A Tensor SVD

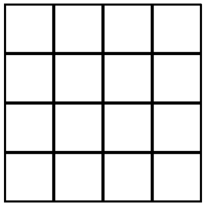
Householder Symposium XVII
June 5, 2008

Carla D. Martin
(with M.E. Kilmer and L. Perrone)

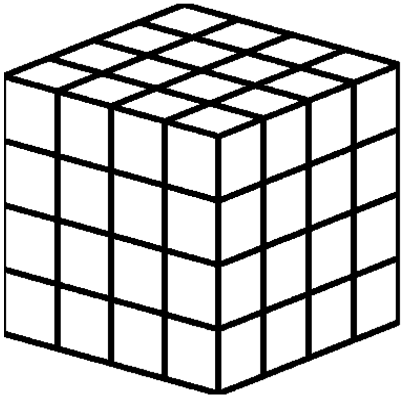
Department of Mathematics and Statistics
James Madison University

What are Tensors?

- Second-order tensor $A = (a_{ij}) \in \mathbb{R}^{n_1 \times n_2}$



- Third-order tensor $\mathcal{A} = (a_{ijk}) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$



- p^{th}-order tensor $\mathcal{A} = (a_{i_1 i_2 \ldots i_p}) \in \mathbb{R}^{n_1 \times \cdots \times n_p}$

Leading Applications

- Chemometrics
- Psychometrics
- Computer Vision
- Signal Processing
- Data Mining
- Neuroscience
- Other applications using multiway data analysis

Motivation: A two-way decomposition

Suppose $U \in \mathbb{R}^{M \times M}$, $V \in \mathbb{R}^{N \times N}$ are orthogonal, and $\Sigma = U^T AV$, then

$$A = U\Sigma V^T = \sum_{i=1}^{M} \sum_{j=1}^{N} \sigma_{ij} u_i v_j^T$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} \sigma_{ij} (u_i \circ v_j)$$

where $u_i = U(:, i)$, $v_j = V(:, j)$

Tensor Decompositions

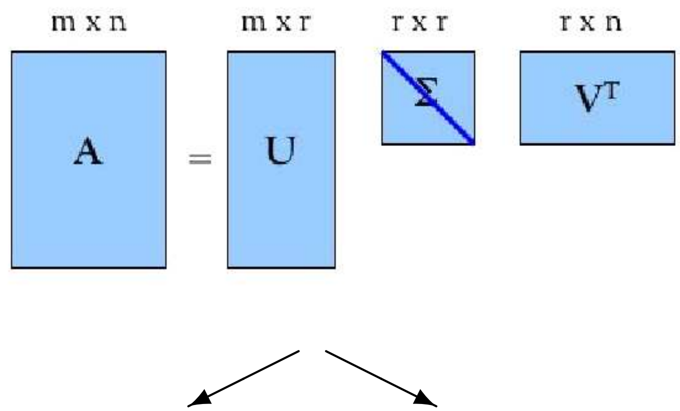
Let $\mathcal{A} \in \mathbb{R}^{M \times N \times P}$

Goal: To find $U \in \mathbb{R}^{M \times M}$, $V \in \mathbb{R}^{N \times N}$, $W \in \mathbb{R}^{P \times P}$, and $\Sigma = (\sigma_{ijk}) \in \mathbb{R}^{M \times N \times P}$ such that

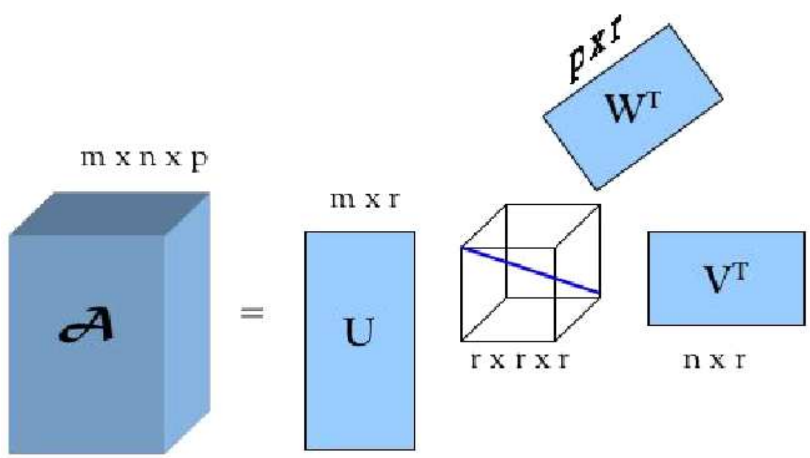
$$\mathcal{A} = \sum_{i=1}^{M} \sum_{j=1}^{N} \sum_{k=1}^{P} \sigma_{ijk} (u_i \circ v_j \circ w_k)$$

