SPREADING THE MOORE - PENROSE PSEUDO INVERSE ON MATRICES EUCLIDEAN SPACES: THEORY AND APPLICATIONS

Volodymyr Donchenko, Fedir Skotarenko

Abstract: *In the paper the development of operating technique for matrices Euclidean spaces is represented. Particularly, within such development transfer of linear operators' technique with preserving properties of closely correspondence to natural subspaces is represented. Also - spectral results, SVD – and Moore – Penrose Pseudo Inverse – technique, theory of orthogonal projectors and grouping operators. Besides, solution of the linear discrimination problem in Euclidean spaces of matrices is represented in the paper. Realization the program of empowering the operating technique in matrices Euclidean space made possible on the basis of putting in circulation of so-called "cortege operators" and, correspondingly - "cortege operations".*

Keywords: *Feature vectors, matrix corteges operators, Single Valued Decomposition for cortege linear operators, linear discrimination.*

ACM Classification Keywords: *G.2.m. Discrete mathematics: miscellaneous, G.2.1 Combinatorics; G.3 Probability and statistics, G.1.6. Numerical analysis, I.5.1.Pattern Recognition; H.1.m. Models and Principles: miscellaneous.*

Introduction

Grouping information problem (GIP), which is fundamental one in applications, appears in two main forms namely: the problem of recovering the function, represented by their observations, and the problem of clustering, classification and pattern recognition. Examples of approaches in the field are represented perfectly in [Kohonen, 2001], [Vapnik, 1998], [Haykin, 2001], [Friedman, Kandel, 2000], [Berry, 2004]. It is opportune to notice, that math modeling is the representation of an object structure by the means of mathematical structuring. A math structure after Georg Cantor is a set plus "ties" between its elements. Only four fundamental types of "ties" (with its combination as fifth one) exist: relations, operations, functions and collections of subsets. Thus, the mathematical description of the object (mathematical modeling) cannot be anything other than representing the object structure by the means of mathematical structuring. It refers fully to so call "complex system". A "complex system" should be understanding and, correspondingly, determined, as an objects with complex structure (complex "ties"). When reading attentively manuals by the theme (see, for example, [Yeates, Wakefield, 2004], [Forster, Hölzl, 2004]) one could find correspondent allusions. "Structure" understanding of the object is reasonable in determining of a "complex systems" instead of defining it as the "objects, consisting of numerous parts, functioning as an organic whole".

In the essence, math modeling is representing by math "parts plus ties" "parts plus ties" of the object in applied field.

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 by R^n .

Euclidean spaces, namely R^n , are the first among math structures rich on ties. Already in the very definition Euclidean spaces offer a range of structural links: from operations to scalar product, norms and limit transitions in various form. Besides, these ones possess highly developed technique of linear operators: with spectral theory, Singular Valued Decomposition (SVD) and Moore – Penrose technique in Euclidean spaces of real valued vectors (in R^n). Regarding uses of R^n we recall linear regression and classification or clasterization problem with necessity of designing an appropriate feature vectors. But there is an urgent need in the application to expand the range of "representatives" of real objects with preserving the wealth structural relationships inherent Euclidean spaces. Matrices Euclidean spaces are a natural extension of the class of Euclidean spaces. Speech recognition with the spectrograms as the representatives and the images in the problem of image processing and recognition are the natural object areas with “matrices representatives”. So, of utmost importance is developing mathematical modeling tools and, in particular, the problems of grouping information problem is the transfer on matrices Euclidean spaces the wealth of technical capabilities of R^n . As to technique designing for the spreading of Euclidean spaces as “environmental” math structure first steps have been made for example, by [Donchenko, 2011], [Donchenko, Zinko, Skotarenko, 2012].

