From Matrix to Tensor:
The Transition to Numerical Multilinear Algebra

Lecture 4. Tensor-Related Singular Value Decompositions

Charles F. Van Loan

Cornell University

The Gene Golub SIAM Summer School 2010
Selva di Fasano, Brindisi, Italy

Where We Are

Lecture 1.	Introduction to Tensor Computations
Lecture 2.	Tensor Unfoldings
Lecture 3.	Transpositions, Kronecker Products, Contractions
Lecture 4.	Tensor-Related Singular Value Decompositions
Lecture 5.	The CP Representation and Tensor Rank
Lecture 6.	The Tucker Representation
Lecture 7.	Other Decompositions and Nearness Problems
Lecture 8.	Multilinear Rayleigh Quotients
Lecture 9.	The Curse of Dimensionality
Lecture 10.	Special Topics

What is this Lecture About?

Pushing the SVD Envelope

The SVD is a powerful tool for exposing the structure of a matrix and for capturing its essence through optimal, data-sparse representations:

$$A \;=\; \sum_{k=1}^{\mathrm{rank}(A)} \sigma_k u_k v_k^T \;\approx\; \sum_{k=1}^{\hat{r}} \sigma_k u_k v_k^T$$

For a tensor $\mathcal{A}$, let's try for something similar...

$$\mathcal{A} \;=\; \sum \text{whatever!}$$

What is this Lecture About?

The Higher Order Singular Value Decomposition (HOSVD)

The HOSVD of a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ involves computing the matrix SVDs of its modal unfoldings $\mathcal{A}_{(1)}, \ldots, \mathcal{A}_{(d)}$. This results in a representation of $\mathcal{A}$ as a sum of rank-1 tensors.

Before we present the HOSVD, we need to understand the mode-k matrix product which can be thought of as a tensor-times-matrix product.

What is this Lecture About?

The Kronecker Product Singular Value Decomposition (KPSVD)

The KPSVD of a matrix A represents A as an SVD-like sum of Kronecker products. Applied to an unfolding of $\mathcal{A}$, it results in a representation of $\mathcal{A}$ as a sum of low rank tensors.

Before we present the KPSVD, we need to understand the nearest Kronecker Product problem.

The Mode-k Matrix Product

Main Idea

A mode-k matrix product is a special contraction that involves a matrix and a tensor. A new tensor of the same order is obtained by applying the matrix to each mode-k fiber of the tensor.

The Mode-k Matrix Product

A mode-1 example when the Tensor $\mathcal{A}$ is Second Order...

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

(The fibers of $\mathcal{A}$ are its columns.)

A mode-2 example when the Tensor $\mathcal{A}$ is Second Order...

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \\ m_{41} & m_{42} & m_{43} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

(The fibers of $\mathcal{A}$ are its rows.)

The matrix doesn't have to be square.

The Mode-k Matrix Product

A Mode-2 Example When $A \in \mathbb{R}^{4\times 3\times 2}$

$$\begin{bmatrix} b_{111} & b_{211} & b_{311} & b_{411} & b_{112} & b_{212} & b_{312} & b_{412} \\ b_{121} & b_{221} & b_{321} & b_{421} & b_{122} & b_{222} & b_{322} & b_{422} \\ b_{131} & b_{231} & b_{331} & b_{431} & b_{132} & b_{232} & b_{332} & b_{432} \\ b_{141} & b_{241} & b_{341} & b_{441} & b_{142} & b_{242} & b_{342} & b_{442} \\ b_{151} & b_{251} & b_{351} & b_{451} & b_{152} & b_{252} & b_{352} & b_{452} \end{bmatrix}$$

$$=$$

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \\ m_{41} & m_{42} & m_{43} \\ m_{51} & m_{52} & m_{53} \end{bmatrix} \begin{bmatrix} a_{111} & a_{211} & a_{311} & a_{411} & a_{112} & a_{212} & a_{312} & a_{412} \\ a_{121} & a_{221} & a_{321} & a_{421} & a_{122} & a_{222} & a_{322} & a_{422} \\ a_{131} & a_{231} & a_{331} & a_{431} & a_{132} & a_{232} & a_{332} & a_{432} \end{bmatrix}$$

Note that (1) $B \in \mathbb{R}^{4\times 5\times 2}$ and (2) $\mathcal{B}_{(2)} = M \cdot \mathcal{A}_{(2)}$.

