УДК 519.685.3

Н.Ф. КИРИЧЕНКО, Г.И. КУДИН

АНАЛИЗ И СИНТЕЗ СИСТЕМ КЛАССИФИКАЦИИ СИГНАЛОВ СРЕДСТВАМИ ВОЗМУЩЕНИЙ ПСЕВДООБРАТНЫХ И ПРОЕКЦИОННЫХ ОПЕРАЦИЙ

Ключевые слова: *синтез систем классификации, псевдообратные и проекционные матрицы*.

ВВЕДЕНИЕ

В работах [1–5] предложены идеи синтеза систем классификации, кластеризации сигналов, основанные на использовании свойств псевдообратных матриц. Ниже предлагается развитие подобных средств синтеза систем, включая анализ линейной разделимости конечных множеств за счет применения ранее полученных результатов по теории возмущения псевдообратных и проекционных матриц [6, 7].

В настоящей работе последовательно описываются основные соотношения для возмущений матриц, а также соответствующих им возмущений псевдообратных и проекционных операций, средства выделения (фильтрации) в конечном множестве точек из R^n отделимых и разделимых подмножеств, оптимизации процессов отделимости и разделимости подмножеств при отборе элементов подмножеств, интерпретации полученных результатов по математическому конструированию (синтезу) систем классификации различной природы сигналов.

1. СВЕДЕНИЯ ИЗ ТЕОРИИ ПСЕВДООБРАТНЫХ И ПРОЕКЦИОННЫХ МАТРИЦ

Рассмотрим способы представления матрицы $A \in R^{m \times n}$, составленной из элементов a_{ij}:

$$A = (a(1) \vdots \ldots \vdots a(n)) \equiv (a_{(1)} \vdots \ldots \vdots a_{(m)})^{\mathrm{T}} \in R^{m \times n}, \tag{1}$$

где $a(j) \in R^m$ — столбцы матрицы A, $a_{(i)} \in R^n$ — векторы соответствующих строк, $i = \overline{1, m}$, $j = \overline{1, n}$.

1.1. Оптимизационная форма определения псевдообратной матрицы. Для матрицы $A \in R^{m \times n}$ псевдообратную матрицу $A^+ \in R^{n \times m}$ можно определить согласно условию

$$\forall b \in R^m, \ A^+ b = \arg \min_{x \in \Omega_A(b)} \| x \|^2, \text{ где } \Omega_A(b) = \arg \min_{x \in R^n} \| Ax - b \|^2. \tag{2}$$

Существует несколько эквивалентных определений псевдообратной матрицы [2].

1.2. Основные проекционные матрицы, связанные с псевдообратной матрицей. При применении псевдообращения важными являются проекционные матрицы, которые определяются и вычисляются с помощью матриц A и A^+:

- проекционные матрицы

$$Z(A) = I_n - A^+ A, \ Z(A^{\mathrm{T}}) = I_m - AA^+ \tag{3}$$

— это соответственно проекторы на подпространство, ортогональное подпространству вектор-строк матрицы A, и на подпространство, ортогональное подпространству вектор-столбцов матрицы A;

- взвешенные проекционные матрицы

$$R(A) = A^+ (A^+)^{\mathrm{T}}, \ R(A^{\mathrm{T}}) = (A^+)^{\mathrm{T}} A^+. \tag{4}$$

1.3. Зависимость псевдообратных и проекционных матриц от добавления произвольных вектор-строк к исходной матрице. Если предположить, что расширение матрицы A (формула (1)) происходит добавлением новой строки $a^{\mathrm{T}} \in R^n$ после $(i-1)$-й строки $i = \overline{2, m+1}$, т.е. образуется матрица

$$A_{i,a} = (a_{(1)} \vdots \ldots \vdots a_{(i-1)} \vdots a \vdots a_{(i)} \vdots \ldots \vdots a_{(m)})^{\mathrm{T}} \in R^{(m+1) \times n}, \tag{5}$$

то при известной псевдообратной матрице $A^+ \in R^{n \times m}$ для рекуррентного вычисления псевдообратной матрицы $A^+_{i,a} \in R^{n \times (m+1)}$ имеют место соотношения — прямые формулы Гревиля.

Прямые формулы Гревиля для псевдообратных матриц. Если неизвестную псевдообратнаую матрицу $A^+_{i,a}$ представить в виде

$$\mathrm{A}^+_{\mathrm{i,a}} = (\mathrm{p}(1) \vdots \ldots \vdots \mathrm{p}(\mathrm{i}-1) \vdots \mathrm{p}(\mathrm{i}) \vdots \mathrm{p}(\mathrm{i}+1) \vdots \ldots \vdots \mathrm{p}(\mathrm{m}+1)) \in \mathrm{R}^{\mathrm{n} \times (\mathrm{m}+1)}, \tag{6}$$

где

$$P_i = (p(1) \vdots \ldots \vdots p(i-1) \vdots p(i+1) \vdots \ldots \vdots p(m+1)), \tag{7}$$

т.е. считать, что матрицу $A^+_{i,a}$ можно получить из матрицы $P_i \in R^{n \times m}$ добавлением после $(i-1)$-го столбца вектора $p(i)$, то для матрицы P_i имеет место формула

$$P_i = (1 - p(i) a^{\mathrm{T}}) A^+. \tag{8}$$

При этом неизвестный вектор $p(i)$ определяется согласно свойствам линейной зависимости вектора a от векторов подпространства вектор-строк матрицы A:

1) если вектор a линейно независим от векторов подпространства вектор-строк матрицы A, т.е. $a^{\mathrm{T}} Z(A) a > 0$, то

$$p(i) = Z(A) a \, \| Z(A) a \|^{-2}; \tag{9}$$

2) если вектор a линейно зависим от векторов подпространства вектор-строк матрицы A, т.е. $a^{\mathrm{T}} Z(A) a = 0$, то

$$p(i) = R(A) a (1 + a^{\mathrm{T}} R(A) a)^{-1}. \tag{10}$$

Следствие. Операция псевдообращения коммутативна с транспонированием, поэтому прямые формулы Гревиля аналогично записываются для варианта расширения матрицы столбцом.

