



УДК 512.64+517.5/512.643.5; 519.6

Ю.Н. Минаев, д-р техн. наук
Национальный авиационный университет
(Украина, 03057, Киев, пр-т Космонавта Комарова, 1,
тел. (044) 2495454, e-mail: min_14@ukr.net),
О.Ю. Филимонова, канд. техн. наук, **Ю.И. Минаева**, канд. техн. наук
Киевский национальный университет строительства и архитектуры
(Украина, 03037, Киев, Воздухофлотский пр-т, 31,
тел. (044) 2486427, 2425462, e-mail: filimonova@nm.ru; jumin@big-mir.net)

Гранулярные, нечетко множественные и тензорно-следовые характеристики многомерных временных рядов

Рассмотрены представления многомерного (многокомпонентного) временного ряда (ВР) в виде $3D$ тензорной модели, последующая тензорная декомпозиция которой с помощью процедур PARAFAC-декомпозиции и высокопорядковой сингулярной декомпозиции (HOSVD) позволяет представить весь ВР (или его составные части — окно, фрагмент, сегмент) в виде некоторой гранулы — подмножества упорядоченных троек, обладающих свойствами, близкими к свойствам нечетких множеств (НМ) второго типа, называемых псевдоНМ второго типа. Показана аналогия, существующая между свойствами следов, сингулярных величин и Φ -нормой стандартной сингулярной декомпозиции ($2D$ матрицы) и HOSVD, применяемой для $3D$ матрицы. Приведены примеры представления многомерных ВР гранулами — подмножествами упорядоченных троек.

Розглянуто представлення багатовимірного (багатокомпонентного) часового ряду (ЧР) у вигляді $3D$ тензорної моделі, подальша тензорна декомпозиція якої за допомогою процедур PARAFAC-декомпозиції і високопорядкової сингулярної декомпозиції (HOSVD) дозволяє представити весь ЧР (або його складові частини — вікно, фрагмент, сегмент) у вигляді деякої гранули — підмножини впорядкованих трійок, які мають властивості, близькі до властивостей нечітких множин (НМ) другого типу, названих псевдоНМ другого типу. Показано аналогію, існуючу між властивостями слідів, сингулярних величин і Φ -нормою стандартної сингулярної декомпозиції ($2D$ матриці) і HOSVD, яка застосовується для $3D$ матриці. Наведено приклади представлення багатовимірних ЧР гранулами — підмножинами впорядкованих трійок.

К л ю ч е в ы е с л о в а: сингулярная декомпозиция, нечетко-множественная гранула, высокопорядковая сингулярная декомпозиция, след матрицы, Φ -норма, временной ряд.

© Ю.Н. Минаев, О.Ю. Филимонова, Ю.И. Минаева, 2017

Основная задача анализа многомерных данных определяется целью, которая поставлена исследователем [1]. В частном случае такой задачей является идентификация аномалий временных рядов (ВР), так как развитие нераспознанных на ранних этапах аномалий приводит к катастрофическим последствиям. Выявление скрытых знаний — основа анализа многомерных данных. Считается, что получение скрытых знаний возможно посредством решения следующих задач:

моделирование структуры данных (здесь структура понимается настолько широко, что ее математическое содержание практически исчезает);

классификация (например, задачи идентификации аномалий, приводимые к задачам классификации, рассматриваемые в условиях неполной информации);

прогнозирование состояния компьютерных систем (КС) (в том числе ранний прогноз аномалии с явления самоподобия в многомерном случае).

Исследование и анализ многомерных данных, в том числе многомерных ВР, исторически связаны с использованием метода главных компонент (МГК) [2]. Однако сложность возникающих задач заставляет искать все более изощренные методы анализа [3], в том числе расширение МГК на многомерные ВР (МВР). Наиболее полно это представлено в работе [4], основное кредо которой сформулировано в ее названии, а именно: извлечение знаний через тензорные (матричные) декомпозиции (ТД).

В работах [5—7] показано, что представление нечетких множеств (НМ) (подмножества упорядоченных пар (ПмУП) $\tilde{x} \triangleq \{x/\mu^x\}$, $\mu^x \rightarrow [0, 1]$, $\tilde{x} \in E = \{e_i\}_{i=1}^n$ в виде тензорных моделей — гранул, позволяет получить скрытые знания в результате использования ТД, в частности сингулярной декомпозиции (SVD), и инвариантов матричных моделей.

Покажем, что применение ТД к МВР также приводит к псевдоНМ-гранулам, что должно способствовать поиску новых знаний, в частности новых признаков идентификации аномалий трафика КС. Следует, однако, заметить, что возможности предлагаемого подхода значительно шире рассматриваемой задачи. Тензорное моделирование данных, в том числе представленных в виде ВР, имеет преимущество, состоящее в том, что инварианты тензорной (матричной) модели не изменяются при стандартных преобразованиях: смещение и повороты осей координат. Это позволяет более эффективно решать задачу обработки многомерных данных.

Гранулы одномерных ВР описаны в работе [8], где показано, что одномерный ВР (окно \rightarrow фрагмент (совокупность окон) \rightarrow сегмент (совокупность фрагментов) и так далее) может быть представлен в виде тен-

зорной модели (ТМ), декомпозиция которой с помощью процедуры SVD позволяет вместо исходного ВР использовать ПМУП, имеющее семантику псевдоНМ. Эта процедура на рис. 1 представлена в виде цепочки: одномерный ВР (ОВР) \rightarrow ТМ \rightarrow процедура SVD \rightarrow формирование ПМУП (псевдоНМ) \rightarrow вычисление дефадзифицированных значений \rightarrow формирование нечеткого утверждения. Она положена в основу постановки задачи анализа одномерного ВР [1]. Дефадзифицированные значения псевдоНМ, вычисленные методом центра тяжести, следующие: исходное ПМУП — 1,22, нормализованное ПМУП — 0,65. Данным значениям ПМУП соответствует нечеткое утверждение: объект обладает свойством находиться в интервале $[0, 1]$.

Постановка задачи имеет некоторые особенности, так как при анализе МВР в значительной степени используются свойства ОВР. Пусть задан одномерный ВР $\mathbf{X}^{(t)} = \{x(t_i)\}, i=1, N$, в виде последовательности его фрагментов $\{X_j^{(t)}(t_i)\}, i=1, I; j=1, f$; каждый фрагмент состоит из окон $^{(k)}x(t_i), i=1, I_x, k=1, f$; f — число окон во фрагменте: $X_j^{(t)} = \bigcup_{k=1, f}^{(k)} x(t_i), \mathbf{X}^{(t)} = \bigcup_{j=1, J} X_j^{(t)}$. С помощью процедуры матрицизации окно (или фрагмент) ВР

представим в виде тензора с матрицей $m \times n$, т.е. $x(t_i) \rightarrow T_x = \text{reshape}(x(t_i), m, n), m \cdot n = I_x$. В общем случае $m \neq n$, однако будем рассматривать случай $m = n$, так как при этом появляются дополнительные возможности нахождения новых знаний (признаков) и существенно уменьшается вычислительная сложность задачи.

В работе [8] на рис. 1 приведен пример реального ВР и двух его фрагментов, один из которых содержит внедренные аномалии, а на рис. 2 и 3 представлены окна фрагментов ВР и их тензорные модели (2D матрицы). Заметим, что тензоры с матрицами разной размерности могут быть приведены к тензорам с матрицами одинаковой размерности с помощью кронекерова произведения (КП) [9]. Анализ многомерных данных в значительной степени основан на анализе 2D матриц (декомпозиции, в частности, SVD), расширенных на многопорядковые тензоры. Аналогично в анализе многомерных ВР используются результаты, полученные для одномерных ВР. Приведем некоторые положения анализа одномерных ВР, необходимые для дальнейшего понимания работы.

Одномерный ВР из N элементов $X = \{x_i\}_{i=1}^N$ представляют в виде ТМ, состоящей из f -фрагментов, в каждом из которых по k окон, состоящих из M элементов каждое. Например, для $N = 300$ реализуем стандартную про-

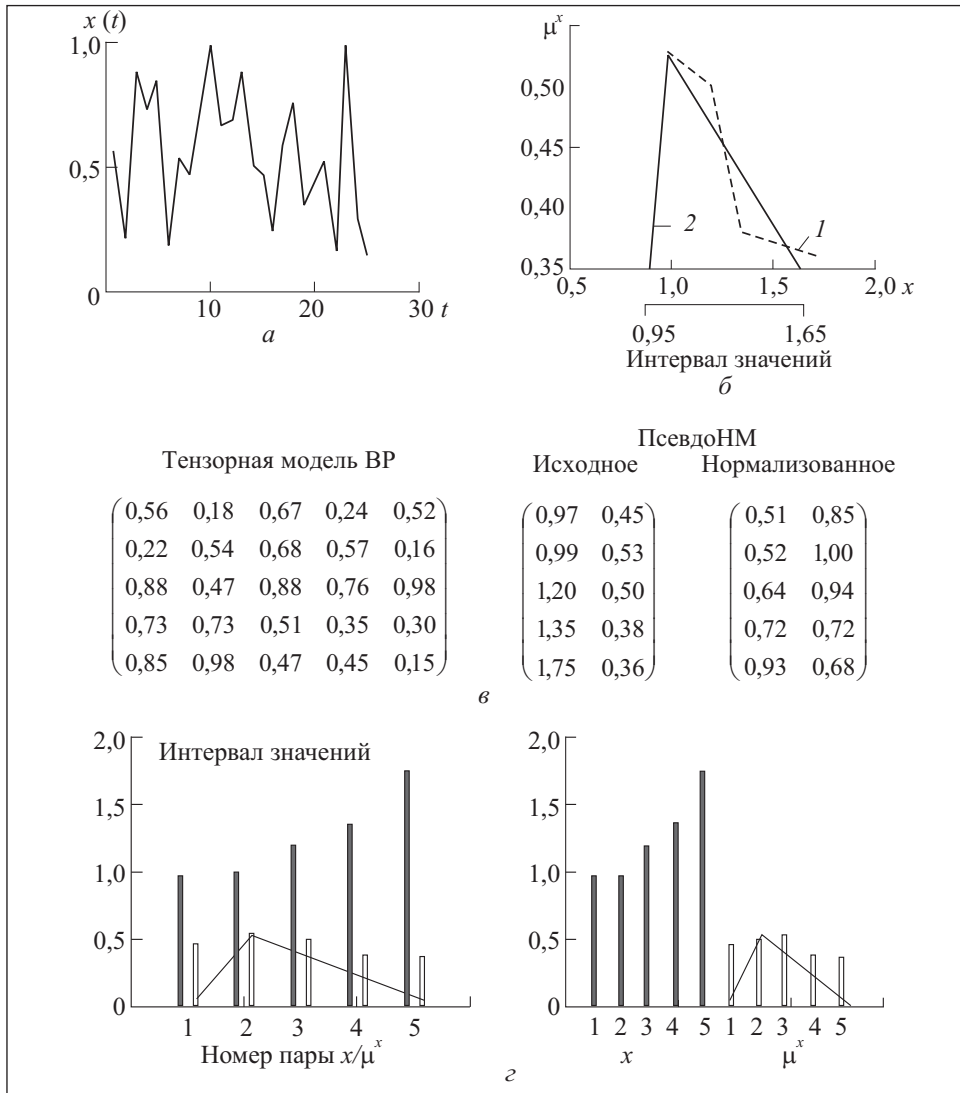


Рис. 1. Представление одномерного BP в виде ПМУП: а — отдельное окно BP; б — графическое представление ПМУП (1 — реальное псевдоНМ, 2 — аппроксимированное псевдоНМ $\|{}^{\pi}FS_1 - {}^{\pi}FS_2\|_F^2 \rightarrow \min$); в — тензорная (2D матрица), модель; г — формы визуального представления ПМУП; ■ — значение; □ — функция принадлежности (ФП)

