DIDACTICS
OF
MATHEMATICS

9(13)

The Publishing House
of Wrocław University of Economics
Wrocław 2012

Reviewers cooperating with the journal
Giovanna Carcano, Salvatore Federico, Marian Matłoka,
Włodzimierz Odyniec, Anatol Pilawski, Achille Vernizzi, Henryk Zawadzki

Copy-editing
Elżbieta Macauley, Tim Macauley, Marcin Orszulak

Proof reading
Barbara Łopusiewicz

Typesetting
Elżbieta Szlachcic

Cover design
Robert Mazurczyk

Front cover painting: W. Tank, Sower
(private collection)

This publication is available at: www. journal.ue.wroc.pl and www.ibuk.pl

Abstracts of published papers are available in the international database
The Central European Journal of Social Sciences and Humanities
http://cejsh.icm.edu.pl

Information of submitting and reviewing paper is available
on the Publishing House's website www.wydawnictwo.ue.wroc.pl

PL ISSN 1733-7941

The original version: printed

Printing: Printing House TOTEM
Print run: 200 copies

Table of contents

D I D A C T I C S O F M A T H E M A T I C S

No. 9 (13) 2012

SOME REASONS WHY WE SHOULD TEACH MATRICES TO STUDENTS OF ECONOMICS

Wojciech Rybicki

Abstract. The paper makes up the second part of the series of articles aimed at establishing the usefulness of matrices in the study of contemporary economic sciences. The series was initiated by the present author in his previous article of this subject (Rybicki, 2010). The items chosen to be presented here concern applications of (operating with) matrices in the field of welfare economics and to the description dynamics of economic systems. The first class of matrices we discuss serves as tools for indicating the inequalities of distributions of (finite) commodity bundles (and as devices to "equalize" these distributions). Other considered families of matrices consist of transition matrices of Markov chains. The presented statements are of an elementary character – they are intended to help students feel (and believe in) some uniformity of the content of lectures on mathematics and economics and (in a wider sense) operations research.

Keywords: commodity bundle, double stochastic matrix, equalization of allocations and distributions, Markov chain, transition matrix, Markov-Chapman-Kołmogorov equations.

1. Introduction

In this article, the considerations on the role of matrices for the education of students of economics are continued. The author began to discuss some observations and impressions concerning this problem in his previous paper (Rybicki, 2010), where the main ideas were laid out and a number of examples were outlined. There we argued that students of economics, as well as the readers of scientific papers qualifying for these (and related) fields, unavoidably encounter the notion of matrix and matrix notations on many occasions during the period of their study (the same concerns also

Wojciech Rybicki
Department of Mathematics and Cybernetics, Wrocław University of Economics, Komandorska Street 118/120, 53-345 Wrocław, Poland.
E-mail: wojciech.rybicki@ue.wroc.pl

some routine operations with matrices). Actually, with very little exaggeration, one may formulate (somewhat jokingly) a warning: "the matrices can be spotted everywhere" (from representations of multidimensional data, through "profound" statistics and econometrics, theory of games, investigations of random phenomena, up to some macroeconomic models).

According to the announcement given in the above mentioned paper, today's subject matter is to present applications of matrices to selected problems from welfare economics and stochastic dynamics of economic systems.

So in the first section the role matrices play as tools for indicating, comparing and measuring inequalities of distributions (of goods, resources or finite commodity bundles) is established. Simultaneously, operations with matrices treated as devices for equalizations of such inequalities (via reallocation procedures) are involved. The special role of double stochastic matrices has to be pointed out, on occasion.

The subsequent parts of the paper are devoted to showing the applications of matrices to modelling "Markov dynamics". The crucial role of transition matrices in this context is noted and accompanied by illustrative examples – chosen mainly from the area of Markov chains with finite or denumerable state spaces. We will mainly discuss the discrete time case, but some remarks on continuous processes and "their" families of matrices will also be given. The correspondence between the shape of the transition matrices with the behavior of processes is pointed out.

Moving on to the end of the introduction, the author wants to point out that he does not pursue the great generality when presenting the chosen notions and mathematical objects. One may easily find some general schemes, "governing" (obeying) the discussed special cases. They are the stochastic integral kernels (Markov kernels), the idea of mixing (of parameterized families of probability distributions), the concept of measurement of randomness of probability distributions (non-parametric) setting, double stochastic operators and some stochastic orders (especially convex, concave and Lorenz orders). Some generalization (just in the spirit of Markov kernels and the theory of mixtures of probability measures) are discussed in the subsequent paper the author has prepared for print in the next issue of Didactics of Mathematics (Rybicki, 2013). The matrices useful in the description of autoregressive schemes, the Youle-Wakker equations and some model of the financial market, will also appear in that article.

