

УДК 519.685.3

Кудін Г.І., к.ф.м.н.
Зінько Т.П., к.т.н.

Синтез кусково гіперплощинних кластерів в просторах векторів ознак

Київський національний університет
імені Тараса Шевченка,
м. Київ, пр-т. Глушкова, 4-д
e-mail: gkudin@ukr.net.ua

Kudin G.I., Ph.D. (Phys.-Math.)
Zinko T.P., Ph.D. (Tech.)

A synthesis of piecewise hyperplane clusters in the spaces of vectors of signs

Taras Shevchenko National University of Kyiv,
Glushkova ave., 4-d, Kyiv
e-mail: gkudin@ukr.net.ua

Досліджується задача побудови засобами псевдо обернення кусково гіперплощинних кластерів для навчальних виборок векторів у просторі ознак, а саме: поділ вибірки на два кусково гіперплощинних кластери так, щоб певна підмножина точок була в околі ланок першого кластеру, а точки іншої підмножини - в околі ланок другого кластеру. Особливість постановки: заздалегідь не відомі ні підмножини, ні складові обох сукупностей гіперплощин.

Ключові слова: Кластеризація інформації, гіперплощини, псевдо обернення матриць, проєкційні оператори.

The algorithms for the construction of a piecewise hyperplane cluster by pseudo-inverse matrices are proposed. Vectors in the space of signs are divided into two pieces of hyperplane clusters such that a certain subset of points is in the vicinity of the first cluster, and the points of another subset - in the vicinity of the second cluster. Feature of the algorithm: no subset, or components of the hyperplane of both clusters are known in advance.

The use of hyperplanes for which the ratio for distances between vectors and hyperplanes is obtained, between sets of vectors and hyperplanes in the spaces of signs, in conjunction with the results of the perturbation theory of pseudo-inverse, projective and R-matrices provides the construction of effective algorithms for solving the problem. The description of the hyperplanes is carried out through solutions (pseudo-solutions) of systems of linear algebraic equations, which are determined by finding the optimal elements of them - the basic matrix, the vector of the right-hand sides.

The article presents the original scientific results of the theory of pseudo-inversion of matrices obtained by various authors over the past ten years.

Keywords: Clusterization of information, hyperplane, pseudo rotations of matrices, projection operators.

Статтю представив д.т.н., проф. Гаращенко Ф.Г.

Вступ

Матрична структура представлення множин векторів ознак в відповідному евклідовому просторі, якими описуються досліджувані об'єкти при розв'язанні різноманітних прикладних задач цифрової обробки інформації, обумовлює широкий спектр можливостей розвиненню концепції псевдо обернення матриць для створення математичних методів кластеризації та класифікації інформації [1]-[4]. Використання елементарних лінійних структур – гіперплощин – з простими співвідношеннями для віддалей між векторами і гіперплощинами, між множинами векторів і гіперплощинами в

просторах ознак в поєднанні з результатами теорії збурення псевдо обернених, проєкційних та R-матриць [1], [2] забезпечує не тільки побудову ефективних алгоритмів розв'язання задач, а також можливості їх оптимізації в рамках конкретних їх постановок. З роботами [3], [4] пов'язується варіант гіперплощинної кластеризації, в рамках якого опис гіперплощин здійснюється через гіперплощини, що є розв'язками (псевдо розв'язками) систем лінійних алгебраїчних рівнянь (в цьому випадку виникає проблема знаходження оптимальних елементів таких систем лінійних алгебраїчних рівнянь – основної матриці, вектора правих