цедуру МатЛаб reshape (g, 5, 5, 3, 4), $f=4, k=3, M=m \cdot m=25 (f \cdot k \cdot M=N)$ [8]. Тензорная модель состоит из 12 2D матриц $gt(:, :, 1:3, 1:4)$ размерностью 5×5 (первая матрица — $gt(:, :, 1, 1)$, вторая — $gt(:, :, 2, 1)$, ..., 12-я — $gt(:, :, 3, 4)$). Массив, имеющий четыре размерности, интерпретируется

стандартным МатЛабом как многомерная матрица или тензор. К его обработке не могут быть применены методы и модели МатЛаб, ориентированные на $2D$ матрицы.

В работе [8] показана возможность представления одномерного ВР (окно, фрагмент (совокупность окон), сегмент (совокупность фрагментов), ВР (совокупность сегментов)) в виде ПМУП, имеющих формальное сходство со стандартным НМ, но не являющимся таковым. Основой такого представления являются процедуры ТД. В случае многомерного ВР использование ПМУП не всегда возможно или целесообразно, поэтому прибегают к представлению такого объекта в виде подмножеств упорядоченных троек (ПМУТ). Эти подмножества называются гранулами.

Формирование ПМУТ реализовано с помощью процедур ТД, в частности PARAFAC-декомпозиции или стандартного HOSVD. Разработка и исследование процедур формирования НМ-тензор-гранул (в общем случае псевдоНМ) для многомерного (многокомпонентного) ВР составляет суть решаемой задачи.

В работе [8] предложено ПМУП, рассматриваемое как псевдоНМ, применять в качестве способа представления окон, фрагментов и сегментов ВР, используя математический аппарат [10]. Для заданной матрицы данных \mathbf{X} (составленной из элементов ВР) необходимо найти n -векторы \mathbf{u}, \mathbf{v} (возможно, но не обязательно, с $\|\mathbf{u}\|=1$), которые минимизируют критерий $\|\mathbf{X}_{-Kr} - \mathbf{u}\mathbf{v}^T\|_F^2 + \mathbf{P}_\lambda(u, v)$, где $\|\mathbf{X}_{-Kr} - \mathbf{u}\mathbf{v}^T\|_F^2 = \text{trace}\{(\mathbf{X} - \mathbf{u}\mathbf{v}^T)(\mathbf{X} - \mathbf{u}\mathbf{v}^T)^T\} = \sum_{i=1}^n \sum_{j=1}^n (x_{ij} - u_i v_j)^2$, $\mathbf{P}_\lambda(u, v) = \sum_{j=1, n; i=1, n} p_\lambda(|u_j|, |v_i|)$ — штрафная функция; $-_{Kr}$ — символ кронекеровой разности.

Известен путь решения данной задачи — аппроксимация исходного множества данных, рассматриваемого как тензор (в общем случае многомерный массив), так называемыми тензорными разложениями [11, 12], в соответствии с которыми SVD $m \times n$ -матрицы имеет вид $\mathbf{X} \cong \sigma_1 u_1 v_1^T + \dots + \sigma_r u_r v_r^T$, где $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ и $\sigma_i = \sqrt{\lambda_i}$; λ_i — собственные значения $\mathbf{X}\mathbf{X}^T$; векторы $u_1, v_1, u_2, v_2, \dots, u_r, v_r$ — ортогональны; $r = \text{rank}(B) \leq \min(m, n)$. Для приближенных вычислений ограничиваются первым элементом, т.е. $\mathbf{X} \cong \sigma_1 u_1 \otimes v_1^T$.

Если ВР аппроксимирован ПМУП типа $\sigma_1 u_1 \otimes v_1^T$, то можно формулировать задачу поиска ближайшего (относительно Ф-нормы) ПМУП $x \otimes \mu_x^T$, которое имеет семантику НМ, т.е. $\mu_x^T \rightarrow [0, 1]$ показывает значимость x в

составе \mathbf{X} : $\min_{x \in X, \mu_x \in M_x} \left\| \sigma_1 u_1 \otimes v_1^T - x \otimes \mu_x^T \right\|_F^2$. Напомним, что в соответствии с

[12, 13] для k -порядкового тензора $\mathfrak{N} \in \mathbb{R}^{d_1 \times \dots \times d_k}$, наилучшая ранг- r аппроксимация состоит в определении векторов $\mathbf{x}_i \in \mathbb{R}^{d_1}$, $\mathbf{y}_i \in \mathbb{R}^{d_2}$, ..., $\mathbf{z}_i \in \mathbb{R}^{d_k}$, $i = 1, \dots, r$, таких, что они минимизируют $\left\| \mathbf{A} - \mathbf{x}_1 \otimes \mathbf{y}_1 \otimes \dots \otimes \mathbf{z}_1 - \dots - \mathbf{x}_r \otimes \mathbf{y}_r \otimes \dots \otimes \mathbf{z}_r \right\|$ или $\arg \min_{\text{rank}(\mathbf{B}) \leq r} \left\| \mathbf{A} - \mathbf{B} \right\|$, где $\left\| \cdot \right\|$ определяет норму на $\mathbb{R}^{d_1 \times \dots \times d_k}$.

Для $k=2$ имеем $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V} = \sum_{i=1}^r \sigma_i u_i \otimes v_i$. Сингулярная декомпозиция

$\mathbf{A} \in \mathbb{R}^{m \times n}$ есть лучшая ранг- r аппроксимация и определяется первыми r термами в приведенной выше сумме [11, 12, 14]. В работе [15] показано, что для любого НМ может быть определено обычное четкое множество (ЧМ), расположенное на наименьшем расстоянии от него, т.е. имеющее наименьшую норму (добавим, по отношению к норме НМ или наименее уклоняющуюся от нормы НМ: $\left\| \mathbf{A} \right\|_F^2 \approx \left\| \tilde{\mathbf{A}} \right\|_F^2$, где $\left\| \cdot \right\|_F^2$ — квадрат фробениусовой нормы; $\left\| \cdot \right\|_F^2 = \text{trace}(\mathbf{A} \cdot \mathbf{A}^T)$). Это означает, что решение, например, задачи $\left\| \mathbf{A} - \tilde{\mathbf{A}} \right\|_F^2 \rightarrow \min$ позволяет на основании НМ $\tilde{\mathbf{A}} = [a_i \mu^{a_i}]_{i=1}^n$, $\mu^{a_i} \rightarrow [0, 1]$, вычислить другое (ближайшее) множество, в том числе и четкое (при некоторых дополнительных ограничениях): $\mathbf{A} = [a_i \lambda^{a_i}]_{i=1}^n$, где $\lambda^{a_i} \in \{0 \text{ или } 1\}$; μ^{a_i} — ФП; λ^{a_i} — характеристическая функция. Следует учитывать, что НМ и ЧМ должны быть представлены единообразно, т.е. в виде матрицы размерностью $2 \times n$ или матрицы, полученной как КП векторов: $(\mathbf{a}, \boldsymbol{\mu}^{\mathbf{a}})$ или $(\mathbf{a}, \boldsymbol{\lambda}^{\mathbf{a}})$ размерностью $n \times n$. Решение задачи минимизации $\psi = \left\| \mathbf{E} - \mathbf{B} \otimes \mathbf{M} \right\|_F^2$ может быть выполнено, если использовать оператор векторизации вес, т.е. оператор поколонкового представления матрицы [16].

В работе [9] показано, что если $\mathbf{A} \in \mathbb{R}^{m \times n}$ при $m = m_1 m_2$ и $n = n_1 n_2$ и если $\mathbf{B} \in \mathbb{R}^{m_1 \times n_1}$, $\mathbf{C} \in \mathbb{R}^{m_2 \times n_2}$, то $\left\| \mathbf{A} - \mathbf{B} \otimes \mathbf{C} \right\|_F = \left\| \mathfrak{R}(\mathbf{A}) - \text{vec}(\mathbf{B}) \text{vec}(\mathbf{C})^T \right\|_F$. Если $\hat{\mathbf{A}} \in \mathfrak{R}(\mathbf{A})$ имеет SVD $\mathbf{U}^T \hat{\mathbf{A}} \mathbf{V} = \Sigma = \text{diag}(\sigma_i)$, где $\mathbf{C} \in \mathbb{R}^{m_2 \times n_2}$, σ_1 — наибольшая сингулярная величина, $\mathbf{U}(:, 1)$ и $\mathbf{V}(:, 1)$ — сингулярные векторы, то матрицы $\mathbf{B} \in \mathbb{R}^{m_1 \times n_1}$ и $\mathbf{C} \in \mathbb{R}^{m_2 \times n_2}$, определенные в виде $\text{vec}(\mathbf{B}) = \sigma_1 \mathbf{U}(:, 1)$ и $\text{vec}(\mathbf{C}) = \mathbf{V}(:, 1)$, минимизируют $\left\| \mathbf{A} - \mathbf{B} \otimes \mathbf{C} \right\|_F$.

Высокопорядковые тензоры. Тензор — обобщение матриц, чьи порядки выше двух; тензор $\mathbf{A} \in \mathbb{R}^{(I_1 \times I_1 \times L \times \dots \times I_p)}$ имеет порядок p и I_1, I_2, \dots, I_p — целые числа, указывающие количество элементов для каждого измерения. Например, временная последовательность, представленная окнами, кото-

рые, в свою очередь, представлены $m \times n$ матрицами, может рассматриваться как тензор третьего порядка при $I_1 = m, I_2 = n, \dots$, и $I_3 = \tau$, где τ — число матриц во фрагменте.

