Krassimir Markov, Vitalii Velychko, Oleksy Voloshin
(editors)

Information Models
of
Knowledge

I T H E A®
K I E V – S O F I A
2010

Krassimir Markov, Vitalii Velychko, Oleksy Voloshin (ed.)
Information Models of Knowledge
ITHEA®
Kiev, Ukraine – Sofia, Bulgaria, 2010
ISBN 978-954-16-0048-1

First edition

Recommended for publication by The Scientific Concil of the Institute of Information Theories and Applications FOI ITHEA
ITHEA IBS ISC: 19.

This book maintains articles on actual problems of research and application of information technologies, especially the new approaches, models, algorithms and methods fot information modeling of knowledge in: Intelligence metasynthesis and knowledge processing in intelligent systems; Formalisms and methods of knowledge representation; Connectionism and neural nets; System analysis and sintesis; Modelling of the complex artificial systems; Image Processing and Computer Vision; Computer virtual reality; Virtual laboratories for computer-aided design; Decision support systems; Information models of knowledge of and for education; Open social info-educational platforms; Web-based educational information systems; Semantic Web Technologies; Mathematical foundations for information modeling of knowledge; Discrete mathematics; Mathematical methods for research of complex systems.

It is represented that book articles will be interesting for experts in the field of information technologies as well as for practical users.

General Sponsor: Consortium FOI Bulgaria (www.foibg.com).

Printed in Ukraine

ISBN 978-954-16-0048-1

C\o Jusautor, Sofia, 2010

Mathematical Foundations

ЕВКЛИДОВЫ ПРОСТРАНСТВА: КОНСТРУКТИВНЫЕ МЕТОДЫ ОПИСАНИЯ БАЗОВЫХ СТРУКТУР И ИХ ИСПОЛЬЗОВАНИЕ

Владимир Донченко

Аннотация: *Рассмотрены базовые структуры стандарных типов евклидовых пространств: числовых векторов и матриц фиксированных размерностей. Под базовыми имеются в виду линейные и нелинейные объекты, принципиальные с точки зрения их использвания в прикладных задачах. Развиты конструктивные методы построения базовых структур, их описания, взаимного перехода и использования на основе систематического развития и применения аппарата сингулярного представления, а также псевдообращения по Муру - Пенроузу. Теория сингулярного разложения и псевдообращения по Муру-Пенроузу распространена на матрицы, как линейные отображения между евклидовыми пространствами матриц. Приведены и доказаны соответствующие теоремы. Развита теория группирующих операторов, важная в прикладном отношении как инструмент выявления и использования групповых свойств объектов в евклидовых пространствах. Кроме того, группирующие операторы позволяют, в частности, прояснить алгебраическую суть расстояния Махаланобиса.*

Ключевые слова*: Псевдообращение по Муру - Пенроузу, сингулярное представление матрицы как матричного отображения, псевдообращение матричного отображения, эллипсы группировки, расстояния соответствия в кластеризации, расстояние Махаланобиса.*

ACM Classification Keywords*: G.3 Probability and statistics, G.1.6. Numerical analysis: Optimization; G.2.m. Discrete mathematics: miscellaneous.*

Вступление

В работе предложена и обоснована концепция базовых структур евклидового пространства, к которым предлагается отнести структуры двух основных видов: линейные и нелинейные. И те и другие структуры евклидова пространства проявляются либо во множественной форме: подпространства и гиперплоскости минимальные эллипсы группировки, либо - в единичной: линейные операторы, квадратичные формы. Отметим, что к базовым нелинейным структурам отнесены, квадратичные формы (в работе – неотрицательно определённые) и «характерные» множества на их основе: эллипсы группирвки, а также минимальные эллипсы группировки: поверхности уровня отвечающие единичному значению соответствующих квадратичных форм. Примечательной особенностью эллипсов группировки является то, что они содержат все векторы набора векторов, и что такие эллипсы можно выбрать «минимальными» для них. Описаны конструктивные методы генерации базовых структур в свзи с порождающей их совокупностью векторов, а также формулы взаимного перехода от одних типов структур к другим: от линейных подпространств и гиперплоскостей к матрицам и наоборот, а также от набора векторов к матрицам квадратичных форм и эллипсам группировки. В числе других рассмотрены конструктивные способы порождения подпространств и гиперплоскостей, а также ортогональных проекторов, связанных с указанными объектами. В том же русле конструктивности лежит рассмотрение концепции группирующих операторов и минимальных эллипсов группировки. Концепция группирующих операторов позволяет использовать в прикладных исследованиях связь того или иного набора векторов с важным видом нелинейных структур евклидовых пространств: минимальными эллипсоидами группировки. И в линейном, и в нейлинейном случае упомянутая конструктивность обеспечивается применением псевдообращения по Муру – Пенроузу (ПдО), а также новыми результатами в этой области, берущими своё начало и опирающимися на фундаментальную работу Н.Ф Кириченко [Кириченко, 1997]. Для расширения возможности аппарата, предложено развитие теории псевдообращения на матричные отображения между евлидовыми пространствами матриц фиксированных размерностей со следовыми скалярными произведениями. Кроме того, на случай отображений между евлидовыми матричными пространствами перенесена теорема о сингулярном разложении матрицы оператора, доказана теорема свёртки о соотношении между сингулярными разложениями одной и той же матрицы в векторном и матричном случаях.

Важность и эффективность применения приведённых результатов определяется возможностью их использования для решения широкого класса прикладных задач, включая линейную регрессию, в том числе векторную, теорию оптимального управления, кластеризацию, прогноз и функциональные сети, являющиеся обобщение искусственных нейронных сетей.

Как отмечалось в работе [Донченко,2009], «структура объекта» ассоциируется с реализацией представления об объекте, как чём-то едином, составленном из взаимодействующих, связанных между собою частей. Математическое описание, модель объекта, реализует описание «структуры» объекта средствами математического описания «связей». Это означает, что структура объекта должна быть передана, отражена в математической модели, средствами математического структурирования. К последним относятся четыре базовых структуры, а также их комбинации. К упомянутым четырём базовым структурам математики отностятся: отношения, функции, операции и наборы подмножеств. К комбинациям относятся, например, структуры линейного и евклидового пространств.

Структура линейного пространства, занимает ведущее место среди важнейших математических структур. В свою очередь, к базовым структурам таких пространств можно отнести линейные и нелинейные объекты, являющиеся принципиальными с точки зрения ипользования. Это, с одной стороны объекты на основе матриц: линейные операторы и квадратичные формы, с другой - структуры, которые можно назвать множественными. К ним можно отнести подпространства и гиперплоскости, а также эллипсоиды, порождаемые неотрицательно определёнными квадратичными формами.

