Structured Matrix Computations from Structured Tensors

Lecture 2. Tensor Iterations, Symmetries, and Rank

Charles F. Van Loan

Cornell University

CIME-EMS Summer School
June 22-26, 2015
Cetraro, Italy

The Setting

- We looked at the nearest Rank-1 tensor problem in Lecture 1.
- For matrices, the nearest Rank-1 problem is an SVD problem. Is the nearest Rank-1 problem related to some tensor SVD?
- In Lectures 3 and 4 we will look at 4 different tensor SVD ideas. Structured SVDs, Schur Decompositions, and QR factorizations will be part of the scene.
- We set the stage for this in Lecture 2 by developing various power methods. It will be an occasion to refine what we know about tensor rank, tensor symmetry, and tensor unfoldings.

Rayleigh Quotient Ideas

Rayleigh Quotient Ideas: The Matrix Case

The Variational Approach to Singular Values and Vectors

The singular values and singular vectors of a general matrix $A \in \mathbb{R}^{m \times n}$ are stationary values and vectors for the multilinear form

$$\psi_A(x, y) \;=\; x^T A y \;=\; \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j$$

subject to the constraints $\| \, x \, \| = \| \, y \, \| = 1$

Let us understand this connection and then push the same idea for tensors.

Rayleigh Quotient Ideas: The Matrix Case

Gradient Calculations

We seek unit vectors x and y that zero the gradient of

$$\tilde{\psi}_A(x, y) = \psi(x, y) - \frac{\lambda}{2}(x^T x - 1) - \frac{\mu}{2}(y^T y - 1)$$

Since

$$\psi_A(x, y) \;=\; \sum_{i=1}^{m} x_i \left(\sum_{j=1}^{n} a_{ij} y_j \right) \;=\; \sum_{j=1}^{n} y_j \left(\sum_{i=1}^{m} a_{ij} x_i \right)$$

it follows that we want

$$\nabla \tilde{\psi}_A(x, y) \;=\; \begin{bmatrix} Ay & - & \lambda x \\ A^T x & - & \mu y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus, $\lambda = \mu = x^T A y = \psi_A(x, y)$.

Rayleigh Quotient Ideas: The Matrix Case

Gradient Calculations

$$\nabla \tilde{\psi}_A(x, y) = \begin{bmatrix} Ay - (x^T Ay)x \\ A^T x - (x^T Ay)y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Singular Values and Vectors

If $U^T AV = \Sigma = \text{diag}(\sigma_i)$ is the SVD of A and $U = [\, u_1 \mid \cdots \mid u_m \,]$ and $V = [\, v_1 \mid \cdots \mid v_n \,]$, then from $AV = U\Sigma$ and $A^T U = V\Sigma^T$ we have

$$\begin{aligned} Av_i &= \sigma_i u_i \\ A^T u_i &= \sigma_i v_i \end{aligned}$$

Since $\sigma_i = u_i^T Av_i$, it follows that

$$\begin{aligned} Av_i - (u_i^T Av_i)u_i &= 0 \\ A^T u_i - (u_i^T Av_i)v_i &= 0 \end{aligned}$$

Rayleigh Quotient Ideas: The Matrix Case

Gradient Calculations

$$\nabla\tilde{\psi}_A(x,y) = \begin{bmatrix} Ay - (x^TAy)x \\ A^Tx - (x^TAy)y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Singular Values and Vector

If $U^TAV = \Sigma = \text{diag}(\sigma_i)$ is the SVD of A and $U = [\, u_1 \mid \cdots \mid u_m \,]$ and $V = [\, v_1 \mid \cdots \mid v_n \,]$, then from $AV = U\Sigma$ and $A^TU = V\Sigma^T$ we have

$$\begin{aligned} Av_i &= \sigma_i u_i \\ A^Tu_i &= \sigma_i v_i \end{aligned}$$

Since $\sigma_i = u_i^TAv_i$, it follows that

$$\begin{aligned} Av_i - (u_i^TAv_i)u_i &= 0 \\ A^Tu_i - (u_i^TAv_i)v_i &= 0 \end{aligned}$$

Rayleigh Quotient Ideas: The Tensor Case

The Singular Values and Vectors of a Tensor: A Definition

If $\mathcal{A} \in \mathbb{R}^{m \times n \times p}$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $z \in \mathbb{R}^p$, then the singular values and vectors of $\mathcal{A}$ are the stationary values and vectors of

$$\psi_{\mathcal{A}}(x, y, z) \;=\; \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk}\, x_i y_j z_k$$

subject to the constraints $\| \, x \, \| = \| \, y \, \| = \| \, z \, \| = 1$

Order-3 tensors are plenty good enough to show the main ideas.
Generalizations to order d tensors are generally pretty obvious.