Just the belonging to the base math structure (Euclidean space) determines advantages of the “vectors” against “corteges” as ordered finite collection of elements. It is noteworthy to say, that real-valued vectors as a variant of Euclidean space is not unique. A space of all matrixes of a fixed dimension is alternative example. As it was mentioned above, the choice of R^n as “environmental” space 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 [Nashed, 1978] (see, also, [Albert, 1972], [Ben-Israel, Greville, 2002]). Outstanding impacts and achievements in this area are due to N. F. Kirichenko (especially, [Kirichenko, 1997], [Kirichenko, 1997], see also [Kirichenko, Lepeha, 2002]). Greville’s formulas: forward and inverse - for pseudo inverse matrixes, 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 [Kirichenko, Lepeha, 2001], [Donchenko, Kirichenko, Serbaev, 2004], [Kirichenko, Crak, Polishuk, 2004], [Kirichenko, Donchenko, Serbaev,2005], [Kirichenko, Donchenko, 2005] [Donchenko, Kirichenko, Krivonos, 2007], [Kirichenko, Donchenko, 2007], [Kirichenko, Krivonos, Lepeha, 2007], [Kirichenko, Donchenko, Krivonos, Crak, Kulyas, 2009].

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.

In the paper the development of operating technique for matrices of Euclidean spaces is represented. Particularly, transfer on these spaces linear operators technique with preserving of close correspondence with “natural subspaces”; spectral results; SVD – and Moore – Penrose Pseudo Inverse – technique M-P PdI); the theory of orthogonal projection and grouping operators. Also, solution of the linear discrimination problem in Euclidean spaces of matrices is represented below in the paper. Realization the program of empowering the operating technique in matrices Euclidean space made possible on the basis of putting in circulation of so-called "cortege operators” and, correspondingly - "cortege operations”.

Matrixes spaces and cortege operators

Theorem 1. (Enhanced spectral theorem) 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_i^2), $(u_i,\lambda_i^2) i=\overline{1,r}$, $r = rank \wp_E$ 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 $\lambda_i^2, i=\overline{1,r} : \lambda_{i-1} \ge \lambda_i > 0$, $\overline{i=2,r}$ such that $\wp_E x = \sum_{i=1}^{r} \lambda_i u_i (v_i, x)_1$, $\wp_E^* y = \sum_{i=1}^{r} \lambda_i v_i (u_i, y)_2$.

Besides, the following relations take place:

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

Matrixes spaces and SVD – technique for cortege operator

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

$$(\alpha,\beta)_{cort} = \sum_{k=1}^{K} (A_k, B_k)_{tr} = \sum_{k=1}^{K} trA_k^T B_k, \ \alpha = (A_1 \vdots \ldots \vdots A_K), \beta = (B_1 \vdots \ldots \vdots B_K) \in R^{(m\times n),K}.$$

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

$$\wp_\alpha y = \sum_{k=1}^{K} y_k A_k, \alpha = (A_1 \vdots \ldots \vdots A_K) \in R^{(m\times n),K}, y = \begin{pmatrix} y_1 \\ \cdots \\ y_K \end{pmatrix} \in R^K \quad (1)$$

2. **Theorem 2.** Range $\Re(\wp_\alpha) = L_{\wp_\alpha}$, which is linear subspace of R^{m*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 $\wp_\alpha$: $\Re(\wp_\alpha) = L_{\wp_\alpha} = L(A_1, \ldots, A_K)$.

3. **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:

$$\wp_\alpha^* X = \begin{pmatrix} (A_1, X) \\ \cdots \\ (A_K, X) \end{pmatrix} = \begin{pmatrix} trX^T A_1 \\ \cdots \\ trX^T A_K \end{pmatrix} \quad (2)$$

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

$$\wp_\alpha^* \wp_\alpha = \begin{pmatrix} trA_1^T A_1, \ldots, trA_1^T A_K \\ \cdots \\ trA_K^T A_1, \ldots, trA_K^T A_K \end{pmatrix} \quad (3)$$

Remark. Matrix defined by (3) is the Gram' matrix for the matrixes - cortege components $A_k, k=\overline{1,K}$ of Euclidean space $R^{m\times n}$.

5. Singular value decomposition (SVD) for a matrix (3) is obvious. That matrix is a classical $K\times K$ matrix: symmetric and positively semi-definite. This SVD is defined by a collection of singularities $(v_i,\lambda_i^2), i,j=\overline{1,r}$:

$$\|v_i\|=1, v_i\perp v_j, i\neq j; i,j=\overline{1,r}; \lambda_1>\lambda_2>...>\lambda_r>0,\ \wp_\alpha^*\wp_\alpha v_i=\lambda_i^2 v_i, i=\overline{1,r}.$$

The operator $\wp_\alpha^*\wp_\alpha$ by itself is determined by the relation $\wp_\alpha^*\wp_\alpha=\sum\limits_{i=1}^{r}\lambda_i^2 v_i v_i^T=\sum\limits_{i=1}^{r}\lambda_i^2 v_i(v_i,\ \cdot\ )$.

