General Introduction of Tensors

Introduction

In **Linear Algebra**, $\boldsymbol{a} \in \mathbb{R}^{I}$ is a vector, $\boldsymbol{A} \in \mathbb{R}^{I \times J}$ is a matrix, and $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is a N-mode tensor. The ith, $\{i, j\}$th and $\{i_1, \ldots, i_N\}$th entries of $\boldsymbol{\mathcal{A}}$ are denoted by $\boldsymbol{a}(i)$, $\boldsymbol{A}(i, j)$ and $\boldsymbol{\mathcal{A}}(i_1, \ldots, i_N)$.

Products

For matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with same size $I \times J$, the **Hadamard** product is the element-wise product as:

$$\boldsymbol{C} = \boldsymbol{A} \circledast \boldsymbol{B} \in \mathbb{R}^{I \times J} = \begin{pmatrix} a_{1,1}b_{1,1} & a_{1,2}b_{1,2} & \cdots & a_{1,J}b_{1,J} \\ a_{2,1}b_{2,1} & a_{2,2}b_{2,2} & \cdots & a_{2,1}b_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I,1}b_{I,1} & a_{I,2}b_{I,2} & \cdots & a_{I,J}b_{I,J} \end{pmatrix},$$

and the **Kronecker** product with $\boldsymbol{D} \in \mathbb{R}^{K \times L}$ is:

$$\boldsymbol{E} = \boldsymbol{A} \otimes \boldsymbol{D} \in \mathbb{R}^{IK \times JL} = \begin{pmatrix} a_{1,1}\boldsymbol{D} & a_{1,2}\boldsymbol{D} & \cdots & a_{1,J}\boldsymbol{D} \\ a_{2,1}\boldsymbol{D} & a_{2,2}\boldsymbol{D} & \cdots & a_{2,1}\boldsymbol{D} \\ \vdots & \vdots & \ddots & \vdots \\ a_{I,1}\boldsymbol{D} & a_{I,2}\boldsymbol{D} & \cdots & a_{I,J}\boldsymbol{D} \end{pmatrix},$$

and **Khatri-Rao** product of $\boldsymbol{A} \in \mathbb{R}^{I \times K}$ and $\boldsymbol{B} \in \mathbb{R}^{J \times K}$ is defined as:

$$\boldsymbol{A} \odot \boldsymbol{B} = [\boldsymbol{a}_1 \otimes \boldsymbol{b}_1, \boldsymbol{a}_2 \otimes \boldsymbol{b}_2, \cdots, \boldsymbol{a}_K \otimes \boldsymbol{b}_K] \in \mathbb{R}^{IJ \times K},$$

where $\boldsymbol{a}_k$ and $\boldsymbol{b}_k$ are the kth column of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

The **Mode-n** product of tensor $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ and matrix $\boldsymbol{B} \in \mathbb{R}^{J \times I_n}$ is defined by

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}} \times_n \boldsymbol{B} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N},$$

whose entries are

$$\boldsymbol{\mathcal{C}}(i_1, \cdots, j, \cdots, i_N) = \sum_{i_n=1}^{I_n} \boldsymbol{\mathcal{A}}(i_1, \cdots, i_n, \cdots, i_N)\boldsymbol{B}(j, i_n).$$

Decompostions

Canonical Polyadic (**CP**): For a N-mode tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, the CP decompostion is defined by

$$\mathcal{X} = \sum_{r=1}^{R} \boldsymbol{u}_r^{(1)} \circ \boldsymbol{u}_r^{(2)} \circ \cdots \circ \boldsymbol{u}_r^{(N)} = [\![\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \cdots, \boldsymbol{U}^{(N)}]\!],$$

where $\boldsymbol{U}^{(n)} = \left[\boldsymbol{u}_r^{(1)}, \boldsymbol{u}_r^{(2)}, \ldots, \boldsymbol{u}_R^{(n)}\right] \in \mathbb{R}^{I_n \times R}, n = 1, \ldots, N$ are factor matrices.

Tucker:

$$\mathcal{X} = \mathcal{G} \times_1 \boldsymbol{U}^{(1)} \times_2 \boldsymbol{U}^{(2)} \cdots \times_N \boldsymbol{U}^{(N)} = [\![\mathcal{G}; \boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \cdots, \boldsymbol{U}^{(N)}]\!],$$

where $\mathcal{G} \in \mathbb{R}^{R_1 \times R_2 \times \cdots \times R_N}$ is called core tensor.