Прямые формулы Гревиля для проекционных матриц. Используя прямые формулы Гревиля (8)–(10) для представления псевдообратной матрицы $A^{+}_{i,a}$, можно получить формулы для вычисления проекционных матриц $Z(A_{i,a})$, $Z(A^{\mathrm{T}}_{i,a})$ и взвешенных проекционных матриц $R(A_{i,a})$, $R(A^{\mathrm{T}}_{i,a})$ (см. (3), (4)):

а) при $a^{\mathrm{T}} Z(A) a > 0$ (вектор a линейно независим):

• для проекционных матриц

$$Z(A_{i,a}) = Z(A) - p(i) p^{\mathrm{T}}(i) \| p(i) \|^{-2} \in R^{n \times n}, \tag{11}$$

$$Z(A^{\mathrm{T}}_{i,a}) = \begin{pmatrix} Z_{11}(A^{\mathrm{T}}) & 0(i-1) & Z_{12}(A^{\mathrm{T}}) \\ 0^{\mathrm{T}}(i-1) & 0 & 0^{\mathrm{T}}(m+1-i) \\ Z_{21}(A^{\mathrm{T}}) & 0(m+1-i) & Z_{22}(A^{\mathrm{T}}) \end{pmatrix} \in R^{(m+1)\times(m+1)}, \tag{12}$$

где матрицы-блоки $Z_{11}(A^{\mathrm{T}}) \in R^{(i-1)\times(i-1)}$, $Z_{21}(A^{\mathrm{T}}) \in R^{(m-i+1)\times(i-1)}$, $Z_{12}(A^{\mathrm{T}}) \in$ $\in R^{(i-1)\times(m-i+1)}$, $Z_{22}(A^{\mathrm{T}}) \in R^{(m-i+1)\times(m-i+1)}$ определяются согласно представлению невозмущенной проекционной матрицы $Z(A^{\mathrm{T}})$:

$$Z(A^{\mathrm{T}}) = \begin{pmatrix} Z_{11}(A^{\mathrm{T}}) & Z_{12}(A^{\mathrm{T}}) \\ Z_{21}(A^{\mathrm{T}}) & Z_{22}(A^{\mathrm{T}}) \end{pmatrix} \in R^{m \times m}, \tag{13}$$

а $0(k) = (0, 0, \ldots, 0)^{\mathrm{T}} \in R^{k}$ — нулевой вектор;

$$R(A_{i,a}) = R(A) - (p(i) a^{\mathrm{T}} R(A) - R(A) a p^{\mathrm{T}}(i)) \| p(i) \|^{-2} +$$
$$+ p(i) p^{\mathrm{T}}(i) (1 + a^{\mathrm{T}} R(A) a)^{-1} \in R^{n \times n}; \tag{14}$$

$$R(A^{\mathrm{T}}_{i,a}) = \begin{pmatrix} R_{11}(A^{\mathrm{T}}) & c(i-1) & R_{12}(A^{\mathrm{T}}) \\ c^{\mathrm{T}}(i-1) & 1 & c^{\mathrm{T}}(m+1-i) \\ R_{21}(A^{\mathrm{T}}) & c(m+1-i) & R_{22}(A^{\mathrm{T}}) \end{pmatrix} \in R^{(m+1)\times(m+1)}, \tag{15}$$

где матрицы-блоки $R_{11}(A^{\mathrm{T}}) \in R^{(i-1)\times(i-1)}$, $R_{21}(A^{\mathrm{T}}) \in R^{(m-i+1)\times(i-1)}$, $R_{12}(A^{\mathrm{T}}) \in$ $\in R^{(i-1)\times(m-i+1)}$, $R_{22}(A^{\mathrm{T}}) \in R^{(m-i+1)\times(m-i+1)}$ — блоки невозмущенной проекционной матрицы $R(A^{\mathrm{T}})$:

$$R(A^{\mathrm{T}}) = \begin{pmatrix} R_{11}(A^{\mathrm{T}}) & R_{12}(A^{\mathrm{T}}) \\ R_{21}(A^{\mathrm{T}}) & R_{22}(A^{\mathrm{T}}) \end{pmatrix} \in R^{m \times m}, \tag{16}$$

а векторы $c(i-1)$, $c(m+1-i)$ являются вектор-блоками вектора c:

$$c(i-1) = (c_1, c_2, \ldots, c_{i-i})^{\mathrm{T}}, \ c(m+1-i) = (c_i, c_{i+1}, \ldots, c_m)^{\mathrm{T}}, \tag{17}$$

$$(c(i-1) \vdots c(m+1-i))^{\mathrm{T}} \equiv (c_1, c_2, \ldots, c_{i-i}, c_i, c_{i+1}, \ldots, c_m)^{\mathrm{T}} \equiv c = -A^{+\mathrm{T}} a; \tag{18}$$

б) при $a^{\mathrm{T}} Z(A) a = 0$ (вектор a линейно зависим):

$$Z(A_{i,a}) = Z(A) \in R^{n\times n}, \tag{19}$$

$$R(A_{i,a}s) = \begin{pmatrix} Z_{11}(A^{\mathrm{T}}_{i,a}) & d(i-1) & Z_{12}(A^{\mathrm{T}}_{i,a}) \\ d^{\mathrm{T}}(i-1) & -1 & d^{\mathrm{T}}(m+1-i) \\ Z_{21}(A^{\mathrm{T}}_{i,a}) & d(m+1-i) & Z_{22}(A^{\mathrm{T}}_{i,a}) \end{pmatrix} \in\in R^{(m+1)\times(m+1)}, \tag{20}$$

