial moment about*** *c (r-th incomplete moment on the right about c) is defined by*

$$\mu^r(t) = \sum_{x=t}^{\infty} (x-c)^r \mathbf{P}[X=x],\ t \in \mathbb{N}. \tag{3}$$

Definition 4 ***The r-th lower partial descending factorial moment about*** *c (r-th incomplete descending factorial moment on the left about c) is defined by*

$$\mu_{(r)}(t) = \sum_{x=0}^{t} (x-c)^{(r)} \mathbf{P}[X=x],\ t \in \mathbb{N}, \tag{4}$$

where $y^{(r)} = y(y-1)(y-2)\cdots(y-r+1)$.

Definition 5 ***The r-th upper partial descending factorial moment about*** *c (r-th incomplete descending factorial moment on the right about c) is defined by*

$$\mu^{(r)}(t) = \sum_{x=t}^{\infty} (x-c)^{(r)} \mathbf{P}[X=x],\ t \in \mathbb{N}. \tag{5}$$

Definition 6 ***The r-th lower partial ascending factorial moment about*** *c (r-th incomplete ascending factorial moment on the left about c) is defined by*

$$\mu_{[r]}(t) = \sum_{x=0}^{t} (x-c)^{[r]} \mathbf{P}[X=x],\ t \in \mathbb{N}, \tag{6}$$

where $y^{[r]} = y(y+1)(y+2)\cdots(y+r-1)$.

Definition 7 ***The r-th upper partial ascending factorial moment about*** *c (r-th incomplete ascending factorial moment on the right about c) is defined by*

$$\mu^{[r]}(t) = \sum_{x=t}^{\infty} (x-c)^{[r]} \mathbf{P}[X=x],\ t \in \mathbb{N}. \tag{7}$$

3 Recurrence relations for lower partial moments of IMPSD

In this Section using method introduced by Gupta (1974) we obtain relations for lower partial moments about c. It is obvious that if $t < s$ then the lower partial moments are independent of a point of inflation s. In this situation recurrence relations for lower partial moments of IMPSD concur with formulae obtained by Gupta (1974). So we consider only the situation when a point of inflation s is less than a limit t.

Theorem 8 *The $r+1$-th lower partial moment about a point c of IMPSD is given by*

$$\mu_{r+1}(t)=\frac{g(\theta)}{g'(\theta)}\frac{d\mu_r(t)}{d\theta}+\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-c\right)\mu_r(t)-\beta(s-c)^r\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right). \tag{8}$$

Proof. Observe that for IMPSD with $s<t$ we have

$$\mu_r(t)=\beta(s-c)^r+\alpha\sum_{x=0}^{t}(x-c)^r\frac{[g(\theta)]^x a(x)}{f(\theta)}. \tag{9}$$

Differentiation (9) with respect to θ gives

$$\begin{aligned}\frac{d\mu_r(t)}{d\theta}&=\alpha\sum_{x=0}^{t}(x-c)^r a(x)\frac{x[g(\theta)]^{x-1}g'(\theta)}{f(\theta)}-\alpha\sum_{x=0}^{t}(x-c)^r a(x)\frac{[g(\theta)]^x f'(\theta)}{[f(\theta)]^2}\\&=\alpha\frac{g'(\theta)}{g(\theta)}\sum_{x=0}^{t}(x-c)^{r+1}a(x)\frac{[g(\theta)]^x}{[f(\theta)]}\\&+\alpha\left(c\frac{g'(\theta)}{g(\theta)}-\frac{f'(\theta)}{f(\theta)}\right)\sum_{x=0}^{t}(x-c)^r a(x)\frac{[g(\theta)]^x}{[f(\theta)]}.\end{aligned}$$

Rearranging (9) we get

$$\alpha\sum_{x=0}^{t}(x-c)^r\frac{[g(\theta)]^x a(x)}{f(\theta)}=\mu_r(t)-\beta(s-c)^r.$$

This gives

$$\begin{aligned}\frac{d\mu_r(t)}{d\theta}&=\frac{g'(\theta)}{g(\theta)}\left[\mu_{r+1}(t)-\beta(s-c)^{r+1}\right]\\&\quad+\left(c\frac{g'(\theta)}{g(\theta)}-\frac{f'(\theta)}{f(\theta)}\right)\left[\mu_r(t)-\beta(s-c)^r\right].\end{aligned}$$

Hence after some simply calculations we get (8). ∎

Putting in (8) $c=0$ we get recurrence relations for lower partial moments of IMPSD.

Corollary 9 *The $r+1$-th lower partial ordinary moment of IMPSD is given by*

$$m_{r+1}(t)=\frac{g(\theta)}{g'(\theta)}\frac{dm_r(t)}{d\theta}+\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}m_r(t)-\beta s^r\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right) \tag{10}$$

Now we derive recurrence relation for lower partial descending factorial moments.

Theorem 10 *The* $r+1$*-th lower partial descending factorial moment about a point* c *of IMPSD is given by*

$$\begin{aligned}\mu_{(r+1)}(t) &= \frac{g(\theta)}{g'(\theta)}\frac{d\mu_{(r)}(t)}{d\theta}+\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-c-r\right)\mu_{(r)}(t)\\ &\quad -\beta(s-c)^{(r)}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right).\end{aligned} \tag{11}$$

Proof. For IMPSD with $s<t$ we have

$$\mu_{(r)}(t)=\beta(s-c)^{(r)}+\alpha\sum_{x=0}^{t}(x-c)^{(r)}\frac{[g(\theta)]^x a(x)}{f(\theta)}. \tag{12}$$

Differentiation (12) with respect to θ gives

$$\begin{aligned}\frac{d\mu_{(r)}(t)}{d\theta} &= \alpha\frac{g'(\theta)}{g(\theta)}\sum_{x=0}^{t}(x-c)^{(r+1)}a(x)\frac{[g(\theta)]^x}{[f(\theta)]}\\ &\quad +\alpha\left((c+r)\frac{g'(\theta)}{g(\theta)}-\frac{f'(\theta)}{f(\theta)}\right)\sum_{x=0}^{t}(x-c)^{(r)}a(x)\frac{[g(\theta)]^x}{[f(\theta)]}.\end{aligned} \tag{13}$$

From (12) it follows

$$\alpha\sum_{x=0}^{t}(x-c)^{(r)}\frac{[g(\theta)]^x a(x)}{f(\theta)}=\mu_{(r)}(t)-\beta(s-c)^{(r)}. \tag{14}$$

Moreover

$$x(x-c)^{(r)}=(x-c)^{(r+1)}+(x-c)^{(r)}(c+r). \tag{15}$$

Combining (14), (15) and (13) we obtain

$$\begin{aligned}\frac{d\mu_{(r)}(t)}{d\theta} &= \frac{g'(\theta)}{g(\theta)}\left[\mu_{(r+1)}(t)-\beta(s-c)^{(r+1)}\right]\\ &\quad +\left((c+r)\frac{g'(\theta)}{g(\theta)}-\frac{f'(\theta)}{f(\theta)}\right)\left[\mu_{(r)}(t)-\beta(s-c)^{(r)}\right].\end{aligned}$$

After some simply calculations we get (11). ■

Putting in (11) $c=0$ we get recurrence relations for lower descending factorial moments of IMPSD.

Corollary 11 *The* $r+1$*-th lower partial descending factorial moment of IMPSD is given by*

$$m_{(r+1)}(t)=\frac{g(\theta)}{g'(\theta)}\frac{dm_{(r)}(t)}{d\theta}+\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-r\right)m_{(r)}(t)-\beta s^{(r)}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right) \tag{16}$$

We end this Section with computation lower partial ascending factorial moments.

Theorem 12 *The $r+1$-th lower partial ascending factorial moment about c of IMPSD is given by*

$$\begin{aligned}\mu_{[r+1]}(t) &= \frac{g(\theta)}{g'(\theta)}\frac{d\mu_{[r]}(t)}{d\theta}+\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-c+r\right)\mu_{[r]}(t)\\ &\quad -\beta(s-c)^{[r]}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right).\end{aligned}\tag{17}$$

Proof. For IMPSD with $s<t$ we have

$$\mu_{[r]}(t)=\beta(s-c)^{[r]}+\alpha\sum_{x=0}^{t}(x-c)^{[r]}\frac{[g(\theta)]^{x}a(x)}{f(\theta)}.\tag{18}$$

Differentiation (18) with respect to θ gives

$$\begin{aligned}\frac{d\mu_{[r]}(t)}{d\theta} &= \alpha\sum_{x=0}^{t}(x-c)^{[r]}a(x)\frac{x[g(\theta)]^{x-1}g'(\theta)}{f(\theta)}\\ &-\alpha\sum_{x=0}^{t}(x-c)^{[r]}a(x)\frac{[g(\theta)]^{x}f'(\theta)}{[f(\theta)]^{2}} = \alpha\frac{g'(\theta)}{g(\theta)}\sum_{x=0}^{t}(x-c)^{[r+1]}a(x)\frac{[g(\theta)]^{x}}{f(\theta)}\\ &+\alpha\left[(c-r)\frac{g'(\theta)}{g(\theta)}-\frac{f'(\theta)}{f(\theta)}\right]\sum_{x=0}^{t}(x-c)^{[r]}a(x)\frac{[g(\theta)]^{x}}{f(\theta)}.\end{aligned}\tag{19}$$

From (18) it follows

$$\alpha\sum_{x=0}^{t}(x-c)^{[r]}\frac{[g(\theta)]^{x}a(x)}{f(\theta)}=\mu_{[r]}(t)-\beta(s-c)^{[r]}.\tag{20}$$

Moreover

$$x(x-c)^{[r]}=(x-c)^{[r+1]}-(x-c)^{[r]}(c-r).\tag{21}$$

Combining (20), (21) and (19) we obtain

$$\begin{aligned}\frac{d\mu_{[r]}(t)}{d\theta} &= \frac{g'(\theta)}{g(\theta)}\left[\mu_{[r+1]}(t)-\beta(s-c)^{[r+1]}\right]\\ &\quad+\left[(c-r)\frac{g'(\theta)}{g(\theta)}-\frac{f'(\theta)}{f(\theta)}\right]\left[\mu_{[r]}(t)-\beta(s-c)^{[r]}\right].\end{aligned}$$

After some simply calculations we get (17). ■

Putting in (17) $c=0$ we get recurrence relations for lower ascending factorial moments of IMPSD.

Corollary 13 *The $r+1$-th lower partial ascending factorial moment of IMPSD is given by*

$$m_{[r+1]}(t)=\frac{g(\theta)}{g'(\theta)}\frac{dm_{[r]}(t)}{d\theta}+\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}+r\right)m_{[r]}(t)-\beta s^{[r]}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right)\tag{22}$$

Remark 14 *From (8) and (16) for $t \to \infty$ we can get relations for complete moments of non-zero IMPSD obtained by Murat and Szynal (1998).*

4 Examples

This Section is devoted to illustrative examples of formulae which we obtained in Section 3. We consider some special cases of IMPSD.

