

АНАЛИТИЧЕСКОЕ ИССЛЕДОВАНИЕ СОСТАВЛЯЮЩИХ ОШИБКИ ДЛЯ РЕШЕНИЯ ДИСКРЕТНЫХ НЕКОРРЕКТНЫХ ЗАДАЧ НА ОСНОВЕ СЛУЧАЙНОГО ПРОЕЦИРОВАНИЯ

Аннотация. Проведено аналитическое исследование зависимости составляющих ошибки восстановления истинного сигнала от числа строк матрицы случайного проектора. Показано, что при возрастании размерности случайного проектора детерминированная составляющая ошибки уменьшается, а стохастическая возрастает. Получены выражения для расчета интервала уровней шума, обеспечивающих наличие глобального минимума ошибки. Аналитические результаты подтверждены численным экспериментом.

Ключевые слова: дискретная некорректная задача, регуляризация, случайные проекции, составляющие ошибки, минимум ошибки.

ВВЕДЕНИЕ

Решение некорректных задач актуально для многих областей науки и техники. Дискретные некорректные задачи [1, 2] возникают, например, при дискретизации интегральных уравнений в спектрометрии, гравиметрии, магнитометрии, электроразведке и т.д.

Известно, что решение дискретной некорректной обратной задачи как задачи наименьших квадратов является неустойчивым, т.е. малым изменениям в векторе данных (измерений) соответствуют большие изменения в векторе решения, при этом велико значение ошибки решения. Для преодоления неустойчивости и, следовательно, повышения точности решения дискретных некорректных задач используют методы регуляризации [1–4]. Однако им присущи недостатки: высокая вычислительная сложность и сложность подбора правильного параметра регуляризации, от которого в значительной мере зависит устойчивость решения. В работах [5–14] для устранения проблемы неустойчивости предложен подход, использующий случайное проецирование [15–21], являющееся разновидностью методов формирования нейросетевых распределенных представлений, которые имеют широкий спектр применений [21–33].

Экспериментальное исследование метода решения дискретных некорректных задач на основе проецирования случайной матрицей [5, 8, 10] показало наличие такой оптимальной размерности (числа строк) данной матрицы, при которой ошибка восстановления истинного сигнала близка к минимальной.

В данной статье проводится аналитическое исследование составляющих ошибки восстановления истинного сигнала — стохастической (возрастающей с увеличением размерности матрицы) и детерминированной (убывающей). Для проецирования используется матрица с ортогональными столбцами, полученная путем преобразования исходной случайной матрицы. Исследуется также существование глобального минимума при промежуточных значениях размерности матрицы, с помощью которой проводится проецирование.

ВОССТАНОВЛЕНИЕ СИГНАЛА С ИСПОЛЬЗОВАНИЕМ СЛУЧАЙНЫХ МАТРИЦ

Во многих практических приложениях преобразование сигнала описывается линейной моделью

$$Ax \approx y, \quad (1)$$

где известны матрица $\mathbf{A} \in \mathfrak{R}^{N \times N}$ и вектор $\mathbf{y} \in \mathfrak{R}^N$, искаженный аддитивным шумом $\boldsymbol{\varepsilon} \in \mathfrak{R}^N$: $\mathbf{y} = \mathbf{y}_0 + \boldsymbol{\varepsilon}$. Компоненты вектора шума $\boldsymbol{\varepsilon}$ — реализации независимых нормально распределенных случайных величин. Требуется оценить вектор сигнала $\mathbf{x} \in \mathfrak{R}^N$.

Задача наименьших квадратов для такой оценки записывается в виде

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2.$$

Восстановление сигнала \mathbf{x} на основе псевдообращения \mathbf{A} получается как $\mathbf{x}^* = \mathbf{A}^+ \mathbf{y}$, где \mathbf{A}^+ — псевдообратная матрица [35, 36]. Отметим, что последняя интенсивно используется при исследовании нейросетей [37–40].

В случае, когда \mathbf{y} содержит шум, ряд сингулярных чисел σ_i матрицы \mathbf{A} плавно убывает к нулю, \mathbf{A} имеет большое число обусловленности $\sigma_{\max} / \sigma_{\min}$, задачу оценки \mathbf{x} называют дискретной некорректной обратной задачей [1]. Решение $\mathbf{x}^* = \mathbf{A}^+ \mathbf{y}$ является неустойчивым и неточным.

Далее описан разработанный подход, основанный на отыскании минимума ошибки решения дискретной некорректной задачи с использованием метода случайного проецирования и обеспечивающий устойчивость решения.

Проанализируем поведение ошибки восстановления сигнала \mathbf{x} , полученной в результате применения подхода случайного проецирования.

Зададим матрицу $\mathbf{G} \in \mathfrak{R}^{N \times N}$, элементы которой — реализации случайной величины с нормальным распределением, нулевым средним и единичной дисперсией. Используя SVD (singular value decomposition) [35, 36], вычислим матрицу $\boldsymbol{\Omega}$ как $\mathbf{G} = \boldsymbol{\Omega} \boldsymbol{\Sigma} \boldsymbol{\Psi}^T$, где $\boldsymbol{\Omega} \in \mathfrak{R}^{N \times N}$ и $\boldsymbol{\Psi} \in \mathfrak{R}^{N \times N}$ — соответственно ортонормированные матрицы левых и правых сингулярных векторов, $\boldsymbol{\Sigma} \in \mathfrak{R}^{N \times N}$ — диагональная матрица сингулярных значений.

Матрицу $\boldsymbol{\Omega}_k$ сформируем, выбрав начальные k строк матрицы $\boldsymbol{\Omega}$: $\boldsymbol{\Omega}_k = [\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_k]$, где $\boldsymbol{\omega}_k$ — строка матрицы $\boldsymbol{\Omega}_k$. Отметим, что для данного исследования важна ортогональность векторов $\boldsymbol{\omega}_k$, но не их соответствие наибольшим сингулярным значениям. Здесь и далее использование индекса k означает, что по крайней мере одна из размерностей матриц равна k .

Для решения дискретной некорректной задачи с помощью проекционного подхода обе части исходного уравнения (1) умножим на матрицу $\boldsymbol{\Omega}_k \in \mathfrak{R}^{k \times N}$. Получим уравнение $\mathbf{F}_k \mathbf{x} = \mathbf{b}_k$, где $\mathbf{F}_k = \boldsymbol{\Omega}_k \mathbf{A}$, $\mathbf{F}_k \in \mathfrak{R}^{k \times N}$, $\mathbf{b}_k = \boldsymbol{\Omega}_k \mathbf{y}$, $\mathbf{b}_k \in \mathfrak{R}^k$. Матрица $\mathbf{F}_k = \boldsymbol{\Omega}_k \mathbf{A}$ имеет структуру

$$\mathbf{F}_k = \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{f}_k \end{bmatrix}, \quad (2)$$

где $\mathbf{F}_{k-1} \in \mathfrak{R}^{(k-1) \times N}$, вектор-строка $\mathbf{f}_k = \boldsymbol{\omega}_k \mathbf{A}$ размерностью k .

Вектор восстановленного сигнала $\mathbf{x} \in \mathfrak{R}^N$ с использованием псевдообращения Мура–Пенроуза [35] матрицы \mathbf{F}_k получим как $\mathbf{x}_k^* = \mathbf{F}_k^+ \mathbf{b}_k$.