2. Double stochastic matrices and a comparison of distributions of commodity bundles

The idea of comparing certain n-vectors of numbers (of equal "lengths" and sums of their coordinates), reflecting the extent of their uniformity ("similarity" of coordinates) goes back to the paper of Muirhead (Muirhead 1903), the seminal theorem of Hardy, Littlewood and Polya (Hardy, Littlewood, Polya, 1929) and Schur (Schur, 1923) (pre)orders (Schmeidler, 1979; Le Breton, 1991; Arnold, 1987; Nermuth, 1993). Some significant elements of the original reasoning and constructions turned out to allow generalizations towards a wide variety of directions; very often they "encountered" other (sub)theories, "interwove" with them and developed with an increased impetus. From a historical perspective, one may note the "prevailing probabilistic share" in this development. Just inside of the "probability and statistics world", various types of stochastic majorization have appeared, as well as the notions of stochastic dominances and stochastic orders, such as convex and Lorenz orders – starting from one-dimensional case and culminating in the Choquet order, defined on spaces of probability measures on locally convex compact (and more general) spaces (Shaked, Shanthikumar, 1993; Mosler, Scarsini (Eds.), 1991; Phelps, 1966; Winkler, 1980). Let us mention (once again) that the related problems appeared in the "socio-economic costume" at the beginning of the 20th century – in the fundamental works of Lorenz (Lorenz, 1905) as well as Dalton (Dalton, 1920).

At the moment we restrict these considerations to the case of measures, supported by finite subsets of (say) Banach space. From now on, we will refer to the selected thoughts and observations made in the paper of Nermuth (Nermuth, 1993) – up to the end of the present section. On this occasion, we note that the above mentioned paper does not contain original mathematical topics. The author himself wrote (Nermuth, 1993): "In this paper I present neither a survey of known theories or else I try to show how different economic theories are connected, not on the surface in the sense that they address closely related economic problems, but in a deeper sense, viz. in their underlying logical structure". So, on account of the fact that matrices appear in that article, as an apparatus used for portraying and elaborating chosen models (examples), the article seems to be a "perfectly tailored" source for drawing some illustrations for our needs.

Let $\boldsymbol{x}=\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ and $\boldsymbol{y}=\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$ be subsets of a space $\boldsymbol{X}$, and let vectors $\mu_x=\{\mu_1, \ldots, \mu_m\}$ and $\nu_y=\{\nu_1, \ldots, \nu_n\}$ be two probability distributions supported by $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively (all μ_K and ν_s are non-negative and each of the vectors μ and ν has coordinated summing to one). The matrix **A** with non-negative entries

$$\mathbf{A}=(a_{ij})\ i=1,...,m;\ \ j=1,..., n \tag{1}$$

is called stochastic (or row-stochastic) if its rows sum to unit. A stochastic matrix **B** such that (additionally) its columns sum to one is called bi-stochastic (or double-stochastic). Denoting by b_i the i-th row of **B** and, analogously, by b^j – the j-th column of it, we may note

$$b_i \cdot e_n^T = 1,\ \ e_m \cdot b^j = 1,\ \ \mu \cdot e_m^T = \nu \cdot e_n^T = 1, \tag{2}$$

where e_n is a row vector consisting of n ones, e_m is a row vector consisting of m ones. The equivalent (but more compact) matrix notations of the first two conditions is

$$\mathbf{B}\cdot e_n^T = e_n^T \quad \text{and} \quad \mathbf{B}^T e_m^T = e_m^T. \tag{2a}$$

Definition 1 (Nermuth, 1993; Schmeidler, 1979). *A distribution $\boldsymbol{\mu}_x$ is less dispersed than ν_y if one of the following (equivalent!) conditions is satisfied*

$$\sum_i \mu_i\ f(\boldsymbol{x}_i) \le \sum_j \nu_j\, f(\boldsymbol{y}_s) \tag{3}$$

for all convex functions f: $\boldsymbol{X} \to \boldsymbol{R}$

$$\mu_x \cdot \mathbf{A} = \nu_x,\ \ x=\mathbf{A}y \text{ for a (row) stochastic matrix } \mathbf{A}. \tag{4}$$

Remarks:

a) The definition also contains the theorem of equivalences of conditions (3) and (4).

b) One may easily check that $\mu_x x = \nu_y$, which means that the distributions μ_x and ν_y have the same expectations.

c) The choice of the level of generality of the universe $\boldsymbol{X}$ was made intentionally: apart from the finiteness of support of distributions considered, we assumed of $\boldsymbol{X}$ to be a Banach space. For a further construction and examples given below, the minimal condition – of an algebraic character – on **X**, is its linear structure. This enables us to model problems, “beyond the real line”, covering the cases of multidimensional objects.

So, consider the case when $\boldsymbol{X}$ serves as a space of “feasible” l-bundles of commodities: ($\boldsymbol{X} = \boldsymbol{R}^l$) allocated for members of set **A** (agents, objects, persons) of (finite) cardinality. In this case one deals with the set $A = \{a_1, \ldots, a_n\}$ and vectors $\boldsymbol{x}_i \in \boldsymbol{X}$ such that $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{il})$, ($i = 1, 2, \ldots, n$), are interpreted as a basket of goods allocated to a “person” (without any loss of generality we may take $A = \{1, 2, \ldots, n\}$). Referring to the general model introduced above we note that have $n = m$ and $\mu = \nu = \left(\frac{1}{n}, \frac{1}{n}, \ldots \frac{1}{n}\right)$ (n-times) are uniform distributions and x_{ih} is the amount of commodity h allocated to person i. A crucial role is played in this case by the matrix **x** of n rows (representing agents) and l columns (representing commodities)

$$\mathbf{x} = (x_{ih}); \quad i = 1, \ldots, n;\ h = 1, \ldots, l. \tag{5}$$

The routine “translation” of the statements of Definition 1, together with the substitution of convex function by its concave counterpart u ($u = -f$) leads to the subsequent specifications. The matrices **x** will be called, in short, allocations.