where $u_i = U(:, i)$, $v_j = V(:, j)$, $w_k = W(:, k)$

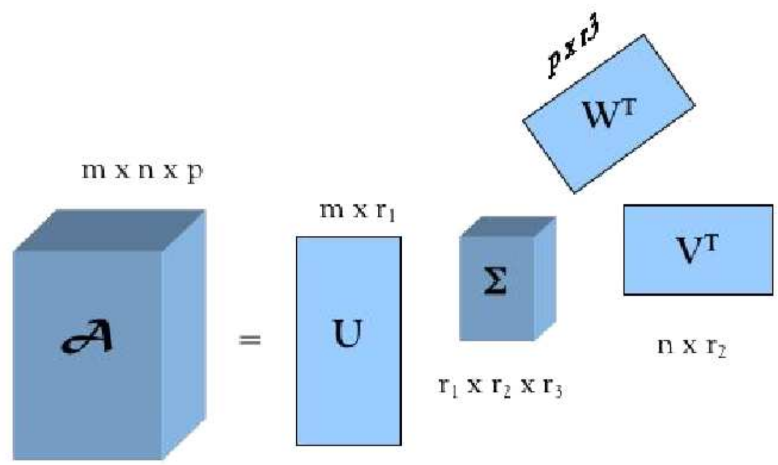
Orthogonal *or* Diagonal for Tensors



Case 1:
Diagonal Σ (CP)



Case 2:
Orthogonal U, V, W



A new Tensor SVD

Questions:

- Is there another rank-revealing “SVD-like” tensor decomposition?
- How does it compare to existing decompositions (HOSVD, PARAFAC, etc.)?
- Is it useful?

Tensor-tensor Multiplication

Contracted product in the *first-mode*:

$$\begin{array}{l} \mathcal{A} \in \mathbb{R}^{L \times M_1 \times N_1} \\ \mathcal{B} \in \mathbb{R}^{L \times M_2 \times N_2} \end{array} \quad \Rightarrow \quad \mathcal{A}\mathcal{B} \in \mathbb{R}^{M_1 \times N_1 \times M_2 \times N_2}$$

$$\boxed{(\mathcal{A}\mathcal{B})_{m_1 n_1 m_2 n_2} = \sum_{\ell=1}^{L} \mathcal{A}_{\ell m_1 n_1} \mathcal{B}_{\ell m_2 n_2}}$$

$$\begin{aligned} m_1 &= 1, \ldots, M_1 & \qquad n_1 &= 1, \ldots, N_1 \\ m_2 &= 1, \ldots, M_2 & \qquad n_2 &= 1, \ldots, N_2 \end{aligned}$$

Tensor-tensor Multiplication

Using contracted product...

- Set of all third-order tensors is not closed
- No notion of inverse possible

New multiplication operation defined that is closed, inverse exists, that gives way to a new way of thinking about the SVD

New tensor-tensor multiplication

$$\begin{array}{l} \mathcal{A} \in \mathbb{R}^{L\times M\times N} \\ \mathcal{B} \in \mathbb{R}^{M\times P\times N} \end{array} \quad \Rightarrow \quad \mathcal{A} * \mathcal{B} \in \mathbb{R}^{L\times P\times N}$$

- Multiplication defined in terms of the tensor “slices”
- Circulant matrices play a role
- Operation is associative
- Can define an inverse
- Set of $N \times N \times N$ invertible tensors form a group under this multiplication

`fold` operator

Given a set of p matrices $A_1, \ldots, A_N \in \mathbb{R}^{L \times M}$,

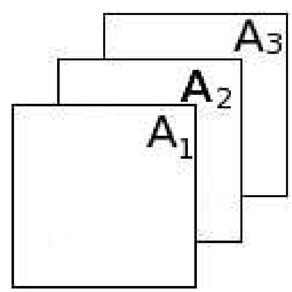
$$\mathcal{A} = \texttt{fold}(\{A_1, \ldots, A_N\}, i) \in \mathbb{R}^{L \times M \times N}$$

creates a tensor $\mathcal{A}$ from the N matrices along the i-th dimension.