Важную роль в прикладных исследованиях играют конструктивные методы описания соответствующих объектов. Принципиальным в этом отношении применительно к базовым линейным структурам является ипользование ортогональных проекторов. Для нелинейных - использование группирующих операторов. И в том и другом случае аппарат псевдообращения позволяет конструктивно строить как те, так и другие. Именно применение этого аппарата является основой для решения основых задач, связанных с построением, описанием и использованием базовых структур линейных, точнее, - евклидовых пространств.

Отметим, что ортогональные проекторы играют важную роль в исчерпывающем исследовании систем линейных алгебраических уравнений (СЛАУ). Принципиально важны они также и в постановке и решении важных оптимизационных задач с квадратичными функционалами качества в евклидовых пространствах. Важная роль ортогональных проекторов в решении прикладных задач в полной мере сохраняется и в связи с задачей построения наилучших квадратичных приближений правой части СЛАУ значениями левой, когда СЛАУ не имеет точных решений. Такие наилучшие приближения называют также псевдорешениями. Конструктивное описание ортогональных проекторов в связи с естественными подпространствами линейного оператора прямо определяется псевдообращением по Муру – Пенроузу [Moore,1920],[Penrose, 1955] (см. также, например,[Алберт, 1977]).

Следует отметить, также, что важную роль в применении аппарата псевдообращения играет сингулярное представление (его называют также SVD - представлением) произвольной $m \times n$ матрицы в специфической форме записи в виде взвешенной суммы тензорных произведений специального набора пар векторов с прозрачным геометрическим содержанием. При этом матрица рассматривается как матрица линейного оператора между двумя пространствами числовых векторов R^n и R^m. Это определяет специфику сингулярного разложения и инерпретацию его составляющих.

Заметим, однако, что одна и та же $m \times n$ матрица A задаёт как линейный оператор $A: R^n \to R^m$ между евклидовыми пространствами числовых векторов, так и линейный оператор $A: R^{n \times p} \to R^{m \times p}$ для произвольного натурального p. Переход к евклидовым пространствам $R^{n \times p}, R^{m \times p}$ матриц со следовыми скалярными произведениями требует учёта специфики этих евклидовых пространств для передачи принципиальных особенностей SVD-представления. В работе приведено SVD-представление для матриц, рассматриваемых как линейные операторы между пространствами матриц, доказана теорема свёртки, связывающие два варианта SVD-представления матрицы: векторного и матричного.

Важным в описании нелинейных структур являются минимальные эллипсы группировки и, как и в линейном случае, конструктивные способы их описания. Говоря о минимальных эллипсах группировки, будем иметь в виду «минимальный эллипс», включающий заданный набор векторов. Как оказывается, матрицей квадратичной формы для такого эллипса является группирущий оператор, конструктивно описываемый средствами псевдообращения. Квадратичная форма, отвечающая группирующему оператору, в искажающем суть такого оператора виде, фигурирует в «статистическом» варианте расстояния Махаланобиса [Mahalanobis,1936] (см. также, например, [McLachlan, Geoffry, 1992]), когда вместо ковариационной матрицы многомерного нормального распределения используется стандарная её оценка на основе выборки векторов $a(1),...,a(n)$.

В заключение отметим, что основные идеи, дух и результаты предлагаемой работы восходят к работам и используют результаты развитой в них теории псевдообращения нашего безвременно ушедшего коллеги, друга и учителя, профессора Н.Ф. Кириченко.

Базовые линейные структуры и связи между ними

В дальнейшем, говоря о евклидовом пространстве R^p, будем иметь в виду множество конечных числовых последовательностей одной и той же длины p, записанных в столбик с покоординатными операциями сложения и умножения на скаляр и суммой покоординатных произведений в качестве скалярного произведения. Именно такой вариант евклидового пространства будем стандартным образом обозначать через R^p, а его элементы – через $a, a^T = (a_1,...,a_p)$. Стандартные ортонормированные базисы, составленные из векторов с единственной единичной компонентой (остальные – нули) на месте с соответствующим номером будут обозначаться для R^m и R^n соответственно через $e(j) \in R^m, j = \overline{1,m}, e_{(i)} \in R^n, i = \overline{i,n}$, а для общего случая - $e_k \in R^p, k = \overline{1,p}$. Оператор A из R^n, в R^m: $A: R^n \to R^m$, задаваемый в ортонормированных базисах $e(j) \in R^m, j = \overline{1,m}, e_{(i)} \in R^n, i = \overline{i,n}$, будем отождествлять с $m \times n$ - матрицей $A = (a_{ij})$ этого оператора в этих базисах. Для матрицы $A = (a_{ij})$ будем использовать также блочное представление по столбцам (столбцовое) и строкам (строчное):

$$A = \begin{pmatrix} a_{(1)}^T \\ \cdots \\ a_{(m)}^T \end{pmatrix} = (a(1) \vdots ... \vdots a(n)), a_{(i)} \in R^n, i = \overline{1,m}, a(j) \in R^m, j = \overline{1,n}.$$

Пространство всех $p \times q$ матриц с покоординатным умножением и сложением, а также «следовым» скалярным произведением будем обозначать $R^{p \times q}$. «Следовое» скалярное произведение $(\cdot,\cdot)_{tr}$ двух матриц опеделяется соотношением

$$(C\cdot,D\cdot)_{tr} = trC^T D = \sum_{i=1}^{q}\sum_{k=1}^{p} c_{ki}d_{ki}, C = (c_{ij}), D = (d_{ij}) \in R^{p\times q}.$$

Линейное подпространство, порождённое системой векторов, $c_k \in R^p, k = \overline{1,K}$, будет обозначаться через $L(c_k, k = \overline{1,K}) \equiv L(c_1..., c_K)$, а линейное подпространство значений линейного оператора $A: R^n \to R^m$ – через L_A, соответственно, для A^T - через L_{A^T}.

Первым из набора свойств, описывающих связи между базовыми структурами, является следующее утверждение.

1. Наборы векторов, подпространства, матрица, составленная из набора векторов. Справедливо следующее утверждение.

$$L_A = L(a(1),...,a(n)).$$

Таким образом, линейное подпространство, порождённое набором векторов, совпадает с подпространством значений матрицы, составленной из векторов набора, как из столбцов.

