Журнал обчислювальної та прикладної математики | 2015, №3(120) | Journal of Computational & Applied Mathematics

УДК 512.64:004.93
MSC 15Bxx, 68T10

CORTEGE CONCEPTION FOR LINEAR OPERATORS AND ITS IMPLEMENTATION FOR MATRIX CORTEGES

Volodymyr Donchenko, Taras Zinko, Fedir Skotarenko
Faculty of Cybernetics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine, E-mail: voldon@unicyb.kiev.ua.

КОНЦЕПЦІЯ КОРТЕЖНОСТІ ДЛЯ ЛІНІЙНИХ ОПЕРАТОРІВ ТА ЇЇ РЕАЛІЗАЦІЯ ДЛЯ МАТРИЧНИХ КОРТЕЖІВ

В. С. Донченко, Т. П. Зінько, Ф. О. Скотаренко
Факультет кібернетики, Київський національний університет імені Тараса Шевченка, Київ, Україна, E-mail: voldon@unicyb.kiev.ua.

Abstract. Concept of cortege linear operators is proposed and implemented for the variant when a cortege is a sequence of matrices with fixed dimensions. The proposed concept lets you transfer means of description the basic linear and nonlinear structures of Euclidean vectors space in matrix Euclidean space preserving the basic content and interpretations. It is principal that the complexity of the problems in matrices spaces is no more complex, than in vector spaces. In both cases the base is conventional eigenvalue problem for matrices.
Keywords: linear operators, matrices, the pseudoinverse, Singular Value Decomposition, cortege.

Резюме. В роботі висунута концепція кортежних операторів, яка реалізована для варіанту, коли кортежем є набір матриць фіксованої розмірності. Запропонована концепція дозволяє перенести конструктивні засоби опису основних лінійних та нелінійних структур евклідового простору числових векторів на матричні евклідові простори із збереженням основних змістів та інтерпретацій у запропонованій на основі „кортежної концепції“ формалізму. Принциповим є те, що складність опису згаданих структур в матричних евклідових просторах не перевищує складності опису відповідників для евклідових просторів числових векторів і зводиться до задачі на власні значення для матриць.
Ключові слова: лінійні оператори, матриці, псевдообернення, сингулярне подання, кортежі.

Вступ

Евклідові простори числових векторів R^n та їхнє використання в прикладних дослідженнях визначається з одного боку — багатством структур в рамках цих просторів, породжених операціями та скалярним добутком, з іншого — багатством конструктивних можливостей в оперуванні таким структурами. До перших: структур в евклідових просторах, які можна називати базовими, — можна, наприклад, віднести лінійні підпростори та гіперплощини (зсунені підпростори). Це ті структури, які можна позначити як множинні, як і еліпсоїди чи еліпсоїдальні циліндри. З іншого боку до базових структур можні віднести матриці лінійних операторів та невід'ємно визначених квадратичних форм. Ці базові утворення в евклідових просторах можна позначити як синглові структури. В евклідових просторах R^n множинні та синглові структури пов'язані між собою. Так з матрицею лінійного оператора природнім чином пов'язуються два підпростори: множина можливих значень та ядро оператора, а множина можливих розв'язків чи псевдорозв'язків СЛАР (системи лінійних алгебраїчних рівнянь) є гіперплощиною. Так само пов'язані між собою еліпсоїдальні циліндри та матриці невід'ємно визначених квадратичних форм, що їх задають. Ефективність та конструктивного використання базових структур значною мірою визначається засобами оперування з ними, зокрема, можливістю переходу від синглових до множинних і навпаки. Так, наприклад множина розв'язків СЛАР: точних чи псевдо — це гіперплощина (зсунутий підпростір). Це множинна структура, її конструктивний опис визначається можливістю її визначення за сингловою структурою — матрицею СЛАР. Такі можливості реалізуються через SVD (сингулярний розклад) та ПдО (псевдообернення) матриці СЛАР. Так само SVD та ПдО дозволяють побудувати ортогональні проектори на підпростори та гіперплощини, породжені наборами векторів: побудувати синглові структури за множинними. Такий перехід реалізує можливість ефективного обчислення відстаней від досліджуваних множинних утворень. Так само, SVD та ПдО, а також оператори на їх основі реалізує можливості ефективного опису матриці квадратичної форми, що задають еліпсоїди, які оптимальним, в певному сенсі, чином „накривають“ заданий набір векторів. Згаданий зв'язок між множинними та сингловими структурами, можливості його конструктивного опису є дуже важливими у прикладних задачах, які можна визначити терміном „задачі групування інформації“(GIP): задач класифікації-кластеризації чи відновлення функції, представленою своїми спостереженнями.

Однак GIP-задачі є важливими не тільки в ситуаціях, коли об'єктами групування є числові вектори. Типовим прикладом згаданої ситуації є задачі розпізнавання аудіо сигналів, які природнім чином представлені матрицею спектрограми та матриці, якими представляються зображення в задачах обробки зображень. Отже, розвиток засобів розв'язання GIP-задач для матричних об'єктів, аналогічних тим, що існують для евклідових просторів числових векторів, є нагальною та важливою задачею для прикладних досліджень. В математичному плані, це означає розвиток методів, які б забезпечували ті ж можливості у використанні базових матричних структур, що й в R^n. Як зазначалося вище, такі можливості для R^n визначаються технікою SVD та ПдО для матриць. Звісно, перенесення згаданої фундаментальної техніки на евклідові простори, відмінні від R^n, зокрема — на матричні евклідові простори, — вимагає зміни погляду на самі ці об'єкти. Це означає, перехід від погляду на SVD та ПдО як на матричні властивості, до формулювання погляду на них як на операторні властивості. Останнє означає, що SVD та ПдО стають способом подання лінійних операторів між слушними евклідовими просторами. Саме це складає основу концепції кортежних операторів для розв'язання задачі розвинення техніки оперування з базовими структурами в матричних евклідових просторах.