Example:

Suppose $A_1, A_2, A_3 \in \mathbb{R}^{L \times M}$. Then

$$\mathcal{A} = \texttt{fold}(\{A_1, A_2, A_3\}, 3) =$$

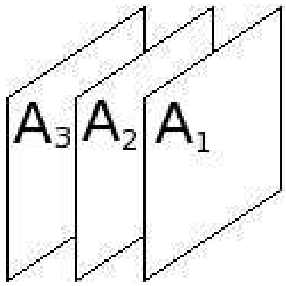


$$\in \mathbb{R}^{L \times M \times 3}$$

`fold` operator

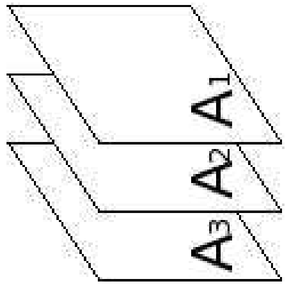
Similarly,

$$\mathcal{A} = \texttt{fold}(\{A_1, A_2, A_3\}, 2) =$$



$$\in \mathbb{R}^{L \times 3 \times M}$$

$$\mathcal{A} = \texttt{fold}(\{A_1, A_2, A_3\}, 1) =$$



$$\in \mathbb{R}^{3 \times L \times M}$$

New tensor-tensor multiplication

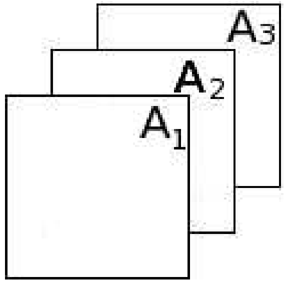
Let $\mathcal{A} \in \mathbb{R}^{L\times M\times N}$ with $L \times M$ faces $A_1, \ldots, A_N$, and $\mathcal{B} \in \mathbb{R}^{M\times P\times N}$ with $M \times P$ faces $B_1, \ldots, B_N$. Then

$$\mathcal{A} * \mathcal{B} = \texttt{fold}\left(\begin{bmatrix} A_1 & A_N & A_{N-1} & \ldots & A_2 \\ A_2 & A_1 & A_N & \ldots & A_3 \\ \vdots & A_2 & \ddots & \ddots & \vdots \\ A_{N-1} & \vdots & \ddots & A_1 & A_N \\ A_N & A_{N-1} & \ldots & A_2 & A_1 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_N \end{bmatrix}, 3 \right)$$

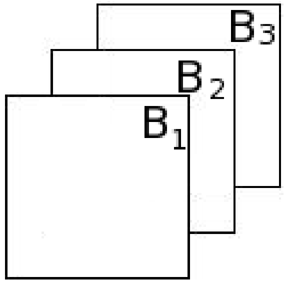
$$\mathcal{A} * \mathcal{B} \in \mathbb{R}^{L\times P\times N}$$

Example

$$\mathcal{A} \in \mathbb{R}^{N \times N \times 3} \quad \Rightarrow \quad \mathcal{A} \quad =$$



$$\mathcal{B} \in \mathbb{R}^{N \times N \times 3} \quad \Rightarrow \quad \mathcal{B} \quad =$$



$$\mathcal{A} * \mathcal{B} = \texttt{fold}\left(\begin{bmatrix} A_1 & A_3 & A_2 \\ A_2 & A_1 & A_3 \\ A_3 & A_2 & A_1 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}, 3 \right)$$

Identity

The $N \times N \times P$ **identity Tensor**, $\mathcal{I}$, is the tensor whose frontal face is the $N \times N$ identity matrix and whose other faces are zeros.

Example:

The $2 \times 2 \times 3$ identity tensor is defined by

$$\mathcal{I} = \texttt{fold}\left(\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}, 3 \right)$$

In general, $\mathcal{A} * \mathcal{I} = \mathcal{I} * \mathcal{A} = \mathcal{A}$

Inverse

Let $\mathcal{A} \in \mathbb{R}^{N \times N \times N}$. Then the tensor inverse of $\mathcal{A}$ is any tensor $\mathcal{B} \in \mathbb{R}^{N \times N \times N}$ such that

$$\mathcal{A} * \mathcal{B} = \mathcal{B} * \mathcal{A} = \mathcal{I}$$

We denote the inverse of $\mathcal{A}$ as $\mathcal{A}^{-1}$.