2 . «*Векторы из матрицы*». Для элементов столбцового и строчного представления матрицы $A \in R^{m\times n}$ справедливы соотношения

$$a(j) = Ae_{(j)}, j = \overline{1,n}, a_{(i)}^T = e^T(i)A, i = \overline{1,m}.$$

3. *Произведение матриц как сумма тензорных произведений их составляющих*. Для произведения произвольных матриц B,C со столбцовым и строчным представлением

$$B = (b(1)\vdots...\vdots b(r)), b(j) \in R^m, j = \overline{1,r}, C = \begin{pmatrix} c_{(1)}^T \\ \cdots \\ c_{(r)}^T \end{pmatrix}, c_{(r)}, c(i) \in R^n, i = \overline{1,r}$$

соответственно, а также диагональной матрицы $\Lambda = diag(\lambda_1,...,\lambda_r)$ справедливо соотношение

$$B\Lambda C = \sum_{i=1}^{r} \lambda_i b(i) c_{(i)}^T.$$

Важной составляющей аппарата конструктивного описания и использования линейных структур является понятие ортогонального проектора (в дальнейшем ОП), определение которого полностью согласуется со стандартным геометрическим представлением об ортогональном проектировании. Общей, основой эффективного использования ортогональных проекторов, как и возможности их конструктивного построения в связи с линейными подпространствами, является наличие двух эквивалентных определений таких проекторов, а также возможность их описание через псевдообращение.

4.*«Геометрическое» определение ОП*. Для разложения $R^p = L + L^{\perp}$ в прямую сумму ортогональных подпространств ортогональным проектором P_L на линейное подпространство $L \subseteq R^p$ называется оператор, определяемый соотношением $P_L x = P_L(x_L + x_{L^\perp}) = x_L$, где

$$x = x_L + x_{L^\perp}, x_L \in L, x_{L^\perp} \in L^\perp \quad (1)$$

– однозначное разложение произвольного вектора $x \in R^p$ по двум составляющим ортогональной суммы. Очевидным образом оператор ортогонального проектирования является линейным оператором.

5. *Дополнительные друг к другу ОП*. Разложение (1) произвольного вектора $x \in R^p$ в силу симметричности относительно ортогональных слагаемых определяет одновременно два ортогональных проектора: $P_L, P_{L^\perp}$ с очевидным соотношением

$$P_L + P_{L^\perp} = E_p$$

где E_p – единичная матрица соответствующей размерности.

6. *Дополнительные друг к другу ОП и соответствующие им подпространства*. Для ортогонального проектора P_L на подпространство L оператор $Z_L \equiv E_p - P_L$ является ортогональным проектором на ортогональное дополнение $L^\perp : Z_L \equiv E_p - P_L = P_{L^\perp}$.

7. *Абстрактное определение ОП*. Для того, чтобы линейный оператор $P : R^p \to R^p$, был оператором ортогонального проектирования необходимо и достаточно, чтобы он был идемпотентным и симметричным оператором. Линейное подпространство L_P, на которое совершается ортогональное проектирование в соответствии с «геометрическим определением» описывается одним из двух соотношений:

$$L_P = \{x : x = Pu, u \in R^p\} = \{x : x = Px, x \in R^p\}.$$

8. *Классическое или векторное сингулярное (SVD-) представление произвольной $m \times n$ матрицы*. Для произвольной $A \in R^{m\times n}$ ранга $r \le \min(m,n)$, рассматриваемой как линейное отображение между евклидовыми пространствами векторов, справедливо следующее представление матрицы в виде взвешенной суммы тензорных произведений векторов

$$A = \sum_{i=1}^{r} \lambda_i u_i v_i^T \tag{2}$$

где

$\lambda_1^2 \ge ... \ge \lambda_r^2 > 0$ общий набор ненулевых собственных чисел матриц $AA^T, A^T A$;

$u_i \in R^m, i = \overline{i,r}$, - ортонормированный набор собственных векторов матрицы AA^T, отвечающих ненулевым собственным числам: $AA^T u_i = \lambda_i^2, \lambda_i^2 > 0, i = \overline{1,r}, u_i^T u_j = \delta_{ij}, i \ne j$;

$v_i \in R^n, i = \overline{1,r}$,-- ортонормированный набор собственных векторов матрицы $A^T A$, отвечающих ненулевым собственным числам: $A^T A v_i = \lambda_i^2 v_i, \lambda_i^2 > 0, i = \overline{1,r}, v_i^T v_j = \delta_{ij}, i \ne j$.

9. *Комментарий к п.8.* Собственно, действие оператора A на вектор $x \in R^n$ в SVD-разложении, представляет собой перенос разложения вектора по ортогональному базису или его части $v_i \in R^n, i = \overline{1,r}$, с координатами $x_i = v_i^T x = (v_i, x)_{R^n}, i = \overline{1,r}$, в одном пространстве в разложение по ортогональному базису $u_i \in R^m, i = \overline{j,r}$, или - его части, в другом пространстве с умножением координат на положительные числа, соответственно, $\lambda_i \in R^m, i = \overline{1,r}$. Поэтому вариантом записи SVD – разложения (2), может быть следующий:

$$Ax = \sum_{i=1}^{r} \lambda_i u_i (v_i, x)_{R^n}, x \in R^n.$$

Матрицы как линейные операторы между матричными евклидовыми пространствами

10. *Матрица $A \in R^{m\times n}$ как линейный оператор из матричного евклидового пространства $R^{n\times p}$ в $R^{m\times p}$*. Произвольная $m \times n$ матрица A может рассматриваться как матрица линейного оператора между двумя евклидовыми пространствами матриц $R^{n\times p}$ и $R^{m\times p}$ со следовым скалярным произведением. Этот

оператор описывается стандартным образом: для произвольной матрицы $X \in R^{n\times p}$ результатом действия оператора является матрица $AX \in R^{m\times p}$.

11. *Матричное сингулярное разложение* $m \times n$ *матрицы* $A \in R^{m\times n}$ *(M-SVD).*

Теорема 1. Справедливо следующее представление матрицы в виде суммы, реализующий перенос разложения по ортонормированным собственным элементам (собственным матрицам) матрицы $A^T A$ на ортонормированные собственные элементы матрицы $A^T A$

$$AX = \sum_{i=1}^{r}\sum_{k=1}^{p} \lambda_i u_i e_k^T (v_i e_k^T, X)_{tr} \tag{3}$$

где

$\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$ - общий набор ненулевых собственных чисел матриц $AA^T, A^T A$;

- $u_i e_k^T \in R^{m\times p}, i = \overline{1,r}, k = \overline{1,p}$, - ортонормированный набор собственных $m \times p$ матриц матрицы AA^T, отвечающих ненулевым собственным числам: $AA^T u_i e_k^T = \lambda_i^2 u_i e_k^T, \lambda_i^2 > 0, i = \overline{1,r}, k = \overline{1,p}, (u_i e_k^T, u_j e_l^T)_{tr} = \delta_{ij}\delta_{kl}, i \neq j, k \neq l$;
- $v_i e_k^T \in R^{n\times p}, i = \overline{1,r}, k = \overline{1,p}$, - ортонормированный набор собственных $n \times p$ матриц матрицы $A^T A$, отвечающих ненулевым собственным числам: $A^T A v_i e_k^T = \lambda_i^2 v_i e_k^T, \lambda_i^2 > 0, i = \overline{1,r}, k = \overline{1,p}, (v_i e_k^T, v_j e_l^T)_{tr} = \delta_{ij}\delta_{kl}, i \neq j, k \neq l$.

