Tensor Decomposition

Kazeto Fukasawa

2018/1/18

Tensor

Definition 1

- A Tensor is a multidimensional array.

Definition 2

- A N-th order tensor is an element of the tensor product of N vector spaces.

We write a tensor like $\chi \in \mathbb{R}^{I_1 \times I_2 \times I_3}$.

Tensor Product

Vector space V, W.

$\otimes$ is bilinear.

Ex. outer product, kronecker product.

- some books define tensor prod is kronecker prod.

Outer Product

$$(a^{(1)} \circ a^{(2)} \circ \cdots \circ a^{(N)})_{i_1 \cdots i_N} = a^{(1)}_{i_1} a^{(2)}_{i_2} \cdots a^{(N)}_{i_N}.$$

- In some books, outer product is defined only in vector case.

Kronecker Product

$\otimes : (\mathbb{R}^{I \times J}, \mathbb{R}^{K \times L}) \rightarrow \mathbb{R}^{IK \times KL}.$

Property of Kronecker Prod

$$(A \otimes B)(C \otimes D) = AC \otimes BD,$$

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger.$$

Khatri-Rao Product

$\odot : (\mathbb{R}^{I\times K}, \mathbb{R}^{J\times K}) \to \mathbb{R}^{IJ\times K}$.
column-wise kronecker product.

$$A \odot B := (\mathbf{a}_1 \otimes \mathbf{b}_1, \cdots, \mathbf{a}_K \otimes \mathbf{b}_K).$$

property of Khatri-Rao Prod

$$A \odot B \odot C = (A \odot B) \odot C = A \odot (B \odot C),$$

$$(A \odot B)^T (A \odot B) = A^T A * B^T B,$$

$$(A \odot B)^\dagger = ((A^T A * B^T B))^\dagger (A \odot B)^T.$$

Hadamard product

$* : (\mathbb{R}^{I \times J}, \mathbb{R}^{I \times J}) \rightarrow \mathbb{R}^{I \times J}$.

Pseudo Inverse

$A \in \mathbb{R}^{I \times J}$,

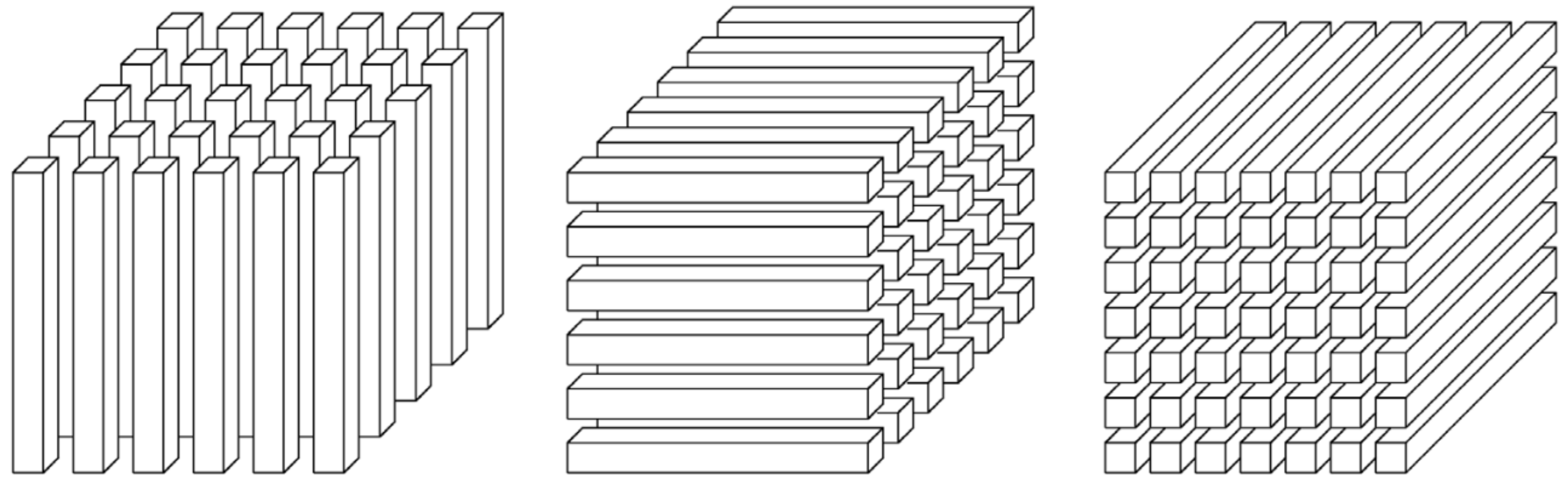
$$A^{\dagger} = (A^T A)^{-1} A.$$

Norm

The norm of a tensor $\chi \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is