Точность решения обратной задачи оценим с помощью ошибки e_k восстановления истинного сигнала \mathbf{x} , вычисляемой как квадрат евклидовой нормы вектора ошибки \mathbf{e}_k :

$$e_k = \|\mathbf{x} - \mathbf{x}_k^*\|^2 = \|\mathbf{e}_k\|^2. \quad (3)$$

Выражение для вектора ошибки \mathbf{e}_k представим как:

$$\mathbf{e}_k = \mathbf{x}_k^* - \mathbf{x} = \mathbf{F}_k^+ \mathbf{b}_k - \mathbf{x} = \mathbf{F}_k^+ \mathbf{\Omega}_k (\mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}) - \mathbf{x} = (\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x} + \mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon}.$$

Усредним квадрат евклидовой нормы вектора ошибки $\|\mathbf{e}_k\|^2$ (3) по реализациям шума $\boldsymbol{\varepsilon}$:

$$e_k^2 = \mathbb{E}\{e_k\} = \mathbb{E}\|(\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2 + \mathbb{E}\|\mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon}\|^2 + 2\mathbb{E}\langle (\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}, \mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon} \rangle.$$

Поскольку компоненты вектора шума $\boldsymbol{\varepsilon}$ — независимые нормально распределенные случайные величины, $\mathbb{E}\{\varepsilon_i \varepsilon_j\} = \mathbb{E}\{\varepsilon_i\} \mathbb{E}\{\varepsilon_j\} = 0$, $\mathbb{E}\{\varepsilon_i \varepsilon_i\} = \sigma^2$, математическое ожидание второго слагаемого e_k^2 :

$$\begin{aligned} \mathbb{E}\|\mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon}\|^2 &= \mathbb{E}\{\boldsymbol{\varepsilon}^T \mathbf{\Omega}_k^T \mathbf{F}_k^{+T} \mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon}\} = \mathbb{E}\left\{ \sum_{i=1}^N \sum_{j=1}^N \varepsilon_i [\mathbf{\Omega}_k^T \mathbf{F}_k^{+T} \mathbf{F}_k^+ \mathbf{\Omega}_k]_{ij} \varepsilon_j \right\} = \\ &= \sum_{i=1}^N \sum_{j=1}^N [\mathbf{\Omega}_k^T \mathbf{F}_k^{+T} \mathbf{F}_k^+ \mathbf{\Omega}_k]_{ij} \mathbb{E}\{\varepsilon_i \varepsilon_j\} = \sigma^2 \text{tr}(\mathbf{\Omega}_k^T \mathbf{F}_k^{+T} \mathbf{F}_k^+ \mathbf{\Omega}_k). \end{aligned}$$

Математическое ожидание третьего слагаемого e_k^2 :

$$2\mathbb{E}\langle (\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}, \mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon} \rangle = \mathbb{E}\{\mathbf{x}^T (\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})^T \mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon}\} = \mathbb{E}\{\mathbf{a}^T \boldsymbol{\varepsilon}\},$$

где $\mathbf{a} = \mathbf{x}^T (\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})^T \mathbf{F}_k^+ \mathbf{\Omega}_k$.

Так как $\mathbb{E}\{\varepsilon_i\} = 0$, то $\mathbb{E}\{\mathbf{a}^T \boldsymbol{\varepsilon}\} = \sum_{i=1}^N a_i \mathbb{E}\{\varepsilon_i\} = 0$ и $2\mathbb{E}\langle (\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}, \mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon} \rangle = 0$.

Итоговое выражение для квадрата евклидовой нормы вектора ошибки имеет вид

$$e_k^2 = \|(\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2 + \sigma^2 \text{tr}(\mathbf{\Omega}_k^T \mathbf{F}_k^{+T} \mathbf{F}_k^+ \mathbf{\Omega}_k).$$

Для матрицы $\mathbf{\Omega}_k$ справедливо $\mathbf{\Omega}_k \mathbf{\Omega}_k^T = \mathbf{I}_k$. С учетом свойства цикличности следа произведения матриц переупорядочим выражение $\mathbf{\Omega}_k^T \mathbf{F}_k^{+T} \mathbf{F}_k^+ \mathbf{\Omega}_k$ и получим:

$$e_{s_k}^2 = \mathbb{E}\|\mathbf{F}_k^+ \mathbf{\Omega}_k \boldsymbol{\varepsilon}\|^2 = \sigma^2 \text{tr}(\mathbf{F}_k^{+T} \mathbf{F}_k^+). \quad (4)$$

Составляющую ошибки (4), включающую дисперсию шума, будем называть стохастической, а составляющую

$$e_{d_k}^2 = \|(\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2, \quad (5)$$

не зависящую от шума, — детерминированной.

Исследуем поведение составляющих (4), (5) ошибки решения дискретной некорректной задачи в зависимости от числа k строк матрицы $\mathbf{\Omega}_k$.

Исследование стохастической составляющей ошибки. В целях исследования поведения стохастической составляющей ошибки (4) в зависимости от k запишем выражение для получения \mathbf{F}_k^+ в рекурсивном виде. Для этого используем представление возмущения псевдообратной матрицы через возмущение исходной матрицы [41]

$$\mathbf{B}^+ - \mathbf{A}^+ = -\mathbf{B}^+ \mathbf{P}_B \mathbf{E} \mathbf{R}_A \mathbf{A}^+ + \mathbf{B}^+ \mathbf{P}_B \mathbf{P}_A^\perp - \mathbf{R}_B^\perp \mathbf{R}_A \mathbf{A}^+, \quad (6)$$

где $\mathbf{B} = \mathbf{A} + \mathbf{E}$, \mathbf{E} — матрица возмущения; $\mathbf{P}_A = \mathbf{A} \mathbf{A}^+$ и $\mathbf{R}_A = \mathbf{A}^+ \mathbf{A}$ — проекторы [41] на подпространство соответственно вектор-столбцов и вектор-строк

матрицы \mathbf{A} ; $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{P}_A$ и $\mathbf{R}_A^\perp = \mathbf{I} - \mathbf{R}_A$ — проекторы на ортогональное дополнение к данным подпространствам соответственно.

Отметим, что для проекторов как идемпотентных матриц $\mathbf{P}^2 = \mathbf{P}$. Покажем, что \mathbf{P}_A , \mathbf{R}_A , \mathbf{P}_A^\perp и \mathbf{R}_A^\perp действительно являются проекторами: используя сингулярное разложение $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^\top$, $\mathbf{A} \in \mathfrak{R}^{k \times N}$, $\mathbf{U} \in \mathfrak{R}^{k \times k}$, $\mathbf{S} \in \mathfrak{R}^{k \times N}$ (диагональная), $\mathbf{V} \in \mathfrak{R}^{N \times N}$, запишем

$$\begin{aligned} \mathbf{P}_A &= \mathbf{A}\mathbf{A}^+ = \mathbf{U}\mathbf{S}\mathbf{V}^\top\mathbf{V}\mathbf{S}^+\mathbf{U}^\top, \quad \mathbf{V}^\top\mathbf{V} = \mathbf{I}_N, \quad \mathbf{S}\mathbf{I}_N\mathbf{S}^+ = \mathbf{I}_k, \\ \mathbf{U}\mathbf{I}_k\mathbf{U}^\top &= \mathbf{I}_k, \quad \mathbf{P}_A^2 = \mathbf{I}_k\mathbf{I}_k, \end{aligned} \quad (7)$$

где $\mathbf{I}_N \in \mathfrak{R}^{N \times N}$, $\mathbf{I}_k \in \mathfrak{R}^{k \times k}$ — единичные матрицы;

$$\begin{aligned} \mathbf{R}_A &= \mathbf{A}^+\mathbf{A} = \mathbf{V}\mathbf{S}^+\mathbf{U}^\top\mathbf{U}\mathbf{S}\mathbf{V}^\top, \quad \mathbf{U}^\top\mathbf{U} = \mathbf{I}_k, \quad \mathbf{S}^+\mathbf{I}_k\mathbf{S} = \begin{bmatrix} \mathbf{I}_k & \mathbf{O}_{k \times N-k} \\ \mathbf{O}_{N-k \times k} & \mathbf{O}_{N-k \times N-k} \end{bmatrix}, \\ \mathbf{R}_A &= \mathbf{V} \begin{bmatrix} \mathbf{I}_k & \mathbf{O}_{k \times N-k} \\ \mathbf{O}_{N-k \times k} & \mathbf{O}_{N-k \times N-k} \end{bmatrix} \mathbf{V}^\top, \end{aligned} \quad (8)$$

$$\begin{aligned} \mathbf{R}_A^2 &= \mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}^+\mathbf{I}_k\mathbf{A} = \mathbf{V}\mathbf{S}^+\mathbf{U}^\top\mathbf{I}_k\mathbf{U}\mathbf{S}\mathbf{V}^\top, \quad \mathbf{U}^\top\mathbf{I}_k\mathbf{U} = \mathbf{I}_k, \quad \mathbf{R}_A^2 = \mathbf{V}\mathbf{S}^+\mathbf{I}_k\mathbf{S}\mathbf{V}^\top, \\ \mathbf{R}_A^2 &= \mathbf{V} \begin{bmatrix} \mathbf{I}_k & \mathbf{O}_{k \times N-k} \\ \mathbf{O}_{N-k \times k} & \mathbf{O}_{N-k \times N-k} \end{bmatrix} \mathbf{V}^\top, \quad \mathbf{R}_A = \mathbf{R}_A^2, \end{aligned} \quad (9)$$

где \mathbf{O} — нулевая матрица, индекс которой соответствует размерности;

$$\begin{aligned} \mathbf{R}_A^{\perp 2} &= (\mathbf{I} - \mathbf{R}_A)(\mathbf{I} - \mathbf{R}_A) = \mathbf{I}^2 - 2\mathbf{R}_A + \mathbf{R}_A^2, \quad \mathbf{R}_A^2 = \mathbf{R}_A, \\ \mathbf{R}_A^{\perp 2} &= \mathbf{I}^2 - 2\mathbf{R}_A + \mathbf{R}_A = \mathbf{I} - \mathbf{R}_A, \quad \mathbf{R}_A^{\perp 2} = \mathbf{R}_A^\perp. \end{aligned} \quad (10)$$

Идемпотентность \mathbf{P}_A^\perp , а именно $\mathbf{P}_A^{\perp 2} = \mathbf{P}_A^\perp$, обосновывается рассуждениями, аналогичными (10).