Definition 1* (Nermuth, 1993). *An allocation* $\boldsymbol{x}$ *is more equal than allocation* $\boldsymbol{y}$ *if one of the following (equivalent) conditions is satisfied*

$$\sum_{i=1}^{n} u(\boldsymbol{x}_i) \geq \sum_{i=1}^{n} u(\boldsymbol{y}_i) \tag{6}$$

for all concave (utility) functions $u : \boldsymbol{R}^l \to \boldsymbol{R}$,

$$\boldsymbol{x} = \mathbf{B}\boldsymbol{y} \quad \textit{for a (certain) double – stochastic matric } \mathbf{B}. \tag{7}$$

Remarks:

a) Both of the conditions (6) and (7) have clear economic (in the spirit of welfare economics) interpretations. More equal allocations are preferred when applying a utilitarian criterion (and operating with an identical utility function, fulfilling a classical requirement – diminishing marginal utility). The “equalization” effect $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ can be achieved with the use of weighted averaging of bundles $(y_1, ..., y_n)$ with weights given, respectively by entries (b_{ih}) of bi-stochastic matrix **B.** So, for baskets $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ we have

$$\boldsymbol{x}_1 = \sum_{j=1}^{n} b_{1j}\boldsymbol{y}_j \text{ and } \boldsymbol{x}_2 = \sum_{j=1}^{n} b_{2j}\boldsymbol{y}_j. \tag{8}$$

b) The explanation of the socio-economic "content" of relations (7) and (8) in mathematical terms (and notions) is contained in the famous characterization of the collection of bi-stochastic matrices as a convex hull of the set of permutation matrices, due to Birkhoff (Birkhoff, 1946; Marshall, Olkin, 1979; Nermuth, 1993). The author announced this topic in his previous paper on reasons for telling students of economics about matrices (Rybicki, 2010).

c) It should be pointed out the role matrices play as "mixing advice". In the above σ-mixing (preceding by permuting) of given objects (bundles of commodities) aimed at "improving" (in a defined sense) allocation. Fixing one of the indices in a "pair-index" numbering terms of a given matrix, and then "integrating" it onto the second ones, results in a new bundle. The above scheme is common to the general theory of mixtures of distributions, Markov "mechanics", kernel (stochastic) dominance and so called double stochastic operators (see i.e. Mosler, Scarsini (Eds.), 1991; Feller, 1966; Szekli, 1995; Schmeidler, 1979). Anticipating further discussion, we merely outline at the time, the general pictures of functioning of the mentioned scheme

$$\mu(A) = \int_X \boldsymbol{B}(y, A)\nu(dy); \quad A \in \mathcal{B}, \tag{9}$$

$$p_j^t = \sum_{i \in S} p^t(i, j) p_i^t; \qquad j \in S. \tag{10}$$

The precise meaning of the above relations (as well as the explanations of symbols appearing in them) will be given later. Now we are only stating that B is a function of two variables: the states from the space S of sets of a Markov process and the sets belonging to σ-field $\mathcal{B}$, S denotes the state space of a Markov process and both of the relations express the way the "new" probability measure is built from the "old" one. One can also take another position and treat the transformations (9) and (10) as mixing performed by "mixers" ν (and p^t) on the families of distributions, parameterized with the help of variables y (or i), respectively).

Let us return to the mainstream of considerations, and restrict – for a moment – the discussion to the one dimensional case: $\boldsymbol{X} = \boldsymbol{R}$. In this case one speaks of income distributions. This branch of the subject constitutes the earliest (historically) part of the field of comparisons of equality (of sets of objects, namely – numbers; Lorenz, 1905; Dalton, 1920; Muirhead, 1903; Hardy, Littlewood, Polya, 1929).

Nermuth (Nermuth, 1993) mentioned the notions of social welfare functions being, at the same time, inequality measures as mappings

$$W : \boldsymbol{X}^n \to \boldsymbol{R} \tag{11}$$

monotonic (increasing) with respect to the above defined relation. Denoting it by symbol $\succ$ we may define (and say – according to points (6) and (7) of definition 1*)

$$\boldsymbol{x} \succ \boldsymbol{y} \quad \text{if } \boldsymbol{x} \text{ is more equal than } \boldsymbol{y}. \tag{12}$$

Then the condition for W to be called an inequality measure may be formulated as a requirement

$$W(\boldsymbol{x}) \geq W(\boldsymbol{y}) \ \text{ whenever } \ \boldsymbol{x} \succ \boldsymbol{y}, \tag{13}$$

which, in turn, is equivalent to concavity W in the Schur sense (Nermuth, 1993; Schmeidler, 1979).

Reasoning along this line, one may introduce the notion of (linear) equalizing redistribution as a linear transformation T assigning to every allocation $\boldsymbol{x}$ a new, more equal allocation $\boldsymbol{y}$. So T acts from $\boldsymbol{X}^{\boldsymbol{n}}$ to $\boldsymbol{X}^{\boldsymbol{n}}$. In the framework of the proceeding part of the paper, T has a matrix representation: there exists such a bi-stochastic matrix $\mathbf{B}$ that $T\boldsymbol{x} = \mathbf{B}x$. Nermuth mentioned a possibility of comparing a grade (or intensity) of equalization: summarizing his concept, T is shown to be more equalizing than T^* whenever $Tx \succ T^*x$. Then, for corresponding double stochastic matrices $\mathbf{B}$ and $\mathbf{B}^*$ (respectively) the following statement holds: there exists the third double-stochastic matrix $\mathbf{B}^{**}$ such that $\mathbf{B} = \mathbf{B}^{**} \cdot \mathbf{B}^*$. It is worth noting the relation of this condition with the measuring of riskiness and intensity of subjective aversion to risks (in the tradition of Kilhstrom-Mirman, 1974).