The Mode-k Matrix Product

In General...

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $M \in \mathbb{R}^{m_k \times n_k}$, then

$$\mathcal{B}(\alpha_1, \ldots, \alpha_{k-1}, i, \alpha_{k+1}, \ldots, \alpha_d)$$

$$=$$

$$\sum_{j=1}^{n_k} M(i,j) \cdot \mathcal{A}(\alpha_1, \ldots, \alpha_{k-1}, j, \alpha_{k+1}, \ldots, \alpha_d)$$

is the mode-k matrix product of M and $\mathcal{A}$.

Multiply All the Mode-k Fibers by M...

$$\mathcal{B}_{(k)} \;=\; M \cdot \mathcal{A}_{(k)}$$

The Mode-k Matrix Product

Notation

The mode-k matrix product of M and $\mathcal{A}$ is denoted by

$$\mathcal{B} \;=\; \mathcal{A} \times_k M$$

Thus, if $\mathcal{B} \;=\; \mathcal{A} \times_k M$, then $\mathcal{B}$ is defined by $\mathcal{B}_{(k)} \;=\; M \cdot \mathcal{A}_{(k)}$.

Order and Dimension

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, $M \in \mathbb{R}^{m_k \times n_k}$, and $\mathcal{B} \;=\; \mathcal{A} \times_k M$, then

$$\mathcal{B} = \mathcal{B}(1{:}n_1, \ldots, 1{:}n_{k-1}, 1{:}m_k, 1{:}n_{k+1}, \ldots, 1{:}n_k).$$

Thus, $\mathcal{B}$ has the same order as $\mathcal{A}$.

MATLAB Tensor Toolbox: **Mode-k Matrix Product Using ttm**

```
n = [2 5 4 7];
A = tenrand(n);
for k=1:4
   M = randn(n(k),n(k));
   % Compute the mode-k product of tensor A
   % and matrix M to obtain tensor B...
   B = ttm(A,M,k);
end
```

Problem 4.1. Write a function `B = Myttm(A,M,k)` that does the same thing as `ttm`. Make effective use of `tenmat`.

Problem 4.2. If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $v \in \mathbb{R}^{n_k}$, then the mode-k vector product is defined by

$$\mathcal{B}(\alpha_1, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_d)$$

$$=$$

$$\sum_{j=1}^{n_k} v(j) \cdot \mathcal{A}(\alpha_1, \ldots, \alpha_{k-1}, j, \alpha_{k+1}, \ldots, \alpha_d)$$

The Tensor Toolbox function `ttv` (for tensor-times-vector) can be used to compute this contraction. Note that

$$\mathcal{B} = \mathcal{B}(1{:}n_1, \ldots, 1{:}n_{k-1}, 1{:}n_{k+1}, \ldots, 1{:}n_k).$$

Thus, the order of $\mathcal{B}$ is one less than the order of $\mathcal{A}$.

Write a MATLAB function `B = Myttv(M,A,k)` that carries out the mode-k vector product. Use the fact that $\mathcal{B}$ is a reshaping of $v^T \mathcal{A}_{(k)}$.