В качестве возмущения матрицы \mathbf{F}_{k-1} будем рассматривать матрицу $\mathbf{E}_k \in \mathfrak{R}^{k \times N}$, содержащую одну ненулевую k -ю строку, которая добавляется на k -м шаге (2):

$$\mathbf{F}_{k-1} + \mathbf{E}_k = \mathbf{F}_k, \quad (11)$$

т.е. матрица возмущения формируется как

$$\mathbf{E}_k = \begin{bmatrix} \mathbf{O}_{k-1} \\ \mathbf{f}_k \end{bmatrix}, \quad (12)$$

где $\mathbf{f}_k = \omega_k \mathbf{A}$, ω_k — строка матрицы $\mathbf{\Omega}_k = [\omega_1, \dots, \omega_k]$, \mathbf{O}_{k-1} — нулевая подматрица размера $((k-1) \times N)$.

Невозмущенной матрице \mathbf{A} в (6) сопоставим $\begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} \in \mathfrak{R}^{k \times N}$, а возмущенной

матрице \mathbf{B} сопоставим матрицу $\mathbf{F}_k = \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{f}_k \end{bmatrix}$, тогда \mathbf{A}^+ и \mathbf{B}^+ в (6) будут соответствовать $[\mathbf{F}_{k-1}^+; \mathbf{0}] \in \mathfrak{R}^{N \times k}$ и $\mathbf{F}_k^+ \in \mathfrak{R}^{N \times k}$, где $\mathbf{0} \in \mathfrak{R}^N$ — нулевой вектор-столбец, а $\mathbf{0}^\top \in \mathfrak{R}^N$ — вектор-строка.

Матрице \mathbf{P}_B в (6) сопоставим $\mathbf{F}_k\mathbf{F}_k^+ \in \mathfrak{R}^{k \times k}$. Матрица $\mathbf{\Omega}_k$ имеет ортонор-

мальные строки, а поскольку последние линейно независимы, ранг этой матрицы равен k . Для невырожденной квадратной матрицы \mathbf{B} ранг $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{A})$ [35], для $\mathbf{F}_k = \mathbf{\Omega}_k \mathbf{A}$ ранг $\text{rank}(\mathbf{F}_k) = \text{rank}(\mathbf{\Omega}_k \mathbf{A}) = k$, отсюда следует линейная независимость строк матрицы $\mathbf{F}_k \in \mathfrak{R}^{k \times N}$. Для матрицы с линейно независимыми строками [35] $\mathbf{F}_k \mathbf{F}_k^+ = \mathbf{I}_k$, что будет использовано далее.

Матрицам $\mathbf{R}_A \in \mathfrak{R}^{N \times N}$, $\mathbf{P}_A^\perp \in \mathfrak{R}^{k \times k}$ и $\mathbf{R}_B^\perp \in \mathfrak{R}^{N \times N}$ в (6) сопоставим соответственно

$$\mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}, \quad (13)$$

$$\mathbf{I}_k - \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}], \quad (14)$$

$$\mathbf{I} - \mathbf{F}_k^+ \mathbf{F}_k. \quad (15)$$

Подставив в (6) матрицы (13)–(15), получим рекурсивное выражение для \mathbf{F}_k^+ :

$$\begin{aligned} \mathbf{F}_k^+ &= [\mathbf{F}_{k-1}^+; \mathbf{0}] - [\mathbf{F}_k^+ \mathbf{E}_k \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} \mathbf{F}_{k-1}^+; \mathbf{0}] + \\ &+ \mathbf{F}_k^+ \mathbf{F}_k \mathbf{F}_k^+ \left(\mathbf{I}_k - \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] \right) - (\mathbf{I} - \mathbf{F}_k^+ \mathbf{F}_k) \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} [\mathbf{F}_{k-1}^+; \mathbf{0}]. \end{aligned} \quad (16)$$

В целях упрощения (16) рассмотрим его слагаемые. Для части третьего слагаемого в (16) произведение

$$\mathbf{F}_k \mathbf{F}_k^+ \left(\mathbf{I}_k - \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] \right) \in \mathfrak{R}^{k \times k} \quad (17)$$

представляет собой матрицу, у которой элемент с индексом (k, k) равен единице, а остальные — нулю, так как $\mathbf{F}_k \mathbf{F}_k^+ = \mathbf{I}_k$ и

$$\mathbf{I}_k - \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] = \mathbf{I}_k - \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}_{k-1} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (18)$$

где $\mathbf{\Theta}_{k-1}$ — нулевая подматрица $((k-1) \times (k-1))$.

Умножая слева (17) на \mathbf{F}_k^+ , получаем

$$\mathbf{F}_k^+ \mathbf{F}_k \mathbf{F}_k^+ \left(\mathbf{I}_k - \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] \right) = \mathbf{F}_k^+ \begin{bmatrix} \mathbf{\Theta}_{k-1} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{bmatrix} = [\mathbf{0}_{k-1}^\top; \mathbf{t}_k], \quad (19)$$

где вектор-столбец \mathbf{t}_k — последний столбец матрицы \mathbf{F}_k^+ .

Рассмотрим четвертое слагаемое в (16)

$$(\mathbf{I} - \mathbf{F}_k^+ \mathbf{F}_k) [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} = [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} - \mathbf{F}_k^+ \mathbf{F}_k [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix}. \quad (20)$$

Преобразуем $\mathbf{F}_k [\mathbf{F}_{k-1}^+; \mathbf{0}]$ из второго слагаемого (20):

$$\begin{aligned} \mathbf{F}_k [\mathbf{F}_{k-1}^+; \mathbf{0}] &= \left(\begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} + \mathbf{E}_k \right) [\mathbf{F}_{k-1}^+; \mathbf{0}] = \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^\top \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] + \mathbf{E}_k [\mathbf{F}_{k-1}^+; \mathbf{0}] = \\ &= \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} + \begin{bmatrix} \mathbf{O}_{k-1} \\ \mathbf{f}_k \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} + \begin{bmatrix} \mathbf{O}_{k-1} & \mathbf{0} \\ \mathbf{f}_k \mathbf{F}_{k-1}^+ & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{f}_k \mathbf{F}_{k-1}^+ & 0 \end{bmatrix}. \end{aligned}$$

В [42] показано, что $\mathbf{F}_k^+ \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{f}_k \mathbf{F}_{k-1}^+ & 0 \end{bmatrix} = [\mathbf{F}_{k-1}^+; \mathbf{0}]$, поэтому второе слагаемое

(20) есть

$$\mathbf{F}_k^+ \mathbf{F}_k [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix} = \mathbf{F}_k^+ \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{f}_k \mathbf{F}_{k-1}^+ & 0 \end{bmatrix} = [\mathbf{F}_{k-1}^+; \mathbf{0}],$$

и, учитывая, что

$$\begin{aligned} & [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix} - \mathbf{F}_k^+ \mathbf{F}_k [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix} = \\ & = [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix} - [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix} = \mathbf{0}_N, \end{aligned}$$

получаем

$$(\mathbf{I} - \mathbf{F}_k^+ \mathbf{F}_k) [\mathbf{F}_{k-1}^+; \mathbf{0}] \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix} [\mathbf{F}_{k-1}^+; \mathbf{0}] = \mathbf{0}_N, \quad (21)$$

т.е. четвертое слагаемое в (16) есть нулевая матрица.