Нотации и индексация. Слайс N -мерного тензора — это $2D$ сечение, определенное фиксацией всех индексов, кроме двух; волокно A N -мерного тензора $1D$ — сечение, определенное для всех индексов, кроме одного. Для тензора третьего порядка \mathbf{A} используют МатЛаб нотацию $\mathbf{A}(k, :, :)$, $\mathbf{A}(:, k, :)$ и $\mathbf{A}(:, :, k)$, чтобы определить k -й горизонтальный, продольный, фронтальный слайсы, и $A(:, i, j)$, $A(i, :, j)$ и $A(i, j, :)$, чтобы определить $(i; j)$ -е направления 1, 2 и 3 волокон. В частности, используют $A^{(k)}$ для представления $\mathbf{A}(:, :, k)$.

Определение 1. Пусть \aleph -тензор размерности $n_1 \times n_2 \times n_3$, тогда \aleph^T — это $n_1 \times n_2 \times n_3$ тензор, полученный транспозицией каждого фронтального слайса и последующим реверсом порядка транспонированного фронтального слайса 2 через n_3 .

Определение 2. Тензор идентичности $\mathbf{I} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ — тензор, в котором первый фронтальный слайс — это $n_1 \times n_1$ матрица идентичности и все остальные фронтальные слайсы есть нули.

Определение 3. f -диагональный тензор — это тензор, у которого каждый фронтальный слайс есть диагональная матрица.

В общем случае многомерный ($3D$) массив может быть представлен совокупностью слайсов; каждый слайс — матрица ($2D$ массив); каждая матрица — множество данных.

Матричное представление тензора высшего порядка. Для однозначности матричное раскрытие тензора [17, 18] определено как матричные представления данного тензора таким образом, что все вектор-столбцы (вектор-строки и др.) упорядочены последовательно, сохранено единообразное упорядочение вектор-столбца (вектор-строки и др.); для третьего порядка процедура раскрытия в соответствии с [17] представлена на рис. 2. Определения матричных раскрытий включают измерения тензора I_1, I_2, I_3 на циклической перестановке. При раскрытии размерности $I_c \times I_a \times I_b$ формально допускают, что индекс i_a изменяется медленнее, чем i_b .

Обобщая сказанное, приведем следующее положение [17]. Пусть N -порядковый тензор имеет вид $\mathbf{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$. Матричное представление (раскрытие) $\mathbf{A}_{(n)} \in \mathbb{R}^{I_n(I_{n+1}I_{n+2} \dots I_N I_1 I_2 \dots I_{n-1})}$ содержит элемент $a_{i_1 i_2 \dots i_N}$ в позиции с номером i_n строки и номером столбца $(i_{n+1} - 1)I_{n+2}I_{n+3} \dots I_N I_1 I_2 \dots I_{n-1} + (i_{n+2} - 1)I_{n+3}I_{n+4} \dots I_N I_1 I_2 \dots I_{n-1} + \dots + (i_N - 1)I_1 I_2 + (i_1 - 1)I_2 I_3 \dots I_N I_1 I_2 (i_2 - 1)I_3 I_4 \dots I_N I_1 I_2 \dots I_{n-1} \dots I_{n-1} + \dots + i_{n-1}$.

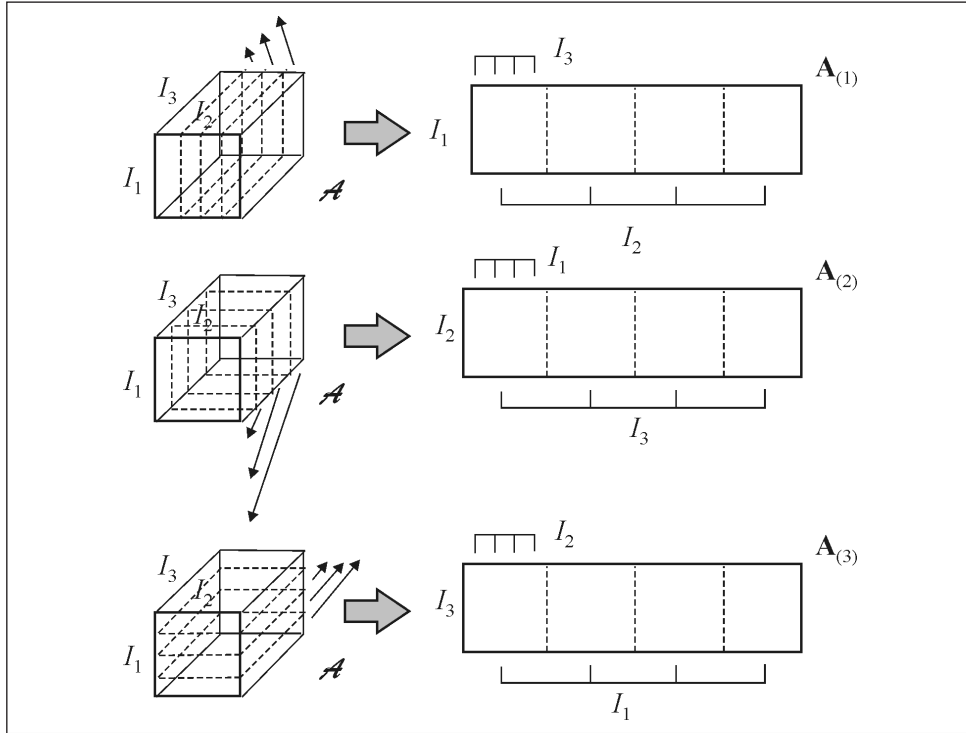


Рис. 2. Раскрытие $(I_1 \times I_2 \times I_3)$ -тензора \mathcal{A} в виде $(I_1 \times I_2 \cdot I_3)$ -матрицы $\mathbf{A}_{(1)}$, $(I_2 \times I_3 \cdot I_1)$ -матрицы $\mathbf{A}_{(2)}$ и $(I_3 \times I_1 \cdot I_2)$ -матрицы $\mathbf{A}_{(3)}$ ($I_1 = I_2 = I_3 = 4$) [17]

Пример [17]. Дан тензор $\mathcal{A} \in \mathbb{R}^{3 \times 2 \times 3}$ с элементами $a_{111} = a_{112} = a_{211} = -a_{212} = 1$, $a_{213} = a_{311} = a_{313} = a_{121} = a_{122} = a_{221} = -a_{222} = 2$, $a_{223} = a_{321} = a_{323} = 4$, $a_{113} = a_{312} = a_{123} = a_{322} = 0$. Тогда матричное раскрытие $\mathbf{A}_{(1)}$ имеет вид

$$\mathbf{A}_{(1)} = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & 2 & 0 \\ 1 & -1 & 2 & 2 & -2 & 4 \\ 2 & 0 & 2 & 4 & 0 & 4 \end{array} \right) \rightarrow \begin{array}{|ccc|} \hline 2 & 2 & 0 \\ 2 & -2 & 4 \\ \hline 1 & 1 & 0 \\ 1 & -1 & 2 \\ 2 & 0 & 2 \\ \hline \end{array}$$

Входные данные — 3D тензор $\mathbf{X} = \{X_{ijk}\}_{i=1}^{n_1} \{j=1}^{n_2} \{k=1}^{n_3}$ — могут быть рассмотрены как $\mathbf{X} = \{X_1, \dots, X_{n_3}\}$, где каждый $X_i \in \mathbf{X}$ есть 2D матрица (изображение, ВР) размерностью $n_1 \times n_2$. Многомерный ВР задан совокупностью своих компонент $\mathbf{X}^{(t)} = \{^{(i)}\mathbf{X}^{(t)}\}_{i=1}^n$, $^{(i)}\mathbf{X}^{(t)} = \{\mathbf{x}^{(i)}(t_j)\}$, $j=1, J$, рассматриваемых в n -мерном пространстве $\mathbb{R}^{x_1 \times x_2 \times \dots \times x_n}$, и представлен в табличном виде:

Номер записи	x_1	x_2	x_n	
1	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	Окно MBP = = поток
2	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	
3	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	
4	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	
5	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	$\langle \dots \rangle$	
·	·	·	·	·	·	
·	·	·	·	·	·	
·	·	·	·	·	·	

Стилизованное представление MBP приведено на рис. 3. Данная форма записи означает, что это может быть MBP, компоненты которого ортогональны, многокомпонентный BP, расположенный в трехмерном (или другом) пространстве. В ТМ MBP игнорируется указанное обстоятельство, все MBP рассматриваются как трехмерные. Здесь работает парадигма: многомерный (трехмерный массив (тензор) или 3D матрица) \rightarrow слайс (2D матрица) или волокно, с которыми в основном проводится реальная работа.

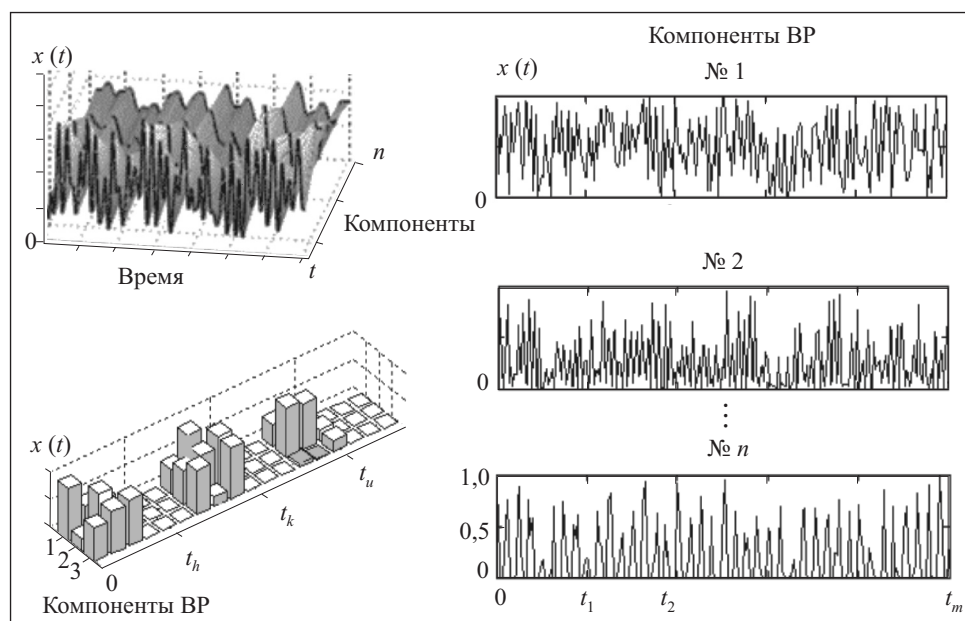


Рис. 3. Стилизованное представление MBP, заданного таблично

Пусть МВР имеет вид фрагмента, представленного покомпонентно:

Номер записи	x_1	x_2	x_3	x_4	x_5
1	0,6895	0,7631	0,8598	0,3423	0,9920
2	0,9231	0,9751	0,1732	0,2338	0,6338
3	0,5832	0,4397	0,7653	0,6545	0,0168
4	0,9875	0,3303	0,9185	0,2614	0,5155
5	0,5070	0,3012	0,4902	0,4496	0,4052
...
70	0,1298	0,7048	0,3867	0,9850	0,5638
71	0,6268	0,8461	0,2434	0,4991	0,2998
72	0,4306	0,5863	0,5993	0,4573	0,3479
73	0,9180	0,0469	0,8406	0,5741	0,7940
74	0,9780	0,8439	0,6605	0,9409	0,1266
75	0,2893	0,5868	0,6538	0,2483	0,8559
...