3. Matrices as devices for modeling Markovian random movements – generalities

There is a common agreement (among economists, as well as researchers acting in other fields) that the majority of phenomena and processes appearing in the real world should be modeled with the use of stochastic processes, for their inherited random (or incompletely determined) character. The simplest models of random dynamics are provided by random (real) sequences taking values from finite or countable sets. Such processes are also called time series or processes with a discrete time – for their variables are "numbered by subsequent moments (or periods) of time". The complete mathematical description of the above process (even with more general spaces of states) is given by defining all their finite dimensional distributions (satisfying the routine consistency conditions of the Kołmogorov-Daniell type, see e.g. Kingman 1972).

Consider a random sequence

$$X = (X_0, X_1, X_2, ...,) \tag{14}$$

consisting of real variables. For simplicity (and aiming not to obscure the main stream of reasoning) we omit the formal restrictions and precise assumptions. One may, for instance, assume that all the functions X_t are defined on the same probability space ($\mathbf{\Omega}$, $\mathcal{B}$, P). But, on the other hand, this assumption will not be exploited later. Nevertheless, one should not forget some tacit (usual) limitation, i.e. all of the appearing subsets ("events") of a real line are assumed to be Borel measurable sets.

Several kinds of questions concerning the "stochastic behavior" of variables of the above process may be of some interest. Firstly, one would want to know some "basic" families of probabilities.

$$P(X_0 \in A_0). \tag{15}$$

$$P(X_t \in A_t) \quad t = 0, 1, 2, \ldots \tag{16}$$

$$P(X_0 \in A_0,\ X_1 \in A_1,\ \ldots,\ X_t \in A_t);\ \ t = 0,1,\ \ldots \tag{17}$$

$$P(X_t \in A_t \,|\, X_{t-1} \in A_{t-1},\ X_{t-2} \in A_{t-2},\ \ldots,\ X_1 \in A_1,\ X_0 \in A_0)\ \ t = 1,2,\ \ldots \tag{18}$$

$$P(X_t \in A_t \,|\, X_{t-1} \in A_{t-1})\ \ t = 1,2,\ \ldots \tag{19}$$

The above expressions refer (respectively) to the initial distribution, the family of all finite-dimensional distributions (the joint probability distributions of the first finite "segments" – collections of variables of process), all the conditional distributions, when conditioning with respect to the "whole past" and merely to the latest period. Remember that, in the general case, calculating the above quantities turns out to be cumbersome. Note also, on occasion, the basic role of conditioning at each stage of calculations

$$\begin{aligned} &P(X_0 \in A_0,\ X_1 \in A_1,\ \ldots,\ X_t \in A_t) = \\ &= P(X_t \in A_t \,|\, X_{t-1} \in A_{t-1},\ X_{t-2} \in A_{t-2},\ \ldots,\ X_0 \in A_0)\cdot \\ &P(X_{t-1} \in A_{t-1} \,|\, X_{t-2} \in A_{t-2},\ \ldots,\ X_0 \in A_0)\cdot \\ &P(X_{t-2} \in A_{t-2} \,|\, X_{t-3} \in A_{t-3},\ \ldots,\ X_6 \in A_0)\cdot\ \ldots\ \cdot \\ &P(X_2 \in A_2 \,|\, X_1 \in A_1, X_0 \in A_0)_1\cdot \\ &P(X_1 \in A \,|\, X_0 \in A_0)\cdot P(X_0 \in A_0). \end{aligned} \tag{20}$$

It should be pointed out that all finite dimensional distributions (of arbitrary finite collections of variables of process) can be obtained from joint distributions of "full initial segments" $(X_0, X_1,\ \ldots,\ X_t)$.

Things are getting much simpler when it is possible "to forget" all but the latest moment in the history of process (there are many phenomena, when such simplifications are justified: the classical random walk, allowing one step moves, "up" or "down", serves as a commonly used example). In such a case the long tail on the right of the equality can be substituted by the (relatively) "shapely" recurrent expression

$$\begin{aligned} &P(X_0 \in A_0,\ \ldots,\ X_t \in A_t) = P(X_t \in A_t \,|\, X_{t-1} \in A_{t-1})\cdot \\ &P(X_{t-1} \in A_{t-1} \,|\, X_{t-2} \in A_{t-2})\cdot\ \ldots\ \cdot \\ &P(X_2 \in A_2 \,|\, X_1 \in A_1)\cdot \\ &P(X_t \in A_t \,|\, X_0 \in A_0)\cdot P(X_0 \in A_0). \end{aligned} \tag{21}$$

Abstracting, for a moment, on the "nature of time" one may formulate the idea of a Markov property as a requirement

$$P(X_t \in A \,|\, X_s = x_s,\ \ s \in \langle 0;\ \tau\rangle,\ \tau < t) = P(X_t \in A \,|\, X_\tau = x_\tau)\,. \tag{22}$$

We mention once again, that the heuristic convention, we have assumed, "allows" us to neglect some formal subtleness, such as requirements on relations to hold "almost sure" (with respect to a given "basic" probability measure), or noting "finite-dimensional" dependences in (22).