Подставив в (16) выражение (19), исключив выражение (21) и обозначив второе слагаемое (16) как

$$[\mathbf{M}_{k-1}; \mathbf{0}] \equiv \mathbf{F}_k^+ \mathbf{E}_k \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} [\mathbf{F}_{k-1}^+; \mathbf{0}], \quad (22)$$

получим

$$\mathbf{F}_k^+ = [\mathbf{F}_{k-1}^+; \mathbf{0}] - [\mathbf{M}_{k-1}; \mathbf{0}] + [\mathbf{O}_{k-1}^T; \mathbf{t}_k]. \quad (23)$$

Тогда

$$\mathbf{F}_k^{+T} \mathbf{F}_k^+ = ([\mathbf{F}_{k-1}^+; \mathbf{0}] - [\mathbf{M}_{k-1}; \mathbf{0}] + [\mathbf{O}_{k-1}^T; \mathbf{t}_k])^T ([\mathbf{F}_{k-1}^+; \mathbf{0}] - [\mathbf{M}_{k-1}; \mathbf{0}] + [\mathbf{O}_{k-1}^T; \mathbf{t}_k]). \quad (24)$$

Раскрываем скобки, учитывая, что произведение $[\mathbf{O}_{k-1}^T; \mathbf{t}_k]^T [\mathbf{F}_{k-1}^+; \mathbf{0}]$, как и $[\mathbf{F}_{k-1}^+; \mathbf{0}]^T [\mathbf{O}_{k-1}^T; \mathbf{t}_k]$, есть нулевая матрица (что следует из правила умножения матриц).

Рассмотрим

$$\begin{aligned} & [\mathbf{M}_{k-1}; \mathbf{0}]^T [\mathbf{F}_{k-1}^+; \mathbf{0}] = (\mathbf{F}_k^+ \mathbf{E}_k \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} [\mathbf{F}_{k-1}^+; \mathbf{0}])^T [\mathbf{F}_{k-1}^+; \mathbf{0}] = \\ & = [\mathbf{F}_{k-1}^+; \mathbf{0}]^T \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix}^T [\mathbf{F}_{k-1}^+; \mathbf{0}]^T \mathbf{E}_k^T \mathbf{F}_k^{+T} [\mathbf{F}_{k-1}^+; \mathbf{0}]. \end{aligned} \quad (25)$$

Поскольку $[\mathbf{F}_{k-1}^+; \mathbf{0}]^T \begin{bmatrix} \mathbf{F}_{k-1} \\ \mathbf{0}^T \end{bmatrix}^T = \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}$, выражение (25) равно $[\mathbf{F}_{k-1}^+; \mathbf{0}]^T \mathbf{E}_k^T \times \mathbf{F}_k^{+T} [\mathbf{F}_{k-1}^+; \mathbf{0}]$.

Матрица $[\mathbf{F}_{k-1}^+; \mathbf{0}]^T = \begin{bmatrix} \mathbf{F}_{k-1}^{+T} \\ \mathbf{0} \end{bmatrix}$ (k -я строка нулевая) умножается на \mathbf{E}_k^T (12), где

только k -й столбец ненулевой, поэтому результат есть нулевая матрица, следовательно, и (25) — нулевая матрица.

Рассмотрим $[\mathbf{F}_{k-1}^+; \mathbf{0}]^T [\mathbf{M}_{k-1}; \mathbf{0}]$ в (24). При вычислении

$$[\mathbf{F}_{k-1}^+; \mathbf{0}]^T [\mathbf{M}_{k-1}; \mathbf{0}] = [\mathbf{F}_{k-1}^+; \mathbf{0}]^T \mathbf{F}_k^+ \mathbf{E}_k \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} [\mathbf{F}_{k-1}^+; \mathbf{0}]$$

матрица $[\mathbf{F}_{k-1}^+; \mathbf{0}] \mathbf{F}_k^+$ (k -й столбец нулевой), умноженная на \mathbf{E}_k (12), где только k -я строка ненулевая, дает нулевую матрицу. Следовательно, произведение $[\mathbf{F}_{k-1}^+; \mathbf{0}]^T [\mathbf{M}_{k-1}; \mathbf{0}]$ в (24) есть нулевая матрица.

Исключая нулевые матрицы-слагаемые в (24), получаем

$$\begin{aligned} \mathbf{F}_k^{+\text{T}} \mathbf{F}_k^+ &= [\mathbf{F}_{k-1}^+ \vdots \mathbf{0}]^{\text{T}} [\mathbf{F}_{k-1}^+ \vdots \mathbf{0}] + [\mathbf{M}_{k-1} \vdots \mathbf{0}]^{\text{T}} [\mathbf{M}_{k-1} \vdots \mathbf{0}] - [\mathbf{M}_{k-1} \vdots \mathbf{0}]^{\text{T}} [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k] - \\ &\quad - [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k]^{\text{T}} [\mathbf{M}_{k-1} \vdots \mathbf{0}] + [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k]^{\text{T}} [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k]. \end{aligned} \quad (26)$$

Рассмотрим структуру матриц, образованных слагаемыми выражения (26).

Сумма первых двух слагаемых дает матрицу размера $k \times k$, у которой подматрица $\mathbf{B}_k \in \mathfrak{R}^{(k-1) \times (k-1)}$ окаймлена нулями:

$$[\mathbf{F}_{k-1}^+ \vdots \mathbf{0}]^{\text{T}} [\mathbf{F}_{k-1}^+ \vdots \mathbf{0}] + [\mathbf{M}_{k-1} \vdots \mathbf{0}]^{\text{T}} [\mathbf{M}_{k-1} \vdots \mathbf{0}] = \begin{bmatrix} \mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+ + \mathbf{M}_{k-1}^{\text{T}} \mathbf{M}_{k-1} & \mathbf{0}_{k-1} \\ \mathbf{0}_{k-1}^{\text{T}} & 0 \end{bmatrix}. \quad (27)$$

Отметим, что матрица $\mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+ + \mathbf{M}_{k-1}^{\text{T}} \mathbf{M}_{k-1}$ в (27) имеет неотрицательные элементы на диагонали, так как диагонали $\mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+$ и $\mathbf{M}_{k-1}^{\text{T}} \mathbf{M}_{k-1}$ неотрицательны.

Третье и четвертое слагаемые выражения (26) есть матрицы следующей структуры:

$$[\mathbf{M}_{k-1} \vdots \mathbf{0}]^{\text{T}} [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k] = \begin{bmatrix} \Theta_{k-1} & \mathbf{c}_{k-1} \\ \mathbf{0}_{k-1}^{\text{T}} & 0 \end{bmatrix}, \quad [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k]^{\text{T}} [\mathbf{M}_{k-1} \vdots \mathbf{0}] = \begin{bmatrix} \Theta_{k-1} & \mathbf{0}_{k-1} \\ \mathbf{c}_{k-1}^{\text{T}} & 0 \end{bmatrix},$$

где $\mathbf{c}_{k-1} \in \mathfrak{R}^{k-1}$, $\Theta_{k-1} \in \mathfrak{R}^{(k-1) \times (k-1)}$.

Пятое слагаемое выражения (26) есть матрица

$$[\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k]^{\text{T}} [\mathbf{O}_{k-1}^{\text{T}} \vdots \mathbf{t}_k] = \begin{bmatrix} \Theta_{k-1} & \mathbf{0}_{k-1} \\ \mathbf{0}_{k-1}^{\text{T}} & d_k \end{bmatrix}, \quad (28)$$

где $d_k = \mathbf{t}_k^{\text{T}} \mathbf{t}_k \geq 0$.

Таким образом, матрицу $\mathbf{F}_k^{+\text{T}} \mathbf{F}_k^+$ (24), (26) можно представить в виде

$$\mathbf{F}_k^{+\text{T}} \mathbf{F}_k^+ = \begin{bmatrix} \mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+ + \mathbf{M}_{k-1}^{\text{T}} \mathbf{M}_{k-1} & \mathbf{c}_{k-1} \\ \mathbf{c}_{k-1}^{\text{T}} & d_k \end{bmatrix}. \quad (29)$$

Из выражений (29), а также (27) и (28) следует, что матрица $\mathbf{F}_k^{+\text{T}} \mathbf{F}_k^+$ имеет неотрицательную диагональ. Заметим, что $\mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+ = \begin{bmatrix} \mathbf{F}_{k-2}^{+\text{T}} \mathbf{F}_{k-2}^+ + \mathbf{M}_{k-2}^{\text{T}} \mathbf{M}_{k-2} & \mathbf{c}_{k-2} \\ \mathbf{c}_{k-2}^{\text{T}} & d_{k-1} \end{bmatrix}$.