Стаття присвячена розгляду адекватних класів операторів між слушними евклідовими просторами, які би уможливили побудову засобів оперування з базовими алгебраїчними структурами в евклідовому просторі матриць фіксованої розмірності, аналогічними тим, що використовуються в R^n із тим самим рівнем конструктивності такого використання.

Згаданим слушним класом лінійних операторів, що дозволяє реалізувати програму створення засобів конструктивного оперування з базовими структурами матричних евклідових просторів є так звані „кортежні" оператори за рядковими кортежами матриць між евклідовими просторами R^n та $R^{m\times n}$. Рівень конструктивності для таких операторів в побудові SVD та ПдО відповідає рівню конструктивності для R^n, оскільки сингулярний розклад таких операторів та відповідно, — ПдО — зводяться до сингулярного подання звичайних матриць. Отже, пропонована техніка реалізується тими ж технічними засобами розв'язання задач на власні значення, що і в класичному випадку.

1. Абстрактний варіант сингулярного подання лінійних операторів між евклідовими просторами

SVD-подання матриці, яке полягає в тому чи іншому варіанті її факторизації, є важливим засобом конструктивного опису та оперування із базовими структурами R^n. Побудова апарату конструктивного оперування для інших типів евклідових просторів потребує наданню SVD-поданню матриці ширшого „операторного" значення, що є предметом теореми 1.

Нижче під сингулярністю оператора розумітиметься пара власне число — власний вектор, через A^* — позначатиметься спряжений до A оператор.

Теорема 1. *(про SVD-подання лінійного оператора над довільними евклідовими просторами). Для довільної лінійного оператора A між евклідовими просторами E_1, E_2 : $A : E_1 \to E_2$ — справедливе наступне його подання у вигляді суми за елементами ненульових сингулярностей операторів AA^*, A^*A*

$$A = \sum_{i=1}^{r} \lambda_i u_i \ell_{v_i} = \sum_{i=1}^{r} \lambda_i u_i (v_i, \cdot), \tag{1}$$

де

– $r = \dim R(A) = \dim L_A$ *розмірність підпростору можливих значень;*
– $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$ — *спільний набір ненульових власних чисел операторів* AA^*, A^*A;
– $u_i \in E_2$, $i = \overline{1,r}$ — *ортонормований набір власних векторів оператора* AA^*, *що відповідають ненульовим власним числам* $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$: $AA^*u_i = \lambda_i^2$, $\lambda_i^2 > 0$, $i = \overline{1,r}$, $u_i^T u_j = \delta_{ij}$, $i \neq j$;
– $v_i \in E_1$, $i = \overline{1,r}$ — *ортонормований набір власних векторів оператора* A^*A, *що відповідають ненульовим власним числам* $\lambda_1^2 \geq ... \geq \lambda_r^2 > 0$: $A^*Av_i = \lambda_i^2 v_i$, $\lambda_i^2 > 0$, $i = \overline{1,r}$, $v_i^T v_j = \delta_{ij}$, $i \neq j$, *крім того,*

$$u_i = \frac{Av_i}{\lambda_i},\ i = \overline{1,r}, \qquad v_i = \frac{A^*u_i}{\lambda_i},\ \ i = \overline{1,r}. \tag{2}$$

Доведення. Доведення здійснюється так само, як для матриць: простори можливих значень оператора та спряженого співпадають із лінійними оболонками відповідних наборів: $v_i \in E_1$, $i = \overline{1,r}$ та $u_i \in E_2$, $i = \overline{1,r}$. Включення в один бік є очевидними, а співпадіння визначається однаковими розмірностями відповідних просторів. Далі перевіряється співпадіння вихідного оператора на $v_i \in E_1$, $i = \overline{1,r}$, а, отже і на $L(v_1, ..., v_r)$. Останній підпростір є рейнджем — $\Re(\mathrm{A}^*)$ (множина можливих значень) для спряженого оператора, а його ортогональне доповнення $L^{\perp}(v_1, ..., v_r)$ — ядром оператора A. Отже, на векторах цього підпростору оператор A приймає нульові значення, а оператор, що визначається сумою — теж, оскільки всі вектори ортонормованого набору $v_i \in E_1$, $i = \overline{1,r}$ є ортогональними до векторів $L^{\perp}(v_1, ..., v_r)$. Оскільки, $E_1 = L(v_1, ..., v_r) + L^{\perp}(v_1, ..., v_r)$, то за лінійністю операторів в лівій та правій частині, отримуємо співпадіння їхніх значень на будь яких векторах — елементах області можливих значень. □

Для матриць заміна спряженості на транспонування, а ℓ_{v_i}, $i = \overline{1,r}$ на v_i^T, $i = \overline{1,r}$ дає „операторний“ варіант теореми про SVD-подання матриці.

Наслідок 1. *Теорема про (SVD подання матриці лінійного оператора* ([1], див. також [2])). Довільна матриця $A \in R^{m \times n}$ рангу $r \leq \min(m, n)$, може бути подана у вигляді

$$A = \sum_{i=1}^{r} \lambda_i u_i v_i^T, \tag{3}$$

за ненульовими наборами сингулярностей $\left(u_i, \lambda_i^2\right)$, $\left(v_i, \lambda_i^2\right)$, $i = \overline{1,r}$ матриць AA^T, A^TA.

Важливою є наступна теорема 2, яка має „обернений“ до теореми 1 характер.

Теорема 2. *Нехай в абстрактних евклідових просторах* E_1, E_2 *задані:*

1) *набір додатних чисел* $\lambda_i^2 : \lambda_1^2 \geq ... \geq \lambda_r^2 > 0$, $i = \overline{1,r}$, $r \leq \min(\dim E_1, \dim E_2)$;
2) *ортонормований набір векторів* $v_i \in E_1, i = \overline{1,r}$;

3) *ортонормований набір векторів* $u_i \in E_2, i = \overline{1, r}$.

