Mathematical Forecasting Methods

Лекция 10

МФТИ

Весна, 2024

План курса и темы

Формальная часть

- Домашнее задание на тензорное разложение (Tucker decomposition) и регрессию HOPLS
- Примерные даты 26 марта - 9 апреля
- Индивидуальные проекты по лабораторным работам прошлых лет
- Примерные даты 16 марта - 7 мая

Темы из курса прошлых лет

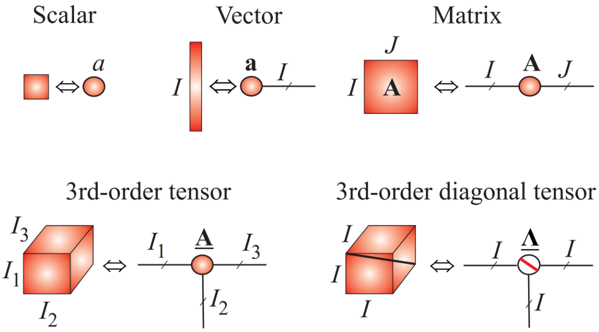
- Tensors and Penrose notation
- Tucker decomposition and alternated least squares
- Higher-order singular values decomposition
- Higher-order PLS

Особенности обозначений

Machine Learning	Quantum Physics
Nth-order tensor	rank-N tensor
high/low-order tensor	tensor of high/low dimension
ranks of TNs	bond dimensions of TNs
unfolding, matricization	grouping of indices
tensorization	splitting of indices
core	site
variables	open (physical) indices
ALS Algorithm	one-site DMRG or DMRG1
MALS Algorithm	two-site DMRG or DMRG2
column vector $\mathbf{x} \in \mathbb{R}^{I \times 1}$	ket $\|\Psi\rangle$
row vector $\mathbf{x}^{\mathrm{T}} \in \mathbb{R}^{1 \times I}$	bra $\langle\Psi\|$
inner product $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{x}$	$\langle\Psi\|\Psi\rangle$
Tensor Train (TT)	Matrix Product State (MPS) (with Open Boundary Conditions (OBC))
Tensor Chain (TC)	MPS with Periodic Boundary Conditions (PBC)
Matrix TT	Matrix Product Operators (with OBC)
Hierarchical Tucker (HT)	Tree Tensor Network State (TTNS) with rank-3 tensors

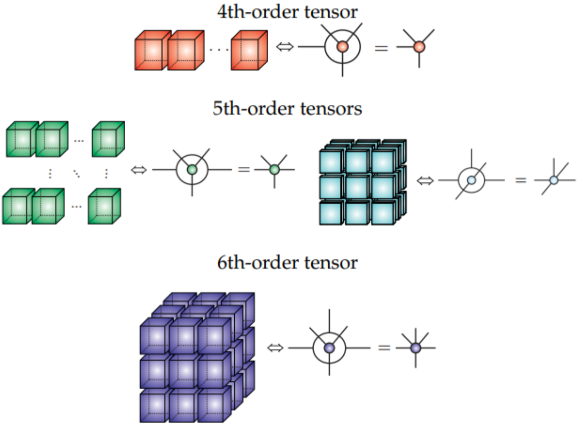
Тензоры и графическое представление

Определение 1. Под тензором $\underline{\mathbf{A}} \in \mathbb{F}^{I_1 \times \ldots \times I_d}$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ будем понимать многомерный массив с элементами $a_{i_1,\ldots,i_d}$, где
d — размерность (иногда порядок или ранг),
i_k — размер k-ой моды

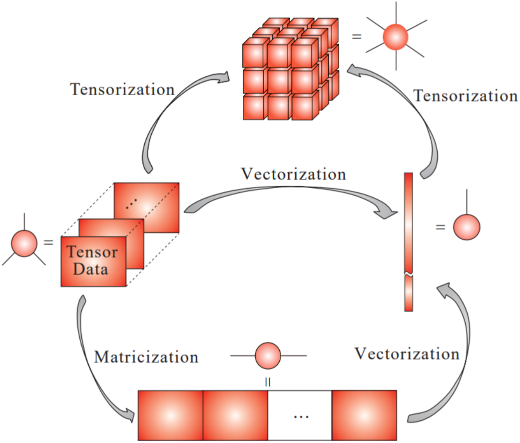


Тензоры и графическое представление

Определение 1. Под тензором $\underline{\mathbf{A}} \in \mathbb{F}^{I_1 \times \dots \times I_d}$, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ будем понимать многомерный массив с элементами $a_{i_1,\dots,i_d}$, где
d — размерность (иногда порядок или ранг),
i_k — размер k-ой моды