The Mode-k Matrix Product

Property 1. (Reshaping as a Matrix-Vector Product)

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, $M \in \mathbb{R}^{m_k \times n_k}$, and

$$\mathcal{B} \;=\; \mathcal{A} \times_k M$$

then

$$\begin{aligned} \text{vec}(\mathcal{B}) \;&=\; \left(I_{n_d} \otimes \cdots \otimes I_{n_{k+1}} \otimes M \otimes I_{n_{k-1}} \otimes \cdots \otimes I_{n_1}\right) \cdot \text{vec}(\mathcal{A}) \\ &=\; \left(I_{n_{k+1} \cdots n_d} \otimes M \otimes I_{n_1 \cdots n_{k-1}}\right) \cdot \text{vec}(\mathcal{A}) \end{aligned}$$

Problem 4.3. Prove this. Hint. Show that if $\mathbf{p} = [1{:}k-1\;\; k+1{:}d\;\; 1]$ then $(\mathcal{A} \times_k M)^{[\,\mathbf{p}\,]} = \mathcal{A}^{[\,\mathbf{p}\,]} \times_1 M$.

The Mode-k Matrix Product

Property 2. (Successive Products in the Same Mode)

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $M_1, M_2 \in \mathbb{R}^{m_k \times n_k}$, then

$$(\mathcal{A} \times_k M_1) \times_k M_2 \;=\; \mathcal{A} \times_k (M_1 M_2).$$

Problem 4.4. Prove this using Property 1 and the Kronecker product fact that $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

The Mode-k Matrix Product

Property 3. (Successive Products in Different Modes)

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, $M_k \in \mathbb{R}^{m_k \times n_k}$, $M_j \in \mathbb{R}^{m_j \times n_j}$, and $k \neq j$, then

$$(\mathcal{A} \times_k M_k) \times_j M_j = (\mathcal{A} \times_j M_j) \times_k M_k$$

Problem 4.5. Prove this using Property 1 and the Kronecker product fact that $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

The Tucker Product

Definition

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $M_k \in \mathbb{R}^{m_k \times n_k}$ for $k = 1{:}d$, then

$$\mathcal{B} \; = \; \mathcal{A} \times_1 M_1 \times_2 M_2 \cdots \times_d M_d$$

is a Tucker product of $\mathcal{A}$ and $M_1, \ldots, M_d$.

MATLAB Tensor Toolbox: **Tucker Products**

```
  function B = TuckerProd(A,M)
% A is a n(1)-by-...-n(d) tensor.
% M is a length-d cell array with
%     M{k} an m(k)-by-n(k) matrix.
% B is an m(1)-by-...-by-m(d) tensor given by
%     B = A x1 M{1} x2 M{2} ... xd M{d}
% where "xk" denotes mode-k matrix product.
  B = A;
  for k=1:length(A.size)
     B = ttm(B,M{k},k);
  end
```

The Tucker Product Representation

Before...

Given $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $M_k \in \mathbb{R}^{m_k \times n_k}$ for $k = 1{:}d$, compute the Tucker product

$$\mathcal{B} \; = \; \mathcal{A} \times_1 M_1 \; \times_2 M_2 \cdots \; \times_d M_d$$

Now...

Given $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, find matrices $U_k \in \mathbb{R}^{n_k \times r_k}$ for $k = 1{:}d$ and tensor $\mathcal{S} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$ such that

$$\mathcal{A} \; = \; \mathcal{S} \times_1 U_1 \; \times_2 U_2 \cdots \times_d U_d$$

is an "illuminating" Tucker product representation of $\mathcal{A}$.

The Tucker Product Representation

Multilinear Sum Formulation

Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $U_k \in \mathbb{R}^{n_k \times r_k}$. If

$$\mathcal{A} \;=\; \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$$

where $\mathcal{S} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$, then

$$\mathcal{A}(\mathbf{i}) \;=\; \sum_{\mathbf{j}=\mathbf{1}}^{\mathbf{r}} \mathcal{S}(\mathbf{j}) U_1(i_1, j_1) \cdots U_d(i_d, j_d)$$