частин, для яких подаються алгоритми їх обчислення). В роботі [2] параметри гіперплощин знаходяться рекурентно згідно функціонала прийнятності наявного розбиття. Функціонали прийнятності - це той чи інший варіант відстані елементу кластера до його гіперплощини, відстань елементу кластера до гіперплощини іншого кластеру. В якості відстаней використовувались евклідові відстані, а в якості міри компактності гіперплощинного кластера – дисперсія елементів кластера відносно своєї гіперплощини. В якості ефективного розв'язання проблеми мінімізації дисперсії елементів кластеру може бути використано варіант гіперплощинного кластеру, який би описувався не одною гіперплощиною, а декількома, тобто будується кусково гіперплощинний кластер. Ця ідея кусково гіперплощинного кластеру витікає з запропонованого підходу розв'язання проблеми кусково лінійної віддільності множини точок в просторі векторів ознак від початку координат, запропонованого в 2007 році професором Кириченком М.Ф. [1]. Викладення результатів по темі кусково гіперплощинної кластеризації інформації подається після короткого опису використовуваних в подальшому математичних результатів стосовно визначення псевдо обернення, формул обчислення псевдо обернених, проєкційних та R-матриць при зміні розмірності вихідних матриць. Приводяться різні варіанти відстаней між елементами векторних просторів.

Основні математичні засоби псевдо обернення, його властивості

В подальшому крім стандартного позначення $m \times n$ матриці за її елементами: $A = (a_{ij})$, використовуватиметься її зображення за стовпчиками чи рядками, відповідно:

$$A = (a(1) : \dots : a(n)), \quad a(j) \in R^n, \quad j = \overline{1, n},$$

$$A = (a_{(1)} \quad \dots \quad a_{(m)})^T, \quad a_{(i)}^T \in R^n, \quad i = \overline{1, m}.$$

Традиційно стандартні базиси в R^n, R^m , складені із координатних ортів позначатимуться відповідно

$$e_{(i)} \in R^n, \quad i = \overline{1, n}, \quad e_{(j)} \in R^m, \quad j = \overline{1, m},$$

$$e_{(i)}^T = (0, \dots, 0, 1, 0, \dots, 0), \quad i = \overline{1, n},$$

$$e^T(j) = (0, \dots, 0, 1, 0, \dots, 0), \quad j = \overline{1, m}.$$

Сингулярний розклад матриці. Будь яка $m \times n$ матриця A рангу $r \leq \min(m, n)$ може бути

записана у вигляді $A = \sum_{i=1}^r \lambda_i u_i v_i^T$, в якому:

$\lambda_1^2 \geq \dots \geq \lambda_r^2 > 0$ спільний набір ненульових власних чисел матриць $AA^T, A^T A$; $v_i \in R^n, i = \overline{1, r}$ - ортонормований набір власних векторів матриці $A^T A$, що відповідають ненульовим власним числам:

$$A^T A v_i = \lambda_i^2 v_i, \quad \lambda_i^2 > 0, \quad i = \overline{1, r}, \quad v_i^T v_j = \delta_{ij}, \quad i \neq j;$$

$u_i \in R^m, i = \overline{1, r}$ - ортонормований набір власних векторів матриці AA^T , що відповідають ненульовим власним числам $\lambda_i^2, i = \overline{1, r}$:

$$AA^T u_i = \lambda_i^2 u_i, \quad i = \overline{1, r}, \quad u_i^T u_j = \delta_{ij}.$$

Визначення та обчислення псевдо оберненої матриці допускає декілька еквівалентних варіантів. Тут використовується варіант Penrose'a - визначення псевдооберненої матриці здійснюється через систему лінійних алгебраїчних рівнянь (СЛАР): $Ax = b$, $x \in R^n, b \in R^m, b \neq 0$. За таких умов псевдо обернена матриця при довільному ненульовому векторі $b \in R^m$ визначається співвідношенням

$$A^+ b = \arg \min_{x \in \text{Arg} \min_x \|Ax - b\|^2} \|x\|, \quad \forall b \in R^m, b \neq 0.$$

Підпростори, ортогональні проєктори. Підпросторами, які розглядаються у зв'язку матрицею $A \in$ множини можливих значень L_A, L_{A^T} матриць A, A^T : відповідно, а ортогональними проєкторами на них - це P -проєктори, позначаються $P(A^T), P(A)$ відповідно, ортогональні до них - називаються Z -проєкторами і позначаються $Z(A), Z(A^T)$ відповідно.