Тензоры и графическое представление



Определение 2. Пусть есть два тензора $\underline{\mathbf{A}} \in \mathbb{R}^{I_1 \times \ldots \times I_d}$ $\underline{\mathbf{B}} \in \mathbb{R}^{J_1 \times \ldots \times J_D}$, тогда назовём их внешним произведением следующий тензор $\underline{\mathbf{A}} \circ \underline{\mathbf{B}} \in \mathbb{R}^{I_1 \times \ldots \times I_d \times J_1 \times \ldots \times J_D}$ с элементами:

$$(\underline{\mathbf{A}} \circ \underline{\mathbf{B}})_{i_1,\ldots,i_d,j_1,\ldots,j_D} = a_{i_1,\ldots,i_d} b_{j_1,\ldots,j_D}.$$

Свёртка тензоров

- (N, 1)-свёртка (contraction) тензоров $\underline{\mathbf{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ и $\underline{\mathbf{B}} \in \mathbb{R}^{J_1 \times \cdots \times J_M}$ (свёртка вдоль N-ой моды первого тензора и 1-ой моды второго), где обязательно $I_N = J_1$, даёт тензор $\underline{\mathbf{C}} \in \mathbb{R}^{I_1 \times \cdots \times I_{N-1} \times J_2 \times \cdots \times J_M}$ с элементами:

$$\underline{\mathbf{C}} = \underline{\mathbf{A}} \times_N^1 \underline{\mathbf{B}} = \underline{\mathbf{A}} \times^1 \underline{\mathbf{B}} = \underline{\mathbf{A}} \bullet \underline{\mathbf{B}}$$

$$c_{i_1,\ldots,i_{N-1},j_2,\ldots,j_M} = \sum_{i_N=1}^{I_N} a_{i_1,\ldots,i_{N-1},i_N} b_{i_N,j_2,\ldots,j_M}$$

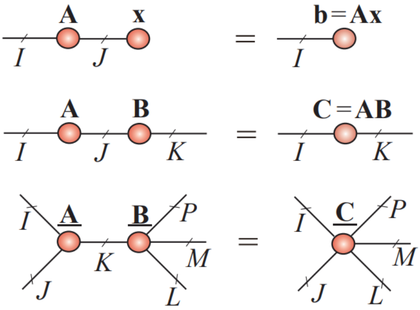
- Можно также определить свёртку тензоров по нескольким индексам одновременно, например, так обозначается свёртка последних 3-ёх мод тензора $\underline{\mathbf{A}}$ и первых трёх мод тензора $\underline{\mathbf{B}}$ (операция требует совпадения соответствующих размерностей):

$$\underline{\mathbf{C}} = \underline{\mathbf{A}} \times_{N,N-1,N-2}^{1,2,3} \underline{\mathbf{B}}$$

$$c_{i_1,\ldots,i_{N-1},j_2,\ldots,j_M} = \sum^{J_1} \sum^{J_2} \sum^{J_3} a_{i_1,\ldots,i_{N-3},j_3,j_2,j_1} b_{j_1,j_2,j_3,\ldots,j_M}$$