4.1 Inflated generalized Poisson distribution with parameters

Suppose X is a discrete random variable whose probability mass function is given by

$$P[X=x] = \begin{cases} \beta + \alpha \frac{b(b+ax)^{x-1}}{x!}\theta^x e^{-\theta(b+ax)}, & x = s \\ \alpha \frac{b(b+ax)^{x-1}}{x!}\theta^x e^{-\theta(b+ax)}, & x \neq s, \end{cases} \tag{23}$$

for $x = 0,1,2,\ldots$; $\theta > 0, |\ \theta a\ | < 1$, $b > 0$.
In this case we have

$$a(x) = \frac{b(b+ax)^{x-1}}{x!},\ g(\theta) = \theta e^{-a\theta},\ f(\theta) = e^{b\theta}.$$

Using (10) we get the following recurrence relation for lower partial ordinary moments for the inflated generalized Poisson distribution

$$m_{r+1}(t) = \frac{\theta}{1-a\theta}\frac{dm_r(t)}{d\theta} + \frac{b\theta}{1-a\theta}m_r(t) - \beta s^r\left(\frac{b\theta}{1-a\theta} - s\right),\ r > 0.$$

Relations for lower partial descending and ascending factorial moments for inflated generalized Poisson distribution we obtain from (16) and (22)

$$m_{(r+1)}(t) = \frac{\theta}{1-a\theta}\frac{dm_{(r)}(t)}{d\theta} + \left(\frac{b\theta}{1-a\theta} - r\right)m_{(r)}(t) - \beta s^{(r)}\left(\frac{b\theta}{1-a\theta} - s\right),$$

$$m_{[r+1]}(t) = \frac{\theta}{1-a\theta}\frac{dm_{[r]}(t)}{d\theta} + \left(\frac{b\theta}{1-a\theta} + r\right)m_{[r]}(t) - \beta s^{[r]}\left(\frac{b\theta}{1-a\theta} - s\right)$$

for $r > 0$.

Putting in (23) $a = 1$ and $b = 0$ we get inflated Poisson distribution with parameter θ. In this case the relations for lower partial ordinary moments and lower partial factorial moments are as follows

$$m_{r+1}(t) = \frac{\theta}{1-\theta}\frac{dm_r(t)}{d\theta} - \beta s^{r+1},$$

$$m_{(r+1)}(t) = \frac{\theta}{1-\theta}\frac{dm_{(r)}(t)}{d\theta} - \beta s^{(r)}s,$$

$$m_{[r+1]}(t) = \frac{\theta}{1-\theta}\frac{dm_{[r]}(t)}{d\theta} - \beta s^{[r]}s.$$

Putting $s = 0$ we can get partial moments for Poisson distribution.

4.2 Inflated generalized negative binomial distribution

Let a random variable X has the following probability mass function

$$P[X=x]=\begin{cases}\beta+\alpha\frac{n\Gamma(n+bx)[\theta(1-\theta)^{b-1}]^{x}}{x!\Gamma(n+bx-x+1)(1-\theta)^{-n}}, & x=s,\\ \alpha\frac{n\Gamma(n+bx)[\theta(1-\theta)^{b-1}]^{x}}{x!\Gamma(n+bx-x+1)(1-\theta)^{-n}}, & x\neq s,\end{cases}$$

for $x=0,1,2,...;\ 0<\theta<1,\ |\ \theta b\ |<1,\ n>0$.
Here

$$a(x)=\frac{n\Gamma(n+bx)}{x!\Gamma(n+bx-x+1)},\ g(\theta)=\theta(1-\theta)^{b-1},\ f(\theta)=(1-\theta)^{-n}.$$

Using (10) we get the following recurrence relation for lower partial ordinary moments for the inflated negative binomial distribution

$$m_{r+1}(t)=\frac{\theta(1-\theta)}{1-b\theta}\frac{dm_r(t)}{d\theta}+\frac{n\theta}{1-b\theta}m_r(t)-\beta s^r\left(\frac{n\theta}{1-b\theta}-s\right),$$

Relations for lower partial descending and ascending factorial moments for inflated generalized negative binomial distribution we obtain from (16) and (22)

$$m_{(r+1)}(t)=\frac{\theta(1-\theta)}{1-b\theta}\frac{dm_{(r)}(t)}{d\theta}+\left(\frac{n\theta}{1-b\theta}-r\right)m_{(r)}(t)-\beta s^{(r)}\left(\frac{n\theta}{1-b\theta}-s\right),$$

$$m_{[r+1]}(t)=\frac{\theta(1-\theta)}{1-b\theta}\frac{dm_{[r]}(t)}{d\theta}+\left(\frac{n\theta}{1-b\theta}+r\right)m_{[r]}(t)-\beta s^{[r]}\left(\frac{n\theta}{1-b\theta}-s\right),$$

We have some special cases:
(a) If $b=0$ then X has inflated binomial distribution with p.f.

$$P[X=x]=\begin{cases}\beta+\alpha\binom{n}{x}\theta^x(1-\theta)^{n-x}, & x=s,\\ \alpha\binom{n}{x}\theta^x(1-\theta)^{n-x}, & x\neq s.\end{cases}$$

In this case

$$a(x)=\binom{n}{x},\ g(\theta)=\frac{\theta}{1-\theta},\ f(\theta)=(1-\theta)^{-n}.$$

Hence we have the following relations for lower partial ordinary and lower partial factorial moments

$$m_{r+1}(t)=\theta(1-\theta)\frac{dm_r(t)}{d\theta}+n\theta m_r(t)-\beta s^r(n\theta-s),$$

$$m_{(r+1)}(t)=\theta(1-\theta)\frac{dm_{(r)}(t)}{d\theta}+(n\theta-r)m_{(r)}(t)-\beta s^{(r)}(n\theta-s),$$

$$m_{[r+1]}(t) = \theta(1-\theta)\frac{dm_{[r]}(t)}{d\theta} + (n\theta + r)\, m_{[r]}(t) - \beta s^{[r]}\,(n\theta - s)\,.$$

(b) If $b = 1$ then X has inflated negative binomial distribution with

$$a(x) = \binom{n+x-1}{x},\ g(\theta) = \theta,\ f(\theta) = (1-\theta)^{-n}.$$

In this cases we obtain the relations for lower partial ordinary and factorial moments

$$m_{r+1}(t) = \frac{\theta(1-\theta)}{1-\theta}\frac{dm_r(t)}{d\theta} + \frac{n\theta}{1-\theta}m_r(t) - \beta s^r\left(\frac{n\theta}{1-\theta} - s\right),$$

$$m_{(r+1)}(t) = \frac{\theta(1-\theta)}{1-\theta}\frac{dm_{(r)}(t)}{d\theta} + \left(\frac{n\theta}{1-\theta} - r\right)m_{(r)}(t) - \beta s^{(r)}\left(\frac{n\theta}{1-\theta} - s\right),$$

$$m_{[r+1]}(t) = \frac{\theta(1-\theta)}{1-\theta}\frac{dm_{[r]}(t)}{d\theta} + \left(\frac{n\theta}{1-\theta} + r\right)m_{[r]}(t) - \beta s^{[r]}\left(\frac{n\theta}{1-\theta} - s\right).$$

From above formulae for $s = 0$ we get relations for partial moments of binomial and negative binomial distribution, respectively.

4.3 Inflated generalized logarithmic series distribution

Suppose X has the probability mass function

$$P[X = x] = \begin{cases} \beta + \alpha\frac{n\Gamma(bx)[\theta(1-\theta)^{b-1}]^x}{x\Gamma(x)\Gamma(bx-x+1)[-ln(1-\theta)]}, & x = s, \\ \alpha\frac{n\Gamma(bx)[\theta(1-\theta)^{b-1}]^x}{x\Gamma(x)\Gamma(bx-x+1)[-ln(1-\theta)]}, & x \neq s, \end{cases} \tag{24}$$

where $x = 1, 2, \ldots, 0 < \theta < 1, 0 < b < \theta^{-1}$.
It can be seen that

$$a(x) = \frac{\Gamma(bx)}{x\Gamma(x)\Gamma(bx-x+1)},\ g(\theta) = \theta(1-\theta)^{b-1},\ f(\theta) = -ln(1-\theta).$$

In this case relations for lower ordinary and factorial moments we get from (10), (16) and (22)

$$\begin{aligned} m_{r+1}(t) &= \frac{\theta(1-\theta)}{1-b\theta}\frac{dm_r(t)}{d\theta} - \frac{\theta}{(1-b\theta)\ln(1-\theta)}m_r(t) \\ &\quad +\beta s^r\left(\frac{\theta}{(1-b\theta)\ln(1-\theta)} + s\right), \end{aligned}$$

$$\begin{aligned} m_{(r+1)}(t) &= \frac{\theta(1-\theta)}{1-b\theta}\frac{dm_{(r)}(t)}{d\theta} - \left(\frac{\theta}{(1-b\theta)\ln(1-\theta)} + r\right)m_{(r)}(t) \\ &\quad +\beta s^{(r)}\left(\frac{\theta}{(1-b\theta)\ln(1-\theta)} + s\right), \end{aligned}$$

$$m_{[r+1]}(t) = \frac{\theta(1-\theta)}{1-b\theta}\frac{dm_{[r]}(t)}{d\theta} - \left(\frac{\theta}{(1-b\theta)\ln(1-\theta)} - r\right) m_{[r]}(t) + \beta s^{[r]}\left(\frac{\theta}{(1-b\theta)\ln(1-\theta)} + s\right).$$

Putting in (24) $b = 1$ we have inflated Fisher's logarithmic series distribution. In this cases the lower partial ordinary and factorial moments fulfil the relations

$$m_{r+1}(t) = \theta\frac{dm_r(t)}{d\theta} - \frac{\theta}{(1-\theta)\ln(1-\theta)} m_r(t) + s^r\left(\frac{\theta}{(1-\theta)\ln(1-\theta)} + s\right),$$

$$m_{(r+1)}(t) = \theta\frac{dm_{(r)}(t)}{d\theta} - \left(\frac{\theta}{(1-\theta)\ln(1-\theta)} + r\right) m_{(r)}(t) + \beta s^{(r)}\left(\frac{\theta}{(1-\theta)\ln(1-\theta)} + s\right),$$

$$m_{[r+1]}(t) = \theta\frac{dm_{[r]}(t)}{d\theta} - \left(\frac{\theta}{(1-\theta)\ln(1-\theta)} - r\right) m_{[r]}(t) + \beta s^{[r]}\left(\frac{\theta}{(1-\theta)\ln(1-\theta)} + s\right).$$

4.4 Inflated lost games distribution

Suppose X has the probability mass function

$$P[X = x] = \begin{cases} \beta + \alpha\frac{a}{2x-a}\binom{2x-a}{x}\frac{[\theta(1-\theta)]^x}{\theta^a}, & x = s, \\ \alpha\frac{a}{2x-a}\binom{2x-a}{x}\frac{[\theta(1-\theta)]^x}{\theta^a}, & x \neq s, \end{cases}$$

for $x = a, a+1, ...;\ a \geq 1,\ 0 < \theta < \frac{1}{2}$.
Then

$$a(x) = \frac{a}{2x-a}\binom{2x-a}{x},\ f(\theta) = \theta^a,\ g(\theta) = \theta(1-\theta).$$

