УДК 004.93

V.S. Donchenko[1], Ph.D, Associate Professor,
F.M. Skotarenko[2], postgraduate student

Development of pseudo inverse technique: cortege operators and applications in grouping information problem

This paper focuses on model resolving problem of patterns recognition is one the manifestation of grouping information problem. To solve this problem, Euclidean space is considered as the one that has the most extensive arsenal of describing relations between elements. R^n. *This technique includes, particularly, spectrum of linear operator (SVD), Moore-Penrose inversion, orthogonal projectors operators, Grouping operators and so on*

Key Words: Feature vectors, matrix corteges operators, Single Valued Decomposition for cortege linear operators, linear discrimination

[1-2] *Taras Shevchenko National University of Kyiv, 03680, Kyiv, Glushkova av., 4d,*
e-mail: voldon@bigmir.net
e-mail: fred.unicyb@gmail.com

Донченко В.С.[1], д.ф.м.-н., профессор,
Скотаренко Ф.М.[2], аспірант

Розвиток апарату псевдо обернення: кортежні оператори і застосування в проблемі групування інформації

У даній роботі наведено модель вирішення проблеми відновлення функції, яка є одною із проблем групування інформації. Для розв’язання даної проблеми розглядається Евклідів простір, як той, що має найбільш багатий арсенал засобів опису зв’язків між елементами. Зокрема, спектр лінійного оператора (СВД), інверсія за Муром-Пенроузом, ортогональні проектори, оператори групування та інші

Ключові слова: вектор функція, кортежні оператори, сингулярний розклад для кортежу лінійних операторів, лінійна дискримінація

[1-2] Київський національний університет імені Тараса Шевченка, 03680, м. Київ, пр.-т Глушкова 4д,
e-mail: voldon@bigmir.net
e-mail: fred.unicyb@gmail.com

Paper submitted by Ph. D., Prof. Garashchenko F.G.

1 Introduction

Grouping information problem (GIP) is fundamental problem in applied investigations. There are two main form of it, 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 is 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 Cantoris 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"). Namely, when reading attentively manuals by the theme (see, for example, [Yeates, Wakefield, 2004], [Forster , Hölzl, 2004]) one could find correspondent allusions. "Structure" understanding is reasonable determining of a "complex systems" instead of defining them 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.

As to technique designing for the Euclidean space $R^{m\times n}$ as "environmental" math structure first steps have been made for example, by [Donchenko, 2011], [Donchenko, Zinko, Skotarenko, 2012]. Speech recognition with the spectrograms as the representative and the images in the problem of image processing and recognition are the natural application areas for the correspond technique.

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.

Just the belonging to the base math structure (Euclidean space) determines advantages of the "vectors" against "corteges". It is noteworthy to say, that this variant of Euclidean space is not unique: the space $R^{m\times n}$ of all matrixes of a fixed dimension $m\times n$ may represent alternative example. The choice of the R^n space as "environmental" 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 [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 technique designing for the Euclidean space $R^{m\times n}$ as "environmental" one see, for example [Donchenko, 2011]. 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 families past or range of convenient 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 in passing through selection features step. For correspond examples see, [*Ivachnenko*,1969] with Ivachnenko's GMDH (Group Method Data Handling), [*Vapnik*, 1998] with Vapnik's Support Vector Machine. Further development of the recurrent technique one may find in *Donchenko* , *Kirichenko,Serbaev*, 2004], [*Kirichenko, Crak, Polischuk*,2004] [*Kirichenko, Donchenko, Serbaev*,2005], [*Kirichenko, Donchenko*,2005] [*Donchenko, Kirichenko* , *Krivonos*, 2007], [*Kirichenko, Donchenko*,2007] ,[*Kirichenko, Krivonos, Lepeha* 2007]. 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 thorough 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: [*Donchenko* , *Kirichenko,Serbaev*, 2004] [*Kirichenko, Crak, Polischuk*, 2004], *[Kirichenko, Donchenko, Serbaev*, 2005]).

1.1. Matrixes 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 $(v_i,\lambda_i^2),\ (u_i,\lambda_i^2) i=\overline{1,r},\ r = rank\wp_E$

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

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

Besides, the following relations take place: $u_i = \lambda_i^{-1}\wp v_i, i=\overline{1,r}$,

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

1.2. SVD – technique for matrixes spaces

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:

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

4. 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:

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

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.

5. 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 $(v_i, \lambda_i^2), i, j = \overline{1,r}$:

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

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

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

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*.

6. 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. $\wp_\alpha^* : R^K \to R^{m\times n}$

Proof. This follows from Theorem 1, and the standard relations 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 $\wp_\alpha$, $\wp_\alpha^*$:

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

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

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

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

1.3. Pseudo Inverse Technique for matrixes Euclidean spaces

Basic operators PdI theory for a cortege operators: pseudo inverse by SVD-representation.

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

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

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

2. Basic operators PdI theory for a cortege operators:: basic orthogonal projectors.

The basic orthogonal projectors PdI-theory are two pairs of orthogonal projectors. The first one is the pair of orthogonal projectors on the pair principal 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:

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

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

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

Correspondingly:

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

3. Basic operators PdI theory for a cortege operators:: basic orthogonal projectors. Grouping operators, denoted below as $R(\wp_\alpha)$, $R(\wp_\alpha^*)$, are also "paired" operators, and are determined by the relations:

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

4. 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

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

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

5. 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.

6. 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:

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

7. 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:

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

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

.

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 $\alpha=(A_1\vdots\ldots\vdots A_K)\in R^{(m\times n),K}$

List of used sources

1. *Albert A.E.* Regression and the Moore-Penrose pseudoinverse.-Academic Press.- 1972. - 180 p.
2. *Adi Ben-Israel, Thomas N.E.* Greville Generalized Inverses. Theory and Applications.- Second Edition.- Springer-Verlag New York, Inc. -2003. - 420 p.
3. *Michael W. Berry* Editor Survey of Text Mining: Clustering, Classification, and Retrieval Springer - Verlag New York, Inc. - 2004. -244 p.
4. *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. – pp. 147-151
5. *Donchenko V. Kirichenko M., Krivonos Yu.* Generalizing of neural nets: Functional nets of special type// International Journal Information Models and Analysis. – N 3. – V.14. – 2007. Kyiv, Ukraine -Sofia, Bulgaria. – pp. 259-266.
6. *Donchenko V. Krak Ju.., Krivonos Yu.* Recurrent procedure in solving the Grouping Information Problem in Applied Mathematics//ITHEA: International Journal Information Theories and Applications. – V.1.– N 1. Kyiv, Ukraine -Sofia, Bulgaria. – 2012. – pp. 62 -77.
7. *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.- pp. 111-124.
8. *Ivakhnenko A.G* and *Muller J.- A.* Present state and new problems of further GMDH development.-SAMS,1995. - pp. 3-16.
9. *Menahem Friedman Abraham Kandel* Introduction to pattern recognition statistical, structural, neural and Fuzzy logic approaches.- World scientific publishing co. Pte. Ltd. - reprinted - London,2000. -329 p.
10. *Forster J. , Hölzl V.* Applied Evolutionary economics and complex systems.-2004. Edward Elgar, UK - 293 p.

Received 27.11.2013