ISSN 1407-5806

COMPUTER MODELLING AND NEW TECHNOLOGIES

Volume 15
No 2

2011

Transporta un sakaru institūts
(Transport and Telecommunication Institute)

Computer Modelling and New Technologies

Volume 15, No.2 – 2011

ISSN 1407-5806
ISSN 1407-5814
(On-line: www.tsi.lv)

Riga – 2011

Host Organization:
Transport and Telecommunication Institute

Supporting Organizations:
Latvian Transport Development and Education Association
Latvian Academy of Sciences
Latvian Operations Research Society

THE JOURNAL IS DESIGNED FOR PUBLISHING PAPERS CONCERNING THE FOLLOWING FIELDS OF RESEARCH:

- mathematical and computer modelling
- mathematical methods in natural and engineering sciences
- physical and technical sciences
- computer sciences and technologies
- material science, solid state physics and chemistry
- nanomaterials, nanodevices and nanotechnologies
- aviation and aerospace technologies
- electronics and telecommunication
- transport and logistics
- economics and management
- innovative education
- social sciences

In journal articles can be presented in English. All articles are reviewed.

EDITORIAL CORRESPONDENCE
Transporta un sakaru institūts (Transport and Telecommunication Institute)
Lomonosova iela 1, LV-1019, Riga, Latvia. Phone: (+371) 67100593. Fax: (+371) 67100535
E-mail: journal@tsi.lv, www.tsi.lv

COMPUTER MODELLING AND NEW TECHNOLOGIES, 2011, Vol. 15, No.2
ISSN 1407-5806, **ISSN** 1407-5814 (on-line: www.tsi.lv)
Scientific and research journal of Transport and Telecommunication Institute (Riga, Latvia)
The journal is being published since 1996.

CONTENTS

Editors' Remarks

From the Poem

THE BALLAD OF EAST AND WEST

****

Oh, East is East, and West is West, and never the twain shall meet,

Till Earth and Sky stand presently at God's great Judgment Seat;

But there is neither East nor West, Border, nor Breed, nor Birth,

When two strong men stand face to face,

tho' they come from the ends of the earth!

****

Rudyard Kipling♣, 1889

The 15th volume No.2 presents the papers on the actual topics such as **Applied Statistics and Operations Research**, **Mathematical Methods**, **Information Processing** and **Mathematical Modelling**.

Our journal policy is directed on the fundamental and applied sciences researches, which are the basement of a full-scale modelling in practice.

This edition is the continuation of our publishing activities. We expect our journal will be interesting for research community, and we are open for collaboration both in research and publishing. This number continues the current 2011 year of our publishing work. We hope that journal's contributors will consider the collaboration with the Editorial Board as useful and constructive.

EDITORS

Yu.N. Shunin

I.V. Kabashkin

♣ **Joseph Rudyard Kipling** (30 December 1865 – 18 January 1936) - an English poet, short-story writer, and novelist chiefly remembered for his celebration of British imperialism, tales and poems of British soldiers in India, and his tales for children. Kipling received the 1907 Nobel Prize for Literature. He was born in Bombay, in the Bombay Presidency of British India. Kipling is best known for his works of fiction, including *The Jungle Book* (a collection of stories which includes "Rikki-Tikki-Tavi"), *Just So Stories* (1902) (1894), *Kim* (1901) (a tale of adventure), many short stories, including "The Man Who Would Be King" (1888); and his poems, including *Mandalay* (1890), *Gunga Din* (1890), *The White Man's Burden* (1899).

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 7–20
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

ADJUSTING MARKOV MODELS TO CHANGES IN MAINTENANCE POLICY FOR RELIABILITY ANALYSIS

Jarosław Sugier

Wrocław University of Technology
Institute of Computer Engineering, Control and Robotics
ul. Janiszewskiego 11/17, 50-372 Wrocław, Poland
jaroslaw.sugier@pwr.wroc.pl

Equipment deterioration can be analysed with discrete state-transition models which are used as a foundation for numerical evaluation of various reliability and operational parameters. In this paper we discuss an approach in which a system with scheduled inspections and possible repair activities is described by the discrete state-transition deterioration model. The model can be translated into a semi-Markov process which, after solving, yields numerous reliability characteristics including average equipment life, its deterioration rate represented by so called life curve, probability of failure within given time horizon, etc. The analysis is performed for particular maintenance policy which has been integrated in the model and by adjusting the model to specific changes in this policy (e.g., modifications of repair frequencies) their consequences for system reliability can be evaluated. In the text we present briefly the methodology behind model creation and concentrate on its one specific aspect: automatic adaptation of the model to the adjusted maintenance policy with modified frequencies of repairs. In particular, it is shown that such adaptation can be reduced to the adjustment of transition probabilities that are found in the model, and such adjustment can be numerically implemented with one of the standard root-finding algorithms. Upon presentation of the adjustment method, its use is illustrated by a description of the computer tool which helps in evaluating reliability and financial effects of changes in maintenance policy of an ageing equipment.

Keywords: *state-transition deterioration model, semi-Markov process, reliability analysis, root-finding algorithm, maintenance analysis*

1. Introduction

Efficient and, at the same time, cost-effective maintenance is an important element of reliable operation in contemporary complex technical systems. Selecting the optimal maintenance strategy must take numerous issues into account and among them reliability and economic factors are often of equal importance. On one side, it is obvious that for successful system operation failures must be avoided and this opts for extensive and frequent maintenance activities. On the other, superfluous maintenance may result in very large and unnecessary cost. Finding a reasonable balance between these two is a key point in reliable system operation.

In order to be able to plan such maintenance appropriate models are necessary that would represent equipment deterioration process and, at the same time, would take into account various maintenance operations. In this paper we present a methodology that assists a person who decides about maintenance activities by evaluating risks and costs associated with choosing different maintenance strategies. Instead of searching for a globally optimal solution to a problem: "what maintenance strategy would lead to the best reliability and dependability parameters of system operation", in this approach different maintenance scenarios can be examined in "what-if" studies and their reliability and economic effects can be compared so that a person managing the maintenance is assisted in making informed decisions ([1–3]).

The method has been presented initially in [1] and its specific extensions were further described in [4–8]. In this work, we summarize the current state of development and concentrate on one important aspect of the methodology: fully automatic adjustment of the model to possible modifications of the maintenance policy which are often required for studies requested by the user. Also we will provide an illustration of the application of the method presenting a computer tool which helps in evaluating reliability and financial effects of changes in maintenance policy of an ageing equipment.

The main contents of the paper is divided into three parts. The first one (section 2) presents general description of the methodology which is based on state-transition deterioration models, semi-Markov processes and the concept of a life curve used for visualisation of equipment ageing. The problem of automatic model adjustment is presented in the second part (section 3) which includes detailed discussion of possible numerical algorithms that can be implemented for probability approximation. Finally, in the last part (section 4) the method of model adjustment is used in a computer tool which uses the concept of a life curve and discounted cost to study the effect of the equipment ageing under different hypothetical maintenance strategies

2. Modelling the Ageing Process

There are numerous factors that have an effect on the ageing process of equipment that undergoes scheduled inspections and maintenance actions. Among them there are various aspects related to its physical characteristics, operating practices, and the maintenance policy. The method described in this paper uses a generic model that assumes that the equipment will deteriorate in time and, if not maintained, will eventually fail. To counteract, the scheduled inspections are performed and if the deterioration process is discovered, preventive maintenance is applied which can restore the condition of the equipment. Such a maintenance activity will return the system to a specific state of deterioration, whereas repair after failure will restore to "as-new" condition ([9–10]). With these assumptions, the maintenance policy components that must be recognized in the model are: monitoring or inspection (how the equipment state is determined), the decision process (which determines the outcome of the decision), and, finally, the maintenance actions – the repairs (or possible decision outcomes).

2.1. The State-Transition Model

One of the approaches that can properly incorporate all the above suppositions about the aging process and maintenance activities is based on state-space (Markov) model ([11–16]). The model consists of the states the equipment can assume in the process, and the possible transitions between them.

The method described in this paper uses a model of the Asset Maintenance Planner (AMP) that was initially developed and implemented by George J. Anders and Henryk Maciejewski ([17–18]). The AMP model is designed for equipment exposed to deterioration but undergoing maintenance at prescribed times. It computes the probabilities, frequencies and mean durations of the states of such equipment. The basic ideas in this approach are the probabilistic representation of the deterioration process through discrete stages, and the provision of a link between deterioration and maintenance.

For structure of a typical AMP model see Figure 1. In the model, the deterioration progress is represented by a chain of deterioration states $D1$ … DK which then leads to the failure state F. In most situations, it is sufficient to represent deterioration by three stages: an initial ($D1$), a minor ($D2$), and a major ($D3$) stage of deterioration ($K = 3$). This last is followed, in due time, by equipment failure (F) which requires extensive repair or replacement.

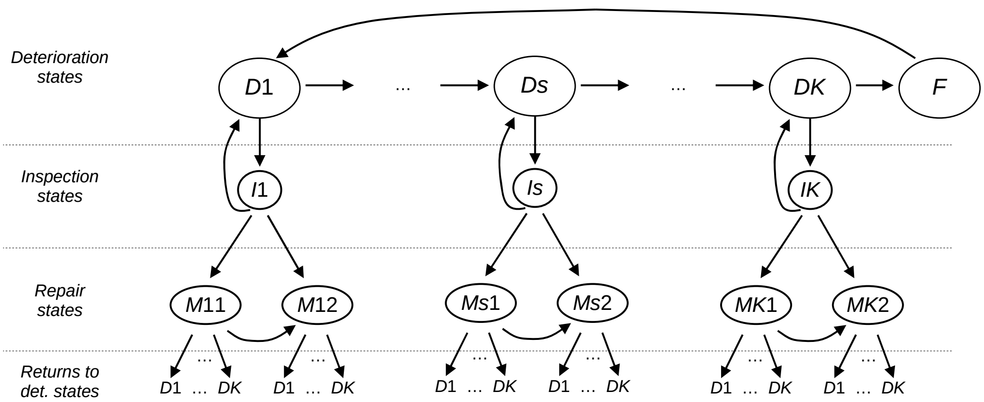


Figure 1. The state-transition model representing the deterioration chain with inspection and repair states (an example with two types of repairs is shown)

In order to slow deterioration and thereby extend equipment lifetime, the operator will carry out maintenance according to some pre-defined policy. In the model of Figure 1, regular inspections (I*s* states) are performed which result in decisions to continue with, e.g., minor ($Ms1$) or major ($Ms2$) repair (more than two types of repairs can be modelled), or to return to the deterioration state Ds without any repair. The expected result of all maintenance activities is a single-step improvement in the deterioration chain; however, it is possible to take into account also cases where no improvement is acquired or even where some damage is done through human error in carrying out the maintenance, which results in returning to the stage of more advanced deterioration.

The choice probabilities (at transitions from inspection states) and the probabilities associated with the various possible outcomes are based on user input and can be estimated, e.g., from historical records or operator expertise. Therefore, creation of the model and then its fine-tuning to some real historical data of equipment operation and maintenance records is a complex task that requires expert intervention.

Mathematically, the state-transition model of Figure 1 can be translated into a semi-Markov process, and then solved by the well-known procedures. The solution will yield all the state probabilities, frequencies and mean durations. Another technique, employed for computing the so-called first passage times (FPT) between states, will provide the average times for first reaching any state from any other state. If the end state of the passage is F, the FPTs are the mean remaining lifetimes from any of the initiating states. For state *D*1 this estimates the expected equipment life for the maintenance policy that has been incorporated in the model.

2.2. Using the Model in Reliability Analysis

A convenient way to represent the deterioration process is by the life curve of the equipment ([9]). Such a curve (see the first graph in Figure 2) shows the relationship between asset condition, expressed in either engineering or financial terms, and time. This concept is easy to comprehend for a non-expert end user (for example, a manager analysing various options of the maintenance policies) who does not need to know all the intrinsic details of state-transition models and Markov processes.

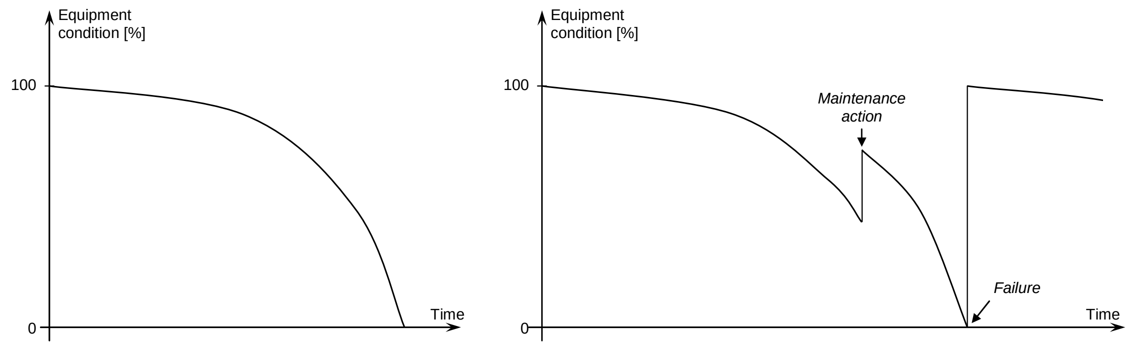


Figure 2. A simple life curve computed from the semi-Markov model (left) and a complex life curve representing some maintenance scenario (right)

A life curve that corresponds to some given Markov model can be created as follows. As pointed out above, computing the average first passage time (FPT) from the first deterioration state (D1) to the failure state (F) yields an average lifetime of the equipment, i.e., the length of the curve. On the other hand, solving the model for the state probabilities makes possible computing the expected state durations, which are then used to determine the shape of the curve, i.e. the rate of deterioration over different phases of equipment wear (some additional decisions are required as to how the deterioration states are mapped to ranges of the asset condition values).

Simple life curves obtained for specific maintenance policies (i.e. specific models) can later be combined in composite life curves which describe possible complex maintenance scenarios ([1], [8]). The second graph in Figure 2 shows an example of such a scenario when a preventive maintenance is performed at some moment in time (restoring the asset condition to approx. 80%) after which the failure occurs leading to equipment replacement. This curve is composed of three segments of simple life curves and they do not need to represent the same models, i.e. a situation can be modelled when a change in maintenance policy takes place, e.g. after the repair.

Furthermore, having the model and its (simple) life curve, one can compute the probability of failure (PoF) within given time period *T* for the equipment which currently is in some specific deterioration condition *AC*. The procedure is as follows:

(1) for the current value of asset condition *AC*, find from the life curve the corresponding deterioration state *Dc* and then compute a state progress *SP* (%), i.e. estimate how long the equipment has already been in the *Dc* state
(2) running FPT analysis on the model, find the probability distribution functions d_c(t) and d_{c+1}(t) of First Passage Time from the current state *Dc* and the subsequent deterioration state *D*(*c* + 1), to the failure state *F*
(3) interpreting the state progress *SP* as a weight which balances the current equipment condition between *Dc* and *D*(*c* + 1), estimate the final value of the probability as:

$$\mathrm{PoF}(\ T\ ) = d_c(\ T\ ) \cdot (1 - \mathrm{SP}) + d_{c+1}(\ T\ ) \cdot SP. \tag{1}$$

When equipment deterioration over the time period T is represented by a complex life curve, then the above procedure must be applied to the every simple curve segment which is included in T.

3. Adjusting the Model to Different Repair Frequencies

In practical applications when different maintenance policies need to be investigated, one of the most important aspect of system operation is appropriate representation of various types and frequencies of repairs that may be included in possible maintenance policies under consideration. In the approach presented in this paper it is the task of the expert to properly incorporate these characteristics of the original ("continue as before") policy in the model, but once such model is available the non-expert end user may wish to analyse various hypothetical policies that are created by simple modifications of the original repair frequencies. For example, the end user may wish to consider an option "what if the minor repair rate is reduced by half" or "what if the minor and major repairs are removed completely but the medium repair is performed twice as often". Such studies require a mechanism that, having the expert-created model representing some original repair policy, would be able to generate a derived model that would correspond to the same equipment (i.e. the same deterioration chain as presented in Figure 1) but submitted to a modified repair policy with different repair frequencies. It should be noted that the same mechanism would be also helpful during construction of the model by the expert when fine-tuning of repair frequencies is needed in order to achieve compliance of the model with some real-world repair policy stored in historical records of equipment operation.

3.1. The Adjustment Procedure

Let's assume that the deterioration model under consideration consists of K deterioration states and R repairs. Also, let P^{sr} denotes probability of selecting maintenance r in state s (assigned to the decision after inspection state Is) and P^{s0} represents probability of returning to state Ds from inspection Is which corresponds to a situation when no maintenance is scheduled as a result of the inspection. The foremost condition that must be met at all times is that in all deterioration states $s = 1 \ldots K$:

$$\mathrm{P}^{s0} + \sum_r \mathrm{P}^{sr} = 1\,. \tag{2}$$

Let F^r represents the frequency of some repair r as it is generated by the model. The problem of model adjustment can be formulated with various assumptions and with different goals in mind but in this approach it is defined as follows:

> Given an initial semi-Markov model M_0, with internal structure representing deterioration, inspection and repair states as described above and producing the initial vector of repair frequencies $\mathbf{F}_0 = [\ \mathrm{F}_0^1,\ \mathrm{F}_0^2\ \ldots\ \mathrm{F}_0^R\ ]$, modify the probabilities P^{sr} assigned to transitions from inspections states Is so that the resulting model generates some requested vector of goal frequencies $\mathbf{F}_\mathrm{G}$.

Usually, the vector $\mathbf{F}_\mathrm{G}$ may represent the observed historical values of the repair frequencies (when an expert user is working on fine-tuning of the initial model) or some hypothetical frequencies (when an end user investigates possible modifications of the repair policy).

There are numerous approaches that can be used in order to accomplish such model adjustment. In the proposed solution, an iterative approximation approach has been chosen in order to preserve an original construction of the model M_0 as mush as possible. In this method a sequence of tuned models $M_0, M_1, M_2, \ldots M_N$ is evaluated in N steps with each consecutive model approximating desired goal with a better accuracy. Starting with $i = 0$, the procedure consists in the following steps:

1° for the current model M_i compute its vector of repair frequencies $\mathbf{F}_i$

2° evaluate an error of M_i as a distance between vectors $\mathbf{F}_\mathrm{G}$ and $\mathbf{F}_i$

3° if the error is within the user-defined limit ε, consider M_i as the final model and stop the procedure ($N = i$); otherwise proceed to the next step

4° create a new model M_{i+1} by *tuning* values of P_i^{sr}, then correct P_i^{s0} according to condition (2)

5° return to the step 1° for the next iteration.

The error computed in step 2° can be expressed in may ways. As the absolute values of repair frequencies may vary in a broad range within one vector $\mathbf{F}_i$, yet the values of all are significant for model evaluation, the relative measures work best in practice:

$$\|\mathbf{F}_G - \mathbf{F}_i\| = \frac{1}{R}\sum_{r=1}^{R}\left|F_i^r / F_G^r - 1\right| \tag{3}$$

or

$$\|\mathbf{F}_G - \mathbf{F}_i\| = \max_r \left|F_i^r / F_G^r - 1\right| . \tag{4}$$

The latter formula is more restrictive: it ensures that any repair does not differ from the goal more than the imposed limit and therefore this version was used in the numerical implementation of the method that is considered in this paper.

3.2. Generation of the Adjusted Model

Of all the steps that compose the iteration run, it is clear that adjusting probabilities P_i^{sr} in step 4° is the heart of the whole method. This is accomplished with the following two assumptions that are introduced not only to simplify the task, but also to improve quality of the result.

The first assumption is related to the fact that although, in general, the P^{sr} probabilities represent $K \cdot R$ free parameters that could be freely modified in order to arrive at the requested goal, their uncontrolled modification can lead to serious deformation of the model and this should be avoided. To this point a restrictive condition is adopted: if the probability of some particular repair must be modified, it is modified *proportionally* in *all* deterioration states, so that during the adjustment the proportion between this repair probabilities over all states remains unchanged and are the same as in the initial model M_0:

$$\forall\ i, r \quad P_0^{1r} : P_0^{2r} : \ldots : P_0^{Kr} \ \sim\ P_i^{1r} : P_i^{2r} : \ldots : P_i^{Kr} . \tag{5}$$

This assumption also significantly reduces dimensionality of the problem, because now only R scaling factors, denoted as the vector $\mathbf{X}_{i+1}= [X_{i+1}^1,\ X_{i+1}^2,\ \ldots\ X_{i+1}^R]$, must be found to compute all new probabilities required to create the model M_{i+1}:

$$P_{i+1}^{sr} = X_{i+1}^r \cdot P_0^{sr}, \quad r = 1\ldots R,\ s = 1\ldots K. \tag{6}$$

Moreover, and this observation leads to the second assumption, although the frequency of a repair r depends on the probabilities of all repairs (modifying probability of one repair changes, among others, state durations in the whole model; thus, it changes the frequency of all states) it can be assumed that, in a case of a single-step small adjustment, its dependence on repairs other than r can be considered negligible and

$$F_i^r = F_i^r\left(X_i^1, X_i^2 \ldots X_i^R\right) \approx F_i^r\left(X_i^r\right). \tag{7}$$

With these two assumptions, generation of a new model in step 4° of the above described procedure is reduced to the problem of solving R non-linear equations in the form of

$$F_i^r\left(X_i^r\right) = F_G^r . \tag{8}$$

This task can be accomplished with one of the standard root-finding algorithms which will be presented in the next point.

One aspect of the procedure requires additional attention, though: applying equation (6) with $X_{i+1} > 1$ may violate condition

$$\sum_r P_{i+1}^{sr} \le 1 \tag{9}$$

in some deterioration state s. This situation needs special tests that would detect such illegal probability values and then reduce them proportionally so that their sum does not exceed 1: a so called *scale-down transformation* needs to be applied. As the case studies show such situations do occur during model tuning towards repair frequencies that are remarkably higher than in the initial model M_0. In its simplest form, the scale-down operation consists in dividing each probability P^{sr} in the offending state s by the sum of all repair probabilities in this state, as they are computed with equation (6) without scaling:

$$P^{sr} = P^{sr} / S_{Ds}\,, \quad S_{Ds} = \sum_{r=1}^{R} P^{sr}\,. \tag{10}$$

This will also lead to $P^{s0} = 0$ which means that every inspection ends with some repair and there are no direct returns from inspection state *Is* to deterioration state *Ds*. Moreover, this obligatory correction mechanism can result in violation of the proportionality rule (5) as an unavoidable side effect. In such cases modification of the model is more serious than the initial assumptions allow but this must be tolerated if the goal repair frequencies requested by the end user are to be achieved.

3.3. Numerical Methods Used for Probability Estimation

With all the assumptions regarding scaling factors and their influence on repair frequencies, generation of a new model is now reduced to the problem of solving R equations in the form of (8). This can be accomplished with one of the standard numerical algorithms for finding roots of a non-linear function. The method described in this paper has been tested with implementation of the following three algorithms: the Newton method working on linear approximation of $F_i^r()$, the secant method and the false position (*falsi*) method.

3.3.1. Newton method on Linear Approximation (NOLA)

In this solution it is assumed that $F_i^r()$ is a linear function defined by points $F_i^r(X_i)$ (obtained after solving the model in step 1°) and $F_i^r(0)$ (which can be assumed to be equal zero). Then simply

$$X_{i+1}^r = F_G^r \,/\, F_i^r\,. \tag{11}$$

Noteworthy advantage of this approach lies in the fact that no other solution than the current frequency $F_i^r(X_i)$ is required to compute the next approximation, so errors of previous steps do not accumulate and convergence is good from the first iteration (the method has no memory effects).

Table 1 and Figure 3 present the details of exemplary model adjustment with this method implemented. The sample model consisted of three deterioration states and three types of repairs: minor, medium and major ($K = R = 3$). The values of goal frequencies has been selected to be ≈ 50% of that in the original repair policy ($\mathbf{F}_G = ½\mathbf{F}_0$) which corresponds to some hypothetical repair policy "what if repair frequencies of this piece of equipment are reduced by half".

Table 1. Sample model adjusted to $\mathbf{F}_G = ½\mathbf{F}_0$ with the NOLA method, $\varepsilon = 1E-4$

i	**Relative freq. of repairs** F_i^r / F_G^r			**Error**	**Scaling factors** X_{i+1}^r		
	***r* = 1**	**2**	**3**		***r* = 1**	**2**	**3**
0	2.007762	2.000581	2.049800	1.049814	0.49810	0.49990	0.48780
1	1.127838	1.230000	1.254300	0.254331	0.88670	0.81300	0.79720
2	1.018676	1.039839	1.044400	0.044398	0.98170	0.96170	0.95750
3	1.002895	1.006516	1.007300	0.007255	0.99710	0.99350	0.99280
4	1.000467	1.001065	1.001200	0.001179	0.99950	0.99900	0.99880
5	1.000076	1.000161	1.000200	0.000183	0.99990	0.99980	0.99980
6	1.000010	1.000032	1.000000	0.000040	1.00000	1.00000	1.00000

As it can be seen from the submitted data, the adjustment process goes smoothly and without any perturbations: the average convergence ratio is nearly constant and the model firmly approaches the required goals for all three frequencies. As the goal vector has been assumed to be 50% of the frequencies in the original (initial) model, the relative error in the first iteration equals to approx. 100% (the frequencies are twice as large as required) and in each subsequent iteration it is reduced by a factor of 4 to even 6. Finally, the imposed accuracy of 1E-4 (0.01%) is reached after 6 steps. It should be noted that such high precision in model tuning has been selected for illustrative purposes in this paper, while in practical engineering cases accuracy of 1% is more than adequate; in the discussed example such level is obtained after just three steps.

3.3.2. The Secant Method

In this standard technique the function is approximated by the secant defined by the last two approximations computed for points X_{i-1}^r, X_i^r; thus the new solution is calculated as:

$$X_{i+1}^r = X_i^r - \frac{X_i^r - X_{i-1}^r}{F_i^r - F_{i-1}^r}\left(F_i^r - F_G^r\right) . \tag{12}$$

After that X_{i-1}^r is discarded and X_{i+1}^r and X_i^r are considered as the pair defining the secant for the next iteration.

To begin the procedure two initial points are needed. In this approach we propose to choose the first point equal to the initial frequency of the model M_0 (X_0^r = 1), while the second point is computed as in the NOLA method: $X_1^r = F_G^r / F_0^r$. Having the pairs X_0^r, X_1^r (r = 1, 2, … R) the procedure starts according to equation (12) in iteration i = 1 and continues so in further steps.

In our exemplary case this method produced not-so constant convergence rate compared to the NOLA method presented in the previous point. Looking at the error values listed in table 2 for every iteration, there are some steps with very good improvement (e.g. error reduction from -0.1 to 0.008 for i = 3) but those are followed by mediocre progress in the next iteration (reduction from 0.008 to 0.006 for i = 4). Nevertheless, this does not exclude the method from application in the discussed system: the required accuracy is achieved in just one additional step (i = 7 compared to 6 for the NOLA method) and this is still an acceptable result.

Table 2. The same adjustment case ($\mathbf{F}_G = ½\mathbf{F}_0$, ε = 1E-4) controlled by the secant method

i	Relative freq. of repairs F_i^r / F_G^r			Error	Scaling factors X_{i+1}^r		
	r = 1	2	3		*r* = 1	2	3
0	2.007762	2.000581	2.049800	1.049814	0.49810	0.49990	0.48780
1	1.127838	1.230000	1.254300	0.254331	0.85360	0.70140	0.66440
2	0.990467	0.918355	0.893500	-0.106504	1.01190	1.11150	1.14910
3	0.995514	1.004032	1.007800	0.007832	1.01050	0.99530	0.99110
4	1.006038	0.999935	0.999600	0.006037	0.99410	1.00010	1.00040
5	1.000219	1.000161	1.000200	0.000218	0.99980	0.99990	0.99990
6	1.000000	1.000129	1.000100	0.000120	1.00000	0.99980	0.99990
7	1.000010	0.999903	1.000000	-0.000084	1.00000	1.00010	1.00000

The graphs in Figure 3 provide an additional explanation of the problem for this particular case. It can be seen that for i = 3 the method proposed actually a very good approximation but with just a little overshoot and this caused problems in the next iteration. This is a known setback of the secant method which is removed in the *falsi* method.

3.3.3. The False Position (Falsi) Method

In this approach X_{i+1}^r is computed as in (12) but the difference lies in choosing points for the next iteration. While in the secant method always X_{i-1}^r is dropped, now X_{i+1}^r is paired with that one of X_i^r or X_{i-1}^r which lies on the opposite side of the root. In this way when (12) is applied the solution is bracketed between X_i^r and X_{i-1}^r (which is the essence of the *falsi* method).

As in 3.3.2, the two initial points are needed but now they must lie on both sides of the root, i.e.

$$\left(F_0^r - F_G^r\right)\cdot\left(F_1^r - F_G^r\right) < 0 \tag{13}$$

Choosing such points may pose some difficulty. To avoid multiple sampling, as in the secant method it is proposed to select X_0^r=1 and to compute X_1^r like in NOLA method, but now with some "overshoot" that would guarantee (13):

$$X_1^r = \left(F_G^r / F_0^r\right)^{\alpha} \tag{14}$$

with a new parameter $\alpha > 1$ controlling the overshoot effect. The overshot must be sufficient to ensure condition (13) but, on the other hand, it should not produce too much of an error as this would deteriorate approximation process during initial steps and would produce extra iterations. If (13) is not met by initial value of X_1^r (14) can be re-applied with an increased value of α, although it should be noted that each such correction requires solving a new M_1 model and in effect this is the extra cost almost equal to that of the whole iteration.

Table 3. The sample case ($\mathbf{F}_G = ½\mathbf{F}_0$, $\varepsilon = 1E\text{-}4$) adjusted by the *falsi* method

i	**Relative freq. of repairs** F_i^r / F_G^r			**Error**	**Scaling factors** X_{i+1}^r		
	***r* = 1**	**2**	**3**		***r* = 1**	**2**	**3**
0	2.007762	2.000581	2.049800	1.049814	0.24810	0.24990	0.23800
1	0.591533	0.686129	0.691200	-0.408464	1.87420	1.71690	1.72780
2	1.067305	1.086871	1.092200	0.092239	0.93400	0.90950	0.90310
3	1.004705	1.005032	1.004800	0.005045	0.99510	0.99430	0.99450
4	1.000238	1.000258	1.000300	0.000313	0.99980	0.99970	0.99960
5	1.000010	1.000000	1.000000	0.000014	1.00000	1.00000	1.00000

Theoretical characteristics of the *falsi* technique are well illustrated by the data shown in Table 3 and Figure 3. Of all three methods presented here, this one generated the requested accuracy with the minimum number of iterations: 6 vs. 7 of NOLA and 8 of secant. The α parameter that controls the overshoot effect in the first approximation was equal 2 and this indeed generated an overestimation in the model M_1 but this was properly and quickly compensated in the steps that followed. As a result, the convergence flow which can be seen in Figure 3 is close to the ideal case with two deviations from the target: an overestimation immediately followed by an underestimation after which the solution arrives at the final goal.