где матрицы-блоки $Z_{kl}(A^{\mathrm{T}}_{i,a})$ определяются матричным равенством

$$\begin{pmatrix} Z_{11}(A^{\mathrm{T}}_{i,a}) & Z_{12}(A^{\mathrm{T}}_{i,a}) \\ Z_{21}(A^{\mathrm{T}}_{i,a}) & Z_{22}(A^{\mathrm{T}}_{i,a}) \end{pmatrix} = Z(A^{\mathrm{T}}) + dd^{\mathrm{T}}(1 + a^{\mathrm{T}}R(A)a) \in R^{m\times m}, \tag{21}$$

проекционная матрица $Z(A^{\mathrm{T}})$ определяется формулой (15), а векторы $d(i-1)$, $d(m+1-i)$ — векторным равенством

$$(d(i-1) \vdots d(m+1-i))^{\mathrm{T}} \equiv (d_1, \ldots, d_{i-i}, d_i, d_{i+1}, \ldots, d_m)^{\mathrm{T}} \equiv$$
$$\equiv -A^{+\mathrm{T}} a(1 + a^{\mathrm{T}} R(A)a)^{-1}; \tag{22}$$

- для взвешенных проекционных матриц:

$$R(A_{i,a}) = R(A) + R(A)aa^{\mathrm{T}}R(A)(1 + a^{\mathrm{T}}R(A)a)^{-1} \in R^{n\times n}, \tag{23}$$

$$\mathrm{R}(\mathrm{A}^{\mathrm{T}}_{\mathrm{i,a}}) = \begin{pmatrix} \mathrm{R}_{11}(\mathrm{A}^{\mathrm{T}}_{\mathrm{i,a}}) & \mathrm{b(i-1)} & \mathrm{R}_{12}(\mathrm{A}^{\mathrm{T}}_{\mathrm{i,a}}) \\ \mathrm{b}^{\mathrm{T}}(\mathrm{i-1}) & 1 & \mathrm{b}^{\mathrm{T}}(\mathrm{m+1-i}) \\ \mathrm{R}_{21}(\mathrm{A}^{\mathrm{T}}_{\mathrm{i,a}}) & \mathrm{b(m+1-i)} & \mathrm{R}_{22}(\mathrm{A}^{\mathrm{T}}_{\mathrm{i,a}}) \end{pmatrix} \in \mathrm{R}^{(\mathrm{m+1})\times(\mathrm{m+1})}, \tag{24}$$

матрицы-блоки $R_{kl}(A^{\mathrm{T}}_{i,a})$ определяются равенством

$$\begin{pmatrix} R_{11}(A^{\mathrm{T}}_{i,a}) & R_{12}(A^{\mathrm{T}}_{i,a}) \\ R_{21}(A^{\mathrm{T}}_{i,a}) & R_{22}(A^{\mathrm{T}}_{i,a}) \end{pmatrix} =$$

$$= R(A^{\mathrm{T}}) + A^{+\mathrm{T}} aa^{\mathrm{T}} A^{+} \| p(i) \|^2 - A^{+\mathrm{T}} (p(i)a^{\mathrm{T}} + ap^{\mathrm{T}}(i))A^{+} \in R^{m\times m}, \tag{25}$$

где взвешенная проекционная матрица $R(A^{\mathrm{T}})$ определяется формулой (17), а векторы $b(i-1)$, $b(m+1-i)$ — векторным равенством

$$(b(i-1) \vdots b(m+1-i))^{\mathrm{T}} \equiv (b_1, \ldots, b_{i-i}, b_i, b_{i+1}, \ldots, b_m)^{\mathrm{T}} \equiv$$
$$\equiv b = A^{+\mathrm{T}} p(i) + \| p(i) \|^2 c, \tag{26}$$

$$p(i) = R(A)a(1 + a^{\mathrm{T}}R(A)a)^{-1}, \tag{27}$$

$$c = -A^{+\mathrm{T}} a. \tag{28}$$

1.4. Зависимость псевдообратных и проекционных матриц от удаления произвольных вектор-строк в исходной матрице. Если матрице $A_{i,a} \in R^{(m+1)\times n}$ (после $(i-1)$-й строки $(i = \overline{2, m+1})$ матрицы A расположена вектор-строка a^{T}) соответствует псевдообратная матрица $A^{+}_{i,a} \in R^{n\times(m+1)}$, то для определения псевдообратной матрицы $A^{+} \in R^{n\times m}$ (строка a^{T} из матрицы $A_{i,a}$ удаляется) имеют место обратные формулы Гревиля.

Обратные формулы Гревиля для псевдообратных матриц. 1. В случае линейной зависимости удаляемой вектор-строки a^{T} от векторов подпространства вектор-строк матрицы $A_{i,a}$, а это определяется выполнением условия

$$a^{\mathrm{T}} p(i) < 1, \tag{29}$$

псевдообратная матрица $A^{+} \in R^{n\times m}$ имеет вид

$$A^{+} = (I_n + p(i) a^{\mathrm{T}} (1 - a^{\mathrm{T}} p(i))^{-1}) P_i . \tag{30}$$

При этом ранг псевдообратной матрицы не меняется, т.е. $\operatorname{rank} A = \operatorname{rank} (A^{\mathrm{T}} \vdots a)^{\mathrm{T}}$.