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

$$v_i^T=(v_{i1},...,v_{iK}), i=\overline{1,r},\ r \text{ - is the rank of } A_k, k=\overline{1,K} \text{ in linear space is } R^{m\times n}.$$

6. **Theorem 5.** Matrices $U_i\in R^{m\times n}: U_i=\dfrac{1}{\lambda_i}\wp_\alpha v_i=\dfrac{1}{\lambda_i}\sum\limits_{k=1}^{K}A_k v_{ik}, i=\overline{1,r}$, 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. $\wp_\alpha^*: R^K\to R^{m\times n}$.

Proof. The result directly follows from the Theorem 1 and standard relationships between singularities of the $\wp_\alpha^*\wp_\alpha$, $\wp_\alpha\wp_\alpha^*$ operators.

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

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

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

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

Pseudo Inverse Technique for matrixes Euclidean spaces

Fundamental operator of Moore - Penrose Pseudo Inverse (M-P PdI) theory is PdI operator: for in the case under consideration. This operator is proposed to be determined by SVD-representation.

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

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

9. **Principal operators PdI theory for a cortege operators: basic OP-operators**.

Two pairs of operators are principal importance in classical M-P PdI theory namely, these are: a) two pares of orthogonal projector operators(OP-operators): on the ranges of operators $\wp_\alpha^*, \wp_\alpha$ correspondingly and on the kernels of these operators; b) grouping operators(G-operators) [Kirichenko, Lepeha, 2002].

As it was mentioned above OP – operators project on the subspaces $\Re(\wp_\alpha) = L_{\wp_\alpha}, \Re(\wp_\alpha^*) = L_{\wp_\alpha^*}$, Ker $\wp_\alpha = \Re^\perp(\wp_\alpha^*) = L^\perp_{\wp_\alpha^*}$, Ker$(\wp_\alpha^*) = \Re^\perp(\wp_\alpha)$ - are determined, in the essence, by the orthogonal projections on two ranges: $\wp_\alpha, \wp_\alpha^*$ correspondingly. These orthogonal projections will be designated in one of two equivalent ways:

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

Two OP- operators: on the kernels of $\wp_\alpha, \wp_\alpha^*$ – being orthogonal complements to the correspond ranges are the compliments to identity operators of ranges OP- operators: as $Ker\wp_\alpha^* = L^\perp_{\wp_\alpha} \subseteq R^{m\times n}, Ker\wp_\alpha = L^\perp_{\wp_\alpha^*} \subseteq R^K$.

These OP –operators on correspondent kernels we denote by $Z(\wp_\alpha), Z(\wp_\alpha^*)$ correspondingly:

$$Z(\wp_\alpha) \equiv P_{L^\perp_{\wp_\alpha^*}}, Z(\wp_\alpha^*) \equiv P_{L^\perp_{\wp_\alpha}},$$

Obviously:

$$Z(\wp_\alpha) \equiv E_K - P(\wp_\alpha), Z(\wp_\alpha^*) \equiv E_{m\times n} - P(\wp_\alpha^*) \tag{4}$$

$E_K, E_{m\times n}$ - identity operators in correspondent spaces.

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

$$P(\wp_\alpha) = \wp_\alpha^+ \cdot \wp_\alpha, P(\wp_\alpha^*) = \left(\wp_\alpha^*\right)^+ \cdot \wp_\alpha^* = \wp_\alpha \cdot \wp_\alpha^+ = \sum_{k=1}^{r} v_k v_k^T .$$

Correspondingly:

$$Z(\wp_\alpha) \equiv E_K - \wp_\alpha^+ \cdot \wp_\alpha, \quad Z(\wp_\alpha^*) \equiv E_{m\times n} - \wp_\alpha \cdot \wp_\alpha^+ = E_K - \sum_{k=1}^{r} v_k v_k^T .$$