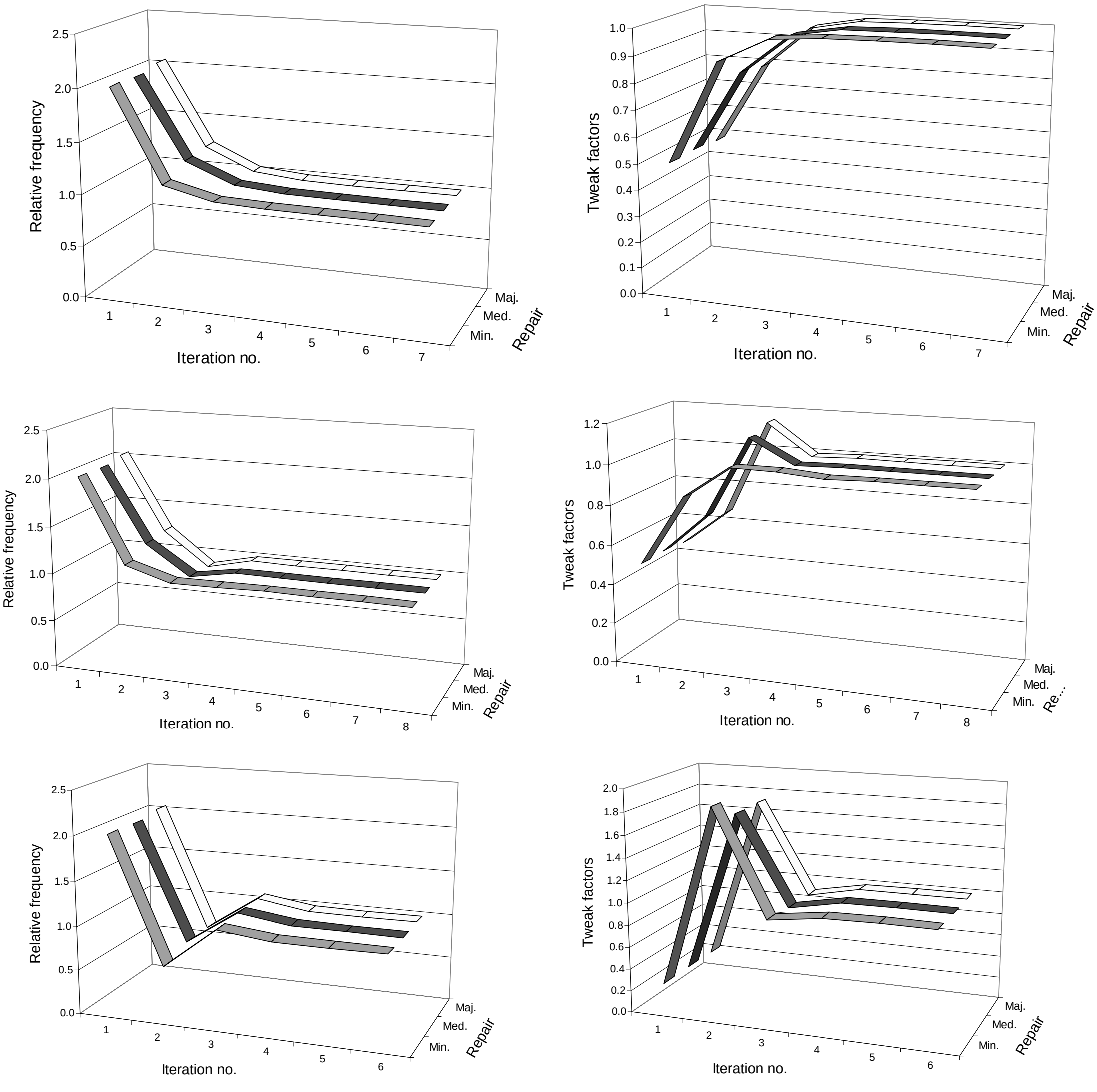


Figure 3. Adjusting the sample model to $\mathbf{F}_G = ½\mathbf{F}_0$ ($\varepsilon = 1E\text{-}4$) with three different numerical methods: NOLA (top), secant (middle) and *falsi* (bottom)

3.3.4. Evaluation of the Methods

Figure 4 illustrates efficiency of the three approximation methods in another typical adjustment cases. For these tests the model used in the previous examples was adjusted to the policies $\mathbf{F}_G = [0, 0, F_0^3]$ (only major repair is performed with the other two types removed) and $\mathbf{F}_G = ¼\mathbf{F}_0$ (frequencies of all repairs reduced to 25%). A short comment should be made about the first of these examples. With $\mathbf{F}_G = [0, 0, F_0^3]$, it is very easy to remove any repair from the model by simply assigning $P^{sr} = 0$ in all the states, effectively increasing the value of P^{s0} at the same time. In such case no adjustment is required as the requested goal ($F^r = 0$) is achieved immediately. Nevertheless, such modification does affect the frequencies of other repairs that are left in the model (F^3 in this case) and the adjustment mechanism is still necessary just to return them to their requested values.

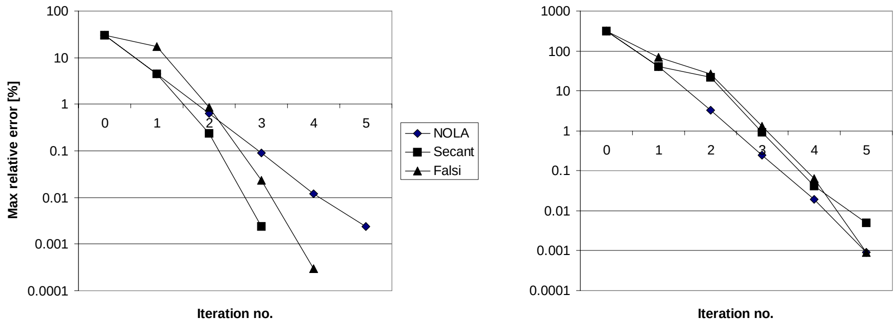


Figure 4. Convergence rate of the three approximation methods for two sample maintenance policies: $\mathbf{F}_G = [0, 0, F_0^3]$ (left) and $\mathbf{F}_G = ¼\mathbf{F}_0$ (right); $\varepsilon = 0.01\%$ in both cases

Comparing the effectiveness of the methods it should be noted that although simplifications of the NOLA solution may seem critical, in practice it works quite well. As it was noted before, due to its simplicity this method has one advantage over its more sophisticated rivals: since computation of the next solution does not depend on previous approximations, selection of the starting point is not so important and the accuracy during the first iterations is often better than in the secant or *falsi* cases. For example, in the case used in 3.3.1 ÷ 3.3.3 the NOLA method reached accuracy of 4.4% already after 2 iterations, while for secant and falsi methods the error after two iterations was, respectively, 11% and 9.2%. Superiority of the latter methods, especially of the *falsi* algorithm, manifests itself in the later stages of the process when the potential problems with an initial selection of the starting points have been diminished.

4. Asset Risk Manager

In this section we will present the "Asset Risk Manager" (ARM) software package which uses the concept of a life curve and discounted cost to study the effect of equipment ageing under different hypothetical maintenance strategies [1]. The results generated by the program are based on Markov models that were presented in the previous sections and in many cases the models were adjusted to the modified maintenance policies as it has been described in section 3.

For the program to generate automatically the life curves, default Markov model for the equipment has to be built and stored in the computer database. This is done through the prior running of the AMP program by an expert user. Therefore, both AMP and ARM programs are closely related, and usually, should be run consecutively.

Implementation details of Markov models, tuning its parameters and all other internal particulars should not be visible to the non-expert end user. All final results are visualized either through an easy-to-comprehend idea of a life curve or through other well-known concepts of financial analysis. Still, prior to running the analysis some expert involvement is needed, largely in preparation, importing and adjusting AMP models.

4.1. User input

A typical study is described through a comprehensive set of parameters that are supplied by a non-expert end user. They fall into three broad categories.

(A) General data. The Markov model of the equipment in question and its current state of deterioration form the primary information that is the starting point to most of ARM computations. The Markov model represents the equipment with present maintenance policy and is selected from a database of imported AMP models which need to be prepared by an expert in advance. Deterioration state, referred to as "Asset Condition" (AC) throughout the ARM, must be supplied by the end user as a percentage of "as-new" condition. Besides, a number of additional general parameters need to be specified, such as the time horizon over which the analysis will be performed, discount and inflation rates for financial calculations etc.

(B) Description of the present maintenance policy. It is assumed that three types of maintenance repairs can be performed: minor, medium and major. These correspond to the appropriate states in Markov model and not all of them must be actually present in the policy. For each repair user supplies its basic attributes, e.g. cost, duration and frequency.

(C) List of alternative actions. These are the hypothetical maintenance policies that decision-maker can chose from. Each action is defined as one of four types:

- continue as before (i.e. do not change the present policy),
- do nothing (i.e. stop all the repairs),
- refurbish,
- replace.

Apart from the first type, every action can be delayed for a defined amount of time. Additionally, for “non-empty” actions (i.e. any of the last two types) user must specify what to do in the period after action; the choices are:

(a) to change type of equipment and / or
(b) to change maintenance policy.

In case of option (b), the new maintenance policy will require a different Markov model which, in turn, may need to be generated by the adjustment algorithm discussed in section 3. Such generation is done automatically on the background of the system operation and the end user does not even know about this fact.

For every action the user must also specify what to do in case of failure: whether to repair or replace the failed equipment, its condition afterwards, cost of this operation etc. Thank to these options a broad range of maintenance situations can be described and then analysed.

The first action in the list is always “Continue as before” and this is the base of reference for all the others. The ARM can be directed to compute life curves, cost curves, or probabilities of failure – for each action independently – and then to visualize computed data in many graphical forms to assist the decision-maker in effective action assessment.

It should be noted that while the need for some action (e.g., overhaul or change in maintenance policy) is identified at the present moment, the actual implementation will usually take place only after a certain delay during which the original maintenance policy is in effect. Using ARM it is possible to analyze the effect of that delay on the cost and reliability parameters.

4.2. Life Curves

As it has been pointed out before, computing the average first passage time (FPT) from the first deterioration state (D1) to the failure state (F) in the Markov model yields the average lifetime of the equipment, i.e. length of its life curve. On the other hand, solving the model for the state probabilities of all consecutive deterioration states makes possible computing the state durations, which in turns determine the shape of the curve. Simple life curves obtained for different maintenance policies are later combined in constructing composite life curves which describe various maintenance scenarios.

For sake of simplicity and consistency, exactly three deterioration states, or levels, are always presented to the end user: minor, medium and major, with adjustable AC ranges. In case of Markov models which have more than three D*s* states, the expert decides how to assign the Markov states to three levels when importing the model.

Figure 5 shows exemplary life curves computed by ARM for typical maintenance situations. In each case the action is delayed for 3 time units (months, for example) and the analysis is performed for a time horizon of 10 time units. In case of failure seen in “Do nothing” action, equipment is repaired and its condition is restored to 85%.

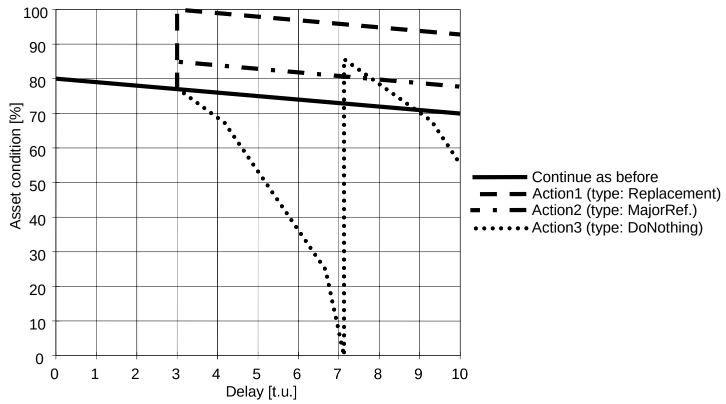


Figure 5. Life curves computed for three different actions (“Action1” … “Action3”) and compared to the present maintenance policy (“Continue as before”)

4.3. Probability of Failure

For a specific action, probability of failure within the time horizon (PoF_{TH}) is a sum of two probabilities: of failure taking place before (PoF_B) and after (PoF_A) the moment of action. It is assumed that failures in these two periods making up the time horizon are independent, so

$$PoF_{TH} = PoF_B + PoF_A - PoF_B \cdot PoF_A \,. \tag{15}$$

To compute $PoF(\,T\,)$ within some time period T, the Markov model for the equipment and the life curve are required. Then, the procedure for computation of PoF_A and PoF_B can be applied as it has been described in section 2.2, equation 1.

For better visualization, rather than finding a single PoF_{TH} value for action defined by the user in input parameters, ARM computes a curve which shows the PoF_{TH} as a function of action delay varying in a range 0 ÷ 200% of user-specified initial value. An example is demonstrated in Figure 6 for "Do nothing" action (user-defined delay = 3 time units), where also the two probability components PoF_B and PoF_A are shown.

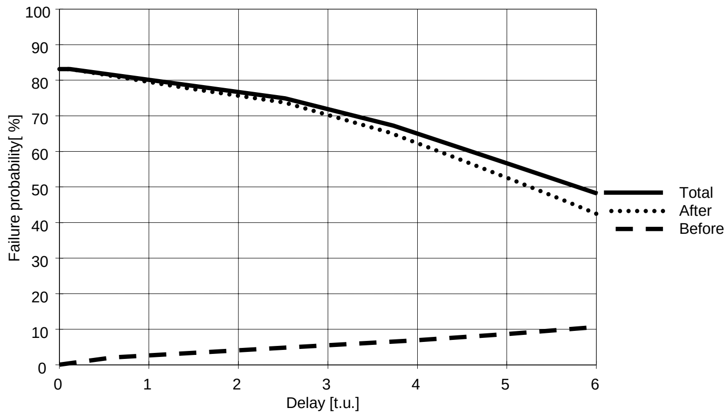


Figure 6. Probability of equipment failure within the time horizon for a modelled action, computed as a function of action delay

A more detailed analysis is shown in Figure 7. In this case we consider a situation when, with initial equipment deterioration estimated as 80% of "as-new" condition, some specific actions – a repair or just a change in maintenance policy – will take place after a 3-year delay while the effects will be evaluated for a 10-year time period. The actions in the scenarios will be as follows:

- adopting "do nothing" policy, which means just stopping all inspections and repairs; in case of failure the equipment will be repaired and its condition restored to 85%,
- replacing the equipment with "as-new" one and then switching to the "do nothing" policy,
- performing a major refurbishment of the equipment which restores its condition to 85% and then continuing with a medium repair only.

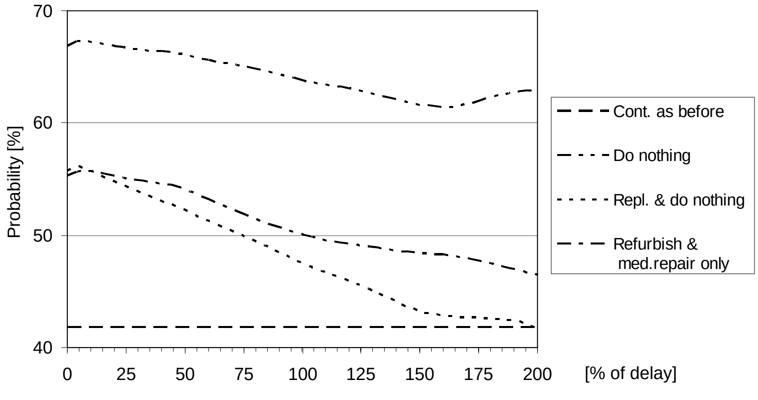


Figure 7. Probability of equipment failure within a period of 10 years for different cases of maintenance scenarios

Probability of failure within the time horizon computed for these strategies is shown in Figure 7. Values on the graphs are presented as functions of the action delay time (100% = 3 years) and they are compared against the probability of failure for the unmodified standard maintenance ("continue as before"). The value of this probability has been computed to be 42%.

It can be seen in case of all three scenarios that, since the new maintenance policy after the action is more or less reduced, the more the action is delayed, the less probable equipment failure becomes. For evident reasons adopting "do nothing" policy leads to the highest values of the failure probability, while replacing the equipment and "doing nothing" afterwards turned out to be a less dangerous strategy (in terms of failure probability) than refurbishing and then keeping only the medium repair. Whether the differences in the economic expenses of these two possible strategies justify this discrepancy in the reliability parameter or not – it remains an open question in further cost analysis and generally depends on the costs associated with the equipment failures.

One interesting observation can be made about the curve for "do nothing" strategy: its decrease is not strictly monotonic and there is a local minimum at the level of 61% for the delay equal to 164% (4.9 years) after which the probability begins to rise slowly. To explain this rise, two components: the probability of failure before and after the action should be investigated and they are shown in Figure 8. In general, these two components behave as expected: the later the action takes place, the higher the probability of failure before and the lower the probability of failure after the action but the rates of these two flows – increasing and decreasing – are not constant and do not sum up into a monotonic decrease. In this case, the probability of failure after the action falls down to some extent slower after the point of 164% and this causes the local minimum in the total probability of failure.

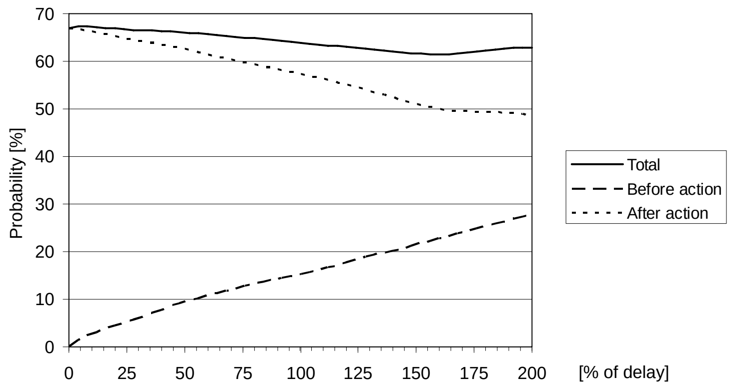


Figure 8. Probability of equipment failure before and after the action for "do nothing" scenario

4.4. Cost Curves

In financial evaluations the costs are expressed as present value (PV) quantities and this approach should also be used in this kind of studies because maintenance decisions on aging equipment include timing, and the time value of money is an important consideration in any decision analysis. The cost difference is often referred to as the Net Present Value (NPV). In the case of maintenance, the NPV can be obtained for several re-investment options which are compared to the "Continue as before" policy.

Cost evaluation for any maintenance scenario involves calculation of the following three fundamental classes of components:

1. Cost of the maintenance activities;
2. Cost of the selected action (i.e. refurbishment or replacement);
3. Cost associated with failures (cost of repairs, system cost, penalties).

To compute the PV, inflation and discount rates are required for the specified time horizon. The cost of maintenance over the time horizon is the sum of the maintenance costs incurred by the original maintenance policy for the duration of the delay period (up to the action), and the costs incurred by the new policy for the remainder of the time horizon (after the action). The costs associated with the equipment failure over the time horizon can be computed similarly except that the failure costs before and after the action should be multiplied by the respective probabilities of failures, and two products added.

Figure 9 presents the plots showing the cost analysis for the scenario "replace and then do nothing". Again (as it has been in the case of probability of failure) the values are visualized as functions of the action delay varying in the range 0 ÷ 200% of user-specified reference value. The cost of replacement ("Action"), although does not depend on the delay, is not constant on the plot due to the PV calculations. It is also evident that delaying the action causes more repairs to be performed as elements of the present repair policy before "do nothing" becomes effective, hence several noticeable leaps appear in the maintenance cost flow.

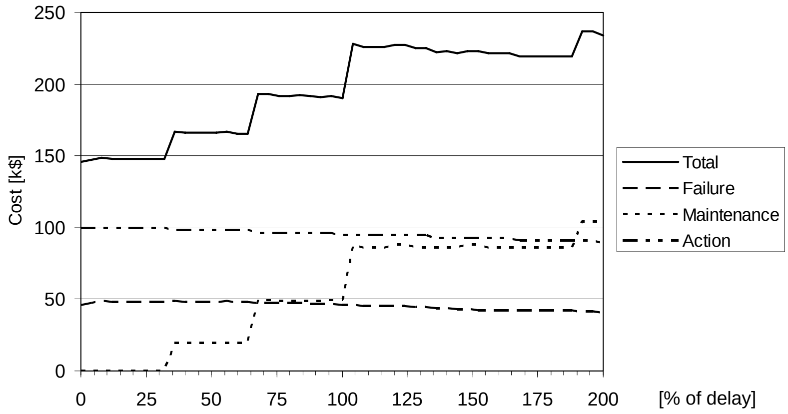


Figure 9. Estimated cost of "replace & do nothing" scenario (total value and the three components)

5. Conclusions

The paper described a modelling methodology which helps in choosing effective yet cost-efficient maintenance policy. Based on semi-Markov models representing the deterioration process, the equipment life curve and other reliability parameters can be evaluated. Once a database of equipment models is prepared, the end-user can perform various studies about different maintenance strategies and compare expected outcomes. As the results are visualized through the relatively simple concept of a life curve, no detailed expert knowledge about internal reliability parameters or configuration is required.

Additionally, the paper presented a method of model adaptation that allows automatic adjustment of the basic model to user-expected changes in maintenance policy. The numerical part of the method can be solved with common root-finding algorithms and it has proved its validity in numerous practical examples. Ability to adjust the models to such changes is crucial in practical studies of various possible maintenance scenarios.

Another important issue is how the adjustment modifies behaviour of the model in addition to reaching the desired repair frequencies and how the model should be constructed in order to accommodate the modifications without undesired side effects. For discussion of these problems please refer to [5–7]. In particular, while there is usually no problem with *reduction* of repair frequencies with the method described in this paper, special care must be taken when *increase* is requested because the probabilities cannot be enlarged indefinitely. In cases when the limit (2) is reached in all states even with $P^{s0} = 0$ and the requested frequencies are still not achieved, more substantial modifications of the model may be necessary. This also leads to specific directions as to how the model should be constructed in order to avoid such situations.

References

1. Anders, G. J., Sugier, J. Risk assessment tool for maintenance selection. In: *Proc. Int. Conf. Dependability of Computer Systems DepCoS – RELCOMEX 2006*. IEEE Comp. Soc. Press, 2006, pp. 306–313.
2. Endrenyi, J., Aboresheid, S., Allan, R. N., Anders, G. J., Asgarpoor, S., Billinton, R., Chowdhury, N., Dialynas, E. N., Fipper, M., Fletcher, R. H., Grigg, C., McCalley, J., Meliopoulos, S., Mielnik, T. C., Nitu, P., Rau, N., Reppen, N. D., Salvaderi, L., Schneider, A., Singh, Ch. The Present Status of Maintenance Strategies and the Impact of Maintenance on Reliability, *IEEE Trans. Power Systems*, Vol. 16, 2001, pp. 638–646.

3. Zhang, T., Nakamura, M., Hatazaki, H. A decision methodology for maintenance interval of equipment by ordering based on element reparation-replacement rate. In: *Proc. Power Engineering Society Summer Meeting*. IEEE, 2002. Vol. 2, pp. 969–974.
4. Sugier, J., Anders, G. J. Modeling changes in maintenance activities through fine-tuning Markov models of ageing equipment. In: *Proc. Int. Conf. Dependability of Computer Systems DepCoS – RELCOMEX 2007*. IEEE Comp. Soc. Press, 2007, pp. 336–343.
5. Sugier, J., Anders, G. J. Verification of Markov models of ageing power equipment. In: *Proc. Int. Conf. Probabilistic Methods Applied to Power Systems PMAPS'2008*. Rincon, Puerto Rico, 2008, pp. 1–6.
6. Sugier, J., Anders, G. J. Modifying Markov models of ageing equipment for modeling changes in maintenance policies. In: *Proc. Int. Conf. Dependability of Computer Systems DepCoS – RELCOMEX 2009*. IEEE Comp. Soc. Press, 2009, pp. 348–355.
7. Sugier, J. Avoiding probability saturation during adjustment of Markov models of ageing equipment. In: *Proc. Int. Conf. Dependability of Computer Systems DepCoS – RELCOMEX 2011*. Berlin Heidelberg: Springer-Verlag, 2011, pp. 205–218.
8. Sugier, J., Anders, G. J. Probabilistic evaluation of deterioration processes with maintenance activities, *Journal of Polish Safety and Reliability Association*, Vol. 1, 2011, pp. 177–184. (Summer Safety and Reliability Seminars SSARS 2011, Polish Safety and Reliability Association, Gdynia)
9. Anders, G. J., Endrenyi, J. Using Life Curves in the Management of Equipment Maintenance. In: *Proc. Int. Conf. Probabilistic Methods Applied to Power Systems PMAPS'2004 Conference*. Ames, Iowa.
10. Hughes, D. T., Russell, D. S. Condition Based Risk Management (CBRM), a Vital Step in Investment Planning for Asset Replacement. In: *Proceedings 3rd IEE International Conference on Reliability of Transmission and Distribution Networks (RTDN 2005). London, UK, 15–17 Feb. 2005*, pp. 261–265.
11. Chan, G. K., Asgarpoor, S. Preventive Maintenance with Markov Processes. In: *Proc. 2001 North American Power Symposium. College Station, TX, October 2001,* pp. 510–515.
12. Endrenyi, J., Anders, G. J., Leite da Silva, A. M. Probabilistic Evaluation of the Effect of Maintenance on Reliability – An Application. *IEEE Transactions on Power Systems*, Vol. 13, 1998, pp. 567–583.
13. Hosseini, M. M., Kerr, R. M., Randall, R. B. An inspection model with minimal and major maintenance for a system with deterioration and Poisson failures, *IEEE Trans. on Reliability*, Vol. 49, 2000, pp. 88–98.
14. Limnios, N., Oprisan, G. *Semi-Markov Processes and Reliability*. Boston, MA,: Birkhauser, 2001.
15. Perman, M., Senegacnik, A., Tuma, M. Semi-Markov Models with an Application to Power-Plant Reliability Analysis, *IEEE Transactions on Reliability*, Vol. 46, 1997, pp. 526–532.
16. Yin, L., Fricks, R. M., Trivedi, K. S. Application of Semi-Markov Process and CTMC to Evaluation of UPS System Availability. In: *Proc. 2002 Annual Reliability and Maintainability Symposium*, pp. 584–591.
17. Anders, G. J., Leite da Silva, A. M. Cost Related Reliability Measures for Power System Equipment, *IEEE Transactions On Power Systems*, Vol. 15, 2000, pp. 654–660.
18. Anders, G. J., Maciejewski, H. Estimation of impact of maintenance policies on equipment risk of failure. In: *Proc. Int. Conf. Dependability of Computer Systems DepCoS – RELCOMEX 2006*. IEEE Comp. Soc. Press, 2006, pp. 351–357.

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 21–29
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

HEURISTIC PREFERENCE RULES IN MULTIROUTE SCHEDULING

A. Ben-Yair[1], D. Golenko-Ginzburg[2]

[1]Center for Reliability and Risk Management
SCE - Shamoon College of Engineering
Beer-Sheva, 84100, Israel
E-mail: avner2740@gmail.com

[2]Department of Industrial Engineering and Management,
Ben-Gurion University of the Negev, Beer-Sheva, 84105, Israel, and
Department of Industrial Engineering and Management
Ariel University Center (AUC) of Samaria, 44837, Israel
E-mail: dimitri@bgu.ac.il

A systematic account and comparative analysis of the efficiency of various statistical decision-making rules in scheduling theory, is undertaken. We will describe various adaptive randomisation preference rules in order to use the latter in industrial scheduling.

Keywords: *Heuristic preference rules; mn-permutations; SPT, LRT, FIFO and PC rules; Randomization preference rules*

1. Introduction

The area of our research centres on developing deterministic and statistical algorithms in job-shop scheduling, i.e., in determining quasi-optimal schedules for multiroute (job-shop) problems with n jobs and m machines [1–17].

The procedures described below for taking decisions in conflict situations (jobs queued up for a single machine) are, in essence, the simplest simulation models with which a quasi-optimal plan can be drawn up for feeding the jobs to the machine for a varied-route problem with n jobs and m machines. Thus, the paper under consideration is a further development of Refs [11–15]. We will henceforth use the terminology of [11].

Suppose that as a result of such a conflict situation at a certain moment of time t there is a queue of $r \le n$ jobs waiting to be fed into machine M. Denote the consecutive numbers of these jobs $q_1, q_2, ..., q_r$.

To devise a definite rule of preference means to work out such an algorithm as will make it possible to determine a numerical characteristic for each of the jobs in the queue J_{q_ξ}, $\xi = \overline{1,r}$. The numerical characteristic is the preference coefficient

$$Q\left\{J_{q_\xi}\right\} = \varphi\left\{T_t, T, t, q_\xi\right\}, \tag{1}$$

where symbol T_t is the initial technological matrix T transformed by moment t (after excluding the operations already done by moment t). Job J_{q_ξ}, which corresponds to optimal value $Q\left\{J_{q_\xi}\right\}$, is selected from the queue and is fed first into the machine.

Depending on the random effect, the preference rules are subdivided into deterministic and random ones. The randomising procedure of preference rule Q centres on the following: first, the preference coefficients of the jobs in the queue $Q\left\{J_{q_\xi}\right\}$, $\xi = \overline{1,r}$, are normalized by

$$Q^*\left\{J_{q\xi}\right\} = \frac{Q\left\{J_{q\xi}\right\}}{\sum\limits_{\eta=1}^{r} Q\left\{J_{q\eta}\right\}}, \tag{2}$$

after which random value γ is simulated by the Monte Carlo method and equi-distributed in interval $[0;1]$. Job $J_{q\xi}$ is fed into the machine, there the relation exists

$$\sum_{\ell=1}^{\xi-1} Q^*\left\{J_{q\ell}\right\} < \gamma \leq \sum_{\ell=1}^{\xi} Q\left\{J_{q\ell}\right\} \tag{3}$$

(sum $\sum\limits_{1}^{0}$ is taken as equal to zero).

The physical idea of randomisation in preference rules consists in the fact that of the jobs demanding to be fed into the machine, the ones fed in usually have a larger preference coefficient. Note that if the use of deterministic preference rules always leads to the drawing up of a determined schedule (the assignment matrix can be determined by rule Q back in the stage preceding the start of the system's work by a preliminary simulation of the system's activity), in the case of randomisation principles, values S_{ij} are random in nature. As a result, in the deterministic version one can investigate the relative efficiency of two alternate preference rules Q_1 and Q_2 by comparing two determined values; the preference rule leading to the value of optimisation criterion K for a concrete technological matrix T is preferable. In a randomisation case, the preference of one rule or another should be decided by comparing statistical samplings (from the general totality of random values of criterion K), and applying the theory of checking statistical hypotheses.

We thus come to the statement of the basic shortcoming of preference rules: a comparative analysis of their efficiency can be made only for a concrete technological matrix T, and in turn, conclusions are to be drawn if T changes can also change. In fact, even if given a large number of existing preference rules proved in practice, one can always devise technological matrix T, for which we can draw up a new preference rule, more efficient than all the existing ones.

Let us now describe the best known and most widely applied preference rules.

2. Non-Combined (Local) Preference Rules

2.1. The Shortest Processing Time (SPT) Rule [16]

What this rule means is that when resolving a conflict situation, preferred is the job in set $Q\left\{J_{q\xi}\right\}$, $\xi = \overline{1,r}$, waiting to be fed into released machine M, which has the shortest operation duration. In other words,

$$Q\left\{J_{q\xi}\right\} = \frac{1}{t_{q\xi j\xi}}, \tag{4}$$

($t_{q\xi j\xi}$ being the duration of doing the $j\xi$-th operation of job $J_{q\xi}$ on machine M).

2.2. The Longest Remaining Time (LRT) Rule [16]

Under the LRT rule, job $J_{q\xi}$, for which the total duration of all remaining operations is the maximum,

$$Q\left\{J_{q\xi}\right\} = \sum_{k=j}^{m} t_{q\xi k}, \tag{5}$$

is fed into machine M first.

2.3. The Rule TSW: SPT, LRT, and "Tense Job" Expectation Rules [13]

Of the many decisions that can be taken at any moment t, the one is taken that corresponds to the minimal change of the lower estimate of the criterion. This estimate is based on the sum of the time for performing the remaining operations. The approximation of the method consists in that the criterion estimate forecasted after each decision is regarded as the true value of the criterion, as a result of which, movement along only one branch of the decision tree takes place and the other branches are not memorized. Let us deal in more detail with realization of the preference rule.

At any moment t of freeing machine M, q "tense" job appears whose total time for performing the remaining operations (the time left for finishing an operation begun is not included in the total) is the maximum. The schedule for doing the remaining operations of the tense job depends strongly on the value of the criterion. If the tense job is in queue for machine S, it is fed in at moment t. If it has not yet been placed in the queue, compute the remaining time t_S^* for doing the tense job's operations till the one to be done on machine M, and compare it with time $t_S^{**} = \min_{\xi} t_{q_\xi j_\xi}$, the briefest operation of the jobs in queue for machine M (the SPT rule). If $t_S^{**} > t_S^*$, the machine stands idle till moment $t + t_S^*$. If $t_S^{**} \le t_S^*$, machine M is fed by job J_{q_η} ($t_S^{**} = t_{q_\eta j_\eta}$). Similarly, rule SPT can be substituted by rule LRT.

2.4. The Rule of Pairwise Comparison PC [10, 13]

For any two jobs J_i and J_k requiring to be operated on freed machine M, denote symbols O_{ij} and $O_{k\ell}$ for the routine operations, and symbols $O_{i,j+1},...,O_{im},O_{k,\ell+1},...,O_{km}$ for the remaining operations of jobs J_i and J_k, respectively.

Let $t_{ij}, t_{k\ell}$ stand for the time to carry out operations $O_{ij}, O_{k\ell}$ and $\tau_{ij}, \tau_{k\ell}$ stand for the time to carry out all the remaining operations of jobs J_i and J_k, respectively. There exist two possible ways of feeding in the jobs: 1) First J_i, then J_k; 2) First J_k, then J_i. For the first way, we must calculate value $F(I) = t_{ij} + \max[\tau_{ij}, t_{k\ell} + \tau_{k\ell}]$, and for the second, $F(II) = t_{k\ell} + \max[\tau_{k\ell}, t_{ij} + \tau_{ij}]$. Say the first is preferable if there holds condition $F(I) < F(II)$. By comparing pairwise all jobs requiring operation on machine M, we can discover the job with the greatest priority.

3. Combined Preference Rules

3.1. The Rule of SPT and Feeding in Tense Jobs (Rule TS) [13–14]

A tense job is assigned to a freed machine if it must be done by moment t. Otherwise the next job is selected by the SPT rule. In keeping with the characteristic of rule (3), if at moment of time t jobs $J_{i_1},...,J_{i_r}$ have completed operations $O_{i_1 j_1},...,O_{i_r j_r}$ and have not yet begun their next operation, while jobs $J_{i_{i+1}},...,J_{i_n}$ are still performing operations $O_{i_{r+1} j_{r+1}},...,O_{i_n j_n}$, job J_{i_ξ} is considered a tense one if $\tau_{i_\xi}(t) = \max_{1 \le k \le n} \tau_{i_k}(t)$, $\tau_{i_k}(t) = \sum_{\gamma = j+1}^{m} t_{i_k \gamma}$ holds.

3.2. The Rule of PC and Starting a Tense Job (Rule TPC) [13–14]

A tense job is assigned first to a freed machine if it must be carried out. Otherwise, the job is chosen by rule PC.

3.3. The FIFO Rule [16]

The FIFO rule, in case a conflict situation appears, directs onto a machine the job from the queue waiting longest to be done, and consequently requiring to be done on that machine ahead of others. In other words, this is the principle of first come, first served.