2. В случае линейной независимости удаляемой вектор-строки a^{T} от векторов подпространства вектор-строк матрицы A, а это определяется выполнением условия

$$a^{\mathrm{T}} p(i) = 1, \tag{31}$$

матрица $A^{+} \in R^{n\times m}$ определяется формулой

$$A^{+} = (I_n - p(i) p^{\mathrm{T}}(i) \| p(i) \|^{-2}) P_i . \tag{32}$$

При этом ранг псевдообратной матрицы падает, т.е. $\operatorname{rank} A = \operatorname{rank} (A^{\mathrm{T}} \vdots a)^{\mathrm{T}} - 1$. Имеют место соотношения

$$\| p(i) \|^2 = R_{ii} (A_{i,a}^{\mathrm{T}}), \tag{33}$$

$$1 - p^{\mathrm{T}}(i) a = Z_{ii} (A_{i,a}^{\mathrm{T}}). \tag{34}$$

Обратные формулы Гревиля для проекционных матриц. Используя обратные формулы Гревиля (30), (32) для представления псевдообратной матрицы $A^{+} \in R^{n\times m}$ при известной псевдообратной матрице $A_{i,a}^{+} \in R^{n\times(m+1)}$ (из матрицы $A_{i,a} \in R^{(m+1)\times n}$ удаляется i-я ($i = \overline{2, m+1}$) вектор-строка a^{T}), можно получить полезные для приложений формулы экономного вычисления проекционных матриц $Z(A)$, $Z(A^{\mathrm{T}})$, $R(A)$, $R(A^{\mathrm{T}})$.

1. Если удаляемая вектор-строка a^{T} линейно зависима от векторов подпространства вектор-строк матрицы $A_{i,a}$, т.е. $a^{\mathrm{T}} p(i) < 1$, то

$$Z(A) = Z(A_{i,a}); \tag{35}$$

$$R(A) = R(A_{i,a}) + p(i) p^{\mathrm{T}}(i) Z_{ii}^{-1} (A_{i,a}^{\mathrm{T}}); \tag{36}$$

$$Z(A^{\mathrm{T}}) = E_i Z(A_{i,a}^{\mathrm{T}}) - E_i h_i h_i^{\mathrm{T}} E_i^{\mathrm{T}} Z_{ii}^{-1} (A_{i,a}^{\mathrm{T}}). \tag{37}$$

Здесь использовано соотношение (34), а также введена в рассмотрение матрица E_i, полученная изъятием i-й ($i = \overline{2, m+1}$) вектор-строки из единичной матрицы I_{m+1}:

$$E_i = (e_1(m+1) \vdots \ldots \vdots e_{i-1}(m+1) \vdots e_{i+1}(m+1) \vdots \ldots \vdots e_{m+1}(m+1))^{\mathrm{T}} \in R^{m\times(m+1)}, \tag{38}$$

$$h_i = -A_{i,a}^{+\mathrm{T}} a; \tag{39}$$

$$R(A^{\mathrm{T}}) = E_i R(A_{i,a}^{\mathrm{T}}) E_i^{\mathrm{T}} + E_i (h_i h_i^{\mathrm{T}} (p^{\mathrm{T}}(i) p(i)) + h_i t_i^{\mathrm{T}} + t_i h_i^{\mathrm{T}}) E_i^{\mathrm{T}}, \tag{40}$$

$$t_i = -A_{i,a}^{+\mathrm{T}} p(i) Z_{ii}^{-1} (A_{i,a}^{\mathrm{T}}). \tag{41}$$

2. Если удаляемая вектор-строка a^{T} линейно независима от подпространства вектор-строк матрицы $A_{i,a}$, т.е. $a^{\mathrm{T}} p(i) = 1$, то

$$Z(A) = Z(A_{i,a}) + p(i) p^{\mathrm{T}}(i) \| p(i) \|^{-2}, \tag{42}$$

$$R(A) = Z(p^{\mathrm{T}}(i)) R(A_{i,a}) Z(p^{\mathrm{T}}(i)), \tag{43}$$

$$Z(p^{\mathrm{T}}(i)) = I_n - p(i) p^{\mathrm{T}}(i) \| p(i) \|^{-2}. \tag{44}$$

С использованием обозначения матрицы E_i (формула (37)) получаем

$$Z(A^{\mathrm{T}}) = E_i Z(A_{i,a}^{\mathrm{T}}) E_i^{\mathrm{T}}, \tag{45}$$

$$R(A^{\mathrm{T}}) = E_i R(A_{i,a}^{\mathrm{T}}) E_i^{\mathrm{T}}. \tag{46}$$

2. ФИЛЬТРАЦИЯ ОТДЕЛИМЫХ ПОДМНОЖЕСТВ ДЛЯ КОНЕЧНОГО МНОЖЕСТВА ТОЧЕК В КОНЕЧНОМЕРНОМ ПРОСТРАНСТВЕ

Пусть $x=(x_1,\ldots,x_m)^{\mathrm T}$ — вектор признаков, который характеризует распознаваемый сигнал (временной процесс или изображение). Для последовательности $x(1),\ldots,x(n)$ в пространстве признаков, элементы которой соответствуют различным реализациям исследуемых сигналов, сформулируем задачу линейной классификации следующим образом. Задано некоторое конечное множество точек

$$\Omega_x=\{x\colon x(j)\in R^m,\ j=\overline{1,n}\},\tag{47}$$

и оно не является линейно отделимым от начала координат, т.е. $\overline{\exists}\, a\in R^m$, $\|a\|>0$, для которого

$$a^{\mathrm T}x(j)\geq 1,\ \ j=\overline{1,n}.\tag{48}$$

Возникает вопрос: какое подмножество точек $\Omega_f\in\Omega_x$ не позволяет этому множеству быть линейно отделимым. При этом желательно, чтобы такое подмножество содержало минимальное количество элеметов. Поставленная задача, с точки зрения ее классического решения, при больших значениях n и m сложная, поэтому предлагается решить ее последовательными исключениями элементов, которые неблагоприятно влияют на линейную отделимость множеств, и тем самым осуществить фильтрацию линейно отделимых подмножеств. Затем покажем полезность использования такого алгоритма фильтрации при синтезе систем классификации сигналов.