12. *Совпадение векторного и матричного варианта SVD: теорема свёртки*.

Теорема 2. Векторный и матричный вариант SVD-разложения одной и той же $m \times n$ матрицы $A \in R^{m\times n}$ совпадают между собой: для произвольной матрицы $X \in R^{n\times p}$

$$\left(\sum_{i=1}^{r} \lambda_i u_i v_i^T\right) X == \sum_{i=1}^{r}\sum_{k=1}^{p} \lambda_i u_i e_k^T (v_i e_k^T, X)_{tr} \tag{4}$$

Принципиальную роль в описании базовых структур евклидовых пространств играет псевдообращение по Муру – Пенроузу [Moore,1920], [Penrose, 1955] как одноместной операции A^+ над прямоугольными матрицами произвольной размерности. В дальнейшем термин «псевдообращение» будем сокращать до ПдО.

С учётом п.10 необходимо различать ПдО для матрицы, как векторного отображения, и ПдО для тойже матрицы, как матричного отображения. Первое, собственно, классическое по Муру-Пенроузу, будем обозначать как ВПдО, второе – как МПдО.

Псевдообращение: векторный и матричный вариант отображения

13.*Определение ВПдО через векторное SVD - представление матрицы*. Для произвольной $m \times n$ - матрицы A её ВПдО A^+ (классическое по Муру-Пенроузу) определяется соотношением

$$A^+ = \sum_{i=1}^{r} \lambda_i^{-1} v_i u_i^T : R^m \to R^n \tag{5}$$

14.*Определение МПдО через матричное SVD - представление матрицы.*

Теорема 3. Для произвольной $m \times n$ - матрицы A её МПдО A_M^+ определяется соотношением

$$A_M^+ X == \sum_{i=1}^{r}\sum_{k=1}^{p} \lambda_i^{-1} v_i e_k^T (u_i e_k^T, X)_{tr} : R^{m\times p} \to R^{n\times p} \tag{6}$$

15.*Совпадение векторного (классического) ПдО и ПдО матричного*.

Теорема 4. Матричный и векторный вариант ПдО для произвольной $m \times n$ - матрицы A совпадают между собой: т.е. совпадают операторы A^+, A_M^+, определяемые соотношениями (5), (6) соответственно: $A_M^+ = A^+$.

Доказательство. Результат очевидным образом вытекает из теоремы свёртки п.12.

Евклидовы пространства, базовые линейные структуры и ПдО

16. *Основные ортогональные проекторы (ОП), связанные с матрицей линейного оператора*. Основными ОП, связанными с матрицей A являются ОП $P(A^T), P(A)$ на подпространства L_A, L_{A^T} соответственно, а также $Z(A^T), Z(A)$ соответственно - на их ортогональные дополнения в R^m, R^n соответственно.. Они определяются соотношениями :

$$P(A^T) = AA^+ = \sum_{i=1}^{r} u_i u_i^T, P(A) = P\left((A^T)^T\right) = A^T (A^T)^+ = A^+ A = \sum_{i=1}^{r} v_i v_i^T,$$

$$Z(A) = E_n - P(A) = E_n - A^+ A, Z(A^T) = E_m - P(A^T) = E_m - A^{T+} A^T = E_m - AA^+.$$

Важность последних соотношений определяется тем, что $L_{A^T}^{\perp}$ является множеством нулей оператора A.

17. *Множество нулей (ядро) оператора* A. Подпространство $L_{A^T}^{\perp}$ является ядром, $KerA$, (множеством нулей) оператора A:

$$L_{A^T}^{\perp} = KerA = Z(A)R^n.$$

18. *Условие совместности системы линейных алгебраических уравнений(СЛАУ)*. Для совместности СЛАУ $Ax = y, y \neq 0$, необходимо и достаточно, чтобы $y^T Z(A^T) y = 0$. В этом случае $A^+ y$ является наименьшим по норме решением. Оно ортогонально к $KerA$, а множество всех решений Ω_y описывается соотношением

$$\Omega_y = A^+ y + Z(A)R^n = \{x : x = A^+ y + Z(A)v, v \in R^n\} \tag{7}$$

19. *Псевдорешения СЛАУ.* Если СЛАУ $Ax = y$ несовместна, т. е $y^T Z(A^T) y > 0$, то множество, определяемое соотношением (7) описывает совокупность всех наилучших квадратичных приближений правой части значениями левой:

$$\Omega_y = A^+ y + Z(A)R^n = \underset{x \in R^n}{Arg \min} \| Ax - y \|^2 \tag{8}$$

Значение невязки для любого наилучшего квадратичного приближения составляет $y^T Z(A^T) y$.

20. Для того, чтобы матричное уравнение (МаЛУ) $AX = Y, A \in R^{m \times n} X \in R^{n \times p}, Y \in R^{m \times p}$, имело решения необходимо и достаточно, чтобы $trY^T Z(A^T) Y = 0$. В этом случае множество Ω_Y определяется соотношением

$$\Omega_Y = \{X : X = A^+ Y + Z(A)V, V \in R^{n \times p}\} \tag{9}$$

В случае, когда МаЛУ несовместно, т.е., когда $trY^T Z(A^T) Y > 0$, множество (9) описывает решение оптимизационной задачи наилучшего квадратического приближения правой части Y значениями левой части AX того же уравнения:

$$\underset{X \in R^{n \times p}}{Arg} \min \| AX - Y \|_{tr}^2 = \Omega_Y = \{X : X = A^+Y + Z(A)V, V \in R^{n \times p}\}. \tag{10}$$

Величина невязки для каждого решения оптимизационной задачи составляет $trY^T Z(A^T)Y$.

*Замечание 1.*Обратим внимание, что множество решений МаЛУ, и множество псевдорешений, когда МаЛУ несовместно, описываются одним и тем же соотношением (9).

Замечание 2. Поскольку при совместности МаЛУ $\min\limits_{X \in R^{n \times p}} \| AX - Y \|_{tr}^2 = 0$ и этот минимум достигается на решениях МаЛУ, то множество Ω_Y из (9) описывает и множество решений и множество псевдорешений как решение одной и той же оптимизационной задачи

$$\Omega_Y = \{X : X = A^+Y + Z(A)V, V \in R^{n \times p}\} = \underset{X \in R^{n \times p}}{Arg} \min \| AX - Y \|_{tr}^2 \tag{11}$$

Замечание 3. Обратим внимание, что матрица A^+Y является решением, точным или псевдо-, МаЛУ. Это решение будем называть базовым.

21. Оптимизационное свойство Пенроуза для матричного ПдО.