At this level of generality, one can comment on the two-parametric family of transition probabilities. Fixing *t*, *τ* we obtain the generalized "matrix", whose entries are numbered by pairs (set *A*, point *x*) and denote the probability of entering the respective sets at the moment t, "starting" from the state x at the moment $\tau < t$: $p_{\tau,t} = \left\{P_{\tau,t}(x, A)\right\};\ x \in S,\ A \in \mathcal{B}_{\boldsymbol{R}}$.

The symbol $\mathcal{B}_{\boldsymbol{R}}$ denotes a family of all Borel subsets of reals and the meaning of other used symbols will became clear in the light of the following notation

$$p_{\tau,t}(x, A) = P(X_t \in A \mid X_\tau = x). \tag{23}$$

This symbol $p_{\tau,t}$ denotes, of course, the probability of entering the set A at a moment t, conditioning the value of the process at an earlier instant τ to be x. One may regard the above explanation to be a pointless repetition of remarks preceding the formula (23) – but the author did it deliberately: to underline the importance of operation of conditioning, from the mathematical (probabilistic) point of view.

Coming back to the case of (discrete time) Markov chains with the finite or denumerable state spaces, we formulate the "original" Markov property

$$\begin{gathered} P(X_t = s_t \mid X_{t-1} = s_{t-1},\ \ldots,\ X_1 = s_1,\ X_0 = s_0) = P(X_t = s_t \mid X_{t-1} = s_{t-1}) \\ s_l \in \boldsymbol{S}, \quad l \geq 0. \end{gathered} \tag{24}$$

Abbreviating the notation, we may write

$$p_t(k, l) = P(X_{t+1} = s_l \mid X_t = s_k), \tag{25}$$

which is simply the probability of transition from the state k at the moment t to the state l at the next moment. So we again have obtained the (one-parameter) family of conditional probabilities – the so called one-step transition probabilities (numbered by discrete points of the time axis).

Assuming, for a moment, the finiteness of a state space $\boldsymbol{S}$ (i.e. $s = n$), one may arrange these quantities in a family of square ($n \times n$) matrices (transitions matrices for subsequent moments of time)

$$P_t = [p_t(k, l)];\ k, l = 1, \ldots, n;\ t = 0, 1, \ldots \tag{26}$$

The most basic applications of such matrices consist in operating with them to calculate subsequent "marginal" (one-dimensional or multidimensional) distributions.

Denoting by p_t the (row) vector – distribution of the variable X_t: $p_t = \left[p_t^1, \ldots, p_t^n \right]$, where

$$p_t^k = P(X_t = s_k) \tag{27}$$

we obtain the identity

$$p_{t+1} = p_t \cdot P_t \tag{28}$$

(the terms of vector p_{t+1} are calculated with the – multiple – use of the formula of total probability).

Iterating the procedure defined by the formula (28) (engaging the subsequent "present" ones and the matrices of one-step transition probabilities), results in the appearance of the expression for the probability distribution (of a variable of process in considerations) at time t. Given the distribution at point s and the product of $t – s$ one-step transition matrices

$$p_t = p_s \cdot (P_s \cdot P_{s+1} \cdot \ldots \cdot P_t). \tag{29}$$

The above equations inform us about the "mechanics" of transitions during the period of (discrete) time $\langle s,t \rangle$. In the "language" of probability transitions matrices, one can "codify" these partial calculations into the compact form

$$P_{s,t} = \prod_{k=0}^{t-s} P_{s+k}; \quad s < t, \tag{30}$$

where the symbol on the left denotes the matrix of (probabilities of) transitions (for all "feasible" pairs of states) from the moment s to the moment t.

The generalization of the above relations leads directly to the basic relations of the theory of Markov-type dependencies. For each triple of time moments (say, on the positive half-line of reals) $s < t < u$, the following equality holds

$$P_{s,t} \cdot P_{t,u} = P_{s,u}. \tag{31}$$

The equation (31) is known as the Chapman-Kolmogorov equation. It should be noted that (31) is valid also in the case of continuous time and that it is a "prototype" for various generalizations: integral representations (the same idea) and semi-group of transition (probability) operators (see Feller, 1966; Szekli, 1995). It is clear, that – on the other hand – the Chapman-Kolmogorov formula itself makes up a generalization of the original Markov requirement (condition) – which will be quoted (and briefly commented on), in the sequel.

Coming back to the models elaborated earlier (discrete time, discrete state-space), let us assume, for a moment, the independence of transition formulas on time. More precisely, one may consider a special case of the models discussed up till now: suppose all the one-step transition probabilities do not depend on the time parameter

$$p_t(k,l) = p(k,l) \quad \forall t \geq 0;\ k,l \in S. \tag{32}$$

This requirement involves, in turn, an analogous condition for one-step transition matrices (which cover simultaneously all "singletons" like (32)). So one may write

$$P_t \equiv P. \tag{33}$$

The corresponding Markov process is called a homogenous Markov chain. The slightly more general formulation of the condition (33) involves many-step transitions

$$P_{s,t} \equiv P_{(t-s)}. \tag{34}$$

The essence of the above conditions makes up the expression of dependence of a functional form of the Markov-type rule, governing the stochastic movement merely on differences between moments (periods) of time, but neglecting their absolute positions. So this is the reason for replacing the ordered pair of subscripts *s*, *t* by their difference in brackets at the right-hand subscript of symbol of matrix of transition probabilities $P_{(t-s)}$. The multiplication of matrices, and increasing them to the subsequent powers, then "produce" the desired representations of multi-stage transitions.