Крім P - та Z - проєкторів, пов'язаних з базисними підпросторами матриці A , важливе застосування мають також оператори $R(A), R(A^T)$.

Зв'язок між P - та Z - проєкторами визначається співвідношеннями:

$$P(A) = A^+ A, \quad P(A^T) = (A^T)^+ A^T = AA^+,$$

$$Z(A) = I_n - A^+ A, \quad Z(A^T) = I_m - AA^+,$$

Звести обчислення псевдо оберненої матриці для довільної матриці до обчислення оберненої відповідної їй деякої квадратної матриці з розмірністю $\min(m, n)$ дозволяють важливі для практичних застосувань такі співвідношення:

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

Прямі формули Гревеля. Важливе місце у застосуваннях, пов'язаних із застосуванням псевдо обернення, посідають формули Гревеля: прямі та обернені.

В прямому варіанті формули Гревеля – це співвідношення, які пов'язують псевдо обернену матрицю, розширену введенням додаткового рядка чи стовпчика, з псевдо оберненою вихідної матриці. Обернені формули – це співвідношення, в яких псевдо обернена матриця для вихідної матриці обчислюється з використанням псевдо оберненої матриці матриці, що отримана з вихідної матриці викреслюванням рядка чи стовпчика.

Припустимо, що розширення матриці A відбувається появою нового рядка $a^T \in R^n$ після $(i-1)$ -го рядка ($i = \overline{2, m+1}$), тобто утворюється матриця

$$A(i, a) = (a_{(1)} : \dots : a_{(i-1)} : a : a_{(i)} : \dots : a_{(m)})^T \in R^{(m+1) \times n}.$$

Тоді при відомій псевдо оберненій матриці $A^+ \in R^{n \times m}$ для рекурентного обчислення псевдо оберненої матриці $A^+(i, a) \in R^{n \times (m+1)}$ мають місце співвідношення - *прямі формули Гревеля* [1], [2].

Якщо шукану псевдо обернену матрицю $A^+(i, a)$ подати в вигляді

$$A^+(i, a) = (p(1) : \dots : p(i-1) : p(i) : p(i+1) : \dots : p(m+1)) \in R^{n \times (m+1)},$$

де $P_i = (p(1) : \dots : p(i-1) : p(i+1) : \dots : p(m+1))$, тобто вважати, що матриця $A^+(i, a)$ може бути отримана з матриці $P_i \in R^{n \times m}$ вставленням після $(i-1)$ -го стовпчика вектора $p(i)$, то для матриці P_i справедлива формула:

$$P_i = (1 - p(i)a^T)A^+.$$

При цьому невідомий вектор $p(i)$ визначається з врахуванням властивості лінійної залежності вектора a від векторів підпростору вектор – рядків матриці A :

1) якщо вектор a лінійно незалежний від векторів підпростору вектор – рядків матриці A , тобто $a^T Z(A)a > 0$, то

$$p(i) = Z(A)a \|Z(A)a\|^{-2};$$

2) якщо вектор a лінійно залежний від векторів підпростору вектор – рядків матриці A , тобто $a^T Z(A)a = 0$, то

$$p(i) = R(A)a(1 + a^T R(A)a)^{-1}.$$

Наслідок. Операція псевдо обернення комутативна з транспонуванням, тому прямі формули Гревеля аналогічно виписуються для варіанту розширення матриці стовпчиком.