Значит, значение стохастической составляющей ошибки (4)

$$\begin{aligned} \sigma^2 \text{tr}(\mathbf{F}_k^{+\text{T}} \mathbf{F}_k^+) &= \sigma^2 \text{tr} \begin{bmatrix} \mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+ + \mathbf{M}_{k-1}^{\text{T}} \mathbf{M}_{k-1} & \mathbf{c}_{k-1} \\ \mathbf{c}_{k-1}^{\text{T}} & d_k \end{bmatrix} = \\ &= \sigma^2 \text{tr}(\mathbf{F}_{k-1}^{+\text{T}} \mathbf{F}_{k-1}^+) + \sigma^2 \text{tr}(\mathbf{M}_{k-1}^{\text{T}} \mathbf{M}_{k-1}) + \sigma^2 d_k \end{aligned}$$

возрастает с увеличением k .

Исследование детерминированной составляющей ошибки. Для исследования зависимости величины детерминированной составляющей ошибки (5) $\|(\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2$ от числа строк матрицы $\mathbf{\Omega}_k$ используем рекурсивное выражение матрицы $\mathbf{F}_k^+ \mathbf{F}_k$, представив квадрат нормы как

$$e_{d_k}^2 = \|(\mathbf{F}_k^+ \mathbf{F}_k - \mathbf{I})\mathbf{x}\|^2 = (\mathbf{F}_k^+ \mathbf{F}_k \mathbf{x} - \mathbf{x})^{\text{T}} (\mathbf{F}_k^+ \mathbf{F}_k \mathbf{x} - \mathbf{x}). \quad (30)$$

Матрица $\mathbf{F}_k^+ \mathbf{F}_k$ является матрицей проектора, что показано в (9), для которого справедливо $(\mathbf{F}_k^+ \mathbf{F}_k)^T \mathbf{F}_k^+ \mathbf{F}_k = \mathbf{F}_k^+ \mathbf{F}_k$. Поэтому из (30) получим

$$e_{d_k}^2 = \mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x} - 2\mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x} + \mathbf{x}^T \mathbf{x} = \mathbf{x}^T \mathbf{x} - \mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x}.$$

Используя (23) и (11), запишем рекурсивное выражение для $\mathbf{F}_k^+ \mathbf{F}_k$:

$$\mathbf{F}_k^+ \mathbf{F}_k = ([\mathbf{F}_{k-1}^+; \mathbf{0}] - [\mathbf{M}_{k-1}; \mathbf{0}] + [\mathbf{O}_{k-1}^T; \mathbf{t}_k])([\mathbf{F}_{k-1}; \mathbf{0}]^T + \mathbf{E}_k^T). \quad (31)$$

Раскроем скобки в выражении (31). Учитывая, что $[\mathbf{O}_{k-1}^T; \mathbf{t}_k][\mathbf{F}_{k-1}; \mathbf{0}]^T = 0$,

$$[\mathbf{F}_{k-1}^+; \mathbf{0}]\mathbf{E}_k^T = [\mathbf{F}_{k-1}^+; \mathbf{0}][\mathbf{O}_{k-1}; \mathbf{f}_k]^T,$$

$$[\mathbf{M}_{k-1}; \mathbf{0}]\mathbf{E}_k = [\mathbf{M}_{k-1}; \mathbf{0}][\mathbf{O}_{k-1}; \mathbf{f}_k]^T = 0,$$

получаем выражение

$$\mathbf{F}_k^+ \mathbf{F}_k = \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} + [\mathbf{O}_{k-1}^T; \mathbf{t}_k]\mathbf{E}_k - [\mathbf{M}_{k-1}; \mathbf{0}][\mathbf{F}_{k-1}; \mathbf{0}]. \quad (32)$$

Обозначим

$$\mathbf{D}_k \equiv [\mathbf{O}_{k-1}^T; \mathbf{t}_k]\mathbf{E}_k - [\mathbf{M}_{k-1}; \mathbf{0}][\mathbf{F}_{k-1}; \mathbf{0}]. \quad (33)$$

Учитывая (19), в первое слагаемое (33) вместо $[\mathbf{O}_{k-1}^T; \mathbf{t}_k]$ подставляем

$\mathbf{F}_k^+ \begin{bmatrix} \mathbf{O}_{k-1}^T & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$ и, преобразуя второе слагаемое с учетом (22), получаем

$$\mathbf{D}_k = \mathbf{F}_k^+ \begin{bmatrix} \mathbf{O}_{k-1}^T & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{E}_k - \mathbf{F}_k^+ \mathbf{E}_k \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}. \quad (34)$$

Учитывая, что во втором слагаемом (34) $(\mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})(\mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}$ (согласно свойству проектора $\mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}$), и заменяя в соответствии с (18)

$\begin{bmatrix} \mathbf{O}_{k-1}^T & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} = \mathbf{I}_k - \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}$, можем преобразовать (34) следующим образом:

$$\mathbf{D}_k = \mathbf{F}_k^+ \mathbf{I}_k \mathbf{E}_k - \mathbf{F}_k^+ \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \mathbf{E}_k - \mathbf{F}_k^+ \mathbf{E}_k \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}. \quad (35)$$

Второе слагаемое в (35)

$$\mathbf{F}_k^+ \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \mathbf{E}_k = \mathbf{F}_k^+ \begin{bmatrix} \mathbf{I}_{k-1} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} [\mathbf{O}_{k-1}; \mathbf{f}_k]^T = 0,$$

поэтому

$$\mathbf{D}_k = \mathbf{F}_k^+ \mathbf{E}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}). \quad (36)$$

Поскольку матрица возмущения \mathbf{E}_k (12) имеет только одну ненулевую строку, произведение $\mathbf{E}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = [\mathbf{O}_{k-1}; \mathbf{f}_k]^T (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})$ в (36) представляет собой матрицу $[\mathbf{O}_{k-1}; \mathbf{m}_k]^T$, где $\mathbf{m}_k = \mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})$. Поэтому можно записать

$$\mathbf{F}_k^+ \mathbf{E}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = [\mathbf{O}_{k-1}; \mathbf{f}_k]^T \begin{bmatrix} \mathbf{O}_{k-1} \\ \mathbf{m}_k \end{bmatrix} = \mathbf{f}_k^+ \mathbf{m}_k. \quad (37)$$

В работе [42] приведено выражение для последнего столбца (\mathbf{f}_k^+) матрицы \mathbf{F}_k^+ :

$$\mathbf{f}_k^+ = \frac{(\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T}{\mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T}, \quad (38)$$

где $\mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T \in \mathfrak{R}^1$. Используя (38) и (37), запишем

$$\mathbf{D}_k = \mathbf{F}_k^+ \mathbf{E}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = \mathbf{f}_k^+ \mathbf{m}_k = \frac{(\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T \mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})}{\mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T}. \quad (39)$$

Используя (8) и представляя \mathbf{I} как \mathbf{VIV}^T , получаем

$$\begin{aligned} \mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} &= \mathbf{VIV}^T - \mathbf{V} \begin{bmatrix} \mathbf{I}_k & \mathbf{O}_{k \times N-k} \\ \mathbf{O}_{N-k \times k} & \mathbf{O}_{N-k \times N-k} \end{bmatrix} \mathbf{V}^T = \\ &= \mathbf{V} \left(\mathbf{I} - \begin{bmatrix} \mathbf{I}_k & \mathbf{O}_{k \times N-k} \\ \mathbf{O}_{N-k \times k} & \mathbf{O}_{N-k \times N-k} \end{bmatrix} \right) \mathbf{V}^T = \mathbf{V} \begin{bmatrix} \mathbf{O}_k & \mathbf{O}_{k \times N-k} \\ \mathbf{O}_{N-k \times k} & \mathbf{I}_{N-k \times N-k} \end{bmatrix} \mathbf{V}^T. \end{aligned}$$

Из последнего выражения видно, что $\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}$ — симметричная матрица, поэтому $(\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})^T$. Как следует из (10), $(\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})^2$. Поэтому можно записать

$$z_k = \mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T = \mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})^T \mathbf{f}_k^T = \mathbf{m}_k \mathbf{m}_k^T \quad (40)$$

и $z_k \geq 0$.

Учитывая (39) и (40), получаем

$$\mathbf{D}_k = \frac{(\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) \mathbf{f}_k^T \mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})}{\mathbf{f}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1})^T \mathbf{f}_k^T} = \frac{\mathbf{m}_k^T \mathbf{m}_k}{\mathbf{m}_k \mathbf{m}_k^T} = \frac{\mathbf{m}_k^T \mathbf{m}_k}{z_k}. \quad (41)$$

Из (41) следует, что \mathbf{D}_k — квадратная матрица, симметричная относительно диагонали.