Rayleigh Quotient Ideas: The Tensor Case

Take a look at $\psi_{\mathcal{A}}(x, y, z)$

Some handy rearrangements o fψ_A:

$$\psi_{\mathcal{A}}(x,y,z) = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} a_{ijk}\, x_i y_j z_k = \sum_{i=1}^{m} x_i \left(\sum_{j=1}^{n}\sum_{k=1}^{p} a_{ijk}\, y_j z_k\right)$$

$$= \sum_{j=1}^{n} y_j \left(\sum_{i=1}^{m}\sum_{k=1}^{p} a_{ijk}\, x_i z_k\right)$$

$$= \sum_{k=1}^{p} z_k \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ijk}\, x_i y_j\right)$$

Before we go after the gradient, let's frame the double summations in linear algebra terms...

Rayleigh Quotient Ideas: The Tensor Case

Suppose $m = 4, n = 3$, and $p = 2$

$$\psi_{\mathcal{A}}(x,y,z) = \sum_{i=1}^{m} x_i \left(\sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk}\, y_j z_k \right)$$

$$= \sum_{i=1}^{m} x_i \begin{bmatrix} a_{i11} & a_{i21} & a_{i31} & a_{i12} & a_{i22} & a_{i32} \end{bmatrix} \begin{bmatrix} z_1 y_1 \\ z_1 y_2 \\ z_1 y_3 \\ z_2 y_1 \\ z_2 y_2 \\ z_2 y_3 \end{bmatrix}$$

$$= x^T \begin{bmatrix} a_{111} & a_{121} & a_{131} & a_{112} & a_{122} & a_{132} \\ a_{211} & a_{221} & a_{231} & a_{212} & a_{222} & a_{232} \\ a_{311} & a_{321} & a_{331} & a_{312} & a_{322} & a_{332} \\ a_{411} & a_{421} & a_{431} & a_{412} & a_{422} & a_{432} \end{bmatrix} z \otimes y$$

Rayleigh Quotient Ideas: The Tensor Case

Suppose $m = 4, n = 3$, and $p = 2$

$$\psi_{\mathcal{A}}(x,y,z) = \sum_{i=1}^{m} x_i \left(\sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk}\, y_j z_k \right)$$

$$= y^T \begin{bmatrix} a_{111} & a_{121} & a_{131} & a_{112} & a_{122} & a_{132} \\ a_{211} & a_{221} & a_{231} & a_{212} & a_{222} & a_{232} \\ a_{311} & a_{321} & a_{331} & a_{312} & a_{322} & a_{332} \\ a_{411} & a_{421} & a_{431} & a_{412} & a_{422} & a_{432} \end{bmatrix} z \otimes x$$

The matrix you see here is the mode-1 unfolding of $\mathcal{A} \in \mathbb{R}^{4\times 3\times 2}$

Rayleigh Quotient Ideas: The Tensor Case

$\psi_{\mathcal{A}}(x, y, z)$ in terms of the Mode-1 Unfolding

$$\psi_{\mathcal{A}}(x, y, z) \;=\; \sum_{i=1}^{m} x_i \left(\sum_{j=1}^{n} \sum_{k=1}^{p} a_{ijk}\, y_j z_k \right) \;=\; x^T \mathcal{A}_{(1)} \cdot z \otimes y$$

$$\mathcal{A}_{(1)} \;=\; \left[\begin{array}{cccccc} a_{111} & a_{121} & a_{131} & a_{112} & a_{122} & a_{132} \\ a_{211} & a_{221} & a_{231} & a_{212} & a_{222} & a_{232} \\ a_{311} & a_{321} & a_{331} & a_{312} & a_{322} & a_{332} \\ a_{411} & a_{421} & a_{431} & a_{412} & a_{422} & a_{432} \end{array} \right]$$