The Tucker Product Representation

Matrix-Vector Product Formulation

Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $U_k \in \mathbb{R}^{n_k \times r_k}$. If

$$\mathcal{A} \;=\; \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$$

where $\mathcal{S} \in \mathbb{R}^{r_1 \times \cdots \times r_d}$, then

$$\text{vec}(\mathcal{A}) \;=\; (U_d \otimes \cdots \otimes U_1) \cdot \text{vec}(\mathcal{S})$$

The Tucker Product Representation

Important Existence Situation

If $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $U_k \in \mathbb{R}^{n_k \times n_k}$ is nonsingular for $k = 1{:}d$, then

$$\mathcal{A} \;=\; \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$$

with

$$\mathcal{S} \;=\; \mathcal{A} \times_1 U_1^{-1} \times_2 U_2^{-1} \cdots \times_d U_d^{-1}.$$

We will refer to the U_k as the inverse factors and $\mathcal{S}$ as the core tensor.

Proof.

$$\begin{aligned} \mathcal{S} \times_1 U_1 \times_2 U_2 \times_3 U_3 \;&=\; \left(\mathcal{A} \times_1 U_1^{-1} \times_2 U_2^{-1} \times_3 U_3^{-1}\right) \times_1 U_1 \times_2 U_2 \times_3 U_3 \\ &=\; \mathcal{A} \times_1 (U_1^{-1}U_1) \times_2 (U_2^{-1}U_2) \times_3 (U_3^{-1}U_3) \;=\; \mathcal{A} \end{aligned}$$

The Tucker Product Representation

Core Tensor Unfoldings

Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and $U_k \in \mathbb{R}^{n_k \times n_k}$ is nonsingular for $k = 1{:}d$. If

$$\mathcal{A} \;=\; \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$$

then the mode-k unfoldings of $\mathcal{S}$ satisfy

$$\mathcal{S}_{(k)} \;=\; U_k^{-1} \mathcal{A}_{(k)} \left(U_d \otimes \cdots \otimes U_{k+1} \otimes U_{k-1} \otimes \cdots \otimes U_1 \right)$$

The Higher-Order SVD (HOSVD)

Definition

Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and that its mode-k unfolding has SVD

$$\mathcal{A}_{(k)} \;=\; U_k \Sigma_k V_k^T$$

for $k = 1{:}d$. Its higher-order SVD is given by

$$\mathcal{A} \;=\; \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$$

where

$$\mathcal{S} \;=\; \mathcal{A} \times_1 U_1^T \times_2 U_2^T \cdots \times_d U_d^T$$

A Tucker product representation where the inverse factor matrices are the left singular vector matrices for the unfoldings $\mathcal{A}_{(1)}, \ldots, \mathcal{A}_{(d)}$.

MATLAB Tensor Toolbox: **Computing the HOSVD**

```
function [S,U] = HOSVD(A)
% A is an n(1)-by-...-by-n(d) tensor.
% U is a length-d cell array with the
%   property that U{k} is the left singular
%   vector matrix of A's mode-k unfolding.
% S is an n(1)-by-...-by-n(d) tensor given by
%    A x1 U{1} x2 U{2} ... xd U{d}

S = A;
for k=1:length(A.size)
    C = tenmat(A,k);
    [U{k},Sigma,V] = svd(C.data);
    S = ttm(S,U{k}',k);
end
```

The Higher-Order SVD (HOSVD)

The HOSVD of a Matrix

If $d = 2$ then $\mathcal{A}$ is a matrix and the HOSVD is the SVD. Indeed, if

$$\begin{aligned} A &= A_{(1)} = U_1 \Sigma_1 V_1^T \\ A^T &= A_{(2)} = U_2 \Sigma_2 V_2^T \end{aligned}$$

then we can set $U = U_1 = V_2$ and $V = U_2 = V_1$. Note that

$$\mathcal{S} = (\mathcal{A} \times_1 U_1^T) \times_2 U_2^T = (U_1^T A) \times_2 U_2 = U_1^T A V_1 = \Sigma_1.$$

The Higher-Order SVD (HOSVD)