Обернені формули Гревеля. Якщо для матриці $A(i, a) \in R^{(m+1) \times n}$ (після $(i-1)$ -го рядка ($i = \overline{2, m+1}$) матриці A вставлено вектор-рядок a^T), відома її псевдо обернена матриця $A^+(i, a) \in R^{n \times (m+1)}$, то для визначення псевдо оберненої матриці $A^+ \in R^{n \times m}$ (рядок a^T з матриці $A(i, a)$ видаляється) справедливі *обернені формули Гревеля* [1], [2]:

1) для випадку лінійної залежності вектор – рядка a^T , який видаляється, від векторів підпростору вектор – рядків матриці $A(i, a)$, а це визначається виконанням умови $a^T p(i) < 1$, то псевдо обернена матриця $A^+ \in R^{n \times m}$ має вид

$$A^+ = (I_n + p(i)a^T(1 - a^T p(i))^{-1})P_i.$$

При цьому ранг псевдо оберненої матриці не змінюється, тобто $\text{rank } A = \text{rank}(A^T : a)^T$.

2) для випадку лінійної незалежності вектор – рядка a^T , який видаляється, від векторів підпростору вектор – рядків матриці $A(i, a)$, а це визначається виконанням умови $a^T p(i) = 1$, матриця $A^+ \in R^{n \times m}$ визначається формулою $A^+ = (I_n - p(i)p^T(i) \|p(i)\|^{-2})P_i$,

При цьому ранг псевдо оберненої матриці падає, тобто: $\text{rank } A = \text{rank}(A^T : a)^T - 1$.

Використовуючи обернені формули Гревеля для представлення псевдо оберненої матриці $A^+ \in R^{n \times m}$ при відомій псевдо оберненій матриці $A^+(i, a) \in R^{n \times (m+1)}$ (з матриці $A(i, a) \in R^{(m+1) \times n}$ видаляється i -тий ($i = \overline{2, m+1}$) вектор – рядок

a^T), можна отримати корисні для застосувань формули економного обчислення збурених проєкційних матриць $Z(A)$, $Z(A^T)$, $R(A)$, $R(A^T)$.

1) для випадку лінійної залежності вектор – рядка a^T , який видаляється, від векторів підпростору вектор – рядків матриці $A(i, a)$, а це визначається виконанням умови $a^T p(i) < 1$, то

$$Z(A) = Z(A(i, a));$$

$$R(A) = R(A(i, a)) + p(i) p^T(i) Z_{ii}^{-1}(A^T(i, a));$$

$$Z(A^T) = E_i Z(A^T(i, a)) - E_i h_i h_i^T E_i^T Z_{ii}^{-1}(A^T(i, a)),$$

Тут введена в розгляд матриця E_i - вона отримана вилученням i -го ($i = \overline{2, m+1}$) вектор-рядка з одиничної матриці I_{m+1} :

$$E_i = \begin{pmatrix} e_1(m+1) : \dots : e_{i-1}(m+1) : \\ : e_{i+1}(m+1) : \dots : e_{m+1}(m+1) \end{pmatrix}^T \in R^{m \times (m+1)}$$

$$h_i = -A_{i,a}^+{}^T a;$$

$$R(A^T) = E_i R(A_{i,a}^T) E_i^T + E_i (h_i h_i^T (p^T(i) p(i)) + h_i t_i^T + t_i h_i^T) E_i^T,$$

$$t_i = -A^+{}^T(i, a) p(i) Z_{ii}^{-1}(A^T(i, a));$$

2) для випадку лінійної незалежності вектор – рядка a^T , який видаляється, від векторів підпростору вектор – рядків матриці $A(i, a)$, а це визначається виконанням умови $a^T p(i) = 1$, то

$$Z(A) = Z(A(i, a)) + p(i) p^T(i) \|p(i)\|^{-2},$$

$$R(A) = Z(p^T(i)) R(A(i, a)) Z(p^T(i)),$$

$$Z(p^T(i)) = I_n - p(i) p^T(i) \|p(i)\|^{-2};$$

Використовуючи позначення матриці E_i (формула вище)

$$Z(A^T) = E_i Z(A^T(i, a)) E_i^T,$$

$$R(A^T) = E_i R(A^T(i, a)) E_i^T.$$

Математичні засоби кусково гіперплощинної кластеризації інформації

Базовий алгоритм поділу вибірки на два гіперплощинних кластери [2].