Преобразуем (32) с учетом (33) и (36) как

$$\mathbf{F}_k^+ \mathbf{F}_k = \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} + \mathbf{F}_k^+ \mathbf{E}_k (\mathbf{I} - \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1}) = \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} + \mathbf{D}_k.$$

Итак, для детерминированной составляющей ошибки (5)

$$e_{d_k}^2 = \mathbf{x}^T \mathbf{x} - \mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x} = \mathbf{x}^T \mathbf{x} - (\mathbf{x}^T \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} \mathbf{x} + \mathbf{x}^T \mathbf{D}_k \mathbf{x}).$$

Используя (41) и неотрицательность z , запишем

$$\mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x} = \mathbf{x}^T \mathbf{F}_{k-1}^+ \mathbf{F}_{k-1} \mathbf{x} + \mathbf{x}^T \frac{\mathbf{m}_k^T \mathbf{m}_k}{z_k^{1/2} z_k^{1/2}} \mathbf{x}. \quad (42)$$

Второе слагаемое в (42) — положительное число, поэтому значение $\mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x}$ возрастает с увеличением k и, следовательно, значение детерминированной составляющей ошибки (где возрастающее значение $\mathbf{x}^T \mathbf{F}_k^+ \mathbf{F}_k \mathbf{x}$ вычитается из константы $\mathbf{x}^T \mathbf{x}$) убывает.

Поведение $\mathbf{F}_k^+ \mathbf{F}_k$ при $k \neq N$ следует из (42). При $k = N$, как следует из (8), $\mathbf{F}_N^+ \mathbf{F}_N = \mathbf{V}\mathbf{V}^T = \mathbf{I}$, поэтому $e_{d_N}^2 = \mathbf{x}^T \mathbf{x} - \mathbf{x}^T \mathbf{F}_N^+ \mathbf{F}_N \mathbf{x} = \mathbf{x}^T \mathbf{x} - \mathbf{x}^T \mathbf{I} \mathbf{x} = 0$, т.е. детерминированная составляющая ошибки убывает до нуля.

Глобальный минимум ошибки. Детерминированная составляющая ошибки убывает, а стохастическая возрастает с увеличением k на интервале $k = 1, \dots, N$. Однако зависимости составляющих ошибки от k не являются выпуклыми, поэтому функция ошибки может не быть выпуклой и не обеспечивать наличие глобального минимума при $2 \leq k \leq N - 1$.

Запишем условия того, что значение функции $e_s(k) + e_d(k)$ при $k = k_0$ имеет глобальный минимум:

$$\sigma^2 e_g(k_0 + i) + e_d(k_0 + i) > \sigma^2 e_g(k_0) + e_d(k_0), \quad (43)$$

где $i = 1, \dots, N - k_0$;

$$\sigma^2 e_g(k_0 - j) + e_d(k_0 - j) > \sigma^2 e_g(k_0) + e_d(k_0), \quad (44)$$

где $j = 1, \dots, k_0 - 1$ и $e_g = \|\mathbf{F}_k^+ \mathbf{\Omega}_k \mathbf{\varepsilon}_0\|^2$, $\mathbf{\varepsilon}_0$ — центрированный вектор шума, сгенерированный с единичной дисперсией и $e_s = \sigma^2 e_g$.

Из условий глобального минимума (43), (44) получаем

$$\sigma^2 > (e_d(k_0) - e_d(k_0 + i)) / (e_g(k_0 + i) - e_g(k_0)) \text{ для } i = 1, \dots, N - k_0,$$

$$\sigma^2 < (e_d(k_0) - e_d(k_0 - j)) / (e_g(k_0 - j) - e_g(k_0)) \text{ для } j = 1, \dots, k_0 - 1.$$

Обозначим

$$\sigma_{kR}^2 = (e_d(k_0) - e_d(k_0 + i)) / (e_g(k_0 + i) - e_g(k_0)) \text{ для } i = 1, \dots, N - k_0,$$

$$\sigma_{kL}^2 = (e_d(k_0) - e_d(k_0 - j)) / (e_g(k_0 - j) - e_g(k_0)) \text{ для } j = 1, \dots, k_0 - 1.$$

Граничное значение дисперсии шума σ_R^2 для (43) (при его превышении значение ошибки минимально в k_0): $\sigma_R^2 = \max_{i=1, \dots, N-k_0} \sigma_{kR}^2$. Граничное значение дисперсии шума σ_L^2 для (44) (при значениях дисперсии меньших σ_L^2 значение ошибки минимально в k_0): $\sigma_L^2 = \min_{j=1, \dots, k_0-1} \sigma_{kL}^2$. Если существуют значения σ^2 , удовлетворяющие неравенствам $\sigma_R^2 < \sigma^2 < \sigma_L^2$, то в точке k_0 существует глобальный минимум ошибки восстановления сигнала.

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

Рассмотрим поведение составляющих ошибки и общей ошибки восстановления сигнала в зависимости от размерности матрицы $\mathbf{\Omega}_k$ для известных дискретных некорректных задач Carasso, Phillips, Delves [43] (рис. 1–3).

Задача Carasso [43] — реконструкция временного профиля источника тепла на основе измерения температуры с фиксированного расстояния. Здесь обратное уравнение теплового потока — интегральное уравнение Вольтерра первого рода $\int_a^x K(x, t) f(t) dt = g(x)$ с интервалом интегрирования $[0, 1]$ и ядром $K(s, t) = f(s - t)$, $t^{3/2} / (2\pi^{1/2} \exp(-1/4t))$. Интегральное уравнение дискретизируется методом квадратур с использованием простой коллокации и правила средней точки с N точками. Конструируется точное решение \mathbf{x}_0 , затем \mathbf{b}_0 вычисляется как $\mathbf{b}_0 = \mathbf{F}\mathbf{x}_0$, т.е. в этой задаче $\mathbf{F}\mathbf{x}_0$ точно равно \mathbf{b}_0 .

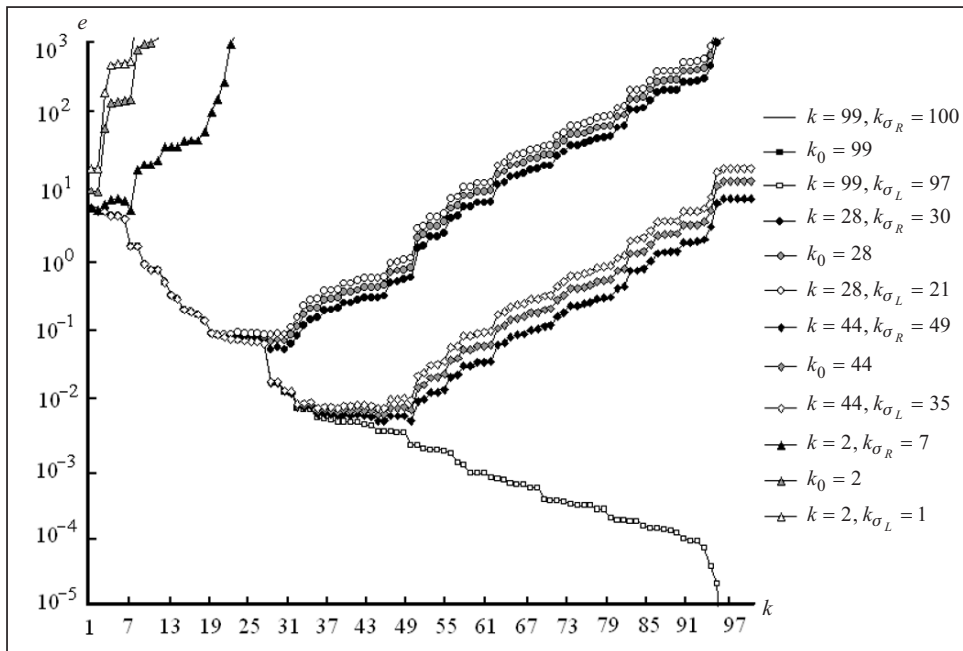


Рис. 1. Зависимость ошибки решения и ее составляющих от k для задачи Carasso при $k_0 = \{2, 28, 44, 99\}$