It follows that $(\mathcal{A} * \mathcal{B})^{-1} = \mathcal{B}^{-1} * \mathcal{A}^{-1}$

Transpose

Let $\mathcal{A} \in \mathbb{R}^{L\times M\times N}$ with faces $A_1, \ldots, A_N \in \mathbb{R}^{L\times M}$. Then

$$\mathcal{A}^T = \texttt{fold}\left(\{A_1^T, A_N^T, \ldots, A_2^T\}, 3\right) \in \mathbb{R}^{M\times L\times N}$$

It follows that $(\mathcal{A} * \mathcal{B})^T = \mathcal{B}^T * \mathcal{A}^T$

Permutation Tensors

Let $\mathcal{P} = (p_{ijk}) \in \mathbb{R}^{N \times N \times P}$.

$\mathcal{P}$ is a **permutation tensor** if there are exactly N entries of unity, such that there is at most one nonzero entry in row i, column j, and slice k.

Example:

A $3 \times 3 \times 2$ permutation tensor:

$$\mathcal{I} = \texttt{fold}\left(\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\}, 3 \right)$$

Frobenius Norm

Let $\mathcal{A} = (a_{ijk}) \in \mathbb{R}^{L \times M \times N}$.

Then the **Frobenius norm** of $\mathcal{A}$ is

$$||\mathcal{A}||_F = \sqrt{\sum_{i=1}^{L} \sum_{j=1}^{M} \sum_{k=1}^{N} a_{ijk}^2}$$

Orthogonality

Let $\mathcal{Q} \in \mathbb{R}^{N \times N \times P}$.

$\mathcal{Q}$ is **orthogonal** if

$$\mathcal{Q}^T * \mathcal{Q} = \mathcal{Q} * \mathcal{Q}^T = \mathcal{I}$$

If $\mathcal{A}$ is a tensor, then it follows that

$$||\mathcal{Q} * \mathcal{A}||_F = ||\mathcal{A}||_F$$

Tensor SVD

Let $\mathcal{A} \in \mathbb{R}^{L \times M \times N}$. Then $\mathcal{A}$ can be factored as

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$$

where $\mathcal{U} \in \mathbb{R}^{L \times L \times N}$ and $\mathcal{V} \in \mathbb{R}^{M \times M \times N}$ are orthogonal tensors and $\mathcal{S} \in \mathbb{R}^{L \times M \times N}$ has diagonal matrix faces.

If $\mathcal{A} \in \mathbb{R}^{N \times N \times N}$,

$$\mathcal{A} = \sum_{i=1}^{N} \mathcal{U}(:, i, :) * \mathcal{S}(i, i, :) * \mathcal{V}(:, i, :)^T$$

Tensor SVD: computation

- Computation of the tensor SVD involves SVDs of block diagonal elements obtained from block diagonalizing the circulant matrix generated by $\mathcal{A}$
- Using the SVDs of the blocks leads to algorithms for compression
- Operation extends recursively to order-p tensors when $p > 3$

One Compression Strategy

Suppose $\mathcal{A} \in \mathbb{R}^{N \times N \times N}$

Can prove that if $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ then

$$\sum_{i=1}^{N} \mathcal{U}(:,:,i) \qquad \text{is orthogonal.}$$

Therefore

$$\sum_{i=1}^{N} \mathcal{A}(:,:,i) = \left(\sum_{i=1}^{N} \mathcal{U}(:,:,i) \right) \left(\sum_{i=1}^{N} \mathcal{S}(:,:,i) \right) \left(\sum_{i=1}^{N} \mathcal{V}(:,:,i) \right)^T$$