10. **Basic operators PdI theory for a cortege operators: grouping operators.**

Grouping operators (G- operators) [Donchenko, Zinko, Skotarenko, 2012], designated below by $R(\wp_\alpha)$, $R(\wp_\alpha^*)$, are also "paired" operators, and are determined by the relations:

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

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

$$R(\wp_\alpha^*)X = \sum_{i=1}^{r} \lambda_i^{-2} U_i (U_i, X)_{tr} = \sum_{i=1}^{r} \lambda_i^{-2} U_i tr U_i^T X = \sum_{i=1}^{r} \lambda_i^{-2} U_i tr X^T U_i, i = \overline{1,r}$$

with

$$U_i = \frac{1}{\lambda_i} \sum_{k=1}^{K} v_{ik} A_k, i = \overline{1,r};$$

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

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

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

$$(X,R(\wp_{\alpha}^{*})X)_{tr}=\sum_{i=1}^{r}\lambda_i^{-4}v_i^T\begin{pmatrix} trA_1^TXtrA_1^TX & trA_2^TXtrA_2^TX & \cdots & trA_1^TXtrA_K^TX \\ trA_2^TXtrA_1^TX & trA_2^TXtrA_2^TX & \cdots & trA_2^TXtrA_K^TX \\ \cdots & \cdots & \cdots & \cdots \\ trA_K^TXtrA_1^TX & trA_K^TXtrA_1^TX & \cdots & trA_K^TXtrA_1^TX \end{pmatrix}v_i=$$

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

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

13. **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:

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

14. **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:

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

Comment to the theorems 10, 11. These theorems give the minimal grouping ellipsoids for the matrixes $A_i, i=\overline{1,K}$. In order to build it one only has to construct cortege operator $\wp_{\alpha}$ by the cortege $\alpha=(A_1\vdots\ldots\vdots A_K)\in R^{(m\times n),K}$.

Linear discrimination problem in matrix Euclidean space $R^{m\times n}$

Linear discrimination problem (LDP) is the problem of separating of two classes, represented by correspond learning sample by appropriate hyper plane, For Euclidean spaces R^m this problem was formulated and successfully solved on the base of PdI technique in [Kirichenko, Lepeha, 2002]. This problem is formulated and solved below on base of PdI technique developed for matrices on the base of cortege operators represented early in this article.

The wording of the problem.

Let $X_1,\ldots,X_K\in R^{m\times n}$ united collection of matrixes from learning sample, represented two classes:

$$X_j\in Kl_1, j\in J_1, X_j\in Kl_2, j\in J_2:\ J_1\cap J_2=\varnothing, J_1\cup J_2=\{1,2,\ldots,K\}.$$

It is necessary to find $\Delta>0$ and design linear functional $A:R^{m\times n}\to R^1$ in such a way that

$$(A,X_j)_{tr}>\Delta, j\in J_1, (A,X_j)_{tr}<-\Delta, j\in J_2$$

We will designate by $\Omega(\Delta)$ the domain of real-valued vector $y^T=(y_1,\ldots,y_K)$ from R^K with the components which satisfy to the next constraints: $y_j>\Delta, j\in J_1, y_j<-\Delta, j\in J_2$

LDP solution.

LGP solution for matrices spaces is because the vector of "discriminating" values $\left((A,X_1)_{tr},...,(A,X_K)_{tr}\right)^T$ of discriminating linear forms $(A,X)_{tr}$ determines the value of conjugate operator to cortege operator $\wp_\chi : \chi = (X_1 \vdots,...,\vdots X_K)$ on argument A.

The next theorem then is valid.

Theorem 12. LDP is equivalent of linear equation problem $\wp_\chi^* X = y$ for cortege operator $\wp_\chi, \chi = (X_1 \vdots,...,\vdots X_K)$ and $(A,X_j) > \Delta, j \in J_1, (A,X_j) < -\Delta, j \in J_2$.

Proof. Indeed, fulfilling of (1) means that vector

$$y : \left((A,X_1),...,(A,X_K)\right) \equiv y^T \tag{5}$$

belongs to $\Omega(\Delta)$.

Theorem 13. Allows to conclude, that (5) is equivalent to solvability the equation $\wp_\chi^* A = y, \chi = (X_1 \vdots,...,\vdots X_K), y \in \Omega(\Delta)$. And, thus, the proof of the theorem is finished.