The Modal Unfoldings of the Core Tensor $\mathcal{S}$

If $\mathcal{A} = \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$ is the HOSVD of $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, then for $k = 1{:}d$

$$\mathcal{S}_{(k)} = U_k^T \mathcal{A}_{(k)} \left(U_d \otimes \cdots \otimes U_{k+1} \otimes U_{k-1} \otimes \cdots \otimes U_1 \right)$$

The Core Tensor $\mathcal{S}$ Encodes Singular Value Information

Since $U_k^T \mathcal{A}_{(k)} V_k = \Sigma_k$ is the SVD of $\mathcal{A}_{(k)}$, we have

$$\mathcal{S}_{(k)} = \Sigma_k V_k^T \left(U_d \otimes \cdots \otimes U_{k+1} \otimes U_{k-1} \otimes \cdots \otimes U_1 \right).$$

It follows that the rows of $S_{(k)}$ are mutually orthogonal and that the singular values of $\mathcal{A}_{(k)}$ are the 2-norms of these rows.

The Higher-Order SVD (HOSVD)

The HOSVD as a Multilinear Sum

If $\mathcal{A} = \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$ is the HOSVD of $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, then

$$\mathcal{A}(\mathbf{i}) = \sum_{\mathbf{j}=\mathbf{1}}^{\mathbf{n}} \mathcal{S}(\mathbf{j}) U_1(i_1, j_1) \cdots U_d(i_d, j_d)$$

The HOSVD as a Matrix-Vector Product

If $\mathcal{A} = \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$ is the HOSVD of $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, then

$$\text{vec}(\mathcal{A}) = (U_d \otimes \cdots \otimes U_1) \cdot \text{vec}(\mathcal{S})$$

Note that $U_d \otimes \cdots \otimes U_1$ is orthogonal.

Problem 4.6. Suppose $\mathcal{A} = \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$ is the HOSVD of $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and that $r \le \min\{n_1, \ldots, n_d\}$. Through MATLAB experimentation, what can you say about the approximation

$$\mathcal{A} \approx \mathcal{S}_r \times_1 U_1(:, 1{:}r) \times_2 U_2(:, 1{:}r) \cdots \times_d U_d(:, 1{:}r)$$

where $\mathcal{S}_r = \mathcal{S}(1{:}r, 1{:}r, \ldots, 1{:}r)$?

Problem 4.7. Formulate an HOQRP factorization for a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ that is based on the QR-with-column-pivoting factorizations

$$\mathcal{A}_{(k)} P_k = Q_k R_k$$

for $k = 1{:}d$. Does the core tensor have any special properties?

The Nearest Kronecker Product Problem

The Problem

Given the block matrix

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1,c_2} \\ \vdots & \ddots & \vdots \\ A_{r_2,1} & \cdots & A_{r_2,c_2} \end{bmatrix} \qquad A_{i_2,j_2} \in \mathbb{R}^{r_1 \times c_1}$$

find $B \in \mathbb{R}^{r_2 \times c_2}$ and $C \in \mathbb{R}^{r_1 \times c_1}$ so that

$$\phi_A(B, C) \;=\; \| \, A - B \otimes C \, \|_F$$

is minimized.

Looks like the nearest rank-1 matrix problem which has an SVD solution.

The Nearest Kronecker Product Problem

The Tensor Connection

The block matrix

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1,c_2} \\ \vdots & \ddots & \vdots \\ A_{r_2,1} & \cdots & A_{r_2,c_2} \end{bmatrix} \qquad A_{i_2,j_2} \in \mathbb{R}^{r_1 \times c_1}$$

corresponds to an order-4 tensor

$$A_{i_2,j_2}(i_1,j_1) \quad \leftrightarrow \quad \mathcal{A}(i_1,j_1,i_2,j_2)$$

and so does $M = (M_{ij}) = B \otimes C$:

$$M_{i_2,j_2}(i_1,j_1) \quad \leftrightarrow \quad \mathcal{B}(i_1,j_1)\mathcal{C}(i_2,j_2)$$