Тоді для оператора $A_{\lambda,u,v}$, *визначеного за згаданими вище елементами співвідношенням*

$$A_{\lambda,u,v} \equiv \sum_{i=1}^{r} \lambda_i u_i \ell_{v_i} \tag{4}$$

його сингулярне подання співпадає із співвідношенням (4), *що його задає.*

Доведення. Доведення випливає з того, що, як свідчить безпосередня перевірка, $(v_i, \lambda_i^2), (\mathrm{u}_i, \lambda_i^2)$, $i = \overline{1, r}$ є повними ортонормованими наборами сингулярностей для операторів $A^*_{\lambda,u,v} A_{\lambda,u,v}$, $A_{\lambda,u,v} A^*_{\lambda,u,v}$, відповідно. Що й доводить теорему. $\square$

2. „Абстрактний“ варіант псевдообернення

В „абстрактному“ варіанті визначення псевдообернення (ПдО): ПдО для лінійних операторів між абстарактними евклідовими просторами, — слушним інструментом є визначення ПдО за SVD-поданням оператора.

Означення 1. ПдО для лінійного оператора A між евклідовими просторами E_1, E_2 називатимемо оператор $A^+ : E_2 \to E_1$, що визначається співвідношенням:

$$A^+ = \sum_{i=1}^{r} \lambda_i^{-1} v_i \ell_{u_i} = \sum_{i=1}^{r} \lambda_i^{-1} v_i (u_i, \cdot). \tag{5}$$

Зауваження. За оберненою теоремою про SVD-подання співвідношення (5) визначає сингулярне подання оператора A^+.

Всі властивості матричного ПдО залишаються справедливими для загального варіанту визначення з точністю до заміни транспонування на спряженість, а реалізації скалярного добутку засобами матричної алгебри — самим скалярним добутком.

Основні властивості „абстрактного“ ПдО

1. $(A^*)^+ = (A^+)^*$.
2. $P_{L_{A^*}} = A^+ A$, $Z(A) = I_{E_1} - A^+ A = P_{KerA}$ — оператори ортогонального проектування на підпростір можливих значень оператора A^* та ядро оператора A, відповідно.
3. $P_{L_A} \equiv A^{*+} A^* = A A^+, Z(A^*) = I_{E_2} - A A^+ = P_{KerA^*}$ оператори ортогонального проектування на підпростір можливих значень оператора та ядро оператора A^*, відповідно.
4. $A^* \left(A A^*\right)^+ = \left(A^* A\right)^+ A^*$.
5. Множина розв'язків чи псевдорозв'язків Ω_y СЛАР $Ax = y$ задається співвідношенням

$$\Omega_y = A^+ y + Z(A) E_1. \tag{6}$$

Якщо $(y, Z(A^*)y) = 0$, множина (6) задає множину розв'язків, якщо $(y, Z(A^*)y) > 0$ — псевдорозв'язків. Умова $(y, Z(A^*)y) = 0$ є необхідною та достатньою для розв'язності СЛАР.

Під псевдорозв'язками мається на увазі розв'язки оптимізаційної задачі пошуку $Arg \min\limits_{x \in R^n} ||Ax - y||^2$ — найкращого квадратичного наближення правої частини СЛАР значеннями лівої частини.

1. Групувальні оператори $R(A)$, $R(A^*)$ визначаються співвідношеннями: $R(A) = A^+ A^{*+}$, $R(A^*) \equiv A^{*+} A^+$.
2. $\underbrace{A^* R(A^*) A}_{A^+ A} = A^+ A = P_{L_{A^*}}$.

3. „Кортежні оператори“

В роботі [2] висунута концепція кортежних операторів, реалізована в згаданій роботі для варіанту матричних рядкових кортежів за матрицями фіксованої розмірності. Ця концепція передбачає, що реалізація програми перенесення техніки оперування з основними базовими структурами в евклідових просторах числових векторів на матричні евклідові простори розв'язується через введення та дослідження слушних лінійних операторів над слушними евклідовими просторами. Такими лінійними операторами є ті, що нижче позначаються терміном „кортежні“.

В подальшому використовуватиметься матричний евклідів простір $R^{m \times n}$ — матриць фіксованої розмірності $m \times n$ з покоординатними операціями додавання, а також множення на скаляр, та з „покоординатним“ визначенням скалярного добутку:

$$(A, B) = \sum_{i=\overline{1,m}, j=\overline{1,n}} a_{ij} b_{ij} : \qquad A = (a_{ij}),\ B = (b_{ij}) \in R^{m \times n}.$$

Цей скалярний добуток називають також „слідовим“, оскільки $(A, B) = = trA^T B$.

Називатимемо впорядкований набір $\alpha = (A_1, ..., A_K)$ матриць фіксованої розмірності: $A_k \in R^{m \times n}$, $k = \overline{1, K}$ — рядковим матричним кортежем (довжини K).

Означення 2. Кортежним оператором $\wp_\alpha$, що задається рядковим матричним кортежем $\alpha = (A_1, ..., A_K)$, називатимемо лінійний оператор $\wp_\alpha : R^K \to R^{m \times n}$, що задається співвідношенням:

$$\wp_\alpha x = \sum_{k=1}^{K} x_k A_k,\ \alpha = (A_1, ..., A_K), \quad x^T = (x_1, ..., x_K) : x \in R^K.$$

Очевидним чином має місце наступна теорема, що є відповідником теореми для R^n, за якою лінійний підпростір значень лінійного оператора є лінійною оболонкою стовпчиків матриці, якою цей оператор задається.