Многомерный ВР, состоящий, например, из 30 записей (строк), имеющий в нотации МатЛаб вид $y(1:30, 1:5)$, где 1:30 — число записей, 1:5 — число компонент (размерность) записи, представляется в виде 3D массива ($10 \times 5 \times 3$ -тензора) фронтальными слайсами как $z = \text{reshape}(y(1:30, 1:5), 10, 5, 3)$, где $y(1:30, 1:5)$ — 2D массив значений; $10 \times 5 \times 3$ — размерность 3D тензора. Могут быть и другие форматы записи, например $25 \times 3 \times 2$, $5 \times 5 \times 6$ или $5 \times 5 \times 3 \times 2$, однако число элементов во всех случаях данного примера должно быть равным 150.

Напомним, что МВР $\mathbf{X}(t)$, где $t = \{t_i\}$, $i = 1, I^t$, рассматривается как совокупность окон равной длины $\{X^{(j)}(t)\}$, $j = 1, J^X$, т.е. окна $X^{(k)}(t)$ и $X^{(l)}(t)$ анализируются во временных интервалах $t_k + \Delta_t$, $t_l + \Delta_t$, где Δ_t — целое число, $\Delta_t \in [1, t^\Delta]$. С помощью процедуры $\text{reshape}(\cdot)$ оконное множество значений ВР преобразуется в тензор второго порядка (матрицу) для одномерного ВР, в тензор третьего порядка (пространственную матрицу) — для МВР, т.е. $\mathbf{X}^{(k)}(t) = \{x^k(t_k), \dots, x^k(t_k + t^\Delta)\} \rightarrow \mathfrak{N}_k$, $\mathbf{X}(t) = \{\cup X^{(k)}(t)\}_{k=1, K} \rightarrow \{\cup \mathfrak{N}_k\}_{k=1, K}$. При этом 3D тензор матрицизируется, в частности, в виде фронтальных слайсов: тензор $\mathfrak{N}_k \rightarrow \{\mathbf{X}_k(:, :, 1), \mathbf{X}_k(:, :, 2), \mathbf{X}_k(:, :, 3)\}$ [12]. Многомерный ВР во многом повторяет свойства ОВР, так как получен с помощью процедуры матрицизации, однако при гранулировании ВР различие между ними является принципиальным, в частности одномерный ВР — псевдоНМ с ФП первого типа, МВР — псевдоНМ с ФП второго типа.

В работе [18] показано различие между стандартным (матричным SVD) и многомерным (высокопорядковым SVD — HOSVD), что важно с точки зрения формирования гранулы МВР. Высокопорядковый SVD пред-

ставляют как процедуру разложения p -порядкового тензора \mathbf{A} на составные части: $\mathbf{A} = \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \dots \times_p \mathbf{U}^{(p)}$, где $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(p)}$ — ортогональные матрицы, содержащие ортонормальные векторы, распределяющие пространство столбцов матрицы $\mathbf{A}_{(i)}$, открываемое при $i = 1, 2, \dots, p$. В работе [18] приведен также пример для $p=3$, в котором дано схематическое изображение процедуры HOSVD, где \mathbf{S} определен как тензор сердцевины, что соответствует обобщению матрицы в стандартном SVD. Обычно \mathbf{S} — полный тензор, но не диагональная матрица, как \mathbf{S} .

HOSVD является двухшаговой процедурой:

1. Вычисляют раскрытую матрицу $\mathbf{A}_{(i)}$, $i = 1, 2, \dots, p$, для тензора \mathbf{A} и стандартный SVD: $\mathbf{A}_{(i)} = \mathbf{U}\mathbf{S}\mathbf{V}^H$. Ортогональная матрица $\mathbf{U}^{(i)}$ определяется как $\mathbf{U}^{(i)} = \mathbf{U}$, т.е. левая матрица SVD.

2. Вычисляют корневой тензор, используя выражение $\mathbf{S} = \mathbf{A} \times_1 \mathbf{U}^{(1)H} \times_2 \mathbf{U}^{(2)H} \dots \times_p \mathbf{U}^{(p)H}$, где H — оператор эрмитовой матричной транспозиции.

Обратим внимание на существование определенной аналогии между процедурами сингулярной декомпозиции $3D$ матриц (тензоров) и $2D$ матриц. В частности, используют формальную аналогию между $2D$ и $3D$ матрицами для сингулярных величин и следов (первых инвариантов) матриц. Пусть $3D$ матрица \mathbf{A} имеет размерность $m_1 \times m_2 \times m_3$ с рангом $\rho = \{\rho_1, \rho_2, \dots, \rho_{m_3}\}$ и сингулярными величинами $\bar{\sigma} = \{\bar{\sigma}_1, \bar{\sigma}_2, \dots, \bar{\sigma}_{m_3}\}$, где $\bar{\sigma}_{i_3} = \{\bar{\sigma}_{i_3 1}, \bar{\sigma}_{i_3 2}, \dots, \bar{\sigma}_{i_3 m_3}\}$, $i_3 = 1, 2, \dots, m_3$, $SS(\mathbf{A}) = \sum_{i_3=1}^{m_3} \sum_{s=1}^{\rho_{i_3}} \sigma_{i_3 s}^2$. Установлены сле-

дующие характеристики (свойства) сингулярных величин \mathbf{A} :

$$1) SS(\mathbf{A}) = \text{trace}(\mathbf{A}^T \mathbf{A});$$

$$2) SS(\mathbf{A}) = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \sum_{i_3=1}^{m_3} a_{i_1 i_2 i_3}^2;$$

3) пусть $3D$ матрица \mathbf{A} разделена вдоль каждого из трех измерений на две части так, что получены восемь подматриц \mathbf{A}_{ijk} с размерностями $p_i \times q_j \times r_k$, $i, j, k = 1, 2$, такими, что $p_1 + p_2 = m_1$, $q_1 + q_2 = m_2$, $r_1 + r_2 = m_3$; тогда $\sum_{i,j,k=1}^2 SS(\mathbf{A}_{i,j,k}) = SS(\mathbf{A})$;

$$4) \text{ матричная норма } \|\mathbf{A}\| \text{ матрицы } \mathbf{A} \text{ определяется как } \|\mathbf{A}\| = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} \sum_{i_3=1}^{m_3} a_{i_1 i_2 i_3}^2.$$

Пусть задана последовательность фрагментов ВР $\mathbf{X}^{(t)} = \{\mathbf{X}^{(i)}\}_{i=1}^n$, ${}^{(i)}\mathbf{X}^{(t)} = \{x^{(i)}(t_j)\}$, $j=1, J$, состоящих из окон $x^{(i)}(t_j)$, рассматриваемых в

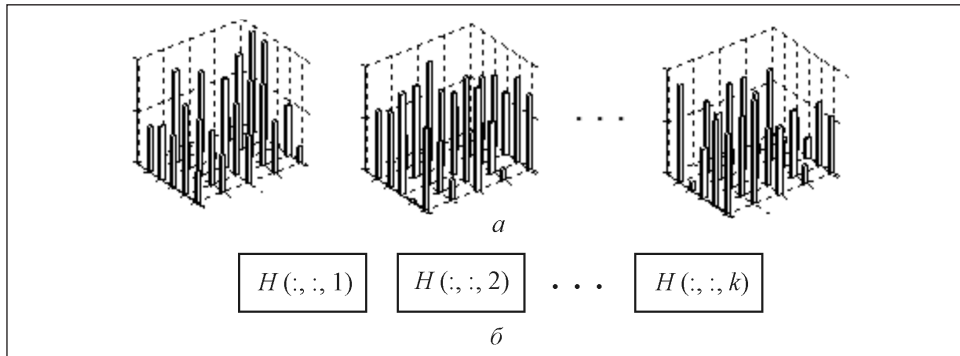


Рис. 4. Формирование тензор-гранулы окна (фрагмента, ВР) для одномерного ВР: *a* — тензорные (матричные) модели последовательных окон ВР; *б* — тензор-окно с матрицей $p \times q$ или $p \times p$ или $q \times q$, ограничение $p \cdot q = f$ или $p \cdot p = f$ или $q \cdot q = f$

двух $\mathbb{R}^{x_1 \times x_2}$ и n -мерном пространстве $\mathbb{R}^{x_1 \times \dots \times x_n}$. Тензорные (матричные) модели окон представлены на рис. 4. С помощью процедуры PARAFAC-декомпозиции или стандартного SVD формируют НМ тензор-гранулы. Последовательность формирования тензор-гранулы для многомерного (многокомпонентного) ВР представлена на рис. 5.

Следует заметить, что способ формирования тензорной НМ-гранулы в виде матриц размерностью $2 \times f$ на основе ПмУП и $f \times f$ на основе КП применяется и в случае ПмУТ:

$$\tilde{\mathbf{X}} = \left\{ x / \underbrace{\{y/v\}}_{\mathbf{M}^x} \right\} = \left\{ \begin{array}{l} \left[\begin{array}{cc} y_1(x_1) & v^{y_1(x_1)} \\ \vdots & \vdots \\ y_k(x_1) & v^{y_k(x_1)} \end{array} \right] \\ \vdots \\ \left[\begin{array}{cc} y_1(x_f) & v^{y_1(x_f)} \\ \vdots & \vdots \\ y_k(x_f) & v^{y_k(x_f)} \end{array} \right] \end{array} \right\} \rightarrow \mathbf{T}_{\tilde{\mathbf{X}}} = (x \otimes [y(x) \otimes v^{y(x)}]),$$

$$\tilde{\mathbf{X}} \in U, \mu^x \rightarrow [0,1], v^{y(x)} \rightarrow [0,1],$$

где $\tilde{\mathbf{X}}$ — результат реализации PARAFAC-декомпозиции 3D тензора, полученного на основании reshape-процедуры представления MBP; \mathbf{M}^x — ПмУП, сформированное на основе PARAFAC-декомпозиции; $\mathbf{T}_{\tilde{\mathbf{X}}}$ — тензор-гранула $\{x_i\}_{i=1}^f$ элементов.