The present circumstances allow us to rewrite the Chapman-Kolmogorov equation in the form:

$$P_{(m+n)} = P_m \cdot P_n, \tag{35}$$

which is closer to the original Markov equation

$$p_{(m+n)}(k,l) = \sum_{j \in S} p_m(k,j)\, p_n(j,l) \tag{36}$$

for all natural m, n and all "triples" of states $k, j, l \in S$. Going a step ahead, we conclude that the matrix of transitions in (say) n step equals the n-th power of one-step transition matrices

$$P_{(n)} = P^{(n)}. \tag{37}$$

After outlining the basic features of applications of a matrix calculus for describing some elementary rules of a kind of random evolutions, we only mention the validity and effectiveness of the so called algebraic theory of Markov chains. This approach exploits merely "pure" algebraic techniques such as e.g. the Perron-Frobenius theory. The classification of states of chains and the distinguishing of special kinds of subsets of the state space (according to their role in the long-term behavior of the process) is strictly connected with the form transition matrices take: especially important are correspondences between sets of recurrent, periodic and transient states and their closure properties with the block-like shape of matrices or the impossibility of their decomposition (the irreducibility of chains and "their" matrices – see Rosenblatt, 1967; Feller, 1966 or Fisz, 1967). The same thing concerns the phenomenon of absorption (by some "barriers") and zeros of corresponding entries of transition matrices, and – in fact – the whole "finite state space" ergodic theory. We are not going in to pursue in this paper these (elementary, and – at the same time – very deep) theories, restricting our considerations to the basic notes on relations linking the description of (the simplest) random dynamics with elementary notions of linear algebra – on the level of definition, which were already presented.

4. Matrices as devices for modeling Markovian movements – some examples

Let us begin with the random sequences of i.i.d. – type[1] taking their values from the finite or countable set of states $S = \{0, 1, 2, \ldots\}$, according to the (fixed) probability distribution $p = (p_1, p_2, \ldots)$, ($P(X_t = s) = p_t$ for all variables X_t of the process in mind). In this case the transition probabilities do not depend on "starting states" (in subsequent moments) – simply, by definition. So all the rows of a transition matrix are the same (say, $p = (p_1, p_2, \ldots)$ and – of course – $p_s \geq 0$ ($\forall s \in S$), $\Sigma p_s = 1$ and

[1] The notation "i.i.d." is in common use as an abbreviation of the phrase: "independent, identically distributed".

$$P = \begin{pmatrix} p_0, & p_1, & p_2, & \cdots \\ p_0, & p_1, & p_2, & \cdots \\ p_0, & p_1, & p_2, & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}; \tag{38}$$

passing to the slightly more interesting (class of) cases is possible in an easy way. One should substitute the "original" i.i.d. moves (or changes) by their cumulative effects: the variables X_s are replaced by their partial sums

$$S_n = \sum_{t=0}^{n} X_t. \tag{39}$$

The simplest process belonging to this class is a model of movement of a particle, walking on the lattice of nonnegative integers. The subsequent steps are assumed to be i.i.d. jumps of a unit "up" or "down" $(X_t \in \{-1, 1\})$ with respective probabilities $p = P(X_t = 1)$ and $q = P(X_t = -1)$. The above regime results in the behavior of process of cumulated sums of individual, subsequent moves. The variables of process S_n do not remain independent, the process itself becomes a Markov chain – the (discrete: time and the state space) random walk. If we allow this walk to vary without bounds, the unbounded random walk is obtained, which makes up a mathematical idealization of (a simplified) physical phenomenon. Before writing the corresponding transition matrix we are going to discuss (very briefly) the other possibilities: there may appear two (or one) "traps" – so called absorbing states (barriers, screens). If a particle attains such a state then it remains there forever, the probability of getting out is zero. These restrictions may be relaxed to admit leaving the screen with positive probability (so called elastic screens) or even guaranteeing ("automatic") escaping to the nearest previous position with probability one (the case of reflecting screens).

There are at least two other stories which may be associated with the above model - one may equally well decide to take the inverse direction of reasoning – the formal models being involved in the stories one would conceive. One of them is the classical problem of a gambler's ruin. The player visiting a casino (or simply – deciding the "fair" or "unfair" coin) performs a sequence of independent repetitions of bets which could result in gains (of say one dollar) or of the loss (of the same size) with, respectively, probabilities p or $1 - p$. The player begins his/her game with the initial capital (say) "a", while the capital of the casino may be assumed to be a "large" (positive, integer) number "$C - a$" or even infinity. The situation of infinite gambles may be modeled by the assumption, that the level zero plays a role of a reflecting barrier (the "benevolent uncle" who lends him/her every time a lacking dollar). The interpretation of the discussed random evolution of great importance may be expressed in terms of fluctuations of a level of capital (or resources) of an insurance company, confronted with random shocks (unavoidable expenditure when claims are reported), disturbing its steady – usually, linear – increase of capital (flows of premiums from policyholders).