Теорема 5. Базовое решение МаЛУ как в точном, так и в псевдоварианте, является единственным наименьшим по норме решением оптимизационной задачи поиска наилучшего квадратичного приближения правой части МаЛУ значениями левой части того же уравнения:

$$A^+Y = \underset{X \in \underset{X \in R^{n \times p}}{Arg} \min \| AX - Y \|_{tr}^2}{Arg} \min \| X \|_{tr} . \tag{12}$$

Доказательство. В силу замечания 2 и множество решений, и множество псевдорешений являются решением одной и той же оптимизационной задачи. Поэтому для любого $X \in \underset{x \in R^n}{Arg} \min \| AX - Y \|_{tr}^2$ в соответствии с (10) справедливо представление

$$X = A^+Y + Z(A)V, V \in R^{n \times N} . \tag{13}$$

Поскольку $Z(A)V \in Z(A)R^{n \times p}$ и, как нетрудно убедиться, для любого $V \in R^{n \times p}$

$$A^+Y \perp_{tr} Z(A)V ,$$

то в соответствии с многомерным вариантом теоремы Пифагора

$$\| X \|^2 = \| A^+Y + Z(A)V \|^2 = \| A^+Y \|^2 + \| Z(A)V \|^2 \geq \| A^+Y \|^2 .$$

Таким образом, для произвольного $X \in \underset{X \in R^{n \times p}}{Arg} \min \| AX - Y \|_{tr}$

$$\| X \|^2 \geq \| A^+Y \|^2 ,$$

и базовое решение A^+Y , псевдо- или точное, является наименьшим по норме решением оптимизационной задачи, определяемой правой частью соотношения (12). Единственность решения оптимизационной задачи вытекает из того,что $\| Z(A)V \|^2 = 0 \Leftrightarrow Z(A)V = 0$.

22. *Линейная зависимость вектора от набора векторов*. Для линейной зависимости вектора $d \in R^m$ от векторов набора $a(j) \in R^m, j = \overline{1,n}$, необходимо и достаточно, чтобы выполнялось соотношение $d^T Z(A^T)d = 0$ с $A = (a(1) \vdots ... \vdots a(n))$. Очевиднім образом, то же условие является необходимым и достаточным для линейной зависимости вектора $d \in R^m$ от столбцов $m \times n$ матрицы A .Отметим также,

что необходимым и достаточным условием линейной зависимости строки $a^T, a \in R^n$ от строк матрицы $A \in R^{m \times n}$ является условие $a^T Z(A) a = 0$.

23. В связи с ограниченностью работы только упомянем другие важные в применениях результаты, которые приведены в работе [Донченко et al, 2010]. Это прямые и обратные формулы Гревиля, а также формулы аналитического возмущения ПдО Н.Ф.Кириченко [Кириченко,1997]. В сущности, в этих результатах речь идёт о формулах, связывающих ПдО изменённой (возмущённой) матрицы с ПдО исходной, а также – с характеристиками возмущения. В формулах Гревиля таким возмущением является добавление (прямые) или вычёркивание (обратные) строки или столбца матрицы. Собственно, Гревилю принадлежит формула, касающаяся добавления строки, да ещё и только для случая независимости добавляемой строки от строк исходной матрицы. Результаты, связанные с добавлением зависимой строки, как и с вычёркиванием произвольного типа строки, принадлежат Н.Ф.Кириченко [Кириченко,1997].

В аналитических формулах возмущения [Кириченко,1997] изменение исходной матрицы происходит аддитивно: через добавление «простейшей» матрицы. В качестве «простейшей» матрицы выступает тензорное произведение ab^T двух векторов $a \in R^m, b \in R^n$. Как и в случае формул Гревиля, вид ПдО возмущённой матрицы определяется тем, являются ли компоненты возмущения зависимыми, или независимыми, от, соответственно, столбцов и строк возмущаемой матрицы. Кроме того, результат зависит также и от того, падает или сохраняется ранг возмущённой матрицы, когда составляющие элементы возмущения зависят от соответствующих составляющих возмущаемой матрицы. Обратим внимание, что все формулы или условия описываются явными аналитическими выражениями. Условия линейной зависимости или независимости для столбцов уже нашли своё отражение в п.22. Для описания зависимости от строк в соответствующем условии необходимо только заменить A^T на A.

Что же касается условий падения или сохранения ранга возмущённой матрицы, когда вектор a зависит от столбцов, а вектор b^T - от строк A, то они имеют вид:

$b^T A^+ a \neq -1$ сохраняется, $b^T A^+ a = -1$ падает.

24. *Квадрат расстояния вектора до гиперплоскости*. Квадрат расстояния $\rho^2(a, \Gamma(b, L_A))$ вектора $a \in R^m$ от гиперплоскости $\Gamma(b, L_A) = b + L_A \subseteq R^m$ определяется соотношением

$$\rho^2(a, \Gamma(b, L_A)) = \min_{y \in \Gamma(b, L_A)} \| a - y \|^2 = (a - b)^T Z(A^T)(a - b).$$

Отметим, что привязка подпространств в этом, как и в других случаях, к множеству значений оператора A, не ограничивает сферу применимости результатов. С помощью п.1 они очевидным образом распространяются на ситуацию, когда подпространство порождается конечной совокупностью векторов.

Базовые нелинейные структуры: группирующие операторы

Важнейшими нелинейными структурами евклидова пространства являются квадратичные формы (в работе - неотрицательно определённые) и отвечающие им эллипсы или эллипсоидальные цилиндры. Среди таких нелинейных структур принципиальными являются матрицы так называемых «группирующих операторов», которые естественным образом связаны с групповыми свойствами набора векторов. Группирующие операторы возникают в связи с набором векторов $a(j) \in R^m, j = \overline{1,n}$, и отвечающей ему матрицей $A = (a(1) \vdots \ldots \vdots a(n))$. Как и ортогональные проекторы, группирующие операторы являются парными. Будем обозначать их, соответственно, $R(A), R(A^T)$.

25. *Определение группирующих операторов*. Группирующие операторы, обозначаемые $R(A), R(A^T)$, определяются соотношениями

$$R(A^T)=A^{+^T}A^+,\ R(A)=A^+A^{+^T}.$$

Их важность для практики и область применения раскрываются в свойствах, приведённых ниже.

26.*Проектирование на нормированный вектор* $u\in R^m:\|u\|=1$, *элементов набора векторов* $a(j)\in R^m$, $j=\overline{1,n}$. Основной результат этого пункта представлен леммой 1 ниже.