Лема 1. *Лінійний підпростір $L(A_k,\ k = \overline{1, K})$, породжений набором матриць $A_k \in R^{m \times n}$, $k = \overline{1, K}$ співпадає із лінійним підпростором $L_{\wp_\alpha}$ можливих значень „кортежного оператора“ $\wp_\alpha$ за кортежем $\alpha = (A_1, ..., A_K)$.*

Лема 2. *Спряженим $\wp^*_\alpha$ до кортежного оператора $\wp_\alpha$ є лінійний оператор $\wp^*_\alpha : R^{m\times n} \to R^K$* [6],

$$\wp^*_\alpha Y = \begin{pmatrix} (A_1, Y) \\ \cdots \\ (A_K, Y) \end{pmatrix} = \begin{pmatrix} trY^T A_1 \\ \cdots \\ trY^T A_K \end{pmatrix} = \begin{pmatrix} \ell_{A_1} Y \\ \cdots \\ \ell_{A_K} Y \end{pmatrix} = \begin{pmatrix} \ell_{A_1} \\ \cdots \\ \ell_{A_K} \end{pmatrix} Y, \quad (7)$$

де $Y \in R^{m\times n}$, ℓ_A — це лінійний функціонал на $R^{m\times n}$, породжений $A \in R^{m\times n}$ у відповідності до співвідношення

$$\ell_A Y = (A, Y),\ Y \in R^{m\times n}.$$

Зауваження. Зазначимо, що співвідношення (7), визначає „стовпчиковий кортежний оператор“ $\wp_\chi : R^{m\times n} \to R^K$, що задається стовпчиковим кортежем χ, ($Y \in R^{m\times n}$):

$$\chi = \begin{pmatrix} \ell_{A_1} \\ \cdots \\ \ell_{A_K} \end{pmatrix} :\ \wp_\chi Y = \begin{pmatrix} (A_1, Y) \\ \cdots \\ (A_K, Y) \end{pmatrix} = \begin{pmatrix} \ell_{A_1} Y \\ \cdots \\ \ell_{A_K} Y \end{pmatrix} = \begin{pmatrix} \ell_{A_1} \\ \cdots \\ \ell_{A_K} \end{pmatrix} Y.$$

Лема 3. *Добуток двох операторів $\wp^*_\alpha \wp_\alpha : R^K \to R^K$ є лінійним оператором що визначається стандартним чином $K \times K$ матрицею F, яка визначається співвідношенням* [6]

$$F \equiv \wp^*_\alpha \wp = \left((A_i, A_j)\right)_{i,j=\overline{1,K}}. \quad (8)$$

Зауважимо, що матриця F, визначена співвідношенням (8), є матрицею Грама набору елементів (матриць) $A_1, ..., A_K$ з евклідового простору $R^{m\times n}$.

Сингулярний розклад оператора $\wp^*_\alpha \wp$ є сингулярним розкладом матриці F, що визначається співвідношенням (8).

Матриця F очевидним чином є симетричною та невід'ємно визначеною. Вона однозначно визначає ортонормований набір ненульових сингулярностей $(v_i, \lambda_i^2),\ i = \overline{1,r}$:

$$||v_i|| = 1,\ v_i \bot v_j,\ i \neq j;\ i,j = \overline{1,r};\ \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_r > 0,$$
$$\wp^*_\alpha \wp_\alpha v_i = F v_i = \lambda_i^2 v_i,\ i = \overline{1,r},\ r = rankF,$$
$$v_i^T = (v_{1i}, ..., v_{Ki}),\ i = \overline{1,r}$$

і має SVD-подання

$$\wp^*_\alpha \wp_\alpha = F = \sum_{i=1}^r \lambda_i^2 v_i v_i^T = \sum_{i=1}^r \lambda_i^2 v_i (v_i, \cdot) = \sum_{i=1}^r \lambda_i^2 v_i \ell_{v_i}. \quad (9)$$

Отже, справедлива така теорема.

Теорема 3. *SVD-подання лінійного оператора $\wp^*_\alpha \wp$ задається співвідношенням* (9).

Наступне твердження є очевидним наслідком теореми 3 та загальної теореми про SVD-подання лінійного оператора над абстрактними евклідовими просторами.

Наслідок 2. Матриці $U_i \in R^{m\times n} : U_i = \frac{1}{\lambda_i} \wp_\alpha v_i = \frac{1}{\lambda_i} \sum_{k=1}^K A_k v_{ki}$, $i = \overline{1,r}$, які визначені сингулярностями $(v_i, \lambda_i^2),\ i = \overline{1,r}$ оператора $\wp^*_\alpha \wp_\alpha$

є елементами повного набору ненульових сингулярностей $(U_i, \lambda_i^2),\ i = \overline{1,r}$ оператора $\wp_\alpha \wp_\alpha^* : R^{m\times n} \to R^{m\times n}$ [6].

Теорема 4. *(Сингулярний розклад кортежного оператора)* [6]. *Набори сингулярностей двох операторів: $\wp_\alpha^* \wp_\alpha,\ \wp_\alpha \wp_\alpha^*$ однозначно визначають сингулярне (SVD-) подання операторів $\wp_\alpha,\ \wp_\alpha^*$:*

$$\wp_\alpha x = \sum_{i=1}^{r} \lambda_i U_i v_i^T x = \sum_{i=1}^{r} \lambda_i U_i (v_i, x) = \left(\sum_{i=1}^{r} \lambda_i U_i \ell_{v_i} \right) x,\ x \in R^K, \qquad (10)$$

$$\wp_\alpha^* Y = \sum_{i=1}^{r} \lambda_i v_i (U_i, Y) = \left(\sum_{i=1}^{r} \lambda_i v_i \ell_{U_i} \right) Y, \quad Y \in R^{m\times n}, \quad r = rankF.$$

Наслідок 3. Варіантом сингулярного подання $\wp_\alpha$ є його запис у вигляді:

$$\wp_\alpha = \sum_{k=1}^{r} \lambda_k U_k v_k^T = \sum_{k=1}^{r} \left(\wp_\alpha v_k \right) \ell_{v_k},\ r = rankF. \qquad (11)$$

Наслідок 4. Набори ненульових сингулярностей $(v_i, \lambda_i^2), (U_i, \lambda_i^2)$, $i = \overline{1,r}$, операторів $\wp_\alpha^* \wp_\alpha,\ \wp_\alpha \wp_\alpha^*$ відповідно пов'язані між собою «симетричними» співвідношеннями:

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

Зауваження. Звернемо увагу, що сингулярне подання кортежного оператора: співвідношення (10) чи (11), — будується на основі розв'язання матричної задачі на власні значення. Отже, таке подання має конструктивність того самого порядку, що й у випадку SVD-подання матриць.