Вычислительный эксперимент проведен для подтверждения эффективности представления многомерного ВР тензорной псевдоНМ-грану-

лой. Известно, что в ряде случаев решение для $3D$ матриц получают на основании метода SVD, разработанного для $2D$ матриц. Тем не менее, методы, предложенные в [17], являются предпочтительными. Стандартный SVD, используемый иногда для уменьшения размерности, имеет свои ограничения. Если \mathbf{F} -матрица разложена с использованием стандартного SVD ($[\mathbf{U} \mathbf{S} \mathbf{V}] = \text{svd}(\mathbf{F})$), произведение первых столбцов левой и правой матриц с первыми элементами диагональной матрицы \mathbf{S} дает наилучшую n -ранговую матрицу аппроксимации для \mathbf{F} . Это свойство не распространяется на порядки больше двух, кроме того, $\sum_{i=1,n} U(:,i)^2 \neq 1$.

Для вычисления наилучшей ранговой аппроксимации требуется альтернативный итеративно изменяемый метод наименьших квадратов (алгоритм ALS)[18, 21]. Аппроксимация, полученная в результате простого округления, в большинстве случаев оказывается практически оптимальной, однако ее применение в качестве псевдоНМ-гранулы ограничено.

Тензоры и PARAFAC-декомпозиция [19]. Тензор — многомерная матрица. В терминологии тензоров каждая размерность названа модой (направлением). Наиболее популярные тензоры — трехмерные, однако существуют приложения, в которых анализируются тензоры высших размерностей.

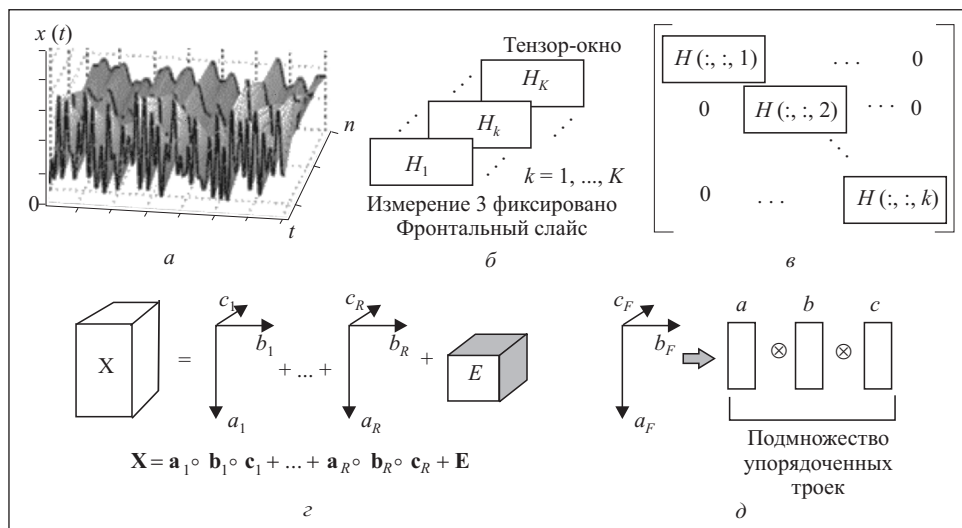


Рис. 5. Формирование тензор-гранулы окна (фрагмента, VP) для МВР: a — многомерный ВР; b — тензор-фрагмент размерностью $p \times q \times k$, где k — число окон во фрагменте; v — тензор-окно — фронтальный слайс тензор-фрагмента; z — тензор-фрагмент — блочно-диагональная матрица; d — диагональный элемент — тензор-окно

PARAFAC-декомпозиция может быть представлена как обобщение матричных разложений, например SVD высших размерностей. PARAFAC-декомпозиция тензора \mathfrak{N} в компоненты \mathbf{F} — это реализация выражения $\mathfrak{N} \approx \sum_{f=1}^F a_f \circ b_f \circ c_f$, где $[a \circ b \circ c](i, j, k) = a(i) b(j) c(k)$ — трехмодовое внешнее произведение (рис. 6). PARAFAC-декомпозиция представляется как триплет матриц \mathbf{A} , \mathbf{B} , \mathbf{C} , f -я колонка которых содержит соответственно a_f, b_f, c_f . Заметим, что тензор $\mathfrak{N} \in \mathbb{R}^{I \times J \times K}$ можно матрицизовать тремя способами: $\mathbf{X}_{(1)}$ размерностью $I \times JK$, $\mathbf{X}_{(2)}$ размерностью $J \times IK$, $\mathbf{X}_{(3)}$ размерностью $K \times IJ$.

Альтернативный алгоритм наименьших квадратов — наиболее популярный алгоритм для построения PARAFAC-декомпозиции. Он состоит из трех шагов. В основе каждого шага должна быть условная коррекция одной из трех матриц-факторов. Например, коррекцию матрицы фактора \mathbf{A} выполняют, сохраняя \mathbf{B} ; \mathbf{C} зафиксированными, она включает вычисление и псевдоинверсию $(\mathbf{C} \bullet \mathbf{B})$ (и соответственно для коррекций \mathbf{B} ; \mathbf{C}). Метод ALS описан в [21].

Пример. Для МВР, представленного на рис. 7 и в табл. 1, необходимо сформировать тензорно-гранулярную модель, т.е. представить МВР в виде тензорной гранулы или псевдоНМ. В табл. 2 представлены фронтальные слайсы 2D матрицы 5×5 , образующие 3D тензора.

Анализ результатов работы алгоритма (табл. 3) позволяет сделать следующие выводы:

стандартная гранула $\{\text{kron}(\mathbf{A}(:, 1), \text{kron}(\mathbf{B}(:, 1), \mathbf{C}(:, 1)'))\} \rightarrow \mathbf{A}(:, 1) \otimes \otimes (\mathbf{B}(:, 1) \otimes \mathbf{C}(:, 1))$ совпадает с матрицированным МВР, представленным в табл. 3;

модифицированная гранула, состоящая из двух субгранул, а именно: первая — $\text{kron}(\text{kron}(A(:, 1), C(1:5, 1)'), \text{B}(:, 1), C(6:10, 1)'))$, вторая — $\text{kron}(B(:, 1), C(6:10, 1)')$, представленная как двойное КП: $[A(:, 1) \otimes C(1:5, 1)] \otimes [(B(:, 1) \otimes C(6:10, 1))]$, является близкой к стандартной.

Стандартная гранула $(\text{kron}(\mathbf{A}(:, 1), \text{kron}(\mathbf{B}(:, 1), \mathbf{C}(:, 1)')))$ является блочной матрицей. В табл. 4 она выделена цветом. Показана возможность представления стандартной PARAFAC гранулы МВР $T_k = (\text{kron}(\mathbf{A}(:, 1), \text{kron}(\mathbf{B}(:, 1), \mathbf{C}(:, 1)')))$ в виде объединения двух субгранул, а именно ${}^{(1)}T_k = \text{kron}(A(:, 1), \text{kron}(B(:, 1), C(1:5, 1)'))$ и ${}^{(2)}T_k = \text{kron}(A(:, 1), \text{kron}(B(:, 1), C(6:10, 1)'))$ в виде ${}^{(1)}T_k \cup {}^{(2)}T_k = T_k$. Напомним, что символом «штрих» в выражениях $\text{kron}(\bullet)$ обозначена транспозиция.

На рис. 8, а, приведена PARAFAC-декомпозиция МВР, представленного в виде трехмерного тензора $5 \times 5 \times 10$, в форме подмножества, состоящего из

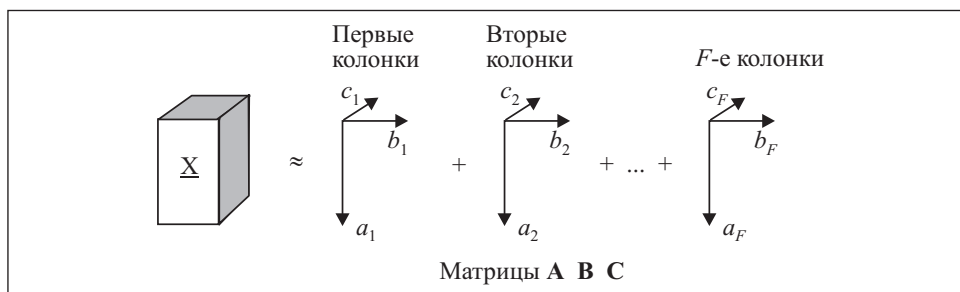


Рис. 6. PARAFAC-декомпозиция трехмерного тензора как сумма F внешних произведений (ранг-1 тензор), подобная ранг- F SVD матрицы

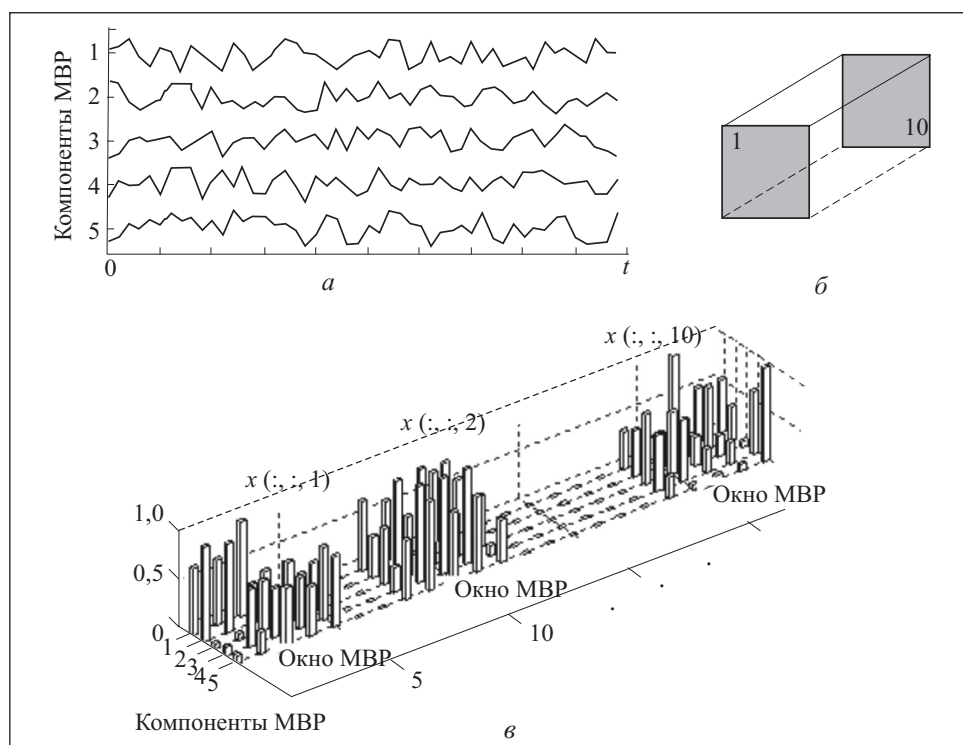


Рис. 7. Исходный MBP: a — 5 компонентов 50 измерений; b — тензорная форма представления MBP $\mathbf{X} = \text{reshape}(z, 5, 5, 10)$; c — 50 сканирований, пятикомпонентный ряд, окно $(1:5, 1:5, 1), \dots, (1:5, 1:5, 10)$

трех векторов. Вектор $\mathbf{C}(:, 1)$ является псевдоФП, так как $C(1:10, 1) \rightarrow [0, 1]$, для $i = 1, 5$ каждой паре элементов $\{a(i, 1) b(i, 1)\}$ соответствует элемент $c(i:5, 1)$; для $i = 6, 10$ каждой паре элементов $\{a(i, 1) b(i, 1)\}$ соответствует элемент $c(i:10, 1)$, что показано на рис. 8, б.