In this case relations for lower ordinary and factorial moments we get from (10), (16) and (22)

$$m_{r+1}(t) = \frac{\theta(1-\theta)}{1-2\theta}\frac{dm_r(t)}{d\theta} + \frac{a(1-\theta)}{1-2\theta} m_r(t) - \beta s^r\left(\frac{a(1-\theta)}{1-2\theta} - s\right),$$

$$m_{(r+1)}(t) = \frac{\theta(1-\theta)}{1-2\theta}\frac{dm_{(r)}(t)}{d\theta} + \left(\frac{a(1-\theta)}{1-2\theta} - r\right) m_{(r)}(t) - \beta s^{(r)}\left(\frac{a(1-\theta)}{1-2\theta} - s\right),$$

$$m_{[r+1]}(t) = \frac{\theta(1-\theta)}{1-2\theta}\frac{dm_{[r]}(t)}{d\theta} + \left(\frac{a(1-\theta)}{1-2\theta} - r\right) m_{[r]}(t) - \beta s^{[r]}\left(\frac{a(1-\theta)}{1-2\theta} - s\right).$$

4.5 Inflated distribution of the number of customers served in a busy period of the queue M/M/1

Let us consider the following probability function

$$P[X=x]=\begin{cases}\beta+\alpha\frac{a}{2x-a}\binom{2x-a}{x}\left[\frac{\theta}{(1+\theta)^2}\right]^x\left(\frac{1+\theta}{\theta}\right)^{-a}, & x=s,\\ \alpha\frac{a}{2x-a}\binom{2x-a}{x}\left[\frac{\theta}{(1+\theta)^2}\right]^x\left(\frac{1+\theta}{\theta}\right)^{-a}, & x\neq s,\end{cases}$$

for $x=a,a+1,...;\ a\geq 1,\ 0<\theta<\frac{1}{2}$.
For $\alpha=1$ and we obtain distribution of the number of customers served in a busy period of the queue M/M/1 considered by A. W. Kemp and C. D. Kemp (1968). The parameter θ is called the traffic intensity. In this case

$$a(x)=\frac{a}{2x-a}\binom{2x-a}{x},\ f(\theta)=\left(\frac{1+\theta}{\theta}\right)^a,\ g(\theta)=\frac{\theta}{(1+\theta)^2}.$$

In this case lower partial moments fulfil recurrence relations

$$m_{r+1}(t)=\frac{\theta(1+\theta)}{1-\theta}\frac{dm_r(t)}{d\theta}+\frac{a}{\theta-1}m_r(t)-\beta s^r\left(\frac{a}{\theta-1}-s\right),$$

$$m_{(r+1)}(t)=\frac{\theta(1+\theta)}{1-\theta}\frac{dm_{(r)}(t)}{d\theta}+\left(\frac{a}{\theta-1}-r\right)m_{(r)}(t)-\beta s^{(r)}\left(\frac{a}{\theta-1}-s\right),$$

$$m_{[r+1]}(t)=\frac{\theta(1+\theta)}{1-\theta}\frac{dm_{[r]}(t)}{d\theta}+\left(\frac{a}{\theta-1}+r\right)m_{[r]}(t)-\beta s^{[r]}\left(\frac{a}{\theta-1}-s\right).$$

5 Recurrence relations for upper partial moments of IMPSD

In this Section using the same method we derive upper partial moments. It is obvious that for $t>s$ the upper partial moments are independent of a point of inflation s and recurrence relations for these moments are the same as formulae obtained by Gupta (1974). So we consider only the situation when a point of inflation s is greater than a limit t.

Theorem 15 *The $r+1$-th upper partial moment about a point c of IMPSD is given by*

$$\mu^{r+1}(t)=\frac{g(\theta)}{g'(\theta)}\frac{d\mu^r(t)}{d\theta}+\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-c\right)\mu^r(t)-\beta(s-c)^r\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}-s\right). \tag{25}$$

Putting in (25) $c=0$ we get recurrence relations for upper partial moments of IMPSD.

Corollary 16 *The $r+1$-th upper partial ordinary moment of IMPSD is given by*

$$m^{r+1}(t) = \frac{g(\theta)}{g'(\theta)}\frac{dm^r(t)}{d\theta} + \frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)}m^r(t) - \beta s^r\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - s\right) \tag{26}$$

Now we give relations for upper partial factorial moments. As in the previous Section, we assume that $s > t$.

Theorem 17 *The $r+1$-th upper partial descending factorial moment about a point c of IMPSD is given by*

$$\begin{aligned}\mu^{(r+1)}(t) &= \frac{g(\theta)}{g'(\theta)}\frac{d\mu^{(r)}(t)}{d\theta} + \left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - c - r\right)\mu^{(r)}(t) \\ &\quad -\beta(s-c)^{(r)}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - s\right).\end{aligned} \tag{27}$$

From (27) for $c = 0$ we obtain recurrence relations for upper descending factorial partial moments of IMPSD.

Corollary 18 *The $r+1$-th upper partial descending factorial moment of IMPSD is given by*

$$\begin{aligned}m^{(r+1)}(t) &= \frac{g(\theta)}{g'(\theta)}\frac{dm^{(r)}(t)}{d\theta} + \left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - r\right)m^{(r)}(t) \\ &\quad -\beta s^{(r)}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - s\right).\end{aligned} \tag{28}$$

Theorem 19 *The $r+1$-th upper partial ascending factorial moment about c of IMPSD is given by*

$$\begin{aligned}\mu^{[r+1]}(t) &= \frac{g(\theta)}{g'(\theta)}\frac{d\mu^{[r]}(t)}{d\theta} + \left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - c + r\right)\mu^{[r]}(t) \\ &\quad -\beta(s-c)^{[r]}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - s\right).\end{aligned} \tag{29}$$

Corollary 20 *The $r+1$-th upper partial descending factorial moment of IMPSD is given by*

$$\begin{aligned}m^{[r+1]}(t) &= \frac{g(\theta)}{g'(\theta)}\frac{dm^{[r]}(t)}{d\theta} + \left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} + r\right)m^{[r]}(t) \\ &\quad -\beta s^{[r]}\left(\frac{f'(\theta)}{f(\theta)}\frac{g(\theta)}{g'(\theta)} - s\right).\end{aligned} \tag{30}$$

Remark 21 *From (25) and (28) for $t \to 0$ we can get relations for complete moments of non-zero IMPSD obtained by Murat and Szynal (1998).*

6 Conclusion

Obtained recurrence relations for partial lower and upper moments generalize and extend formulae for moments established by Gupta, Gupta and Thripati (1986). They also complement formulae given by Gupta, Gupta and Thripati (1995) and and Murat and Szynal (1998).

References

Antile J. M., Asymmetry, partial moments, and production risk, "American Journal of Agricultural Economics", 2010 vol. 92(5), pp. 1294—1309.

Arrow K. J., Karlin S., Scarf H., Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, Stanford, California, 1958.

Britney R. R., Winkler R. L., Bayesian Point Estimation Under Various Loss Functions, "Proceedings of the American Statistical Association", 1968, pp. 356-364.

Crepon B., E. Duguet E., Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity, "Journal of Econometrics", 1997 vol. 79, pp. 355-378.

DeGroot M. H., Some Problems of Optimal Stopping, "Journal of the Royal Statistical Society B", 1968 vol. 30, pp. 108-122.

Gupta R.C., Modified power series distributions and some of its applications, Sankhy$\tilde{a}$ ser. B, 1974 vol. 35, pp. 288-298.

Gupta P. L., Gupta R. C., Tripathi R. C., Incomplete moments modified power series distributions with applications, "Communications in Statistics-Theory and Methods", 1986 vol. 15(3), pp. 999-1015.

Gupta P. L., Gupta R. C., Tripathi R.C., Inflated modified power series distribution, "Communications in Statistics-Theory and Methods", 1995 vol. 24(9), 2355-2374.

Gurmu S., Semi-parametric estimation of hurdle regression models with an application to medicaid utilization, "Journal of Applied Econometrics", 1997 vol. 12, pp. 225-242.

Gurmu S., Trivedi P., Excess zeros in count models for recreational trips, "Journal of Business and Economic Statistics", 1996. vol. 14, pp. 469-477.

Hamburg M., Matlack W. F., Maximizing Insurance Buyers' Utility, "Management Science", 1968 vol. 14, pp. 294-301.

Hayes R. H., Optimal Strategies for Divestiture, "Operations Research", 1969 vol. 17, pp. 292-310.

Horowitz I., Decision Making and the Theory of the Firm, Holt, Rinehart and Winston, Inc., New York, 1970.

Lambert D., Zero-inflated Poisson regression with an application to defects in manufacturing, "Technometrics", 1992 vol. 34, pp. 1-14.

Miaou S. P., The relationship between truck accidents and geometric design of road sections. Poisson versus negative binomial regressions, "Accident Analysis and Prevention", 1994 vol. 26, pp. 471-482.

Murat M., Szynal D., Non –Zero-Inflated Modified Power Series Distributions, "Communications in Statistics-Theory and Methods", 1998 vol. 27(12), pp. 3047-3064.

Nawrocki D. N., Optimal Algorithms And Lower Partial Moment: Ex Post Results, "Applied Economics", 1991 vol. 23(3), pp. 465-470.

Pandey K. N., Generalized inflated Poisson distribution, "Journal of Scientific Research-Banaras Hindu University", 1964-65 vol. XV(2), pp. 157-162.

Ridout M., Demétrio C.G.B., Hinde J., Models for count data with many zeros, Invited paper presented at the Nineteenth International Biometric Conference, Capetown, South Africa, 1998.

Shonkwiler J., Shaw W., Hurdle count-data models in recreation demand analysis, "Journal of Agricultural and Resource Economics", 1996 vol. 21, pp. 210-219.

Stone B.K., A General Class of Three Parameter Risk Measures, "The Journal of Finance", 1973 vol. 28(3), pp. 675-685.

Welsh A., Cunningham R., Donnelly C., Lindennmayer D., Modelling the abundance of rare species-statistical-models for counts with extra zeros, "Ecological Modelling", 1996 vol. 88, pp. 297-308.

Ernest Nieznaj [1]

A note on mixed moments of random variables governed by Poisson random measure

Keywords: Poisson random measure, mixed moments, Bell numbers

Abstract

Let $(E, \mathscr{E}, \mu)$ be a measurable space with $\sigma-$ finite measure μ and π Poisson random measure with intensity measure μ. We assume that X_k and Y_k are random variables of the form

$$X_k(\omega) = \int_E f_k(x)\pi(dx;\omega), \quad Y_k(\omega) = \int_E f_k(x)\widetilde{\pi}(dx;\omega),$$

for $k = 1, \ldots, n$ and $\widetilde{\pi}(x;\omega) := \pi(x;\omega) - \mu(dx)$. If mixed moments of X_k and Y_k are finite then we provide formulae for calculating

$$\mathbb{E}(X_1X_2\ldots X_n), \quad \mathbb{E}(Y_1Y_2\ldots Y_n). \tag{1}$$

In E. Nieznaj (2011) these formulae were proved for characteristic functions, however in this note we prove (1) in a slightly different way. In particular when we take $f_k(x) = f(x)$, $k = 1, \ldots, n$ we get formulas for $\mathbb{E}X^n$ and $\mathbb{E}Y^n$ proved in B. Bassan, E. Bona (1990) and N. Privault (2012).

1 Introduction and main results

Poisson random measure

Recall that X has Poisson distribution $\mathscr{P}(\lambda)$ with intensity $\lambda > 0$, if

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

We define also distributions $\mathscr{P}(0)$, $\mathscr{P}(\infty)$ by

$$\mathbb{P}(X = 0) = 1, \quad \mathbb{P}(X = \infty) = 1.$$

Throughout this note we assume that $(E, \mathscr{E}, \mu)$ is a measurable space with $\sigma-$finite measure μ.