3.1. Псевдообернення для „кортежних операторів“

ПдО кортежного оператора визначається у відповідності до „абстрактного“ визначення ПдО: визначення за співвідношенням (3) для лінійних операторів між абстрактними евклідовими просторами.

Означення 3. (*ПдО кортежного оператора через його SVD*) [6]. ПдО: $\wp_\alpha^+,\ \wp_\alpha^{*+}$ відповідно — визначатиметься співвідношеннями:

$$\wp_\alpha^+ Y = \sum_{k=1}^{r} \lambda^{-1} v_k \left(U_k, Y \right) \equiv \left(\sum_{k=1}^{r} \lambda^{-1} v_k \ell_{U_k} \right) Y, \quad Y \in R^{m\times n},$$

$$\left(\wp_\alpha^* \right)^+ x = \sum_{k=1}^{r} \lambda_k^{-1} U_k v_k^T x = \left(\sum_{i=1}^{r} \lambda_k^{-1} U_k \ell_{v_k} \right) x, \quad x \in R^K.$$

3.2. Ортогональні проектори основних підпросторів для „кортежних операторів“

Фундаментальними підпросторами лінійних операторів є: множина можливих значень (рейндж) для самого оператора та спряженого до нього; їхні ортогональні доповнення, які є ядрами суміжних операторів. ПдО теорія за Муром-Пенроузом для матриць надає можливості конструктивного опису ортогональних проекторів згаданих підпросторів в евклідових просторах числових векторів. Ці результати повністю переносяться на кортежні оператори.

Фундаментальні підпростори кортежного оператора та спряженого до нього: відповідні пари множина можливих значень $\Re$ (рейндж) — ядро — позначатимуться відповідно: $\Re(\wp_\alpha) \equiv L_{\wp_\alpha} \subseteq R^{m\times n}$, $Ker\wp_\alpha \subseteq R^K$ та $\Re(\wp_\alpha^*) \equiv L_{\wp_\alpha^*} \subseteq R^K$, $Ker\wp_\alpha^* \subseteq R^{m\times n}$. Ортогональні проектори пари оператор - спряжений позначатимуться відповідно $P_{L_{\wp_\alpha}}, P_{L_{\wp_\alpha^*}}$. Оскільки ядро оператора є ортогональним доповненням до рейнджу спряженого, то ортогональні проектори $Z(\wp_\alpha) \equiv P_{Ker(\wp_\alpha)}$, $Z(\wp_\alpha^*) \equiv P_{Ker\wp_\alpha^*}$ ядер операторів $\wp_\alpha, \wp_\alpha^*$, відповідно є доповненнями до тотожних E_K, $E_{m\times n}$ операторів у відповідних евклідових просторах.

Лема 4. *Ортогональні проектори фундаментальних підпросторів операторів $\wp_\alpha$, $\wp_\alpha^*$ визначаються співвідношеннями $P_{L_{\wp_\alpha^*}}$, $P_{L_{\wp_a}}$*

$$P_{L_{\wp_\alpha^*}} = \wp_\alpha^{*+}\wp_\alpha, \quad P_{L_{\wp_a}} = \wp_\alpha^{+}\wp_\alpha^*, \tag{12}$$

$$Z(\wp_\alpha) \equiv E_K - \wp_\alpha^{*+}\wp_\alpha, \quad Z(\wp_\alpha^*) \equiv E_{m\times n} - \wp_\alpha^{+}\wp_\alpha^*. \tag{13}$$

Доведення. Справедливість (12) випливає з того, що

$$\wp_\alpha^{*+}\wp_\alpha = \sum_{k=1}^{r} v_k \ell_{v_k}, \wp_\alpha^{+}\wp_\alpha^* = \sum_{k=1}^{r} U_k \ell_{U_k}. \tag{14}$$

Як нескладно переконатися, співвідношення (14) визначають симетричні та ідемпотентні оператори, тобто ці оператори є ортогональними проекторами. Очевидним чином їхні рейнджі співпадають відповідно із $L(v_1, ..., v_r)$ та $L(U_1, ..., U_r)$, які у свою чергу співпадають з $\Re(\wp_\alpha^*) = L_{\wp_\alpha^*}$, $\Re(\wp_\alpha) = P_{L_{\wp_a}}$, відповідно. □

3.3. Властивості псевдообернення для „кортежних операторів“

Як зазначалося вище, ПдО кортежного оператора має властивості, що фігурують у визначенні ПдО за Муром чи за Пенроузом для матриць.

Зокрема, виконуються такі співвідношення:

1. Аксіоми переваги та ортогональних проекторів за Муром:

$$\begin{cases} \wp_\alpha^+\wp_\alpha\wp_\alpha^+ = \wp_\alpha^+, \ \ \wp_\alpha\wp_\alpha^+\wp_\alpha = \wp_\alpha \ - \ \text{"аксіома" переваги}, \\ P_{L_{\wp_\alpha}} = \wp_\alpha^+\wp_\alpha, \ \ P_{L_{\wp_\alpha^*}} = \wp_\alpha\wp_\alpha^+ - \text{"аксіома" ортогональних проекторів}. \end{cases}$$

2. „Аксіома оптимальності“: визначення псевдооберненої матриці за Пенроузом: $\wp_\alpha^+ Y = \underset{x\in Arg\min||\wp_\alpha x - Y||^2}{\arg\min} ||x||^2$, $Y \in R^{m\times n}$, $Y \neq 0$.