Theorem 14. LDP is solvable if there exists $y_* \in \Omega(\Delta) \subseteq R^K$ and correspond solution is determined by the equality

$$A = \wp_\chi^{*+} y_* \tag{6}$$

Corollary 1. LDP is solvable if there exists $y_* \in \Omega(\Delta) \subseteq R^K$ for which the next condition is fulfilled

$$y_*^T Z(\wp_\alpha) y_* = y_*^T (E_K - \sum_{k=1}^{r} v_k v_k^T) y_* = 0$$

and correspond solution is still determined by (6).

Theorem 15. LDP is equivalent to quadratic optimization problem for quadratic form

$$y^T Z(\wp_\alpha) y = y^T (E_K - \sum_{k=1}^{r} v_k v_k^T) y$$

In domain $\Omega(\Delta) \subseteq R^K$. If the solution y_* of the optimization problem in the domain gives minimum, that equal zero, then matrix A - LDP solution, is determined by equality

$$A = \wp_\chi^{*+} y_* .$$

Linear Discrimination Problem: algorithm

When saying about the algorithm of matrix LDP problem with united for two classes collection of matrices $X_1,...,X_K \in R^{m \times n}$ then it starts with the first step".

1-st step: calculation of Gram' matrix for collection of matrices $X_1,...,X_K \in R^{m \times n}$:

$$F = \begin{pmatrix} trA_1^T A_1,...,trA_1^T A_K \\ \cdots \\ trA_K^T A_1,...,trA_K^T A_K \end{pmatrix} = \begin{pmatrix} (A_1,A_1)_{tr},...,(A_1,A_K)_{tr} \\ \cdots \\ (A_K,A_1)_{tr},...,(A_K,A_K)_{tr} \end{pmatrix}.$$

And then:

2-nd step: computing the singularities $(v_k, \lambda_k^2), \lambda_k^2 > 0,\ \ k = \overline{1,r},\ \ \ r = rankF$;

3-d step: calculating the matrixes $E_K - \sum_{k=1}^{r} v_k v_k^T$ of quadratic form;

4-th step: calculating the minimum of quadratic form $y^T \left(E_K - \sum_{k=1}^{r} v_k v_k^T \right) y$ in domain $\Omega(\Delta) \subseteq R^K$ (numerical methods) and correspondent argument y_*;

4-th step: verification the conditions of zero value of minimum: if $y_*^T \left(E_K - \sum_{k=1}^{r} v_k v_k^T \right) y_* = 0$:

5-th step:

- If condition is fulfilled: $y_*^T \left(E_K - \sum_{k=1}^{r} v_k v_k^T \right) y_* = 0$, then computing linear form A of LDP-solution accordingly to relation $A = \wp_\chi^{*+} y_*$;
- If condition is not fulfilled, then LDP is unsolvable.

Matrix Linear Discrimination Problem in clasterization

Theorem 11 can be applied to problems of matrix clustering just in the same way as in [Donchenko, Krak, Krivonos, 2007] it has been done for R^m.

Conclusion

Conception of enriching the standard considering the "representatives" in Applied Math to be the feature vectors: elements from Euclidean space, - has been further developed in the paper (see, also, [Donchenko, Zinko, Skotarenko, 2012]). Using matrixes as the "representatives" of the real objects is main idea of the conception. This mean, that matrix instead vector represents all principal features of the objects in applied fields. Support of this concept requires the development of technologies handling with matrixes similar techniques operating with vectors from Euclidean spaces. SVD-technique as well as PIMP - technique are the priority among them. The results of such type are represented in the paper. These results demanded a generalization of matrix algebra and transforming it in algebra of matrix and vector cortege as well as definition and using the linear cortege operator. Correspond results are represented in the paper of the authors [Donchenko, Zinko, Skotarenko, 2012]. Using that handling technique for matrix features ("matrix feature vectors") make it possible to put and fully solute the Linear Discrimination problem for two collection of matrixes. Corresponded solution uses standard SVD and PIMP for Gramian matrix of united collections and solution of quadratic optimization in a domain of appropriate. Thus, the development of matrix technique manages to reduce to existing technique for real valued vectors. Solution of Linear Discrimination Problem for matrixes is similar to corresponded result for real-valued vectors in [Kirichenko, Krivonos, Lepeha, 2007] or [Donchenko, Krak, Krivonos, 2012]. The two obvious application areas are worth mentioning within the context of the application of these results. These are: speech recognition and image processing. Matrixes naturally represent the objects under consideration, namely, spectrograms and digital images.