Розглядається задача гіперплощинної кластеризації, тобто задача представлення

наявної навчальної вибірки множини векторів $x(j) \in R^m$, $n = \overline{1, n}$ у вигляді двох частин - підмножин (кластерів)- на основі припущення, що структура цих підмножин (кластерів) в векторному просторі ознак може бути наближена гіперплощинами. Алгоритм полягає в послідовному виконанні рекурентних кроків, на кожному з яких здійснюється корегування параметрів гіперплощин $L^{(k)}(1)$, $L^{(k)}(2)$ (без обмежень загальності розглядається задача поділу навчальної вибірки на два кластери), побудованих на попередньому кроці за умов вихідного розбиття довільним способом. Це корегування параметрів гіперплощин полягає у вилученні із сукупності векторів, що породжують гіперплощини, тих, що мають найгірші значення відстаней відповідності. Вилучені вектори зараховується у альтернативний кластер. За новим розбиттям векторів будується нова пара гіперплощин $L^{(k+1)}(1)$, $L^{(k+1)}(2)$, за якими здійснюється новий рекурентний крок.

Математична постановка задачі забезпечується формальним введенням поняття відстані $\rho(x(j), L(A, b))$ від точки $x(j) \in R^m$ до деякої гіперплощини

$$L(A, b) = \{x \in R^n : x = A^+ y + Z(A) z, z \in R^n\},$$

яка визначається як множина розв'язків (псевдо розв'язків) системи алгебраїчних рівнянь

$$Ax = b, A \in R^{s \times m}, b \in R^s,$$

де A і b є відповідно матричним і векторним параметрами розглянутої гіперплощини.

Базові формули гіперплощинної кластеризації інформації [3], [4]. Відстань $\rho(x(j), L(A, b))$ від точки $x(j)$ до гіперплощини $L(A, b)$ визначається відповідно до співвідношення

$$\rho^2(x(j), L(A, b)) = \|A^+(b - Ax(j))\|^2, \|x\|^2 = x^T x.$$

Сума квадратів відстаней множини точок $x(j)$, $j = \overline{1, n}$ до гіперплощини $L(A, b)$

$$\rho^2(\{x : x(j), j = \overline{1, n}\}, L(A, b)) =$$

$$\sum_{j=1}^n (b - Ax(j))^T R(A^T) (b - Ax(j)) =$$

$$= \text{tr} R(A^T) \sum_{j=1}^n (b - Ax(j))(b - Ax(j))^T.$$

При заданих значеннях $A, x(j), j = \overline{1, n}$ має місце співвідношення для оптимального значення вектора правої частини системи алгебраїчних рівнянь, визначальної для гіперплощини

$$b_{opt} = A\hat{x}, \quad \hat{x} = \frac{1}{n} \sum_{j=1}^n x(j).$$

Відстань $\rho(\{x: x(j), j = \overline{1, n}\}, L(A, b))$ множини точок $x(j), j = \overline{1, n}$ до гіперплощини $L(A, b)$ при оптимальному виборі вектора b має наступне значення

$$\rho^2 \left(\text{tr } A^+ A \tilde{X} \tilde{X}^T \right)^{\frac{1}{2}},$$

де

$$\tilde{X} = (\tilde{x}(1) \dots \tilde{x}(n)), \quad \tilde{x}(j) = x(j) - \hat{x}, \quad j = \overline{1, n}.$$

Оптимальний вибір матриці $A \in R^{s \times m}$ визначається матрицею

$$A_{opt} = (u_{m-s+1} \dots u_m),$$

при цьому $\text{tr } A_{opt}^+ A_{opt} \tilde{X} \tilde{X}^T = \sum_{j=m-s+1}^r \lambda_j^2,$

$$(u_1, \dots, u_m)^T (u_1, \dots, u_m) = I_m.$$

Алгоритм складається з наступних кроків.

