Computational Multilinear Algebra

Charles F. Van Loan

Department of Computer Science

Cornell University

Outline

- Review the Kronecker Product (KP) Operation
- Explain Why the KP is Increasingly Important in Scientific Computing
- Illustrate Some Nicely Solved KP Problems
- Discuss KP Methods for Decomposing 3-dimensional Tensors (Arrays)

Review of the Kronecker Product

Definition

$$\left[\begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \end{array}\right] \otimes \left[\begin{array}{ccc} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{array}\right] = \left[\begin{array}{ccc|ccc} b_{11}c_{11} & b_{11}c_{12} & b_{11}c_{13} & b_{12}c_{11} & b_{12}c_{12} & b_{12}c_{13} \\ b_{11}c_{21} & b_{11}c_{22} & b_{11}c_{23} & b_{12}c_{21} & b_{12}c_{22} & b_{12}c_{23} \\ b_{11}c_{31} & b_{11}c_{32} & b_{11}c_{33} & b_{12}c_{31} & b_{12}c_{32} & b_{12}c_{33} \\ \hline b_{21}c_{11} & b_{21}c_{12} & b_{21}c_{13} & b_{22}c_{11} & b_{22}c_{12} & b_{22}c_{13} \\ b_{21}c_{21} & b_{21}c_{22} & b_{21}c_{23} & b_{22}c_{21} & b_{22}c_{22} & b_{22}c_{23} \\ b_{21}c_{31} & b_{21}c_{32} & b_{21}c_{33} & b_{22}c_{31} & b_{22}c_{32} & b_{22}c_{33} \end{array}\right]$$

$$\left.\begin{array}{ll} B & m_1\text{-by-}n_1 \\ & \\ C & m_2\text{-by-}n_2 \end{array}\right\} \text{ then } B \otimes C \text{ is a} \left\{\begin{array}{l} (m_1 m_2)\text{-by-}(n_1 n_2) \text{ matrix of scalars} \\ \\ m_1\text{-by-}n_1 \text{ block matrix with } m_2\text{-by-}n_2 \text{ blocks} \end{array}\right.$$

Properties

Quite predictable:

$$\begin{aligned}
(B \otimes C)^T &= B^T \otimes C^T \\
(B \otimes C)^{-1} &= B^{-1} \otimes C^{-1} \\
(B \otimes C)(D \otimes F) &= BD \otimes CF \\
B \otimes (C \otimes D) &= (B \otimes C) \otimes D
\end{aligned}$$

Think twice:

$$B \otimes C \neq C \otimes B$$

$$B \otimes C \;=\; (\text{Perfect Shuffle})(C \otimes B)(\text{Perfect Shuffle})^T$$

The Perfect Shuffle $S_{p,q}$

$$S_{3,4}\left[\begin{array}{c}0\\1\\2\\3\\\hline 4\\5\\6\\7\\\hline 8\\9\\10\\11\end{array}\right] = \left[\begin{array}{c}0\\4\\8\\\hline 1\\5\\9\\\hline 2\\6\\10\\\hline 3\\7\\11\end{array}\right] \equiv \begin{bmatrix}0&4&8\\1&5&9\\2&6&10\\3&7&11\end{bmatrix} \xrightarrow{S_{3,4}} \begin{bmatrix}0&1&2&3\\4&5&6&7\\8&9&10&11\end{bmatrix}$$

Takes the length-pq “card deck” x, splits it into p piles of length-q each, and then takes one card from each pile in turn until the deck is reassembled.

Example: $(2 \times 2) \otimes (2 \times 3)$

$$B \otimes C = \left[\begin{array}{ccc|ccc} b_{11}c_{11} & b_{11}c_{12} & b_{11}c_{13} & b_{12}c_{11} & b_{12}c_{12} & b_{12}c_{13} \\ b_{11}c_{21} & b_{11}c_{22} & b_{11}c_{23} & b_{12}c_{21} & b_{12}c_{22} & b_{12}c_{23} \\ \hline b_{21}c_{11} & b_{21}c_{12} & b_{21}c_{13} & b_{22}c_{11} & b_{22}c_{12} & b_{22}c_{13} \\ b_{21}c_{21} & b_{21}c_{22} & b_{21}c_{23} & b_{22}c_{21} & b_{22}c_{22} & b_{22}c_{23} \end{array}\right]$$