(1,1) (2,1) (3,1) (1,2) (2,2) (3,2)

Rayleigh Quotient Ideas: The Tensor Case

$\psi_{\mathcal{A}}(x, y, z)$ in terms of the Mode-2 Unfolding

$$\psi_{\mathcal{A}}(x, y, z) \;=\; \sum_{j=1}^{n} y_j \left(\sum_{i=1}^{m} \sum_{k=1}^{p} a_{ijk}\, x_i z_k \right) \;=\; y^T \mathcal{A}_{(2)} z \otimes x$$

$$\mathcal{A}_{(2)} \;=\; \begin{bmatrix} a_{111} & a_{211} & a_{311} & a_{411} & a_{112} & a_{212} & a_{312} & a_{412} \\ a_{121} & a_{221} & a_{321} & a_{421} & a_{122} & a_{222} & a_{322} & a_{422} \\ a_{131} & a_{231} & a_{331} & a_{431} & a_{132} & a_{232} & a_{332} & a_{432} \end{bmatrix}$$

$$\quad (1,1) \quad (2,1) \quad (3,1) \quad (4,1) \quad (1,2) \quad (2,2) \quad (3,2) \quad (4,2)$$

Rayleigh Quotient Ideas: The Tensor Case

$\psi_{\mathcal{A}}(x, y, z)$ in terms of the Mode-3 Unfolding

$$\psi_{\mathcal{A}}(x,y,z) \;=\; \sum_{k=1}^{p} z_k \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ijk}\, x_i y_j \right) \;=\; z^T \mathcal{A}_{(3)} \cdot y \otimes x$$

$$\mathcal{A}_{(3)} = \left[\begin{array}{cccccccccccc} a_{111} & a_{211} & a_{311} & a_{411} & a_{121} & a_{221} & a_{321} & a_{421} & a_{131} & a_{231} & a_{331} & a_{431} \\ a_{112} & a_{212} & a_{312} & a_{412} & a_{122} & a_{222} & a_{322} & a_{422} & a_{132} & a_{232} & a_{332} & a_{432} \end{array} \right]$$

(1,1) (2,1) (3,1) (4,1) (1,2) (2,2) (3,2) (4,2) (1,3) (2,3) (3,3) (4,3)

Rayleigh Quotient Ideas: The Tensor Case

Important Skill: Framing a Tensor Computation in Matrix Terms

$$
\begin{aligned}
\psi_{\mathcal{A}}(x,y,z) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} a_{ijk}x_i y_j z_k \\
&= x^T \mathcal{A}_{(1)}(z \otimes y) \\
&= y^T \mathcal{A}_{(2)}(z \otimes x) \\
&= z^T \mathcal{A}_{(3)}(y \otimes x)
\end{aligned}
$$

With these characterizations we can readily compute the stationary vectors and values of this function subject to the constraint that $\parallel x \parallel = \parallel y \parallel = \parallel z \parallel = 1$.

The Gradient Computations

Set the gradient of

$$\tilde{\psi}_{\mathcal{A}}(x,y,z) \;=\; \psi_{\mathcal{A}}(x,y,z) - \frac{\lambda}{2}(x^Tx-1) - \frac{\mu}{2}(y^Ty-1) - \frac{\tau}{2}(z^Tz-1)$$

to zero. Conclude that $\lambda = \mu = \tau = \psi(x,y,z)$. If unit vectors x, y, and z to satisfy

$$\nabla\tilde{\psi}_{\mathcal{A}} \;=\; \left[\begin{array}{ccc} \mathcal{A}_{(1)}(z\otimes y) & - & \psi_{\mathcal{A}}(x,y,z)x \\ \mathcal{A}_{(2)}(z\otimes x) & - & \psi_{\mathcal{A}}(x,y,z)y \\ \mathcal{A}_{(3)}(y\otimes x) & - & \psi_{\mathcal{A}}(x,y,z)z \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right]$$

then $\sigma = \psi(x,y,z)$ is a singular value of $\mathcal{A}$ and x, y, and z the associated singular vectors.

How can we solve this (highly structured) system of nonlinear equations?