1. На першому кроці: первісна сукупність з n векторів довільним чином поділяється на дві частини $x(i_j(k)), j = \overline{1, n_k}, k = 1, 2$:

$$\begin{aligned} \{i_1(1), \dots, i_{n_1}(1)\} \cap \{i_1(2), \dots, i_{n_2}(2)\} &= \emptyset, \\ \{i_1(1), \dots, i_{n_1}(1)\} \cup \{i_1(2), \dots, i_{n_2}(2)\} &= \{1, 2, \dots, n\}. \end{aligned}$$

2. На другому кроці за кожною з наявних частин розбиття будуються гіперплощини $L^{(0)}(A_k, b_k), k = 1, 2$. Визначаються $\hat{x}^{(0)}(k), k = 1, 2$ як середні за елементами відповідної сукупності, вводяться матриці

$$\tilde{X}^{(0)}(k) = (\tilde{x}(i_1(k)) \dots \tilde{x}(i_{n_k}(k))), \quad k = 1, 2,$$

будується їхній SVD-розклад:

$$\tilde{X}^{(0)}(k) = \sum_{i=1}^{r_k} u_i(k) \lambda_i(k) v_i^T(k), \quad k = 1, 2, \quad \text{де}$$

r_1, r_2 – ранги відповідних матриць, обчислюються оптимальні параметри гіперплощин первинного розбиття.

3. На третьому кроці проводиться корегування параметрів гіперплощин первинного розбиття. Воно здійснюється за відстанями

кожного з векторів до побудованих гіперплощин та вилученням найгірших - застосування засобів псевдо обернення забезпечує обчислення таких відстаней явними формулами. В результаті корегування розбиття деякі вектори залишаються серед векторів розбиття, за якими побудовані гіперплощини, інші – зараховується в альтернативний кластер.

4. Оновлені елементи розбиття є основою нового рекурентного кроку: алгоритм повторює свої кроки, починаючи з другого.

5. Зупинка алгоритму може здійснюватись після виконання певної кількості кроків, чи після того, як досягається не поліпшуваний певний рівень відстані кожного з векторів відповідного елемента розбиття до відповідної гіперплощини.

Задача синтезу двох кусково гіпер-площинних кластерів, алгоритм її розв'язування

Під задачею синтезу кусково гіперплощинних кластерів для навчальної вибірки $x(j) \in R^m, j = \overline{1, n}$ векторів у просторі ознак розуміється задача поділу цієї вибірки на два кластери таким чином, щоб точки першого кластера були розташовані ближче до деякої сукупності гіперплощин, яка породжується цією підмножиною, а точки іншої підмножини - другого кластера – ближче до іншої сукупності гіперплощин. Особливістю постановки розглянутої задачі є та обставина, що *заздалегідь не відомі ні підмножини, ні складові обох сукупностей гіперплощин*.

Отже, на початку процесу кусково гіперплощинної кластеризації припускається, без обмеження загальності щодо кількості підмножин, що вектори $x(1), \dots, x(n)$ з простору ознак R^m можуть належати одній з двох гіперплощин $L^{(0)}(A_k, b_k), k = 1, 2$ фіксованої розмірності s ($s < m$). Якщо в процесі рекурентного поліпшення гіперплощинної кластеризації виявляється, що отримані кластери не відповідають раніше встановлених критеріїв (наприклад, незадовільна гіперплощинна концентрація векторів кластеру, тощо), то виникає потреба вплинути на форму кластеру і модифікувати її стає можливим побудовою кусково лінійних кластерів. Для розглядуваної задачі побудови кусково гіперплощинного кластеру це означає доповнення кожної з гіперплощин $L^{(0)}(A_k, b_k), k = 1, 2$, які попередньо визначені в рамках базової гіперплощинної кластеризації, додатковими ланками – гіперплощинами.