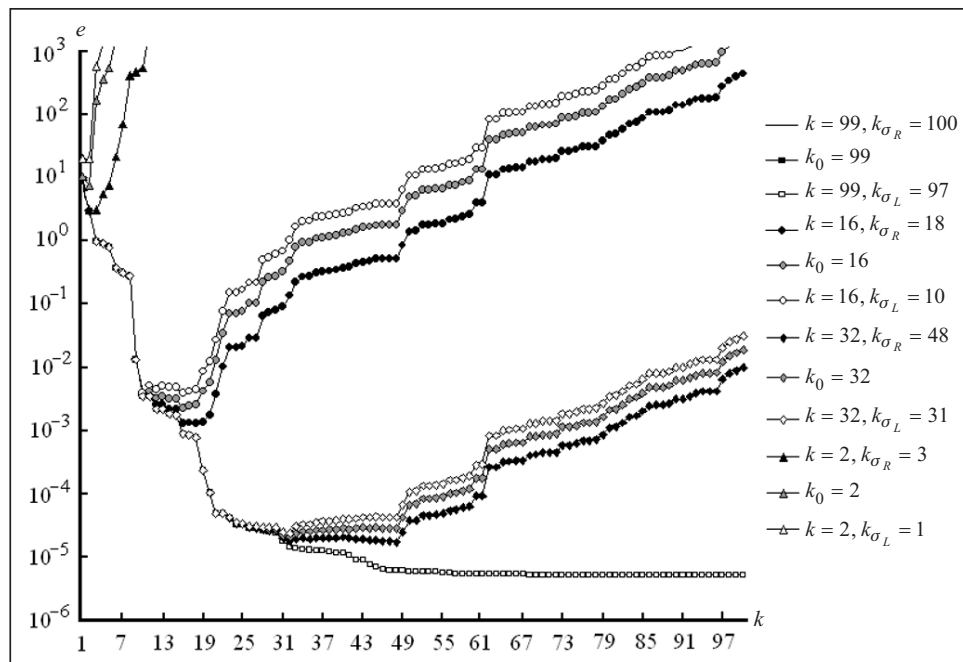


Рис. 2. Зависимость ошибки решения и ее составляющих от k для задачи Phillips при $k_0 = \{2, 16, 32, 99\}$

Задача Phillips [43] — решение уравнения Фредгольма первого рода $\int_a^b K(s, t)f(t)dt = g(s)$, где f — решение, K — ядро и правая часть g имеют вид соответственно $f(t) = 1 + \cos(\pi t/3)$ при $|t| < 3$, $f(t) = 0$ при $|t| \geq 3$; $K(s, t) = f(s-t)$; $g(s) = (6 - |s|)(1 + 0.5 \cos(\pi s/3)) + (9/2\pi) \sin(\pi |s|/3)$. Интервал интегрирования $[-6, 6]$. Дискретизация осуществляется методом Галеркина. Размер матрицы \mathbf{F} должен быть кратен 4. В этой задаче $\mathbf{F}\mathbf{x}_0$ не точно равно \mathbf{b}_0 .

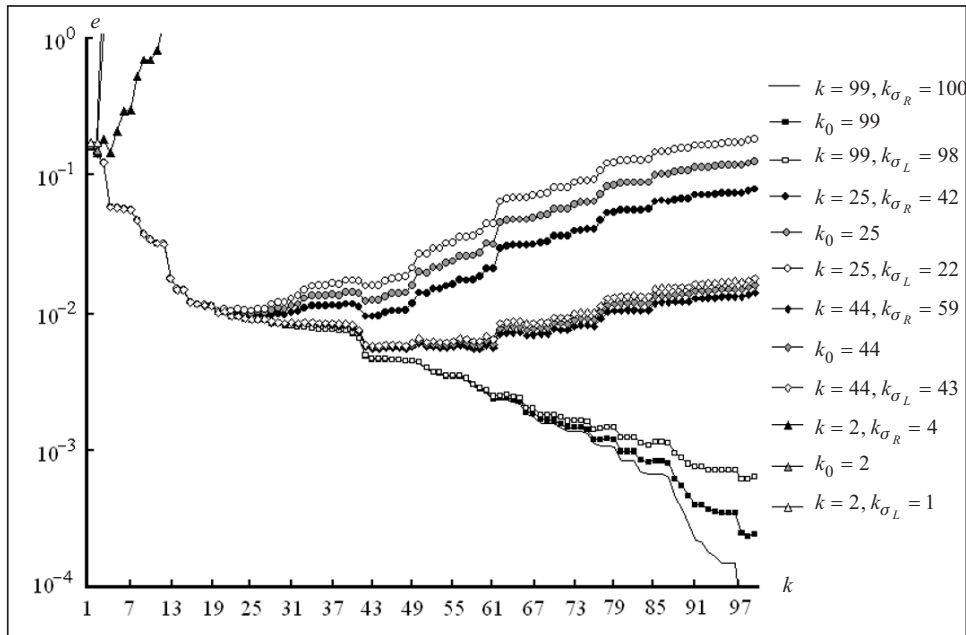


Рис. 3. Зависимость ошибки решения и ее составляющих от k для задачи Delves при $k_0 = \{2, 25, 44, 99\}$

Задача Delves [43] — вычисление второй производной. Проводится дискретизация аналитически заданного уравнения Фредгольма первого рода с ядром K , которое является функцией Грина второй производной с правой частью g и решением f : $K(s, t) = s(t-1)$ при $s < t$, $K(s, t) = t(s-1)$ при $s \geq t$, $g(s) = (s^3 - s)/6$, $f(t) = t$. В этой задаче $\mathbf{F}\mathbf{x}_0$ точно равно \mathbf{b}_0 .

Дискретное представление ядер и правых частей для задач Carasso, Phillips, Delves получено с использованием программ Regularization Tools [43].

Уровни шума выберем следующим образом. В каждой задаче для всего диапазона k вычислим граничные значения дисперсий по формулам (σ_R^2) и (σ_L^2) . Уровни шума, лежащие в интервалах (σ_R^2, σ_L^2) , обеспечивают глобальный минимум при соответствующем значении k .

Вычисления σ_R^2 (σ_L^2) и моделирование суммарной ошибки проводились без усреднения. Для ряда σ_R^2 стохастическую составляющую ошибки получали следующим образом: $e_s = \sigma_R^2 e_g$. Соответственно для σ_L^2 : $e_s = \sigma_L^2 e_g$, т.е. $e_s = \|\mathbf{F}_k^+ \mathbf{\Omega}_k \sigma \mathbf{\epsilon}_0\|^2 = \sigma^2 \|\mathbf{F}_k^+ \mathbf{\Omega}_k \mathbf{\epsilon}_0\|^2 = \sigma^2 \|\mathbf{e}_g\|^2 = \sigma^2 e_g$.

Суммируя стохастическую составляющую с детерминированной, получаем суммарную ошибку, представленную графиками (см. рис. 1–3), из которых видно, что детерминированная составляющая убывает, стохастическая возрастает, а глобальный минимум достигается в выбранных k_0 .

При граничных значениях дисперсии $\sigma^2 = \sigma_R^2$, $\sigma^2 = \sigma_L^2$ зависимость ошибки (от k) имеет два локальных минимума, значения ошибки в которых равны. При значениях дисперсии, удовлетворяющих $\sigma_R^2 < \sigma^2 < \sigma_L^2$, зависимость ошибки имеет один глобальный минимум.

ЗАКЛЮЧЕНИЕ

Проведенное аналитическое исследование составляющих ошибки восстановления истинного сигнала показало, что при возрастании размерности k случайного проектора детерминированная составляющая ошибки уменьшается, а стохастическая возрастает.

Кроме того, получены выражения, позволяющие рассчитать существование и значения интервала уровней шума, обеспечивающих (если интервал имеется) наличие глобального минимума ошибки в заданной точке k_0 .

Аналитические результаты подтверждены для дискретных некорректных задач Carrasso, Phillips, Delves. Направлением дальнейших исследований является разработка критерия, позволяющего оценить размерность k случайного проектора, при которой можно достичь минимума ошибки, не используя вектора сигнала x , неизвестного в прикладных задачах.

