MATRIX "FEATURE VECTORS" AND GROUPING OPERATORS IN PATTERN RECOGNITION

© Volodymyr Donchenko, Taras Zinko

Taras Shevchenko National University of Kyiv
The Faculty of Cybernetics
e-mail: *voldon@bigmir.net, tzinko@ukr.net*

Abstract*. Problem of grouping information: recovering function, represented by its observations, and the of classification (problem) clusterization problem, — is of great importance for applied research. Choice of math object which represent the object under investigations largely determines the effectiveness: scalars, vectors or objects of other kinds. Such choice is determined by the richness of mathematical structures within which "representatives" are investigated. Euclidean spaces R^n are common in this choosing. Euclidean spaces of $R^{m\times n}$ of all $m \times n$ matrices are natural as a math structure for "representatives", but the handling technique for such spaces is poorer in comparison with vector space. Just the development of the technique handling" for Euclidean space of $R^{m\times n}$, including SVD and Moore-Penrose inversion for the linear operators, constructive construction of orthogonal projectors and grouping operators for matrix spaces is the subject of the article. Important "grouping statements" about minimal ellipsoid, which covers elements of fixed sequence of matrices in $R^{m\times n}$ is represented. This statement generalize correspondent results for real valued vectors. "Grouping statements" is proposed to be the base for constructing correspondence distance in solving clusterization problem.*

Introduction

The problem of grouping the information (grouping problem) is the fundamental problem of applied investigations. It appears in various forms and manifestations. All of them eventually are reduced to two forms. Namely, these are: the problem of recovering the function represented by their observations and the problem of clustering, classification and pattern recognition. State of art in the field is represented perfectly in [23, 25, 11, 10, 3].

It's opportune to mark what the information regarding the object or a collection of similar object is exposed to aggregating is. It is of principal importance that an object is considered as a set of its main components and fundamental for the object ties between them. Such consideration and only this one enable application of the math in object description, namely, for math modelling. It is due the fact that after Georg Cantor the objects of investigation in math (math structures) are the sets plus "ties" between its elements. There are only four (may be, five) fundamental mathematical means to describe these "ties". Namely, these are: relations, operations, functions and collections of subsets (or combinations of mentioned above). Thus, the mathematical description of the object (mathematical modelling) can not be anything other than representing the object structure by the means of mathematical structuring. It is applicable to the full extent to that objects which indicated by the term "complex system". A "complex system" should be understanding and, correspondingly, determined, as an objects with complex structure (complex "ties"). Namely, when reading attentively manuals by the theme (see, for example, [9, 26]) one could find correspondent allusions. It is reasonable understanding of "complex systems" instead of the its understanding as the "objects, consisting of numerous parts, functioning as an organic whole".

So, math modelling is designing in math "parts plus ties", which reproduce "part plus ties" in reality.

So it is principal question in math modelling which math objects represents "part" of the object and which the "ties" ones. The math object — representative should be chosen in such a way that variety of math structuring means were sufficient to convey the object structure.

It is commonly used approach for designing objects — representative to construct them as an finite ordered collection of characteristics: quantitative (numerical) or qualitative (non numerical). Such ordered collection of characteristics is determined by term cortege in math. Cortege is called vector when its components are numerical. In the function recovering problem objects — representatives are vectors and functions are used as a rule to design correspond mathematical "ties". In clustering and classification problem the collection may be both qualitative and quantitative. In last case correspond collection is called feature vector. It is reasonable to note that term "vector" means more, than simply ordered numerical collection. It means that curtain standard math "ties" are applicable to them. These "ties" are adjectives of the math structure called Euclidean space denoted be R^n. Namely these are: linear operations (addition and scalar multiplying), scalar product and correspond norm and distance.