Алгоритм синтезу кусково гіперплощинних кластерів. На підставі вище викладеної ідеї синтезу кусково гіперплощинних кластерів алгоритм такого синтезу можна представити так:

1. *Формування одноланкового кластеру* (номер ланки в кластері – індекс в дужках, номер множини - індекс знизу):

1) Створити початкове розбиття вихідної множини $\Omega = \{x: x(j), j = \overline{1, n}\}$ на дві підмножини $\Omega \equiv \Omega^0(1) = \Omega_1^0(1) \cup \Omega_2^0(1), \Omega_1^0(1) \cap \Omega_2^0(1) = \emptyset$ (доцільно засобом k -means).

2) Побудувати гіперплощини

$$A_1(1)x = b_1(1), L(A_1(1), b_1(1)) \text{ для } \Omega_1^0(1),$$

$$A_2(1)x = b_2(1), L(A_2(1), b_2(1)) \text{ для } \Omega_2^0(1).$$

Обчислити оптимальні $A_{1opt}(1), b_{1opt}(1)$ для $\Omega_1^0(1)$ і $A_{2opt}(1), b_{2opt}(1)$ для $\Omega_2^0(1)$.

3) Обновити $\Omega_1^0(1)$ і $\Omega_2^0(1)$ по відповідних відстанях

$$(b_{opt} - A_{opt}x(j))^T R(A_{opt}^T)(b_{opt} - A_{opt}x(j)) -$$

сформувані нові множини $\Omega_1^1(1), \Omega_2^1(1)$ із точок $\Omega^0(1)$, для яких виконуються відповідно умови, а саме:

$x(j) \in \Omega_1^0(1)$, якщо

$$\begin{aligned} & \|A_{1opt}^+(1)(b_{1opt}(1) - A_{1opt}(1)x(j))\|^2 \leq \\ & \leq \|A_{2opt}^+(1)(b_{2opt}(1) - A_{2opt}(1)x(j))\|^2, \end{aligned}$$

$x(j) \in \Omega_2^0(1)$, якщо

$$\begin{aligned} & \|A_{1opt}^+(1)(b_{1opt}(1) - A_{1opt}(1)x(j))\|^2 > \\ & > \|A_{2opt}^+(1)(b_{2opt}(1) - A_{2opt}(1)x(j))\|^2. \end{aligned}$$

Лінійна залежність чи незалежність векторів, що можуть бути вилучені з множин, дає можливість спростити вигляд формул для відстаней цих векторів від відповідних гіперплощин, бо доцільно використовувати співвідношення, які отримані у вигляді обернених формул Гревлія.

4) Перейти на виконання операції 2 з новими підмножинами $\Omega_1^i(1), \Omega_2^i(1), i = 1, 2$. Верхній індекс позначає кількість ітерацій на етапі першої ланки.

5) Зупинка алгоритму на етапі першої ланки кластеру може завершитись після виконання певної кількості кроків або після того, як відстані кожного з векторів відповідного елемента

розбиття до відповідної гіперплощини не поліпшуються.

6) Перевіряється критерій ефективності здійсненої гіперплощинної кластеризації (компактності першої ланки кластеру, тощо). При невиконанні критерію перейти на формування другої ланки кластеру.

2. *Формування другої ланки кластеру* (на прикладі критерію компактності кластеру).

1) З отриманих в процесі побудови першої ланки розбиттів вихідної множини двох множин - $\Omega_1^{l_1}(1), \Omega_2^{l_1}(1)$, де l_1 - кількість ітерацій в рамках першої ланки, вилучаються вектори, які знаходяться від гіперплощин на віддалі більшій від бажаної. Отримуються множини $\Omega^0(2) = \Omega_1^0(2) \cup \Omega_2^0(2), \Omega_1^0(2) \cap \Omega_2^0(2) = \emptyset$.

2) Побудувати гіперплощини $L(A_k(2), b_k(2))$ для $\Omega_k^0(2), k = 1, 2$, яка визначається як множина розв'язків (псевдо розв'язків) систем алгебраїчних рівнянь

$$A_k(2)x = b_k(2), \quad k = 1, 2.$$

3) Обчислити оптимальні $A_{kopt}(2), b_{kopt}(2)$ для $L(A_k(2), b_k(2)), k = 1, 2$.