[1] Department of Mathematics, Technical University of Lublin, Nadbystrzycka 38, 20-618 Lublin, *e-mail: e.nieznaj@pollub.pl*

Following J. Zabczyk (2004), J.F.C. Kingman (1993) and K. Sato (1999) we recall that a Poisson random measure on $(E,\mathscr{E})$ with intensity measure μ, defined on a probability space $(\Omega,\mathscr{F},\mathbb{P})$, is a mapping

$$\pi : \Omega \times \mathscr{E} \to \mathbb{N} = \{0,1,2,\ldots\}$$

such that the following conditions hold:

(i) the random variable $\pi(A)$ has Poisson distribution with intensity $\mu(A)$, for any $A \in \mathscr{E}$,

(ii) if $A_1, A_2, \ldots, A_m$ are disjoint then $\pi(A_1),\ldots,\pi(A_m)$ are independent and

$$\pi(A_1 \cup \ldots \cup A_m) = \pi(A_1) + \ldots + \pi(A_m),$$

(iii) for every $\omega \in \Omega$, $\pi(\cdot;\omega)$ is a measure on E.

We can also think of π as a collection of random variables $\{\pi(A), A \in \mathscr{E}\}$ that satisfy conditions (i) - (iii). The construction and properties of such a measure is e.g. in J. Zabczyk (2004), J.F.C. Kingman (1993) and K. Sato (1999). We define also the compensated Poisson measure

$$\widetilde{\pi}(A) := \pi(A) - \mu(A), \quad A \in \mathscr{E},\ \mu(A) < \infty.$$

Note that $\mathbb{E}\pi(A) = \mu(A)$, $\mathbb{E}\widetilde{\pi}(A) = 0$, where $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$.

Properties of the integral with respect to π and $\widetilde{\pi}$ are contained in the following theorem, which is comming from J. Zabczyk (2004) (see pp. 64).

Theorem 1 *Let $f(x)$ be a measurable function on $(E,\mathscr{E})$. Then the following hold*

(i) if

$$\int_E |f(x)|\pi(dx) < +\infty, \quad \mathbb{P}-a.s.$$

then for $\lambda \in \mathbb{R}^1$

$$\mathbb{E}\exp\left(i\lambda \int_E f(x)\pi(dx)\right) = \exp\left(-\int_E (1 - e^{i\lambda f(x)})\mu(dx)\right) \tag{2}$$

(ii) if $f \in L^1(E)$ then

$$\mathbb{E}\int_E |f(x)|\pi(dx) < +\infty$$

and

$$\mathbb{E}\int_E f(x)\widetilde{\pi}(dx) = 0,$$

(iii) if $f \in L^1(E) \cap L^2(E)$ *then*

$$\mathbb{E}\left|\int_E f(x)\widetilde{\pi}(dx)\right|^2 = \int_E f^2(x)\mu(dx). \tag{3}$$

From the above theorem we conclude that if $f \in L^1(E) \cap L^2(E)$ then

$$\mathbb{E}\exp\left(i\lambda \int_E f(x)\widetilde{\pi}(dx)\right) \tag{4}$$

$$= \exp\left(-\int_E \left(1 - e^{i\lambda f(x)} + i\lambda f(x)\right)\mu(dx)\right), \quad \lambda \in \mathbb{R}.$$

Main results

We assume that we are given random variables

$$X_k(\omega) = \int_E f_k(x)\pi(dx;\omega), \quad k = 1,\dots n,$$

where $f_1,\dots,f_n$ are deterministic measurable functions on $(E,\mathscr{E})$ such that $X_1,\dots,X_n$ are well defined. For example Theorem 1 can be used to define these random variables. We have the following theorem, see Lemma 1 in E. Nieznaj (2011) and B. Bassan, E. Bona (1990).

Theorem 2 *Assume that the expectation of* $X_1 \dots X_n$ *is finite, then for* $n \geqslant 2$ *we have*

$$\mathbb{E}(X_1 X_2 \dots X_n)$$

$$= \sum_I \int \prod_{i\in\{I_1\}} f_i(x)\mu(dx)\dots\int \prod_{i\in\{I_p\}} f_i(x)\mu(dx)$$

where the sum runs over all partitions $I = \{\{I_1\},\dots,\{I_p\}\}$ *of the set* $\{1,2,\dots,n\}$.

We write shortly

$$\int f d\mu := \int_E f(x)\mu(dx).$$

For $n = 2$ we have, see also J.F.C. Kingman (1993),

$$\mathbb{E}(X_1X_2) = \int f_1 d\mu \int f_2 d\mu + \int f_1 f_2 d\mu$$

and $n = 3$

$$\mathbb{E}(X_1X_2X_3) = \int f_1 d\mu \int f_2 d\mu \int f_3 d\mu + \int f_1 d\mu \int f_2 f_3 d\mu$$

$$+\int f_2 d\mu \int f_1 f_3 d\mu + \int f_3 d\mu \int f_1 f_2 d\mu + \int f_1 f_2 f_3 d\mu.$$

Next we consider integration with respect to compensated Poisson measure, so let

$$Y_k(\omega) = \int_E f_k(x)\widetilde{\pi}(dx;\omega), \quad k = 1,\dots n,$$

where $f_1,\dots,f_n$ are measurable functions such that $Y_1,\dots,Y_n$ are well defined, see (3).

We can state now the next theorem, see Lemma 2 in E. Nieznaj (2011) and Proposition 3.3 in N. Privault (2012).

Theorem 3 *Assume that the expectation of $Y_1\dots Y_n$ is finite, then for $n \geqslant 2$ we have*

$$\mathbb{E}(Y_1Y_2\dots Y_n)$$
$$= \sum_I \int \prod_{i\in\{I_1\}} f_i(x)\mu(dx)\dots\int \prod_{i\in\{I_p\}} f_i(x)\mu(dx)$$

where the sum runs over all partitions $I = \{\{I_1\},\dots,\{I_p\}\}$ of the set $\{1,2,\dots,n\}$ such that $|\{I_k\}| \geqslant 2$ for all $k = 1,\dots,p$.

Again for example for $n = 2$ we have $\mathbb{E}(Y_1Y_2) = \int f_1 f_2 d\mu = \mathrm{cov}(Y_1,Y_2)$, $n = 3$

$$\mathbb{E}(Y_1Y_2Y_3) = \int f_1 f_2 f_3 d\mu,$$

and $n = 4$

$$\mathbb{E}(Y_1Y_2Y_3Y_4) = \int f_1 f_2 d\mu \int f_3 f_4 d\mu + \int f_1 f_3 d\mu \int f_2 f_4 d\mu$$
$$+\int f_1 f_4 d\mu \int f_2 f_3 d\mu + \int f_1 f_2 f_3 f_4 d\mu.$$

Remark 4 *As we have mentioned Theorems 1 and 2 were proved in E. Nieznaj (2011) for characteristic functions. So if we take $f_k(x) = \chi_{A_k}(x)$, $A_k \in \mathscr{E}$ then*

$$X_k(\omega) = \int_E \chi_{A_k}(x)\pi(dx;\omega) = \pi(A_k),$$
$$Y_k(\omega) = \pi(A_k) - \mu(A_k), \quad k = 1,\dots,n.$$

and using Theorem 2 we obtain

$$\mathbb{E}\left(\pi(A_1)\pi(A_2)\dots\pi(A_n)\right) = \sum_I \mu(\bigcap_{i\in\{I_1\}} A_i)\dots\mu(\bigcap_{i\in\{I_p\}} A_i), \tag{5}$$

where the sum runs over all partitions of $\{1,\dots,n\}$.

2 Proofs of Theorems 2, 3

Proof of Theorem 2

Substituting in (2) $\lambda = 1$ and $f(x) = t_1 f_1(x) + \dots t_n f_n(x)$ we have

$$\Phi(\mathbf{t}) = \mathbb{E}\exp(i(t_1 X_1 + \dots t_n X_n)) = \exp(g(\mathbf{t})),$$

where $\mathbf{t} = (t_1, \dots, t_n)$ and

$$g(\mathbf{t}) = -\int_E \left(1 - e^{i(t_1 f_1(x) + \dots + t_n f_n(x))}\right)\mu(dx), \quad \mathbf{t} \in \mathbb{R}^d.$$

Observe that $\Phi^{(1)}_{t_1} = \Phi g^{(1)}_{t_1}$, $\Phi^{(2)}_{t_1 t_2} = \Phi[g^{(1)}_{t_1} g^{(1)}_{t_2} + g^{(2)}_{t_1 t_2}]$ and (note that this is the place where partitions of sets appear)

$$\Phi^{(3)}_{t_1 t_2 t_3} = \Phi[g^{(1)}_{t_1} g^{(1)}_{t_2} g^{(1)}_{t_3} + g^{(1)}_{t_1} g^{(2)}_{t_2 t_3}$$

$$+ g^{(1)}_{t_2} g^{(2)}_{t_1 t_3} + g^{(1)}_{t_3} g^{(2)}_{t_1 t_2} + g^{(3)}_{t_1 t_2 t_3}].$$

Now it follows by induction that for $n \geqslant 4$ we have

$$\Phi^{(n)}_{t_1 t_2 \dots t_n} = \Phi \sum_I g^{(|I_1|)}_{I_1} g^{(|I_2|)}_{I_2} \dots g^{(|I_p|)}_{I_p}, \tag{6}$$

where the sum runs over all partitions $I = \{\{I_1\}, \dots, \{I_p\}\}$ of the set $\{t_1, t_2, \dots, t_n\}$, $|\{I_k\}|$ denotes the number of elements of $\{I_k\}$ and $g^{(|I_k|)}_{I_k}$ is derivative of $g(t_1, \dots, t_n)$ with respect to indices contained in $\{I_k\}$. Since

$$g^{(1)}_{t_l}(\mathbf{t}) = i \int_E e^{i(t_1 f_1(x) + \dots + t_n f_n(x))} f_l(x)\mu(dx),$$

we have

$$g^{(1)}_{t_l}(\mathbf{0}) = i \int_E f_l(x)\mu(dx), \quad l = 1, \dots n.$$

and in general

$$g^{(k)}_{t_{l_1} \dots t_{l_k}}(\mathbf{0}) = i^k \int_E f_{l_1}(x) \dots f_{l_k}(x)\mu(dx),$$

for any $l_1, \dots, l_k \in \{1, \dots, n\}$. Using the formula

$$i^n \mathbb{E}(X_1 \dots X_n) = \Phi^{(n)}_{t_1 t_2 \dots t_n}(\mathbf{0}), \quad n \geqslant 1,$$

we conclude the statement of the theorem. □

Proof of Theorem 3

As in the previous proof let

$$\widetilde{\Phi}(\mathbf{t}) = \mathbb{E}\exp(i(t_1Y_1 + \ldots t_nY_n)) = \exp(\widetilde{g}(\mathbf{t})),$$

where

$$\widetilde{g}(\mathbf{t}) = -\int_E (1 - e^{i(t_1f_1(x)+\ldots+t_nf_n(x))}$$
$$+t_1f_1(x) + \ldots t_nf_n(x))\mu(dx), \quad \mathbf{t} \in \mathbb{R}^d.$$