Rules 4.1–4.3, following, employ as a parameter of the preference function the total time that the operations prior to the conflict situation waited in queues to be handled. Experimental calculations have shown that using such a parameter for several practical problems makes the preference function more efficient [14].

4. Total Time Heuristic Preference Algorithms

4.1. A Combination of LRT and Total Waiting Time [13–14]

Use symbol $T_t(J_i)$ for the total time of waiting in queues for job J_i at moment t. In other words, $T_t(J_i) = t - \sum_{j=1}^{j_i(t)} t_{ij}$, where $j_i(t)$ is the last operation of job J_i completed by moment t. Suppose that at moment t jobs $J_{i_1},...,J_{i_r}$ are to be done on one and the same machine. In such a case, the machine is fed job J_{i_ξ}, $1 \le \xi \le r$, for which the sum $T_t\left(J_{i_\xi}\right) + LRT\left[J_{i_\xi}\right]$ is maximized.

4.2. A Combination of FIFO, Total Waiting Time, and LRT [13, 14]

If at moment t jobs $J_{i_1},...,J_{i_r}$ are to be done on freed machine M, job J_{j_ξ}, first in the queue, is fed in. If there are several jobs of that kind $J_{j_1},...,J_{j_\ell}$, job J_{j_ξ}, $\xi = \overline{1,\ell}$, is fed in, the sum $T_t\left(J_{j_\xi}\right) + LRT\left[J_{j_\xi}\right]$ being maximized for it.

4.3. A Combination of Waiting for a Tense Job, Total Waiting Time, and LRT [14]

If at moment t jobs $J_{i_1},...,J_{i_r}$ are waiting to be done on the machine simultaneously, the tense job is found and moment t^* is determined for it to be fed in. If $\Delta t = t^* - t$ is no less than the time of the forthcoming operation for a part in the queue of jobs $J_{j_1},...,J_{j_\ell}$, job J_{j_ξ} is chosen and fed in, according to rule 4.1. Otherwise, the machine is idle and there is a wait for the tense job.

4.4. A Combination of Waiting for a Tense Job and FIFO

The procedure of preference rule 10 remains unchanged, except that of the set of jobs $J_{j_1},...,J_{j_\ell}$, the job first in queue is fed into the machine. If there are several such jobs, apply rule 4.1.

All the preference rules described above can be applied both in deterministic and randomized versions.

5. Using *mn*-Permutations

Note that, due to [15], the *mn*-permutations described can be regarded as priority vectors, the operation being given the greater priority the farther left the corresponding number in the *mn*-permutation is situated. This makes it possible to simulate various priority rules.

We will demonstrate how the most widespread priority rules, SPT and LRT, can be obtained by such an approach.

Transform initial matrix T as follows: Say

$$\tau_{ij} = \sum_{k=j+1}^{m} t_{ik}\ ,\ \tau_{im} = t_{im}\ ,\ i = \overline{1,n}\ ,\ j = \overline{1,m-1}\ .$$

Write all the elements of matrix $\|\tau_{ij}\|$ one after the other, in a row, so that element τ_{ij} stands at the $[(j-1)m + j]$-th place, after which write the row in descending order. Based on the indexes obtained of the elements of that row, the *mn*-permutations are determined with the help of the auxiliary algorithm of an actual schedule, described below. The schedule obtained will coincide with the one obtained by rule LRT (with accuracy of an ambiguous selection of operations by the LRT rule itself, when several operations have priority under that rule).

The SPT rule can be simulated as follows: Form two *mn*-length vectors $A1$ and $A2$, filling their elements consecutively from left to right. To determine the consecutive element $A1(k)$, take the leftmost element not crossed out in every row of matrix T and choose the minimum among them, which will then be crossed out of the matrix. Let the element crossed out have index (i, j). Then assume $A1(k)$ equals t_{ij}, and $A2(k) = (i-1)m + j$. Then, on the basis of the auxiliary algorithm, devise an actual schedule, using vector $A2$ as the initial mn-permutation. The resulting schedule coincides with the one obtained by rule SPT (as in the case of LRT, to an ambiguity accuracy resulting from the SIO rule itself).

The work of the auxiliary algorithm is as follows. Assume that in keeping with the procedure described in [15], at the k-th step of determining a possible schedule based on an *mn*-permutation, operation O_{ij} is assigned, with completion time equal to t_{ij}, and to be done on the M-th machine. Let job J_i be ready at moment τ_i to do operation O_{ij}. Look from left to right, beginning with moment τ_i, at all the idle intervals of the M-th machine. As soon as we find an idle interval with a duration more than zero, assign the corresponding operation O_{ij}. The starting moments of several operations can move to the right by a value no greater than t_{ij}.

This is how the last can be performed: Let operation O_{ij} be fed in during a certain idle interval of the i-th machine, and after the idleness, operation $O_{i_1 j_1}$ is to be done, to which the element at the j-th place, $j < k$, in the *mn*-permutation corresponds. In this case, the k-th element of the *mn*-permutation takes the j-th place, and the elements numbered $j, j+1, ..., k-1$ move one place to the right. Then return to the $(j+1)$-th element of the mn-permutation, and with it continue assigning the next operations.

From what has been said, it follows that an individual preference rule Q, the most efficient of the existing "rules bank" $\{Q_i\}$ must correspond as well to each individual technological matrix T. We believe such a pseudo-optimal rule should be built on the principles of self-edification with the help of a balanced probability combination of rules $\{Q_i\}$. We must make use of methods of search for the optimal balanced combination $\{\delta_i\}$ to make possible full use of the information provided by initial technological matrix T.

Suppose that there exists a data bank of preference rules consisting of d rules Q_p, $p = \overline{1,d}$, and we know how long it takes to do all the jobs with the help of each rule.

In keeping with the results of [11–15] say that we apply rule $Q = \sum_{p=1}^{d} \delta_p Q_p$ to technological matrix T if at the moment when a routine conflict situation arises, rule Q_p is applied to the queue of jobs with probability δ_p. Algorithmically, this means that every time, a rule Q_ξ, $\xi = \overline{1,d}$, is chosen, such that for its ordinal number relation

$$\sum_{p=1}^{\xi-1} \delta_p < \eta \le \sum_{p=1}^{\xi} \delta_p\ , \tag{6}$$

holds, where η is the routine value of a random variable equi-distributed in interval $[0;1]$. Rule Q_ξ is applied (randomly or determined) to the queue of jobs according to the method described above.

Now we take up the development of an algorithm for forming an optimal set of weights $\{\delta_p\}$. The algorithm described below employs the procedure aimed at a random search, and consists of the following stages:

Stage I. Simulate d-dimensional vector $\vec{\Sigma}$ equi-distributed on a d-dimensional hyper-surface $\sum_{p=1}^{d} x_p = 1$, $0 \le x_p \le 1$, and denote the vector coordinates by symbols $\delta_1, \delta_2, ..., \delta_d$, [13].

Note that the order of enumerating coordinates δ_p in sequence $\{\delta_p\}$ does not change from now.

Stage II. Use the Monte Carlo method to simulate a random variable η, equi-distributed in interval $[0;1]$.

Stage III. Determine ordinal number $\xi \in \{\overline{1,d}\}$, such that relation (6) holds.

Stage IV. Draw up a calendar plan schedule for launching jobs on the machines in keeping with a periodical repetition of the Stages II$\Rightarrow$ III procedure and a set of values determined at Stage I for coordinates of vector $\vec{\Sigma}$ for each conflict situation. In other words, simulate the work of the system by deterministic or randomised rule $Q = \sum_{p=1}^{d} \delta_p Q_p$.

Stage V. Calculate the value of optimisation criterion K for a single realization of the model for drawing up a calendar schedule.

Stage VI. Repeat the procedure of Stages II-V a sufficiently large number of times (N times) to obtain representative statistics, and subsequent calculation of mathematical expectation $E[K]$ for simulated values $\delta_1, ..., \delta_d$. Denote the value obtained by symbol $E\{K[\delta_p Q_p]_T\}$, which signifies the estimate of the efficiency of the balanced preference rule $Q = \sum_{p=1}^{d} \delta_p Q_p$ for the initial technological matrix T.

Stage VII. Compare values $A = E\{K[\delta_p Q_p]\}$ and $B = \min_p \{K[Q_p]_T\}$. If inequality $A < B$ holds, it means we have determined (for initial matrix T) a better preference rule $Q_{d+1} = \sum_p \delta_p Q_p$ than those in the data bank, and we turn to Stage X of the algorithm. If $A \ge B$, go to the next stage.

Stage VIII. Counter η_1 of the number of consecutive unsuccessful tests at the initial point of search now functions. If counter $<\eta_1>$ shows no more than the limit number of unsuccessful tests $C^I_{\max}$, turn to Stage I. This means that the procedure of Stages I-VI is again repeated, in order to provide a better balance rule Q_{d+1}. If $<\eta_1> > C^I_{\max}$, transfer to the next stage of the algorithm.

Stage IX. For the initial point of a local search, take vector-point $\vec{\Sigma}^{(1)} = \delta_p^{(1)}$, where $\delta_p^{(1)} = 0$ when $p \ne \gamma$, and $\delta_p^{(1)} = 1$ when $p = \gamma$, $\gamma \in \{\overline{1,d}\}$, $E\{K[Q_\gamma]_T\} = B$. Further, turn to Stage XI, preliminarily sending the coordinate values of vector-point $\vec{\Sigma}^{(1)}$ and of $B = K_{\min}$ to a special array Z.

Stage X. Take vector-point $\vec{\Sigma}^{(1)} = \{\delta_p^{(1)}\}$ as the initial point of the local random search for an extremum, with $E\left\{K\left[\delta_p^{(1)} Q_p\right]_T\right\} = A$. The coordinates of vector-point $\vec{\Sigma}^{(1)}$ and value $A = K_{\min}$ are sent to array Z.

Stage XI. Make a local random search by

$$\vec{\Sigma}^{(2)} = \vec{\Sigma}^{(1)} + \vec{X}\,, \tag{7}$$

where $\vec{X}$ is a random d-dimensional vector-point, whose coordinates x_p satisfy relations $\sum_{p=1}^{d} x_p = 0$ and $\sum_{p=1}^{d} x_p^2 = D$, where D is a previously accepted length of the search step.

Stage XII. Correct vector-point $\vec{\Sigma}^{(2)}$ if it leads beyond the region of $0 \le \delta_p^{(2)} \le 1$, $p = \overline{1,d}$. There are several ways to effect that correction, in particular by reducing search step D (by halving, for instance) till restrictions $0 \le \delta_p^{(2)} \le 1$ are satisfied.

Take note of the fact that the search procedure can also be followed in other ways, including heuristic elements, for example, by relation

$$\delta_p^{(2)*} = \delta_p^{(1)} + D \cdot sign\,\alpha_p\,, \tag{8}$$

where α_p, $p = \overline{1,d}$, are independent values of a random variable equi-distributed in interval $[-1;+1]$, and subsequent correction

$$\delta_p^{(2)**} = \begin{cases} \delta_p^{(2)*} & when\ 0 \le \delta_p^{(2)*} \le 1 \\ 1 & when\ 1 < \delta_p^{(2)*} \\ 0 & when\ 0 > \delta_p^{(2)*} \end{cases}, \tag{9}$$

$$\delta_p^{(2)*} = \delta_p^{(2)**} \cdot \frac{1}{\sum_p \delta_p^{(2)**}}\,. \tag{10}$$

Stage XIII. Perform the procedure of Stages II-VI (N times), preliminarily clearing counter η_1 and subsequently calculating value $E\left\{K\left[\delta_p^{(2)} Q_p\right]_T\right\}$.

Stage XIV. Compare value $E\left\{K\left[\delta_p^{(2)} Q_p\right]_T\right\}$ with the contents of the cell in array Z, where the minimal value of criterion $K_{\min}$ is stored. If the new estimate is less than value $K_{\min}$, transfer to Stage XVI. If not, move on to the next stage.

Stage XV. Counter η_2 of the consecutive unsuccessful search steps now functions. If $<\eta_2>$ is less than limit number $C_{\max}^{II}$, transfer to Stage XI for a repeated realization of a random search step from the previous vector-point $\vec{\Sigma}^{(1)}$. Otherwise, turn to Stage XVIII.

Stage XVI. Clear counter η_2 and then compare values

$$\Delta_1 = \left| \frac{\left\{K\left[\delta_p^{(2)} Q_p\right]_T\right\} - K_{\min}}{K_{\min}} \right| \text{ and } \Delta_{\min}\,,$$

where $\Delta_{\min}$ is a pregiven relative error of the convergence of the optimisation criterion for a local random search. If $\Delta_1 \le \Delta_{\min}$, turn to Stage XIX; otherwise advance to Stage XVII.

Stage XVII. Send values of vector-point $\vec{\Sigma}^{(2)}$ and $E\left\{K\left[\delta_p^{(2)} Q_p\right]_T\right\}$ to array Z, subsequently transferring to Stage XI. In other words, the new vector-point $\vec{\Sigma}^{(2)}$ takes the place of the previous one $\vec{\Sigma}^{(2)} \Rightarrow \vec{\Sigma}^{(1)}$, and the random search is carried out from the new point. Value $E\left\{K\left[\delta_p^{(2)} Q_p\right]_T\right\}$ now becomes minimal value $K_{\min}$.

Stage XVIII. Choose and print the local optimum, an optimal set of balance coefficients $\left\{\delta_p^{(opt)}\right\}$ from array Z, and the corresponding value $K_{\min}$. After this, proceed to Stage XX.

Stage XIX. Choose and print the local optimum $\left\{\delta_p^{(2)}\right\}$ and the value of optimisation criterion $K_{\min} = E\left\{K\left[\delta_p^{(2)} Q_p\right]_T\right\}$. Further, turn to the next stage of the algorithm.

Stage XX. Counter η_3 of the determined local optimums works. If $<\eta_3>$ exceeds the limit number $C_{\max}^{III}$, proceed to Stage XXII. Otherwise, go to the next stage of the algorithm.

Stage XXI. Clear counters η_1 and η_2 and array Z. Further, return to Stage I.

Stage XXII. Choose an overall optimum, set $\left\{\delta_p^*\right\}$, furnishing a minimum of local optimums built on Stages XVIII–XIX, after which the work of the algorithm terminates.

Note that such an adaptive algorithm, as we see it, especially efficient in case it is necessary to perform operative control of the process of feeding jobs to machines at random durations of separate operations O_{ij}, i.e., at the stage of actual production. In such a case, at moment t for each conflict situation, jobs must be chosen for machines according to the optimal set $\left\{\delta_p^*\right\}$ determined at moment t. The durations t_{ij} of the remaining operations O_{ij} are equated to their mathematical expectations, and the reduced matrix T_t, including only the operations still to be done, is taken as the technological matrix T. If another conflict situation arises, we do the same, each time acting on decisions taken, i.e., working out the appropriate control action.

References

1. Adams, E. Balas and D. Zawack. The shifting bottleneck procedure for job-shop scheduling, *Mgmt. Sci*., Vol. 34, 1988. pp. 391–401.
2. Barker, J. R. and B. M. Graham. Scheduling the general job-shop, *Mgmt. Sci*., Vol. 31, No 5, 1985, pp. 594–598.
3. Blackstone, J. H. Jr., Philipps, D. T., Hogg, D. L. A state-of-the-art survey of dispatching rules for manufacturing job-shop operations, *Int. J. Prod. Res*., Vol. 20, 1982, pp. 27–45.
4. Carlier, J. and E. Pinson. An algorithm for solving the job-shop problem, *Mgmt. Sci*., Vol. 35, 1989, pp. 164–176.
5. Coffman, E. G. *Computer and Job-Shop Scheduling Theory*. New York: Wiley, 1976.
6. Conway, R. W., Maxwell, W. L. and L. W. Miller. *Theory of Scheduling*. Addison-Wesley, 1967.
7. Dell'Amico, M., Trubian, M. Applying tabu-search to the job-shop scheduling problem, *Annals of Operations Research*, Vol. 41, 1993, pp. 231–252.
8. Eilon, S. Production scheduling. In: *K. B. Haley (ed.), OR '78*. North-Holland, 1979, pp. 237–266.

9. French, S. *Sequencing and Scheduling*. New York: Wiley, 1982.
10. Gere, W. S. Heuristics in Job Shop Scheduling, *Mgmt. Sci*., Vol. 13, No 3, 1966, pp. 167–190.
11. Golenko-Ginzburg, D., Ben-Yair, A. and D. Greenberg. Undirected random search in multi-route scheduling problems, *Communications in Dependability and Quality Management*, Vol. 7, No 2, 2004, pp. 46–56.
12. Golenko-Ginzburg, D. I., Ljubkin, S. M., Rezer, V. S. and S. L. Sitniakovski. Using metric spaces in optimum scheduling, *Automation and Remote Control*, Vol. 63, No 9, 2002, pp. 1515–1523.
13. Golenko-Ginzburg, D. I. *Statistical Models in Production Control*. Moscow: Statistika, 1973. (In Russian)
14. Golenko-Ginzburg, D., Kats, V. *Priority rules in job-shop scheduling. Proceedings of the 14th Israel Congress on Advanced Technologies on Engineering, Management and Production*. Tel-Aviv, June 7–9, 1994, 41E–47E.
15. Gonik, A., Golenko-Ginzburg, D. Deterministic job-shop scheduling algorithms via permutation spaces, *Communications in Dependability and Quality Management*, Vol. 7, No 1, 2004, pp. 36–42.
16. *Industrial Scheduling / E. M. Muth, G. L. Thompson (Eds.)*. New Jersey: Prentice-Hall, Englewood Cliffs, 1963.
17. Panwalker, R., Iskander, W. A survey of scheduling rules, *Oper. Res*., Vol. 25, No 1, 1977, pp. 45–59.

Received on the 21st of July 2011

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 30–40
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

NEW APPROACH TO PATTERN RECOGNITION VIA COMPARISON OF MAXIMUM SEPARATIONS

K. N. Nechval[1], N. A. Nechval[2], M. Purgailis[2], V. F. Strelchonok[3] G. Berzins[2], M. Moldovan[4]

[1]Transport Department, Transport and Telecommunication Institute
Lomonosova Street 1, LV-1019, Riga, Latvia
E-mail: konstan@tsi.lv

[2]Statistics Department, EVF Research Institute, University of Latvia
Raina Blvd 19, LV-1050, Riga, Latvia
E-mail: nechval@junik.lv

[3]Informatics Department, Baltic International Academy
Lomonosova Street 4, LV-1019, Riga, Latvia
E-mail: str@apollo.lv

[4]Australian Institute of Health Innovation, University of New South Wales
Level 1 AGSM Building, Sydney NSW 2052, Australia.
E-mail: m.moldovan@unsw.edu.au

Fisher's linear discriminant analysis is a widely used multivariate statistical technique with two closely related goals: discrimination and classification. The technique is very popular among users of discriminant analysis. Some of the reasons for this are its simplicity and un-necessity of strict assumptions. In its original form, proposed by Fisher, the method assumes equality of population covariance matrices, but does not explicitly require multivariate normality. However, optimal classification performance of Fisher's discriminant function can only be expected when multivariate normality is present as well, since only good discrimination can ensure good allocation. In practice, we often are in need of analysing input data samples, which are not adequate for Fisher's classification rule, such that the distributions of the groups are not multivariate normal or covariance matrices of those are different or there are strong multi-nonlinearities.

This paper proposes a new approach to pattern recognition based on comparison of maximum separations in the input data samples, where maximum separations are determined via Fisher's separation criterion. The approach represents the improved pattern recognition procedure that allows one to take into account the cases which are not adequate for Fisher's classification rule. Moreover, it allows one to classify sets of multivariate observations, where each of the sets contains more than one observation. For the cases, which are adequate for Fisher's classification rule, the proposed approach gives the results similar to that of Fisher's classification rule. Illustrative examples are given.

Keywords: *Fisher's separation criterion, input data samples, maximum separation, pattern recognition*

1. Introduction

Humans have developed highly sophisticated skills for sensing their environment and taking actions according to what they observe, e.g., recognizing a face, understanding spoken words, reading handwriting, distinguishing fresh food from its smell. We would like to give similar capabilities to machines. A pattern is an entity, vaguely defined, that could be given a name, e.g., fingerprint image, handwritten word, human face, speech signal, etc. Pattern recognition is the study of how machines can observe the environment, learn to distinguish patterns of interest, make sound and reasonable decisions about the categories of the patterns. We are often influenced by the knowledge of how patterns are modelled and recognized in nature when we develop pattern recognition algorithms. Research on machine perception also helps us gain deeper understanding and appreciation for pattern recognition systems in nature. Yet, we also apply many techniques that are purely numerical and do not have any correspondence in natural systems. Pattern recognition techniques find applications in many areas: machine learning, statistics, mathematics, computer science, biology, etc. There are many sub-problems in the design process. Many of these problems can indeed be solved. More complex learning, searching and optimisation algorithms are developed with advances in computer technology. There remain many fascinating unsolved problems.

The aim of pattern recognition is to classify data (patterns) based on either a priori knowledge or on statistical information extracted from the patterns. The patterns to be classified are usually groups of measurements or observations, defining points in an appropriate multidimensional space. Many pattern recognition methods can be decomposed into two stages: discrimination followed by classification.

In some cases, the decomposition is explicit while in others it is a matter of interpretation. Discrimination and classification represent multivariate techniques concerned with separating distinct sets of objects (or observations) and allocating new objects (observations) to previously defined groups. There exist situations in which one may be interested in (1) discrimination: separating, say, two classes of objects or (2) classification: assigning a new object to one of two classes (or both).

The most popular separation criterion of establishing rules for discrimination and classification of patterns is the Fisher's discriminant (separation) ratio. Fisher's idea was to transform the ($p \geq 2$) multivariate observations $\mathbf{x}$ to univariate observations y such that the y's derived from populations π_1 and π_2 were separated as much as possible. Fisher suggested taking linear combinations of $\mathbf{x}$ to create y's because they are simple enough functions of the $\mathbf{x}$ to be handled easily. Fisher's approach does not assume that the populations are normal. It does, however, implicitly assume that the population covariance matrices are equal, because a pooled estimate of the common covariance matrix is used. A fixed linear combination of the $\mathbf{x}$'s takes the values $y_{11}, y_{12}, \ldots, y_{1n_1}$ for the n_1 observations of the $\mathbf{x}^{n_1} = (\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n_1})$ from the first population π_1 and the values $y_{21}, y_{22}, \ldots, y_{2n_2}$ for the n_2 observations of the sample $\mathbf{x}^{n_2} = (\mathbf{x}_{21}, \ldots, \mathbf{x}_{2n_2})$ from the second population π_2. The separation of these two sets of univariate y's is assessed in terms of the difference between the sample means $\bar{y}_1$ and $\bar{y}_2$ expressed in standard deviation units. The separation criterion proposed by Fisher is given by

$$I(\mathbf{u}) = \frac{(\bar{y}_1 - \bar{y}_2)^2}{s_y^2} = \frac{(\mathbf{u}'\bar{\mathbf{x}}_1 - \mathbf{u}'\bar{\mathbf{x}}_2)^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}} = \frac{[\mathbf{u}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}}, \tag{1}$$

and has to be maximized with respect to $\mathbf{u}$ (transformation vector), where

$$s_y^2 = \frac{\sum_{j=1}^{n_1}(y_{1i} - \bar{y}_1)^2 + \sum_{j=1}^{n_1}(y_{2i} - \bar{y}_2)^2}{n_1 + n_2 - 2} = \frac{\sum_{j=1}^{n_1}(\mathbf{u}'\mathbf{x}_{1j} - \mathbf{u}'\bar{\mathbf{x}}_1)^2 + \sum_{j=1}^{n_2}(\mathbf{u}'\mathbf{x}_{2j} - \mathbf{u}'\bar{\mathbf{x}}_2)^2}{n_1 + n_2 - 2}$$

$$= \frac{\sum_{j=1}^{n_1}[\mathbf{u}'(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)]'[\mathbf{u}'(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)] + \sum_{j=1}^{n_2}[\mathbf{u}'(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)]'[\mathbf{u}'(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)]}{n_1 + n_2 - 2}$$

$$= \frac{\sum_{j=1}^{n_1}[(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)'\mathbf{u}]'[(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)'\mathbf{u}] + \sum_{j=1}^{n_2}[(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)'\mathbf{u}]'[(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)'\mathbf{u}]}{n_1 + n_2 - 2}$$

$$= \frac{\sum_{j=1}^{n_1}\mathbf{u}'[(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)']\mathbf{u} + \sum_{j=1}^{n_2}\mathbf{u}'[(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)']\mathbf{u}}{n_1 + n_2 - 2}$$

$$= \frac{\mathbf{u}'[(n_1 - 1)\mathbf{S}_1]\mathbf{u} + \mathbf{u}'[(n_2 - 1)\mathbf{S}_2]\mathbf{u}}{n_1 + n_2 - 2} = \mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u} \tag{2}$$

is the pooled estimate of the variance. The sample means vectors and covariance matrices are determined by

$$\bar{\mathbf{x}}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1}\mathbf{x}_{1j}, \qquad \mathbf{S}_1 = \frac{1}{n_1 - 1}\sum_{j=1}^{n_1}(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)' \tag{3}$$

and

$$\bar{\mathbf{x}}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2}\mathbf{x}_{2j}, \qquad \mathbf{S}_2 = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)'. \tag{4}$$

Since it is assumed implicitly that the populations π_1 and π_2 have the same covariance matrix $\mathbf{\Sigma}$, the sample covariance matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ are combined (pooled) to derive a single unbiased estimate of $\mathbf{\Sigma}$,

$$\mathbf{S}_{\text{pooled}} = \frac{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2}{n_1 + n_2 - 2}. \tag{5}$$

Theorem 1. (*Separation criterion: maximization*). Let $\mathbf{S}_{\text{pooled}}$ be positive definite and $(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$ be a given vector. Then the optimal transformation vector is given by

$$\mathbf{u}^* = \arg\left[\max_{\mathbf{u}\neq\mathbf{0}} I(\mathbf{u})\right] = \arg\left[\max_{\mathbf{u}\neq\mathbf{0}} \frac{[\mathbf{u}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}}\right] = \mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \tag{6}$$

with the maximum attained

$$I(\mathbf{u}^*) = \max_{\mathbf{u}\neq\mathbf{0}} \frac{[\mathbf{u}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}} = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2). \tag{7}$$

Proof. By the extended Cauchy-Schwarz inequality [1],

$$[\mathbf{u}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]^2 \le [\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}][(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]. \tag{8}$$

Because $\mathbf{u} \neq 0$ and $\mathbf{S}_{\text{pooled}}$ is positive definite, $\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u} > 0$. Dividing both sides of the inequality (8) by the positive scalar $\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}$ yields the upper bound

$$\frac{[\mathbf{u}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}}\mathbf{u}} \le (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2). \tag{9}$$

Taking the maximum over $\mathbf{u}$ gives Equation (7) because the bound is attained for $\mathbf{u} = \mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$. This completes the proof.

Fisher's linear discriminant analysis has been successfully used as dimensionality reduction technique to many classification problems, such as face recognition and multimedia information retrieval. The Fisher discriminant criterion is the benchmark for the linear discrimination in multidimensional space [2]. The criterion purpose of the Fisher linear discriminant for pattern analysis is to find an optimal discriminant direction based on the Fisher criterion so that the projected set of training samples on it has the maximal ratio of between-class distance to within-class distance [3–4]. Sammon extended the Fisher linear discriminant method to the optimal discriminant plane in 1970 [5]. Then Foley and Sammon [6] further extended this in 1975 and proposed the optimal set of discriminant vectors by which the well-known Foley-Sammon Transform (FST) can be constituted. Their important result has attracted many researchers' attention in the field of pattern recognition [7–8] and has been used in many pattern classification applications [9–10].

2. Fisher's Approach to Pattern Recognition

2.1. Case of Two Populations

In this case, Fisher's approach to pattern recognition is as follows:

Discrimination. At this step, the optimal transformation vector (Fisher's solution), which allows one to maximally separate the two populations (π_1 and π_2), is determined by (6). The maximum separation in the samples is given by (7).

Classification. At this step, Fisher's solution to the separation problem (Fisher's discriminant function) is used to classify new observations as follows:

Allocate $\mathbf{x}_0$ to π_1 if

$$y_0 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}\mathbf{x}_0 \ge \frac{\bar{y}_1 + \bar{y}_2}{2} = \frac{(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)}{2}; \tag{10}$$

Allocate $\mathbf{x}_0$ to π_2 if

$$y_0 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}\mathbf{x}_0 < \frac{\bar{y}_1 + \bar{y}_2}{2} = \frac{(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)}{2}. \tag{11}$$

Thus, the allocation rule based on Fisher's discriminant function to classify $\mathbf{x}_0$ is given by

$$(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}\mathbf{x}_0 \begin{cases} \geq [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)]/2, & \text{then} \quad \mathbf{x}_0 \in \pi_1, \\ < [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{S}_{\text{pooled}}^{-1}(\bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2)]/2, & \text{then} \quad \mathbf{x}_0 \in \pi_2. \end{cases} \tag{12}$$

A pictorial representation of Fisher's procedure for two p-variate populations with p=2 is given in Figure 1.

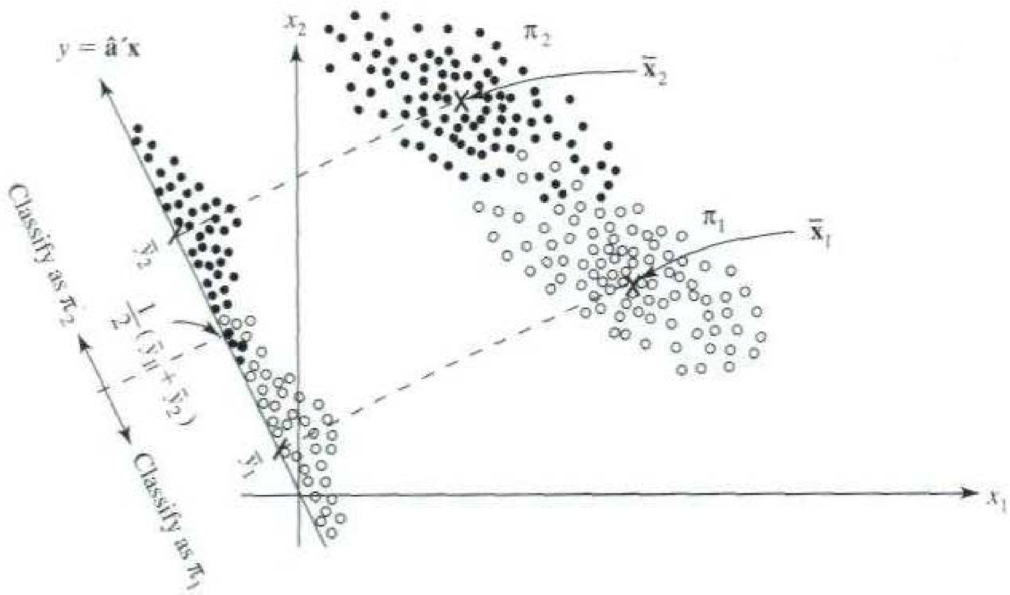


Figure 1. A pictorial representation of Fisher's procedure for two p-variate populations with p = 2

The misclassification probabilities are shown on Figure 2.

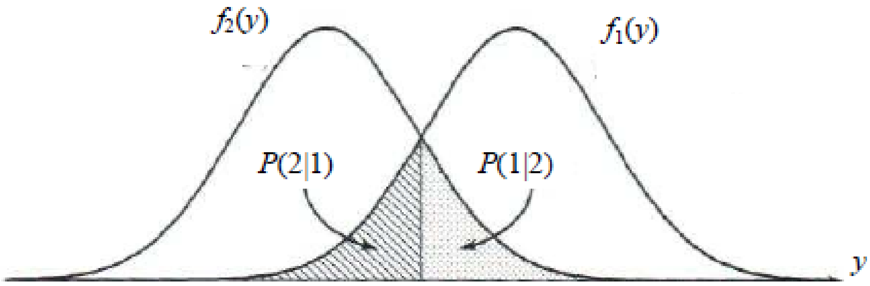


Figure 2. The misclassification probabilities based on Y

Thus, Fisher's approach to pattern recognition has optimality properties (in particular, the minimum total probability of misclassification) only if the underlying distributions of the groups of univariate observations (in terms of the random variable Y) obtained via transformation of the groups of multivariate observations (in terms of the random variable $\mathbf{x}$) are symmetrical with equal variances.

2.2. Case of Several Populations

In this case, Fisher's approach to pattern recognition consists in the following:

Discrimination. At this step, the optimal transformation vector (Fisher's solution), which allows one to maximally separate the ($m > 2$) populations (π_1, …, π_m), is determined by

$$\mathbf{u}^* = \arg\max_{\mathbf{u}\neq\mathbf{0}} I(\mathbf{u}), \tag{13}$$

where

$$I(\mathbf{u}) = \frac{\mathbf{u}'\mathbf{B}\mathbf{u}}{\mathbf{u}'\mathbf{W}\mathbf{u}} \tag{14}$$

represents the criterion of separation in the input data samples,

$$\mathbf{B} = \sum_{i=1}^{m} n_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' \tag{15}$$

is the *sample between groups matrix* which includes the sample sizes n_i, sample mean vectors

$$\bar{\mathbf{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{ij}, \quad i = 1, \ldots, m, \tag{16}$$

and the "overall average"

$$\bar{\mathbf{x}} = \sum_{i=1}^{m} n_i \bar{x}_i \bigg/ \sum_{i=1}^{m} n_i, \tag{17}$$

which is the $p \times 1$ vector average taken over all the sample observations in the set of input data samples,

$$\mathbf{W} = \sum_{i=1}^{m} \sum_{j=1}^{n_i} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)' = \left(\sum_{i=1}^{m} n_i - m \right) \mathbf{S}_{\text{pooled}} \tag{18}$$

is the *sample within group matrix*.