В роботі [7] наведено який вигляд згаданих операторів ортогонального проектування та їхнє використання у „відстанях відповідності" для розв'язання задач розпізнавання жестів дактильної мови.

3. Квадрат відстані від підпростору $L(A_1, ..., A_K)$, $A_k \in R^{m\times n}$, $k = \overline{1,K}$: $\rho^2(Y, L(A_1, ..., A_K)) = (Y, \wp_\alpha \wp_\alpha^+ Y) = \sum_{k=\overline{1,K}} (Y, U_k)^2$, $Y \in R^{m\times n}$, $\alpha = (A_1, ..., A_K)$.

4. Квадрат відстані від гіперплощини $\Gamma(M, L(A_1, ..., A_K)) \equiv \equiv M + L(A_1, ..., A_K) : M, A_k \in R^{m\times n},\ k = \overline{1,K}$: $\rho^2(Y, \Gamma(M, L(A_1, ..., A_K))) = = (Y - M, \wp_\alpha \wp_\alpha^+ Y - M) = \sum_{k=\overline{1,K}} (Y - M, U_k)^2$.

$$Y \in R^{m\times n},\ \alpha = (A_1, ..., A_K).$$

5. Множина розв'язків чи псевдорозв'язків Ω_Y СЛАР $\wp_\alpha x = Y$ задається співвідношенням

$$\Omega_Y = \wp_\alpha^+ Y + Z(\wp_\alpha) R^K. \tag{15}$$

Якщо $(Y, Z(\wp_\alpha^*)Y) = 0$, множина (6) задає множину розв'язків, якщо $(Y, Z(\wp_\alpha^*)Y) > 0$ — псевдорозв'язків. Умова $(Y, Z(\wp_\alpha^*)Y) = 0$ є необхідною та достатньою для розв'язності СЛАР.

3.4. Групувальні оператори та еліпсоїди групування

Нагадаймо, що у випадку евклідового простору числових векторів, групувальні оператори задають структури другого порядку: еліпсоїди чи мінімальні еліпсоїди групування, — та матриці квадратичних форм, що згадані еліпсоїди визначають. Під еліпсоїдами групування: простими чи мінімальними, центральними чи нецентральними, — розуміють еліпсоїди, що оптимально «накривають» вектори із заданої послідовності, зокрема — векторів навчальної вибірки. Еліпсоїди можуть мати центром початок координат: центральні, чи інший центр — нецентральні. Оптимальність еліпсоїдів полягає у тому, що їхні осі максимізують суму квадратів проекцій елементів групування на відповідні осі за послідовної їх побудови за згаданим принципом та виключення із області визначення задачі оптимізації.

Концепція кортежних операторів дозволяє повністю перенести всі твердження про групування для векторів на набори матриць. Для матричних наборів виявляються вірними всі твердження про еліпси групування в евклідових просторах R^n. Ці твердження про групування є предметом теорем, наведених нижче.

Так само, як і твердженнях про групування для евклідових просторів числових векторів основними засобами групування, оператори, які зважаючи на їх властивості, природно називати групувальними. Їх також називають зваженими проекційними.

Означення 4. Групувальними для кортежного оператора називатимемо оператори $R(\wp_\alpha)$, $R(\wp_\alpha^*)$, що визначатимуться співвідношеннями:

$$R(\wp_\alpha) = \wp_\alpha^+ \wp^{+*}, \quad R(\wp_\alpha^*) = \wp^{+*} \wp_\alpha^+.$$

Лема 5. *Кожен з двох групувальних операторів може бути поданий за елементами SVD-подання кортежного оператора* $\wp_\alpha$ *у вигляді*

$$R(\wp_\alpha^*) = \wp^{+*}\wp_\alpha^+ = \sum_{i=\overline{1,r}} \frac{U_i \ell_{U_i}}{\lambda_i^2} = \sum_{i=\overline{1,r}} \frac{U_i(U_i,\cdot)}{\lambda_i^2},$$

$$R(\wp_\alpha) = \sum_{i=\overline{1,r}} \frac{v_i \ell_{v_i}}{\lambda_i^2} = \sum_{i=\overline{1,r}} \frac{v_i v_i^T}{\lambda_i^2}.$$

Лема 6. *Сума квадратів довжин проекцій елементів кортежу* $\alpha = (A_1 \vdots \ldots \vdots A_K)$, $A_k \in R^{m\times n}$, $k = \overline{1,K}$ *на одновимірний підпростір, що задаеться нормованою матрицею* $U : U \in R^{m\times n} : ||U|| = 1$ *— визначаеться співвідношенням:*

$$\sum_{k=1}^{r} ||P_{L(U)} A_k||^2 = (U, \wp_\alpha \wp_\alpha^* U)\,. \tag{16}$$

Доведення. Дійсно,

$$P_{L(\mathrm{U})} A_k = (U, A_k) U \Rightarrow ||P_{L(\mathrm{U})} A_k||^2 = |(U, A_k)|^2,$$

а, отже, $\sum_{k=1}^{K} ||P_{L(\mathrm{U})} A_k||^2 = \sum_{k=1}^{K} |(U, A_k)|^2 = \left\| \begin{pmatrix} (U, A_1) \\ \cdots \\ (U, A_K) \end{pmatrix} \right\|^2 = ||\wp_\alpha^* U||^2 =$
$= (\wp_\alpha^* U, \wp_\alpha^* U) = (U, \wp_\alpha \wp_\alpha^* U)\,.$ □