A Higher-Order Power Method

The Power Method for Matrices

The Gradient Equations

$$\begin{aligned} Ay &= \sigma \cdot x \\ A^T x &= \sigma \cdot y \end{aligned}$$

where $\sigma = \psi_A(x, y) = x^T Ay$.

Iterate...

y a given unit vector

Repeat Until Happy

$$\begin{aligned} \tilde{x} &= Ay,\ x \leftarrow \tilde{x}/\| \tilde{x} \| \\ \tilde{y} &= A^T x,\ y \leftarrow \tilde{y}/\| \tilde{y} \| \\ \sigma &= \psi_A(x, y) \end{aligned}$$

Same as power method applied to $A^T A$.In the limit, σuv^T converges to the closest rank-1 to A.

The Higher-Order Power Method for Tensors

The Gradient Equations

$$A_{(1)} \cdot (z \otimes y) \;=\; \sigma \cdot x$$

$$A_{(2)} \cdot (z \otimes x) \;=\; \sigma \cdot y$$

$$A_{(3)} \cdot (y \otimes x) \;=\; \sigma \cdot z$$

where $\sigma = \psi_A(x, y, z) = x^T A_{(1)} \cdot (z \otimes y) = y^T A_{(2)} \cdot (z \otimes x) = z^T A_{(3)} \cdot (y \otimes x)$

Iterate...

y and z given unit vectors
Repeat Until Happy

$$\begin{aligned}
\tilde{x} &= \mathcal{A}_{(1)}(z \otimes y), \quad x = \tilde{x}/\| \tilde{x} \| \\
\tilde{y} &= \mathcal{A}_{(2)}(z \otimes x), \quad y = \tilde{y}/\| \tilde{y} \| \\
\tilde{z} &= \mathcal{A}_{(3)}(y \otimes x), \quad z = \tilde{z}/\| \tilde{z} \| \\
\sigma &= \psi_{\mathcal{A}}(x, y, z)
\end{aligned}$$

The Higher-Order Power Method for Tensors

A Tempting Repetition

Step 1. Compute closest $\sigma_1 \cdot u_1 \circ v_1 \circ w_1$ to $\mathcal{A}$.

Step 2. Compute closest $\sigma_2 \cdot u_2 \circ v_2 \circ w_2$ to

$$\mathcal{A} - \sigma_1 u_1 \circ v_1 \circ w_1 .$$

Step 3. Compute closest $\sigma_3 \cdot u_3 \circ v_3 \circ w_3$ to

$$\mathcal{A} - \sigma_1 \cdot u_1 \circ v_1 \circ w_1 - \sigma_2 \cdot u_2 \circ v_2 \circ w_2 .$$

Step r. Compute closest $\sigma_r \cdot u_r \circ v_r \circ w_r$ to

$$\mathcal{A} - \sum_{k=1}^{r-1} \sigma_k \cdot u_k \circ v_k \circ w_k .$$

This does not render a "best" approximation to $\mathcal{A}$ as it does for matrices. So maybe we better look more closely at sums of rank-1 tensors and rank.

The Idea of Tensor Rank

$\mathcal{A} \in \mathbb{R}^{2\times 2\times 2}$ as a minimal sum of rank-1 tensors.

Challenge: Find thinnest possible $X, Y, Z \in \mathbb{R}^{2\times r}$ so

$$\begin{bmatrix} a_{111} \\ a_{211} \\ a_{121} \\ a_{221} \\ a_{112} \\ a_{212} \\ a_{122} \\ a_{222} \end{bmatrix} = \sum_{k=1}^{r} z_k \otimes y_k \otimes x_k$$

where $X = [\, x_1 | \cdots | x_r \,]$, $Y = [\, y_1 | \cdots | y_r \,]$, and $Z = [\, z_1 | \cdots | z_r \,]$ are column partitionings.

The minimizing r is the **tensor rank.**

$\mathcal{A} \in \mathbb{R}^{2\times 2\times 2}$ as a minimal sum of rank-1 tensors.

A Surprising Fact

If $\begin{bmatrix} a_{111} \\ a_{211} \\ a_{121} \\ a_{221} \\ a_{112} \\ a_{212} \\ a_{122} \\ a_{222} \end{bmatrix} = \texttt{randn(8,1)}$, then $\begin{cases} \text{rank} = 2 & \text{with prob } 79\% \\ \text{rank} = 3 & \text{with prob } 21\% \end{cases}$

This is very different from the matrix case where $A = \texttt{randn(n,n)}$ implies $\text{rank}(A) = n$ with probability 100%.