The maximum separation in the input data samples can be obtained as follows. Maximizing $I(\mathbf{u})$ by taking the derivative with respect to $\mathbf{u}$ and setting it to 0,

$$\frac{dI(\mathbf{u})}{d\mathbf{u}} = \frac{d}{d\mathbf{u}} \left(\frac{\mathbf{u}'\mathbf{B}\mathbf{u}}{\mathbf{u}'\mathbf{W}\mathbf{u}} \right) = 0, \tag{19}$$

we have

$$(\mathbf{W}^{-1}\mathbf{B} - \lambda \mathbf{I})\mathbf{u} = 0, \tag{20}$$

where

$$\lambda = \frac{\mathbf{u}'\mathbf{B}\mathbf{u}}{\mathbf{u}'\mathbf{W}\mathbf{u}}. \tag{21}$$

Thus, we deal with a standard eigenvalue problem. Now, we must solve

$$\left| \mathbf{W}^{-1}\mathbf{B} - \lambda \mathbf{I} \right| = 0 \tag{22}$$

for the r nonzero eigenvalues ($\lambda_1, \ldots, \lambda_r$) of $\mathbf{W}^{-1}\mathbf{B}$, where $r \le \min(m - 1, p)$. Then the corresponding normalized eigenvectors are obtained by solving

$$(\mathbf{W}^{-1}\mathbf{B} - \lambda_k \mathbf{I})\mathbf{u}_k^* = 0, \quad k = 1, \ldots, r, \tag{23}$$

and scaling the results such that

$$(\mathbf{u}_k^*)'\mathbf{S}_{\text{pooled}}\mathbf{u}_k^* = 1, \quad k = 1, \ldots, r, \tag{24}$$

Thus, the maximum separation in the input data samples based on the criterion (14) is given by

$$\frac{\sum_{k=1}^{r} (\mathbf{u}_k^*)'\mathbf{B}\mathbf{u}_k^*}{\sum_{k=1}^{r} (\mathbf{u}_k^*)'\mathbf{W}\mathbf{u}_k^*} = \frac{\sum_{k=1}^{r} (\mathbf{u}_k^*)' \left[\sum_{i=1}^{m} n_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})' \right] \mathbf{u}_k^*}{\sum_{k=1}^{r} (\mathbf{u}_k^*)' \left[\sum_{i=1}^{m} \sum_{j=1}^{n_i} (\bar{\mathbf{x}}_{ij} - \bar{\mathbf{x}}_i)(\bar{\mathbf{x}}_{ij} - \bar{\mathbf{x}}_i)' \right] \mathbf{u}_k^*}. \tag{25}$$

Classification. At this step, Fisher's solution to the separation problem (Fisher's discriminant function) is used to classify new observations as follows:

Allocate $\mathbf{x}_0$ to π_l, $l\in\{1, \ldots, m\}$ if

$$\sum_{k=1}^{r}(y_{0k}-\bar{y}_{lk})^2=\sum_{k=1}^{r}[(\mathbf{u}_k^*)'(\mathbf{x}_0-\bar{\mathbf{x}}_l)]^2\le\sum_{k=1}^{r}[(\mathbf{u}_k^*)'(\mathbf{x}_0-\bar{\mathbf{x}}_i)]^2 \text{ for all } i\in\{1,\ldots,m\},\ i\neq l, \tag{26}$$

where $\mathbf{u}_k^*$ is defined in (21), $\bar{y}_{lk}=(\mathbf{u}_k^*)'\bar{\mathbf{x}}_l$ and $r \le \min(m-1, p)$.

3. Pattern Recognition Based on Comparing of Maximum Separations in Data Samples

3.1. Case of Two Populations

Discrimination. At this step, based on (7) we determine the maximum separations in the two input data samples taking into account a new observation $\mathbf{x}_0$.

Classification. In order to classify a new observation $\mathbf{x}_0$ to one of the two populations (π_1, π_2), we take decision based on comparison of the maximum separations in the two source data samples as follows:

Allocate $\mathbf{x}_0$ to π_1,

if the maximum separation $(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)'\mathbf{S}^{-1}_{\text{pooled}(1)}(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)$ ($\mathbf{x}_0$ has been added to the sample $\mathbf{x}^{n_1}$ from π_1)

$\ge$ the maximum separation $(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})'\mathbf{S}^{-1}_{\text{pooled}(2)}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})$ ($\mathbf{x}_0$ has been added to the sample $\mathbf{x}^{n_2}$ from π_2), (27)

where

$$\bar{\mathbf{x}}_{1(0)}=\frac{1}{n_1+1}\sum_{j=0}^{n_1}\mathbf{x}_{1i},\qquad \mathbf{S}_{1(0)}=\frac{1}{n_1}\sum_{j=0}^{n_1}(\mathbf{x}_{1i}-\bar{\mathbf{x}}_{1(0)})(\mathbf{x}_{1i}-\bar{\mathbf{x}}_{1(0)})',\qquad \mathbf{x}_{10}\equiv\mathbf{x}_0, \tag{28}$$

$$\bar{\mathbf{x}}_{2(0)}=\frac{1}{n_2+1}\sum_{j=0}^{n_2}\mathbf{x}_{2i},\qquad \mathbf{S}_{2(0)}=\frac{1}{n_2}\sum_{j=0}^{n_2}(\mathbf{x}_{2i}-\bar{\mathbf{x}}_{2(0)})(\mathbf{x}_{2i}-\bar{\mathbf{x}}_{2(0)})',\qquad \mathbf{x}_{20}\equiv\mathbf{x}_0, \tag{29}$$

and

$$\mathbf{S}_{\text{pooled}(1)}=\frac{n_1\mathbf{S}_{1(0)}+(n_2-1)\mathbf{S}_2}{n_1+n_2-1},\qquad \mathbf{S}_{\text{pooled}(2)}=\frac{(n_1-1)\mathbf{S}_1+n_2\mathbf{S}_{2(0)}}{n_1+n_2-1}. \tag{30}$$

Allocate $\mathbf{x}_0$ to π_2,

if the maximum separation $(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)'\mathbf{S}^{-1}_{\text{pooled}(1)}(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)$ ($\mathbf{x}_0$ has been added to the sample $\mathbf{x}^{n_1}$ from π_1)

$<$ the maximum separation $(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})'\mathbf{S}^{-1}_{\text{pooled}(2)}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})$ ($\mathbf{x}_0$ has been added to the sample $\mathbf{x}^{n_2}$ from π_2). (31)

Thus, the allocation rule based on comparison of the maximum separations in the two input data samples to classify $\mathbf{x}_0$ is given by

$$\frac{\max\limits_{\mathbf{u}\neq\mathbf{0}}\dfrac{[\mathbf{u}'(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}(1)}\mathbf{u}}}{\max\limits_{\mathbf{u}\neq\mathbf{0}}\dfrac{[\mathbf{u}'(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}(2)}\mathbf{u}}}=\frac{(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)'\mathbf{S}^{-1}_{\text{pooled}(1)}(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)}{(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})'\mathbf{S}^{-1}_{\text{pooled}(2)}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})}=\frac{\bar{y}_{1(0)}-\bar{y}_2}{\bar{y}_1-\bar{y}_{2(0)}}\begin{cases}\ge 1, & \text{then } \mathbf{x}_0\in\pi_1,\\ <1, & \text{then } \mathbf{x}_0\in\pi_2.\end{cases} \tag{32}$$

3.2. Case of Several Populations

3.2.1. Version I (Paired-comparison procedure)

Discrimination. At this step, based on (7) we determine the maximum separations in the input data samples taking into account a new observation $\mathbf{x}_0$.

Classification. In order to classify $\mathbf{x}_0$ to one of several, say, m ($m > 2$) populations, we use the ratio of the maximum separations in the input data samples as follows. At first, any two data samples from populations, say, π_k and π_r, respectively, are considered.

The allocation rule based on the ratio of the maximum separations in the above data samples to classify $\mathbf{x}_0$ is given by

$$\frac{[(\overline{\mathbf{x}}_{k(0)}-\overline{\mathbf{x}}_r)'\mathbf{S}^{-1}_{\text{pooled}(k)}(\overline{\mathbf{x}}_{k(0)}-\overline{\mathbf{x}}_r)\,|\,(\mathbf{x}_0,\mathbf{x}^{n_k}),\mathbf{x}^{n_r}]}{[(\overline{\mathbf{x}}_k-\overline{\mathbf{x}}_{r(0)})'\mathbf{S}^{-1}_{\text{pooled}(r)}(\overline{\mathbf{x}}_k-\overline{\mathbf{x}}_{r(0)})\,|\,\mathbf{x}^{n_k},(\mathbf{x}_0,\mathbf{x}^{n_r})]}\begin{cases}\geq 1, & \text{then} \quad \mathbf{x}_0\in\pi_k,\\ <1, & \text{then} \quad \mathbf{x}_0\in\pi_r.\end{cases} \tag{33}$$

where

$$\overline{\mathbf{x}}_{k(0)}=\frac{1}{n_k+1}\sum_{i=0}^{n_k}\mathbf{x}_{ki},\qquad \mathbf{S}_{k(0)}=\frac{1}{n_k}\sum_{i=0}^{n_k}(\mathbf{x}_{ki}-\overline{\mathbf{x}}_{k(0)})(\mathbf{x}_{ki}-\overline{\mathbf{x}}_{k(0)})',\qquad \mathbf{x}_{k0}\equiv\mathbf{x}_0, \tag{34}$$

$$\overline{\mathbf{x}}_{r(0)}=\frac{1}{n_r+1}\sum_{i=0}^{n_r}\mathbf{x}_{ri},\qquad \mathbf{S}_{r(0)}=\frac{1}{n_r}\sum_{i=0}^{n_r}(\mathbf{x}_{ri}-\overline{\mathbf{x}}_{r(0)})(\mathbf{x}_{ri}-\overline{\mathbf{x}}_{r(0)})',\qquad \mathbf{x}_{r0}\equiv\mathbf{x}_0, \tag{35}$$

and

$$\mathbf{S}_{\text{pooled}(k)}=\frac{n_k\mathbf{S}_{k(0)}+(n_r-1)\mathbf{S}_{\text{r}}}{n_k+n_r-1},\qquad \mathbf{S}_{\text{pooled(r)}}=\frac{(n_k-1)\mathbf{S}_k+n_r\mathbf{S}_{r(0)}}{n_k+n_r-1}. \tag{36}$$

If $\mathbf{x}_0\in\pi_k$, then the population π_r is eliminated from further consideration. If ($m-1$) populations are so eliminated, then the remaining population (say, kth) is the one to which the observation $\mathbf{x}_0$ being classified belongs.

3.2.2. Version II

Discrimination. At this step, based on (25) we determine the maximum separations in the input data samples taking into account a new observation $\mathbf{x}_0$.

Classification. At this step, the maximum separation in the input data samples is used to classify $\mathbf{x}_0$ as follows:

Allocate $\mathbf{x}_0$ to π_l, $l\in\{1, \ldots, m\}$ if

$$\frac{\sum_{k=1}^{r}(\mathbf{u}^*_{k(l)})'\mathbf{B}_l\mathbf{u}^*_{k(l)}}{\sum_{k=1}^{r}(\mathbf{u}^*_{k(l)})'\mathbf{W}_l\mathbf{u}^*_{k(l)}}\geq\frac{\sum_{k=1}^{r}(\mathbf{u}^*_{k(s)})'\mathbf{B}_i\mathbf{u}^*_{k(s)}}{\sum_{k=1}^{r}(\mathbf{u}^*_{k(s)})'\mathbf{W}_i\mathbf{u}^*_{k(s)}}\quad \text{for all } s\in\{1,\ldots,m\},\ s\neq l, \tag{37}$$

where $\mathbf{x}_0$ has been added to the sample $\mathbf{x}^{n_l}$ from π_l, $\mathbf{u}^*_{k(l)}$ ($k = 1, \ldots, r$) are defined via (12), and $r \leq \min(m-1, p)$.

4. Examples

4.1. Example 1

This example is adapted from a study [11] concerned with the detection of hemophilia A carriers. To construct a procedure for detecting potential hemophilia A carriers, blood samples were assayed for two groups of women and measurements on the two variables,

$$X_1 = \log_{10}(\text{AHF activity}) \text{ and } X_2 = \log_{10}(\text{AHF-like antigen}) \tag{38}$$

recorded (see Table 1). ("AHF" denotes antihemophilic factor.)

Table 1. Hemophilia data

Group 1: Noncarriers (π_1)			Group 2: Obligatory carriers (π_2)		
No.	$\log_{10}$(AHF activity)	$\log_{10}$(AHF –like antigen)	No.	$\log_{10}$(AHF activity)	$\log_{10}$(AHF –like antigen)
1	-0.23	-0.3	1	-0.45	0.015
2	-0.18	-0.3	2	-0.43	-0.095
3	-0.13	-0.3	3	-0.42	-0.12
4	-0.16	-0.24	4	-0.41	-0.25
5	-0.025	-0.2	5	-0.38	-0.28
6	-0.12	-0.08	6	-0.35	-0.015
7	-0.075	-0.14	7	-0.34	0.1
8	-0.02	-0.15	8	-0.33	-0.13
9	-0.13	-0.05	9	-0.24	0.28
10	-0.08	-0.055	10	-0.24	0.15
11	-0.025	-0.09	11	-0.26	0.08
12	-0.06	-0.04	12	-0.26	-0.075
13	0	-0.08	13	-0.25	-0.04
14	0.05	-0.08	14	-0.22	-0.015
15	0.07	-0.1	15	-0.22	0.024
16	0.03	0.11	16	-0.21	-0.04
17	0.05	0	17	-0.175	-0.09
18	0.04	-0.03	18	-0.2	0.25
19	0.1	0	19	-0.19	0.175
20	0.075	0.02	20	-0.075	0.17
21	0.055	0.05	21	-0.015	0.15
22	0.06	0.1	22	-0.03	0.0135
23	0.09	0.09	23	-0.025	0.08
24	0.1	0.05			
25	0.11	0.035			
26	0.1	0.125			
27	0.12	0.125			
28	0.14	0.07			
29	0.21	0.11			

The first group of $n_1 = 29$ women were selected from a population of women who did not carry the hemophilia gene. This group was called the *normal* group. The second group of $n_2 = 23$ women was selected from known hemophilia A carriers (daughters of hemophiliacs, mothers with more than one hemophilic son, and mothers with one hemophilic son and other hemophilic relatives). This group was called the *obligatory carriers*. The pairs of observations (x_1, x_2) for the two groups are plotted on Figure 3. Also estimated contours are shown containing 50% and 95% of the probability for bivariate normal distributions centred at $\overline{\mathbf{x}}_1$ and $\overline{\mathbf{x}}_2$, respectively. Their common covariance matrix was taken as the pooled sample covariance matrix $\mathbf{S}_{\text{pooled}}$. In this example, bivariate normal distributions seem to fit the data fairly well.

The following information is given below:

$$\overline{\mathbf{x}}_1 = \begin{bmatrix} 0.00569 \\ -0.04655 \end{bmatrix}, \quad \overline{\mathbf{x}}_2 = \begin{bmatrix} -0.24870 \\ 0.01467 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_{\text{pooled}}^{-1} = \begin{bmatrix} 121.6071 & -75.6327 \\ -75.6327 & 104.7043 \end{bmatrix}. \tag{39}$$

It follows from (6) that the optimal transformation vector is

$$\mathbf{u}^* = \mathbf{S}_{\text{pooled}}^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) = \begin{bmatrix} 121.6071 & -75.6327 \\ -75.6327 & 104.7043 \end{bmatrix}\begin{bmatrix} 0.2544 \\ -0.0612 \end{bmatrix} = \begin{bmatrix} 35.5657 \\ -25.6504 \end{bmatrix} \tag{40}$$

with

$$\overline{y}_1 = (\mathbf{u}^*)'\overline{\mathbf{x}}_1 = [35.5657 \quad -25.6504]\begin{bmatrix} 0.00569 \\ -0.04655 \end{bmatrix} = 1.3964 \tag{41}$$

and

$$\overline{y}_2 = (\mathbf{u}^*)'\overline{\mathbf{x}}_2 = [35.5657 \quad -25.6504]\begin{bmatrix} -0.24870 \\ -0.01467 \end{bmatrix} = -9.2214, \tag{42}$$

where the midpoint between these means is

$$\frac{\overline{y}_1 + \overline{y}_2}{2} = \frac{1.3964 - 9.2214}{2} = -3.9125. \tag{43}$$

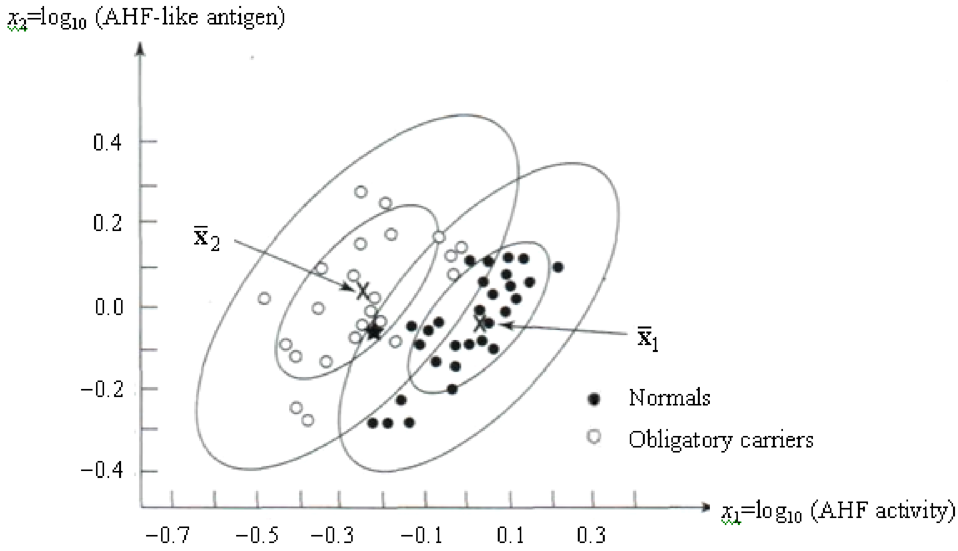


Figure 3. Scatter plots of [log$_{10}$ (AHF activity), log$_{10}$ (AHF-like antigen)] for the normal group and obligatory hemophilia A carriers

For instance, measurements of AHF activity and AHF-like antigen on a woman who may be a hemophilia A carrier give x_1 = –0.210 and x_2 = –0.044. Should this woman be classified as π_1 (normal) or π_2 (obligatory carrier)?

Using Fisher's approach, we obtain

$$y_0 = (\mathbf{u}^*)'\mathbf{x}_0 = [35.5657 \quad -25.6504]\begin{bmatrix}-0.210\\-0.044\end{bmatrix} = -6.34 < \frac{\bar{y}_1+\bar{y}_2}{2} = -3.9125. \tag{44}$$

Using the proposed approach based on the ratio of the maximum separations in the input data samples, we obtain

$$\frac{\max\limits_{\mathbf{u}\neq\mathbf{0}} \dfrac{[\mathbf{u}'(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}(1)}\mathbf{u}}}{\max\limits_{\mathbf{u}\neq\mathbf{0}} \dfrac{[\mathbf{u}'(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})]^2}{\mathbf{u}'\mathbf{S}_{\text{pooled}(2)}\mathbf{u}}} = \frac{(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)'\mathbf{S}^{-1}_{\text{pooled}(1)}(\bar{\mathbf{x}}_{1(0)}-\bar{\mathbf{x}}_2)}{(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})'\mathbf{S}^{-1}_{\text{pooled}(2)}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_{2(0)})} = \frac{\bar{y}_{1(0)}-\bar{y}_2}{\bar{y}_1-\bar{y}_{2(0)}} = \frac{1.029+8.84}{1.4+9.782} = 0.88 < 1. \tag{45}$$

Applying either (42) or (43), we classify the woman as π_2, an obligatory carrier. Thus, Fisher's approach and the proposed one give the same result in the above case.

4.2. Example 2

Consider the observations on $p = 2$ variables from $m = 3$ populations. The input data samples are given below.

π_1 ($n_1 = 3$) $\qquad$ π_2 ($n_2 = 3$) $\qquad$ π_3 ($n_3 = 3$)

$$\mathbf{x}^{n_1} = \begin{bmatrix}-2 & 5\\0 & 3\\-1 & 1\end{bmatrix}; \qquad \mathbf{x}^{n_2} = \begin{bmatrix}0 & 6\\2 & 4\\1 & 2\end{bmatrix}; \qquad \mathbf{x}^{n_3} = \begin{bmatrix}1 & -2\\0 & 0\\-1 & -4\end{bmatrix}. \tag{46}$$

We found that

$$\bar{\mathbf{x}}_1 = \begin{bmatrix}-1\\3\end{bmatrix}; \qquad \bar{\mathbf{x}}_2 = \begin{bmatrix}1\\4\end{bmatrix}; \qquad \bar{\mathbf{x}}_3 = \begin{bmatrix}0\\-2\end{bmatrix}; \qquad \bar{\mathbf{x}} = \begin{bmatrix}0\\5/3\end{bmatrix}; \tag{47}$$

$$\mathbf{B} = \sum_{i=1}^{3} 3(\bar{\mathbf{x}}_i-\bar{\mathbf{x}})(\bar{\mathbf{x}}_i-\bar{\mathbf{x}})' = \begin{bmatrix}6 & 3\\3 & 62\end{bmatrix}; \tag{48}$$

$$\mathbf{W} = \sum_{i=1}^{3}\sum_{j=1}^{n_i}(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)' = \left(\sum_{i=1}^{3} n_i - 3\right)\mathbf{S}_{\text{pooled}} = \begin{bmatrix} 6 & -2 \\ -2 & 24 \end{bmatrix}; \tag{49}$$

$$\mathbf{W}^{-1} = \begin{bmatrix} 0.171429 & 0.014286 \\ 0.014286 & 0.042857 \end{bmatrix}; \quad \mathbf{W}^{-1}\mathbf{B} = \begin{bmatrix} 1.071429 & 1.4 \\ 0.214286 & 2.7 \end{bmatrix}. \tag{50}$$

To solve for the $r \le \min(m-1, p) = \min(2, 2) = 2$ nonzero eigenvalues of $\mathbf{W}^{-1}\mathbf{B}$, we must solve

$$\left|\mathbf{W}^{-1}\mathbf{B} - \lambda\mathbf{I}\right| = \begin{vmatrix} 1.071429 - \lambda & 1.4 \\ 0.214286 & 2.7 - \lambda \end{vmatrix} = 0. \tag{51}$$

We find that $\lambda_1 = 2.867071$ and $\lambda_2 = 0.904357$. Then the corresponding normalized eigenvectors are obtained by solving

$$(\mathbf{W}^{-1}\mathbf{B} - \lambda_k\mathbf{I})\mathbf{u}_k^* = 0, \quad k = 1, 2, \tag{52}$$

and scaling the results such that

$$(\mathbf{u}_k^*)'\mathbf{S}_{\text{pooled}}\mathbf{u}_k^* = 1, \quad k = 1,2. \tag{53}$$

Thus, we obtain

$$\mathbf{u}_1^* = [0.386 \quad 0.495], \quad \mathbf{u}_2^* = [0.938 \quad -0.112]. \tag{54}$$

Using Fisher's approach to classify the new observation $\mathbf{x}_0' = [1 \quad 3]$ in accordance with (26), we have

$$\sum_{k=1}^{2}(y_{0k} - \bar{y}_{1k})^2 = \sum_{k=1}^{2}[(\mathbf{u}_k^*)'(\mathbf{x}_0 - \bar{\mathbf{x}}_1)]^2 = (1.871 - 1.099)^2 + (0.602 + 1.274)^2 = 4.11536, \tag{55}$$

$$\sum_{k=1}^{2}(y_{0k} - \bar{y}_{2k})^2 = \sum_{k=1}^{2}[(\mathbf{u}_k^*)'(\mathbf{x}_0 - \bar{\mathbf{x}}_2)]^2 = (1.871 - 2.366)^2 + (0.602 - 0.49)^2 = 0.257569, \tag{56}$$

$$\sum_{k=1}^{2}(y_{0k} - \bar{y}_{3k})^2 = \sum_{k=1}^{2}[(\mathbf{u}_k^*)'(\mathbf{x}_0 - \bar{\mathbf{x}}_3)]^2 = (1.871 + 0.99)^2 + (0.602 - 0.224)^2 = 8.328205. \tag{57}$$

Since the minimum of $\sum_{k=1}^{2}(y_{0k} - \bar{y}_{lk})^2$ occurs when $l = 2$, we allocate x_0 to population π_2. It will be noted that the approach proposed in this paper gives in the above case the same result. The situation, in terms of the classifiers $\bar{y}_i$ is illustrated on Figure 4.

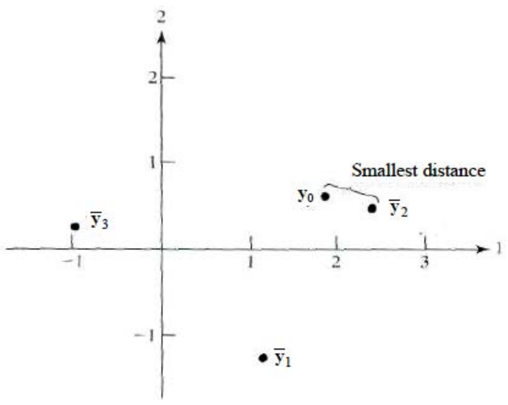


Figure 4. The points y_0, $\bar{y}_1$, $\bar{y}_2$, and $\bar{y}_3$ in the classification plane

5. Conclusions

The methodology described here can be extended in several different directions to handle various problems of pattern recognition that arise in practice (in particular, the problem of change-point detection in a sequence of multivariate observations).

Acknowledgments

This research was supported in part by Grant No. 06.1936, Grant No. 07.2036, Grant No. 09.1014, and Grant No. 09.1544 from the Latvian Council of Science and the National Institute of Mathematics and Informatics of Latvia.

References

1. Rao, C. R. *Linear Statistical Inference and its Applications*. New York: John Wiley, 1965.
2. Fisher, R. The Use of Multiple Measurements in Taxonomic Problems, *Ann. Eugenics*, Vol. 7, 1936, pp. 178–188.
3. Tian, Q., Fainman, Y., Gu, Z. H. and S. Lee. Comparison of Statistical Pattern-Recognition Algorithms for Hybrid Processing, I. Linear-Mapping Algorithm, *J. Opt. Soc. Am. A*, Vol. 5, 1988, pp. 1655–1669.
4. Tian, Q., Fainman, Y. and S. Lee. Comparison of Statistical Pattern-Recognition Algorithms for Hybrid Processing, II: Eigenvector-Based Algorithm, *J. Opt. Soc. Am. A*, Vol. 5, 1988, pp. 1670–1682.
5. Sammon, J. W. An Optimal Discriminant Plane, *IEEE Trans. Computer*, Vol. C-19, 1970, pp. 826–829.
6. Foley, D. and J. Sammon. An Optimal Set of Discriminant Vectors, *IEEE Trans. Computer,* Vol. C-24, 1975, pp. 281–289.
7. Guo, Y. F., Shu, T. T., Yang, J. Y., Li, S. J. Feature Extraction Method Based on the Generalized Fisher Discriminant Criterion and Facial Recognition, *Pattern Analysis & Applications*, Vol. 4, 2001, pp. 61–66.
8. Guo, Y. F., Li, S. J., Yang, J. Y., Shu, T. T., Wu, L. D. A generalized Foley-Sammon Transform (GFST) Based on Generalized Fisher Discriminant Criterion and its Application to Face Recognition, *Pattern Recognition Letters*, Vol. 24, 2003, pp. 147–158.
9. Belhumeur, P., Hespanha, J. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, 1997, pp. 711–720.
10. Salton, G. and M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, 1983.
11. Bouma, B. N., et al. Evaluation of the Detection Rate of Hemophilia Carriers, *Statistical Methods for Clinical Decision Making*, Vol. 7, 1975, pp. 339–350.

Received on the 21st of July 2011

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 41–52
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

INVESTIGATION OF ONE MATHEMATICAL MODEL FOR THE MAXIMISATION OF DIVISIBLE PRODUCTION VOLUME UNDER LIMITED RESOURCES

Sh. E. Guseynov[1,2], J. Baranova[1]

[1]Transport and Telecommunication Institute
Lomonosova 1, Riga, LV-1019, Latvia
E-mail: bavarde@inbox.lv
[2]Institute of Mathematical Sciences and Information Technologies
14 Liela Street, Liepaya, LV-3401, Latvia
E-mail: sh.e.guseinov@inbox.lv

The economic-mathematical optimising problem regarding producing the maximum volume of divisible production (for example the production of diet bread) is formulated and investigated in circumstances of limited resources. A mathematical model of the investigated problem, based on Leontiev's model is developed. Analytical decisions regarding the construction of the model are taken, and a self-regulation method continuously dependent on all initial data is chosen as the method for finding the optimum output. A software product is developed for carrying out of numerical experiments with real initial economic data. The constructed mathematical model has practical value, as any divisible production (loose or not loose) can be used; for example, bakery products, heaters for walls and floors, paper products, cement, paints etc.

Keywords: *Leontiev's model, the duality theory, optimum planning of manufacture*

1. Introduction

The economic crisis that started in 2007 is the largest and the most painful experienced for the whole history of the independent Republic of Latvia. It has forced not only the government, but also the majority of enterprises to accept that they have overestimated economic activity and the possibility of implementing development plans in the short-term as well as in the long-term.

Thus, companies, which occurred, are in a very complicated economic situation, have had to search for new methods of optimising their work in order to remain competitive in the Latvian market. They have had to try to maximise volume of output with limited resources, limited funding and other restrictions, while considering the current price of manufactured production, the appropriate volume of production, human resources etc. In this paper the problem of maximising the volume of output of divisible production (for example, the production of diet bread) will be investigated in the circumstances of limited resources and limited funding. In order to achieve the aim, the following problems are dealt with: working out a mathematical model of the considered problem; making analytical decisions regarding the constructed model; working out a software product for a computer realisation of the constructed model; carrying out numerical calculations by means of the developed software product on an example of a problem of producing bread; carrying out an economic-mathematical analysis of the developed model, its analytical decisions and the results of numerical experiments.

The Latvian bread market has been chosen as the industrial branch, which will be the area of investigation.

There is great demand among the Latvian population for grain, as despite the depressing consequences of the economic crisis, bread always has been and remains the one of basic needs of the population of Latvia.

2. Formalisation of the Consideration Problem: *Construction of the Qualitative Model*

Let us formulate a verbal statement of the consideration problem: suppose that next year (or month or any calendar time) a business that manufactures grain products plans to launch a qualitatively new product X (for example, diet bread). In order to manufacture this product, $n \in N$ kinds of raw materials are required. Analysts working for the company have established the following initial information:

- The quantity of the raw materials necessary for manufacture individual quantity of the product X, is equal to $a_i \in \left[a_i^{\{1\}}, a_i^{\{2\}}\right]$, where $\left[a_i^{\{1\}}, a_i^{\{2\}}\right]$ $\left(i=\overline{1,n}\right)$ there is a segment within which the variation of used quantity i $\left(i=\overline{1,n}\right)$ is limited, the raw materials are comprehensible i $\left(i=\overline{1,n}\right)$ to manufacture the individual quantity X;
- The remains of the i $\left(i=\overline{1,n}\right)$ raw materials in the enterprise at the end of current year are equal to b_i $\left(i=\overline{1,n}\right)$;
- Cost of individual quantity i $\left(i=\overline{1,n}\right)$ of the raw materials is equal to c_i $\left(i=\overline{1,n}\right)$;
- The quantity of financial resources, which the enterprise can spend on the purchase of additional volumes of raw materials during the current year, is equal to FR (financial resources).

It is required:

1) To define what kinds and what quantity of raw materials should be bought by the enterprise to provide the maximum volume of output of the product X next year;
2) To construct dependence of the maximum volume of output of the product X on the size of financial resources FR provided that parameter n, $\mathsf{a}=\{a_i\}_{i=\overline{1,n}}$, $\mathsf{b}=\{b_i\}_{i=\overline{1,n}}$ and $\mathsf{c}=\{c_i\}_{i=\overline{1,n}}$ problems accept values $n=3$; $\mathsf{a}=(0.2;0.6;0.3)$; $\mathsf{b}=(4;18;12)$; $\mathsf{c}=(3;1;2)$;
3) To calculate optimum volumes of the purchase of raw materials at various values of financial resources: $FR=3000\,LVL$; $1200\,LVL$; $3600\,LVL$, when $\mathsf{a}=(0.2;0.6;0.3)$; $\mathsf{b}=(4;18;12)$; $\mathsf{c}=(3;1;2)$.

3. The Mathematical Basis of the Consideration Problem and its Relationship with the Leontiev's`Model

Questions of the use and distribution of material and manpower are extremely important, as decisions regarding these in many respects define the production efficiency. Research into some problems connected with a given circle of questions is possible by means of the theory of the interbranch balance of Leontiev (for instance see [1, 2]) (V. V. Leontiev is the Nobel Prize Winner for Economics in 1973 for developing the method of "expense-release" and for its application to important economic problems).