Таким образом, ТД МВР приводит к ПмУТ, что имеет определенные аналогии с НМ второго типа, в которых ФП является также НМ. ПсевдоНМ второго типа — гранула ПмУТ для МВР приведена в табл. 5.

Форма представления $\mathbf{A}(:, 1) \otimes \mathbf{B}(:, 1) \otimes \mathbf{C}(:, 1) = \mathbf{C}(1:5, 1) \cup \mathbf{C}(6:10, 1)$ в работе [22] рассмотрена в виде полутензорного произведения, для которого разработана отдельная алгебра. На рис. 9 представлены субгранулы в виде, подчеркивающим их природу как ПмУТ, имеющих определенную аналогию с НМ второго типа.

Представление окна (фрагмента, сегмента и др.) ВР в виде гранулы (ПмУТ) наиболее рационально, если указанные элементы образуют тензоры

Таблица 1

Номер измерения						Номер измерения					
	x_1	x_2	x_3	x_4	x_5		x_1	x_2	x_3	x_4	x_5
1	0.8302	1.3343	0.9354	0.9771	0.9836	26	1.1552	0.9561	0.9890	1.0329	0.7125
2	0.3285	0.5279	0.3701	0.3866	0.3892	27	0.4571	0.3783	0.3913	0.4087	0.2819
3	0.2183	0.3508	0.2459	0.2569	0.2586	28	0.3037	0.2514	0.2600	0.2716	0.1873
4	0.1341	0.2155	0.1511	0.1578	0.1589	29	0.1866	0.1544	0.1597	0.1668	0.1151
5	0.9496	1.5262	1.0700	1.1177	1.1251	30	1.3213	1.0936	1.1312	1.1815	0.8149
6	1.1364	1.8263	1.2804	1.3375	1.3463	31	1.5812	1.3087	1.3537	1.4138	0.9752
7	0.4496	0.7226	0.5066	0.5292	0.5327	32	0.6256	0.5178	0.5356	0.5594	0.3859
8	0.2988	0.4802	0.3366	0.3516	0.3540	33	0.4157	0.3441	0.3559	0.3717	0.2564
9	0.1835	0.2950	0.2068	0.2160	0.2174	34	0.2554	0.2114	0.2186	0.2283	0.1575
10	1.2998	2.0890	1.4646	1.5299	1.5400	35	1.8086	1.4969	1.5484	1.6172	1.1155
11	0.9300	1.4946	1.0479	1.0946	1.1018	36	1.2940	1.0710	1.1078	1.1570	0.7981
12	0.3680	0.5914	0.4146	0.4331	0.4360	37	0.5120	0.4238	0.4383	0.4578	0.3158
13	0.2445	0.3929	0.2755	0.2878	0.2897	38	0.3402	0.2816	0.2913	0.3042	0.2098
14	0.1502	0.2414	0.1692	0.1768	0.1779	39	0.2090	0.1730	0.1789	0.1869	0.1289
15	1.0638	1.0796	0.1986	1.2520	1.2603	40	1.4801	1.2251	1.2672	1.3535	0.9129
16	1.2240	1.9672	1.3792	1.4406	1.4502	41	1.7031	1.4096	1.4581	1.5229	1.0504
17	0.4843	0.7784	0.5457	0.5700	0.5738	42	0.6739	0.5577	0.5769	0.6026	0.4156
18	0.3218	0.5172	0.3626	0.3788	0.3813	43	0.4478	0.3706	0.3833	0.4004	0.2762
19	0.1977	0.3177	0.2227	0.2327	0.2342	44	0.2751	0.2277	0.2355	0.2459	0.1696
20	1.4001	2.2502	1.5775	1.6479	1.6588	45	1.9481	1.6124	1.6678	1.7419	1.2015
21	0.9396	1.5101	1.0587	1.1059	1.1133	46	1.3074	1.0821	1.1193	1.1691	0.8064
22	0.3718	0.5975	0.4189	0.4376	0.4405	47	0.5173	0.4282	0.4429	0.4626	0.3191
23	0.2470	0.3970	0.2784	0.2908	0.2927	48	0.3437	0.2845	0.2943	0.3074	0.2120
24	0.1518	0.2439	0.1710	0.1786	0.1798	49	0.2112	0.1748	0.1808	0.1888	0.1302
25	1.0748	1.7274	2.2110	1.2650	1.2734	50	1.4955	1.2378	1.2803	1.3372	0.9224

Таблица 2

$x(:, :, 1) =$					$x(:, :, 2) =$				
0.6850	0.6787	0.9746	0.3677	0.1406	0.7130	0.6230	0.0039	0.7288	0.5546
0.9795	0.9169	0.4022	0.1123	0.3227	0.4192	0.9027	0.9327	0.8949	0.1984
0.0519	0.0705	0.5154	0.6087	0.5098	0.5796	0.6067	0.2404	0.4400	0.5308
0.0917	0.5925	0.5101	0.5274	0.7456	0.2552	0.9953	0.9841	0.9832	0.0979
0.0811	0.2280	0.5935	0.5073	0.7261	0.6104	0.9185	0.7058	0.7828	0.4415
$x(:, :, 3) =$					$x(:, :, 4) =$				
0.3092	0.0584	0.8628	0.6017	0.0196	0.2914	0.5525	0.9224	0.8555	0.4974
0.7876	0.3001	0.3897	0.1972	0.4079	0.2232	0.5074	0.4992	0.2049	0.0884
0.7945	0.0476	0.4427	0.5757	0.6140	0.2259	0.6030	0.8386	0.4780	0.5821
0.5572	0.9083	0.4187	0.9669	0.2623	0.2832	0.6615	0.9195	0.3812	0.0211
0.7015	0.2870	0.9789	0.6758	0.8414	0.8292	0.5866	0.6060	0.4242	0.0047
$x(:, :, 5) =$					$x(:, :, 6) =$				
0.5541	0.2491	0.1116	0.2355	0.5242	0.5922	0.3977	0.9201	0.8997	0.1066
0.1120	0.9642	0.3943	0.7570	0.3992	0.5702	0.9457	0.4575	0.6302	0.2963
0.6982	0.3478	0.8533	0.2322	0.8994	0.5825	0.8765	0.6616	0.3468	0.8516
0.7511	0.4739	0.7393	0.4609	0.1714	0.3413	0.9884	0.5214	0.3058	0.7294
0.2498	0.8610	0.7625	0.0354	0.0247	0.5596	0.3914	0.9829	0.9012	0.6428
$x(:, :, 7) =$					$x(:, :, 8) =$				
0.2682	0.8809	0.2788	0.4020	0.2056	0.8308	0.2581	0.6951	0.1519	0.2402
0.7281	0.8563	0.5741	0.5389	0.2934	0.2975	0.7341	0.7568	0.6839	0.3542
0.8922	0.5995	0.8279	0.1897	0.2561	0.8772	0.4374	0.2666	0.2180	0.6077
0.5099	0.4822	0.7046	0.4347	0.3126	0.7549	0.3125	0.6112	0.7637	0.4129
0.5531	0.0120	0.3419	0.4087	0.5487	0.3252	0.9037	0.1723	0.0669	0.6161
$x(:, :, 9) =$					$x(:, :, 10) =$				
0.4318	0.0304	0.8433	0.5329	0.6755	0.3945	0.2259	0.9925	0.5542	0.5469
0.4159	0.6433	0.2426	0.0540	0.3271	0.4819	0.3554	0.3879	0.6391	0.3505
0.7158	0.8246	0.3700	0.6223	0.9843	0.7229	0.6421	0.2991	0.2129	0.0594
0.8765	0.7786	0.9646	0.6652	0.5123	0.5792	0.6430	0.2476	0.2264	0.6470
0.2622	0.4753	0.4691	0.7595	0.8174	0.2225	0.0521	0.0504	0.0665	0.9731

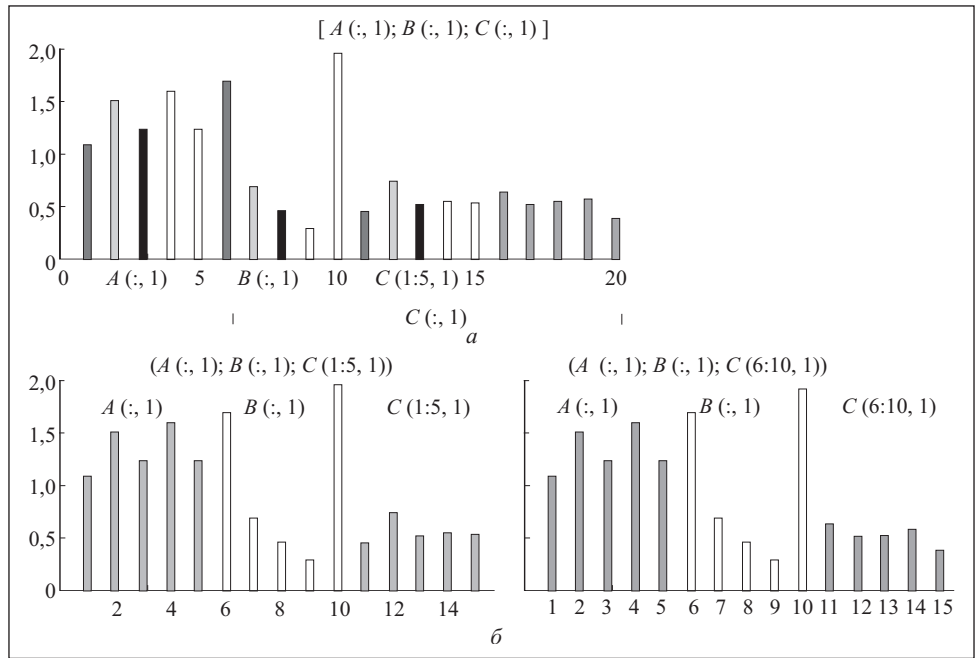


Рис. 8. PARAFAC-декомпозиция МВР в виде трехмерного тензора $5 \times 5 \times 10$ (а) и субгранулы ПМУТ (аналогично НМ второго типа) (б)