The n-Rank

For $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, the n-rank of $\mathcal{X}$, denoted by $\operatorname{rank}_n(\mathcal{X})$ is the column rank of $\mathcal{X}_{(n)}$ which is the size of the vector space spanned by the mode-n fibers. If $R_n = \operatorname{rank}_n(\mathcal{X}),\ n = 1, \ldots, N$, we can say that $\mathcal{X}$ is a rank-$(R_1, R_2, \ldots, R_N)$ tensor

HOSVD

The method **HOSVD** is convincing generalization of the matrix SVD and capable of computing the left singular vectors of $\mathcal{X}_{(n)}$. When $R_n < \text{rank}_n(\mathcal{X})$ for one or more n, the decompostion is called truncated HOSVD.

<u>HOSVD</u>():

Input: $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots I_N}$ **and** n-rank $(R_1, R_2, \cdots, R_N)$

Output: $\mathcal{G}, \boldsymbol{U}^{(\boldsymbol{1})}, \boldsymbol{U}^{(\boldsymbol{2})}, \cdots, \boldsymbol{U}^{(\boldsymbol{N})}$

for $n = 1, \cdots, N$ **do**

  $\boldsymbol{U}^{(\boldsymbol{n})} \leftarrow R_n$ Leading left singular vectors of $\mathcal{X}_{(n)}$

// End For

return $\mathcal{G} \leftarrow \mathcal{X} \times_1 \boldsymbol{U}^{(1)T} \times_2 \boldsymbol{U}^{(2)T} \cdots \times_N \boldsymbol{U}^{(N)T}$

HOOI

A good refinement of HOSVD method is the following constrained optimiziation problem:

$$\min_{\mathcal{G}, \boldsymbol{U}^{(\boldsymbol{1})}, \boldsymbol{U}^{(\boldsymbol{2})}, \cdots, \boldsymbol{U}^{(\boldsymbol{N})}} \left\| (\mathcal{X} - [\![\mathcal{G}; \boldsymbol{U}^{(\boldsymbol{1})}, \boldsymbol{U}^{(2)}, \cdots, \boldsymbol{U}^{(\boldsymbol{N})}]\!]) \right\|_F^2$$
$$\text{s.t. } \mathcal{G} \in \mathbb{R}^{R_1, R_2, \cdots, R_N}, \ \boldsymbol{U}^{(\boldsymbol{n})} \in \mathbb{R}^{I_n \times R_n}, \ \boldsymbol{U}^{(\boldsymbol{n})\boldsymbol{T}} \boldsymbol{U}^{(\boldsymbol{n})} = \boldsymbol{I}_{R_n \times R_n}.$$

where after some mathematical operations, the problem boils down to

$$\max_{\boldsymbol{U}^{(\boldsymbol{n})}} \left\| \mathcal{X} \times_0 \boldsymbol{U}^{(1)T} \cdots \times_N \boldsymbol{U}^{(N)T} \right\|_F^2$$
$$\text{s.t. } \boldsymbol{U}^{(n)} \in \mathbb{R}^{I_n \times R_n}, \ \boldsymbol{U}^{(n)T} \boldsymbol{U}^{(n)} = \boldsymbol{I}_{R_n \times R_n},$$

This formulation called **HOOI** method which has a protocol described in

<u>HOOI</u>():

Input: $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots I_N}$ **and** n-rank $(R_1, R_2, \cdots, R_N)$

Output: $\mathcal{G}, \boldsymbol{U}^{(\boldsymbol{1})}, \boldsymbol{U}^{(\boldsymbol{2})}, \cdots, \boldsymbol{U}^{(\boldsymbol{N})}$

Initialize $\boldsymbol{U}^{(\boldsymbol{n})} \in \mathbb{R}^{I_n \times R_n}$ **for** $n = 1, \ldots, N$ using HOSVD,

repeat

for $n = 1, \cdots, N$ **do**

  $\mathcal{Y} \leftarrow \mathcal{X} \times_1 \boldsymbol{U}^{(1)T} \cdots \times_{n-1} \boldsymbol{U}^{(n-1)T} \times_{n+1} \boldsymbol{U}^{(n+1)T} \cdots \times_N \boldsymbol{U}^{(N)T}$

  $\boldsymbol{U}^{(\boldsymbol{n})} \leftarrow R_n$ Leading left singular vectors of $\mathcal{Y}_{(n)}$

// End for

until fit ceases **to** improve **or** maximum iterations exhausted.

return $\mathcal{G} \leftarrow \mathcal{X} \times_1 \boldsymbol{U}^{(1)T} \times_2 \boldsymbol{U}^{(2)T} \cdots \times_N \boldsymbol{U}^{(N)T}$

Bibliography