A strong hint that tensor rank is decidely more complicated than matrix rank. What are the "full rank" 2-by-2-by-2 tensors?

The 79/21 Property

Connection to a Generalized Eigenvalue Problem

If the a_{ijk} are `randn`, then

$$\det\left(\begin{bmatrix} a_{111} & a_{121} \\ a_{211} & a_{221} \end{bmatrix} - \lambda \begin{bmatrix} a_{112} & a_{122} \\ a_{212} & a_{222} \end{bmatrix}\right) = 0$$

has real distinct roots with probability 79% and complex conjugate roots with probability 21%. The sum-of-rank-ones expansion for $\mathcal{A}$ involves the generalized eigenvectors of this problem.

Yet another example of turning a tensor problem into a matrix problem.

Symmetry

What is a Symmetric Tensor?

The Order-3 Definition: $\mathcal{C} \in \mathbb{R}^{n\times n\times n}$

$$c_{ijk} = \begin{cases} c_{ikj} \\ c_{jik} \\ c_{jki} \\ c_{kij} \\ c_{kji} \end{cases}$$

It just means that permuting the indices does not change the value.

Symmetric Tensors

$\mathcal{C} \in \mathbb{R}^{3\times 3\times 3}$

There are 10 values to specify...

(1,1,1)
(1,1,2),(1,2,1),(2,1,1)
(1,1,3),(1,3,1),(3,1,1)
(2,2,2)
(2,2,1),(2,1,2),(2,2,1)
(2,2,3),(2,3,1),(2,2,3)
(3,3,3)
(3,3,1),(3,1,3),(3,3,1)
(3,3,2),(3,2,3),(3,3,2)
(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)

Symmetric Tensors

The Modal Unfoldings are all the Same

$$C_{(1)} = \begin{bmatrix} c_{111} & c_{121} & c_{131} & c_{112} & c_{122} & c_{132} & c_{113} & c_{123} & c_{133} \\ c_{211} & c_{221} & \color{red}{c_{231}} & c_{212} & c_{222} & c_{232} & c_{213} & c_{223} & c_{233} \\ c_{311} & c_{321} & c_{331} & c_{312} & c_{322} & c_{332} & 3_{113} & c_{323} & c_{333} \end{bmatrix}$$

$$C_{(2)} = \begin{bmatrix} c_{111} & c_{211} & c_{311} & c_{112} & c_{212} & c_{312} & c_{113} & c_{213} & c_{313} \\ c_{121} & c_{221} & \color{red}{c_{321}} & c_{122} & c_{222} & c_{322} & c_{123} & c_{223} & c_{323} \\ c_{131} & c_{231} & c_{331} & c_{132} & c_{232} & c_{332} & 3_{113} & c_{233} & c_{333} \end{bmatrix}$$

$$C_{(3)} = \begin{bmatrix} c_{111} & c_{211} & c_{311} & c_{121} & c_{221} & c_{321} & c_{131} & c_{231} & c_{331} \\ c_{112} & c_{212} & \color{red}{c_{312}} & c_{122} & c_{222} & c_{322} & c_{132} & c_{232} & c_{332} \\ c_{113} & c_{213} & c_{313} & c_{123} & c_{223} & c_{323} & 3_{133} & c_{233} & c_{333} \end{bmatrix}$$

Symmetric Tensors

Rank-1 Symmetric Tensors

If $x \in \mathbb{R}^n$, then

$$\mathcal{C} \;=\; x \circ x \circ x$$

is a symmetric rank-1 tensor. This is obvious since

$$c_{ijk} = x_i x_j x_k.$$

Note that

$$\text{vec}(x \circ x \circ \circ x) \;=\; x \otimes x \otimes x$$

Symmetric Rank

An order-3 symmetric tensor $\mathcal{C}$ has symmetric rank r if there exists $x_1, \ldots, x_r \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}^r$ such that

$$\mathcal{C} \;=\; \sum_{k=1}^{r} \sigma_k \cdot x_k \circ x_k \circ x_k$$

and no shorter sum of symmetric rank-1 tensors exists. Symmetric rank is denoted by $\text{rank}_S(\mathcal{C})$.