The Nearest Kronecker Product Problem

The Objective Function in Tensor Terms

$$
\begin{aligned}
\phi_A(B, C) &= \| A - B \otimes C \|_F \\
&= \sqrt{\sum_{i_1=1}^{r_1} \sum_{j_1=1}^{c_1} \sum_{i_2=1}^{r_2} \sum_{j_2=1}^{c_2} \mathcal{A}(i_1, j_1, i_2, j_2) - \mathcal{B}(i_2, j_2)\mathcal{C}(i_1, j_1)}
\end{aligned}
$$

We are trying to approximate an order-4 tensor with a pair of order-2 tensors.

Analogous to approximating a matrix (an order-2 tensor) with a rank-1 matrix (a pair of order-1 tensors.)

The Nearest Kronecker Product Problem

Reshaping the Objective Function (3-by-2 case)

$$\left\| \left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ \hline a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{array}\right] - \left[\begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{array}\right] \otimes \left[\begin{array}{cc} c_{11} & c_{12} \\ c_{21} & c_{22} \end{array}\right] \right\|_F$$

$$=$$

$$\left\| \left[\begin{array}{cccc} a_{11} & a_{21} & a_{12} & a_{22} \\ \hline a_{31} & a_{41} & a_{32} & a_{42} \\ \hline a_{51} & a_{61} & a_{52} & a_{62} \\ \hline a_{13} & a_{23} & a_{14} & a_{24} \\ \hline a_{33} & a_{43} & a_{34} & a_{44} \\ \hline a_{53} & a_{63} & a_{54} & a_{64} \end{array}\right] - \left[\begin{array}{c} b_{11} \\ b_{21} \\ b_{31} \\ b_{12} \\ b_{22} \\ b_{32} \end{array}\right] \left[\begin{array}{cccc} c_{11} & c_{21} & c_{12} & c_{22} \end{array}\right] \right\|_F$$

The Nearest Kronecker Product Problem

Minimizing the Objective Function (3-by-2 case)

It is a nearest rank-1 problem,

$$\phi_A(B,C) = \left\| \begin{bmatrix} a_{11} & a_{21} & a_{12} & a_{22} \\ \hline a_{31} & a_{41} & a_{32} & a_{42} \\ \hline a_{51} & a_{61} & a_{52} & a_{62} \\ \hline a_{13} & a_{23} & a_{14} & a_{24} \\ \hline a_{33} & a_{43} & a_{34} & a_{44} \\ \hline a_{53} & a_{63} & a_{54} & a_{64} \end{bmatrix} - \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \\ b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & c_{12} & c_{22} \end{bmatrix} \right\|_F$$

$$= \| \tilde{A} - \text{vec}(B)\text{vec}(C)^T \|_F$$

with SVD solution:

$$\tilde{A} = U\Sigma V^T$$

$$\text{vec}(B) = \sqrt{\sigma_1}\, U(:,1)$$

$$\text{vec}(C) = \sqrt{\sigma_1}\, V(:,1)$$

The Nearest Kronecker Product Problem

The "Tilde Matrix"

$$A = \left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ \hline a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{array}\right] = \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{array}\right]$$

implies

$$\tilde{A} = \left[\begin{array}{cccc} a_{11} & a_{21} & a_{12} & a_{22} \\ \hline a_{31} & a_{41} & a_{32} & a_{42} \\ \hline a_{51} & a_{61} & a_{52} & a_{62} \\ \hline a_{13} & a_{23} & a_{14} & a_{24} \\ \hline a_{33} & a_{43} & a_{34} & a_{44} \\ \hline a_{53} & a_{63} & a_{54} & a_{64} \end{array}\right] = \left[\begin{array}{c} \text{vec}(A_{11})^T \\ \text{vec}(A_{21})^T \\ \text{vec}(A_{31})^T \\ \text{vec}(A_{12})^T \\ \text{vec}(A_{22})^T \\ \text{vec}(A_{32})^T \end{array}\right].$$

The Nearest Kronecker Product Problem

Solution via Lanczos SVD

Since only the big "singular value triple" (σ_1, u_1, v_1) associated with $\tilde{A}$'s SVD is required, we can use Lanczos SVD. Note that if A is sparse, then $\tilde{A}$ is sparse.