Analougusly we have

$$\widetilde{g}^{(1)}_{t_l}(\mathbf{t}) = i\int_E f_l(x)(e^{i(t_1f_1(x)+\ldots+t_nf_n(x))} - 1)\mu(dx),$$

therefore (here is the only difference between proofs of Theorem 2 and 3)

$$\widetilde{g}^{(1)}_{t_l}(\mathbf{0}) = 0, \quad l = 1, \ldots n.$$

For $k \geqslant 2$ we have

$$\widetilde{g}^{(k)}_{t_{l_1}\ldots t_{l_k}}(\mathbf{0}) = i^k \int_E f_{l_1}(x)\ldots f_{l_k}(x)\mu(dx), \tag{7}$$

for any $l_1, \ldots, l_k \in \{1, \ldots, n\}$. □

3 Application of Theorem 3

Let $f \in L^2(E)$. Recall that $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space where π is defined and $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$. Using isometry (3) we can define

$$X_f := \int_E f(x)\widetilde{\pi}(dx)$$

as an element of $L^2(\Omega)$. Namely, let $\{f_n, n \geqslant 1\}$ be a sequence of measurable functions such that $f_n \in L^1(E) \cap L^2(E)$, $n \geqslant 1$, and

$$\lim_{n\to\infty} \int_E (f(x) - f_n(x))^2 \mu(dx) = 0.$$

Thanks to (3) we have

$$\mathbb{E}(X_{f_n} - X_{f_m})^2 = \int_E |f_n(x) - f_m(x)|^2 \mu(dx),$$

hence $\{X_{f_n}, n \geqslant 1\}$ is a Cauchy sequence in $L^2(\Omega)$ and we define X_f as its limit in that space, i.e.

$$X_f := \text{l.i.m.}_{n\to\infty} X_{f_n}.$$

Additionally L^2 convergence implies

$$\mathbb{E}X_f = 0, \quad \mathbb{E}X_f^2 = \int_E f^2(x)\mu(dx).$$

Moreover (4) is also true for $f \in L^2(E)$, e.g. by estimation

$$|e^{ix} - 1 - ix| \leqslant \frac{1}{2}x^2, \quad x \in \mathbb{R}.$$

By $L^\infty(E)$ we denote the Banach space of measurable functions with the norm

$$||f||_{L^\infty(E)} = \sup \text{ess}_E |f(x)|.$$

The following result is a consequence of Theorem 3.

Theorem 5 *Let $f \in L^2(E) \cap L^\infty(E)$. Then $\mathbb{E}X_f^n < +\infty$ for $n \geqslant 1$.*

Proof. Observe that for $\alpha \geqslant 0$

$$\int_E |f(x)|^{2+\alpha}\mu(dx) \leqslant ||f||_{L^\infty(E)}^\alpha \int_E |f(x)|^2\mu(dx) < +\infty,$$

hence by (7) we have

$$|\tilde{g}_t^{(k)}(\mathbf{0})| < +\infty, \quad k \geqslant 2.$$

The above implies $\mathbb{E}X_f^n < +\infty$ for every $n \geqslant 1$.□

Remark 6 *In the proof of Theorem 5 we have used the following theorem (see I.I. Gikhman, A.V. Skorokhod 1969, Theorem 2, p. 7): if X is a random variable, $\varphi(t) = \mathbb{E}\exp(itX)$, $t \in \mathbb{R}$ and $|\varphi^{2n}(0)| < +\infty$ for some $n \geqslant 1, n \in \mathbb{N}$ then $\mathbb{E}X^{2n} < +\infty$ and*

$$i^{2n}\mathbb{E}X^{2n} = \varphi^{2n}(0).$$

4 Bell numbers and partitions of a set

Bell numbers

Recall that the Bell number B_n is defined as the number of partitions of a set consisting of n elements, see e.g. G.C. Rota (1964). We define also $B_0 = 1$. For example $B_1 = 1, B_2 = 2$ and the following relation is satisfied

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k, \quad n \geqslant 1.$$

Taking $A_k = A$, $A \in \mathscr{E}$, $k = 1, \ldots, n$ in (5) and $\mu(A) = \lambda$ we have

$$\mathbb{E}X^n = T_n(\lambda), \quad n \geqslant 1,$$

where $X \sim \mathscr{P}(\lambda)$ and

$$B_n(x) := \sum_{k=1}^{n} S(n,k)x^k, \quad x \in \mathbb{R}, \quad n \geqslant 1,$$

are called the Bell polynomials (or sometimes Touchard polynomials), see N. Privault (2011) for more details. The Stirling number $S(n,k)$ denotes the number of partitions of $\{1,\ldots,n\}$ into k non-empty subsets. So when $X \sim \mathscr{P}(1)$, then

$$\mathbb{E}X^n = \sum_{k=1}^{n} S(n,k) = B_n, \quad n \geqslant 1. \tag{8}$$

Table 1. Stirling numbers and Bell polynomials, $B_n = B_n(1)$

n,k	1	2	3	4	5	$B_n(\lambda)$	B_n
1	1	-	-	-	-	λ	1
2	1	1	-	-	-	$\lambda+\lambda^2$	2
3	1	3	1	-	-	$\lambda+3\lambda^2+\lambda^3$	5
4	1	7	6	1	-	$\lambda+7\lambda^2+6\lambda^3+\lambda^4$	15
5	1	15	25	10	1	$\lambda+15\lambda^2+25\lambda^3+10\lambda^4+\lambda^5$	52

Remark 7 *In the literature* (8) *or sometimes the expression for $n-th$ moment of* $\mathscr{P}(1)$

$$B_n = \frac{1}{e}\sum_{k=0}^{\infty}\frac{k^n}{k!}$$

is called "Dobinski's formula" (see G.C.Rota, 1964). Bell numbers can be also defined by

$$e^{e^x-1} = \sum_{n=0}^{+\infty}\frac{B_n}{n!}x^n, \quad B_0 := 1.$$

Additionally the Bell polynomials satisfy the following relation

$$B_{n+1}(x) = x\sum_{k=0}^{n}\binom{n}{k}B_k(x), \quad n \geqslant 1,$$

with $B_0(x) = 1$, $B_1(x) = x$.

Central moments of Poisson distribution

Now let's have a look at Theorem 3. We will find the recurrence formula for the number of partitions of $\{1,\ldots,n\}$ into subsets containing at least 2 elements.

This in fact will be the relation between central moments of Poisson distribution. Let $X \sim \mathscr{P}(\lambda)$. We introduce the following notation

$$E_n(\lambda) = \mathbb{E}(X-\lambda)^n, \quad n \geqslant 1, \tag{9}$$

with $E_0 = 1$. The characteristic function of $X - \lambda$ equals

$$\phi(t) = \mathbb{E}e^{it(X-\lambda)} = e^{(e^{it}-it-1)\lambda}, \quad t \in \mathbb{R}. \tag{10}$$

We write

$$\phi(t) = e^{f(t)}, \quad f(t) = (e^{it} - it - 1)\lambda.$$

Since $\phi'(t) = \phi(t)f'(t)$, therefore using the Leibniz formula we get

$$\phi^{(n+1)}(t) = \sum_{k=0}^{n} \binom{n}{k} \phi^{(n-k)}(t) f^{(k+1)}(t), \ n \geqslant 1. \tag{11}$$

Moreover we have

$$f'(0) = 0, \quad \text{and} \quad f^{(n)}(0) = i^n\lambda, \quad n \geqslant 2.$$

Finally by (11) we obtain the relation

$$E_{n+1}(\lambda) = \lambda \sum_{k=1}^{n} \binom{n}{k} E_{n-k}(\lambda), \quad n \geqslant 1, \tag{12}$$

where $E_0(\lambda) = 1$, $E_1(\lambda) = 0$. For example we have $E_2(\lambda) = \lambda,\ E_3(\lambda) = \lambda$, $E_4 = 3\lambda^2 + \lambda,\ E_5(\lambda) = 10\lambda^2 + \lambda$. Let us denote $E_n := E_n(1)$ for $n \geqslant 1$.

Table 2. Comparison of B_n and E_n

n	$B_n = \mathbb{E}X^n$	$\mathbb{E}(X-1)^n$	n	$B_n = \mathbb{E}X^n$	$\mathbb{E}(X-1)^n$
1	1	0	6	203	41
2	2	1	7	877	162
3	5	1	8	4140	715
4	15	4	9	21147	3425
5	52	11	10	115975	17722

References

Bassan B., Bona E., Moments of stochastic processes governed by Poisson random measures, "Comment. Math. Univ. Caroline", 1990 vol. 31 (2), pp. 337-343, (1990).

Gikhman I.I., Skorokhod A.V., Introduction to the Theory of Random Processes, W.B. Saunders Company, 1969.
Kingman J. F. C., Poisson Processes, Oxford, 1993.
Nieznaj E., On the superdiffusive behavior of a passive tracer in a Poisson shot-noise field, "Z. Angew. Math. Phys.", 2011 vol. 62, pp. 223-231.
Privault N., Moments of Poisson Stochastic Integrals with random integrands, "Prob. and Math. Stat.", 2012 vol 32, pp. 227-239.
Privault N., Generalized Bell polynomials and the combinatorics of Poisson central moments, The Electronic Journal of Combinatorics 18, 2011.
Rota G. C., The number of partitions of a set, "Am. Math. Month.", 1964 vol. 71(5), pp. 498 - 504.
Sato K., Levy Processes and Infinitely Divisible Distributions, Cambridge Studies in Advanced Mathematics 68, 1999.
Zabczyk J., Top*ics in Stochastic Processes*, PISA, 2004.

Dariusz Majerek, Wojciech Rosa [1]

Modeling of financial markets using structural equations

Keywords: structural equation, factor analysis, stock indices, exchange rates

Abstract

The work is a kind of example of using of structural equations in economics. The article uses both confirmatory and exploratory variants of structural models. The use of this tool was dictated by the need to discover or confirm the hypothetical dependencies on the financial market. Considering the individual markets and exchange rates, we can using structural modeling indicate the strength of the interaction between those factors. Some of the models in the work does not have a counterpart in the classical statistical methods, such as regression analysis, due to the strong correlation between the explanatory variables. They were also shown structural models describing the relationship in terms of time. Database were exchange rates and stock indices traded on markets around the world in the period from 01.01.1999 to 20.10.2011. All models are characterized by a strong convergence and seem to explain actually exists dependencies on the financial markets.

1 Introduction

At first we give some arguments justifying the use of structural equation in the analysis of financial markets.

Nowadays analyzing exchange of currency, stocks, stock indices and commodities traded on markets often ask ourselves whether a particular instrument goes up or down, and possibly by how much. To answer this question analysts use mainly two methods separately or complementarily, namely fundamental analysis or technical analysis.

Supporters of the latter are looking for systems based on indicators, oscillators, candlesticks formations and harmonic patterns or price action.

In all these cases, the systems are upgraded by the analysis of the relationships between pairs of instruments strongly interacting. Since the all values measured on the stock markets of all the world are closely related (for example WIG20 is strongly correlated with S&P500) to finding these relationships really can help

[1] Department of Mathematics, Technical University of Lublin, Nadbystrzycka 38, 20-618 Lublin, *e-mail: d.majerek@pollub.pl, roswoj@gmail.com*

to take the appropriate position in the market. If we analyze the relationship between the two instruments, a simple correlation should suffice. But when we examine the flows between several instruments we need to apply more advanced tool, namely the structural equation.