Note, there may be a shorter sum so

$$\mathcal{C} \;=\; \sum_{k=1}^{\tilde{r}} \tilde{\sigma}_k \cdot \tilde{x}_k \circ \tilde{y}_k \circ \tilde{z}_k$$

Symmetric Tensors: Interesting Aside about Rank

Fact

If $\mathcal{C} \in \mathrm{C}^{n\times\cdots\times n}$ is an order-d symmetric tensor, then with probability 1

$$\mathrm{rank}_S(\mathcal{C}) \;=\; \begin{cases} f(d,m)+1 & \text{if } (d,n) = (3,5),(4,3),(4,4), \text{ or } (4,5) \\ \\ f(d,n) & \text{otherwise} \end{cases}$$

where

$$f(d,n) \;=\; \text{ceil}\left(\frac{\begin{pmatrix} n+d-1 \\ d \end{pmatrix}}{n} \right)$$

Symmetric Tensor Rank is "more tractable" than General Tensor Rank.

Rayleigh Quotient Ideas

Symmetric Matrix Eigenvalues

If C is a symmeytric matrix, then the stationary values of

$$\phi_C(x) \;=\; x^T C x$$

subject to the constraint that $\| x \|_2 = 1$ are the eigenvalues of C. The associated stationary vectors are eigenvectors.

Symmetric Tensor Eigenvalues

If $\mathcal{C}$ is a symmetric tensor, then the stationary values of

$$\phi_{\mathcal{C}}(x) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} c_{ijk} x_i x_j x_k = x^T C_{(1)}(x \otimes x)$$

subject to the constraint that $\| x \|_2 = 1$ are the eigenvalues of $\mathcal{C}$. The associated stationary vectors are eigenvectors.

Rayleigh Quotients: Symmetric Tensor Case

Symmetric Higher-Order Power Method

Initialize unit vector x.

Repeat Until Happy

$$\tilde{x} \;=\; \mathcal{C}_{(1)}(x \otimes x)$$

$$x = \tilde{x}/\parallel \tilde{x} \parallel$$

Sample Convergence Result

If the order of $\mathcal{C}$ is even and M is a square unfolding, then the iteration converges if M is positive definite.

The SVD - SymEig Connection

The "Sym" of a Matrix

$$\text{sym}(A) \;=\; \left[\begin{array}{cc} 0 & A \\ A^T & 0 \end{array}\right] \in \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)}$$

The SVD of A Relates to the EVD of sym(A)

If $A = U \cdot \text{diag}(\sigma_i) \cdot V^T$ is the SVD of $A \in \mathbb{R}^{n_1 \times n_2}$, then for $k = 1\text{:rank}(A)$

$$\left[\begin{array}{cc} 0 & A \\ A^T & 0 \end{array}\right] \left[\begin{array}{c} u_k \\ \pm v_k \end{array}\right] \;=\; \pm\sigma_k \left[\begin{array}{c} u_k \\ \pm v_k \end{array}\right]$$

where $u_k = U(:,k)$ and $v_k = V(:,k)$.

The above SVD-related power method was essentially traditional power method applied to finding the largest eigenvector of sym(A).

The SVD - SymEig Connection

The "Sym" of a Matrix

$$\text{sym}(A) \; = \; \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \in \mathbb{R}^{(n_1+n_2)\times(n_1+n_2)}$$

The SVD of A Relates to the EVD of sym(A)

If $A = U \cdot \text{diag}(\sigma_i) \cdot V^T$ is the SVD of $A \in \mathbb{R}^{n_1 \times n_2}$, then for $k = 1{:}\text{rank}(A)$

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u_k \\ \pm v_k \end{bmatrix} = \pm\sigma_k \begin{bmatrix} u_k \\ \pm v_k \end{bmatrix}$$

where $u_k = U(:,k)$ and $v_k = V(:,k)$.

Let us look at the analog of this for tensors. Need transposition

Tensor Transposition: The Order-3 Case

Six possibilities...