To construct a mathematical model of the problem, formulated in section 2, we shall take some important facts from the Leontiev's model (for instance see [1–5]). We shall compare every number $l_i \geq 0$ $\left(i=\overline{1,n}\right)$ of the i $\left(i=\overline{1,n}\right)$ branch, expressing necessary expenses of material and manpower at an individual intensity of given technological process (i.e. branches). Depending on the purpose of modelling, the number i $\left(i=\overline{1,n}\right)$ can be measured either in units of quantity of used material resources, or in man-hours.

If it is possible to designate a vector of used material and/or manpower through $\mathsf{l}=(l_1,...,l_n)$ it is possible to present technology in a modified Leontiev's model in the pair (l, A) where the matrix $A=\{a_{ij}\}_{i,j=\overline{1,n}}$ describes technology at an individual intensity of work of all branches, and is the main matrix of the Leontiev model

$$\mathsf{x}-\mathsf{x}A=\mathsf{c}, \tag{1}$$

where $\mathsf{x}=(x_1,...,x_n)$ is a vector of intensity (total release), and x_i $\left(i=\overline{1,n}\right)$ means volume of total output i $\left(i=\overline{1,n}\right)$ of the branch (i.e. intensity i of the branch); c is the unused material and/or manpower.

Let us notice that in Leontiev's model (1), the basic question is the optimum planning of manufacture for the period $[T_s, T_e]$. However, as is visible from system/model (1), this basic question is formulated in the contrary way: at the set vector c of the final consumption, it is required to define a necessary vector of total release. In other words, in Leontiev's model (1) it is necessary to solve the system of equations at the set vector c and matrix A.

By consideration of the statistical variant of Leontiev 's model (1), which does not consider the background of the manufacturing process, in particular, the limitations of material resources, it is natural to assume that the unique limiting factor of manufacture is the total amount $L > 0$ of industrial resources. If the operating mode of the economic system described by Leontiev 's model (1) is a set of a vector of intensity x in a new situation, not any vector of intensity $x \geq 0$, which is also a vector of total release, will be admissible (for instance see [2, 4]). Really, the volume of expenditure of industrial resources (material and labour) at work according to a vector of intensity x is equal to (x, l). Therefore, any non-negative vector x satisfying the inequality is admissible $(x, l) \leq L$.

Thus, it is impossible to bring attention to the question of satisfying any final demand $c \geq 0$, namely, the system decision

$$\begin{cases} x - xA = c, \\ (x, l) \leq L, \end{cases} \tag{2}$$

does not exist at any non-negative vector c. In this connection the important fact is the definition of the structure of a final demand. For this purpose we shall formulate the following extreme problem: let the vector $c \geq 0$ set, just a part of the final demand, not the whole final demand. Then for the optimum plan it is possible to make a task

$$\begin{cases} V \to \max, \\ x - xA \geq Vc, \\ (x, l) \leq L, \\ x \geq 0, \end{cases} \tag{3}$$

that can be interpreted as aspiration to maximise the quantity of let out "complexes" c.

The maintenance of the problem (3) is the rational distribution of industrial resources. Thus, if the technological matrix A is productive, the Leontiev's model is more exactly productive (it means that the non-negative vector of intensity x exists for any vector $c \geq 0$ of a final demand (for instance see [1, 3, 4]) the problem (3) is admissible. In reality, we take, for example, $V = 1$. Owing to the efficiency of the Leontiev's model (we often speak simply about efficiency of a technological matrix A in (1)), the system of equations (1) has the non-negative decision $x_{opt} \geq 0$. Then, as $L > 0$, there is such $\lambda > 0$ that $(\lambda x_{opt}, l) < L$. Hence $(\lambda x_{opt}, l)$ is admissible for the problem (3). Then the fact that the set of all admissible vectors is limited and closed (i.e. a compact set), we understand that problem (3) has the decision. However, it is important to underline that problem (3), without raising duality theories, is not a trivial problem, and finding its decision through direct methods is problematic enough. With that end, in view (the purpose of finding the decision of the mathematical statement (3)) several further facts from the theory of the duality of problems of linear programming are mentioned.

Dual problem to a direct problem (3) is the following problem (for instance see [4, 7]):

$$\begin{cases} L \cdot q \to \min, \\ lq \geq (I - A) \cdot p, \\ (c, p) \geq 1, \\ p \geq 0, \;\; q \geq 0, \end{cases} \tag{4}$$

where $I \in \mathbb{R}^1(n, n)$ is an individual matrix; $p \in \mathbb{R}_+^n$; $q \in \mathbb{R}_+^1$.

In the dual problem (4), vector p corresponds to the first restriction of the direct problem (3), the number q corresponds to the second restriction of problem (3) and, according to the general theory of linear programming, they are interpreted as objectively caused estimations of vector products and $c \geq 0$ total industrial (material and labour) resources L.

If a technological matrix A is indivisible (for instance see [7]), any vector x entering into the decision (x, V) of a direct problem (3) is strictly positive $x \geq 0$. Really, from the efficiency of the matrix and A it indivisibly follows that $(I - A)^{-1} > 0$. Then from the first restriction $x - xA \geq Vc$ of problem (3) it follows that

$$x \geq Vc(I - A)^{-1}. \tag{5}$$

From (3.2.5) it directly follows that at $c \geq 0$, $c \not\equiv 0$, it is $x \geq 0$.

We have already shown that in the decision (x, V) of the direct problem (3) $V > 0$. It appears from this under the theorem of duality mentioned below the result follows:

$$\begin{cases} lq = (I - A)\mathrm{p}, \\ (c, \mathrm{p}) = 1. \end{cases} \tag{6}$$

Theorem of duality *([7]). If a direct problem*

$$\begin{cases} (\bar{c}, \bar{x}) \to \max, \\ \bar{x}\bar{A} \leq \bar{\mathrm{b}}, \\ \bar{x} \geq 0 \end{cases} \tag{7}$$

and a dual problem to it

$$\begin{cases} (\bar{b}, \bar{\mathrm{y}}) \to \min, \\ \bar{A}\bar{\mathrm{y}} \geq \bar{\mathrm{c}}, \\ \bar{\mathrm{y}} \geq 0 \end{cases} \tag{8}$$

are admissible (remembering that a problem of linear programming is called admissible if the set of the vectors satisfying essential restrictions and restrictions on a sign of this problem, is not empty) (7) and (8) has optimum basic decisions and $\bar{x}_{opt}$ *and* $\bar{y}_{opt}$ *accordingly, and* $(\bar{c}, \bar{x}_{opt}) = (\bar{b}, \bar{y}_{opt})$. *Besides,*

if $\left(\bar{x}_{opt}\bar{A}\right)_i < \bar{b}_i$ *as soon as* $\left(\bar{y}_{opt}\right)_i = 0$;

if $\left(\bar{A}\bar{y}_{opt}\right)_j > \bar{c}_j$ *as soon as* $\left(\bar{x}_{opt}\right)_j = 0$.

The duality converse theorem (for instance see [7]) takes place also:

Duality converse theorem. *If* $\bar{x}_{opt}$ *and* $\bar{y}_{opt}$ *satisfy the restrictions of (to essential restrictions and restrictions on a sign) problems (7) and (8) accordingly, and*

if $\left(\bar{x}_{opt}\right)_i = 0$ *as soon as* $\left(\bar{A}\bar{y}_{opt}\right)_j > \bar{c}_i$;

if $\left(\bar{y}_{opt}\right)_j = 0$ *as soon as* $\left(\bar{x}_{opt}\bar{A}\right)_i < \bar{b}_i$

then the vector $\bar{x}_{opt}$ *is an optimum basic decision of a direct problem (7) and the vector* $\bar{y}_{opt}$ *is an optimum basic decision of a dual problem (8).*

So, we shall return to result (6): we get that

$$p = q \cdot (I - A)^{-1} l, \tag{9}$$

and substituting this value of a vector p in $(c,p)=1,$ we get

$$q=\frac{1}{\left((I-A)^{-1}l,c\right)}. \tag{10}$$

Having united (9) and (10), we get the definitive expression for the vector p:

$$p=\frac{(I-A)^{-1}l}{\left((I-A)^{-1}l,c\right)}. \tag{11}$$

Now we shall return to system (2): if the problem (2) has the decision $x,$ then

$$x=c(I-A)^{-1}. \tag{12}$$

As has already been noted above, industrial expenses are expressed by number $(x,l),$ which taking into account (12), looks like $(x,l)=\left(c(I-A)^{-1},l\right)=\left(c,(I-A)l\right).$

Thus, the industrial expenses necessary for the release of a vector of a final demand c are equal $\left(c,(I-A)l\right).$ According to this the vector $l_{total}=(I-A)l$ expresses the vector of the full industrial expenses whose i-th coordinate describes full industrial expenses of the i-th technological process (i.e. i-th branch).

If we interpret the vector p as a vector of the prices for products, and the number q as a salary based on the person-hour, the dual problem (4) consists of defining these sizes to minimise the general wage funds, provided that the net profit $p_i-(a_i,p)$ of the i-th technological process was not positive where through a_i $\left(i=\overline{1,n}\right)$ is designated i-th $\left(i=\overline{1,n}\right)$ line of a technological matrix $A.$

The given treatment of a problem (4) is resulted in baccalaureate work to show importance and some kind of universality of a problem investigated in baccalaureate work. However in the given baccalaureate work we shall not deal with this problem, but suggest an interesting and important interpretation. This could be the theme of separate research. In sense of interests and importance one more interpretation of a problem (4) is mentioned.

As appears from the duality theory (see the duality theorems), decisions (x,V) and (p,q) self-duality problems (3) and (4) respectively satisfy to the equation:

$$V=qL. \tag{13}$$

As $(c,p)=1,$ the volume V is equal to the total cost of the vector production Vc at the prices $p.$ If positive components of the vector c correspond to products of consumer demand, the equation (13) in this case expresses equality of supply and demand in cost expression, namely, the price of the produced volume of the end production is equal to the total sum of the money received by the people who made this production, as a salary.

If it is possible to designate a vector of full labour expenses through $l_{total},$ where $l_{total}=(I-A)^{-1}l,$ and if it is possible to designate $\gamma=1/\left(l_{total},c\right),$ then the equality (11) it is possible to copy in a kind $p=\gamma l_{total}.$

The basic conclusion, concerning the given interpretation of the problem/model (4), consists of the following: the vector of the prices for production is directly proportional to the vector of full labour expenses. This conclusion is widely known in economics. Many outstanding scientists and economists have considered similar models, which include in explicit form factors of expenses of live work (for instance see [8] and the list of corresponding literature).

4. Mathematical Statement of the Consideration Model

In this paragraph, after necessary preparation in the form of the facts and concepts stated in the section 3, we pass to the direct working out of the mathematical model of the problem formulated in the section 2.

So, using the approaches determined in the section 3, methodology, the facts and results, we can design the following mathematical model which is similar to the models (3) of the previous section:

$$\begin{cases} W\left(x\right) = V \to \max; \\ \begin{cases} x_i - a_i \cdot V \geq -b_i \quad \forall i = \overline{1,n}; \\ \sum_{i=1}^{n} c_i \cdot x_i = FR; \end{cases} \\ x_i \geq 0 \quad \forall i = \overline{1,n}, \end{cases} \tag{14}$$

where through the V planned volume of output of a product is designated; through x_i $\left(i = \overline{1,n}\right)$ the volume of purchase of the i-th $\left(i = \overline{1,n}\right)$ raw materials is designated; other parameters $\left\{a_i,\ c_i,\ b_i\left(i = \overline{1,n}\right);\ FR;\ n\right\}$ are the initial data and describe the same economic indicators, as in the section 2 of qualitative model.

4.1. Solution of the Constructed Model

Now we pass to the decision of the mathematical statement (14) investigated problems. The decision of a problem (14) we name such $(n+1)$-th size vector $(V^*, x_1^*, \ldots, x_n^*)$, which corresponds to all restrictions (14) and provides a maximum of value V^* among all possible values of criterion function $W\left(x\right)$ among all possible volumes of let out grain production.

Firstly, we shall construct dual to problem (14). For this purpose, we shall copy the problem directly (i.e. initially) (14) in the following form:

$$\begin{cases} W\left(x\right) = FR \cdot V + 0 \cdot x_1 + \ldots + 0 \cdot x_n \to \max; \\ \begin{cases} a_1 \cdot V - x_1 \leq b_1; \\ a_2 \cdot V - x_2 \leq b_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n \cdot V - x_n \leq b_n; \\ c_1 \cdot x_1 + c_2 \cdot x_2 + \ldots + c_n \cdot x_n = FR; \end{cases} \quad \begin{array}{|l} y_1 \\ y_2 \\ \ldots \\ y_n \\ y_{n+1} \end{array} \\ V \in R^1;\ x_i \geq 0,\ i = \overline{1,n}, \end{cases} \tag{15}$$

where through y_j $\left(j = \overline{1, n+1}\right)$ the dual variables are designated.

Now it is obvious that to problem (15) the following problem of linear programming is dual:

$$\begin{cases} \overline{W}\left(x\right) = \sum_{j=1}^{n} b_j \cdot y_j + FR \cdot y_{n+1} \to \min; \\ \sum_{j=1}^{n} a_j \cdot y_j = FR; \\ 0 \leq y_j \leq c_j \cdot y_{n+1},\ j = \overline{1,n}; \\ y_{n+1} \in R^1. \end{cases} \tag{16}$$

We should find out whether the initial problem has (14) decision, and if the answer is positive to find it. In other words, we should find out:

1) whether the set of admissible decisions of problem (14) is not empty;

2) whether the criterion function on $W(x) = V$ this set of admissible decisions if it is not empty and limited from above $W(x) = V$;

3) in the case of affirmative replies to the previous two points, to find the decision $(V^*, x_1^*, ..., x_n^*)$.

To answer these questions, at first from (14) we notice that if there is a decision $(V^*, x_1^*, ..., x_n^*)$ it should be carried out

$$\begin{cases} x_i^* \ge a_i \cdot V^* - b_i \quad \forall i = \overline{1,n}; \\ x_i^* \ge 0 \quad \forall i = \overline{1,n}, \end{cases} \tag{17}$$

Let us show that in these obligatory conditions for performance (17) strict inequalities cannot take place. Really, we shall assume the opposite, i.e. we shall assume that there is at least a number $i = i_0 \in \{1, 2, ..., n\}$, for which

$$\begin{cases} x_{i_0}^* > a_{i_0} \cdot V^* - b_{i_0}; \\ x_{i_0}^* > 0. \end{cases} \tag{18}$$

From the theorems of duality mentioned in the previous paragraph, it is clear that mutually dual problems (14) and (16) can either have a simultaneous decision or have no simultaneous decision. Therefore, assuming that the vector $\left(V^*, x_1^*, ..., x_n^*\right)$ is the decision of a direct problem (14) by default we assume that there is a vector $\left(y_1^*, y_2^*, ..., y_{n+1}^*\right)$, which is the decision of a dual problem (16). Furthermore, the following useful theorem from the duality theory is required.

Theorem. *If only one of the sets of admissible decisions of mutually dual problems*

$$\begin{cases} <\overline{c}, \overline{x}> = \sum_{i=1}^{n} \overline{c_i} \cdot \overline{x_i} \to \max; \\ \begin{cases} \sum_{j=1}^{n} \overline{a_{ij}} \cdot \overline{x_j} \le \overline{b_i}, \quad i = \overline{1,k}; \\ \sum_{j=1}^{n} \overline{a_{ij}} \cdot \overline{x_j} = \overline{b_i}, \quad i = \overline{k+1,m}; \end{cases} \\ \overline{x_j} \ge 0, \quad j = \overline{1,s}, \end{cases} \qquad \begin{cases} <\overline{b}, \overline{y}> = \sum_{i=1}^{m} \overline{b_i} \cdot \overline{y_i} \to \min; \\ \begin{cases} \sum_{j=1}^{m} \overline{a_{ij}} \cdot \overline{y_j} \ge \overline{c_j}, \quad j = \overline{1,s}; \\ \sum_{j=1}^{m} \overline{a_{ij}} \cdot \overline{y_i} = \overline{c_j}, \quad j = \overline{s+1,n}; \end{cases} \\ \overline{y_i} \ge 0, \quad i = \overline{1,k}; \\ \overline{y_i} \in R^1, \quad i = \overline{k+1,n} \end{cases}$$

is not empty, then

1) if $\overline{X} \ne \emptyset, \overline{Y} = \emptyset \Rightarrow \sup_{\overline{X}} <\overline{c}, \overline{x}> = \infty;$

2) if $\overline{X} = \emptyset, \overline{Y} \ne \emptyset \Rightarrow \inf_{\overline{Y}} <\overline{b}, \overline{y}> = -\infty,$

where the $\overline{X}$ *is a set of admissible decisions of a direct problem;* $\overline{Y}$ *is a set of admissible decisions of a dual problem.*

Let's return to the assumption (18). From (18) and the just formulated theorem it follows that

$$\begin{cases} y_{j_0}^* = c_{j_0} \cdot y_{n+1}^*; \\ y_{j_0}^* = 0. \end{cases}$$

Therefore, $y_{n+1}^* = 0$ and as appears from the condition $0 \le y_j \le c_j \cdot y_{n+1}$, $j = \overline{1,n}$ of the problem (16), we can write $y_1^* = y_2^* = \ldots = y_n^* = y_{n+1}^* = 0$. It is impossible, for in this case from the equation $\sum_{j=1}^{n} a_j \cdot y_j = FR$, which is essential restriction of the dual problem (16), followed that $FR \equiv 0$, but this result contradicts a condition of statement of an investigated problem. Hence, in (17) there will be no number for which an inequality would be strictly carried out. Therefore, there is only a case of performance of conditions (17) as equality. It means that

$$x_i^* \ge \max\left\{0;\ a_i \cdot V^* - b_i\right\}\ \ \forall i = \overline{1,n}. \tag{19}$$

Thus the formula (19) finds the optimum values of a required vector $X = \left(x_1, \ldots, x_n\right)^T$. Now it is necessary to find optimum (maximum) volume V^*. For this purpose we shall substitute in (19) the essential restriction $\sum_{i=1}^{n} c_i \cdot x_i = FR$ of the initial problems (14):

$$\sum_{i=1}^{n} c_i \cdot \max\left\{0, a_i \cdot V^* - b_i\right\} = FR. \tag{20}$$

We have received the equation for the definition of required volume V^*, i.e. the required volume V^* is a root of the equation (20) if only the root exists and is unique.

To find out a question of existence of uniqueness of an initial root V^* in (20) we shall designate

$$\alpha_i \overset{def}{\equiv} \frac{b_i}{a_i},\ \ a_i \ne 0,\ \left(i = \overline{1,n}\right) \tag{21}$$

Without loss of generality it is possible to assume that the sequence $\left\{\alpha_i\right\}_{i=\overline{1,n}}$ is not an increasing sequence: $\alpha_1 \le \alpha_2 \le \ldots \le \alpha_n$. Really, if it is not valid, then from the sequence $\left\{b_i\right\}_{i=\overline{1,n}}$ we can make up new sequence $\left\{\overline{b_i}\right\}_{i=\overline{1,n}} : \overline{b_1} \le \overline{b_2} \le \ldots \le \overline{b_n}$, by rearrangement of terms of the old sequence. Similarly, from the sequence $\left\{a_i\right\}_{i=\overline{1,n}}$ we can make up new sequence $\left\{\overline{a_i}\right\}_{i=\overline{1,n}} : \overline{a_1} \ge \overline{a_2} \ge \ldots \ge \overline{a_n}$ by means of designations $\alpha_i \overset{def}{\equiv} \frac{\overline{b_i}}{\overline{a_i}},\ \overline{a_i} \ne 0,\ \left(i = \overline{1,n}\right)$. Obtained sequences $\left\{b_i\right\}_{i=\overline{1,n}}$ and $\left\{a_i\right\}_{i=\overline{1,n}}$ are non-increasing sequences.

Considering the designation (21) in the equation (20) we shall receive

$$\sum_{i=1}^{n} a_i \cdot c_i \cdot \max\left\{0, V^* - \alpha_i\right\} = FR. \tag{22}$$

The left member of the equation (22) is a continuous function concerning argument V^*, and this function as an argument function V^* strictly increases on the half-interval $\left[\alpha_1; +\infty\right)$, accepting values from 0 to $+\infty$. Hence, the root V^* of the equation (22) exists and is unique. To find this root that exists and is unique, we should define the number at first m $(m \le n)$. It is obvious that the required number m can be defined from a condition

$$m = \max\left\{ i \cdot \sum_{j=1}^{n} c_j \cdot \max\left\{0; a_j \cdot \alpha_i - b_j\right\} < FR,\ \ j = \overline{1,n} \right\}. \tag{23}$$

Since $\sum_{j=1}^{n} c_j \cdot \max\left\{0; a_j \cdot \alpha_{n+1} - b_j\right\} \ge FR,$ then we have the following:

1) if $m < n$ then $V^* \in \left(\alpha_m; \alpha_{m+1}\right]$;

2) if $m = n$ then $V^* \in \left(\alpha_m; +\infty\right)$ and $\alpha_{m+1} = \infty.$

Consequently, on the half-interval $\left(\alpha_m; \alpha_{m+1}\right]$ the right-hand side of the equation (20) has the following appearance:

$$\sum_{i=1}^{m} c_i \cdot \left(a_i \cdot V^* - b_i\right),$$

and therefore the equation (20) on the half-interval $V^* \in \left(\alpha_m; \alpha_{m+1}\right]$ accepts the more simple form

$$\sum_{i=1}^{m} c_i \cdot \left(a_i \cdot V^* - b_i\right) = FR.$$

It appears from this we have received

$$V^* \cdot \sum_{i=1}^{m} c_i \cdot a_i = FR + \sum_{i=1}^{m} c_i \cdot b_i,$$

i.e. the required maximum volume of output of grain production V^* is calculated by the formula

$$V^* = \frac{\sum_{i=1}^{m} c_i \cdot b_i + FR}{\sum_{i=1}^{m} c_i \cdot a_i}. \tag{24}$$

Considering (24) in (19), we shall receive definitive expression for the definition of the required optimum volumes of the purchase of raw materials x_i^* $\left(i = \overline{1,n}\right)$:

$$x_i^* = \begin{cases} a_i \cdot \dfrac{\sum_{j=1}^{m} c_j \cdot b_j + FR}{\sum_{j=1}^{m} c_j \cdot a_j} - b_i, \ \ \forall i = \overline{1,m} \\ 0, \ \ \forall i = \overline{m+1,n}. \end{cases} \tag{25}$$

Thus, the formulas (23)–(25) completely define the optimum decision of the assigned problem (14). In the section 2, when a verbal statement of an investigated problem was formulated, it was required to solve more two additional problems, namely, the problems 2) and 3). Now, after having definitive results in the form of (23)–(25), without much difficulty we can solve these additional problems "manually".

4.2. Solution of the Subproblems 2) and 3) from the Section 2

Using the formula (24) we shall find dependence of optimum volume V^* on financial resources FR at the initial data $n = 3$; $a = (0.2; 0.6; 0.3)$; $b = (4; 18; 12)$; $c = (3; 1; 2)$:

$$V^*(FR) = \begin{cases} \dfrac{12 + FR}{0,6}, & если \quad m = 1; \\ \dfrac{30 + FR}{1,2}, & если \quad m = 2; \\ \dfrac{54 + FR}{1,8}, & если \quad m = 3. \end{cases}$$

Apparently, from this formula, the dependence of optimum output V^* on available financial resources FR for each value m $(m \le n)$ is linear and, consequently, for all values $1 \le m \le n$ this dependence is piecewise linear. For our initial data the following output-resource curve dependence detailed on Figure 1:

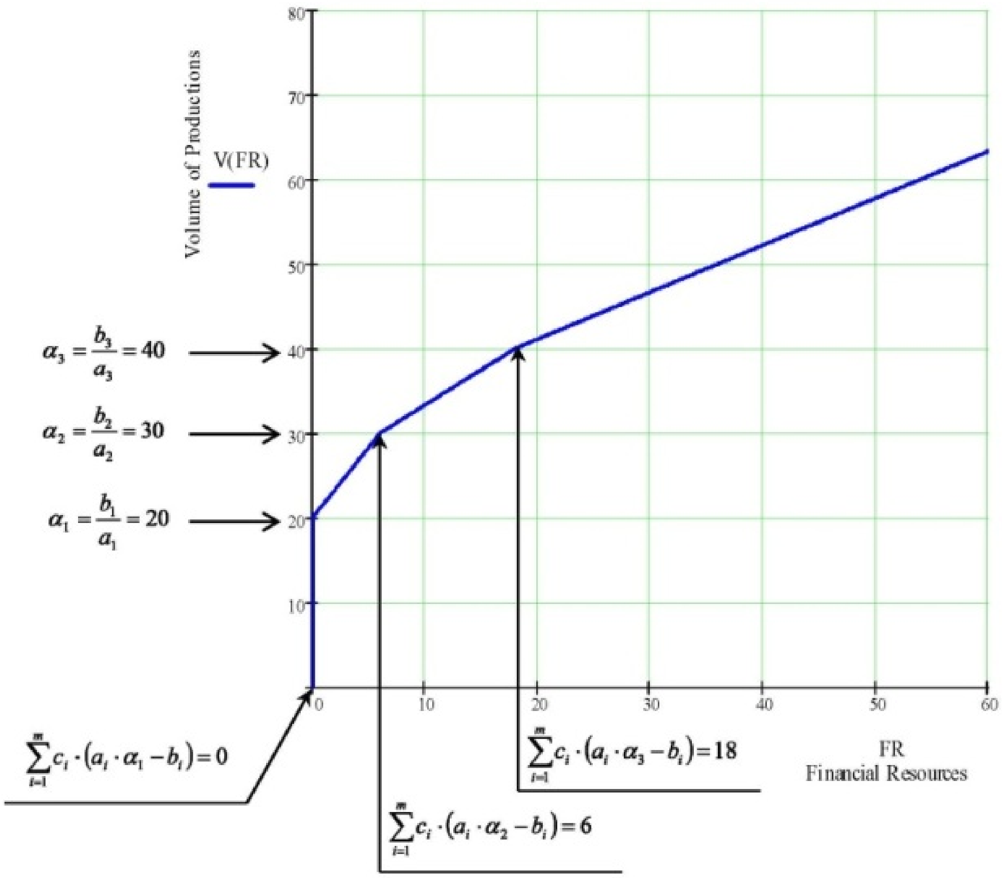


Figure 1. The curve dependence of optimum volume and financial resources

It is important to stress that the software product is developed for a computer decision of the problem investigated in the given work. The results of the application of the developed software product for the decision of problems 1) and 2) from the section 2 are given on Figures 2–5.

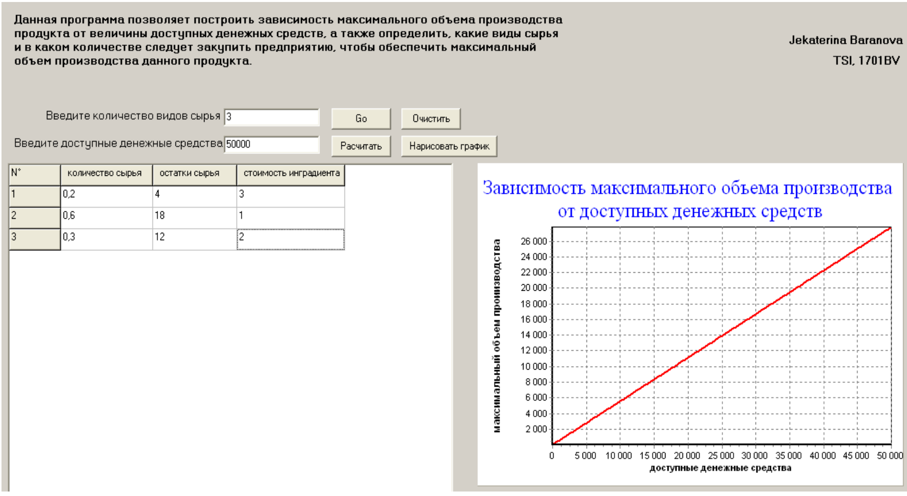


Figure 2. Software product for definition of optimum volume under various financial resources

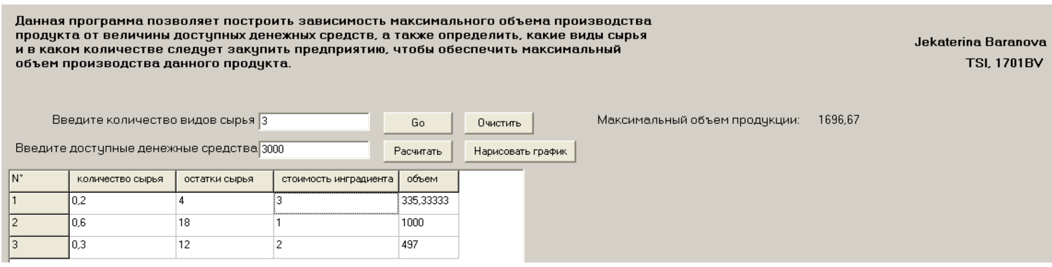


Figure 3. Optimum volume of purchase of raw materials at financial resources $FR = 3000\ LVL$

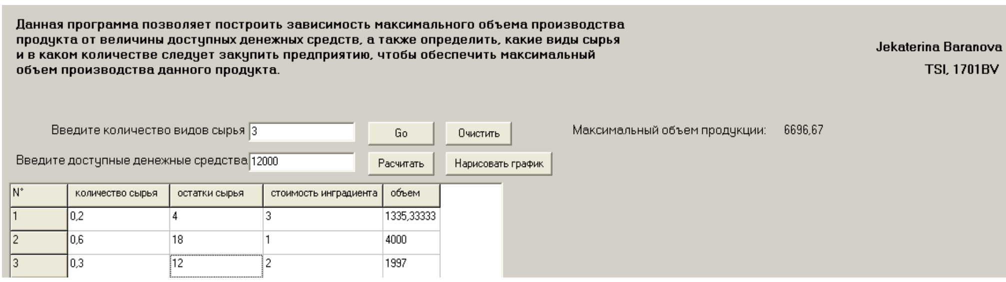


Figure 4. Optimum volume of purchase of raw materials at financial resources $FR = 12000\ LVL$

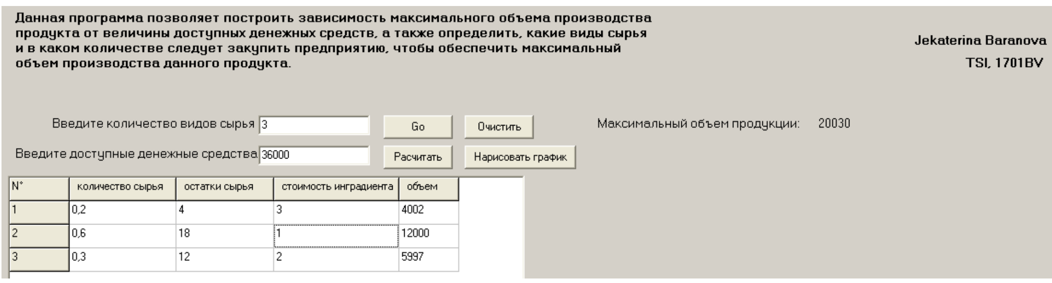


Figure 5. Optimum volume of purchase of raw materials at financial resources : $FR = 36000\ LVL$

5. Conclusions

In this paper the research and the analysis of the bread market in Latvia is conducted. The determining factor that influences the level of the price of bread is revealed as the volume of harvested grain and its price, not only in Latvia, but also throughout the European Union. The tendency of traditional products such as the white loaf and rye bread to dominate in this market has been noticed, but there could be demand for a healthy diet product; grey bread or bread with bran or every possible additive. The leader in this market – '*Hanzasmaiznica*' is revealed. The characteristic of the company, its activity, and the basic industrial indicators are given, and also an analysis of the financial performance of this company for 2006–2008 is conducted. In general, the company analysis shows that the dynamics of profitability of the '*Hanzasmaiznica*' business in 2006–2008 are positive, and growth in practically all analysed indicators is observed.

The mathematical model of the task in view is developed and its analytical decision is found. The analytical method defines the dependence of optimum output on financial resources. Besides, the formula is found in the closed form, allowing a definition of the optimum volume of purchase of additional raw materials at various values of financial resources.

A software product considerably facilitates the decision of the investigated problem of optimum planning of the manufacture of divisible production (computer realisations of the constructed model have been developed for not only grain products). By means of the developed software product numerical experiments have been conducted, and the high precision analysis of the received numerical results coincides with the results found by means of developed analytical formulas.