It is noteworthy to say, that this variant of Euclidean space R^n is not unique: the space $R^{m\times n}$ of all matrices of a fixed dimension $m \times n$ represents alternative example. The choice of the R^n space as "environmental" math structure is determined by perfect technique developed for manipulation with vectors. These include classical matrix methods and classical linear algebra methods. SVD-technique and methods of Generalized or Pseudo Inverse according Moore-Penrose are comparatively new elements of linear matrix algebra technique [24] (see, also, [1, 2]). Outstanding impacts and achievements in this area are due to N.F Kirichenko (especially, [13, 18], see also [19]). Greville's formulas:forward and inverse -for pseudo inverse matrices, formulas of analytical representation for disturbances of pseudo inverse, - are among them. Additional results in the theme as to further development of the technique and correspondent applications one can find in [7, 19, 20, 21, 15, 6, 14, 22, 17].

As to technique designing for the Euclidean space $R^{m\times n}$ as "environmental" one see, for example [5]. Speech recognition with the spectrograms as the representative and the images in the problem of image recognition are the natural application area for the correspond technique.

As to the choice of the collection (design of cortege or vector) it is necessary to note, that good "feature" selection (components for feature vector or cortege or an arguments for correspond functions) determines largely the efficiency of the problem solution.

As noted above, the efficiency of problem solving group, the choice of representatives of right: space arguments or values of functions and suitable characteristics for features vectors. This phase in solving the grouping information problem must be a special step of the correspondent algorithm. Experience showed the effectiveness of recurrent procedures is largely determined just by successful selection of features vector. For correspond examples see,[12] with Ivachnenko's GMDH (Group Method Data Handling), [25] with Vapnik's Support Vector Machine. Further development of the recurrent technique one may find in [7, 20, 21, 15, 6, 14, 22]. The idea of nonlinear recursive regressive transformations (generalized neuron nets or neurofunctional transformations) due to Professor N. F. Kirichenko is represented in the works referred earlier in its development. Correspondent technique has been designed in this works separately for each of two its basic form f the grouping information problem. The united form of the grouping problem solution is represented here in further consideration. The fundamental basis of the recursive neurofunctional technique include the development of pseudo inverse theory in the publications mentioned earlier first of all due to Professor N.F. Kirichenko and his disciples.

The essence of the idea mentioned above is in the choice of the primary collection and changing it if necessary by standard recursive procedure. Each step of the procedure include detecting of insignificant components, excluding or purposeful its changing, control of efficiency of changes has been made. Correspondingly, the means for implementing the correspondent operations of the step must be designed. Methods of neurofunctional transformation (NfT) (generalized neural nets, nonlinear recursive regressive transformation: [7, 20, 21]).

1. Development of Pseudo Inverse Technique for matrices Euclidean spaces

The following are results that transfer basic features of describing the basic structures of Euclidean spaces [5] matrix Euclidean spaces. These are, first of all General Single Valued Decomposition (SVD) theorem and then determination of Pseudo Inverse (PdI) and designing the constructive methods for manipulating with basic structures within matrixes spaces on the base of the Pseudo Inverse. Such transfer make it necessary to introduce special objects and tools for handling them. Namely, these are matrix corteges and corteges operations.

First theorem below is the advanced form of SVD theorem for Euclidean spaces, which one can find in [5].

2. Matrices spaces and cortege operators

Theorem 1. *For an arbitrary linear operator between a pair of Euclidean spaces* $(E_i, (,)_i)$, $i = 1, 2$*:* $\wp_E : E_1 \to E_2$*, the collection of singularities* (v_i, λ_l^2), (u_i, λ_l^2), $i = \overline{1, r}$, $r = rank\wp_E$ *exists for the operators* $\wp_E^* \wp : E_1 \to E_1$, $\wp\wp_E^* : E_2 \to E_2$ *correspondingly, with a common for both operators* $\wp_E^*\wp, \wp\wp_E^*$ *set of Eigen values* $\lambda_l^2, i = \overline{1, r} : \lambda_{i-1} \geq \lambda_i > 0$, $\overline{i = 2, r}$ *such that*

$$\wp_E x = \sum_{i=1}^{r} \lambda_i u_i (v_i, x)_1, \qquad \wp_E^* y = \sum_{i=1}^{r} \lambda_i v_i (u_i, y)_2.$$