Let us start with describing a model with two absorbing screens located at points 1 and u (the integer greater than 1). The (one-step) transition probabilities matrix of the discussed above random walk, can be easily deduced from the assumptions. So we have

$$p(1,1) = p(u,u) = 1, \quad \text{and for } 2 \le i \le u-1$$

$$p(i,j) = \begin{cases} p & \text{if } j = i+1 \\ q & \text{if } j = i-1 \\ 0 & \text{if } |i-j| > 1 \end{cases} .$$

Hence the transition matrix of this homogenous (in time) chain is a square matrix of dimension u and it takes a form:

$$P = \begin{bmatrix} 1 & 0 & 0 & \dots\dots & \dots & \dots & 0 \\ q & 0 & p & 0\dots & \dots & \dots & 0 \\ 0 & q & 0 & p & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}. \tag{40}$$

On the other extreme, one may find (an unbounded) random walk with countable space of states (numbered by positive integers), where the state "1" makes so called reflecting screen p (1, 2) = 1. There is not, in the present case, any upper limit for the described movements, but the remaining transition probabilities are the same as in the previous example

$$P = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & \dots \\ q & 0 & p & 0 & \dots & \dots \\ 0 & q & 0 & p & \dots & \dots \\ 0 & 0 & q & 0 & p & \dots \\ \dots & \dots & \dots & \dots & \dots & \end{bmatrix}. \tag{41}$$

If we allow a particle to stay at the state 0 with probability q and escape to the state 1 with probability $p = 1 - q$ $(0 < q < 1)$, then the chain obtained in this fashion is said to have an elastic screen at 1. Doing the same with (say) level u leads to random walk with two elastic screens – this is again a bounded, homogeneous Markov chain, whose transitions are governed by the u x u matrix

$$P = \begin{bmatrix} q & p & 0 & \dots & \dots & 0 \\ q & 0 & p & \dots & \dots & 0 \\ 0 & q & 0 & p & \dots & 0 \\ 0 & \dots & q & 0 & p & 0 \\ \dots & \dots & \dots & \dots & q & p \end{bmatrix}. \tag{42}$$

In the terminology of (repeated) games, the players decide (with probabilities p or q, respectively) to give a chance of continuation of gambling for their opponents. Analogously: banks admit (conditionally) allowing the insurance company to continue its functioning.

A more realistic case is described by the matrix (Feller 1966):

$$P = \begin{bmatrix} q_0 & p_0 & 0 & 0 & \dots & \dots \\ q_1 & 0 & p_1 & 0 & \dots & \dots \\ 0 & q_2 & 0 & p_2 & \dots & \dots \\ 0 & 0 & q_2 & 0 & p_2 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}, \tag{43}$$

which corresponds to random walk on the positive half-line (the state-space consists of nonnegative integers). In this case the probabilities of transitions depend on the numbers (“magnitudes”) of states. They change – the (homogeneous in time) chain has lost the property of being homogeneous in the state space but, of course, $p_k + q_k = 1, \;\; p_k \geq 0, \;\; q_k \geq 0; \;\; k \in \{0, 1, \dots\}$. Such chains play an important role as (discrete) models in the theory of birth and death processes.

5. Concluding remarks

In the article, several classes of application of matrices in the field of economics were exhibited. Students of economics encounter such problems at various stages of their high-school education. Continuing the beginnings undertaken in his previous paper, the author aimed to extract the essential, decisive "mechanics, hidden between square brackets" – the commonly known pictures. The content enlightened in the paper may be seen as two-fold: (i) the matrices as devices equalizing "unjust" allocations (or distributions) of goods, services or resources; (ii) matrices as generators of rules of (stochastic, Markov-type) movements. The models of the first type appear in the context of welfare economics, the second inevitably enter the branches of operation research (the "classics": queuing, renewal and reliability theories, as well as the problems of (insurance) risk processes, obligatorily present in the modern courses conducted for students of economics. We actually restricted the modeling to the simplest (discrete time and state-spaces) cases.

What should be concluded from the above consideration – in respect of the methods of teaching students (familiarizing them with elementary mathematical notions and convincing them as to the fairness and effectiveness of the "culture" of mathematical treating phenomena of quantitative economics – in the order of learning other subjects)? First of all, it would be desirable for them to note the multi-sided, but – in a sense – unique nature of matrices as operators acting in the "linearized world": contracting, extending, equalizing, mixing, comparing and rotating the systems of vital significance, and (at the same time) – generators of "infinite sequences of one step random moves" of processes. The theme of comparing will be presented in the paper (Rybicki, 2013). A notable part of considerations concerned basic relations linking the description of (the simplest) random dynamics with elementary notions of linear algebra – on the definition level. In the author's own opinion, the above mentioned – mathematically trivial – findings are of some didactic (methodological) importance: they illustrate the linear character of Markovian stochastics, the "mystery of randomness" is resolved as a sequence of multiplying subsequent probability vectors by stochastic matrices. One may – somewhat provocatively – say that such "randomness is nothing but linearity"! But, on the other hand, one may be aware of the fundamental danger hidden in such logical abbreviations and abstractions. The simplicity may easily degenerate into crudeness (or, at least – charlatanry) which in turn, leads directly to lying to the beginners (first course students of economics), which – of course – is an inadmissible practice (especially in the "positive" exposition of the subject!). The idea of randomness must not be "lost" or "deleted" – teachers are expected to suggest effective ways of (and tools for) solving problems and to reveal the various perspectives of looking at them, but nothing like the negation of models before presenting them.