Problem 4.8. Write a MATLAB function

```
[B,C] = NearestKP(A,r1,c1,r2,c2)
```

that solves the nearest Kronecker product problem where $A \in \mathbb{R}^{m\times n}$ with $m = r_1 r_2$, $n = c_1 c_2$ and A is regarded as an r_2-by-c_2 block matrix with r_1-by-c_1 blocks. Make effective use of MATLAB's sparse SVD solver `svds`.

Problem 4.9. Write a MATLAB function

```
[B,C] = NearestTensor(A)
```

where $\mathcal{A}$ is an order-4 tensor, $\mathcal{B}$ is an order-2 tensor, and $\mathcal{C}$ is an order-2 tensor such that

$$\phi_{\mathcal{A}}(\mathcal{B},\mathcal{C}) \;=\; \sqrt{\sum_{\mathbf{i}} |\mathcal{A}(\mathbf{i}) - \mathcal{B}(i_3,i_4)\mathcal{C}(i_1,i_2)|^2}$$

is minimized. Use `tenmat` to set up $\tilde{A}$ and use `svd`.

The Kronecker Product SVD (KPSVD)

Theorem

If

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1,c_2} \\ \vdots & \ddots & \vdots \\ A_{r_2,1} & \cdots & A_{r_2,c_2} \end{bmatrix} \qquad A_{i_2,j_2} \in \mathbb{R}^{r_1 \times c_1}$$

then there exist $U_1, \ldots, U_{r_{KP}} \in \mathbb{R}^{r_2 \times c_2}$, $V_1, \ldots, V_{r_{KP}} \in \mathbb{R}^{r_1 \times c_1}$, and scalars $\sigma_1 \geq \cdots \geq \sigma_{r_{KP}} > 0$ such that

$$A \;=\; \sum_{k=1}^{r_{KP}} \sigma_k U_k \otimes V_k .$$

The sets $\{\mathrm{vec}(U_k)\}$ and $\{\mathrm{vec}(V_k)\}$ are orthonormal and r_{KP} is the Kronecker rank of A with respect to the chosen blocking.

The Kronecker Product SVD (KPSVD)

Constructive Proof

Compute the SVD of $\tilde{A}$:

$$\tilde{A} \;=\; U\Sigma V^T \;=\; \sum_{k=1}^{r_{KP}} \sigma_k u_k v_k^T$$

and define the U_k and V_k by

$$\begin{aligned} \mathrm{vec}(U_k) &= u_k \\ \mathrm{vec}(V_k) &= v_k \end{aligned}$$

for $k = 1{:}r_{KP}$.

$$U_k = \mathtt{reshape}(u_k, r_2, c_2),\; V_k = \mathtt{reshape}(v_k, r_1, c_1)$$

Problem 4.10. Suppose

$$H = \begin{bmatrix} F_{11} \otimes G_{11} & F_{12} \otimes G_{12} \\ F_{21} \otimes G_{21} & F_{22} \otimes G_{22} \end{bmatrix}$$

where $F_{ij} \in \mathbb{R}^{m \times m}$ and $G_{ij} \in \mathbb{R}^{m \times m}$. Regarded as a $2m$-by-$2m$ block matrix, what is the Kronecker rank of H.

Problem 4.11. Suppose

$$H = \begin{bmatrix} F_1 \\ \vdots \\ F_p \end{bmatrix} \begin{bmatrix} G_1 & \cdots & G_q \end{bmatrix}$$

where the F_i and G_i are all m-by-m. Regarded as a p-by-q block matrix, what is the Kronecker rank of H?