Besides, the following relations take place:

$$u_i = \lambda_i^{-1} \wp v_i, \quad i = \overline{1, r},$$

$$v_i = \lambda_l^{-1} \wp_E^* u_i, \quad i = \overline{1, r}.$$

3. SVD — technique for matrices spaces

We denote by $R^{(m \times n), K}$ – Euclidean space of all matrices K-corteges from $m \times n$ matrices: $\alpha = (A_1 \vdots ... \vdots A_K) \in R^{(m\times n),K}$ with a "natural" component wise trace inner product:

$$(\alpha, \beta)_{cort} = \sum_{k=1}^{K} (A_k, B_k)_{tr} = \sum_{k=1}^{K} tr A_k^T B_k,$$

$$\alpha = (A_1 \vdots ... \vdots A_K), \beta = (B_1 \vdots ... \vdots B_K) \in R^{(m\times n),K}.$$

We also denote by $\wp_\alpha : R^K \to R^{m \times n}$a linear operator between the Euclidean space, determined by the relation:

$$\wp_\alpha y = \sum_{k=1}^{K} y_k A_k, \alpha = (A_1 \vdots \dots \vdots A_K) \in R^{(m\times n),K}, \tag{1}$$

$$y = \begin{pmatrix} y_1 \\ \cdots \\ y_K \end{pmatrix} \in R^K.$$

Theorem 2. *Range $\Re(\wp_\alpha) = L_{\wp_\alpha}$, which is linear subspace of $R^{m\times n}$, is the subspace spanned on the components of cortege $\alpha = (A_1 \vdots \ldots \vdots A_K) \in R^{(m\times n),K}$, that determines $\wp_\alpha$:*

$$\Re(\wp_\alpha) = L_{\wp_\alpha} = L(A_1, \ldots, A_K).$$

Theorem 3. *Conjugate for the operator, determined by* (1) *is a linear operator, which, obviously, acts in the opposite direction: $\wp_\alpha^* : R^{m\times n} \to R^K$, and defined as:*

$$\wp_\alpha^* X = \begin{pmatrix} trA_1^T X \\ \cdots \\ trA_K^T X \end{pmatrix} = \begin{pmatrix} trX^T A_1 \\ \cdots \\ trX^T A_K \end{pmatrix}.$$

Theorem 4. *A product of two operators $\wp_\alpha^* \wp_\alpha : R^K \to R^K$ is a linear operator, defined by the matrix from the next equation:*

$$\wp_\alpha^* \wp = \begin{pmatrix} trA_1^T A_1, ..., trA_1^T A_K \\ \cdots \\ trA_K^T A_1, ..., trA_K^T A_K \end{pmatrix}. \tag{2}$$

Remark. Matrix defined by (2) is the 'Gram' matrix for the elements of the cortege $\alpha = (A_1 \vdots \ldots \vdots A_K) \in R^{(m\times n),K}$, which determines the operator.

Singular value decomposition for a matrix (2) is obvious, as it is the classical matrix: symmetric and positive semi-definite, on vector Euclidean R^K. It is defined by a collection of singularities

$$||v_i|| = 1, v_i \perp v_j, i \neq j;\ i, j = \overline{1, r};\ \lambda_1 > \lambda_2 > ... > \lambda_r > 0,$$

$$\wp_\alpha^* \wp_\alpha v_i = \lambda_i^2 v_i,\ i = \overline{1, r}.$$

The operator $\wp_\alpha^* \wp_\alpha$ by itself and is determined by the relation

$$\wp_\alpha^* \wp_\alpha = \sum_{i=1}^{r} \lambda_i^2 v_i v_i^T = \sum_{i=1}^{r} \lambda_i^2 v_i (v_i,\ \cdot\ ).$$

Each of the row – vectors $v_i^T, i = \overline{1, r}$ will be written by their components:

$$v_i^T = (v_{i1}, ..., v_{iK}),\ i = \overline{1, r},$$

i.e. $v_{ik},\ i = \overline{1, r},\ k = \overline{1, K}$ is the component with the number k of a vector v with a number I.