Из ранее полученных результатов [1] известно, что для множества точек Ω_x существует такая линейная операция $a^{\mathrm T}x$, что $\forall x\in\Omega_x$ имеет место неравенство

$$a^{\mathrm T}x\geq 1,\tag{49}$$

если выполняется условие

$$\min_{\mathrm y\in\Omega_{\mathrm y}}\mathrm y^{\mathrm T}\mathrm Z(\mathrm X)\mathrm y=\mathrm y_*^{\mathrm T}\mathrm Z(\mathrm X)\mathrm y_*=0,\tag{50}$$

где

$$\Omega_y=\{y\colon y=(y(1),\ldots,y(n))^{\mathrm T},\ y(j)\geq 1, j=\overline{1,n}\},\tag{51}$$

$$X=(x(1)\vdots\ldots\vdots x(n))\equiv(x_{(1)}\vdots\ \ldots\vdots x_{(m)})^{\mathrm T}\in R^{m\times n}.\tag{52}$$

При этом искомая операция определяется согласно формуле

$$a=(X^{\mathrm T})^{+}y_*,\tag{53}$$

а расстояние опорной гиперплоскости

$$a^{\mathrm T}x=1\tag{54}$$

от начала координат принимает значение

$$h(a)=(y_*^{\mathrm T}R(X)y_*)^{-1/2}.\tag{55}$$

Если же соотношение (50) не выполняется, т.е

$$y_*^{\mathrm T}Z(X)y_*>0,\tag{56}$$

то поиск элемента $x(k)$, устранение которого из множества Ω_x приведет или к линейной отделимости множества точек $\Omega_x\setminus x(k)=\{x\colon x(j)\in R^m,\ j=\overline{1,n},\ j\neq k\}$, или к улучшению перспективы линейной отделимости этого множества в смысле минимизации значения левой части неравенства (56), сводится к определению значения k согласно условию

$$\mathrm k=\arg\min_{\mathrm i=\overline{1,\mathrm n}}\ \min_{\mathrm y_{\mathrm i}\in\Omega_{\mathrm y_{\mathrm i}}}\mathrm y_{\mathrm i}^{\mathrm T}\mathrm Z(\mathrm X_{\mathrm i})\mathrm y_{\mathrm i},\tag{56}$$

где

$$\Omega_{y_i} = \{ y_i : y_i = (y(1), \ldots, y(i-1), y(i+1), \ldots, y(n))^{\mathrm{T}}, \; y(j) \geq 1, \; j = \overline{1, n}, j \neq i \}, \quad (57)$$

$$X_i = (x(1) \vdots \ldots \vdots x(i-1) \vdots x(i+1) \vdots \ldots \vdots x(n)). \quad (58)$$

Используя результаты предыдущего раздела, можно сделать вывод, что при изъятии из множества Ω_x элемента $x(i)$, линейно независимого от других элементов этого множества, т.е. удовлетворяющего условию

$$1 - x^{\mathrm{T}}(i) X^{+\mathrm{T}} e_i(n) = 0, \quad (59)$$

$e_i(n)$ — i-й орт в пространстве R^n, имеет место соотношение

$$y_i^{\mathrm{T}} Z(X_i) y_i = y^{\mathrm{T}} Z(X) y. \quad (60)$$

Поэтому на основании формулы (36), описывающей возмущение проекционной матрицы $Z(X)$ при изъятии из множества Ω_x элемента $x(i)$, линейно зависимого от других элементов этого множества, условие выбора значения k, при котором достигается наилучшее изъятие элемента из множества Ω_x для фильтрации отделимого подмножества, представляется следующим образом:

$$k = \arg \min_{i \in \Omega_i} \min_{y_j \geq 1, \; j = \overline{1, n}, \; j \neq i} y_{i,0}^{\mathrm{T}} (Z(X) -$$

$$- \mathrm{X}^{+} \mathrm{x}(\mathrm{i}) \mathrm{x}^{\mathrm{T}}(\mathrm{i}) \mathrm{X}^{+\mathrm{T}} (1 - \mathrm{x}^{\mathrm{T}}(\mathrm{i}) \mathrm{X}^{+\mathrm{T}} \mathrm{e}_{\mathrm{i}}(\mathrm{n}))^{-1}) \mathrm{y}_{\mathrm{i},0}, \quad (61)$$

$$\Omega_i = \{ i : 1 - x^{\mathrm{T}}(i) X^{+\mathrm{T}} e_i(n) > 0, \; i = \overline{1, n} \}, \quad (62)$$

$$y_{i,o} = (y(1), \ldots, y(i-1), 0, y(i+1), \ldots, y(n))^{\mathrm{T}} \in R^n). \quad (63)$$

Если подмножество $\Omega_x \setminus x(k)$ удовлетворяет условию линейной отделимости, то процесс фильтрации множества Ω_x к свойству линейной отделимости завершается. В противном случае для множества $\Omega_x \setminus x(k)$ осуществляется тот же алгоритм выявления наиболее перспективного на изъятие элемента, что и для исходного множества Ω_x.

Предположим, что линейная отделимость отфильтрованного подмножества возникает за счет последовательного изъятия из исходного множества нескольких его членов, т.е. $\Omega_x \setminus \Omega_x(1)$, где $\Omega_x(1) = \{x : x = x(k_1), \ldots, x(k_{n_1 - 1}), x(k_{n_1})\}$ — линейно отделимое подмножество, а предыдущее подмножество $\Omega_x \setminus \{x : x = x(k_1), \ldots, x(k_{n_1 - 1})\}$ — линейно неотделимое.

Определение 1. Множество Ω_x будем называть 2-кусочно-линейно отделимым, если $\Omega_x \setminus \Omega_x(1)$, $\Omega_x(1)$ — линейно отделимые подмножества.