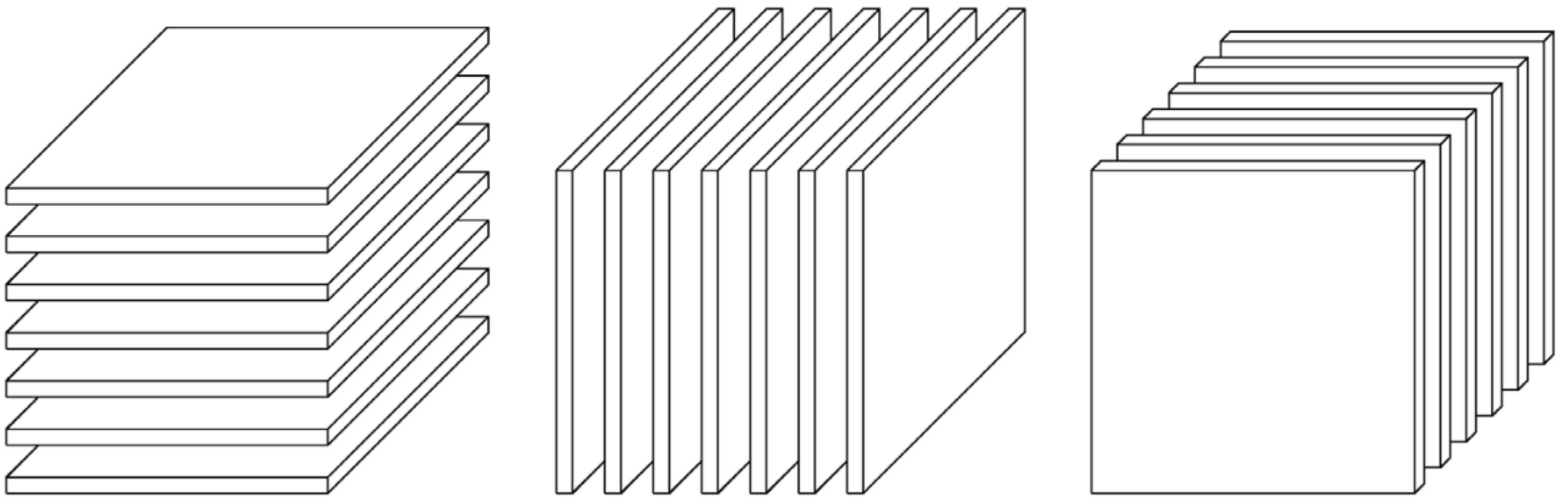
$$\|\chi\| = \sqrt{\sum_{i_1}^{I_1} \cdots \sum_{i_N}^{I_N} x^2_{i_1 \cdots i_N}}.$$

Fibers



(a) Mode-1 (column) fibers: $\mathbf{x}_{:jk}$ (b) Mode-2 (row) fibers: $\mathbf{x}_{i:k}$ (c) Mode-3 (tube) fibers: $\mathbf{x}_{ij:}$

Slices



(a) Horizontal slices: $\mathbf{X}_{i::}$ (b) Lateral slices: $\mathbf{X}_{:j:}$ (c) Frontal slices: $\mathbf{X}_{::k}$ (or $\mathbf{X}_k$)

Rank-One Tensors

A N-way tensor $\chi \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is rank one if