Bibliography

[Albert, 1972] Albert A. E. Regression and the Moore-Penrose pseudoinverse.-Academic Press.- 1972.- 180 стор.

[Ben-Israel, Greville,2002] Adi Ben-Israel, Thomas N. E. Greville Generalized Inverses. Theory and Applications.- Second Edition.- Springer-Verlag New York, Inc. -2003.-420 p

[Berry, 2004] Michael W. Berry Editor Survey of Text Mining: Clustering, Classification, and Retrieval Springer-Verlag New York, Inc. 2004. 244 p.

[Donchenko, Kirichenko, Krivonos, 2007] Donchenko V., Kirichenko M., Krivonos Yu. Generalizing of neural nets: Functional nets of special type// International Journal Information Theories and Applications. – N 3. –V.14. – 2007. – P. 259-266.

[Donchenko, Kirichenko, Serbaev, 2004] Donchenko V.S. Kirichenko M.F., Serbaev D.P. Recursive regression transformation and dynamical systems// Proceedings: of the Seventh International Conference “Computer Data analysis and Modeling: robustness and computer intensive methods”. – V.1. – September 6-10, 2004. – Minsk. – P. 147-151

[Donchenko, Zinko, Skotarenko, 2012] Donchenko V, Zinko T, Skotarenko F. “Feature Vectors” in Grouping Information Problem in Applied Mathematics: Vectors and Matrixes. - Vitalii Velichko, Alexey Voloshin, Krassimir Markov(ed.)- Problems of Computer Intellectualization.- Institute of Cybernetics NASU, ITHEA.-Kyiv, Ukraine -Sofia, Bulgaria.- 2012.- P.111-124.

[Forster, Hölzl, 2004] John Forster and Verner Hölzl Applied Evolutionary economics and complex systems.-2004-- 293 p.

[Friedman, Kandel, 2000] Menahem Friedman Abraham Kandel Introduction to pattern recognition statistical, structural, neural and Fuzzy logic approaches world scientific publishing co. Pte. Ltd.-reprinted 2000.-329 p.

[Haykin, 2001] Haykin S. Neural Networks, A Comprehensive Foundation. Prentice Hall, Upper Saddle River, New Jersey 07458. – 1999.– 842 p

[Ivachnenko, 1995] Ivachnenko A.G and Muller J. A. Present state and new problems of further GMDH development.-SAMS, 1995.- pp.3-16.

[Kirichenko, 1997] N. F. Kirichenko. Analytical Representation of Perturbation of Pseudoinverse Matrices. Cybernetics and Systems Analysis. Vol.33, Number 2, P. 230-239, March-April 1997.

[Kirichenko, Donchenko, Serbaev, 2005] Kirichenko N. F., Donchenko V.S., Serbaev D.P. Nonlinear recursive nonlinear Transfomations: Dynamic sestems and Optimizations.//Cybernetics and Systen Ansalysis.– V.41, №3.– May 2005.– p.364-373.

[Kohonen, 2001] Kohonen T. Self-Organizing Maps -3-d ed. - Tokyo: Springer, 2001, 501 p.

[Nashed, 1976] Nashed M. Zuhair, Votruba G.F., editors. A Unified Operator Theory of Generalized Inverse.//Proceedings of an Advanced Seminar Sponsored by the Mathematical Research Center, The University of Wisconsin, Madison, October 8-10, 1973. – New York, Academic Press, 1976.–1200 p.

[Vapnik, 1998] Vapnik, V. N. Statistical Learning Theory. New York: Wiley, 1998

[Yeates, Wakefield, 2004] Donald Yeates, Tony Wakefield Systems Analysis And Design. Pearson Education. 2004. - 499 p.

Authors' Information

***Volodymyr Donchenko** – Professor, Taras Shevchenko National University of Kyiv, Volodymyrs’ka street, Kyiv, 03680, Ukraine; e-mail: voldon@bigmir.net.*

***Fedir Skotarenko** - Post - graduate student, Taras Shevchenko National University of Kyiv.*