Определение 2 (рекуррентное определение). Множество Ω_x называется s-кусочно-линейно отделимым, если существует такое $(s-1)$-кусочно-линейно отделимое подмножество $\Omega_x(1) \in \Omega_x$, что $\Omega_x \setminus \Omega_x(1)$ — линейно отделимое подмножество.

Пусть Ω_x — s-кусочно-линейно отделимое множество, тогда подмножества

$$\Omega_x \setminus \Omega_x(1), \ldots, \Omega_x(s-2) \setminus \Omega_x(s-1) \Omega_x(s-1) \quad (64)$$

линейно отделимы и в качестве меры s-кусочно-линейной отделимости рассмотрим величину

$$h = \arg \min_{i = \overline{1, s}} \max_{y_i \in \Omega_y(i)} (y_i^{\mathrm{T}} R(X(i)) y_i)^{1/2}, \quad (65)$$

где

$$\Omega_y(i) = \{ y_i : y_i = (y_i(1), \ldots, y_i(n_i))^{\mathrm{T}}, \; y_i(j) \geq 1, \; j = \overline{1, n_i}, \; y_i^{\mathrm{T}} Z(X_i) y_i = 0 \}, \quad (66)$$

$X(i) = (x(i,1) \vdots \ldots \vdots x(i, n_i)) \in R^{m \times n_i}$ — матрица, образованная элементами i-го подмножества из последовательности подмножеств (64).

Мера s-кусочно линейной отделимости — величина h — при реализации алгоритма может быть недостаточно большой, поэтому возможен процесс увеличения величины h транспортировкой элементов из подмножества $\Omega_x \setminus \Omega_x(1)$ в подмножество $\Omega_x(1)$. Например, в условиях 2-кусочно-линейной отделимости, когда выполняется неравенство $h_1 < h_2$, можно увеличить h_1 и $h = \min\{h_1, h_2\}$. Действительно, если дополнительно проанализировать вектор-столбцы матрицы $X(1)$ — элементы подмножества $\Omega_x \setminus \Omega_x(1)$, проверить выполнение условий их линейной независимости от подпространства вектор-столбцов матрицы $X(1)$

$$1 - x^{\mathrm{T}}(1, j) X^{+\mathrm{T}}(1) e_j(n_1) = 0, \; j = \overline{1, n_1}, \; X(1) = (x(1,1) \vdots \ldots \vdots x(1, n_1)) \in R^{m \times n_1}, \quad (67)$$

то вектор-столбцы матрицы $X(1)$ будут удовлетворять этому условию. Удаление такого вектор-столбца (пусть это будет вектор $x(1, n_1)$) из подмножества $\Omega_x \setminus \Omega_x(1)$ и последующего присоединения его к подмножеству $\Omega_x(1)$ сохраняет свойство подмножества $\Omega_x \setminus \Omega_x(1)$ на отделимость, вместо матрицы $X(1) \in R^{m \times n_1}$ рассматривается матрица $X_1(1) \in R^{m \times (n_1 - 1)}$, что обуславливает пересмотр параметра линейной операции $a^{\mathrm{T}}(1)x$, когда $\forall x \in \Omega_x \setminus \Omega_x(1)$ имело место неравенство

$$a^{\mathrm{T}}(1)x \ge 1. \quad (68)$$

Появляется возможность увеличения h_1, если после выполнения условия

$$\min_{y \in \Omega_y(1)} y^{\mathrm{T}} Z(X_1(1)) y = y_{1*}^{\mathrm{T}} Z(X_1(1)) y_{1*} = 0, \quad (69)$$

где

$$\Omega_y(1) = \{y \colon y = (y(1), \ldots, y(n_1 - 1))^{\mathrm{T}}, \; y(j) \ge 1, \; j = \overline{1, (n_1 - 1)}\}, \quad (70)$$

определяется новый параметр линейной операции $a_1(1)$ согласно формуле

$$a_1^{\mathrm{T}}(1) = X_1^{\mathrm{T}+}(1)) y_{1*}, \quad (71)$$

а затем и новое расстояние опорной гиперплоскости

$$a_1^{\mathrm{T}}(1)x = 1 \quad (72)$$

от начала координат

$$h_1(1) = (y_{1*}^{\mathrm{T}} R(X_1(1)) y_{1*})^{-1/2}, \quad (73)$$

и при этом $h_1 < h_1(1)$. Если при выполнении условий линейной отделимости подмножества $\Omega_x(1)$, расширенного вектор-столбцом $x(1, n_1)$, вычисления для него по схеме (69)–(73) дают измененное значение величины $h_2 - h_2(1)$, которое обеспечивает выполнение неравенства

$$\min\{h_1(1), h_2(1)\} = h(1) > h = \min\{h_1, h_2\}, \quad (74)$$

то цель улучшения качества линейной отделимости достигнута.

Если при удалении из матрицы $X(1)$ вектор-столбцов, линейно независимых от подпространства ее вектор-столбцов, качество линейной отделимости не улучшается, то осуществляется необходимый перебор вектор-столбцов матрицы $X(1)$ для последующей транспортировки в подмножество $\Omega_x(1)$. Расчеты будут осуществляться с помощью формул (36), (45).

При необходимости существенного улучшения качества s-кусочно-линейно отделимого ($s > 2$) множества Ω_x пересчет допускает рекуррентное представление согласно формулам предыдущего раздела. Алгоритм может быть реализован на определении в теле цикла наименее отдаленной от начала координат гиперплоскости $a^{\mathrm{T}}(j)x = 1, \; j \in \overline{1, (s_1 - 1)}$, затем осуществляется вычислительная процедура улучшения качества линейной отделимости удалением из матрицы $X(j)$ вектор-столбцов с последующей их транспортацией в матрицу $X(j+1)$. Необходимые вычисления могут осуществляться по схеме, предложенной выше, для улучшения качества линейной отделимости дважды кусочно-линейно отделимого множества.