Свёртка тензоров

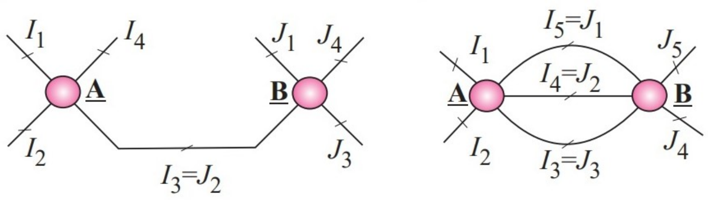
Привычные операции над матрицами можно переписать и проиллюстрировать диаграммами как свёртки соответствующих тензоров:



$$\sum_{k=1}^{K} a_{i,j,k}\, b_{k,l,m,p} = c_{i,j,l,m,p}$$

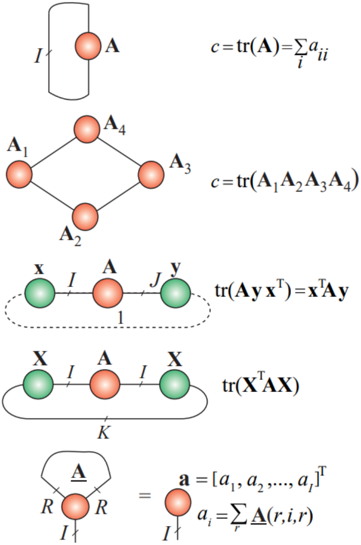
Свёртка тензоров

Примеры свёрток тензоров более высокого порядка и свёрток по нескольким индексам одновременно:



- **Слева:** Свёртка двух тензоров 4-ого порядка вдоль 3-ей моды тензора $\underline{\mathbf{A}}$ и 2-ой моды тензора $\underline{\mathbf{B}}$ даёт тензор 6-ого порядка $\underline{\mathbf{C}} = \underline{\mathbf{A}} \times_3^2 \underline{\mathbf{B}} \in \mathbb{R}^{I_1 \times I_2 \times I_4 \times J_1 \times J_3 \times J_4}$ с элементами $c_{i_1,i_2,i_4,j_1,j_3,j_4} = \sum_{i_3} a_{i_1,i_2,i_3,i_4} b_{j_1,i_3,j_3,j_4}$.
- **Справа:** Свёртка двух тензоров 5-ого порядка вдоль мод 3,4, 5 тензора $\underline{\mathbf{A}}$ и мод 1,2,3 тензора $\underline{\mathbf{B}}$ даёт тензор 4-ого порядка $\underline{\mathbf{C}} = \underline{\mathbf{A}} \times_{5,4,3}^{1,2,3} \underline{\mathbf{B}} \in \mathbb{R}^{I_1 \times I_2 \times J_4 \times J_5}$.

Свёртка тензоров



Другие операции над тензорами

- Произведение n-ой моды тензора $\underline{\mathbf{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ и матрицы $\mathbf{B} \in \mathbb{R}^{J \times I_n}$ даёт тензор $\underline{\mathbf{C}} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ с элементами:

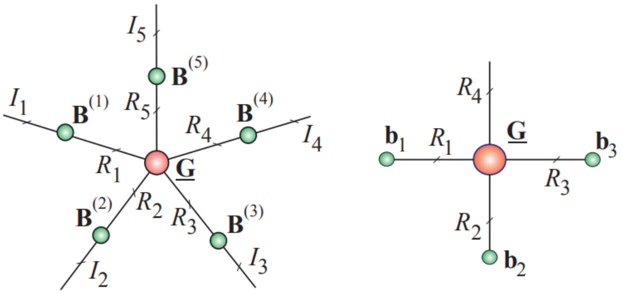
$$\underline{\mathbf{C}} = \underline{\mathbf{A}} \times_n^2 \mathbf{B} = \underline{\mathbf{A}} \times_n \mathbf{B}, \quad c_{i_1,\ldots,i_{n-1},j,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} a_{i_1,\ldots,i_n,\ldots,i_N} b_{j,i_n}$$

- Мультилинейное произведение (произведение Такера) тензора $\underline{\mathbf{G}}$ и матриц $\mathbf{B}^{(n)}$:

$$\underline{\mathbf{C}} = \left[\underline{\mathbf{G}}; \mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(N)}\right] := \underline{\mathbf{G}} \times_1 \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \times_3 \cdots \times_N \mathbf{B}^{(N)}$$