Лемма 1. Для произвольного набора $a(j)\in R^m, j=\overline{1,n}$, с матричным представлением $A=(a(1)\vdots...\vdots a(n))$ и произвольного нормированного вектора $u\in R^m:\|u\|=1$, справедливо равенство:

$$\sum_{j=1}^{n}a^T(j)uu^Ta(j)=u^TAA^Tu \qquad (14)$$

Доказательство. Действительно, принимая во внимание связь п.2 векторов набора $a(j)\in R^m, j=\overline{1,n}$, со своим матричным представлением: $a(j)=Ae_{(j)}, j=\overline{1,n}$, имеем:

$$\sum_{j=1}^{n}a^T(j)uu^Ta(j)=\sum_{j=1}^{n}e^T{}_{(j)}A^Tuu^TAe_{(j)}=\sum_{j=1}^{n}u^TAe_{(j)}e^T_{(j)}A^Tu=u^TA\left[\sum_{j=1}^{n}e_{(j)}e^T_{(j)}\right]A^Tu\,.$$

Остаётся только заметить, что $\sum_{j=1}^{n}e_{(j)}e^T_{(j)}=E_n$, где E_n- единичная матрица в R^n.

Замечание 4. Левая часть соотношения (14) леммы 1 представляет собою сумму квадратов проекций векторов набора $a(j)\in R^m, j=\overline{1,n}$, на нормированный вектор $u\in R^m:\|u\|=1$.

27.Проектирование на элементы $u_i\in R^m, i=\overline{1,r}$, SVD –разложения матричного представления A набора векторов $a(j)\in R^m$, $j=\overline{1,n}$. Основной результат этого пункта представлен леммой 2, приведённой ниже.

Лемма 2. Для произвольного набора $a(j)\in R^m, j=\overline{1,n}$, с матричным представлением $A=(a(1)\vdots...\vdots a(n))$ имеет место соотношение:

$$\sum_{j=1}^{n}a^T(j)u_iu_i^Ta(j)=\lambda_i^2, i=\overline{1,r}\,.$$

Доказательство вытекает из леммы 1 предыдущего пункта и из п.8, в котором наборы $u_i\in R^m, v_i\in R^n, i=\overline{1,r}, r=rankA$, определяются как ортонормированные наборы собственных векторов матриц AA^T, A^TA, отвечающих общему набору ненулевых собственных чисел $\lambda_i^2>0, i=\overline{1,r}$.

28. Группирующие операторы: эллипсы группировки набора векторов $a(j)\in R^m, j=\overline{1,n}$. Основное утверждение пункта – теорема 6 ниже.

Теорема 6. Пусть $a(j)\in R^m, i=\overline{1,n}$, произвольный набор векторов из R^n с матричным представлением $A=(a(1)\vdots\ldots\vdots a(n)), rankA=r\le\min(m,n)$. Тогда все векторы набора принадлежат внутренности эллипса, точнее: эллипсоидального цилиндра, определяемого уравнением

$$x^TR(A^T)x=r, x\in R^m,$$

где, $R(A^T)$, группирующий оператор: $R(A^T)=A^{+^T}A^+$.

Доказательство. Рассмотрим квадраты проекций векторов набора $a(j), j=\overline{1,n}$, на каждый из векторов $u_i, i=\overline{1,r}$, SVD-представления (2) матрицы A. Принимая во внимание ортонормированность набора $u_i, i=\overline{1,r}$, и обозначая квадраты проекций через $\left\|\Pr_{u_i} a(j)\right\|^2, i=\overline{1,r},\ j=\overline{1,n}$, очевидным образом имеем

$$\left\|\Pr_{u_i} a(j)\right\|^2 = a^T(j)u_i u_i^T a(j), j=\overline{1,n}, i=\overline{1,r}\ .$$

Суммирование по всем векторам набора $a(j), j=\overline{1,n}$, и применение леммы 2 даёт

$$\sum_{j=1}^{n} a^T(j)u_i u_i^T a(j) = u_i^T AA^T u_i = \lambda_i^2, i=\overline{1,r}\ .$$

Таким образом, после деления обеих частей последнего соотношения на, соответственно, $\lambda_i^2, i=\overline{1,r}$, имеем

$$\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \frac{u_i^T AA^T u_i}{\lambda_i^2} = 1, i=\overline{1,r}\ ,$$

т.е.

$$\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = 1, i=\overline{1,r}\ .$$

Свернув (просуммировав) последнее равенство по $i=\overline{1,r}$, получаем

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \sum_{i=1}^{m} \frac{u_i^T AA^T u_i}{\lambda_i^2} = r\ .$$

Поменяв порядок суммирования в двойной сумме, имеем

$$\sum_{j=1}^{n}\sum_{i=1}^{m} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \sum_{j=1}^{n} a^T(j)\sum_{i=1}^{m} \frac{u_i u_i^T}{\lambda_i^2} a(j) = r\ .$$

Далее, приняв во внимание, что

$$\sum_{i=1}^{r} \frac{u_i u_i^T}{\lambda_i^2} = A^{+^T} A^{+} = R(A^T)\ ,$$

получаем окончательно

$$\sum_{j=1}^{n} \frac{a^T(j)u_i u_i^T a(j)}{\lambda_i^2} = \sum_{j=1}^{n} a^T(j)\sum_{i=1}^{m} \frac{u_i u_i^T}{\lambda_i^2} a(j) = \sum_{j=1}^{n} a^T(j)A^{+^T}A^{+}a(j) = \sum_{j=1}^{n} a^T(j)R(A^T)a(j) = r\ .$$

Таким образом,

$$\sum_{j=1}^{n} a^T(j)R(A^T)a(j) = r \tag{15}$$

Поскольку $R(A^T)$ – симметричная, неотрицательно определённая матрица, то следствием соотношения (15) является одновременное выполнение неравенств

$$a^T(j)R(A^T)a(j) \le r, j=\overline{1,n} \tag{16}$$

Та же симметричность и неотрицательная определённость позволяет сделать вывод, что уравнение

$$x^T R(A^T)x = r, x \in R^m \tag{17}$$

определяет эллипс, точнее: эллипсоидальный цилиндр, в R^m с длинами $\frac{1}{\lambda_i\sqrt{r}}, i=\overline{1,r}$ нетривиальных полуосей. Напомним, что $r = rankA \le \min(m,n)$.

Таким образом, выполнение неравенства (16) для всех векторов набора $a(j)\in R^m, j=\overline{1,n}$, означает их одновременную принадлежность внутренности эллипсоидального цилиндра с уравнением (17), и доказательство теоремы завершено.

Замечание 5. В действительности неравенство (16) может давать существенное загрубление «радиуса» эллипса. Так, при векторах $a(j)\in R^m, j=\overline{1,n}$, близких к ортогональным, очевидным образом, константу в правой части (16) можно выбрать близкой к 1.

29. *Усиление результата об эллипсах группировки.*

Теорема 7. Все векторы набора $a(j)\in R^m, i=\overline{1,n}$, с матричным представлением $A=\left(a(1)\vdots\ldots\vdots a(n)\right)$ принадлежат внутренности эллипсоидального цилиндра

$$x^T R(A^T)x = r_{\max}^2, r_{\max}^2 \le r = rankA, x\in R^m, r_{\max}^2 = \max_{j=\overline{1,n}} a^T(j)R(A^T)a(j),$$

который будем называть минимальным эллипсом группировки для рассматриваемого набора векторов.