$$\chi = \mathbf{a}^{(1)} \circ \cdots \circ \mathbf{a}^{(N)}.$$

Each element is

$$x_{i_1 \cdots i_N} = a_{i_1}^{(1)} \cdots a_{i_N}^{(N)}.$$

multi linearity

CP Decomposition

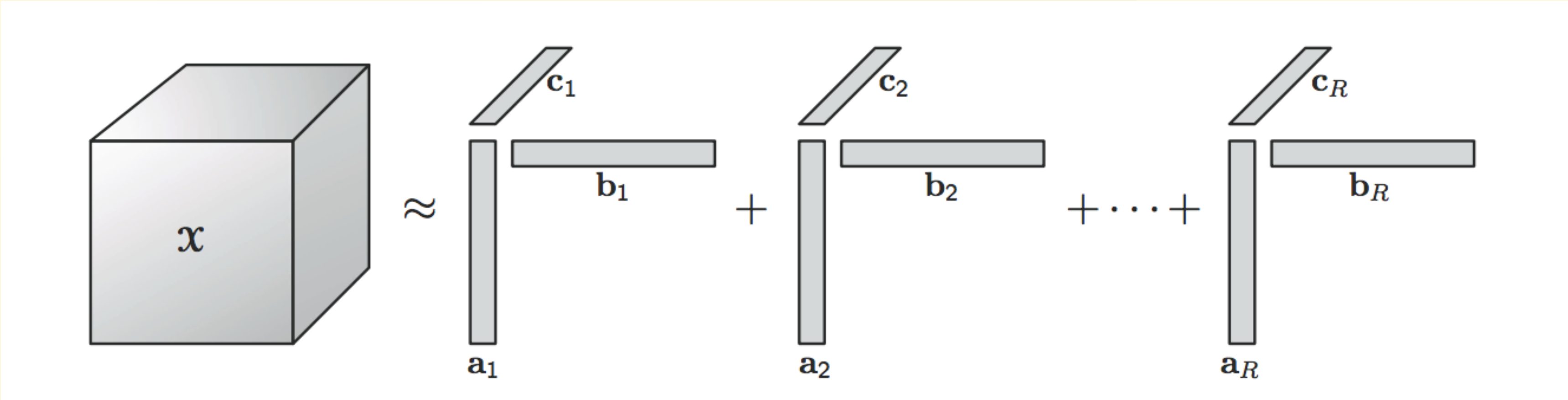
Let
$\chi \in \mathbb{R}^{I_1 \times I_2 \times I_3}$,
$R, I_1, I_2, I_3 \in \mathbb{N}$,
$A \in \mathbb{R}^{I_1 \times R}$, $B \in \mathbb{R}^{I_2 \times R}$, $C \in \mathbb{R}^{I_3 \times R}$, $\lambda \in \mathbb{R}^{R}$.

Then CP decomposition is

$$\min \|\chi - \hat{\chi}\|,$$

$$\hat{\chi} = \sum_{r=1}^{R} \lambda_r \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r .$$

CP Decomposition



ALS Algorithm

- Alternative Least Square (ALS) Algorithm.

procedure CP-ALS($\mathcal{X}$,R)
  initialize $\mathbf{A}^{(n)} \in \mathbb{R}^{I_n \times R}$ for $n = 1, \ldots, N$
  repeat
    for $n = 1, \ldots, N$ **do**
      $\mathbf{V} \leftarrow \mathbf{A}^{(1)\mathsf{T}}\mathbf{A}^{(1)} * \cdots * \mathbf{A}^{(n-1)\mathsf{T}}\mathbf{A}^{(n-1)} * \mathbf{A}^{(n+1)\mathsf{T}}\mathbf{A}^{(n+1)} * \cdots * \mathbf{A}^{(N)\mathsf{T}}\mathbf{A}^{(N)}$
      $\mathbf{A}^{(n)} \leftarrow \mathbf{X}_{(n)}(\mathbf{A}^{(N)} \odot \cdots \odot \mathbf{A}^{(n+1)} \odot \mathbf{A}^{(n-1)} \odot \cdots \odot \mathbf{A}^{(1)})\mathbf{V}^{\dagger}$
      normalize columns of $\mathbf{A}^{(n)}$ (storing norms as $\boldsymbol{\lambda}$)
    end for
  until fit ceases to improve or maximum iterations exhausted
  return $\boldsymbol{\lambda}, \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)}$
end procedure

What is ALS

fix factor matrices exept for one, such as A, using matricisation, we get

$$\hat{X}_{(1)} = A(C \odot B)^T.$$

then objective function is

$$\min_A \|X_{(1)} - A(C \odot B)^T\|_F.$$

What is ALS

The optimal solution is

$$\hat{A} = X_{(1)}\{(C \odot B)^T\}^\dagger.$$

using these properties,

$$(X^T)^\dagger = (X^\dagger)^T,$$

$$(A \odot B)^\dagger = ((A^T A * B^T B))^\dagger (A \odot B)^T,$$

we have

$$\hat{A} = X_{(1)}(C \odot B)(C^T C * B^T B)^\dagger.$$

Uniqueness Condition

CP decomposition is unique if Kruscal Condition is satisfied.

Kruscal Condition
sum of rank of factor matricise is over 2R+2.

Sample Code

on monitor

Summary

A tensor was created by outer product of vectors like CP decomposition.
This is the meaning of Definition 2️⃣.

Application Fields

- Psychometrics
- Chemometrics 🔬
- Neurosicence 🐵
- Text Analysis ✏️
- Data Mining
- Physics 🍎
- etc.

Reference

- Kolda, T. G., & Bader, B. W. (2009). Tensor decompositions and applications. SIAM review, 51(3), 455-500.
- Introduction to Tensor Decompositions and their Applications in Machine Learning, arXiv.
-