Theorem 5. *Matrices $U_i \in R^{m\times n} : U_i = \frac{1}{\lambda_i} \wp_\alpha v_i = \frac{1}{\lambda_i} \sum_{k=1}^{K} A_k v_{ik},\ i = \overline{1, r}$, defined by the singularities $(v_i, \lambda_i^2),\ i = \overline{1, r}$ of the operator $\wp_\alpha^* \wp_\alpha$ are elements of a complete collection of singularities $(U_i, \lambda_i^2),\ i = \overline{1, r}$ of the operator $\wp_\alpha^* : R^K \to R^{m\times n}$.*

Theorem 6. *(Singular Value Decomposition (SVD) for cortege operator). Singularity of two operators $\wp_\alpha^*\wp_\alpha, \wp_\alpha\wp_\alpha^*$, obviously determine the SVD for $\wp_\alpha$, $\wp_\alpha^*$:*

$$\wp_\alpha y = \sum_{i=1}^{r} \lambda_i U_i v_i^T y, y \in R^K,$$

$$\wp_\alpha^* X = \sum_{i=1}^{r} \lambda_i v_i (U_i, X)_{tr}, X \in R^{m\times n}.$$

Corollary 1. *A variant is a SVD for the operator $\wp_\alpha$ is represented by the next relation:*

$$\wp_\alpha = \sum_{k=1}^{r} \lambda_k U_k v_k^T = \sum_{k=1}^{r} \left(\wp_\alpha v_k\right) v_k^T.$$

4. Pseudo Inverse Technique for matrices Euclidean spaces

Basic operators of Pseudo Inverse (PdI-operators) theory for a cortege operators are namely pseudo inverse by itself for linear operator, orthogonal projectors on fundamental subspaces of linear operators and grouping operators which also often called by "weighted projection" operators.

Theorem 7. *The pseudo inverse operators for $\wp_\alpha$, $\wp_\alpha^*$ are determined, correspondingly, by the relations*

$$\wp_\alpha^+ X = \sum_{k=1}^{r} \lambda^{-1} v_k \left(U_k, X\right)_{tr} = \sum_{k=1}^{r} \lambda^{-2} v_k \left(\wp_\alpha v_k, X\right)_{tr}, \ \forall X \in R^{m\times n},$$

$$\left(\wp_\alpha^*\right)^+ y = \sum_{i=1}^{r} \lambda^{-1} U_i v_i^T y, \ \forall y \in R^K.$$

The basic orthogonal projectors PdI-theory are two pairs of orthogonal projectors. The first one is the pair of orthogonal projectors on the pair fundamental subspaces of $\wp_\alpha, \wp_\alpha^* : \Re(\wp_\alpha) = L_{\wp_\alpha}, \Re(\wp_\alpha^*) = L_{\wp_\alpha^*}$ – their ranges. These orthogonal projections will be designated in one of two equivalent ways:

$$P(\wp_\alpha^*) \equiv P_{L_{\wp_\alpha}} = P_{(A_1,\ldots,A_K)}, \ L_{L_{\wp_\alpha}} \subseteq R^{m\times n}, \ P(\wp_\alpha) \equiv P_{L_{\wp_\alpha^*}}, \ L_{\wp_\alpha^*} \subseteq R^K.$$