The second argument in favor of using this tool is that the structural equation can be used both to discover and confirm links occurred.

2 Structural equation model

The first to be interested in the analysis of paths[2] was Sewell Wright in the twenties of the last century. However, at the beginning of the present century, this method began to be widely used due to increased computing power, which is necessary to calculate the parameters of the model.

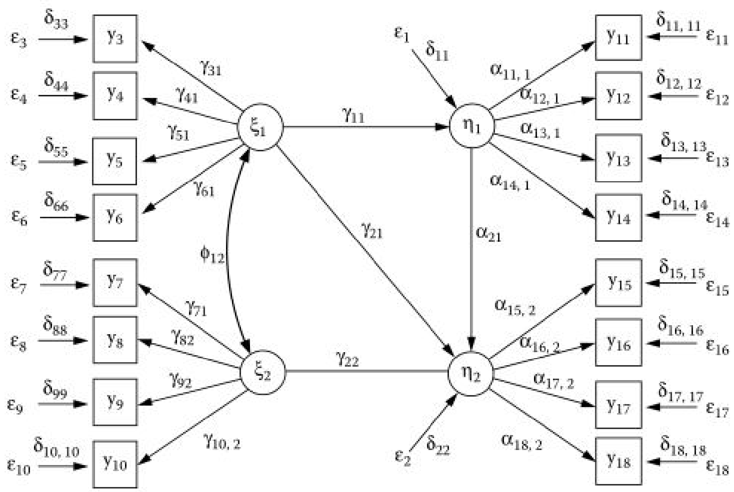

Figure 1: **A structural equation model represented by a path diagram. Squares are observable variables. Circles are latent variables. Disturbances, although latent variables, are represented without circles. One-headed arrows represent causal paths. Two-headed curves represent covariance. Associated with each causal path is a structural parameter.**

In the 1930s, the economist John Maynard Keynes (1936) developed models of the economy using systems of simultaneous linear equations relating one set of variables to another set of variables.

Structural Equation Modeling is a very powerful and widely used method of multivariate techniques. It includes the measurement and structural model. It

[2]Name comes from the graphical diagram describing the studied phenomenon.

provides easy interpretation, the ability to predict, discover hidden factors, the confirmation of hypothetical models (confirmatory analysis). The essence of this method is to compare the observed covariance matrix $\mathscr{S}$ of the implied covariance matrix $\Sigma(\theta)$[3]

$$\Sigma(\theta) = \mathscr{S}, \tag{1}$$

where θ is the parameter vector. Construction of an appropriate model of SEM[4] is to restore the covariance matrix based on the given model. SEM allows testing of many types of theoretical models. They can be interpreted as a combination of two statistical methods: factor analysis and multiple regression.

Factor analysis can be divided into: Exploratory Factor Analysis[5] and Confirmatory Factor Analysis[6]. The aim of the first is to find the number of latent variables and their relation with observable variables. The second method allows you to test the hypothetical relationships between observable variables and latent variables.

Speaking of latent variables we mean the phenomena observed, but not making a quantitative assessment, such as *satisfaction with a flight*. Mentioned variable can be measured only by the component variables that the *satisfaction of a flight* affected, such as cost and comfort of the flight (whether there is turbulence, and served meals, etc). In the case of the financial markets the latent variables can be for example the condition of the U.S. economy, Japanese economy, etc., which are difficult to evaluate. The role of observable[7] variables may play USDJPY, EURJPY, EURUSD or indices S&P500, NIKKEI, DJIA itp.

2.1 General form of the structural model

The general structural model can be represented as

$$\eta = B\eta + \Gamma\xi + \zeta \tag{2}$$

where η is the vector of endogenous latent factors, ξ is the vector of exogenous latent factors, ζ is the vector of residuals, B is the matrix of coefficients of the structural η given η and Γ is the matrix of structural coefficients η given ξ.

The structural model consists also equations describing the relationship between latent and observable variables:

$$y = \Delta_y \eta + \varepsilon \tag{3}$$

$$x = \Delta_x \xi + \delta, \tag{4}$$

[3]Calculated from the model.
[4]Structural Equation Model
[5]EFA
[6]CFA
[7]or manifest variables

where y is a vector of observable endogenous variables, x is a vector of observable exogenous variables, Δ_y and Δ_x are matrices of factor loadings and ε, δ are vectors of measurement errors, respectively of y and x.

The correct form of the model may result from the current knowledge of the studied phenomenon, then we use a model of the CFA. In case you do not know the potential cause-and-effect relationships, you can use the EFA to make such findings.

2.2 Estimating of model parameters

Estimating the parameters of the model is to minimize the of discrepancy function. To estimate model parameters we can use several methods:

- Ordinary Least Squares Method (OLS) - most commonly used to calculate the initial model parameters;
- Maximum Likelihood Method (ML) - the most common method of estimating the model parameters;
- Unweighted Least Squares Method (ULS) - can be applied even if the matrices S and Σ are not positively defined;
- Generalized Least Squares Method (GLS) - allows minor deviation from a multivariate normal distribution of endogenous and exogenous variables;
- Weighted Least Squares Method (WLS) - method more resistant to deviate from a multivariate normal distribution of input and output variables than the GLS method.

2.3 Model Evaluation

Tests or indicators to evaluate the fit of the model can be divided into two groups (see D. Hooper, J. Coughlan and M.R. Mullen, 2008):

- Absolute fit indices: χ^2, RMSEA[8], GFI[9], RMR[10], SRMR[11]
- Relative fit indices: NFI[12], CFI[13], TLI[14].

[8] Root Mean Square Error of Approximation
[9] Goodnes of Fit Index
[10] Root Mean Square Residual Index
[11] Standardized Root Mean Square Residual Index
[12] Normed Fit Index
[13] Comparative Fit Index
[14] Tucker-Lewis Index

In the case of if the above indicators signal the lack of fit of the model to the data, make another model specification. This operation can be stopped if we have a satisfactory level of fit.

3 Modeling of financial markets using structural equations

3.1 The data

In the analyzed data sheet has been placed exchange rates from the European Central Bank (ECB) from 01.01.1999 to 20.10.2011, and indexes of major stock markets in the world (FT-SE 100, DAX, S&P 500, CAC 40, DJIA, Nikkei 225, NASDAQ Composite). Also included variable NEXT(USD) represents the rate EURUSD lagging by one day in compared to other quotes. It will be used to build a predictive model. In the same way has been added the variable NEXT(GBP). Missing data were deleted in cases for the model. In addition, data were divided into five sub-periods - each with 571 measurements (the total was 2855 cases). Thus was fulfilled the assumption of sample size for simple models based on a minimum sample size of $N = 500$. The division into sub-samples of the time due to a hypothetical assumption that over the nearly 13 years the force of impacts concerned markets is likely to be changed (which was confirmed by later studies). This is due to various external factors, such as terrorist attacks, stock market crashes or ecological disasters. As a result of such events, decreases or increases the force of impact of the local currency and stock exchanges and other global currency exchange. According to one of the main assumptions of SEM, any observable variables just before including to the model has been standardized.

The estimated parameters of the following models, are given with an accuracy of two decimal places, as in the case of the correlation matrix. Statistics of the fit of the model (ie, χ^2, RMSEA and CFI) were given to three decimal places.

3.2 Confirmatory analysis

Starting the search based on the position of the researcher, who has no idea about the hidden structures (and those that believed not confirmed by confirmatory analysis), we performed an exploratory factor analysis (T. Asparouhov and B. Muthen, 2009). However, the use of EFA brings some trouble described in the following example: researchers found three hidden factors in the data structure. After the EFA is able to determine the effect of each observable variable on the latent variable. It would, however, remove from the model the relationships that are relatively very small (actually not significant). You may then find that the model is no longer fit to the sample covariance matrix. How to find the hidden depending on the desired matrix?

In our research, the answer to this problem, we used the script created in an **R** (see J. Fox, 2006), which examines all possible combinations of variables by the set structure. The program is looking for the optimal variables for confirmatory analysis model contains four indicators reported x_1, x_2, x_3, x_4 and two factors hidden *factor1*, *factor2*. The best variables are selected on the basis of critical significance test p χ^2. Diagrams for structural models were made in the program LISREL (see K.G. Joreskog and M. Thillo, 1970).

For example, if you are looking for the variables in the model shown in Figure 2, we create four-dimensional array that stores the value of p test χ^2 for each considered model (of course, the entire array will not be saved). At the end of, we sort the results descending relative to p, so we can only deal with models similar to the observed situation and the researcher should only do substantive assessment of received dependencies.

With this algorithm, we can examine all kinds of hidden structures of two factors, each with two indicator variables, measure the time of analysis, do the analysis for each of the sub-periods, consider only significant models, exclude Heywood's cases (see S. Kolenikov and K.A. Bollen, 2007). This script deals with all types of variable data seeking confirmation of the following structure of measurement model (Figure 2).

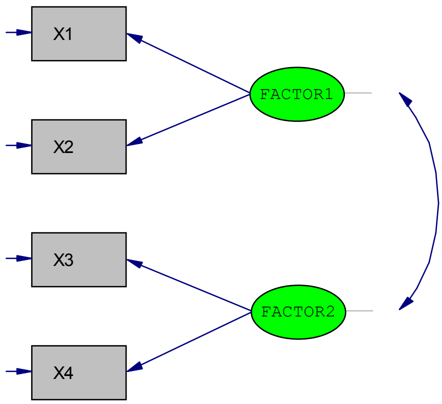

Figure 2: **Measurement model**

Table 1. Correlation matrix of first confirmatory model

	GBP	FTSE100	CHF	DJIA
GBP	1,00	0,27	0,18	0,28
FTSE100	0,27	1,00	0,64	1,00
CHF	0,18	0,64	1,00	0,63
DJIA	0,28	1,00	0,63	1,00

Source: Own elaboration based on ECB data

3.3 The first confirmatory model

Using the mentioned script has been estimated the following model (from 26.02.2004 to 6.09.2006):

$$\begin{cases} \text{FACTOR1} = 1,00 \cdot \text{GBP} + 3,62 \cdot \text{FT-SE100}, \\ \text{FACTOR2} = 1,00 \cdot \text{CHF} + 1,55 \cdot \text{DJIA}, \\ \text{Cov(FACTOR1,FACTOR2)} = 0,18. \end{cases}$$

$$\chi^2 = 0,000; df = 1; p = 0,995; \text{RMSEA} = 0,000; \text{CFI} = 1,000.$$

The first factor consists two variables: GBP and FT-SE100, and in the second factor there are hidden: CHF and DJIA. The model is a very good fit to the observed sample, as evidenced by the low value of χ^2 statistic, the high value of p, low value of RMSEA and high value of CFI. For this reason, and because the scale of measurement of the latent factors was determined by GBP and CHF, correlation hidden factors is approximately equal to the correlation GBP and CHF (0.18). Increased loadings of FT-SE100 and DJIA (respectively 3.62 and 1.55), shows that the indices are highly correlated, but also that they are strongly correlated with the GBP and CHF, than the latter between each other.