Лема 7. Одновимірні підпростори, породжені векторами ненульового набору сингулярностей оператора $\wp_\alpha \wp_\alpha^*$ є розв'язками оптимізаційних задач про максимізацію суми квадратів проекцій на одновимірні підпростори, породжені нормованими матрицями:

$$\begin{gathered} U_1 = \underset{U \in R^{m\times n}: ||U||=1}{\arg\max} \sum_{k=1}^{K} ||P_{L(\mathrm{U})} A_k||^2, \\ \max_{U \in R^{m\times n}: ||U||=1} \sum_{k=1}^{K} ||P_{L(\mathrm{U})} A_k||^2 = \lambda_1^2, \end{gathered} \tag{17}$$

$$\begin{gathered} U_i = \underset{U \in R^{m\times n}: ||U||=1, U \perp U_1, U \perp U_{i-1},}{\arg\max} \sum_{k=1}^{K} ||P_{L(\mathrm{U})} A_k||^2,\ i = \overline{1,r},\ r = rankF, \\ \max_{U \in R^{m\times n}: ||U||=1, U \perp U_1, U \perp U_{i-1}} \sum_{k=1}^{K} ||P_{L(\mathrm{U})} A_k||^2 = \lambda_i^2,\ i = \overline{1,r},\ r = rankF. \end{gathered} \tag{18}$$

Доведення. Те, що на $U = U_i$, $i = \overline{1,r}$ значення суми квадратів проекцій визначається відповідно значеннями λ_i^2, $i = \overline{1,r}$ перевіряється безпосередньою підстановкою. Оптимальність випливає із розгляду розкладу будь якого $U : U \in R^{m\times n} : ||U|| = 1$ за базисом в $R^{m\times n}$, побудованим розширенням U_i, $i = \overline{1,r}$. □

Теорема 5. *Нехай* $A_k \in R^{m\times n}$, $k = \overline{1,K}$ *довільний набір матриць із* $R^{m\times n}$, $\alpha = (A_1, ..., A_K)$ *— матричний кортеж, а* $\wp_\alpha$ *— кортежний оператор, який йому відповідає* [2]. *Тоді всі матриці набору* $A_k \in R^{m\times n}$, $k = \overline{1,K}$ *належать внутрішності еліпса, точніше: еліпсоїдального циліндра, який визначається рівнянням*

$$\sum_{i=1}^{r} \frac{(U_i, X)^2}{(\lambda_i \sqrt{r})^2} \leq 1, \tag{19}$$

чи, що еквівалентно

$$\left(X, r^{-1} R(\wp_\alpha^*) X\right) = 1,\ r = rankF,\ X \in R^{m\times n}, \tag{20}$$

де $R(\wp_\alpha^*)$*– групувальний оператор* $R(\wp_\alpha^*) = \wp^{+*}\wp_\alpha^+$.

Доведення. З (16)–(18) випливає, що

$$||P_{L(\mathrm{U}_i)} A_k||^2 \leq \lambda_i^2, \quad i = \overline{1,r},\ r = rankF,\ k = \overline{1,K},$$

тобто

$$||P_{L(U_i)} A_k||^2 = (U_i A_k)^2 \leq \lambda_i^2,\ i = \overline{1,r},\ k = \overline{1,K},$$

що еквівалентно

$$\frac{(U_i A_k)^2}{\lambda_i^2} \leq 1,\ i = \overline{1,r},\ k = \overline{1,K}.$$

Сумуючи в останньому співвідношенні по i для кожного із фіксованих значень $k = \overline{1,K}$, маємо

$$\sum_{i=1}^{r} \frac{(U_i A_k)^2}{\lambda_i^2} \leq r.$$

чи

$$\sum_{i=1}^{r} \frac{(U_i A_k)^2}{(\lambda_i \sqrt{r})^2} \leq 1,\ k = \overline{1,K}. \tag{21}$$

Зазначимо, що

$$sum_{i=1}^{r} \frac{(U_i X)^2}{r\lambda_i^2} = \left(X, r^{-1} R(\wp_\alpha^*) X\right),$$

а, отже

$$\sum_{i=1}^{r} \frac{(U_i X)^2}{r\lambda_i^2} \leq 1 \Leftrightarrow \left(X, r^{-1} R(\wp_\alpha^*) X\right) \leq 1. \tag{22}$$

Оскільки нерівності в (22) визначають еліпсоїдальний циліндр з осями $U_i \in R^{m\times n}$, $i = \overline{1,r}$ та довжинами напівосей, відповідно, $\sqrt{r}\lambda_i$, $i = \overline{1,r}$, то виконання A_k, $k = \overline{1,K}$ співвідношень (22) означає, що всі вони належать внутрішності згаданих еліпсоїдальних циліндрів, і доведення теореми завершено. □

Будемо називати еліпсоїд, що визначається нерівністю (19) чи (20) центральним еліпсоїдом групування.

Важливими у задачах групування матриць є також еліпсоїди, які позначаються терміном „нецентральні еліпсоїди групування“: ті що мають центром середнє $\bar{A}$ елементів набору A_k, $k = \overline{1,K}$, а кортежний оператор визначається за кортежем $\tilde{\alpha} = (\tilde{A}_1, ..., \tilde{A}_K)$, $\tilde{A}_k = A_k - \bar{A}$, $k = \overline{1,K}$.

Означення 5. Нецентральними еліпсоїдом групування для кортежного оператора $\wp_\alpha$: $\alpha = (A_1, ..., A_K)$ — називатимемо еліпсоїд , що визначається співвідношенням

$$\sum_{i=1}^{r} \frac{(\tilde{U}_i, X - \bar{A})^2}{\left(\tilde{\lambda}_i \sqrt{r}\right)^2} \leq 1 \Leftrightarrow \left(X - \bar{A}, r^{-1} R(\wp_{\tilde{\alpha}}^*)(X - \bar{A})\right) \leq 1,\ X \in R^{m\times n}. \quad (23)$$

Теорема 6. *Всі матриці-елементи матричного кортежу* $\alpha = (A_1, \ldots, A_K)$ *належать нецентральному еліпсоїду групування* (23) *з* $(\tilde{U}_i, \tilde{\lambda}_i^2)$, $i = \overline{1, r}$, *що є ненульовими сингулярностями кортежного оператора* $\wp_{\tilde{\alpha}}$, $\tilde{\alpha} = (\tilde{A}_1, ..., \tilde{A}_K)$.