Группирующие операторы и расстояние Махаланобиса

Группирующие операторы с точностью до скалярного множителя неожиданным образом оказываются связанными с расстоянием Махаланобиса [Mahalanobis,1936] (см. также, например, [McLachlan, Geoffry, 1992]), и проясняют его суть. Напомним, что расстояние Махаланобиса (в дальнейшем РаМ) является одним из способов определения степени принадлежности $\rho^2(x,N)$ неслучайного вектора $x\in R^m$ многомерному $N(a,B)$ нормальному распределению, и определяется соотношением

$$\rho^2(x,N) = (x-a)^T B^{-1}(x-a). \tag{18}$$

Напомним также, что $a\in R^m$ является математическим ожиданием, а симметричная неотрицательно определённая матрица $B: B\in R^{m\times m}, B^T=B, B\ge 0$, является матрицей ковариаций распределения. Неслучайный вектор $x\in R^m$ интерпретируется как возможный элемент выборки.

Как правило, при использовании РаМ предполагается, что наблюдаемое значение может относиться к одному из нескольких $N_k, k=\overline{1,K}$ нормальных распределений $N(a_k,B_k), k=\overline{1,K}$. В этом случае для каждого из возможных распределений определяется своё РаМ. В этом случае рассматривается набор расстояний вида (18):

$$\rho^2(x,N_k) = (x-a_k)^T B_k^{-1}(x-a_k), \tag{19}$$

а отнесение вектора $x\in R^m$ к одному из K возможных нормальных распределений осуществляется по миниммуму РаМ'ов из (19).

Отметим для сравнения, что средний квадрат расстояния от неслучайного вектора $x\in R^m$ до $N(a,B)$ распределённой случайной величины ξ определяется выражением:

$$\mathrm{M}\|\xi - x\|^2 = trB + \|x-a\|^2,$$

а квадратичная форма $(x-a)^T B(x-a)$ описывает дисперсию случайной величины $(x-a)^T\xi$.

Таким образом, квадратичная форма (18) не является адекватной с точки зрения описания естественных квадратических характеристик нормально распределённого случайного вектора на основе среднего.

Единственным обоснованием использования расстояния Махаланобиса яаляется присутствие квадратичной формы из (18) в плотности распределения $f(x), x \in R^m$, определяемой соотношением

$$f(x) = (2\pi)^{-\frac{m}{2}} |B|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(x-a)^T B^{-1}(x-a)\}, x \in R^m .$$

Заканчивая знакомство с РаМ, отметим, что при его практическом использовании вместо параметров a, B используются их оценки $\hat{a}, \hat{B}$ на основе выборки $a(1),...,a(n) \in R^m$, определяемые соотношениями

$$\hat{a} = \bar{a} = \frac{1}{n}\sum_{i=1}^{n} a(i) , \hat{B} = \frac{1}{n}\sum_{i=1}^{n}(a(i)-\bar{a})(a(i)-\bar{a})^T . \quad (20)$$

В этои случае РаМ приобретает статистический вид

$$\hat{\rho}^2(x,N) = (x-\hat{a})^T \hat{B}^{-1}(x-\hat{a}). \quad (21)$$

Теорема 8. Статистический вид РаМ может быть представлен в виде

$$\hat{\rho}^2(x,N) = n(x-\bar{a})^T R(\tilde{A}^T)(x-\bar{a}) ,$$

где $\tilde{A} = (a(1)-\bar{a} \vdots ... \vdots a(n)-\bar{a})$.

Доказательство. Действительно, из (20) непосредственно следует, что $\hat{a} = \bar{a}$. Из тех же соотношений, с использованием результата п.3 вытекает, что

$$\hat{B} = \frac{1}{n}\sum_{i=1}^{n}(a(i)-\bar{a})(a(i)-\bar{a})^T = \frac{1}{n}\tilde{A}\tilde{A}^T \quad (22)$$

Если матрица $\hat{B}$ невырожденная и определяется соотношением (22), то $\hat{B}^{-1} = n(\tilde{A}\tilde{A}^T)^{-1}$. Из п.8 векотрного SVD-разложения для матрицы $\tilde{A}$ вытекает, что

$$\tilde{A}\tilde{A}^T = \sum_{i=1}^{n} \lambda_i^2 u_i^T u_i \quad (23)$$

Из (23) в свою очередь следует, что матрица $(AA^T)^{-1}$, если она существует, определяется соотношением

$$(\tilde{A}\tilde{A}^T)^{-1} = \sum_{i=1}^{m} \frac{u_i^T u_i}{\lambda_i^2} \quad (24)$$

Как нетрудно убедиться, правая часть в (24) совпадает с $R(\tilde{A}^T)$. Таким образом, окончательно имеем:

$$\hat{B}^{-1} = nR(\tilde{A}^T) ,$$

что и завершает доказательство теоремы.

Очевидна, таким образом, эвристическая основа расстояния Махаланобиса и его жёсткая теоретико вероятностная привязка. Последнее означает, что вне теории вероятностей и математической статистики и вне связи с многомерным нормальным распределением говорить о расстоянии Махаланобиса не имеет смысла.

В то же время группирующие операторы предоставляют возможности выявления групповых свойств векторов на основе минимальных эллипсов группировки в любой ситуации.

Применения: кластеризация

Примеры применения базовых линейных структур в линейной регрессии, в линейных системах управления с дискретным временем, а также для специального класса функциональных сетей, обобщающих искусственные нейронные сети, можно найти в работах [Кириченко, Донченко,2005] , [Донченко et al 2010].

Специально остановимся на применении базовых структур в задачах кластеризации. Отметим, что ПдО расширяет возможности кластеризации, позволяя эффективно погружать классифицируемые объекты в подходящие подпространства или гиперплоскости п.1 или минимальные эллипсы группировки. Теория ПдО даёт возможность связывать подпространство, порождённое набором векторов, как и минимальный эллипс группировки, с подходящей матрицей, строящейся конструктивно и в аналитическом виде. Если объект связывается с гиперплоскостью, то её смещение – это, как правило, среднее по векторам порождающей совокупности, а подпространство – это подпространство значений матрицы, построенной из центрированных средним векторов порождающей совокупности, как из столбцов. В случае руппирования минимальным эллипсом группировки его центр – среднее по векторам совокупности, а матрица квадратичной формы – это, с точностью до транспонирования, группирующий оператор для матрицы, столбцовым представлением которой являются ценрированные средним векторы исходной совокупности. Результаты п.25 обеспечивают возможность конструктивного вычисления расстояний от объектов (подпространств или гиперплоскостей), ассоциируемых с порождающей совокупностью в то время, как п. 29 – констрруктивно строить расстояния соответствия на основе минимальных эллипсов группировки. Недостаток места не даёт возможности остановиться на результатх, обеспечивающих рекуррентность вычисления необходимых характеристик. С подробностями можно ознакомиться, например, в работах: [Кириченко, Донченко, 2007], [Кириченко, Донченко, 2008]).