The Kronecker Product SVD (KPSVD)

Nearest rank-r

If $r \leq r_{KP}$, then

$$A_r = \sum_{k=1}^{r} \sigma_k U_k \otimes V_k$$

is the nearest matrix to A (in the Frobenius norm) that has Kronecker rank r.

The Tensor Outer Product Operation

Applied to Order-1 Tensors...

$$\mathcal{A} = b \circ c \qquad b \in \mathbb{R}^{n_1},\ c \in \mathbb{R}^{n_2}$$

means

$$\mathcal{A}(i_1, i_2) \;=\; b(i_1)c(i_2)$$

Note that

$$\text{tenmat}(\mathcal{A}, [1], [2]) \;=\; bc^T$$

The Tensor Outer Product Operation

Applied to Order-2 Tensors...

$$\mathcal{A} = B \circ C \qquad B \in \mathbb{R}^{n_1 \times r_1},\ C \in \mathbb{R}^{n_2 \times r_2}$$

means

$$\mathcal{A}(i_1, i_2, i_3, i_4) \ = \ B(i_1, i_2) C(i_3, i_4)$$

Note that

$$\text{tenmat}(A, [\,3\ 1\,], [\,4\ 2\,]) \ = \ B \otimes C$$

The Tensor Outer Product Operation

More General...

$$\mathcal{A} = B \circ C \circ D \qquad B \in \mathbb{R}^{n_1 \times r_1},\ C \in \mathbb{R}^{n_2 \times r_2},\ D \in \mathbb{R}^{n_3 \times r_3}$$

means

$$\mathcal{A}(i_1, i_2, i_3, i_4) \; = \; B(i_1, i_2)C(i_3, i_4)D(i_5, i_6)$$

Note that

$$\text{tenmat}(A, [\,5\;3\;1\,], [\,6\;4\;2\,]) \; = \; B \otimes C \otimes D$$

The Tensor Outer Product Operation

The HOSVD as a Sum of Rank-1 Tensors

If $\mathcal{A} = \mathcal{S} \times_1 U_1 \times_2 U_2 \cdots \times_d U_d$ is the HOSVD of $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, then

$$\mathcal{A}(\mathbf{i}) = \sum_{\mathbf{j}=\mathbf{1}}^{\mathbf{n}} \mathcal{S}(\mathbf{j}) U_1(i_1, j_1) \cdots U_d(i_d, j_d)$$

implies

$$\mathcal{A} = \sum_{\mathbf{j}=\mathbf{1}}^{\mathbf{n}} \mathcal{S}(\mathbf{j}) \cdot U_1(:, j_1) \circ \cdots \circ U_d(:, j_d)$$

The Tensor Outer Product Operation

The KPSVD...

If

$$\texttt{tenmat}(\mathrm{A},[\,3\;1\,],[\,4\;2\,]) \;=\; \sum_i \sigma_i \cdot B_i \otimes C_i$$

then

$$\mathcal{A} \;=\; \sum_i \sigma_i \cdot \mathcal{B}_i \circ \mathcal{C}_i$$

Summary of Lecture 4.

Key Words

- The Mode-k Matrix Product is a contraction between a tensor and a matrix that produces another tensor.
- The Higher Order SVD of a tensor $\mathcal{A}$ assembles the SVDs of $\mathcal{A}$'s modal unfoldings.
- The Nearest Kronecker product problem involves computing the first singular triple of a permuted version of the given matrix.
- The Kronecker Product SVD characterizes a block matrix as a sum of Kronecker products. By applying it to an unfolding of a tensor $\mathcal{A}$, an outer product expansion for $\mathcal{A}$ is obtained.
- The outer product operation "$\circ$" is a way of combining an order-d_1 tensor and an order-d_2 tensor to obtain an order-$(d_1 + d_2)$ tensor.