References

1. Leontyev, V. V. *General Economic Problems of Interbranch Analysis. Vol. 1.* (Selected publications in 3 volumes). Moscow: Economics, 2006. 407 p. (In Russian)
2. Leontyev, V. V. *Special Researches on the Basis of 'Costs-Production' Methodology. Vol. 2.* (Selected publications in 3 volumes). Moscow: Economics, 2006. 543 p. (In Russian)
3. Leontyev, V. V. *Research of American Economics Structure: Theoretical and Empirical Analysis by Costs-Production Scheme.* Moscow: Gostekhizdat, 1958. 640 p. (In Russian)
4. Ashmanov, S. A. *Mathematical Models and Methods in Economics*. Moscow: MGU Edition, 1980. 201 p. (In Russian)
5. Marvin, Yu. P. *Mathematical Methods and Models in Economics*. Moscow: Higher School, 2007. 424 p. (In Russian)
6. Vasin, A. A., Morozov, V. V. *Theory of Games and Models of Mathematical Economics*. Moscow: "MAKS" Press, 2005. 272 p. (In Russian)
7. Sukharev, A. G., Tomokhov, A. V., Fedorov, V. V. *Optimization Methods Training Course*. Moscow: Physmathlit, 2005. 368 p. (In Russian)

Received on the 25th of March, 2011

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 53–58
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

PROJECTOR SYSTEMS

A. N. Kovantsov

Riga Technical University,Department of Engineering Mathematics
Kalku Str. 1, Riga LV-1658, Latvia
Phone: 67089333. E-mail: info@rtu.lv

Systems of evolution differential equations are applied in various problems of imitation modelling. But any computational solution needs the analytical evaluations. In this connection the projector technique is very effective. The solution of the system of the nonlinear differential equations considerably can be simplified, if the system is transformed to Jordan form, i.e. to such type, when the system consists of blocks of independent system of functions.

Keywords: *nonlinear differential equations, projector technique, Jordan matrix*

1. Introduction

Both in the theory and in many practical questions exclusively important role is played with problems of the decision of systems of the ordinary equations. One of such problems consists in splitting the given system of the equations on blocks, i.e. in such subsystems, each of which contains smaller number of the unknown functions entering simultaneously and under a sign of a derivative, and in the right parts.

Abundantly clear, that integration of such subsystems is a simpler problem, than integration of initial system. Therefore there is a question on a finding of ways of transformation of such systems to the split kind.

Such problem has been solved for a case of linear systems of the differential equations. It has appeared that splitting of such systems to equivalently reduction of a square matrix to Jordan to the form. To each cell Jordan there corresponds the certain block of the split system.

One of such ways of splitting has been offered to K. G. Valeev [1]. It has been entered into consideration special matrixes – linear projectors [2]. Also the iterative way of the approached finding of two supplementing each other up to nonsingular matrixes of projectors from factors of the right parts of the set system of the equations has been found.

That it is available there has been a way approached dichotomy systems. The given way easily enough gave in to machining. However, the geometry of occurring processes at such splitting remained accomplished obscure. Obscure there was also a fact of transferring of such splitting on nonlinear systems.

2. The Projector of the Ordinary Differential Equations

Let the system of the linear differential equations with constant factors (1) is given

$$\frac{dy^i}{dt} = a_j^i \, y^j \, . \tag{1}$$

Having found two projectors of a matrix, system (1) it is possible to split on two blocks then to each of the received blocks to apply the same operation (dichotomy). The further splitting will appear impossible as soon as the spectrum of any matrix will be reduced to one own to number [3].

Let the splitting transformation, leading system to a block kind, translates all Euclid space E_n of dimension n which coordinates of points are values of unknown functions of system (1) in some-dimensional plane which is passing through the beginning of coordinates. Designing occur some projecting plane E_{n-k} .

The equations of such plane are easy for receiving from the formulas defining splitting of system (1). It has appeared, that each projector generates in Euclid space E_n system of the projecting planes forming projecting network of carry at which it is easy to define the first square-law form.

If splitting of system (1) to carry out by means of linear transformations, that each of such transformations is broken into two singular transformations and generates singular matrixes which are projectors of system (1). It is natural to set the task of transferring these statements on a nonlinear case.

By analogy to a linear case it can be made as follows. It has been noticed, that elements of linear projectors are private derivatives of new coordinates on old. We shall consider nonlinear independent system of the differential equations

$$\frac{dy^i}{dt} = f^i(y^j), \quad f^i(0,...,0) = 0. \tag{2}$$

Let's consider nonlinear transformations:

$$y^i = \sum_{k=1}^{n} \Psi_k{}^i(z^k), \quad \Psi_i^i(z^i) = z^i, \quad \Psi_k{}^i = 0 . \tag{3}$$

Making matrixes (already functional) which elements are private derivatives of new coordinates on old and demanding that these matrixes satisfied to conditions of projectors, we shall come to system of the differential equations in private derivatives of the first order on function costing in the right parts of equality (2).

Let's make it so. The split system is presented by the equations.

$$\frac{dz^i}{dt} = \mathrm{X}^i(z^j) \quad \mathrm{X}^i(0) = 0 . \tag{4}$$

Differentiating (3) on and considering (4), we receive a general view of functions f^i in variables z^j:

$$f^i = \alpha_k{}^i \mathrm{X}^k, \quad \alpha^i{}_k = \Psi'^i{}_k(z^k), \quad \alpha^I{}_I = 1. \tag{5}$$

Equality (5) shows that functions f^i depend on n^2 functions of one argument $(n-1)n$ – functions (have considered equality) and functions. Having solved the equations (3) rather and having brought the found values in the right parts of equality (5). Let's find a general view of functions in variables. Differentiating (5) on and accepting considering that according to (4), we shall receive the following

$$\frac{\partial y^i}{\partial z^k} = \alpha_k{}^i ,$$

let's receive

$$\frac{\partial f^i}{\partial y^j}\alpha^j{}_k = \alpha'^i_k \mathrm{X}^k + \alpha^i{}_k \mathrm{X}'^k . \tag{6}$$

We differentiate (7) on (in this case $\alpha^j{}_k, \alpha'^i_k, \mathrm{X}^k, \mathrm{X}'^k$ behave as constants):

$$\frac{\partial^2 f^i}{\partial y^j \partial y^l}\alpha^j{}_k \alpha^l{}_p = 0 . \tag{7}$$

Let $\dfrac{\partial^2 f^i}{\partial y^j \partial y^l} = a^i{}_{jl} = a^i{}_{lj}$.

At fixed y^i the equations $a^i{}_{jl} x^j x^l = 0$ define n hyper-quadric in $n(n-1)$ measured projective space P_{n-1}. Equation (7), i.e. $a^i{}_{jl}\alpha^j{}_k \alpha^l{}_p = 0$ means a polar interlinking of points $M_k(\alpha^j{}_k), M_p(\alpha^l{}_p)$ concerning mentioned hyper-quadric. To find the solution of system (7) is means to find the general auto-polar reference point concerning data hyper-quadric.

This reference point should be non-degenerative, as it is offered transformation (4), $\left|\dfrac{\partial y^i}{\partial z^k}\right| = \left|\alpha^i{}_k\right| \neq 0$. If to carry space P_{n-1} to this reference point all the equations hyper-quadric will accept an initial kind $\alpha'^i_{ii}(x'^j) = 0$.

Hence, the problem (the decision of the equations (7)) is reduced to simultaneous reduction of n square-law forms $a^i{}_{jl} x^j x^l$ an initial kind. It is known, however, that this problem is possible only for two square-law forms, and one of these two forms should be positively certain. Let the form $\varphi^1 = a_{jl}{}^1 x^j x^l$ is positively certain (we can assume, that those it is in a point (0,0..., 0). Then by virtue of a prospective continuity of functions $\dfrac{\partial^2 f^i}{\partial y^i \partial y^l}$ it will be positively certain and in a vicinity of the specified point).

Let's take any second form, for example $\varphi^2 = a_{jl}{}^2 x^j x^l$. Let $\alpha_1{}^i$ – one of tops of a required general auto-polar reference point hyper-quadric $\varphi^1 = 0, \quad \varphi^2 = 0.$ Then polar hyper-planes of this point concerning specified hyper-quadric $a_{jl}{}^1 \alpha_1{}^j x^l = 0, \quad a_{jl}{}^2 \alpha_1{}^j x^l = 0$ should coincide

$$\frac{a_{jl}{}^1 \alpha_1{}^j}{a_{j2}{}^2 \alpha_1{}^j} = \frac{a_{j2}{}^1 \alpha_1{}^j}{a_{j2}{}^2 \alpha_1{}^j} = \dots = \frac{a_{jn}{}^1 \alpha_1{}^j}{a_{jn}{}^2 \alpha_1{}^j} = S .$$

(Through the S size of the general attitude) is designated. From here $(a_{jl}{}^1 - s a_{jl}{}^2)\alpha_1{}^j = 0$. As all $\alpha_1{}^j$ simultaneously are not equalled to zero,

$$\left| a_{jl}{}^1 - S a_{jl}{}^2 \right| = 0 . \tag{8}$$

Let's name this equation the characterize equation of two linear operators $x'_j = a_{jl}{}^1 x^l, \quad x'' = a_{jl}{}^2 x^l$. If $x'_j = S x''_j$, x^l refers to as a double vector of linear operators. For definition of double operators we have system of the linear equations

$$(a_{jl}{}^1 - S a_{jl}{}^2) x^l = 0. \tag{9}$$

From here, excepting from consideration zero-vectors, we come to the characteristic equation (8). Generally set of such vectors forms so-called Jacoby variety of two hyper-quadric $\varphi^1 = \varphi^2 = 0$. In a case considered by us Jacoby the variety represents - (reference point). Obviously, each simple S_i root of the characteristic equation (8) there corresponds only one point $\alpha^l{}_i$ defined by the equations $(a^1{}_{jl} - S_i a^2{}_{jl})\alpha_i{}^l = 0$.

Assuming, that all roots of the characteristic equation are various, we shall receive a unique general auto-polar reference point with tops $\alpha_p{}^l$ concerning two quadric $\alpha_p{}^l = \alpha_p{}^l(y^j)$. At presence of multiple roots of the characteristic equation last decision will be not the only thing.

In case of when any of square-law forms $a_{jl}{}^i x^j x^l$ is not positively certain, presence of the general auto-polar reference point at them also is possible, but at performance of some additional conditions. Let (13) – some general reference point of hyper-surfaces $\varphi^1 = \varphi^2 = 0$ (unique if roots of the characteristic equation are various). As coordinates $\alpha_p{}^l$ it turns out to within the general multipliers we shall choose these multipliers so that $\alpha_i{}^i = 1$.

Each of points $\alpha_p{}^l$ should depend only from z^p that is why differentiation $\alpha_p{}^l = \alpha_p{}^l(y^j)$ from z^q on gives a zero:

$$\frac{\partial \alpha_p{}^l}{\partial y^j} \alpha_q{}^j = 0. \tag{10}$$

To such conditions should satisfy functions f^i that the system (1) supposed projector splitting. However these conditions generally are not sufficient. Additional conditions we shall find from (6). Let $\left(\beta_i{}^k\right)$ – a matrix, return to a matrix $\left(\alpha_k{}^i\right)$: $\beta_l{}^i \alpha_i{}^k = \delta_l{}^k, \quad (\beta_l{}^i = \dfrac{\partial z^i}{\partial y^l})$.

Then from (6) it is received

$$\mathrm{X}^k = \beta_i{}^k f^i .$$

Replacing here an index on an index and differentiating on z^q $(q \neq p)$, we shall receive

$$\left(\frac{\partial \beta_i{}^p}{\partial y^l} + \beta_i{}^p \frac{\partial f^i}{\partial y^l}\right)\alpha_q{}^l = 0. \tag{11}$$

This equation should be attached to equation (10).

Theorem 1. Let the system of the differential equations is given

$$\frac{dy^i}{dt} = f^i(y^j)\ . \tag{12}$$

Let for functions conditions (10) and (11) where $\alpha_k{}^i$ are under formulas (5), and $(\beta_i{}^k)$ – a matrix, return to a matrix $(\alpha_k{}^i)$ are satisfied. Then splitting of system (12) can be found by means of quadratures.

The proof.

Really

$$\alpha_p{}^l = \psi'_p{}^l = \frac{\partial \psi_p{}^l}{\partial y^i}\alpha_p{}^i, \quad (\psi'_p{}^l = \frac{d\psi_p{}^l}{dz^q}), \tag{13}$$

$$0 = \frac{d\psi_p{}^l}{dz^q} = \frac{\partial \psi_p{}^l}{\partial y^i}\alpha_q{}^i, \quad (\frac{d\psi_p{}^l}{dz^q} = 0,\ p \neq q)\ .$$

They are the n algebraic equations from unknown $\frac{\partial \psi_p{}^l}{\partial y^i}$ (p, l – are fixed). The system can be written down in the form of:

$$\alpha_q{}^i \frac{\partial \psi_p{}^l}{\partial y^i} = \delta_p{}^q \alpha_q{}^l\ . \tag{14}$$

Then, $\frac{\partial \psi_p{}^l}{\partial y^i} = \delta_p{}^q \alpha_q{}^l \beta_i{}^q$,

or

$$\frac{\partial \psi_p{}^l}{\partial y^i} = \alpha_p{}^l \beta_i{}^p\ . \tag{15}$$

It is easy to see, that the system (21) is quite integrated.

Really

$$I \equiv \frac{\partial^2 \psi_p{}^l}{\partial y^i \partial y^j} - \frac{\partial^2 \psi_p{}^l}{\partial y^j \partial y^i} = \frac{\partial \alpha_p{}^l}{\partial y^j}\beta_i{}^p + \alpha_p{}^l \frac{\partial \beta_i{}^p}{\partial y^j} - \frac{\partial \alpha_p{}^l}{\partial y^i}\beta_j{}^p - \alpha_p{}^l \frac{\partial \beta_j{}^p}{\partial y^i}\ . \tag{16}$$

Differentiating (10), we receive $\frac{\partial \beta_l{}^i}{\partial y^i}\alpha_i{}^k + \beta_l{}^i \frac{\partial \alpha_i{}^k}{\partial y^j} + 0,$ whence $\frac{\partial \beta_l{}^i}{\partial y^j} = -\beta_k{}^i \beta_l{}^h \frac{\partial \alpha_h{}^k}{\partial y^j}$.

Let's bring it in (16):

$$I \equiv \frac{\partial \alpha_p{}^l}{\partial y^j}\beta_i{}^p - \alpha_p{}^l \beta_k{}^p \beta_i{}^h \frac{\partial \alpha_h{}^k}{\partial y^j} - \frac{\partial \alpha_p{}^l}{\partial y^i}\beta_j{}^p + \alpha_p{}^l \beta_k{}^p \beta_j{}^h \frac{\partial \alpha_h{}^k}{\partial y^i}. \tag{17}$$

Let's add to equality (10) equality

$$\frac{\partial \alpha_p{}^l}{\partial y^j}\alpha_p{}^j = \frac{d\alpha_p{}^l}{dz^p} = \alpha'_p{}^l\ . \tag{18}$$

Equations (10) and (18) can be united $\alpha_q{}^j \dfrac{\partial \alpha_p{}^l}{\partial y^j} = \delta_p{}^q \alpha'_p{}^l$.

From here

$$\frac{\partial \alpha_p{}^l}{\partial y^j} = \delta_p{}^q \alpha'_q{}^l \beta_j{}^q = \alpha'_p{}^l \beta_j{}^p \ . \tag{19}$$

Let's bring (19) in (17):

$$I \equiv \alpha'_p{}^l \beta_j{}^p \beta_i{}^p - \alpha_p{}^l \beta_k{}^p \beta_i{}^h \alpha'_h{}^k \beta_j{}^h - \alpha'_p{}^l \beta_i{}^p \beta_j{}^p - \alpha_p{}^l \beta_k{}^p \beta_j{}^h \alpha'_h{}^k \ \beta_i{}^h \equiv 0 \ ,$$

and it also proves full that the system (21) is quite integrated systems (21). The right parts of this system depend only from $y^1,..., y^n$ and do not depend from $\psi_p{}^l$ the decision can be found by means of quadratures. According to usual procedure we shall take the equation $\dfrac{\partial \psi}{\partial y^1} = a_1$. Integrating, we receive

$$\psi = \int_0^{y^1} a_1 dy^1 + A_2(y^2,..., y^n) \ , \tag{20}$$

where A_2 is while unknown function of the specified arguments. We differentiate (20) on y^2, having considered, that $\dfrac{\partial \psi}{\partial y^2} = a_2$, $\quad \dfrac{\partial a_1}{\partial y^2} = \dfrac{\partial a_2}{\partial y^1}$,

we receive

$$\frac{\partial A_2}{\partial y^2} = a_2 - \int_0^{y^1} \frac{\partial a_2}{\partial y^1} dy^1 \tag{21}$$

The right part (21) depends only from $y^2,..., y^n$. Integrating, we find

$$A_2 = \int_0^{y^2} \left[a_2 - \int_0^{y^1} \frac{\partial a_2}{\partial y^1} dy^1 \right] dy^2 + A_3(y^3,..., y^n) \ . \tag{22}$$

We differentiate (22) on, having considered, that

$$\frac{\partial A_2}{\partial y^3} = \frac{\partial \psi}{\partial y^3} - \int_0^{y^1} \frac{\partial a_1}{\partial y^3} dy^1 = a_3 - \int_o^{y^1} \frac{\partial a_3}{\partial y^1} dy^1 \ ,$$

we receive $a_3 - \displaystyle\int_0^{y^1} \frac{\partial a_3}{\partial y^1} dy^1 = \int_0^{y^2} \left[\frac{\partial a_2}{\partial y^3} - \int_0^{y^1} \frac{\partial^2 a_2}{\partial y^1 \partial y^3} dy^1 \right] dy^2 + \frac{\partial A_3}{\partial y^3}$, whence

$$\frac{\partial A_3}{\partial y^3} = a_3 - \int_0^{y^1} \frac{\partial a_3}{\partial y^1} dy^1 - \int_0^{y^2} \left[\frac{\partial a_2}{\partial y^3} - \int_0^{y^1} \frac{\partial^2 a_2}{\partial y^1 \partial y^3} dy^1 \right] dy^2 \ .$$

Integrating, we find

$$A_3 = \int_0^{y^3} \left\{ a_3 - \int_0^{y^1} \frac{\partial a_3}{\partial y^1} dy^1 - \int_0^{y^2} \left[\frac{\partial a_2}{\partial y^3} - \int_0^{y^1} \frac{\partial^2 a_2}{\partial y^1 \partial y^3} dy^1 \right] dy^2 \right\} dy^3 + A^4(y^4,..., y^n) \ .$$

And so on. The theorem is proved. Equality (4) become

$$y^i = \sum_{k=1}^{n} \psi_k{}^i (y^1,..., y^n) \ . \tag{23}$$

But on construction $\alpha_i^i = \dfrac{d\psi_i^i}{dz^i} = 1$ (not summarize). Hence, $\psi_i^i = z^i$, and then, from (23), it is found

$$
\begin{array}{l}
z^1 = y^1 - \psi_2{}^1(y^k) - \psi_3{}^1(y^k) - \ldots - \psi_n{}^1(y^k),\\
z^2 = y^2 - \psi_1{}^2(y^k) - \psi_3{}^2(y^k) - \ldots - \psi_n{}^2(y^k),\\
-------------------------\\
z^n = y^n - \psi_1{}^n(y^k) - \psi_2{}^n(y^k) - \ldots - \psi_{n-1}{}^n(y^k).
\end{array}
\tag{24}
$$

We have found formulas of the direct transformation, giving full project splitting. To receive formulas of return transformation (4) (namely, these formulas enable us to receive from (3) the equations (5)), it is necessary to solve the equations (24) y^k rather. Technically, it can represent, certainly, significant difficulties.

The second theorem project splitting is proved as follows.

Theorem 2. Formulas direct project splitting are by means of quadratures.

The note. It is of interest to write down to write down the characteristic equation (8) in variables z^i. With this purpose we shall write $\dfrac{\partial^2 f^i}{\partial y^j \partial y^l}$ out values of private derivatives in the specified variables. We differentiate (5):

$$\frac{\partial f^i}{\partial y^j} = (\alpha_k{}^i \mathrm{X}^k)' \frac{\partial z^k}{\partial y^j} = (\alpha_k{}^i \mathrm{X}^k)' \beta_j{}^k ,$$

$$\frac{\partial^2 f^i}{\partial y^j \partial y^l} = (\alpha_k{}^i \mathrm{X}^k)'' \beta_j{}^k \beta_l{}^k + (\alpha_k{}^i \mathrm{X}^k)' \frac{\partial \beta_j{}^k}{\partial y^l} .$$

But, differentiating (14), we receive $\dfrac{\partial \beta_j{}^h}{\partial y^l} \alpha_h{}^t + \beta_i{}^a \alpha'_a{}^t \dfrac{\partial z^a}{\partial y^l} = 0$, whence $\dfrac{\partial \beta_j{}^h}{\partial y^l} = -\alpha'_a{}^t \beta_t{}^h \beta_j{}^a \beta_l{}^a$.

Hence, $\dfrac{\partial^2 f^i}{\partial y^l \partial y^j} = (\alpha_k{}^i \mathrm{X}^k)'' \beta_j{}^k \beta_l{}^k - (\alpha_a{}^i \mathrm{X}^a)' \beta_t{}^a \;\; \alpha'_k{}^t \beta_j{}^k \beta_l{}^k$.

Thus,

$$\frac{\partial^2 f^1}{\partial y^j \partial y^l} - S \frac{\partial^2 f^1}{\partial y^j \partial y^l} = \lambda_k \beta_j{}^k \beta_l{}^k ,$$

where

$$\lambda_k = (\alpha_k{}^1 X^k)'' - (\alpha_a{}^1 X^a)' \beta_t{}^a \alpha'_k{}^t - S\left[(\alpha_k{}^2 X^k)'' - (\alpha_a{}^2 X^a)' \beta_t{}^a \alpha'_k{}^t\right].$$

Hence

$\left|\dfrac{\partial^2 f^1}{\partial y^j \partial y^l} - S \dfrac{\partial^2 f^1}{\partial y^j \partial y^l}\right| = \lambda_{1\ldots}\lambda_n \left|\beta_j{}^k\right|^2$, and the characteristic equation becomes $\lambda_1 \quad \lambda_2 \ldots \lambda_n = 0.$

References

1. Valeev, V. K. *Linear differential equations (splitting of solutions).* Kiev: Znanie, 1978. 47 p. (In Russian)
2. Kovantsov, A. N. Reducing of systems of usual differential equations to the block form using the projectors and induced surfaces. In: *Computational and Applied Mathematics.* Kiev, 1981, pp. 108–117. (In Russian)
3. Maljcev, A. I. *Foundations of Linear Algebra.* Moscow: Gostechizdat, 1956. 434 p. (In Russian)

Received on the 21st of July, 2011

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 59–67
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

EUCLIDEAN RINGS AND CRYPTOSYSTEMS WITHOUT REPETITION

A. Ten, V. Ten

Kazakh-British Technical University
Tole bi 59, Almaty, Republic of Kazakhstan
E-mail: tenvd@mail.ru

Kazakh National University
Masanchy 39/47, Almaty, Republic of Kazakhstan
E-mail: vanzeroten@mail.ru

Cryptosystems without repetitions are constructed. It means that the algorithm of encryption allows encrypting any text so that after encryption all elements of alphabet are distinct.

Keywords: *Euclidean ring, cryptosystem, plaintext, cipher text, secret key, partial key, alphabet*

1. Introduction

We use in our study some terms from [1]. The goal of research is to construct cryptosystems, which satisfy the following properties:

(A) Encrypting algorithm encrypts plaintext of any length so that after encryption with one fixed key all letters of cipher text are distinct, i.e. the cipher text has no repetitions. The length of alphabet is bounded by the length of the key.

(B) Encrypting algorithm encrypts plaintext of any length so that after encryption with one fixed key all letters of cipher text are distinct, i.e. the cipher text has no repetitions. The length of alphabet is not bounded.

In properties (A) and (B) the important role play words "of any length" and "fixed key", as it is well known there exist another ciphers without repetition, for example CBC and CTR are algorithms without repetitions [2]. But if the key is fixed the length of a cipher text is bounded.

2. Fundamental Theorem

Let K be the Euclidean ring [3] with the following property (E_3). For any $a,b,c \in K$ if $\delta(a) < \delta(c),\ \delta(b) < \delta(c)$, $a - b \equiv 0 \ (\mathrm{mod}\, C)$, then $a = b$, where $C = cK$ is the ideal of the ring K generated by the element c.

Remark 1. Examples of rings that satisfy condition (E_3) are the ring of integers and rings of polynomials over fields. The example of the ring without condition (E_3) is the ring of Gauss integers [3].

Algorithms of encryption and decryption follow from *Theorem 1.*

Theorem 1. Let K be the Euclidean ring, p be the element of the ring K. Then for any element $k \in K$ such that k is not invertible, $\delta(k) < \delta(p)$ and for any element $s \in K$ such that

$$(s,k) = e\ ,\ (s,p) = e\ , \tag{1}$$

there exist elements a_k , $\bar{k} \in K$ that satisfy the following conditions:

1) $\delta(a_k) < \delta(p)$.

2) $\overline{k} = a_k s$, $k \equiv \overline{k}\ (\bmod P)$ where P is the ideal of the ring K generated by the element p.

3) $\delta(p) \le \delta(\overline{k})$.

4) If $k_1 \ne k_2$, $\delta(k_1), \delta(k_2) < \delta(p)$ then $\overline{k_1} \ne \overline{k_2}$ and inverse if $\overline{k_1} \ne \overline{k_2}$ then $k_1 \ne k_2$.

Proof. Let p be the element of the ring K, k be the element of K that satisfies the condition $\delta(k) < \delta(p)$. As $(s, p) = e$ then by the Chinese remainder theorem [4] there exists the element $b_k \in K$ such that

$$b_k \equiv 0(\bmod S), \quad b_k \equiv k(\bmod P), \tag{2}$$

where S is the ideal of the ring K generated by the element s and $b_k = kus$ where element u may be obtained from the equality $us + vp = e$ (it is well known that we may find elements u, v with the help of Euclidean algorithm). Let a_k be the remainder after dividing of the element $c_k = ku$ by p. It is known that the element b_k that satisfies conditions (2) doesn't determine uniquely but from the condition $(s, p) = e$ it follows that any two elements $c_k^{'} = ku'$, $c_k^{''} = ku''$ have the same remainders after dividing by p, i.e. a_k is determined uniquely. Really let us suppose that elements $b_k^{'}, b_k^{''}$ satisfy conditions (2). Then we have $b_k^{'} - b_k^{''} = c_k^{'} s - c_k^{''} s \equiv 0(\bmod P)$, it follows from this that $c_k^{'} - c_k^{''} \equiv 0(\bmod P)$, but as $\delta(a_k^{'}), \delta(a_k^{''}) < \delta(p)$ and $a_k^{'} - a_k^{''} \equiv c_k^{'} - c_k^{''} \equiv 0\ (\bmod P)$ then from condition (E_3) we have $a_k^{'} = a_k^{''}$.
No problems to see, that

$$b_k = c_k s \equiv a_k s(\bmod P) \equiv k(\bmod P). \tag{3}$$

Let $\overline{k} = a_k s$ then from (3) it follows that $\overline{k}$ satisfies the item 2). The inequality $\delta(p) < \delta(\overline{k})$ is true as otherwise we have $\overline{k} = a_k s \equiv k\ (\bmod P)$, $\delta(k) < \delta(p), \delta(\overline{k}) < \delta(p)$, it follows from condition (E_3) that $\overline{k} = a_k s = k$, but this is impossible as $(s, k) = e$ and so the item 3) is true. It follows from (2) that if $\delta(k_1), \delta(k_2) < \delta(p)$ and $k_1 \ne k_2$ then $\overline{k_1} \ne \overline{k_2}$ and inverse if $\overline{k_1} \ne \overline{k_2}$, then $k_1 \ne k_2$, so the item 4) is true.

Now we are ready to consider applications of *Theorem 1* to cryptography.

3. Algorithms of Encryption and Decryption

(C_1) Algorithm of encryption: let p be an element of the Euclidean ring K, k be any element of K such that $\delta(k) < \delta(p)$, s be any element of K that satisfies conditions (1). Then using theorem 1 we obtain the element $\overline{k}$. So we have: p is the secrete key, k is the plaintext, $\overline{k}$ is the cipher text.

(D_1) Algorithm of decryption: $k = \overline{k}\ (\bmod p)$, i.e. k is the remainder after dividing $\overline{k}$ by p.

Remark 2. It is very easy to change secret keys in our cryptosystem. Really if we change the secret key p by p_1 then after encryption we obtain new value of the cipher text $\overline{k_1}$, but after decryption we have $k = \overline{k_1}\ (\bmod p_1)$.

Definition 1. Element $s \in K$ is called the partial key of the pair (p, k) if it satisfies conditions (1).

Now let K be the Euclidean ring, p be the element of the ring K, $k_1, k_2, ..., k_m$ be some sequence of elements of K such that $\delta(k_i) < \delta(p)$. Suppose that the ring K satisfies the following condition: (I_s) The set $S = \{s \in K, s\ \ is\ \ prime, \delta(s) > \delta(p)\}$ is infinite.

No problems to see that any element from S may be considered as a partial key of any pair (p,k_i). Now let $s_1,s_2,...,s_m$ be the sequence of distinct elements from S, then s_i is the partial key of the pair $(p,k_i), i=1,2,...,m$ and from theorem 1 it follows that if we use encrypting algorithm (C_1) to plaintexts $k_1,k_2,...,k_m$ with corresponding partial keys $s_1,s_2,...,s_m$ then all cipher texts $\overline{k_1},\overline{k_2},...,\overline{k_m}$ are distinct.

(C_2) Algorithm of encryption: let p be an element of the ring K, $k_1,k_2,...,k_m$ be some sequence of elements of K such that $\delta(k_i)<\delta(p), i=1,...,m$, $s_1,s_2,...,s_m$ be the sequence of distinct elements from S. We encrypt element k_i using partial key s_i by the algorithm (C_1). As the result all encrypting texts $\overline{k_1},\overline{k_2},...,\overline{k_m}$ are distinct.

(D_2) Algorithm of decryption: $k_i=\overline{k_i}\ (\bmod p),\ i=1,2...,m$.

Theorem 2. Let $K=Z$ be the ring of integers. Then the cryptosystem with algorithm of encryption (C_2) and algorithm of decryption (D_2) satisfies property (A).

Proof. It is well known that Z is the Euclidean ring with the function $\delta(z)=|z|,\ z\in Z$ and Z satisfies condition (I_s). Let p be the secret key of our cryptosystem then from theorem 1 it follows that the number of distinct plaintexts is less then $2|p|$. So our cryptosystem satisfies property (A).

Theorem 3. If $K=Q[x]$ be the polynomial ring over the field of rational numbers Q, then the cryptosystem with algorithms of encryption and decryption (C_2), (D_2) satisfies properties (A), (B).

Proof. The ring $Q[x]$ is the Euclidean ring with the function $\delta(f)=\deg f,\ f\in Q[x]$ and it satisfies condition (I_s). Let $p\in Q[x]$ be the secret key of our cryptosystem and $\deg p\geq 2$, then it is clear that the set of polynomials $g\in Q[x]$ such that $\deg g<\deg p$ is infinite. So our cryptosystem satisfies property (B).

Now we consider the following kind of attack against our cryptosystems. Let we have some set of plaintexts $k_1,k_2,...,k_n$ and the set of corresponding cipher texts $\overline{k_1},\overline{k_2},...,\overline{k_n}$, p is the secret key. From the decrypting algorithm it follows that there exist elements $d_1,\dots,d_n\in K$ such that $a_i=\overline{k_i}-k_i=d_i\ p,\quad i=1,2,...,n$.

Let us denote the greatest common divisor of elements $x_1,x_2,...,x_n\in K$ as $(x_1,\dots,x_n)$. Then as p is the prime element we have $(a_1,\dots,a_n)=dp$, $d=(d_1,\dots,d_n)$. And this attack may be effective for finding of secret key.

So we modify our cryptosystem as follows.

(C_3) Algorithm of encryption: let $n\in K, \delta(n)<\delta(p)$.

Let $\overline{\overline{k}}=(q;r)\in K\times K$, where $\overline{k}=nq+r$ and $\overline{k}\neq r$ as $\delta(\overline{k})>\delta(p)$. Let's consider two maps $E_p:K\to K$ and $\overline{E}_n:K\to K\times K$ which are determined by the following way:

$$E_p(k)=\overline{k},\ \overline{E}_n(\overline{k})=(q,r). \tag{4}$$

Then,

$$\overline{\overline{k}}=\overline{E}_n(E_p(k)) \tag{5}$$

and we have: k is the plaintext, $\overline{\overline{k}}$ is the cipher text. First step: we use the algorithm (C_2), second step: we use the map $\overline{E_n}$.

(D_3) Algorithm of decryption:

$$[nq + r](\bmod p) = k\,. \tag{6}$$

Here we have two secret keys: p and n.

No problems to see that if the cryptosystem with algorithms of encryption (C_2) and decryption (D_2) satisfies property (A) or (B), then the cryptosystem with algorithms of encryption (C_3) and decryption (D_3) satisfies the same properties.

Let $K = Z$ be the ring of integers, then we call the cryptosystem with algorithm of encryption (C_3) and algorithm of decryption (D_3) EZ-cryptosystem.

4. Ez-Cryptosystem

The formulation of the theorem 1 for the ring of integers is:

*Theorem 1** Let p be the integer number. Then for any $k \in Z$ such that $1 < |k| < |p|$ and for any $s \in Z$ such that $(s,k) = 1$, $(s,p) = 1$ there exist $a_k, \overline{k} \in Z$ that satisfy the following conditions:

1) $|a_k| < |p|$;
2) $\overline{k} = a_k s$, $k = \overline{k}\ (\bmod p)$;
3) $|p| \le |\overline{k}|$;
4) if $k_1 \ne k_2$ then $\overline{k_1} \ne \overline{k_2}$ and inverse if $\overline{k_1} \ne \overline{k_2}$ then $k_1 \ne k_2$.