The second pair is a pair of orthogonal projectors onto the orthogonal complement $L_{\wp_\alpha}^\perp \subseteq R^{m\times n}$, $L_{\wp_\alpha^*}^\perp \subseteq R^K$ of the first pair of the subspaces. The complements, namely, are the Kernels of the correspondent operators. Each of these projectors will be denoted in one of two equivalent ways:

$$Z(\wp_\alpha) \equiv P_{L_{\wp_\alpha^*}^\perp}, \ Z(\wp_\alpha^*) \equiv P_{L_{\wp_\alpha}^\perp},$$

obviously:

$$Z(\wp_\alpha) \equiv E_K - P(\wp_\alpha),\ Z(\wp_\alpha^*) \equiv E_{m\times n} - P(\wp_\alpha^*). \tag{3}$$

In accordance with the general properties of PdI, the next properties are valid:

$$P(\wp_\alpha) = \wp_\alpha^+ \cdot \wp_\alpha,\ P(\wp_\alpha^*) = (\wp_\alpha^*)^+ \cdot \wp_\alpha^* = \wp_\alpha \cdot \wp_\alpha^+.$$

Correspondingly:

$$Z(\wp_\alpha) \equiv E_K - \wp_\alpha^+ \cdot \wp_\alpha, \quad Z(\wp_\alpha^*) \equiv E_{m\times n} - \wp_\alpha \cdot \wp_\alpha^+.$$

Grouping operators, denoted below as $R(\wp_\alpha),\ R(\wp_\alpha^*)$, are also "paired" operators, and are determined by the relations:

$$R(\wp_\alpha) = \wp_\alpha^+ \left(\wp_\alpha^+\right)^* = \wp_\alpha^+ \left(\wp_\alpha^*\right)^+, \quad R(\wp_\alpha^*) = (\wp_\alpha^*)^+ \left((\wp_\alpha^*)^+\right)^* = \left(\wp_\alpha^+\right)^* \wp_\alpha^+.$$

Theorem 8. *Grouping operators for the cortege operators $\wp_\alpha$, $\wp_\alpha^*$ can be represented by the next expression:*

$$R(\wp_\alpha^*)X = \sum_{k=1}^r \lambda_k^{-2} U_k (U_k, X)_{tr} = \sum_{k=1}^r \lambda_k^{-2} U_k tr U_k^T X = \sum_{k=1}^r \lambda_k^{-2} U_k tr X^T U_k,$$

and the quadratic form $(X, R(\wp_\alpha^*)X)_{tr}$ is determined by the relation:

$$(X, R(\wp_\alpha^*)X)_{tr} = \sum_{k=1}^r \lambda_k^{-2} (U_k, X)_{tr}^2,$$

where

$$\wp_\alpha^+ X = \sum_{k=1}^r \lambda^{-1} v_k \left(U_k, X\right)_{tr} = \sum_{k=1}^r \lambda^{-2} v_k \left(\wp_\alpha v_k, X\right)_{tr},$$

$$\left(\wp_\alpha^*\right)^+ y = \sum_{i=1}^r \lambda^{-1} U_i v_i^T y.$$

Theorem 9. *Quadratic form $(X, R(\wp_\alpha^*)X)_{tr}$ may be written as:*

$$(X, R(\wp_\alpha^*)X)_{tr} = \sum_{i=1}^r \lambda_i^{-4} v_i^T \begin{pmatrix} trA_1^T X trA_1^T X & trA_2^T X trA_2^T X & \cdots & trA_1^T X trA_K^T X \\ trA_2^T X trA_1^T X & trA_2^T X trA_2^T X & \cdots & trA_2^T X trA_K^T X \\ \cdots & \cdots & \cdots & \cdots \\ trA_K^T X trA_1^T X & trA_K^T X trA_1^T X & \cdots & trA_K^T X trA_1^T X \end{pmatrix},$$

$$v_i = \sum_{i=1}^r \lambda_i^{-4} \left\{ v_i^T \begin{pmatrix} trA_1^T X \\ \cdots \\ trA_K^T X \end{pmatrix} \right\}^2 = \sum_{i=1}^r \lambda_i^{-4} \left\{ v_i^T \wp_\alpha^* X \right\}^2.$$

Importance of grouping operators is determined by their properties, represented by the next two theorems.