Таблица 3

A =			C =			A(:, 1)	C(:, 1)
1,0801	-0,3546	-0,3005	0,4513	0,5199	-0,3919	1,0801	0,4513
1,4783	0,1718	-0,5313	0,7253	-0,2143	-0,8547	1,4783	0,7253
1,2098	-0,1320	-0,2831	0,5085	0,6083	-0,4512	1,2098	0,5085
1,5924	0,0186	-0,5627	0,5311	0,4585	-0,6217	1,5924	0,5311
1,2224	-0,2102	-0,3401	0,5347	-0,0706	-0,5493	1,2224	0,5347
B =						B(:, 1)	
1,7034	0,6653	-3,2216	0,6279	0,5848	-0,6123	1,7034	0,6279
0,6740	0,6424	0,4402	0,5197	0,0202	-0,5094	0,6740	0,5197
0,4478	-2,5874	0,9314	0,5376	-0,1715	-0,5038	0,4478	0,5376
0,2751	-1,9726	1,1620	0,5615	0,4791	-0,4762	0,2751	0,5615
1,9484	0,1183	-4,4186	0,3873	0,2693	-0,2727	1,9484	0,3873

с одинаковыми размерностями по всем направлениям ($t \times t \times t$). Однако в силу технических или технологических причин это не всегда возможно, так как требует различных модификаций гранулы, т.е. изменения размерностей триплета $\{A, B, C\}$. Поэтому предложен способ представления гранулы, состоящий в разрежении компоненты $C(:, 1)$ таким образом, чтобы все компоненты триплета — вектора $\{A(:, 1), B(:, 1), C^m(:, 1)\}$ — имели одинаковое число элементов. Эти модификации гранулярных представлений МВР для рассматриваемого примера $\{A(:, 1) — \text{размерность } 1 \times 5, B(:, 1) — 1 \times 5, C^m(:, 1) — 1 \times 5 \text{ вместо } 1 \times 10\}$ могут быть использованы в

Таблица 4

0,8302	1,3343	0,9354	0,9771	0,9836	1,1552	0,9561	0,9890	1,0329	0,7125
0,3285	0,5279	0,3701	0,3866	0,3892	0,4571	0,3783	0,3913	0,4087	0,2819
0,2183	0,3508	0,2459	0,2569	0,2586	0,3037	0,2514	0,2600	0,2716	0,1873
0,1341	0,2155	0,1511	0,1578	0,1589	0,1866	0,1544	0,1597	0,1668	0,1151
0,9496	1,5262	1,0700	1,1177	1,1251	1,3213	1,0936	1,1312	1,1815	0,8149
1,1364	1,8263	1,2804	1,3375	1,3463	1,5812	1,3087	1,3537	1,4138	0,9752
0,4496	0,7226	0,5066	0,5292	0,5327	0,6256	0,5178	0,5356	0,5594	0,3859
0,2988	0,4802	0,3366	0,3516	0,3540	0,4157	0,3441	0,3559	0,3717	0,2564
0,1835	0,2950	0,2068	0,2160	0,2174	0,2554	0,2114	0,2186	0,2283	0,1575
1,2998	2,0890	1,4646	1,5299	1,5400	1,8086	1,4969	1,5484	1,6172	1,1155
0,9300	1,4946	1,0479	1,0946	1,1018	1,2940	1,0710	1,1078	1,1570	0,7981
0,3680	0,5914	0,4146	0,4331	0,4360	0,5120	0,4238	0,4383	0,4578	0,3158
0,2445	0,3929	0,2755	0,2878	0,2897	0,3402	0,2816	0,2913	0,3042	0,2098
0,1502	0,2414	0,1692	0,1768	0,1779	0,2090	0,1730	0,1789	0,1869	0,1289
1,0638	1,7096	1,1986	1,2520	1,2603	1,4801	1,2251	1,2672	1,3235	0,9129
1,2240	1,9672	1,3792	1,4406	1,4502	1,7031	1,4096	1,4581	1,5229	1,0504
0,4843	0,7784	0,5457	0,5700	0,5738	0,6739	0,5577	0,5769	0,6026	0,4156
0,3218	0,5172	0,3626	0,3788	0,3813	0,4478	0,3706	0,3833	0,4004	0,2762
0,1977	0,3177	0,2227	0,2327	0,2342	0,2751	0,2277	0,2355	0,2459	0,1696
1,4001	2,2502	1,5775	1,6479	1,6588	1,9481	1,6124	1,6678	1,7419	1,2015
0,9396	1,5101	1,0587	1,1059	1,1133	1,3074	1,0821	1,1193	1,1691	0,8064
0,3718	0,5975	0,4189	0,4376	0,4405	0,5173	0,4282	0,4429	0,4626	0,3191
0,2470	0,3970	0,2784	0,2908	0,2927	0,3437	0,2845	0,2943	0,3074	0,2120
0,1518	0,2439	0,1710	0,1786	0,1798	0,2112	0,1748	0,1808	0,1888	0,1302
1,0748	1,7274	1,2110	1,2650	1,2734	1,4955	1,2378	1,2803	1,3372	0,9224

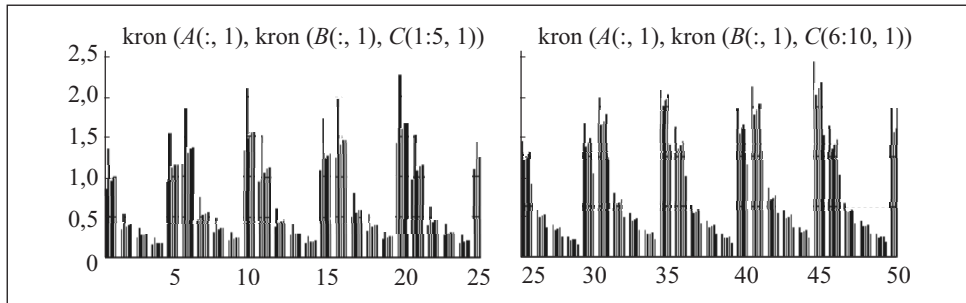


Рис. 9. Субгруппы, моделирующие ВР, представленные в виде ПМУТ

Таблица 5

A (:, 1)	B (:, 1)	C (:, 1)			
		C (1:5, 1)		C (6:10, 1)	
1.0801	1.7034	1	0.4513	6	0.6279
1.4783	0.6740	2	0.7253	7	0.5197
1.2098	0.4478	3	0.5085	8	0.5376
1.5924	0.2751	4	0.5311	9	0.5615
1.2224	1.9484	5	0.5347	10	0.3873

Таблица 6

SVD (kpr0)	SVD (kpr1)	SVD (kpr2)	Ф-норма
13.9806	9.3600	10.3849	norm (kpr0, fro) 13.9806
0.0000	0.0000	0.0000	norm (kpr1, fro) 9.3600
0.0000	0.0000	0.0000	norm (kpr2, fro) 10.3849
0.0000	0.0000	0.0000	
0.0000	0.0000	0.0000	
0.0000			
0.0000			
0.0000			
0.0000			
0.0000			

приложениях. В табл. 6 приведены вычисленные Φ -нормы гранул $krg_0 = \text{kron}(A(:,1), \text{kron}(B(:,1), C(:,1)'))$; $pr_1 = \text{kron}(A(:,1), \text{kron}(B(:,1), 1:2:10, 1)'))$; $kpr_2 = \text{kron}(A(:,1), \text{kron}(B(:,1), C(2:2:10, 1)'))$. Сравнение модифицированных гранул с исходной свидетельствует о том, что они близки, т.е. адекватно представляют МВР.

Выводы

Анализ многомерных ВР — сложная задача. Попытка свести ее к анализу последовательности автономных одномерных рядов в условиях гипертрофированных объемов не позволяет получить первичную (скрытую) информацию, существующую в многомерных ВР. Совместное применение ТД и гранулирования ВР в виде ПМУП (ПМУТ) для анализа многомерных ВР должно помочь в поиске новых знаний, в частности новых признаков идентификации аномалий ВР (например, трафика КС).

Вычисление наилучшей ранговой аппроксимации тензорных моделей ВР целесообразно выполнять с использованием альтернативного итеративно изменяемого метода наименьших квадратов, положенного в основу PARAFAC-декомпозиции, которая может быть представлена как обобщение матричных разложений, например HOSVD. Тензорная декомпозиция МВР приводит к гранулам формата ПМУТ, которые имеют определенные аналогии с НМ второго типа, в которых ФП представляет собой также НМ.

СПИСОК ЛИТЕРАТУРЫ

1. Эбсенсен К. Анализ многомерных данных. Избранные главы. / Пер. с англ. С.В. Кучерявского; под ред. О.Е. Родионовой. — Черноголовка: Изд-во ИПХВ РАН, 2005. — 160 с.
2. Dobos L., Abonyi J. On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation // *Chemical Engineering Science*. — 2012. — № 75. — P. 96—105.
3. Ringberg H., Soule Au., Rexford J., Diot Ch. Sensitivity of PCA for Traffic Anomaly Detection // *SIGMETRICS'07*, June 12—16, 2007, San Diego, California, USA. Copyright 2007 ACM 978-1-59593-639-4/07/0006
4. Skillicorn D. Data Mining and Knowledge Discovery Series. Understanding Complex Datasets. Data Mining with Matrix Decompositions. — Chapman & Hall/CRC, 2007. — 257 p.
5. Минаев Ю.Н., Филимонова О.Ю., Минаева Ю.И. Структурированные гранулы нечеткого множества в задачах гранулярного компьютеринга // *Электрон. моделирование*. — 2015. — 37, № 1. — С. 77—95.
6. Минаев Ю.Н., Филимонова О.Ю., Минаева Ю.И. Кронекеровы (тензорные) модели нечетко-множественных гранул // *Кибернетика и системный анализ*. — 2014. — 50, № 4. — С. 42—52.
7. Минаев Ю.Н., Филимонова О.Ю., Минаева Ю.И. Тензорные модели НМ-гранул и их применение для решения задач нечеткой арифметики // *Искусственный интеллект*. — 2013. — № 2. — С. 22—31.