Then we have the following algorithms of encryption (C_1) and algorithm of decryption (D_1).

Encryption: Let p be the integer number. If $k \in Z, |k| < |p|$, $s \in Z$ and s satisfies the conditions: $(s,p) = 1$, $(s,k) = 1$, s is the partial key. Then we obtain cipher text $\vec{k}$ as follows:

Step 1: we find u from the equality $us + vp = 1$, it is well known that u may be evaluated from generalized Euclidean algorithm [5].

Step 2: we find a_k as the remainder of division of ku by p.

Step 3: $\vec{k} = a_k s$.

Decryption: $k = \overline{k}(\bmod p)$.

4.1. Algorithms of Generation of Partial Keys

The principle role for algorithm of encryption (C_2) plays partial keys. We offer two algorithms of generation of partial keys.

I. Let $\Sigma = \{k_1, k_2, ..., k_m, k_i \in Z\}$ and $K = \max\{|k_i|, i = 1,2,...,m\}$. Then no problems to see that the set $S = \{s \in Z, s \text{ is prime}, (s,p) = 1, |s| > \max(K, \frac{p}{2})\}$ is the set of partial keys for all pairs (p,k_i), $i = 1,2,...,m$ and period of S is infinitive.

II. Let we have the finite set of partial keys $S = \{s_1, s_2, ..., s_l\}$ of pairs (p,k_i), $i = 1,2,...,m$. The following algorithm allows us to extend S to obtain the set of partial keys of any period N, $N \ge l$. Really if we consider any product $s_1^{n_1} s_2^{n_2} ... s_l^{n_l}$, $n_1 + n_2 + ... + n_l \ge 1$ then it is clear that all these products are partial keys for all pairs (p,k_i), $i = 1,2,...,m$. So if we add to the set S the set of $(N - l)$ distinct products $s_1^{n_1} s_2^{n_2} ... s_l^{n_l}$, $n_1 + n_2 + ... + n_l \ge 2$, we obtain the set of partial keys $\overline{S}$ and period of $\overline{S}$ is equal to N.

4.2. Security of Ez-Cryptosystem

Let's consider algorithm of encryption (C_3) with the finite set of partial keys $S = \{s_1, s_2, ..., s_l\}$. Algorithm of encryption guaranties that any text of the length l has no repetitions after encrypting. Suppose that our text has the length more then l. Then cipher text has repetitions if and only if the distance between two copies of some element of our alphabet is multiple of l.

Example. Let's consider the text

HOW DO YOU DO

a) Let $p = 1433, n = 1000$ be secret keys, $S = \{1553,1559\}$ be the set of partial keys

letter	**plaintext**	**cipher text**
H	104	(594; 799)
O	111	(54; 565)
W	119	(316; 812)
SPACE	32	(35; 857)
D	100	(1855; 835)
O	111	(54; 565)
SPACE	32	(1038; 957)
Y	121	(763; 910)
O	111	(613; 435)
U	117	(480; 172)
SPACE	32	(1038; 957)
D	100	(391; 309)
O	111	(613; 435)

We have three pairs of repetitions, these positions are: 2) and 6); 7) and 11); 9) and 13).

b) Let $p = 1433, n = 1000$, $S = \{1553,1559,1567,1571,1579,1583\}$

letter	**plaintext**	**cipher text**
H	104	(594; 799)
O	111	(54; 565)
W	119	(1090; 632)
SPACE	32	(32; 991)
D	100	(1117; 933)
O	111	(500; 228)
SPACE	32	(1038; 957)
Y	121	(763; 910)
O	111	(922; 963)
U	117	(50; 272)
SPACE	32	(186; 322)
D	100	(1513;348)
O	111	(613; 435)

Here we have $|S| = 6$ but the sequence of cipher texts has no repetitions. Well known methods of statistic analysis which are based on frequency distribution of letters and combinations of letters of languages. As we saw if EZ-cryptosystem has repetitions, then the law of distribution of frequencies of these repetitions is differ of the law of distribution of frequencies we consider before. So these methods of statistic analysis can not be used against EZ-cryptosystem. And of course they can not be used if cipher text has no repetitions. This allows us to return to the encryption of language alphabets such as English, Russian, etc.

Let's p, n be secret keys of EZ-cryptosystem, k be some plaintext, $\overline{\overline{k}}$ – the corresponding cipher text. Then we have from (4), (5), (6) – $E_p(k) = \overline{k}$, $\overline{E}_n(\overline{k}) = (q;r)$, $\overline{\overline{k}} = \overline{E}_n(E_p(k))$, $\overline{k}(\mathrm{mod}\, p) = [qn + r](\mathrm{mod}\, p) = k$.

Let's suppose that k and $\overline{\overline{k}} = (q;r)$ are known. Our problem is to find p and n or any information about them. Let's denote: $p = x$, $n = y$, $\overline{k} = tqn + r = qx + r = z$. So we have the following equations:

$$z(\mathrm{mod}\, x) = k; \tag{7}$$

$$z = qy + r. \tag{8}$$

It is well known [6] that the equation (7) with unknowns x, y has infinite number of integer solutions, but if x or z is fixed then the equation (7) has the unique solution. So let's determine values of z from the equation (8). The equation (8) is the equation of the straight line with integer coefficients, so it has infinite number of integer solutions. It is clear that if y is bounded then z is bonded and inverse. But from theorem 1^* it follows that $|z|$ has only lower bounder, as $|z| > p$ and has no upper bounder, so we can not find z even using brute force attack. No problems to show that if we consider the set of plaintexts $k_1, k_2, ..., k_m$ and corresponding to them cipher texts $\overline{\overline{k_1}}, \overline{\overline{k_2}}, ..., \overline{\overline{k_m}}$, then the problem of finding of p and n reduces to the problem of researching of solutions of the linear systems of m equations with $m+1$ unknowns. As it is known [7] either such system has infinite number of solutions or it has no solutions. From theorem 1^* it follows that the system has the solution, so it has infinite number of solutions. Now we consider some characteristics of EZ-cryptosystem.

4.3. Characteristics of Ez-Cryptosystems

1. EZ-cryptosystem is the block cryptosystem,
2. The space of keys is infinite.
3. Secrete keys p and n are keys of encryption and decryption.
4. If p is fixed, then the space of plaintexts PT is infinite but the length of alphabet is not more then $2(|p|-2)$, the space of cipher texts CT is infinite.
5. EZ-cryptosystem allows encrypting short alphabets.
6. It is very easy to change secret keys p and n.

Now we want to offer two problems.

1)Let we have the following English text:

Welcome to our cryptosystem!

Plaintexts k :

W e l c o m e t o o u r c r y p t o
87 101 108 99 111 109 101 32 116 111 32 111 117 114 32 99 114 121 112 116 111
s y s t e m !
115 121 115 116 101 109 33

Cipher texts $\overline{\overline{k}}$:

W e l c o m e
(470,94761) (603, 429882) (809,250201) (1439,511146) (463,284397) (112,116505) (908,156193)
t o o u r
(737,175299) (323,138907) (551,375851) (428,444831) (802,439816) (703,533823) (46,441693)

c r y p t o
(1056,309933) (873,316704) (914,556215) (18,412601) (402, 24117) (2247,563330) (2004,507674)
s y s t e m !
(247,59923) (2577,512514) (833,275150) (2424,156617) (889,333334) (17,214663) (1418,292369)

So analyst has the following information:

a) Algorithm of encryption (C_3) and algorithm of decryption (D_3).

b) Plaintexts.

c) Cipher texts.

Problem.

Find secret keys p and n.

2) Let our English text consists of ten letters b: **bbbbbbbbbb**

Plaintexts k :

b	b	b	b	b	b	b	b	b	b
98	98	98	98	98	98	98	98	98	98

Ciphertexts $\overline{\overline{k}}$:

b	b	b	b	b	b	b
(3970,170)	(223447,2664)	(94863,2470)	(294212,57)	(12525,2205)	(106691,1006)	(126598,142)

b	b	b
(215619,907)	(228592,700)	(150978,785)

Problem.

Find secret keys p and n.

Instead of Conclusions

If somebody will find an algorithm (see Supplement) of effective attack against our cryptosystem and inform us we'll be grateful to him.

References

1. Salomaa, A. *Public-key cryptography*. Berlin-Heidelberg: Springer-Verlag, 1990.
2. *Recommendation for Block Cipher Modes of Operation.* NIST Special Publication 800-38A. Technology Administration US Department of Commerce, 2001.
3. Van der Waerden, B. L. *Algebra 1*. Berlin-Heidelberg: Springer-Verlag, 1971.
4. Lang, S. *Algebra*. N.-Y.: Addison-Wesley Publishing Company, 1965.
5. Proskuryakov, I. V. Numbers and Polynomials. Moscow: Prosvescheniye, 1965. (In Russian)
6. Buchshtab, A. A. *Theory of Numbers*. Moscow: Fiz-Mat. Publishing, 1960. (In Russian)
7. Lipschutz, S. *Linear algebra*. N.-Y.: McGraw-Hill Book Company, 1987.

Supplement. Listing of algorithm of encryption (C_3)

```
private void encrypt_Click(object sender, EventArgs e)
      {
        textBox1.Text = "";
        textBox2.Text = "";
        str2 = "";
        str3 = "";
        int[] s = new int[] { 1553, 1559, 1567, 1571, 1579,
                    1583, 1597, 1601, 1607, 1609,
                              1613, 1619, 1621, 1627, 1637,
                    1657, 1663, 1667, 1669, 1693,
                              1697, 1699, 1709, 1721, 1723,
                    1733, 1741, 1747, 1753, 1759,
                    1777, 1783, 1787, 1789, 1801,
                    1811, 1823, 1831, 1847, 1861,
                    1867, 1871, 1873, 1877, 1879,
                    1889, 1901, 1907, 1913, 1931,
                    1933, 1949, 1951, 1973, 1979,
                    1987, 1993, 1997, 1999, 2003,
                    2011, 2017, 2027, 2029, 2039,
                    2053, 2063, 2069, 2081, 2083,
                    2087, 2089, 2099, 2111, 2113,
                    2129, 2131, 2137, 2141, 2143,
                              2153, 2161, 2179, 2203, 2207,
                    2213, 2221, 2237, 2239, 2243,
                         2251, 2267, 2269, 2273, 2281,
                    2287, 2293, 2297, 2309, 2311};
        int p = Convert.ToInt32(textBox3.Text);
        int n = Convert.ToInt32(textBox4.Text);
```

```
int i = 0, j = 0;
int[] k;
int[] k1;
string str1 = richTextBox1.Text;
int l = str1.Length;
while (i < l)
{
   for (j = 0; j < 100; j++)
   {
      if (i < l)
      {
         int a, b, q, r, u1 = 1, v1 = 0, u2 = 0, v2 = 1, t,
            q1, r1;
         k = new int[l];
         k1 = new int[l];
         char[] ch = str1.ToCharArray();
         k[i] = (int)(ch[i]);
         a = s[j];
         b = p;
         while (b != 0)
         {
            a = u1 * s[j] + v1 * p;
            b = u2 * s[j] + v2 * p;
            q = a / b;
            r = a % b;
            a = b;
            b = r;
            t = u2;
            u2 = u1 - q * u2;
            u1 = t;
            t = v2;
            v2 = v1 - q * v2;
            v1 = t;
         }
         int m = k[i] * u1;
         int ak = m % p;
         if (ak < 0)
         {
           while (ak < 0){ak = ak + p;}
         }
         k1[i] = ak * s[j];
         r1 = k1[i] % n;
         q1 = (k1[i] - r1) / n;
         str2 = str2 + q1.ToString() + "%";
         str3 = str3 + r1.ToString() + "%";
         i++;
      }}}
textBox1.Text = str2;
textBox2.Text = str3;      }
```

Listing of algorithm of decryption (D_3)

```
private void decrypt_Click(object sender, EventArgs e)
      {
        int p = Convert.ToInt32(textBox4.Text);
        int n = Convert.ToInt32(textBox3.Text);
        int intStr1 = 0;
        int intStr2 = 0;
        string str1 = "";
        string str2 = "";
        str1 = textBox1.Text;
        str2 = textBox2.Text;
        string str3 = "";
        string[] strbuf1;
        string[] strbuf2;
        int[] k;
        int[] k1;
        char ch;
        strbuf1 = str1.Split('%');
        strbuf2 = str2.Split('%');
        int l = strbuf1.Length;
        for (int i = 0; i < l - 1; i++)
        {
           k = new int[l];
```

```
        k1 = new int[l];
        try
        {
           intStr1 = Convert.ToInt32(strbuf1[i]);
           intStr2 = Convert.ToInt32(strbuf2[i]);
        }
        catch (FormatException ex)
        {
           Console.WriteLine("Код не является числом: " +
                          strbuf1[i]);
           Console.WriteLine("Код не является числом: " +
                          strbuf2[i]);
        }
        k1[i] = n * intStr1 + intStr2;
        k[i] = k1[i] % p;
        ch = (char)(k[i]);
        str3 = str3 + ch.ToString();
     }
     richTextBox1.Text = str3; }
```

Received on the 21st of July, 2011

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 68–71
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

THE METHOD OF SPEECH SIGNAL INTONATION SYNTHESIS BASED ON SPLINE APPROXIMATION

M. Kalimoldaev[a], Ye. Amirgaliyev[b], R. Musabayev[c]

Institute of Informatics and Control Problems
Pushkin 125, Almaty, Kazakhstan
a)E-mail: mnk@ipic.kz
b)E-mail: amir_ed@mail.ru
c)E-mail: rmusab@gmail.com

The synthesis method of speech signal intonation component based on the spline – mathematically calculated curves, which smoothly connect separate control points of an intonation contour, is described. This method has been used during the realization of speech signal compilative synthesis system. The parametrical description of the synthesized speech signal in the UPL language in aggregate with the method of approximation of its intonation component by spline improves the quality of approximation in comparison with the existing methods.

Keywords: *speech synthesis, spline approximation, fundamental frequency, intonation contour, compilative synthesis*

1. Introduction

An important component part of the most text-to-speech synthesis systems is the intonation synthesis module. This module is intended for generation of an intonation contour and its subsequent applying to a synthesized signal. Naturalness of the synthesized signal is substantially determined by intonation contour definition quality. During the synthesis process it is important to realize the smooth control of speech signal parameters. Otherwise the synthesized speech will have unnatural sounding. Thus during the implementation of speech synthesis and recognition systems, the problem of smooth intonation processes modelling is actual. The given article describes the synthesis method of speech signal intonation component based on the spline – mathematically calculated curves, which smoothly connect separate control points of an intonation contour. This method is being used during the realization of speech signal compilative synthesis system, which has been developed at the Institute of Informatics and Control Problems of the Ministry of Education and Science of the Republic of Kazakhstan.

Classic works of a number of foreign scientists: G. Fant [2], J. Flanagan [3], S. Furui [4], P. Taylor [5], X. Huang [6] are well-known. Similar questions are also considered in the works of the Byelorussian and Russian scientists: B. M. Lobanov [1], V. N. Sorokin [7].

2. The Statement of the Problem

For the synthesis of a speech signal on the compilative principle, it is necessary to obtain preliminary the formalized description of its phonetic and intonation properties. As a part of the given description it is necessary to specify intonation characteristics for all phonemes. Their composition includes also a set of control points of parametrical curves. The parameters of the neighbour phonemes should be smoothly matched under these conditions. Thus the development of the specialized language is defined as a goal. This language will help to make the preliminary description of phonetic and intonation properties of a synthesized speech signal. Another necessary task is an algorithmization of the process of smooth parametrical curves calculation. The dynamics of change of the controlled parameters will be defined with their help.

3. The Proposed Solution

On Figures 1–3 the basic stages of the speech signal synthesis by the compilative principle have been shown. The given approach is realized with the application of smooth parametrical curves defined by the limited set of the control points. The result of synthesis is shown on Figure 4.

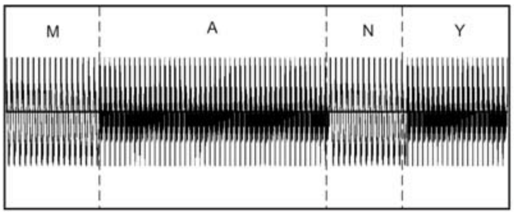


Figure 1. The initial speech signal after matching and concatenation

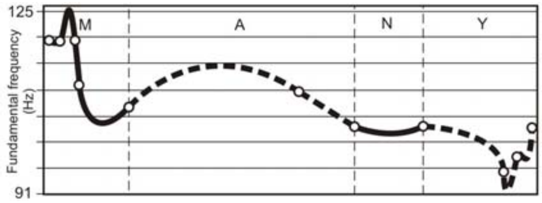


Figure 2. Applying of the fundamental frequency contour

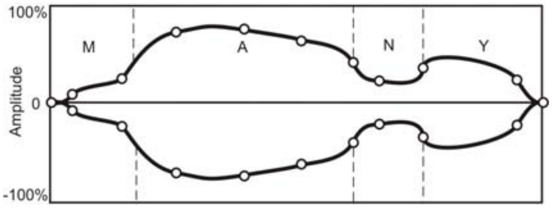


Figure 3. Applying of the amplitude envelopes

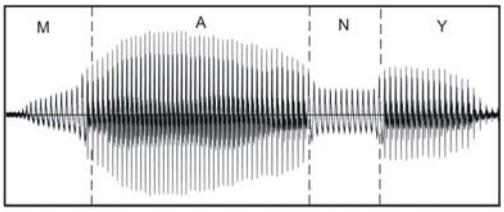


Figure 4. The result of the synthesis

For the achievement of the qualitative synthesis it is important to regulate smoothly the following parameters of a speech signal:

1. The contour of fundamental frequency – is a main intonation component of the speech (Figure 2).

2. Amplitude envelopes which primary function is the dynamic regulation of the amplitude level of a signal (4). The joint increase of the amplitude and the frequency of a signal leads to the increase of its loudness.

Under the compilative synthesis [6] on the basis of base fragments of speech the sound signal is given the necessary form by the method of various algorithmic manipulations. The defined form of a speech signal can be dependent on the set of various factors: a language, specific features of a voice, a synthesized text, a required intonation, a speed and loudness of pronunciation, etc.

The ready-made and normalized by phonemes duration and by the common amplitude level smoothly connected out of various fragments speech signal is submitted to the input of the parameters regulation system (Figure 1). Depending on required intonation characteristics the fundamental frequency contour is formed and applied to the initial speech signal (Figure 2). Then the amplitude envelopes are applied to the signal (Figure 3).

For the curve definition the limited set of the control points are particularized. Their optimum allocation is selected for the best approximation of the controlled parameter assumed function. Initially the extrema of an approximated function are selected as the control points.

As a part of the selection task definition of the control points optimum allocation it was required to estimate the values of their coordinates in the given ranges by the search method in sense of a criterion minimum (deviation square). The criterion is as follows (1):

$$K = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\frac{Y_j^* - Y_j}{Y_j^*}\right)^2}, \tag{1}$$

where Y_j^*, Y_j – accordingly the values of an approximated function and the obtained values in a process of curve calculation; N – samples amount.

The algorithm of custom point calculation of smooth parametrical curve is given below.

The input data:

- A – the set of control points specified by their coordinates $(X;Y)$;
- Ax – the element X component of the set A;
- Ay – the element Y component of the set A;
- T – specifies the position of the calculated point on the curve, $t \in [0,1]$.

The output data:

- X – the coordinate of the calculated point on the X-axis;
- Y – the coordinate of the calculated point on the Y-axis.

Notation:

- $f(g) = g^3 - g$;
- *i, j* – loop variables;
- *Num* – the amount of the elements in the set *A*;
- *dT* – the increment value of *T* the for every element of the set *A*;
- *dX* – the increment value by the *X*-axis;
- *Px, Py, Wx, Wy, D* – the sets with *Num* amount of the elements;
- *div* – the integer division operation.

1. Initialisation: *Num = Length(A) - 1*
2. For each particle *i = 1..Num - 1*

   $D_i = 4$

   $W_x = 6 \cdot ((Ax_{i+1} - Ax_i) - (Ax_i - Ax_{i-1}))$; $W_y = 6 \cdot ((Ay_{i+1} - Ay_i) - (Ay_i - Ay_{i-1}))$

   EndFor
3. $Px_0 = 0$, $Py_0 = 0$, $Px_{Num} = 0$, $Py_{Num} = 0$
4. For each particle *i = 1..Num-2*

   $Wy_{i+1} = Wy_{i+1} - Wy_i \cdot 0.25$; $Wx_{i+1} = Wx_{i+1} - Wx_i \cdot 0.25$

   $D_{i+1} = D_{i+1} - 0.25$

   EndFor
5. For each particle *i = Num-1..1* do

   $Px_i = \frac{Wx_i - Px_{i+1}}{D_i}$; $Py_i = \frac{Wy_i - Py_{i+1}}{D_i}$

   EndFor
6. $X = Ax_0$; $Y = Ay_0$
7. $dX = Ax_{Num} - Ax_0$
8. If $dX > 0$ then

   Begin

   $dT = \frac{1}{Num}$

   for *i = 0..Num-1* do

   If $(dT \cdot i \le T)$ and $(dT \cdot (i+1) \ge T)$ then

   Loop termination

   $T = (T - (T \quad div \quad dT) \cdot dT) \cdot Num$

   $X = T \cdot Ax_{i+1} + (1-T) \cdot Ax_i + \frac{f(T) \cdot Px_{i+1} + f(1-T) \cdot Px_i}{6}$

   $Y = T \cdot Ay_{i+1} + (1-T) \cdot Ay_i + \frac{f(T) \cdot Py_{i+1} + f(1-T) \cdot Py_i}{6}$

   End.

For the solution of the speech signal synthesis problem the Unified Phonetic representation Language (UPL) has been developed. In fact the given language is an extended phonetic transcription. The required characteristics of the speech signal are described in the UPL language. On their basis the compiler chooses the most suitable elements of compilation and performs the subsequent generation of the speech signal.

The synthesized signal in the Unified Phonetic representation Language can be described in the following form:

Phoneme1Stress1(Duration1;Allophone1;[Pitch1];{Amplitude1})
Phoneme2Stress2(Duration2;Allophone2;[Pitch2];{Amplitude2})
...
PhonemeNStressN(DurationN;AllophoneN;[PitchN];{AmplitudeN}),

where *PhonemeN* is the mnemonic representation of a phoneme, *StressN* is the stress attribute, *DurationN* is the sound duration of a phoneme in milliseconds, *AllophoneN* is the number of allophone realization of a phoneme, *PitchN* and *AmplitudeN* are fundamental frequency and amplitude contours respectively, which are defined as follows: $X_1, Y_1; X_2, Y_2; ...; X_m, Y_m$, where X_m is the relative coordinate of *m*-th control point in the interval from the start (0) till the end (1) of sound duration of $X_m \in [0;1]$ phoneme, Y_m – is the value of the fundamental frequency ($X_m \in [0;+\infty]$) or amplitude respectively ($X_m \in [0;1]$). The example of representation of the word "many" in the Unified Phonetic Language may be considered:

PAU(160;1)
M(43;1199;[*0,98;0.534,99.46;1,101.28*];{*0,0;0.6,0.1;1,0.2*})
EH1(82;1;[*0,101.28;0.5,106;1,103.46*];{*0,0.2;0.5,0.21;1,0.2*})
N(78;1;[*0,102.92;0.452,104.38;1,106.74*];{*0,0.2;0.5,0.1;1,0.2*})
IY1(127;1184;[*0,106.74;0.473,108.81;1,102.03*];{*0,0.2;0.5,0.21;1,0*})
PAU(160;1).

4. Conclusions

The comparison of the method proposed in the given work with the method of linear interpolation, which is used in the most of existing speech synthesis systems [6] has been conducted. The estimation was performed by the criterion of a deviation square sum minimum by the analogy with the formula (1) between calculated values according to the both methods and the natural etalon contour. As a result, for the method of linear interpolation, the average criterion is 0.25. For the proposed method the value of the average criterion is 0.07. Thus, the parametrical description of the synthesized speech signal in the UPL language in aggregate with the method of approximation of its intonational component by splines improves the quality of approximation in comparison with the existing methods.

The developed UPL language can be successfully used in the process of definition and description of various phonetic and intonation forms of the speech. All the initial data is described using the unified language representation, which allows implementing the flexible intersystem interaction and solving the speech synthesis task at a qualitative level.

The proposed approach can be also used in the speech recognition systems and in the systems of speaker identification by his intonation features.

References

1. Fant, G. *Speech Acoustics and Phonetics.* Dordrecht: Kluwer Academic Publishers, 2004. 333 p.
2. Flanagan, J. L. *Speech analysis, synthesis and perception.* Springer-Verlag, 1972. 467 p.
3. Furui, S. *Digital Speech Processing, Synthesis and Reccognition.* Marcel Dekker, 2001. 289 p.
4. Taylor, P. *Text to Speech Synthesis.* University of Cambridge, 2007. 597 p.
5. Huang, X., Acero, A., Hon, H.-W. *Spoken Language Processing: A Guide to Theory, Algorithm and System Development.* Prentice Hall, 2001. 472 p.
6. Lobanov, B. M., Tsirulnik, L. I. *Computer synthesis and speech cloning.* Minsk: Belorussian Science, 2008. 344 p. (In Russian)
7. Sorokin, V. N. *Speech synthesis.* Moscow: Science, 1992. 392 p. (In Russian)

Received on the 25th of July, 2011

Computer Modelling and New Technologies, 2011, Vol.15, No.2, 72–75
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

DIFFUSION PROCESS IN QUASI-ONE-DIMENSIONAL STRUCTURES

J.-R. Kalnin

Ventspils University College
Inženieru Street 101a, LV-3601, Ventspils, Latvia
E-mail: simts@latnet.lv

The effective diffusion coefficient in two-phase one-dimensional model with the periodical distribution of inclusions in the effective medium approximation is calculated and generalization about a quasi-one-dimensional case is formed.

Keywords: *effective diffusion, effective medium, method*

1. Introduction

In several papers the expression for the effective diffusion coefficient in the scope of generalized Maxwell-Garnett theory has been considered [1–3]. Different concentrations of diffusing particles in matrix and inclusions require corresponding conditions on the boundary matrix-inclusion.

It has been ad-hoc assumed that a concentration jump is equal to the average concentration ratio. Further in the text we show that this result strictly follows from the effective medium approximation and that one-dimensional approach can be used in solving quasi-one-dimensional problems after performing some computer simulation.

2. Effective Medium Approximation in One-Dimensional Diffusion Model

Let us consider one-dimensional heterogeneous medium: periodically distributed regions with the diffusion coefficient D_1 in matrix and the diffusion coefficient D_2 in inclusions. In the effective medium approximation we replace one-dimensional sample by a representative element consisting of a matrix and one inclusion, which are embedded into the effective media with the diffusion coefficient D_{eff} (Fig. 1).

The corresponding concentrations are c_{eff}, c_1, c_2 and c_3.

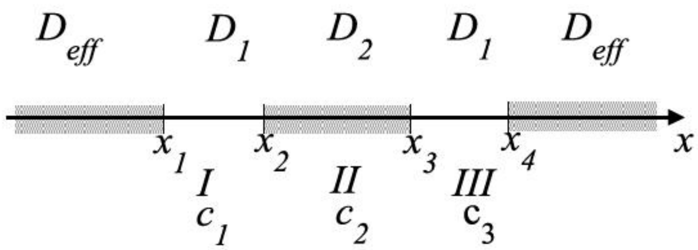


Figure 1. Representative element (I-II-III) in one-dimensional effective media

The external concentration field $c_{eff} = -gx$ with a constant gradient g is applied in the sample. Applying the solution of the stationary one-dimensional diffusion equation gives

$$
\begin{aligned}
&c_{eff} = -gx, \\
&c_1 = \alpha_1 x + \beta_1, \quad (region\ I) \\
&c_2 = \alpha_2 x + \beta_2, \quad (region\ II) \\
&c_3 = \alpha_3 x + \beta_3, \quad (region\ III).
\end{aligned}
\tag{1}
$$

We choose the boundary conditions in the form

$$
\begin{aligned}
&c_{eff}\big|_{x=x_1} = \frac{1}{\chi} c_1\big|_{x=x_1}, \\
&c_1\big|_{x=x_2} = \frac{1}{\alpha} c_2\big|_{x=x_2}, \\
&c_2\big|_{x=x_3} = \alpha\, c_3\big|_{x=x_3}, \\
&c_3\big|_{x=x_4} = \chi\, c_{eff}\big|_{x=x_4}
\end{aligned}
\tag{2}
$$

and

$$
\begin{aligned}
&D_{eff} \frac{\partial c_{eff}}{\partial x}\bigg|_{x=x_1} = D_1 \frac{\partial c_1}{\partial x}\bigg|_{x=x_1}, \\
&D_1 \frac{\partial c_1}{\partial x}\bigg|_{x=x_2} = D_2 \frac{\partial c_2}{\partial x}\bigg|_{x=x_2}, \\
&D_2 \frac{\partial c_2}{\partial x}\bigg|_{x=x_3} = D_1 \frac{\partial c_3}{\partial x}\bigg|_{x=x_3}, \\
&D_1 \frac{\partial c_3}{\partial x}\bigg|_{x=x_4} = D_{eff} \frac{\partial c_{eff}}{\partial x}\bigg|_{x=x_4}.
\end{aligned}
\tag{3}
$$

The coefficient α in Eq. (2) characterizes the concentration jump on the boundary inclusion – matrix and the coefficient χ denotes that there also exists a concentration jump on the boundary matrix-effective medium. D_{eff} can be determined from the systems of equations (2) and (3) inserting into them (1) and demanding the existence of nontrivial solution.

To simplify this equation, we assume that $x_4 - x_1 = 1$, $x_3 - x_2 = f$, $x_2 = (1 - f)/2$ and $x_3 = (1 + f)/2$. The obtained result is then

$$
D_{eff} = \frac{D_1 \chi}{\left((1-f) + f \dfrac{D_1}{D_2 \alpha}\right)}
\tag{4}
$$

with

$$
\chi = \frac{1}{\big((1-f) + \alpha f\big)}.
\tag{5}
$$

In order to determine the coefficient α, one additional equation is needed. We get this equation on the condition that the average particles concentration in the representative region should be equal to the average particles concentration in the effective medium of the same length. Thus, we have

$$
\overline{c}_{eff} = \int_{x_1}^{x_4} c_{eff}\,dx = \int_{x_1}^{x_2} c_1 dx + \int_{x_2}^{x_3} c_2 dx + \int_{x_3}^{x_4} c_3 dx = \overline{c_1} + \overline{c_2} + \overline{c_3}.
\tag{6}
$$

Inserting c_1, c_2 and c_3 from (1) into (6), we receive

$$
\alpha = \frac{\overline{c_1}}{\overline{\overline{c_2}}}.
\tag{7}
$$

Eq. (6) has been postulated in our papers [1, 2].

Finally, for the effective diffusion coefficient D_{eff} we have

$$
D_{eff} = \frac{D_1}{\left(1 - f + \dfrac{\overline{c_2}}{\overline{\overline{c_1}}} f\right)\left((1-f) + \dfrac{D_1 \overline{c_1}}{D_2 \overline{\overline{c_2}}} f\right)}.
\tag{8}
$$

If $\overline{c_1} = \overline{c_2}$, we get a well known result

$$D_{eff} = \frac{D_1 D_2}{D_1 f + D_2 (1-f)} . \tag{9}$$

3. Quasi-One-Dimensional Diffusion Example

We consider there quasi-one-dimensional diffusion in the model sample of channels and spikes with the diffusion coefficient *D* (Fig. 2).

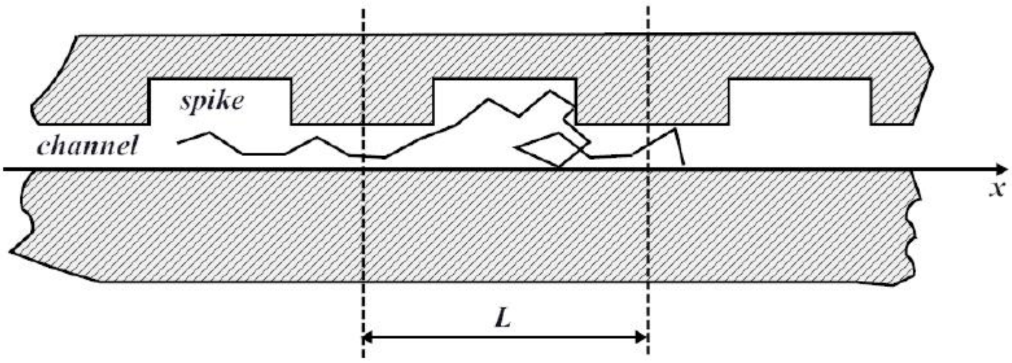


Figure 2. Quasi- one-dimensional sample, *L* – period

One representative element is shown on Figure 3.