СПИСОК ЛИТЕРАТУРЫ

1. Hansen P.C. Rank-deficient and discrete ill-posed problems. Numerical aspects of linear inversion. — Philadelphia: SIAM, 1998. — 247 p.
2. Tikhonov A.N., Arsenin V.Y. Solution of ill-posed problems. — Washington: V.H. Winston, 1977. — 231 p.
3. Chernodub A.N. Training neuroemulators using multicriteria extended Kalman filter and pseudoregularization for model reference adaptive neurocontrol // Proc. of IV IEEE Intern. Congr. on Ultra Modern Telecommunications and Control Systems (ICUMT), St. Petersburg, Russia, October. — 2012. — P. 391–396.
4. Chernodub A.N. Training neural networks for classification using the extended Kalman filter: A comparative study // Optical Memory and Neural Networks. — 2014. — **23**, N 2. — P. 96–103.
5. Revunova E.G., Rachkovskij D.A. Using randomized algorithms for solving discrete ill-posed problems // International Journal Information Theories and Applications. — 2009. — **16**, N 2. — P. 176–192.
6. Ревунова Е.Г., Рачковский Д.А. Повышение точности решения обратной задачи с использованием случайных проекций // 15th Intern. Conf. Knowledge-Dialogue-Solution. — 2009. — P. 93–98.
7. Rachkovskij D.A., Revunova E.G. Intelligent gamma-ray data processing for environmental monitoring // Intelligent Data Analysis in Global Monitoring for Environmental and Security. — 2009. — P. 124–145.
8. Revunova E.G. Study of error components for solution of the inverse problem using random projections // Mathematical Machines and Systems. — 2010. — N 4. — P. 33–42.
9. Revunova E.G. Using model selection criteria for solving discrete ill-posed problems by randomized algorithms // The 4th International Workshop on Inductive Modelling (IWIM'2011), Kyiv. — 2011. — P. 89–97.
10. Rachkovskij D.A., Revunova E.G. Randomized method for solving discrete ill-posed problems // Cybernetics and Systems Analysis. — 2012. — **48**, N 4. — P. 621–635.
11. Пат. 100288 Украина. Способ определения физических величин по результатам их косвенных измерений / В.И. Гриценко, Д.А. Рачковский, Е.Г. Ревунова. — Оpubл. 10.12.2012, Бюл. № 23.
12. Revunova E.G., Rachkovskij D.A. Stable transformation of a linear system output to the output of system with a given basis by random projections // The 5th International Workshop on Inductive Modelling (IWIM'2012) Kyiv. — 2012. — P. 37–41.
13. Ревунова Е.Г. Подход рандомизации в задачах восстановления сигналов по результатам не прямых измерений // Кибернетика и вычислительная техника. — 2013. — Вып. 173. — С. 35–46.
14. Revunova E.G. Randomization approach to the reconstruction of signals resulted from indirect measurements // The 4th International Conference on Inductive Modelling (ICIM'2013) Kyiv. — 2013. — P. 203–208.
15. Johnson W.B., Lindenstrauss J. Extensions of Lipschitz mappings into a Hilbert space // Contemporary Mathematics. — 1984. — **26**. — P. 189–206.
16. Sarlos T. Improved approximation algorithms for large matrices via random projections // Proc. of the 47th Annual IEEE Symp. on Foundations of Computer Science. — 2006. — P. 143–152.
17. Halko N., Martinsson P.G., Tropp J.A. Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions // SIAM Review. — 2011. — **53**, N 2. — P. 217–288.
18. Rachkovskij D.A., Misuno I.S., Slipchenko S.V. Randomized projective methods for construction of binary sparse vector representations // Cybernetics and Systems Analysis. — 2012. — **48**, N 1. — P. 146–156.
19. Rachkovskij D.A. Vector data transformation using random binary matrices // Cybernetics and Systems Analysis. — 2014. — **50**, N 6. — P. 960–968.
20. Rachkovskij D.A. Formation of similarity-reflecting binary vectors with random binary projections // Cybernetics and Systems Analysis. — 2015. — **51**, N 2. — P. 313–323.

21. Гриценко В.И., Рачковский Д.А., Гольцев А.Д., Лукович В.В., Мисуно И.С., Ревунова Е.Г., Слипченко С.В., Соколов А.М., Талаев С.А. Нейросетевые распределенные представления для интеллектуальных информационных технологий и моделирования мышления // Кибернетика и вычислительная техника. — 2013. — Вып. 173. — С. 7–24.
22. Kussul E.M., Baidyk T.N., Lukovich V.V., Rachkovskij D.A. Adaptive neural network classifier with multfloat input coding // 6-th Intern. Conf. "Neural Networks and their Industrial and Cognitive Applications", Neuro-Nimes'93. — 1993. — P. 25–29.
23. Lukovich V.V., Goltsev A.D., Rachkovskij D.A. Neural network classifiers for micromechanical equipment diagnostics and micromechanical product quality inspection // 5th European Congress on Intelligent Techniques and Soft Computing EUFIT'97. — 1997 — **1**. — P. 534–536.
24. Kussul E.M., Kasatkina L.M., Rachkovskij D.A., Wunsch D.C. Application of random threshold neural networks for diagnostics of micro machine tool condition // Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. — 1998. — **1**. — P. 241–244.
25. Kussul E.M., Rachkovskij D.A., Wunsch D.C. The random subspace coarse coding scheme for real-valued vectors // IEEE International Joint Conference on Networks (IJCNN'99). — 1999. — **1**. — P. 450–455.
26. Frolov A.A., Rachkovskij D.A., Husek D. On informational characteristics of Willshaw-like auto-associative memory // Neural Network World. — 2002. — **12**, N 2. — P. 141–158.
27. Markman A.B., Rachkovskij D.A., Misuno I.S., Revunova E.G. Analogical reasoning techniques in intelligent counterterrorism systems // International Journal Information Theories and Applications. — 2003. — **10**, N 2. — P. 139–146.
28. Rachkovskij D.A. Some approaches to analogical mapping with structure-sensitive distributed representations // Journal of Experimental & Theoretical Artificial Intelligence. — 2004. — **16**, N 3. — P. 125–144.
29. Мисуно И.С., Рачковский Д.А., Слипченко С.В. Векторные и распределенные представления, отражающие меру семантической связи слов // Математические машины и системы. — 2005. — № 3. — С. 50–67.
30. Rachkovskij D.A., Slipchenko S.V. Similarity-based retrieval with structure-sensitive sparse binary distributed representations // Computational Intelligence. — 2012. — **28**, N 1. — P. 106–129.
31. Rachkovskij D.A., Kussul E.M., Baidyk T.N. Building a world model with structure-sensitive sparse binary distributed representations // Biologically Inspired Cognitive Architectures. — 2013. — **3**. — P. 64–86.
32. Dedieu J.-P., Novitsky D. Symplectic methods for the approximation of the exponential map and the newton iteration on riemannian submanifolds // Journal of Complexity. — 2005. — **21**. — P. 487–501.
33. Reznik A.M., Galinskaya A.A., Dekhtyarenko O.K., Nowicki D.W. Preprocessing of matrix QCM sensors data for the classification by means of neural network // Sensors and Actuators B. — 2005. — **106**. — P. 158–163.
34. Nowicki D.W., Dekhtyarenko O.K. Averaging on Riemannian manifolds and unsupervised learning using neural associative memory // Proc. ESANN 2005, Bruges, Belgium, April. — 2005. — **1**. — P. 27–29.
35. Хорн Р., Джонсон Ч. Матричный анализ. — М.: Мир, 1989. — 655 с.
36. Demmel J. W. Applied numerical linear algebra. — Philadelphia: SIAM, 1997. — 419 p.
37. Nowicki D., Verga P., Siegelmann H. Modeling reconsolidation in kernel associative memory // PLoS ONE 8(8): e68189. — 2013. — doi:10.1371/journal.pone.0068189.
38. Nowicki D, Siegelmann H. Flexible kernel memory // PLoS ONE 5(6): e10955. — 2010. — doi:10.1371/journal.pone.0010955.
39. Personnaz L., Guyon I., Dreyfus G. Collective computational properties of neural networks: New learning mechanisms // Physical Review A. — 1986. — **34**, N 5. — P. 4217–4228.
40. Кириченко Н.Ф., Резник А.М., Щетинюк С.П. Псевдообращение матриц в проблеме проектирования ассоциативной памяти // Кибернетика и системный анализ. — 2001. — № 3. — С. 18–27.
41. Stewart G.W. On the perturbation of pseudo-inverses, projections and linear least squares problems // SIAM Review. — 1977. — **19**, N 4. — P. 634–662.
42. Кириченко Н.Ф. Аналитическое представление возмущений псевдообратных матриц. // Кибернетика и системный анализ. — 1997. — № 2. — С. 98–107.
43. Hansen P.C. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems // Numer. Algorithms. — 1994. — **6**. — P. 1–35.

Поступила 17.09.2014