4) Обновити $\Omega_k^0(2), k = 1, 2$ по відстані вектора $x(j) \in \Omega^0(2)$ до кожної з гіперплощин $L(A_k(2), b_k(2)), k = 1, 2$,

$$\begin{aligned} & \rho^2(x: x(j), L(A_{kopt}(2), b_{kopt}(2))) = \\ & = (b_{kopt}(2) - A_{kopt}(2)x(j))^T R^T(A_{kopt}(2)) \\ & (b_{kopt}(2) - A_{kopt}(2)x(j)). \end{aligned}$$

(Дії подібні до пункту 3 побудови першої ланки).

5) Перейти на виконання операції 2 з новими підмножинами $\Omega_1^j(1), \Omega_2^j(1), j = 1, 2, \dots$. Верхній індекс j позначає кількість ітерацій на етапі другої ланки.

6) Зупинка алгоритму на етапі другої ланки кластеру може завершитись після того, як відстані кожного з векторів відповідного елемента розбиття до відповідної гіперплощини не поліпшуються.

7) Перевіряється критерій ефективності здійсненої гіперплощинної кластеризації (наприклад, вимога рівня компактності ланки кластеру). При виконанні критерію побудова кластеру завершується, а при невиконанні здійснюється перехід на формування третьої ланки кластеру.

Висновки

Розглянута концепція кусково гіперплощинної кластеризації скінченних сукупностей векторів евклідового простору. Запропоновано алгоритми розв'язання задачі в класичній

постановці – розподіл навчальної вибірки по непересічних кластерах. Приведено необхідні результати теорії псевдо обернених та проєкційних матриць. оптимальних параметрів розглядуваних гіперплощин.

Список використаних джерел

1. Кириченко Н.Ф. Анализ и синтез систем классификации сигналов средствами возмущений псевдообратных и проекционных операций / Кириченко Н.Ф., Кудин Г.И. // Кибернетика и системный анализ. – 2009. – №3. – С. 47-57.
2. Кириченко Н.Ф. Псевдообращение в задачах кластеризации / Кириченко Н.Ф., Донченко В.С. // Кибернетика и системный анализ. – 2007. – №5. – С. 73 – 92.
3. Кириченко Н.Ф. Синтез систем нейрофункциональных преобразователей в решении задач классификации / Кириченко Н.Ф., Кривonos Ю.Г., Лепеха Н.П. // Кибернетика и системный анализ. – 2007. – №3. – С. 47-57
4. Кириченко Н.Ф. Оптимизация синтеза гиперплоскостных кластеров и нейрофункциональных преобразований в системах классификации сигналов / Кириченко Н.Ф., Кривonos Ю.Г., Лепеха Н.П. // Кибернетика и системный анализ. – 2008. – №6. – С. 107 – 125.

References

1. KIRICHENKO, N.F. and KUDIN, G.I. (2009) “Analysis and synthesis of signal classification systems by perturbing pseudoinverse and projection operations”, *Cybernetics and System Analysis*, 45 (4), pp. 613-622.
2. KIRICHENKO, N.F. and DONCHENKO, V.S., (2007) “Pseudoinverse in clustering problems”, *Cybernetics and System Analysis*, 43 (4), pp. 527-541.
3. KIRICHENKO, N.F., KRIVONOS, YU.G., and LEPEKHA, N.P. (2007) “Synthesis of systems of neurofunctional transformations in classification problems”, *Cybernetics and System Analysis*, 43 (3), pp. 353-361.
4. KIRICHENKO, N.F., KRIVONOS, YU.G., and LEPEKHA, N.P. (2008) “ Synthesis of systems of neurofunctional transformations in classification problems”, *Cybernetics and System Analysis*, 44 (6), pp. 832-839.

Надійшла до редколегії 19.08.2017