3. ФИЛЬТРАЦИЯ РАЗДЕЛИМЫХ ПОДМНОЖЕСТВ ДЛЯ КОНЕЧНОГО МНОЖЕСТВА ТОЧЕК В ПРОСТРАНСТВЕ ТОЧЕК ПРИ КАСКАДНОЙ ДИХОТОМНОЙ КЛАССИФИКАЦИИ СИГНАЛОВ

При исследовании задач распознавания образов с помощью линейных дискриминантных операций на этапе обучения системы классификации по обучающей последовательности в пространстве признаков возникает необходимость проверки возможности линейного разделения данных на классы и организации более сложной структуры разделения, чем линейная, обеспечивающая реализацию классификации для линейно неразделимых на классы данных из обучающей последовательности. Здесь идея фильтрации отделимых подмножеств, изложенная в предыдущем разделе, распространяется на исследование задачи линейной и кусочно-линейной разделимости точек в пространстве признаков из обучающей последовательности.

Пусть для точек $x(j) \in R^m$, $j = \overline{1, n}$ (в целях формирования над компонентами $x(j)$ неоднородных линейных операций будем предполагать, что $x_m(j) = 1$), точки $x(i_k) \in R^m$, $k = \overline{1, n_1}$, находятся в первом классе Ω_{x1}, т.е.

$$x(i_k) \in \Omega_{x1},\ \Omega_{x1} = \{x : x = x(i_1), \ldots, x(i_{n_1 - 1}), x(i_{n_1})\}, \tag{75}$$

точки $x(i_s) \in R^m$, $s = \overline{1, n_2}$, — во втором классе, т.е.

$$x(j_s) \in \Omega_{x2},\ \Omega_{x2} = \{x : x = x(j_1), \ldots, x(j_{n_2 - 1}), x(j_{n_2})\}. \tag{76}$$

При этом $\Omega_x = \Omega_{x1} \cup \Omega_{x2}$, $n = n_1 + n_2$. Тогда эти два класса будут называться линейно разделимыми, если выполняются соотношения $\exists\, a \in R^{m+1}$, $0 < \|a\| < \infty$, для которых

$$a^{\mathrm{T}} \begin{pmatrix} x(i_k) \\ 1 \end{pmatrix} \ge 1,\ k = \overline{1, n_1},\ a^{\mathrm{T}} \begin{pmatrix} x(j_s) \\ 1 \end{pmatrix} \le -1,\ s = \overline{1, n_2}. \tag{77}$$

Необходимое и достаточное условие линейной разделимости этих точек представляется в виде

$$\min_{y \in \Omega_y} y^{\mathrm{T}} Z \begin{pmatrix} X \\ \cdots \\ J_n^{\mathrm{T}} \end{pmatrix} y = y_*^{\mathrm{T}} Z \begin{pmatrix} X \\ \cdots \\ J_n^{\mathrm{T}} \end{pmatrix} y_* = 0, \tag{78}$$

где

$$X = (x(1) \vdots \ldots \vdots x(n)) \in R^{m \times n}. \tag{79}$$

$$\Omega_y = \{ y : y = (y(1), \ldots, y(n))^{\mathrm{T}},\ y(i_k) \ge 1, k = \overline{1, n_1},\ y(j_s) \le -1, s = \overline{1, n_2} \}, \tag{80}$$

$$J_n = (1, 1, \ldots, 1)^{\mathrm{T}} \in R^n,$$

и сам вектор $a \in R^{m+1}$, удовлетворяющий условию (77), принимает значение

$$a = \begin{pmatrix} X \\ \cdots \\ J_n^T \end{pmatrix}^{\mathrm{T}+} y_* . \tag{81}$$

Расстояние между найденными гиперплоскостями

$$a^{\mathrm{T}} \begin{pmatrix} x \\ 1 \end{pmatrix} = 1,\ a^{\mathrm{T}} \begin{pmatrix} x \\ 1 \end{pmatrix} = -1 \tag{82}$$

определяется выражением

$$h = 2(\|a\| - a_{m+1}^2)^{-1/2} \equiv 2 \left(\left(y_*^{\mathrm{T}} R \begin{pmatrix} X \\ \cdots \\ J_n^T \end{pmatrix} y_* \right) - \left(y_*^{\mathrm{T}} \begin{pmatrix} X \\ \cdots \\ J_n^T \end{pmatrix}^{+} e(m+1) \right)^2 \right)^{-1/2}, \tag{83}$$

$$e(m+1) = (0, 0, \ldots, 0, 1)^{\mathrm{T}} \in R^{m+1}.$$

Если же условие (78) не выполняется, т.е. разделяющих гиперплоскостей между рассматриваемыми множествами $x(i_k) \in R^m$, $k=\overline{1,n_1}$ и $x(j_s) \in R^m$, $s=\overline{1,n_2}$, не существует, то целесообразно использовать средство фильтрации отделимых подмножеств до свойства кусочно-линейной разделимости точек в пространстве признаков. А именно, из обучающей последовательности осуществляется такое изъятие элементов, при котором достигается наибольшее уменьшение величины

$$y_*^{\mathrm{T}} Z \begin{pmatrix} X \\ \cdots \\ J_n^{\mathrm{T}} \end{pmatrix} y_* .$$