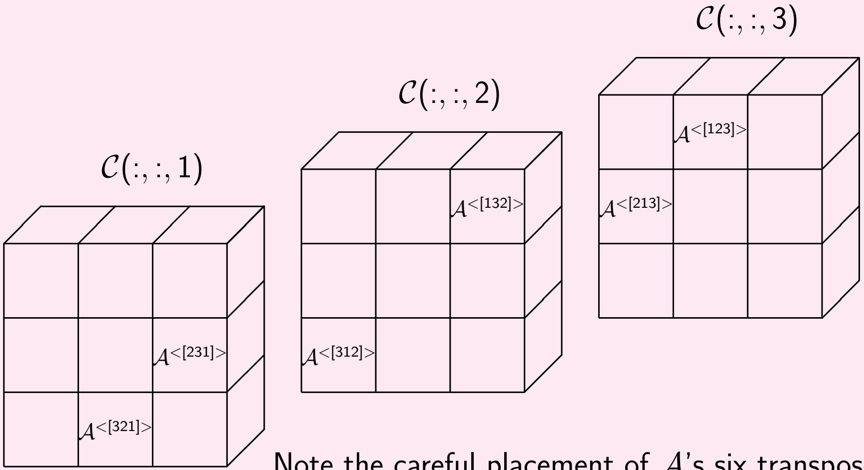
If $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, then there are $6 = 3!$ possible transpositions identified by the notation $\mathcal{C}^{<[i\ j\ k]>}$ where $[i\ j\ k]$ is a permutation of $[1\ 2\ 3]$:

$$\mathcal{B} = \left\{ \begin{array}{c} \mathcal{C}^{<[1\ 2\ 3]>} \\ \mathcal{C}^{<[1\ 3\ 2]>} \\ \mathcal{C}^{<[2\ 1\ 3]>} \\ \mathcal{C}^{<[2\ 3\ 1]>} \\ \mathcal{C}^{<[3\ 1\ 2]>} \\ \mathcal{C}^{<[3\ 2\ 1]>} \end{array} \right\} \quad \Longrightarrow \quad \left\{ \begin{array}{c} b_{ijk} \\ b_{ikj} \\ b_{jik} \\ b_{jki} \\ b_{kij} \\ b_{kji} \end{array} \right\} = c_{ijk}$$

for $i = 1{:}n_1,\ j = 1{:}n_2,\ k = 1{:}n_3$.

Symmetric Embedding of a Tensor $\mathcal{C} = \text{sym}(\mathcal{A})$

An Order-3 Example...



Note the careful placement of $\mathcal{A}$'s six transposes

Connecting $\mathcal{A}$ and sym($\mathcal{A}$)

Algorithms

Interesting connections between power methods with $\mathcal{A}$ and power methods with sym($\mathcal{A}$)

Analysis

If

$$\left\{ \sigma, \begin{bmatrix} u \\ v \\ z \end{bmatrix} \right\}$$

is a stationary pair for sym($\mathcal{A}$) then so are

$$\left\{ \sigma, \begin{bmatrix} u \\ -v \\ -z \end{bmatrix} \right\}, \quad \left\{ -\sigma, \begin{bmatrix} u \\ -v \\ z \end{bmatrix} \right\}, \quad \left\{ -\sigma, \begin{bmatrix} u \\ v \\ -z \end{bmatrix} \right\}$$

Symmetric Tensors: Interesting Aside about Rank

Interesting Possible Connection

Easy:

$$d! \, \text{rank}(\mathcal{A}) \;\leq\; \text{rank}_S(\text{sym}(\mathcal{A}))$$

Equality is hard or perhaps not true. But if it could be established, then we would have new insight into the tensor rank problem.

Problem E2. Suppose $A \in \mathbb{R}^{m \times n^2}$ with $m > n^2$. Develop an alternating least squares solution framework for minimizing $\| A(x \otimes y) - b \|_2$ where $b \in \mathbb{R}^{n^2}$ and $x, y \in \mathbb{R}^n$.

Problem A2. Same notation as E2. What is the gradient of

$$\phi(x, y) = \frac{1}{2} \| A(x \otimes y) - b \|_2^2$$

Closing Remarks

Bridging the Gap via Tensor Unfoldings

A Common Framework for Tensor Computations...

1. Turn tensor $\mathcal{A}$ into a matrix A.

2. Through matrix computations, discover things about A.

3. Draw conclusions about tensor $\mathcal{A}$ based on what is learned about matrix A.