The considerations of the paper are to be continued in the accompanying paper of the author, submitted to print in the same issue of Didactics of Mathematics (Rybicki, 2012). The mentioned article (entitled "Further examples of the appearance of matrices (and the role they play) in the course of education of economists") is devoted to introducing families of transition probability matrices and to stressing the role of intensity matrices – for processes in continuous time; some remarks on "Poissonian mechanics and its matrices" are to be made; elementary facts from the "finite states" ergodic theory are reported and the role of eigenvectors in the problems of the choice of the structure of inputs, given the technology (matrix), is demonstrated.

6. Summary

The subject matter of the paper is to present applications of matrices to the selected problems from welfare economics and stochastic dynamics of economic systems. In the first section, the role matrices play as tools for indicating, comparing and measuring inequalities of distributions (of goods, resources or finite commodity bundles) is established. Simultaneously, operations with matrices treated as devices for equalizations of such inequalities (via reallocation procedures) are involved. The special role of double stochastic matrices has been pointed out. Subsequent parts of the paper are devoted to show the applications of matrices to modeling "Markov dynamics". The key role of transition matrices in this context is noted and accompanied by illustrative examples – chosen from the area of Markov chains with finite or denumerable state spaces. The discrete time case is discussed, but some remarks on continuous processes and "their" families of matrices are also given. The strict connections between the shape of the transition matrices and the behavior of processes is pointed out.

References

Arnold B. (1987). *Majorization and the Lorenz Order: A Brief Introduction*. Springer Verlag. Berlin–Heidelberg–New York–Paris–Tokyo.

Birkhoff G. (1946). *Tres observaciones sobre al algebra lineal*. Univ. Nac. Tucuman Rev. Ser A.5. Pp. 147-151.

Dalton H. (1920). *The measurement of the inequality of incomes*. Economic Journal 30. Pp. 348-361.

Feller W. (1966). *Wstęp do rachunku prawdopodobieństwa*. PWN. Warszawa.

Fisz M. (1967). *Rachunek prawdopodobieństwa i statystyka matematyczna*. PWN. Warszawa.

Hardy G., Littwood J., Polya G. (1929). *Some Simple Inequalities Satisfying by Convex Functions*. Mess. of Math. 58. Pp. 145-152.

Kilhstrom R.E., Mirman L.J. (1974). *Risk Aversion with Many Commodities*. Journal of Economic Theory. Vol. 8. Str. 361-388.

Kingman J.F.C. (1972). *Regenerative Phenomena*. J. Willey & Sons Ltd. London–New York–Sydney–Toronto.

Le Breton M. (1991). *Stochastic Orders In Welfare Economics*. In: K. Mosler, M. Scarsini (Eds.). *Stochastic Orders and Decision Under Risk*. Inst. of Math. Statistics. Lecture Notes – Monograph Series. Vol. 19. Pp. 190-206.

Lorenz M. (1905). *Methods of measuring the concentration of wealth*. Publ. of the American Statistical Association 9. Pp. 209-219.

Marshall A.W., Olkin J. (1979). *Inequalities: Theory of Majorization and Application.* Academic Press. New York.

Mosler K., Scarsini M. (Eds.) (1991). *Stochastic Orders and Decision Under Risk*, Inst. Of Math. Statistics. Lecture Notes – Monograph Series. Vol. 19. Hayward. California.

Muirhead R. (1903). *Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters*. Proceedings of the Edinburg Math. Society 21. Pp. 144-157.

Nermuth M. (1993). *Different Economic theories with the Same Formal Structures: Risk, Inequality, Information Structures etc*. In: E. Diewert, K. Spremann, F. Stehling (Eds.). *Mathematical Modelling*. Springer Verlag. Pp. 271-277.

Phelps R. (1966). *Lectures on Choquet's Theorem*, D. Van Nostrand Company Inc. Princeton. New Jersey. Toronto–New York–London.

Rosenblatt M. (1967). *Procesy stochastyczne*. PWN. Warszawa.

Rybicki W. (2010). *Kilka powodów dla których opowiadamy studentom ekonomii o macierzach*. Didactics of Mathematics 7(11). Wrocław University of Economics. Pp. 109-126.

Rybicki W. (2012). *Further Examples of the Appearance of Matrices (and Role They Play) in the Course of the Economic Education*. Didactics of Mathematics (this issue).

Rybicki W. (2013). *The Role of Matrices in the Contemporary Education of Students of Economics – Further Remarks and Examples of Applications* – the paper prepared for the publication in Didactics of Mathematics 10(14).

Schmeidler D. (1979). *A Bibliographical Note on Theorem of Hardy*. Littlewood and Polya. Journal of Economic Theory 20. Pp. 125-128.

Schur J. (1923). *Über eine klasse von mittel-bildungen mit anwendungen die determinaten*. Theorie Sitzungsber Berlin Math. Gesalschaft 22. Pp. 9-20.

Shaked M., Shanthikumar J. (1993). *Stochastic Orders and Their Applications*. Academic Press. New York.

Szekli R. (1995). *Stochastic Ordering and Dependence in Applied Probability*. Springer Verlag. New York.

Winkler G. (1980). *Choquet Order and Simplices*. Springer-Verlag. Berlin–Heidelberg–New York–Tokyo.