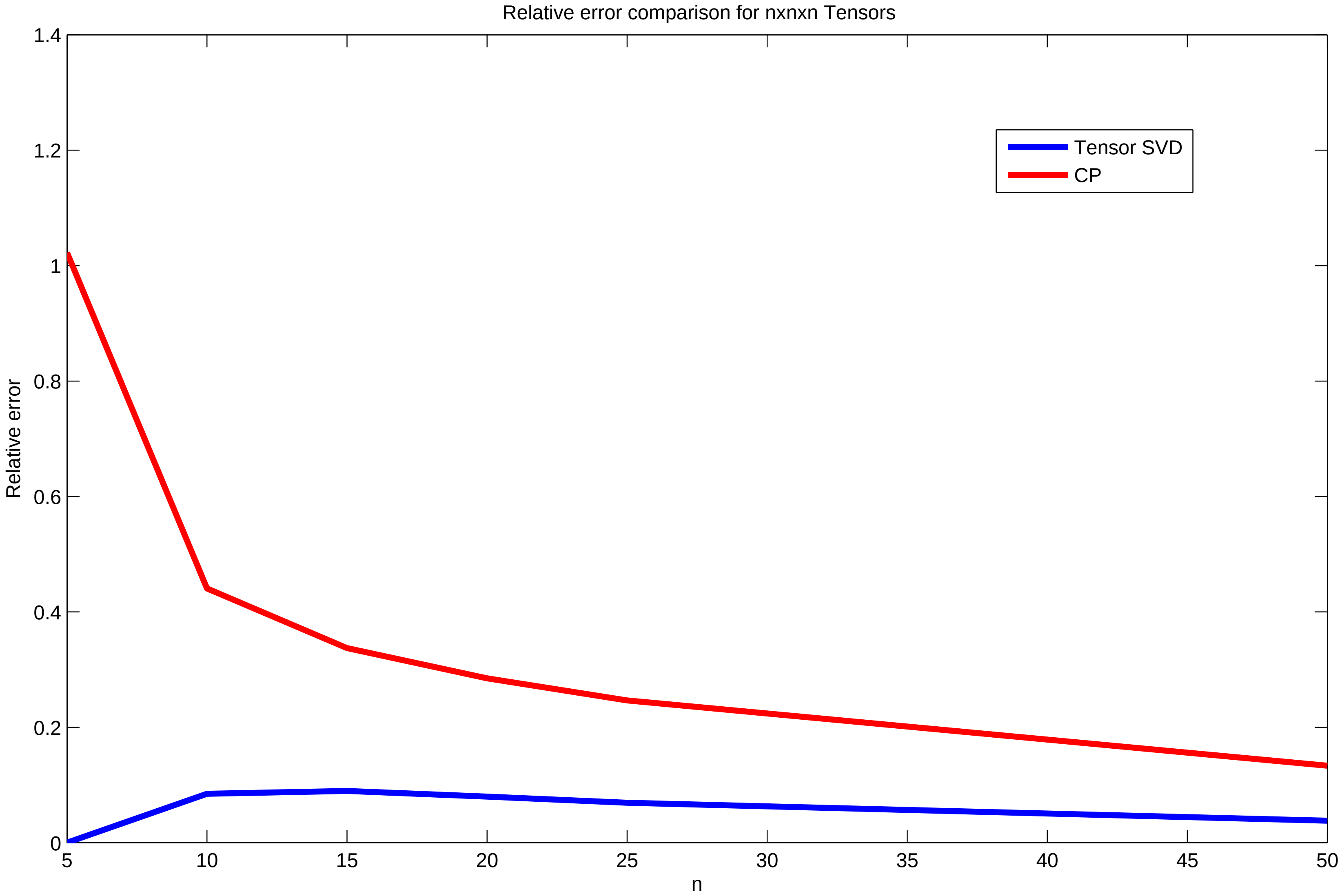
One Compression Strategy

$$\sum_{i=1}^{N} \mathcal{A}(:,:,i) = \left(\sum_{i=1}^{N} \mathcal{U}(:,:,i)\right) \left(\sum_{i=1}^{N} \mathcal{S}(:,:,i)\right) \left(\sum_{i=1}^{N} \mathcal{V}(:,:,i)\right)^T$$

- Take rank-k approximation and "build" back approximation to the tensor $\mathcal{A}$
- Can rewrite "compressed" tensor $\mathcal{A}_c$ as sum of outer products:

$$\mathcal{A}_c = \sum_{i=1}^{k} \sum_{j=1}^{k} \tilde{U}(:,i) \circ \tilde{V}(:,j) \circ \mathcal{M}(i,j,:)$$

CP and Tensor SVD



Tensor QR

Let $\mathcal{A} \in \mathbb{R}^{L\times M\times N}$. Then $\mathcal{A}$ can be factored as

$$\mathcal{A} = \mathcal{Q} * \mathcal{R}$$

where $\mathcal{Q} \in \mathbb{R}^{L\times L\times N}$ is an orthogonal tensor and $\mathcal{R} \in \mathbb{R}^{M\times M\times N}$ has upper triangular matrix faces.

$$\mathcal{A} = \mathcal{Q} * \mathcal{R}$$

A Tensor SVD

Questions from before:

- Is there another rank-revealing "SVD-like" tensor decomposition?
- How does it compare to existing decompositions (HOSVD, PARAFAC, etc.)?
- **Is it useful?**

Thank you!

carlam@math.jmu.edu