3.4 The second confirmatory model - volatility impacts in time

The postulated interactions volatility in exchange rates and stock market indices has been confirmed by the following two models. Both concern the relationship between the JPY and NIKKEI (indicators for the first latent variable) and DJIA and DAX (indicators for the second latent variable). For the period from 26.02.2004 to 6.09.2006 of the model has been estimated as follows:

$$\begin{cases} \text{FACTOR1} = 1,00 \cdot \text{JPY} + 1,24 \cdot \text{NIKKEI}, \\ \text{FACTOR2} = 1,00 \cdot \text{DJIA} + 0,97 \cdot \text{DAX}, \\ \text{Cov(FACTOR1,FACTOR2)} = 0,77. \end{cases}$$

Table 2. Correlation matrix of the second confirmatory model (based on data from 26.02.2004 to 6.09.2006)

	JPY	NIKKEI	DJIA	DAX
JPY	1,00	0,72	0,77	0,73
NIKKEI	0,72	1,00	0,95	0,93
DJIA	0,77	0,95	1,00	0,88
DAX	0,73	0,93	0,88	1,00

Source: Own elaboration based on ECB data

$$\chi^2 = 1,885; df = 1; p = 0,170; \mathrm{RMSEA} = 0,039; \mathrm{CFI} = 0,999.$$

The same model for the period from 31.03.2009 to 18.10.2011 has been estimated, however, with other parameters:

$$\begin{cases} \mathrm{FACTOR1} = 1,00 \cdot \mathrm{JPY} - 1,59 \cdot \mathrm{NIKKEI}, \\ \mathrm{FACTOR2} = 1,00 \cdot \mathrm{DJIA} + 1,06 \cdot \mathrm{DAX}, \\ \mathrm{Cov(FACTOR1,FACTOR2)} = -0,59. \end{cases}$$

$$\chi^2 = 1,338; df = 1; p = 0,247; \mathrm{RMSEA} = 0,024; \mathrm{CFI} = 1,000.$$

Correlation matrix explains the situation. In these sub-periods changed correla-

Table 3. Correlation matrix of the second confirmatory model (based on data from 31.03.2009 to 18.10.2011)

	JPY	NIKKEI	DJIA	DAX
JPY	1,00	-0,60	-0,58	-0,64
NIKKEI	-0,60	1,00	0,94	0,99
DJIA	-0,58	0,94	1,00	0,92
DAX	-0,64	0,99	0,92	1,00

Source: Own elaboration based on ECB data

tion JPY with the other variables - from an average of 0.74 to an average of -0.61. While the other variables remain strongly dependent JPY began to play the opposite role than before in this structure. In both models DAX has a similar coefficient to DJIA - close to one. This means that, despite the passage of time and evolving the relationship in markets these variables remain in the same relationship. Otherwise: changes in global markets, which concerned the observed variables in this model, probably had a similar effect on the DAX and DJIA.

Similar properties (ie variability of estimated parameters) showed other examined models, but most estimates in other sub-periods (than the current sub-period

Table 4. Correlation matrix of the first structural model

	JPY	CAD	NEXT(USD)	NEXT(GBP)
JPY	1,00	0,85	0,51	0,52
CAD	0,85	1,00	0,60	0,57
NEXT(USD)	0,51	0,60	1,00	0,70
NEXT(GBP)	0,52	0,57	0,70	1,00

Source: Own elaboration based on ECB data

for which the model was estimated) were divergent. This approach encourages the consideration of structural time series.

3.5 The first structural model

On the basis of previously performed factor analysis, the researcher can verify structural model (Figure 3). Similarly as for confirmatory analysis, the following

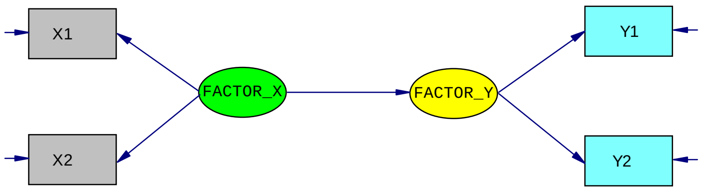

Figure 3: Example of a structural model

model has been found. It describes the relationship in foreign exchange markets, between JPY, CAD and lagged GBP and USD. Model covers the period from 2009-03-31 to 2011-10-18:

$$\begin{cases} \text{FACTOR-X} = 1,00 \cdot \text{JPY} + 1,13 \cdot \text{CAD}, \\ \text{FACTOR-Y} = 1,00 \cdot \text{NEXT(USD)} + 1,00 \cdot \text{NEXT(GBP)}, \\ \text{Cov(FACTOR-X,FACTOR-Y)} = 0,34. \end{cases}$$

$$\chi^2 = 1,338; df = 1; p = 0,091; \text{RMSEA} = 0,057; \text{CFI} = 0,998.$$

In this model, all variables enter with similar weights to the latent variables. RMSEA value is slightly above the threshold value of 0.05, however, other indicators suggest a good fit of the model. In this structure, the relationship between FACTOR-X and FACTOR-Y is weaker (0.34), than would be the result of the correlation matrix. CAD enters to a latent variable with a larger load than JPY, so it is more correlated with the NEXT(USD) and NEXT(GBP). We interpret this as a slightly greater impact of CAD on the tomorrow's courses of USD and GBP than JPY. An image of this situation we can find in the correlation matrix. An important advantage of using SEM to evaluate the impact of two variables on two other variables (in this case, the influence of two currencies rates on other currencies lagged by one day), as opposed to analysis of pairs of variables, such as using normal correlation matrix.

3.6 The second structural model

A more sophisticated structural regression model is shown in Figure 4. For the analyzed data samples failed to find the following set of variables for which the proposed model is significant - model covers the period from 26.02.2004 to 6.09.2006:

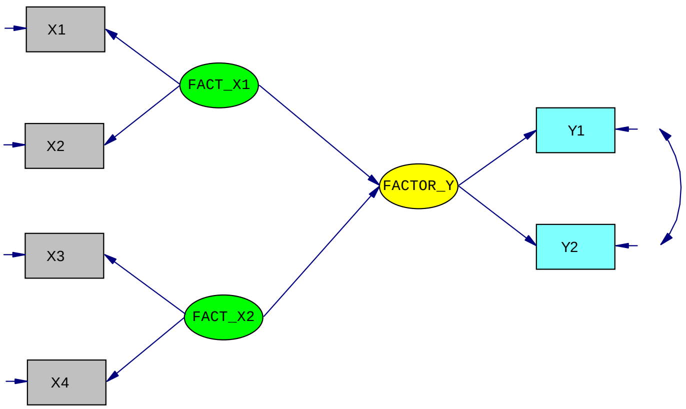

Figure 4: Example of structural model with two exogenous latent variables

$$\begin{cases} \text{FACTOR-X1} = 1{,}00 \cdot \text{CHF} + 1{,}45 \cdot \text{CAC40}, \\ \text{FACTOR-X2} = 1{,}00 \cdot \text{DAX} + 1{,}09 \cdot \text{NASDAQ}, \\ \text{FACTOR-Y} = 1{,}00 \cdot \text{NEXT(JPY)} + 0{,}50 \cdot \text{NEXT(GBP)}, \\ 0{,}94 \cdot \text{FACTOR-X1} + 0{,}21 \cdot \text{FACTOR-X2} = \text{FACTOR-Y}, \\ \text{Cov(FACTOR-X1,FACTOR-X2)} = 0{,}60, \\ \text{ErrCov(NEXT(JPY),NEXT(GBP))} = 0{,}01. \end{cases}$$

$$\chi^2 = 6{,}336; df = 5; p = 0{,}275; \text{RMSEA} = 0{,}022; \text{CFI} = 0{,}999.$$

In this model, indicator variables for the two exogenous latent variables are respectively CHF and CAC40 and DAX and Nasdaq. Endogenous hidden variable (which can be interpreted as the expected situation in the currency markets)

Table 5.Correlation matrix of the second structural model

	CHF	CAC40	DAX	NASDAQ	NEXT(JPY)	NEXT(GBP)
CHF	1,00	0,63	0,61	0,66	0,55	0,19
CAC40	0,63	1,00	0,87	0,95	0,77	0,36
DAX	0,61	0,87	1,00	0,87	0,73	0,33
NASDAQ	0,66	0,95	0,87	1,00	0,81	0,38
NEXT(JPY)	0,55	0,77	0,73	0,81	1,00	0,38
NEXT(GBP)	0,19	0,36	0,33	0,38	0,38	1,00

Source: Own elaboration based on ECB data

is indicated by two (weakly correlated) variables: NEXT(JPY) and NEXT(GBP). The loadings of exogenous variables are aligned (close to 1), but the impact of latent variable on the other is much weaker than the first. We have a lower coefficient of Nasdaq relative to the value of the coefficient of the CAC40. Nasdaq is strongly correlated than CAC40 with each of the endogenous variables. However, its weight in the model is much lower by lower loading of latent variable (FACTOR-X2), and the less impact of FACTOR-X2 to FACTOR-Y. Such conclusion is possible only on the basis of the correlation matrix. The model indicates a very low (0.01) correlation error of NEXT(GBP) and NEXT(JPY). This means that if the values of the endogenous variables deviate, the changes were not related. A higher correlation between these variables can be assumed that the model lacks a variable or variables necessary to explain the variability of endogenous factor.

3.7 The third structural model

This model is based on multiple regression. Figure 5 shows the estimated parameters of the model with two latent variables. If the investigator determines the measuring scale latent variables, by determining the individual impact indicator is chosen separately for each latent variable (for example, JPY and NIKKEI), such a model has zero degrees of freedom and it is not possible to estimate the fit. However, in this way, researchers can get initial estimates of parameters and setting one parameter (for example, GBP), fit indices are then possible to calculate (one degree of freedom). Measures of fit:

```
Goodness of Fit Statistics
Degrees of Freedom = 1
Minimum Fit Function Chi-Square = 0.0078 (P = 0.93)
Normal Theory Weighted Least Squares Chi-Square = 0.0078 (P = 0.93)
Chi-Square Difference with 0 Degree of Freedom = 0.00 (P = 1.00)
Estimated Non-centrality Parameter (NCP) = 0.0
90 Percent Confidence Interval for NCP = (0.0 ; 0.69)
Minimum Fit Function Value = 0.00
Population Discrepancy Function Value (F0) = 0.0
90 Percent Confidence Interval for F0 = (0.0 ; 0.00024)
Root Mean Square Error of Approximation (RMSEA) = 0.0
```