Как уже было отмечено, в качестве расстояний соответствия в задачах кластеризации могут быть использованы как расстояния от гиперплоскостей, так и расстояния на основе минимальных эллипсов группировки. Последние дают возможность строить расстояния соответствия в связи с использованием базовой нелинейной структуры. В сущности, речь идёт о том, что результат теоремы 7 можно использовать для определения расстояния соответствия, точнее его квадрата, обозначим его $\rho^2(x,Kl)$, между вектором $x \in R^m$ и кластером Kl, порожденным обучающей выборкой $a(j) \in R^m, j = \overline{1,n}$.

Обозначим через $\bar{a}$ среднее обучающей выборки $a(j) \in R^m, j = \overline{1,n}$, а через $\tilde{A}$ матричное представление для векторов $\tilde{a}(j) \in R^m, j = \overline{1,n}$, первоначального набора, центрированных средним:

$$\bar{a} = \frac{1}{n}\sum_{j=1}^{n} a(j),$$

$$\tilde{a}(j) = a(j) - \bar{a},$$

$$\tilde{A} = (\tilde{a}(j) \vdots \ldots \vdots \tilde{a}(j)).$$

Тогда квадрат расстояния, определяемый соотношением

$$\rho^2(x,Kl) = \frac{1}{r_{\max}^2}(x-\bar{a})^T R(\tilde{A}^T)(x-\bar{a}),$$

является расстоянием, определяемым минимальным эллипсом группировки

$$(x-\bar{a})^T R(\tilde{A}^T)(x-\bar{a}) = r_{\max}^2 \qquad (25)$$

Уравнение (25) определяет минимальный эллипс группировки векторов $\tilde{a}(j) \in R^m, j = \overline{1,n}$ ’ $\bar{a}$.

При наличии набора кластеров $Kl_l, l = \overline{1,L}$, расстояния соответствия определяются, соотношениями

$$\rho^2(x,Kl_l) = \frac{1}{r_{l,\max}^2}(x-\bar{a}_l)^T R(\tilde{A}_l^T)(x-\bar{a}_l), l = \overline{1,L}, \qquad (26)$$

с очевидной детализацией обозначений в связи с употреблением соответствующего индекса. Отнесение к тому или иному кластеру осуществляется по минимуму расстояний (26).

Примеры эффективного применения расстояний соответствия вида (26) можно найти, например, в работах [Кириченко, Донченко, 2008], [Кириченко, Донченко, 2007], [Донченко, Омардибирова 2005].

Заключение

В работе предложена и обоснована концепция базовых структур евклидового пространства, как линейных, так и нелинейных. Изложены конструктивные способы описания и взаимного перехода от одних типов структур к другим. В числе других рассмотрены конструктивные способы порождения, описания и использования базовых структур. Упомянутая конструктивность обеспечивается применением как классических результатов ПдО, так и новыми результатами в этой области. В частнсти, в работе приведена теорема о сингулярном разложении матрицы как линейного оператора между матричными пространствами, доказана теорема свёртки, обеспечивающая эквивалентность векторного и матричного сингулярных и ПдО-матриц на основе соответствующих сингулярных разложений. Рассмотрены применения полученных результатов для построения кластеризации с разными вариантами расстояний соответствия.

Литература

[Донченко et al, 2010] Донченко В.,Кривонос Ю., Омардибирова В. Базовые структуры евклидовых пространств: конструктвные методы описания и использования/ New Trends in Classification and Data Mining. – ITHEA, Sofia, Bulgaria. -2010.- ISBN 978-954-16-0042-9.- P. 155-170..

[Донченко. 2009] Донченко В.С. Неопределённость и математические структуры в прикладных исследованиях/ Human aspects of Artificial Intelligence International Book Series Information science & Computing.– Number 12. – Supplement to International Journal “Information technologies and Knowledge”. –Volume 3.–2009. – P. 9-18.

[Донченко, Омардибирова 2005] Донченко В.С., Омардибирова В.Н. Технология классификации электронных документов с использованием теории возмущения псевдообратных матриц// Proceedings of the XI-th International Conference “Knowledge-Dialogue-Solution”. – June 20-30, Varna, 2005. – Volume 1. – C.223-226.

[Кириченко, Донченко, 2008] В.С Кириченко Н.Ф. Донченко В.С. Гиперплоскости в «множествах и расстояниях соответствия»: кластеризация / Artificial Intelligence and Decision Making.– International book series “INFORMATION SCIENCE&COMPUTING”, Number 7.– Sofia 2008.– P. 25-36.

[Moore, 1920] Moore E.H. On the reciprocal of the general algebraic matrix // Bulletin of the American Mathematical Society. – 26, 1920. – P.394 -395.

[Penrose, 1955] Penrose R. A generalized inverse for matrices // Proceedings of the Cambridge Philosophical Society 51, 1955. – P.406-413.

[Алберт, 1977] Алберт А. Регрессия, псевдоинверсия, рекуррентное оценивание. – М.: Наука. – 1977.– 305 с.

[Кириченко, 1997] Кириченко Н.Ф. Аналитическое представление псевдообратных матриц //Киб. и СА.- №2. –1997.– С.98-122.

[Кириченко, Донченко,2005] Кириченко М.Ф., Донченко В.С. Задача термінального спостереження динамічної системи: множинність розв’язків та оптимізація//Журнал обчислювальної та прикладної математики. – 2005. –№5– С.63-78.

[Кириченко, Донченко, 2007] Кириченко Н.Ф., Донченко. В.С. Псевдообращение в задачах кластеризации// Киб. и СА.- №4, 2007– С.98-122.

[Кириченко, Донченко, 2008] Кириченко Н.Ф. Донченко В.С. Гиперплоскости в «множествах и расстояниях соответствия»: кластеризация / Artificial Intelligence and Decision Making.– International book series “INFORMATION SCIENCE&COMPUTING”, Number 7.– Sofia 2008. – P. 25-36.

[Mahalanobis,1936] Mahalanobis, P C. On the generalized distance in statistics.//Proceedings of the National Institute of Sciences of India.- 1936.-2 (1).-P. 49–55.

[McLachlan, Geoffry, 1992] McLachlan, Geoffry J. Discriminant Analysis and Statistical Pattern Recognition. Wiley Nescience.-1992.- ISBN 0471691151.

Информация об авторе

Владимир С. Донченко *– профессор; Киевский национальный университет имени Тараса Шевченко, факультет кибернетики, Украина, e-mail: voldon@unicyb.kiev.ua*