- Произведение n-ой моды тензора $\underline{\mathbf{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ и вектора $\mathbf{b} \in \mathbb{R}^{I_n}$ даёт тензор $\underline{\mathbf{C}} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N}$ с элементами:

$$\underline{\mathbf{C}} = \underline{\mathbf{A}} \times_n^1 \mathbf{b} = \underline{\mathbf{A}} \bar{\times}_n \mathbf{b}, \quad c_{i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} a_{i_1,\ldots,i_n,\ldots,i_N} b_{i_n}$$



- **Слева:** Мультилинейное произведение тензора $\underline{\mathbf{G}} \in \mathbb{R}^{R_1 \times \cdots \times R_5}$ и пяти матриц $\mathbf{B}^{(n)} \in \mathbb{R}^{I_n \times R_n}$ $(n = 1, \ldots, 5)$ даёт тензор $\underline{\mathbf{C}} = \left[\underline{\mathbf{G}}; \mathbf{B}^{(1)}, \ldots, \mathbf{B}^{(5)}\right] =$ $= \underline{\mathbf{G}} \times_1 \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \times_3 \cdots \times_5 \mathbf{B}^{(5)} \in \mathbb{R}^{I_1 \times \cdots \times I_5}$.
- **Справа:** Мультилинейное произведение тензора $\underline{\mathbf{G}} \in \mathbb{R}^{R_1 \times R_2 \times R_3 \times R_4}$ и трёх векторов $\mathbf{b}_n \in \mathbb{R}^{R_n}$ $(n = 1, 2, 3)$ даёт вектор $\mathbf{c} = \underline{\mathbf{G}} \bar{\times}_1 \mathbf{b}_1 \bar{\times}_2 \mathbf{b}_2 \bar{\times}_3 \mathbf{b}_3 \in \mathbb{R}^{R_4}$.

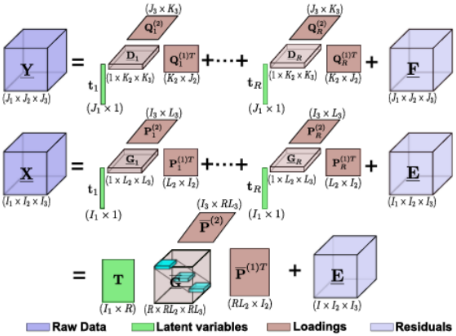
Тензоры как мультилинейное отображение

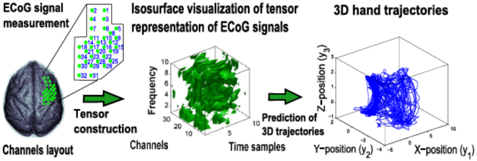
- Напоминание из курса линейной алгебры: любой матрице $\mathbf{A} \in \mathbb{R}^{R_1 \times R_2}$ можно поставить в соответствие:
    1. отображение из пространства $\mathbb{R}^{R_1}$ в пространство $\mathbb{R}^{R_2}$: $v \in \mathbb{R}^{R_1} \to Av \in \mathbb{R}^{R_2}$
    2. отображение из пространства $\mathbb{R}^{R_2}$ в пространство $\mathbb{R}^{R_1}$: $w \in \mathbb{R}^{R_2} \to A^T w \in \mathbb{R}^{R_1}$
    3. билинейную функцию $(\cdot, \cdot)_{\mathbf{A}} : \mathbb{R}^{R_1} \times \mathbb{R}^{R_2} \to \mathbb{R}$: $(v, w) \in \mathbb{R}^{R_1} \times \mathbb{R}^{R_2} \to w^T Av \in \mathbb{R}$
- По аналогии, тензору $\underline{\mathbf{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ порядка N можно сопоставлять различные отображения с помощью операции произведения Такера (мультилинейного произведения), например:
    - отображение из $\mathbb{R}^{I_2} \times \cdots \times \mathbb{R}^{I_N}$ в $\mathbb{R}^{I_1}$:

$$(v^{(2)}, \ldots, v^{(N)}) \to \left[\underline{\mathbf{A}}; I, v^{(2)}, \ldots, v^{(N)}\right] \in \mathbb{R}^{I_1}$$

- Таким образом для тензоров можно определить понятия нормы, сингулярных чисел и т.д. (будет рассмотрено подробнее в дальнейшем).

Пример 1. HOPLS





Пример 1. Tucker decompositions