8. Минаев Ю.Н., Филимонова О.Ю., Минаева Ю.И. Нечетко множественные характеристики одномерных временных рядов // Электрон. моделирование. — 2016. — **38**, № 6. — С. 45—66.
9. Van Loan Ch. The ubiquitous Kronecker product // Journal of Computational and Applied Mathematics. — 2000. — № 123. — P. 85—100.
10. Shen H., Huang J.Z. Sparse principal component analysis via regularized low rank matrix approximation / Journal of Multivariate Analysis. — 2008. — № 99. — P. 1015—1034.
11. Bader B.W., Kolda T.G. Tensor Decompositions, the MATLAB Tensor Toolbox, and Applications to Data Analysis. Tensor Decompositions Multilinear operators for higher-order decompositions. Technical Report SAND2006-2081, Sandia National Laboratories, April, 2006. Albuquerque, New Mexico 87185 and Livermore, California 94550. — 39 p. Интернет-ресурс — <http://csmr.ca.sandia.gov/~tgkolda/>
12. De Silva V., Lim L. Tensor Rank and the Ill-Posedness of the best Low-Rank Approximation Problem // Institute for Computational and Mathematical Engineering, Stanford University, Stanford. CA 94305-9025. — Интернет-ресурс: arxiv.org/pdf/math/0607647.
13. Воеводин В.В. Энциклопедия линейной алгебры. Электронная система ЛИНЕАЛ / В.В. Воеводин, Вл.В. Воеводин. — СПб. : БХВ-Петербург, 2006. — 544 с.
14. Kibangou A.I.Y. Tensor decompositions and Applications. An overview and some contributions. GIPSA-N_CS. March 17, 2009. — 88 p. — Интернет-ресурс: <http://www.sandia.gov/~tgkolda/.../TensorReview.pdf>
15. Кофман А. Введение в теорию нечетких множеств / Пер. с франц. — М. : Радио и связь, 1982. — 432 с.
16. Van Loan C.F., Pitsianis C.F. Approximation with Kronecker Products. M.S. Moonen et al. (eds.). Linear Algebra for Large Scale and Real-Time Applications. Kluwer Publishers. — 1993. — P. 293—314.
17. De Lathauwer L., De Moor B., Vandewalle J. On the best rank-1 and rank-(R1,R2, ..., RN) Approximation of Hinger-Order Tensors // SIAM J. MATRIX ANAL. APPL. Society for Industrial and Applied Mathematics. — 2000. — Vol. 21, No. 4. — P. 1324—1342.
18. Costantini R., Sbaiz L., Süsstrunk S. Higher Order SVD Analysis for Dynamic Texture Synthesis // IEEE Transactions on image processing. — 2008. — Vol. 17, No. 1. — P. 42—52.
19. Papalexakis E., Faloutsos Ch., Sidiropoulos N., Harpale A. Large Scale Tensor Decompositions: Algorithmic Developments and Applications. — Bulletin of the IEEE Computer Society Technical Committee on Data Engineering. — 2013. — P. 59—67.
20. Kamalja K.K., Khangar N.V. Singular Value Decomposition for Multidimensional Matrices. Research article. Int. Journal of Engineering Research and Applications, V. 3, Issue 6, Nov - Dec 2013, P. 123—129.
21. Sidiropoulos D., Giannakis G.B., Bro R. Blind PARAFAC. Receivers for DS-CDMA Systems // IEEE Transactions on Signal Processing March. — 2000. — V. 48, № 3. — P. 810—823.
22. Cheng D., Qi H., Xue A.A. Survey on Semi-tensor Product of Matrices // Jrl Syst Sci and Complexity. — 2007. — N 20. — P. 304—322.

REFERENCES

1. Esbensen, K. (2005), *Analiz mnogomernyh dannyh. Izbrannye glavy* [Analysis of multidimensional data. Selected chapters], Transl. from Engl. S.V. Kucheryavskiy; Ed O.E. Rodionova, Izdatelstvo IPHV RAN, Chernogolovka, Russia.
2. Dobos, L. and Abonyi, J. (2012), “On-line detection of homogeneous operation ranges by dynamic principal component analysis based time-series segmentation”, *Chemical engineering science*, Vol. 75, pp. 96-105.

3. Ringberg, H., Soule, Au., Rexford, J. and Diot, Ch. (2007), "Sensitivity of PCA for traffic anomaly detection", *SIGMETRICS '07*, June 12-16, 2007, San Diego, California, USA. Copyright 2007 ACM 978-1-59593-639-4/07/0006
4. Skillicorn, D. (2007), Data mining and knowledge discovery series. Understanding complex datasets. Data mining with matrix decompositions, Chapman & Hall/CRC.
5. Minaev, Ju.N., Filimonova, O.Ju. and Minaeva, Ju.I. (2015), "Structured granules of fuzzy set in the problems of granular computing", *Elektronnoe modelirovanie*, Vol. 37, no. 1, pp.77-95.
6. Minaev, Ju.N., Filimonova, O.Ju. and Minaeva, Ju.I. (2014), "Kronecker (tensor) models of fuzzy-set granules", *Kibernetika i sistemnyj analiz*, Vol. 50, no. 4, pp. 42-52.
7. Minaev, Ju.N., Filimonova, O.Ju. and Minaeva, Ju.I. (2013), "Tensor models of NM-granules and their use for solving problems of fuzzy arithmetic", *Iskusstvennyj intellekt*, no. 2, pp. 22-31.
8. Minaev, Ju.N., Filimonova, O.Ju. and Minaeva, Ju.I. (2016), "Nonfuzzy multiple characteristics of unidimensional time series (TS)", *Elektronnoe modelirovanie*, Vol. 38, no. 6, pp. 45-66.
9. Van Loan, Ch. (2000), "The ubiquitous Kronecker product", *Journal of computational and applied mathematics*, Vol. 123, pp. 85-100.
10. Shen, H. and Huang, J.Z. (2008), "Sparse principal component analysis via regularized low rank matrix approximation", *Journal of multivariate analysis*, Vol. 99, pp. 1015-1034.
11. Bader, B.W. and Kolda, T.G. (2006), Tensor decompositions, the MATLAB tensor toolbox, and applications to data analysis. Tensor decompositions multilinear operators for higher-order decompositions. Technical Report SAND2006-2081, Sandia National Laboratories, April 2006. Albuquerque, New Mexico 87185 and Livermore, California 94550, available at: - <http://csmr.ca.sandia.gov/~tgkolda/>
12. De Silva, V. and Lim, L. Tensor rank and the ill-posedness of the best low-rank approximation problem, Institute for Computational and Mathematical Engineering, Stanford University, Stanford. CA 94305-9025, available at: arxiv.org/pdf/math/0607647.
13. Voevodin, V.V. and Voevodin, Vl.V. (2006), *Entsiklopedia lineynoi algebrы. Elektronnaya sistema LINEAL* [Encyclopedia of linear algebra. Electron system LINEAL], BHV-Peterburg, St-Petersburg, Russia.
14. Kibangou, Al.Y. (2009), Tensor decompositions and applications. An overview and some contributions, GIPSA-N_CS. March 17, 2009, available at: www.sandia.gov/~tgkolda/.../TensorReview.pdf
15. Kofman, A. (1982), *Vvedenie v teoriyu nechetkikh mnozhestv* [Introduction in the theory of fuzzy sets], Transl. from French, Radio i svyaz, Moscow, Russia.
16. Van Loan, C.F. and Pitsianis, C.F. (1993), Approximation with Kronecker products, Eds M.S. Moonen et al., Linear algebra for large scale and real-time applications, Kluwer Publishers, Dordrecht, the Netherlands.
17. De Lathauwer, L., De Moor, B. and Vandewalle, J. (2000), "On the best rank-1 and rank-(R_1, R_2, \dots, R_N) approximation of higher-order tensors", *SIAM J. MATRIX ANAL. APPL.* Society for Industrial and Applied Mathematics, Vol. 21, no. 4, pp. 1324-1342.
18. Costantini, R., Sbaiz, L. and Süsstrunk, S. (2008), "Higher order SVD analysis for dynamic texture synthesis", *IEEE Transactions on Image Processing*, Vol. 17, no. 1, pp. 42-52.
19. Papalexakis, E., Faloutsos, Ch., Sidiropoulos, N. and Harpale, A. (2013), "Large scale tensor decompositions: Algorithmic developments and applications", *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, pp. 59-67.
20. Kamalja, K.K. and Khangar, N.V. (2013), "Singular value decomposition for multidimensional matrices. Research article", *Int. Journal of engineering research and applications*, Vol. 3, Iss. 6, pp. 123-129.
21. Sidiropoulos, D., Giannakis, G.B. and Bro R. (2000), "Blind PARAFAC. Receivers for DS-CDMA Systems", *IEEE Transactions on Signal Processing*, Vol. 48, no. 3, pp. 810-823.
22. Cheng, D., Qi, H. and Xue, A.A. (2007), "Survey on semi-tensor product of matrices", *J. syst. sci. and complexity*, Vol. 20, pp. 304-322.

Yu.N. Minaev, O.Yu. Filimonova, J.I. Minaeva

GRANULAR, FUZZY SET AND TESOR-TRACE
CHARACTERISTICS OF MULTIDIMENSIONAL TIME SERIES

The questions of representation of multidimensional (multicomponent) time series (TS) as a 3D tensor model; its following tensor decomposition with the use of procedures of PARAFAC-decomposition and higher-order singular decomposition (HOSVD) allows us to represent the entire TS (or its components — window, a fragment, a segment) in the form of a certain granule — a subset of ordered triples with properties similar to those of the second type FS called the second type psevdofS.

The analogy is shown between the properties of the existing traces, singular values and F-norm of standard singular decomposition (2D matrix) and HOSVD used for 2D matrix. Examples showing the representation of multidimensional TS by granules — subsets of ordered triples.

Key words: singular decomposition, fuzzy multiple granule, higher-order singular decomposition, trace of a matrix, F-norm, time series.

Поступила 15.02.16;
после доработки 30.11.16

МИНАЕВ Юрий Николаевич, д-р техн. наук, профессор кафедры компьютерных систем и сетей Национального авиационного университета Украины. В 1959 г. окончил Харьковский политехнический ин-т. Область научных исследований — интеллектуальный анализ данных, применение интеллектуальных технологий в системах принятия решений.

ФИЛИМОНОВА Оксана Юрьевна, канд. техн. наук, доцент Киевского национального университета строительства и архитектуры. В 1989 г. окончила Киевский инженерно-строительный ин-т. Область научных исследований — интеллектуальный анализ данных.

МИНАЕВА Юлия Ивановна, канд. техн. наук, доцент кафедры основ информатики Киевского национального университета строительства и архитектуры, который окончила в 2008 г. Область научных исследований — интеллектуальный анализ данных.