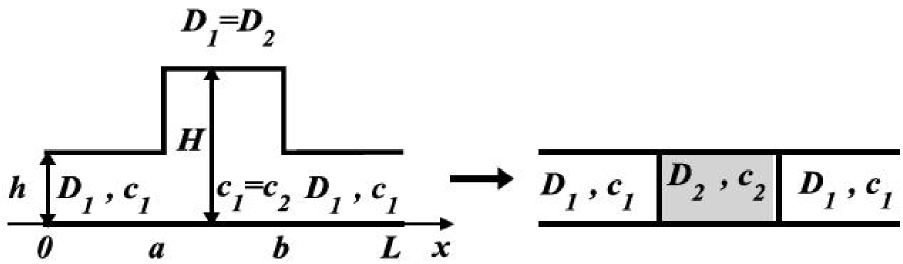


Figure 3. On the left – representative element; the spike begins at *a* and ends at *b*, *H* – the height of spike, *h* – channel height, *L* – the length of representative element. On the right – equivalent quasi-one-dimensional model

A particle diffusing through the sample takes a long time in spikes. If the height of the spike grows, the diffusing particle will be trapped in the spike for an extended time. It is obvious that the effective diffusion coefficient will decrease because of spikes. We can approximate the particle diffusion in the spike (as soon as we are interested in the one-dimensional diffusion through the sample) as one-dimensional diffusion with the different diffusion coefficient, which we denote D_2. The diffusion coefficient in the channel is denoted further by D_1. Thus, we replace the representative element by some effective quasi-one-dimensional media (Fig. 3).

The concentration of diffusing particles in the region (*a, b*) will differ from that in the channel. We denote this concentration by c_2. (Henceforth, we denote average concentrations as c_1 and c_2.)

Now we can calculate effective diffusion coefficient through the sample as

$$\frac{D_{eff}}{D_1} = \frac{1}{\left(1 - f + \frac{c_2}{c_1} f\right)\left((1-f) + \frac{D_1 c_1}{D_2 c_2} f\right)} , \tag{10}$$

where $f = (b - a) / L$. We can assume that $c_2/c_1 = H / h$, then D_{eff} depends on H and h, D_{eff} also depends on the form of spike. A computer simulation is a possible way to obtain the D_{eff} and D_2 values. The results received in the computer simulation, may be used for the estimation of effective diffusion coefficients in practical cases.

We have simulated the diffusion of particles through a quasi-one-dimensional sample using the Monte Carlo method. The simulation has been carried out in the representative element with the periodic boundary conditions along the x axes. Inside the representative volume, particles are reflected from the channel and spike borders. The mean free path of particle l has been chosen $l = 0.2$. The particle performs two-dimensional random walking with the equal probability to jump to any direction. D_{eff} is calculated by formula:

$$D_{eff} = \frac{< x^2 >}{2N\tau}, \tag{11}$$

where $< x^2 >$ is the mean square displacement of a diffusing particle, N – the number of jumps, τ – jump time, $D_1 = l^2 / (4\tau)$. The number of diffusing particles is 10000, and one particle history has been varied from 50000 to 600000 jumps long. The estimated error is 3%. The results of computer simulation are presented in Fig. 4.

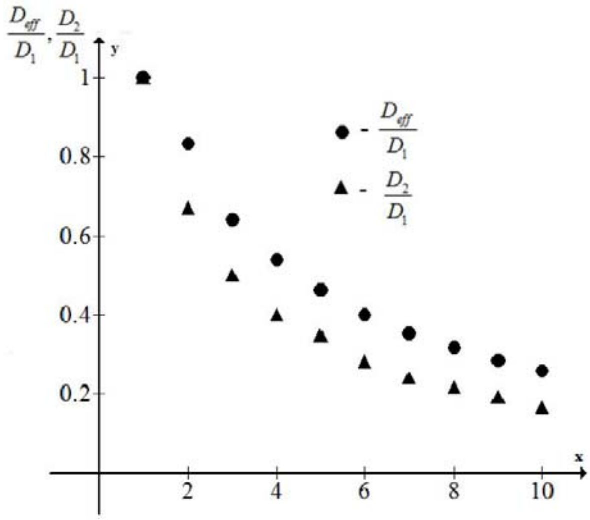


Figure 4. Results of computer simulation. D_{eff}/D_1 (circles) and D_2/D_1 (triangles) dependence on H/h. $h = 1$, $f = 0.4$, $a = 3$, $b = 7$, $L = 10$

It is seen from Fig. 4, that D_{eff} and D_1 are strongly correlated.

4. Conclusions

A closed formula for the effective diffusion coefficient in the heterogeneous media with the periodically distributed inclusions has been developed. It has been shown that the previously[1,2] accepted ad-hoc assumption about concentrations on the boundary matrix-inclusion follows in a self consistent way from the effective medium approximation. Simple quasi-one-dimensional models can be considered as one-dimensional in an effective manner. Simple computer simulation can help to determine the necessary parameters for the construction of practically usable effective diffusion coefficients.

References

1. Kalnin, J.-R., Kotomin, E.A., Maier, J. Calculations of the effective diffusion coefficient for inhomogeneous media, *J. Phys. Chem. Solids,* Vol. 63, 2002, pp. 449–456.
2. Jamnik, J., Kalnin, J.-R., Kotomin, E.A., Maier, J. Generalised Maxwell-Garnett equation: application to electrical and chemical transport, *Physical Chemistry Chemical Physics,* Vol. 8(11), 2006, pp. 1310–1314.

Received on the 22nd of July, 2011

Computer Modelling & New Technologies, 2011, Volume 15, No. 2
Transport and Telecommunication Institute, Lomonosov Str.1, Riga, LV-1019, Latvia

Authors' index

Yuri N. Shunin (born in Riga, March 6, 1951)

- Vice-Rector on Academic Issues (Information Systems Management Institute), professor, Dr.Sc.Habil., Member of International Academy of Refrigeration
- Director of Professional Study Programme Information Systems (Information Systems Management Institute)
- Director of Master Study Programme Computer systems (Information Systems Management Institute)
- **University studies**: Moscow physical and technical institute (1968–1974)
- Ph.D. (physics & mathematics) on solid state physics (1982, Physics Institute of Latvian Academy of Sciences), Dr.sc.habil (physics & mathematics) on solid state physics (1992, Ioffe Physical Institute of Russian Academy of Sciences)
- **Publications:** 430 publications, 1 patent
- **Research activities**: solid state physics, physics of disordered condensed media, amorphous semiconductors and glassy metals, semiconductor technologies, heavy ion induced excitations in solids, mathematical and computer modelling, system analysis

Igor V. Kabashkin (born in Riga, August 6, 1954)

- Vice-rector for Research and Development Affairs of Transport and Telecommunication Institute, Professor, Director of Telematics and Logistics Institute
- PhD in Aviation (1981, Moscow Institute of Civil Aviation Engineering) Dr.sc.habil. in Aviation (1992, Riga Aviation University), Member of the International Telecommunication Academy, Member of IEEE, Corresponding Member of Latvian Academy of Sciences (1998)
- **Publications:** 445 scientific papers and 67 patents
- **Research activities:** information technology applications, operations research, electronics and telecommunication, analysis and modelling of complex systems, transport telematics and logistics

Juris-Roberts Kalnin (born in Riga, Latvia, 1942)

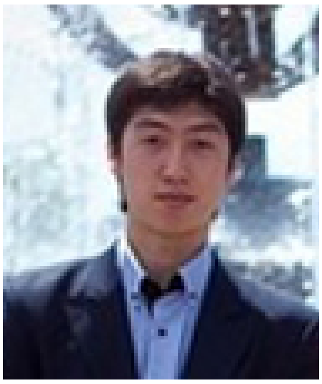
- Dr.habil.phys, Associate Professor, leading researcher,
  Ventspils University College
- **University Studies:** 1960–1965 M.S.+B.S. degrees, Dept. of Theoretical Physics, University of Latvia
- **Research activities:**
  Theoretical physics. Transport in heterogeneous media and composite materials. Kinetics of bimolecular reactions in condensed matter with focus on many-particle processes. The kinetics of the generation, annealing, migration of defects in solids under irradiation. Large-scale computer simulations of processes in condensed matter

Sharif E. Guseynov (born in January 13, 1967)

- Leading Researcher, Professor
- Institute of Mathematical Sciences and Information Technologies, University of Liepaja
- **University studies:** M. V. Lomonosov Moscow State University, Department of Computing Mathematics and Cybernetics
- **Research activities:** Main Research Fields: Inverse and Ill-posed Problems, Partial Differential Equations, Equations of Mathematical Physics, Integral Equations, Operator Equations, Nonlinear Physics, Mathematical Modelling, Functional Analysis and Applications, Computational Methods, Mathematical Economics, Optimal Control and Optimization Methods
- **Publications:** 198 papers

Vladimir D. Ten (born in June 24, 1948)

- Associate Professor of Kazakh-British Technical University, the Chair of Computing Engineering, Almaty
- **Education:**
  1965–1970 – Student of Novosibirsk State University, Department of Math & Mech., Novosibirsk
  1973–1976 – Ph.D. student at the Chair of Higher Algebra of Moscow State University, Moscow
  Ph.D. Dissertation: "Polynomial Subsets in Algebras." (Moscow State University)
- **Research activities**:
  Ring theory. Varieties of Associative rings and Algebras. Invariant Theory of Associative Rings and Algebras under Action of Finite Group of Automorphisms, Rings with Involution
  Mathematical Methods in Cryptography. Applications of Ring Theory and Commutative Algebra in Cryptography



Alexander V. Ten (born in December 31, 1987)

- Manager of IT-department of Halyk-Bank, Almaty
- **Education:**
  2004–2008 – Student of Kazakh National University, Department of Math & Mech., Almaty
  2008–2010 – Master-student of Kazakh National University, Department of Math & Mech., Almaty
  Master Degree Dissertation: "Semi-definable Cryptosystem without Repetition"

Alexander N. Kovantsov (born in 1949)

- Riga Technical University, Department of Engineering Mathematics , Associate Professor, Dr.math (1993)
- **University studies:**
  1971 – Kiev State University, Faculty of Mathematics and Mechanics
  Since 1999 – Information System Management Institute; Assoc. Professor of Natural Sciences and Computer Technologies Department
  Riga Aviation University (RAU) Higher Mathematics Department:
  1987 – Associate Professor RAU
  1985 – Senior lecturer RAU
  1975 – Assistant RAU

Dimitri Golenko-Ginzburg (born in Moscow, November 24, 1932)

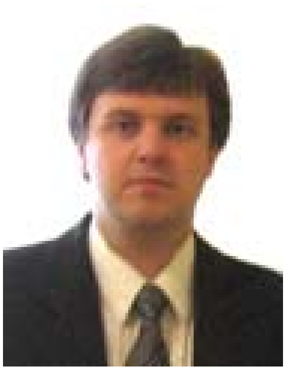
- Professor, Industrial Engineering and Management Department, Ariel University Center of Samaria, Ariel, Israel
- Professor, Industrial Engineering and Management Department, Ben-Gurion University of the Negev, Beer-Sheva, Israel (1988–2004)
- Full Professor (tenured position), Institute of National Economics, Uzbekistan Ministry of Higher Education, Tashkent (1977–1979)
- Full Professor (tenured position), Moscow Economico-Statistical Institute, USSR Ministry of Higher Education, Moscow (1967–1977)
- **University study**: Moscow State University, Department of Mathematics (1954–1958); Moscow Institute of National Economics, Department of Economics (1950–1954)
- Ph.D. (Applied Mathematics) on simulating probability processes on computers (1962, Moscow Physico-Technical Institute)
- **Publications:** 15 books, about 500 refereed articles and refereed letters in scientific journals
- **Research activities**: production planning and control, planning and controlling network projects, industrial scheduling, managing reliability and safety

Avner Ben-Yair (born in Moscow, May 19, 1961)

- Lecturer, Industrial Engineering and Management Department, SCE – Sami Shamoon College of Engineering, Beer-Sheva, Israel (2002–2010)
- Health and Safety-at-Work Engineer, Baran Engineering Co., Israel (1996–1999)
- Health and Safety-at-Work Engineer, AVX Israel, Jerusalem (1986–1996)
- **University study**: Ben-Gurion University of the Negev, Beer-Sheva, Israel (1999–2001); Moscow Polygraphic Institute, Department of Mechanical Engineering (1979–1985)
- Ph.D. (Industrial Engineering and Management) on harmonization models in strategic management and safety engineering (2004, Ben-Gurion University of the Negev)
- **Publications:** 90 refereed articles and refereed letters in scientific journals
- **Research activities**: system performance and effectiveness, reliability and failure analysis, fault tree analysis, trade-off optimization models for organization systems, risk analysis and contingency planning, maintainability and hazard analysis techniques, economic aspects of safety, production planning, scheduling and control, strategic management, network models structure and project scheduling, cost optimization and PERT-COST models

Nicholas A. Nechval

- University of Latvia, Professor
- **University study:** Riga Civil Aviation Engineers Institute (Faculty of Electrical Engineering) cum laude, 1965
- Ph.D. degree in automatic control and systems engineering, Riga Aviation University, Riga, Latvia, 1969**;** Dr.habil.sc.ing. (radio engineering), Riga Aviation University (RAU), 1993
- **Research activities:** Mathematics, Stochastic Processes, Pattern Recognition, Digital Radar Signal Processing, Operations Research, Statistical Decision Theory, Adaptive Control
- **Professional Activities and Memberships:** Scientific Society (on problem of Protection from Biodeterioration), Russian Academy of Sciences, (since 1987); Latvian Association of Professors, (since 1997)



Konstantin Nechval (born in March 5,1975)

- Mg.sc.ing., Lecturer of Transport and Telecommunication Institute
- **University study:** Bachelor of science in engineering (RAU, 1996). Engineer qualification in maintenance of aircraft and engines (RAU,1998); Master degree in engineering (RAU,1999), PhD studies in Riga Technical University (2000)
- **Research activities:** Operation Systems: Dos, Linux, Windows; Programming: Delphi , BASIC, HTML, PHP; LISP, C# ; Aplications: MatLab, MatCad, Matematica, StatGraphics, Derive, AutoCad, MS Office, Open Office

Jekaterina A. Baranova (born in Riga, July 6, 1987)

- **University study**: Dipl.-Math. at University of Latvia (2006–2011), Bachelor`s Degree in Economics at Transport and Telecommunication Institute (2007–2010)
- **Publications**: 1 publication
- **Scientific activities**: Reliability and Statistics in Transportation and Communication (RelStat`10). 20–23 October 2010

Maksat Kalimoldaev (born in June 18, 1957)

- Professor, Dr.sc.habil. (phys&math)
- Head of Laboratory of Mathematical Modelling and Cybernetics, Institute of Informatics and Control Problems
- **University studies:** Kazakhstan State University (1980)
- **Publications:** 80 papers

Yedilhan N. Amirgaliyev (born in March 1, 1959)

- Leading Researcher, Dr.sc.habil.eng., Assist. Professor
- Laboratory of Images Recognition and Decision Making, Institute of Informatics and Control Problems
- **University studies:** Kazakhstan State University
- **Publications:** 60 papers

Rustam Musabayev (born in July 1, 1984)

- Scientific secretary, Cand.Tech.Sci.
- Laboratory of Pattern Recognition and Decision Making, Institute of Informatics and Control Problems
- **University studies:** Kazakh National Technical University (2006)
- **Publications:** 40

CUMULATIVE INDEX

COMPUTER MODELLING and NEW TECHNOLOGIES, volume 15, No. 2, 2011

(Abstracts)

J. Sugier. Adjusting Markov Models to Changes in Maintenance Policy for Reliability Analysis, *Computer Modelling and New Technologies*, vol. 15, No 2, 2011, pp. 7–20.

Equipment deterioration can be analysed with discrete state-transition models which are used as a foundation for numerical evaluation of various reliability and operational parameters. In this paper we discuss an approach in which a system with scheduled inspections and possible repair activities is described by the discrete state-transition deterioration model. The model can be translated into a semi-Markov process which, after solving, yields numerous reliability characteristics including average equipment life, its deterioration rate represented by so called life curve, probability of failure within given time horizon, etc. The analysis is performed for particular maintenance policy which has been integrated in the model and by adjusting the model to specific changes in this policy (e.g., modifications of repair frequencies) their consequences for system reliability can be evaluated. In the text we present briefly the methodology behind model creation and concentrate on its one specific aspect: automatic adaptation of the model to the adjusted maintenance policy with modified frequencies of repairs. In particular, it is shown that such adaptation can be reduced to the adjustment of transition probabilities that are found in the model, and such adjustment can be numerically implemented with one of the standard root-finding algorithms. Upon presentation of the adjustment method, its use is illustrated by a description of the computer tool which helps in evaluating reliability and financial effects of changes in maintenance policy of an ageing equipment.

Keywords: *state-transition deterioration model, semi-Markov process, reliability analysis, root-finding algorithm, maintenance analysis*

A. Ben-Yair, D. Golenko-Ginzburg. Heuristic Preference Rules in Multiroute Scheduling, *Computer Modelling and New Technologies*, vol. 15, No 2, 2011, pp. 21–29.

A systematic account and comparative analysis of the efficiency of various statistical decision-making rules in scheduling theory, is undertaken. We will describe various adaptive randomisation preference rules in order to use the latter in industrial scheduling.

Keywords: *Heuristic preference rules; mn-permutations; SPT, LRT, FIFO and PC rules; Randomization preference rules*

K. N. Nechval, N. A. Nechval, M. Purgailis, V. F. Strelchonok, G. Berzins, M. Moldovan. New Approach to Pattern Recognition Via Comparison of Maximum Separations, *Computer Modelling and New Technologies*, vol. 15, No 2, 2011, pp. 30–40.

Fisher's linear discriminant analysis is a widely used multivariate statistical technique with two closely related goals: discrimination and classification. The technique is very popular among users of discriminant analysis. Some of the reasons for this are its simplicity and un-necessity of strict assumptions. In its original form, proposed by Fisher, the method assumes equality of population covariance matrices, but does not explicitly require multivariate normality. However, optimal classification performance of Fisher's discriminant function can only be expected when multivariate normality is present as well, since only good discrimination can ensure good allocation. In practice, we often are in need of analysing input data samples, which are not adequate for Fisher's classification rule, such that the distributions of the groups are not multivariate normal or covariance matrices of those are different or there are strong multi-nonlinearities.

This paper proposes a new approach to pattern recognition based on comparison of maximum separations in the input data samples, where maximum separations are determined via Fisher's separation criterion. The approach represents the improved pattern recognition procedure that allows one to take into account the cases which are not adequate for Fisher's classification rule. Moreover, it allows one to classify sets of multivariate observations, where each of the sets contains more than one observation. For the cases, which are adequate for Fisher's classification rule, the proposed approach gives the results similar to that of Fisher's classification rule. Illustrative examples are given.

Keywords: *Fisher's separation criterion, input data samples, maximum separation, pattern recognition*

COMPUTER MODELLING and NEW TECHNOLOGIES, 15.sējums, Nr. 2, 2011
(Anotācijas)

J. Sugiers. Markova modeļu pielāgošana uzturēšanas politikas izmaiņām drošuma analīzei, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 7.–20. lpp.

Iekārtu nodilums var tikt analizēts ar diskrētiem valsts-tranzīta modeļiem, kas tiek lietoti kā bāze dažādu drošuma un ekspluatācijas parametru skaitliskai izvērtēšanai. Šajā rakstā autori izskata pieeju, kurā sistēma ar plānotām pārbaudēm un iespējamiem labošanas pasākumiem tiek aprakstīta ar diskrētu valsts-tranzīta nodiluma modeli. Modelis var tikt pārvērsts daļējā Markova procesā, kurš pēc atrisināšanas dod daudz uzticamības iezīmes, ieskaitot vidējo iekārtas kalpošanas, tās bojāšanās ātrumu, ko pārstāv tā saucamā dzīves līkne, bojājuma varbūtība noteiktā laika periodā, utt. Analīze tiek veikta konkrētai izpildes politikai, kas ir integrēts modelī, un, piemērojot modeli konkrētām izmaiņām šajā politikā (piemēram, remonta frekvenču izmaiņas), var tikt novērtētas tās konsekvences sistēmas drošumam. Tekstā autori īsumā parāda metodoloģiju bez modeļa radīšanas un koncentrējas uz tā īpašo aspektu: modeļa automātisko adaptāciju pielāgotajai uzturēšanas politikai ar labojumu modificētām frekvencēm. Līdz ar piemērošanas metodes prezentāciju, tā lietošana tiek ilustrēta ar datorizēta līdzekļa aprakstu, kas palīdz drošuma novērtēšanai un finansiālās ietekmes uz novecošanās iekārtu uzturēšanas politikas izmaiņām.

Atslēgvārdi: *valsts-tranzīts, nodiluma modelis, daļējs Markova process, drošuma analīze, saknes atrašanas algoritms, uzturēšanas analīze*

A. Ben-Jears, D. Golenko-Ginzburgs. Heiristiskie priekšrokas noteikumi daudzmaršrutu plānošanā, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 21.–29. lpp.

Rakstā tiek veikta dažādu statistisko lēmumu pieņemšanas noteikumu plānošanas teorijā efektivitātes sistemātisks aprēķins un salīdzinoša analīze. Autori apraksta dažādus pielāgošanās randomizācijas priekšnoteikumus, lai tos izmantotu rūpnieciskajā plānošanā.

Atslēgvārdi: *heiristiskie priekšrokas jeb izvēles noteikumi, mn-permutācijas, SPT, LRT, FIFO un PC noteikumi, randomizācijas priekšrokas jeb izvēles noteikumi*

K. N. Nečvals, N. A. Nečvals, M. Purgailis, V. F. Strelčonoks, G. Bērziņš, M. Moldovans. Jauna pieeja modeļa atpazīšanai, izmantojot maksimuma dalījumu salīdzināšanu, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 30.–40. lpp.

Fišera lineārā diskriminanta analīze tiek plaši izmantota daudzfaktoru statistiskā metode ar diviem cieši saistītiem mērķiem: diskriminācija un klasifikācija. Metode ir ļoti populāra diskriminanta analīzes lietotāju vidū. Daži iemesli tam ir to vienkāršība un striktu pieņēmumu nevajadzība. Tās oriģinālajā veidā, ko ierosina Fišers, metode paredz metode paredz iedzīvotāju kovariācijas matricas vienlīdzību, bet nav pilnībā neprasa daudzfaktoru saskaņu ar normu. Tomēr Fišera diskriminanta funkciju optimāla klasifikācijas izpilde var notikt tikai tad, kad daudzfaktoru normalitāte ir klāt, kā arī, jo tikai laba diskriminācija var nodrošināt labu sadalījumu.

Šajā rakstā autori ierosina jaunu modeļa atpazīšanas pieeju, balstītu uz maksimuma sadalījumu salīdzinājumu ievades datu paraugos, kur maksimuma dalījumi tiek noteikti, izmantojot Fišera dalījuma kritēriju. Pieeja ir uzlabota modeļa atpazīšanas procedūra, kas ļauj ņemt vērā gadījumus, kas nav piemēroti Fišera klasifikācijas noteikumiem. Attiecībā uz gadījumiem, kas ir piemēroti Fišera klasifikācijas noteikumiem piedāvātā pieeja dod rezultātus, līdzīgus Fišera klasifikācijas noteikumiem. Rakstā tiek doti ilustratīvi piemēri.

Atslēgvārdi: *Fišera dalījuma kritērijs, ievades datu piemēri, maksimuma dalījums, modeļa atpazīšana*

Š. Guseinovs, J. Baranova. Viena matemātiskā modeļa izpēte dalāmas produkcijas apjoma palielināšanai ar ierobežotiem resursiem, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 41.–52. lpp.

Ekonomiski matemātiski optimizējoša problēma, kas skar sadalāmās ražošanas maksimālo daudzumu (piemēram, diētiskās maizes ražošana) tiek formulēta un izpētīta ierobežotu resursu apstākļos. Rakstā tiek izstrādāts izpētītās problēmas matemātiskais modelis, balstīts uz Leontjeva modeli. Modeļa konstruēšanā tiek pieņemti analītiski lēmumi, un tiek izvēlēta pašregulācijas metode, kas pastāvīgi atkarīga no visiem sākotnējiem datiem un kā metode optimāla rezultāta atrašanai. Programmēšanas produkts ir izstrādāts, lai veiktu skaitlisko eksperimentu ar īstiem sākotnēji apstiprinātiem ekonomiskiem datiem. Uzbūvētajam matemātiskajam modelim ir praktiska vērtība un kā jebkura sadalāmā ražošana (brīvai vai ne tik brīvai) var tikt lietota; piemēram, maizes izstrādājumi, sienu un grīdu sildītāji, papīra produkcija, cements, krāsas, utt.

Atslēgvārdi: *Leontjeva modelis, dualitātes teorija, optimāli plānota ražošana*

A. Kovancovs. Projektora sistēmas, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 53.–58. lpp.

Diferenciālvienādojumu attīstības sistēmas tiek piemērotas imitācijas modelēšanas dažādās problēmās. Bet jebkurš skaitļošanas risinājums prasa analītisku izvērtēšanu. Šajā sakarā projektora tehnika ir ļoti efektīva. Nelineāras diferenciālvienādojumu sistēmas risinājums ievērojami var tikt vienkāršots, ja sistēma tiek transformēta Džordana formā, t.i., tādā veidā, kad sistēma sastāv no neatkarīgas sistēmas funkciju blokiem.

Atslēgvārdi: *nelineārs diferenciālvienādojums, projektora tehnika, Džordana matrica*

A. Tens, V. Tens. Eiklīda gredzeni un kriptosistēmas bez atkārtošanās, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 59–67. lpp.

Tiek konstruētas kriptosistēmas bez atkārtošanās. Tas nozīmē, ka šifrēšanas algoritms atļauj jebkura teksta šifrēšanu tā, ka pēc šifrēšanas visi alfabēta elementi ir skaidri.

Atslēgvārdi: *Eiklīda gredzeni, kriptosistēma, vienkāršs teksts, ciparu teksts, slepena atslēga, alfabēts*

M. Kalimondajevs, J. Amirgalijevs, R. Musabajevs, Runas signāla intonācijas sintēzes metode balstīta uz splainu tuvināšanos, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 68–71. lpp.

Runas signāla intonācijas komponenta sintēzes metode balstīta uz līklīniju – matemātiski aprēķinātas līknes, kuras vienmērīgi savieno atsevišķus intonācijas kontūra kontroles punktus. Šī metode tiek lietota, kad tiek īstenota runas signāla kompilācijas sintēzes sistēma. Sintezētā runas signāla parametriskais atspoguļojums *UPL* valodā, kopumā ar tās intonācijas komponentu pēc līklīnijas aproksimācijas metodi, uzlabo tuvināšanās kvalitāti salīdzinājumā ar esošajām metodēm.

Atslēgvārdi: *runas sintēze, splainu tuvināšanās, fundamentālā frekvence, intonācijas kontūrs, kompilācijas sintēze*

J.-R. Kalniņš. Difūzijas process kvazi-viendimensijas struktūrās, *Computer Modelling and New Technologies,* 15.sēj., Nr.2, 2011, 72–75. lpp.

Rakstā tiek aprēķināts efektīvs difūzijas koeficients divpakāpju viendimensionālā modelī ar ieslēgumu periodisku sadalījumu efektīvas vides aproksimācijā un tiek veidots vispārinājums par kvazi-viendimensionālo gadījumu.

Atslēgvārdi: *efektīva difūzija, efektīva vide, metode*

COMPUTER MODELLING
&
NEW TECHNOLOGIES

ISSN 1407-5806 & ISSN 1407-5814(on-line)

Host Organization:
Transport and Telecommunication Institute

Supporting Organizations:
Latvian Transport Development and Education Association
Latvian Academy of Sciences
Latvian Operations Research Society

THE JOURNAL IS DESIGNED FOR PUBLISHING PAPERS CONCERNING THE FOLLOWING FIELDS OF RESEARCH:

- mathematical and computer modelling
- mathematical methods in natural and engineering sciences
- physical and technical sciences
- computer sciences and technologies
- material science, solid state physics and chemistry
- nanomaterials, nanodevices and nanotechnologies
- aviation and aerospace technologies
- electronics and telecommunication
- transport and logistics
- economics and management
- innovative education
- social sciences

In journal articles can be presented in English. All articles are reviewed.

EDITORIAL CORRESPONDENCE

Transporta un sakaru institūts (Transport and Telecommunication Institute)
Lomonosova iela 1, LV-1019, Riga, Latvia. Phone: (+371) 67100593. Fax: (+371) 67100535
E-mail: journal@tsi.lv, www.tsi.lv

COMPUTER MODELLING AND NEW TECHNOLOGIES, 2011, Vol. 15, No.2
ISSN 1407-5806, **ISSN** 1407-5814 (on-line: www.tsi.lv)
Scientific and research journal of Transport and Telecommunication Institute (Riga, Latvia)
The journal is being published since 1996.

PREPARATION OF CAMERA-READY TYPESCRIPT: COMPUTER MODELLING AND NEW TECHNOLOGIES

1. In order to format your manuscript correctly, see the Page Layout Guideline for A4 (21 cm x 29,7 cm) paper size. Page Layout should be as follows: Top – 3 cm, Bottom – 3 cm, Left – 3 cm, Right – 3 cm.
2. Maximum length for the article is 10 pages.
3. **Application of other Styles with the exception of Normal is impossible!**
4. ***Articles*** should be Times New Roman typeface, single-spaced.
5. The article should include:
   - title;
   - author's name(s) and information (organisation, city, country, present address, phones, and e-mail addresses);
   - abstract (100–150 words);
   - keywords (max. about six);
   - introduction – clearly explaining the nature of the problem, previous work, purpose and contribution of the research;
   - description of proper research;
   - conclusion section (this is mandatory) which should clearly indicate advantages, limitations and possible applications;
   - references.

   Attention! First name, last name, the title of the article, abstract and keywords must be submitted in the English and Latvian languages (in Latvian it is only for Latvian authors) as well as in the language of the original (when an article is written in different language).
6. The text should be in clear, concise English (or other declared language). Please be consistent in punctuation, abbreviations, spelling (***British English***), headings and the style of referencing.
7. ***The title of the article*** – 14 point, UPPERCASE, style Bold and centered.
8. ***Author's names*** – centered, type size 12 point, Upper and lower case, style Bold Italic.
9. ***Author's information*** – 10 point, Upper and lower case, style Italic, centered.
10. ***Abstract and keywords*** – 8 point size, style Normal, alignment Justify.
11. ***The first level Headings*** – 11 point, Upper and lower case, style Bold, alignment Left. Use one line space before the first level Heading and one line space after the first level Heading.
12. ***The second level Headings*** – 10 point, Upper and lower case, style Bold, alignment Left. One line space should be used before the second level Heading and 1/2 line space after the second level Heading.
13. ***The third level Headings*** – 10 point, Upper and lower case, style Italic, alignment Left. One line space should be used before the second level Heading and 1/2 line space after the third level Heading.
14. ***Text*** of the article – 10 point, single-spaced, alignment Justify.
15. The set of ***formulas*** on application of fonts, signs and a way of design should be uniform throughout the text. The set of formulas is carried out with use of editors of formulas MS Equation 3.0 or MathType. The formula with a number – the formula itself should be located on the left edge of the text, but a number – on the right. Font sizes for equations are: 11pt – full, 7pt – subscripts/superscripts, 5pt – sub- subscripts/superscripts, 16pt – symbols, 11pt – subsymbols.
16. All ***Figures*** – must be centered. Figure number and caption always appear below the Figure, type size 8 point.

Figure 1. This is figure caption

Diagrams, Figures and Photographs – must be of high quality, в in format *.TIFF, *.JPG, *.BMP with resolution not less than 300 dpi. Also formats *.CDR, *.PSD are possible. Combination of Figures in format, for instance, *.TIFF with elements of the in-built Figure Editor in MS Word is prohibited.

17. ***Table Number and Title*** – always appear above the Table. Alignment Left. Type size 8 point. Use one line space before the Table Title, one line space after the Table Title and 1/2 line space after the Table.

Table 1. This is an example of a Table

Heading	**Heading**	**Heading**
Text	Text	Text
Text	Text	Text

18. ***References*** in the text should be indicated by a number in square brackets, e.g. [1].
    References should be numbered in the order cited in the manuscript. The correct format for references is the following:

Article: author, title, journal (in italics), volume and issue number, year, inclusive pages

Example: 1. Amrahamsson, M., Wandel, S. A Model of Tearing in Third – Party Logistics with a Service Parts Distribution Case Study, *Transport Logistics*, Vol. 1, No 3, 1998, pp. 181-194.

Book: author, title (in Italics), location of publishers, publishers, year

Example: 2. Kayston, M. and Fried, W. R. *Avionic Navigation Systems.* New York: John Wiley and Sons Inc, 1969.

Conference Proceedings: author; title of an article; proceedings (in italics); title of a conference, date and place of a conference; publishing house, year, pages.

Example: 3. Canales Romero J. A First Step to Consolidate the European Association of Aerospace Students in Latvia (Presented by the Munich Local Group). In: *Research and Technology – Step into the Future: Program and Abstracts. Research and Academic Conference, Riga, Latvia, April 7-11, 2003, Transport and Telecommunication Institute.* Riga: TTI, 2003, p. 20.

19. ***Authors Index***
    Editors form the author's index of a whole Volume. Thus, all contributors are expected to present personal colour photos with the short information on the education, scientific titles and activities.
20. ***Acknowledgements***
    Acknowledgements (if present) mention some specialists, grants and foundations connected with the presented paper. The first page of the contribution should start on page 1 (right-hand, upper, without computer page numbering). Please paginate the contributions, in the order in which they are to be published. Use simple pencil only.
21. **Articles poorly produced or incorrectly formatted may not be included in the proceedings**.