Доведення. Доведення безпосередньо випливає з теореми 5, оскільки для елементів кортежу $\tilde{\alpha}$ та кортежного оператора $\wp_{\tilde{\alpha}}$ виконується співвідношення

$$\sum_{i=1}^{r} \frac{(\tilde{U}_i, \tilde{A}_k)^2}{\left(\tilde{\lambda}_i \sqrt{r}\right)^2} \leq 1, k = \overline{1, r},$$

чи

$$\sum_{i=1}^{r} \frac{(\tilde{U}_i, A_k - \bar{A})^2}{\left(\tilde{\lambda}_i \sqrt{r}\right)^2} \leq 1,\ k = \overline{1, r},$$

що й доводить теорему. □

Зауваження. Наведені вище «групувальні теореми» можна підсилити, підставивши замість r число: $r_{\min} \leq r$.

3.5. Мінімальні еліпсоїди групування для „кортежних операторів“

Визначимо $r_{\min} \leq r$, $\tilde{r}_{\min} \leq r$ відповідно співвідношеннями

$$r_{\min} = \max_{k=\overline{1,K}} (A_k, R(\wp_\alpha^*) A_k), \quad \tilde{r}_{\min} = \max_{k=\overline{1,K}} (\tilde{A}_k, R(\wp_{\tilde{\alpha}}^*) \tilde{A}_k). \quad (24)$$

Теорема 7. *Всі матриці набору* $A_k \in R^{m\times n}$, $k = \overline{1, K}$ *належать внутрішності еліпсів, точніше еліпсоїдальних циліндрів — центральних чи нецентральних, — що задаються співвідношеннями, відповідно* [2]*:*

$$\begin{gathered} \left(X, r_{\min}^{-1} R(\wp_\alpha^*) X\right) \leq 1, \\ \left(X - \bar{A}, \tilde{r}_{\min}^{-1} R(\wp_{\tilde{\alpha}}^*)(X - \bar{A})\right) \leq 1. \end{gathered} \quad (25)$$

Означення 6. Еліпсоїди групування, що визначаються співвідношеннями (23) називатимемо мінімальними еліпсами групування.

Висновок

Оперування із базовими лінійними та нелінійними структурами в матричних евклідових просторах $R^{m\times n}$ потребує розвинення адекватних конструктивних засобів такого оперування. Практика використання згаданих структур в R^n значною мірою базується на SVD-поданні матриць та псевдооберненні (ПдО) матриць за Муром-Пенроузом, тому побудова засобів оперування в $R^{m \times n}$ потребує слушних відповідників SVD та ПдО. Таким відповідниками є SVD та ПдО твердження для слушних лінійних операторів, якими, як продемонстровано вище, є „кортежні оператори“ за рядковими кортежами матриць. Концепція розвитку засобів оперування із базовими структурами для матричних евклідових просторах може бути позначена терміном „концепція кортежності“. Для описаного в роботі класу кортежних операторів побудоване SVD-подання, що дозволило побудувати на основі такого подання теорію псевдообернення з конструктивними можливостями отримання формульних виразів як для ортогональних проекторів базових підпросторів оператора, так і для групувальних операторів. Згадана вище конструктивність означає, що побудова згаданих об'єктів зводиться до розв'язання задачі на власні значення для матриць операторів в R^n. Наявність засобів оперування з базовими структурами в матричних просторах суттєво розширює можливості використання таких просторів в математичному моделюванні, оскільки в багатьох важливих в прикладному відношенні задачах математичного моделювання саме матриці з'являються як представники аналізованих об'єктів.

Література

1. Kirichenko N. F. Analytical representation of perturbation of pseudoinverse matrices // Cybernetics and system analysis. — Vol. 33, № 2. –March-April, 1997. — P. 230–239.
2. Donchenko V., Zinko T., Skotarenko F. „Feature vectors“in grouping information problem in applied mathematics: vectors and matrixes // Problems of computer Intellectualization. ITHEA, Kiyv, Ukraine-Sofia, Bulgaria, 2012. — P. 111–124.
3. 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. — P. 155–170.
4. Кириченко Н. Ф. Аналитическое представление возмущений псевдообратных матриц // Кибернетика и системный анализ. — К., 1997. № 2. — С. 98–107.
5. Donchenko V., Nazaraga I., Tarasova O. Matrixes least squares method and examples of its application // International Journal Information Technologies & Knowledge, 2013. — Vol. 7, № 4. — P. 325–336.
6. Донченко В. С. Евклидовы пространства: конструктивные методы описания базовых структур и их использование // Information Models of Knowledge. — Editors: Krassimir Markov, Vitalii Velichko, Oleksy Voloshin — ITHEA. — Kiev, Ukraine — Sofia, Bulgaria. Number 19, 2010. — P. 362–376.
7. Голік А. О., Донченко В. С. Застосування еліпсоїдальної та ортогональної відстаней відповідності у задачах розпізнавання мови та жестів. // Вісник Київського національного університету імені Тараса Шевченка. Серія: фіз.-мат. науки, — 2013, Вип. 1. — С. 146–155.
8. Kirichenko N. F., Donchenko V. S., Serbaev D. P. Nonlinear recursive nonlinear transformation: dynamic systems and optimization // Cybernetics and system analysis. — Vol. 41, № 3. May 2005. — P. 364–373.