Поиск элемента $x(l)$, устранение которого из множества Ω_x приведет или к линейной разделимости множества точек $\Omega_x \setminus x(l) = \{x : x(j) \in R^m, j=\overline{1,n}, j \neq l\}$ или к улучшению перспективы линейной разделимости этого множества, сводится к определению значения l согласно условию

$$\mathrm{l} = \arg \min_{\mathrm{q}=\overline{1,\,\mathrm{n}}} \min_{\mathrm{y}_\mathrm{q} \in \Omega_{\mathrm{y}_\mathrm{q}}} \mathrm{y}_\mathrm{q}^{\mathrm{T}} \mathrm{Z} \begin{pmatrix} \mathrm{X}_\mathrm{q} \\ \cdots \\ \mathrm{J}_{\mathrm{n}-1}^{\mathrm{T}} \end{pmatrix} \mathrm{y}_\mathrm{q} , \tag{84}$$

где

$$\begin{gathered} \Omega_{y_q} = \{ y_q : y_q = (y(1), \ldots, y(q-1), y(q+1), \ldots, y(n))^{\mathrm{T}} , \\ y(i_k) \geq 1,\ i_k \neq q,\ k=\overline{1,n_1},\ y(j_s) \geq 1,\ j_s \neq q,\ s=\overline{1,n_2} \}, \end{gathered} \tag{85}$$

$$X_q = (x(1) \vdots \ldots \vdots x(q-1) \vdots x(q+1) \vdots \ldots \vdots x(n)) . \tag{86}$$

Как и в предыдущем разделе, изымать из множества Ω_x для последующей транспортировки в подмножество $\Omega_x(1)$ целесообразно элемент $x(l)$, который линейно зависим от подпространства вектор-столбцов матрицы $(X^{\mathrm{T}} \vdots J_N)^{\mathrm{T}}$.

Предположим, что линейная разделимость отфильтрованного подмножества Ω_x возникает за счет последовательного изъятия из исходного множества нескольких его членов, т.е. подмножество $\Omega_x \setminus \Omega_x(1)$, где $\Omega_x(1) = \Omega_{x1}(1) \cup \Omega_{x2}(1)$, линейно разделимое, но обладает ли таким свойством подмножество $\Omega_x(1)$ — вопрос дополнительных исследований. Введем понятие s-кусочно-линейно разделимого множества.

Определение 3. Множество Ω_x будем называть 2-кусочно-линейно разделимым, если $\Omega_x \setminus \Omega_x(1)$, $\Omega_x(1)$ — линейно разделимые подмножества.

Определение 4 (рекуррентное определение). Множество Ω_x называется s-кусочно-линейно разделимым, если существует такое $(s-1)$-кусочно-линейно разделимое подмножество $\Omega_x(1) \in \Omega_x$, что $\Omega_x \setminus \Omega_x(1)$ — линейно разделимое подмножество.

Если множество Ω_x — s-кусочно-линейно разделимое, то подмножества

$$\Omega_x \setminus \Omega_x(1), \ldots, \Omega_x(s-2) \setminus \Omega_x(s-1) \tag{87}$$

кусочно-линейно разделимы, для каждого из них будут определены векторы $a(1), \ldots, a(s)$ — векторы-нормали соответствующих им разделяющих гиперплоскостей, а также величины h_i, $i=\overline{1,s}$, — расстояние между найденными гиперплоскостями. Для улучшения качества кусочно-линейного разделения множества Ω_x, т.е. увеличения меры s-кусочно-линейной разделимости

$$h = \min_{i=\overline{1,s}} \{h_i\}, \tag{88}$$

несложно адаптировать вычислительную схему улучшения качества s-кусочно-линейной разделимости множества.

Для s-кусочно-линейно разделимых подмножеств

$$\Omega_x \setminus \Omega_x(1), \ldots, \Omega_x(s-2) \setminus \Omega_x(s-1), \Omega_x(s-1) \tag{89}$$

и соответствующих им векторов $a(1), \ldots, a(s)$ средство классификации по обучающей последовательности в пространстве признаков представляется системой с каскадной структурой.

ЗАКЛЮЧЕНИЕ

Методы псевдообращения, в том числе теория возмущения псевдообратных и проекционных матриц, позволяют построить конструктивные, явные схемы средств выделения в конечном множестве дискретных точек отделимых и разделимых подмножеств, оптимизировать качество таких прцессов перебором элементов подмножеств.

Предложенный подход синтеза систем классификации сигналов позволяет в пределах кусочно-линейной модели успешно решать практические задачи, обеспечивая при этом эффективную вычислительную реализацию соответствующих алгоритмов.

СПИСОК ЛИТЕРАТУРЫ

1. Кириченко Н.Ф., Кривонос Ю.Г., Лепеха Н.П. Оптимизация синтеза гиперплоскостных кластеров и нейрофункциональных преобразований в системах классификации сигналов // Кибернетика и системный анализ. — 2008. — № 6. — С. 107–124.
2. Кириченко Н.Ф., Кривонос Ю.Г., Лепеха Н.П. Синтез систем нейрофункциональных преобразователей в решении задач классификации // Там же. — 2007. — № 3. — С. 47–57.
3. Кириченко Н.Ф., Донченко В.С. Псевдообращение в задачах кластеризации // Там же. — 2007. — № 4. — С. 73–92.
4. Кириченко Н.Ф., Крак Ю.В., Полищук А.А. Псевдобратные и проекционные матрицы в задачах синтеза функциональных преобразователей // Там же. — 2004. — № 3. — С. 116–129.
5. Верченко А.П., Кириченко Н.Ф., Лепеха Н.П. Рекуррентные средства кластеризации в применении к задачам сегментации изображений // Проблемы управления и информатики. — 2005. — № 2. — С. 62–71.
6. Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц // Кибернетика и системный анализ. — 1997. — № 2. — С. 98–107.
7. Кириченко Н.Ф., Лепеха Н.П. Применение псевдообратных и проекционных матриц к исследованию задач управления, наблюдения и идентификации // Там же. — 2002. — № 4. — С. 107–124.
8. Kohonen T. Self-organizing maps. — Berlin: Springer, 2001. — 501 p.

Поступила 06.11.2008