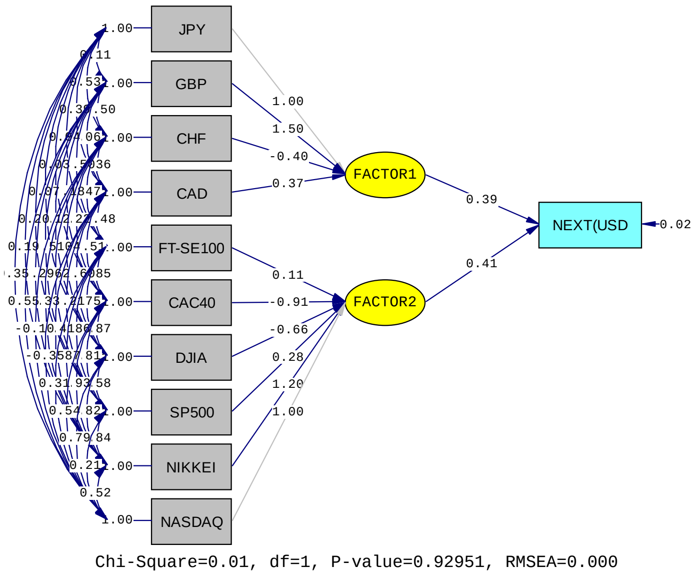

Figure 5: Diagram of the third structural model

```
90 Percent Confidence Interval for RMSEA = (0.0 ; 0.016)
P-Value for Test of Close Fit (RMSEA < 0.05) = 1.00

Expected Cross-Validation Index (ECVI) = 0.065
90 Percent Confidence Interval for ECVI = (0.065 ; 0.065)
ECVI for Saturated Model = 0.046
ECVI for Independence Model = 13.77

Chi-Square for Independence Model with 55 Degrees of Freedom = 39296.33
Independence AIC = 39318.33
Model AIC = 240.01
Saturated AIC = 132.00
Independence CAIC = 39394.86
Model CAIC = 1074.91
Saturated CAIC = 591.20

Normed Fit Index (NFI) = 1.00
Non-Normed Fit Index (NNFI) = 1.00
Parsimony Normed Fit Index (PNFI) = 0.018
Comparative Fit Index (CFI) = 1.00
Incremental Fit Index (IFI) = 1.00
Relative Fit Index (RFI) = 1.00

Critical N (CN) = 2421602.39

Root Mean Square Residual (RMR) = 0.00
Standardized RMR = 0.00
Goodness of Fit Index (GFI) = 1.00
Adjusted Goodness of Fit Index (AGFI) = 1.00
Parsimony Goodness of Fit Index (PGFI) = 0.015

Time used: 0.016 Seconds
```

Correlation matrix (for standardized variables):

Covariance Matrix

	NEXT(USD	JPY	GBP	CHF	CAD	FT-SE100
NEXT(USD	1.00					
JPY	0.66	1.00				
GBP	0.75	0.11	1.00			
CHF	-0.07	0.53	-0.50	1.00		
CAD	0.24	0.39	0.06	0.36	1.00	

FT-SE100	-0.25	0.04	-0.50	0.47	-0.48	1.00
CAC40	0.00	0.03	-0.18	0.22	-0.51	0.85
DJIA	0.23	0.07	0.12	0.04	-0.60	0.75
SP500	-0.16	0.20	-0.51	0.62	-0.21	0.86
NIKKEI	0.06	0.19	-0.29	0.33	-0.41	0.87
NASDAQ	0.71	0.35	0.55	-0.10	-0.35	0.31

Covariance Matrix

	CAC40	DJIA	SP500	NIKKEI	NASDAQ
CAC40	1.00				
DJIA	0.87	1.00			
SP500	0.81	0.58	1.00		
NIKKEI	0.93	0.82	0.84	1.00	
NASDAQ	0.54	0.79	0.21	0.52	1.00

As you can see both the factor containing the exchange rates as well as factor including stock indices have a similar impact on the NEXT(USD). Undoubted advantages of this model are: lack of a variable USD as exogenous variable, no splitting into sub-periods, the model covers the whole aspect of time, grouping the exogenous variables into currencies and stocks. This model explain 98 % endogenous variable.

4 Summary

As you can see describing financial dependencies by structural equation modeling has many advantages. Combining two methods: linear regression and factor analysis you can obtain big structures describing complicated relationships between financial instruments.

Above models shows us that number of shares quoted on financial markets is characterized by strong dependency. The proposed models are the starting point for discussions about the use of structural equation in the analysis of financial markets. It can be modified by adding additional exogenous variables but given that the relationship between further exchange rates are strong, adding more does not bring much to the model. However, the addition of the endogenous variables can cause unidentification of the model.

Based on the above considerations it can be concluded that the application of structural modeling to explore and confirm the observed dependencies on the financial markets is justified.

References

Konarski R., *Modele równań strukturalnych - Teoria i praktyka*, Warszawa, Wydawnictwo Naukowe PWN, 2009.
Asparouhov T., Muthén B., *Exploratory structural equation modeling*, "Structural Equation Modeling", 2009 vol. 16, pp. 397-438.
Kolenikov S., Bollen K. A., Testing Negative Error Variances: Is a Heywood Case a Symptom of Misspecification University of Missouri & University of North Carolina, 2007.

Fox J., *Structural Equation Modeling With the sem Package in R,* "Structural Equation Modeling", 2006 vol. 13(3), pp. 465-486.

Hooper D., Coughlan J., Mullen M. R., *Structural Equation Modelling: Guidelines for Determining Model Fit*, "The Electronic Journal of Business Research Methods", 2008 vol. 6 (1), pp. 53-60.

Jöreskog K. G., Thillo M., *LISREL: A General Computer Program for Estimating a Linear Structural Equation System Involving Multiple Indicators of Unmeasured Variables*, Princeton, NJ: Educational Testing Service, 1970.

A Data sheet

Table 1: Sample data from the spreadsheet data from the European Central Bank, including data from 1.01.1999 to 20.10.2011.

Date	EURO	USD	FT-SE10(	CAC40	JPY	CHF	SP500	NIKKEI	DJIA	DAX	NASDAQ
1999-01-01	4.0925	3.5025			2.5531	3.0929					
1999-01-04	4.0670	3.4500	4147.5	5879.4	2.5156	3.0404	13415.0	1228.1	5252.4	9184.3	9809.2
1999-01-05	4.0240	3.4130	4200.8	5958.2	2.4948	3.0581	13232.0	1244.8	5253.9	9311.2	9891.1
1999-01-06	4.0100	3.4140	4294.8	6148.8	2.4871	3.0493	13468.0	1272.3	5443.6	9545.0	10233.8
1999-01-07	4.0195	3.4550	4230.7	6101.2	2.4862	3.1087	13536.0	1269.7	5323.2	9537.8	10693.6
1999-01-08	4.0380	3.4640	4245.4	6147.2	2.5021	3.1026	13391.0	1275.1	5392.8	9643.3	10722.7
1999-01-11	4.0370	3.4880	4201.9	6085.0	2.5058	3.1929	13368.0	1263.9	5270.6	9619.9	10634.3
1999-01-12	4.0340	3.5040	4100.7	6033.6	2.5097	3.1123	13361.0	1239.5	5200.1	9474.7	10711.6
1999-01-13	4.2100	3.5860	3958.7	5850.1	2.6442	3.1979	13403.0	1234.4	4931.8	9349.6	10273.8
1999-01-14	4.1300	3.5440	3997.1	5820.2	2.5914	3.1112	13738.0	1212.2	4912.8	9120.9	10183.1
1999-01-15	4.1500	3.5700	4054.8	5941.0	2.5947	3.1503		1243.3	4960.2	9340.6	10147.4
1999-01-18	4.0700	3.5050	4151.7	6123.9	2.5420	3.0619	13805.0		5050.4		10402.5
1999-01-19	4.1030	3.5330	4116.0	6027.6	2.5671	3.1066	13770.0	1252.0	5073.2	9355.2	10290.1
1999-01-20	4.0840	3.5280	4190.0	6105.6	2.5451	3.1022	14028.0	1256.6	5143.1	9335.9	10314.9
1999-01-21	4.1310	3.5710	4154.0	6022.3	2.5773	3.1693	14245.0	1235.2	5156.7	9264.1	10048.6
1999-01-22	4.1580	3.5940	4019.3	5861.2	2.6013	3.1519	14154.0	1225.2	5019.3	9120.7	9738.5
1999-01-25	4.1600	3.5890	4050.8	5880.9	2.6051	3.1456	14208.0	1234.0	4982.5	9203.3	9499.5
1999-01-26	4.1790	3.6070	4071.3	5885.7	2.6090	3.1775	14382.0	1252.3	4986.8	9324.6	9509.8
1999-01-27	4.1605	3.6090	4098.1	5876.4	2.5870	3.1537	14450.0	1243.2	5061.2	9200.2	9719.7
1999-01-28	4.1395	3.6285	4199.7	5872.5	2.5647	3.1291	14342.0	1265.4	5096.4	9281.3	9361.0
1999-01-29	4.1650	3.6590	4251.8	5896.0	2.5839	3.1532	14499.0	1279.6	5160.0	9358.8	9506.9
1999-02-01	4.2350	3.7350	4303.9	6012.4	2.6319	3.2383	14465.0	1273.0	5190.8	9345.7	9599.6
1999-02-02	4.2350	3.7350	4243.6	6013.0	2.6463	3.3134	14349.0	1262.0	5166.9	9274.1	9502.7
1999-02-03	4.2020	3.7050	4188.8	5940.3	2.6295	3.3098	14161.0	1272.1	5085.7	9366.8	9419.9
1999-02-04	4.1520	3.6860	4167.4	5939.9	2.5920	3.2528	14086.0	1248.5	5077.9	9304.5	9438.7
1999-02-05	4.1730	3.6970	4147.3	5855.3	2.6102	3.2579	13898.0	1239.4	5080.8	9304.2	9190.2
1999-02-08	4.1930	3.7290	4154.0	5834.9	2.6184	3.2824	13992.0	1243.8	5027.2	9291.1	9139.6
1999-02-09	4.2530	3.7530	4038.5	5779.9	2.6579	3.2717	13902.0	1216.1	4904.4	9133.0	9244.5
1999-02-10	4.2430	3.7405	4001.9	5770.2	2.6590	3.2443	13952.0	1223.6	4796.8	9177.3	9076.3
1999-02-11	4.2515	3.7570	4072.3	5888.5	2.6629	3.2873		1254.0	4839.3	9363.5	9146.8
1999-02-12	4.2310	3.7630	4060.4	5950.7	2.6460	3.2793	13973.0	1230.1	4888.7	9274.9	9425.4
1999-02-15	4.2620	3.7930	4065.2	6023.2	2.6714	3.2968	14054.0		4879.6		9402.4
1999-02-16	4.2755	3.8240	4052.3	6108.6	2.6763	3.2413	14232.0	1241.9	4904.7	9297.0	
1999-02-17	4.2780	3.8020	3985.5	6078.4	2.6786	3.2074	14158.0	1224.0	4810.1	9195.5	
1999-02-18	4.2800	3.8110	4039.4	6074.9	2.6793	3.1813	14146.0	1237.3	4845.1	9298.6	
1999-02-19	4.2610	3.8180	4130.5	6031.2	2.6650	3.1822	14098.0	1239.2	4802.4	9340.0	9254.1
1999-02-22	4.2610	3.8750	4165.9	6069.9	2.6663	3.1901	14256.0	1272.1	4845.2	9552.7	9229.3
1999-02-23	4.2410	3.8655	4208.0	6155.2	2.6584	3.1905	14500.0	1271.2	4987.6	9544.4	9434.0

From the Preface:

In view of the number of books which already exist on the subject, one can ask, with a reason, whether another book about, probability in action' is really needed. In particular, what is our excuse for writing this one? Our answer is this. We want to bring the results of the recent research to the reader without delay. We will utilize mechanism that allows a rapid publication in style of arXiv.org of Cornell University. Last but not least, this book has to substantiate our activity in the last two years. These are the goals of this book.

...we chose the applications of probability methods as Ariadna's thread making the papers selection for this volume (...). A natural determination to understand the world around which sometimes takes a form of scientific curiosity is a motivation to pose intriguing questions, state problems and hypothesis. If the inspiration comes from the reality outside world of mathematics and the results can be verified in practice, then such work fits into our cathegory of broad applications.

Referee's report:

I strongly recommend this volume to publication.

Prof. dr hab. J. Kozicki
Institute of Mathematics, UMCS

ISBN 978-83-7947-036-5
9 788379 470365