Theorem 10. *For any $A_i,\ i=\overline{1,K}$ of $\alpha=(A_1\vdots\ldots\vdots A_K)\in R^{(m\times n),K}$ the next inequalities are fulfilled:*

$$(A_i, R(\wp_\alpha^*)A_i)_{tr}\le r,\ i=\overline{1,K},\ r=rank\wp_\alpha.$$

Theorem 11. *For any $A_i,\ i=\overline{1,K}$ of $\alpha=(A_1\vdots\ldots\vdots A_K)\in R^{(m\times n),K}$ the next inequalities are fulfilled:*

$$(A_i, R(\wp_\alpha^*)A_i)_{tr}\le r_{\min}\le r,\ i=\overline{1,K},\ r=rank\wp_\alpha,$$
$$r_{\min}=\min_{i=\overline{1,n}}(A_i, R(\wp_\alpha^*)A_i)_{tr}\le r_{\min}\le r,\ i=\overline{1,K},\ r=rank\wp_\alpha.$$

Note. Statement of theorem 11 is equivalent to that one ellipsoid

$$\frac{1}{r_{\min}}(X, R(\wp_\alpha^*))_{tr}\le 1 \tag{4}$$

is minimal to cover all matrices $A_i,\ i=\overline{1,K}$ of cortege $\alpha=(A_1\vdots\ldots\vdots A_K)\in R^{(m\times n),K}$.

Definition 1. Ellipsoid, defined by (4) we will call the minimum grouping ellipsoid for matrices collection $A_i,\ i=\overline{1,K}$.

5. Grouping operators and correspondence distances clasterization problems with feature matrix

The results, represented earlier one can apply to solve the grouping information problem in applied math with matrices 'representatives': matrices "feature vectors" or simply — "feature matrices". Indeed, in many important applied researches the objects under investigations are naturally represented by matrices. Spectrograms in speech recognition or digital images in image processing are appropriate examples of such situation. Important means for solving the clasterization problem is constructing and using of appropriate correspondence distance $\rho(X,Kl)$ from a cluster Kl, represented by learning sample of matrices $A_i,\ i=\overline{1,K}$. Such distance one can construct using characteristics of the minimal grouping ellipsoid from theorem 10, 11, built for cortege operator $\wp_\alpha$, generated by the $A_i,\ i=\overline{1,K}$ with $\alpha=(A_1\vdots\ldots\vdots A_n)$:

$$\rho^2(X,Kl)=\frac{1}{r_{\min}}(X, R(\wp_\alpha^*)X)_{tr}, r_{\min}=\min_{i=\overline{1,n}}(A_i, R(\wp_\alpha^*)A_i)_{tr}\le r.$$

Conclusion

Development of the technique for manipulating with the basic structures of Euclidean spaces within matrices spaces is represented. This technique include General SVD theorem and Moore-Penrose pseudo inverse technique for matrices spaces. Designing the technique demanded introduction matrices corteges and of special cortege operators associated with them.

References

1. Albert, A. 1972. *Regression and the Moore-Penrose pseudoinverse.* New York: Academic Press.
2. Ben-Israel, A. and Greville, T. N. E. 2003. *Generalized inverses.* New York: Springer.
3. Berry, M. W. 2004. *Survey of text mining.* New York: Springer.
4. Bublik, B. and Kirichenko, N. 1975. *Osnovyi teorii upravleniya.* Kyiv: Vischa shkola.
5. Donchenko, V. 2011. Evklidovyi prostranstva chislovyih vektorov i matrits: konstrutkivnyie metodyi opisaniya bazovyih struktur i ih ispolzovanie. *Information technologies & Knowledge*, 5 (3), pp. 203–216.
6. Donchenko, V., Kirichenko, M. and Krivonos, Y. 2007. Generalizing of neural nets: functional nets of special type. Institute of Information Theories and Applications FOI ITHEA.
7. Donchenko, V., Kirichenko, M. and Serbaev, D. 2004. "Recursive regression transformation and dynamical systems", paper presented at Computer Data analysis and Modeling: robustness and computer intensive methods, Minsk, September 6-10. Minsk: pp. 147–151.
8. Donchenko, V., Zinko, T. and Skotarenko, F. 2012. "Feature vectors in grouping information problem in applied mathematics: vectors and matrices", paper presented at Problems of Computer Intellectualization, Kyiv, Ukraine-Sofia Bulgaria: NASU, V. M. Glushkov Institute of Cybernetics, ITHEA, Kyiv: pp. 111–124.
9. Foster, J. and Helzl, W. 2004. *Applied evolutionary economics and complex systems.* Cheltenham: Elgar.
10. Friedman, M. and Kandel, A. 1999. *Introduction to pattern recognition.* London: Imperial College Press.
11. Haykin, S. S. 1999. *Neural networks.* Upper Saddle River, N.J.: Prentice Hall.
12. Ivahnenko, O. 1969. Samoobuchayuschiesya sistemyi raspoznavaniya i avtomaticheskogo upravleniya. Kyiv: Tehnika.
13. Kirichenko, M. 1997. Analytical Representation of Perturbation of Pseudoinverse Matrices. *Cybernetics and Systems Analysis*, 33 (2), pp. 98–107.
14. Kirichenko, M. and Donchenko, V. 2007. Pseudoinversion in clustering problems. *Cybernetics and Systems Analysis*, 4, pp. 73–92.
15. Kirichenko, M. and Donchenko, V. 2005. Zadacha terminal"noho sposterezhennya dynamichnoyi systemy: mnozhynnist rozv'yazkiv ta optymizaciya. *Journal of Computational & Applied Mathematics*, 5, pp. 63–78.
16. Kirichenko, M., Donchenko, V. and Serbaev, D. 2005. Nonlinear recursive nonlinear Transfomations: Dynamic systems and Optimizations. *Cybernetics and System Analysis*, 41 (3), pp. 364–373.

17. Kirichenko, M., Donchenko, V., Krivonos, Y., Krak, Y. and Kulyas, A. 2009. *Analiz ta syntez sytuacij v systemax pryjnyattya rishen*. Kyiv: Naukova dumka.
18. Kirichenko, N. 1997. Analiticheskoe predstavlenie vozmuscheniy psevdoobratnyih matrits. *Cybernetics and Systems Analysis*, 2.
19. Kirichenko, N. and Lepeha, N. 2002. Primenenie psevdoobratnyih i proektsionnyih matrits k issledovaniyu zadach upravleniya, nablyudeniya i identifikatsii. *Cybernetics and Systems Analysis*, 4, pp. 107–124.
20. Kirichenko, N. F., Krak, Y. V. and Polischuk, A. 2004. Psevdoobratnyie i proektsionnyie matritsyi v zadachah sinteza funktsionalnyih preobrazovateley. *Cybernetics and Systems Analysis*, 3, pp. 116–129.
21. Kirichenko, N., Donchenko, V. and Serbaev, D. 2005. Nelineynyie rekursivnyie regressionnyie preobrazovateli: dinamicheskie sistemyi i optimizatsiya. *Cybernetics and Systems Analysis*, 3, pp. 58–68.
22. Kirichenko, N., Krivonos, Y. and Lepeha, N. 2007. Sintez sistem neyrofunktsionalnyih preobrazovateley v reshenii zadach klassifikatsii. *Cybernetics and Systems Analysis*, 3, pp. 47–57.
23. Kohonen, T. 2001. Self-organizing maps. Berlin: Springer.
24. Nashed, M. Z. and Votruba, G. F. 1976. *A unified operator theory of generalized inverses.*
25. Vapnik, V. N. 1998. *Statistical learning theory*. New York: Wiley.
26. Wakefield, T. 2004. *Systems Analysis and Design*. Pearson Education UK.