Reorder rows [1 3 2 4] and reorder columns [1 4 2 5 3 6]:

$$S_{2,2} A S_{2,3}^T \; = \; = \; C \otimes B \; = \left[\begin{array}{cc|cc|cc} b_{11}c_{11} & b_{12}c_{11} & b_{11}c_{12} & b_{12}c_{12} & b_{11}c_{13} & b_{12}c_{13} \\ b_{21}c_{11} & b_{22}c_{11} & b_{21}c_{12} & b_{22}c_{12} & b_{21}c_{13} & b_{22}c_{13} \\ \hline b_{11}c_{21} & b_{12}c_{21} & b_{11}c_{22} & b_{12}c_{22} & b_{11}c_{23} & b_{12}c_{23} \\ b_{21}c_{21} & b_{22}c_{21} & b_{21}c_{22} & b_{22}c_{22} & b_{21}c_{23} & b_{22}c_{23} \end{array}\right]$$

Inheriting Structure

$$\text{If } B \text{ and } C \text{ are} \left\{ \begin{array}{l} \text{nonsingular} \\ \text{lower(upper) triangular} \\ \text{banded} \\ \text{symmetric} \\ \text{positive definite} \\ \text{stochastic} \\ \text{Toeplitz} \\ \text{permutations} \\ \text{orthogonal} \end{array} \right\} \quad \text{then } B \otimes C \text{ is } \left\{ \begin{array}{l} \text{nonsingular} \\ \text{lower(upper)triangular} \\ \text{block banded} \\ \text{symmetric} \\ \text{positive definite} \\ \text{stochastic} \\ \text{block Toeplitz} \\ \text{a permutation} \\ \text{orthogonal} \end{array} \right\}$$

Factoring $B \otimes C$

If you have the { LU, Cholesky, QR} factorization of B and C,

then you have the { LU, Cholesky, QR} factorization of $B \otimes C$...

$$B \otimes C \;=\; (P_B^T L_B U_B) \otimes (P_C^T L_C U_C) \;=\; (P_B \otimes P_C)^T (L_B \otimes L_C)(U_B \otimes U_C)$$

$$B \otimes C \;=\; (G_B G_B^T) \otimes (G_C G_C^T) \;=\; (G_B \otimes G_C)(G_B \otimes G_C)^T$$

$$B \otimes C \;=\; (Q_B R_B) \otimes (Q_C R_C) \;=\; (Q_B \otimes Q_C)(R_B \otimes R_C)$$

Factoring $B \otimes C$

If you have the $\{$ Eigenvalue, Singular Value$\}$ decomposition of B and C,

then you *sort of* have the $\{$ Eigenvalue, Singular Value$\}$ decomposition of $B \otimes C$...

$$B \otimes C \; = \; (Q_B \Lambda_B Q_B^T) \otimes (Q_C \Lambda_C Q_C^T) \; = \; (Q_B \otimes Q_C)(\Lambda_B \otimes \Lambda_C)(Q_B \otimes Q_C)^T$$

$$B \otimes C \; = \; (U_B \Sigma_B V_B^T) \otimes (U_C \Sigma_C V_C^T) \; = \; (U_B \otimes U_C)(\Sigma_B \otimes \Sigma_C)(V_B \otimes V_C)^T$$

“Sort of”

$$\begin{bmatrix} \times & 0 \\ 0 & \times \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} \times & 0 & 0 \\ 0 & \times & 0 \\ 0 & 0 & \times \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \left[\begin{array}{ccc|ccc} \times & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \times & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Unhappy Factoring of $B \otimes C$

If you have the CS decomposition of B and C,

then you *do not* have the CS decomposition of $B \otimes C$.

The CS decomposition says that the blocks of an orthogonal matrix have related SVDs:

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}^T$$

$$C = \mathrm{diag}(\cos(\theta_1), \ldots, \cos(\theta_m)) \qquad S = \mathrm{diag}(\sin(\theta_1), \ldots, \sin(\theta_m))$$

where U_1, U_2, V_1, and V_2 are orthogonal.

The vec Operation

- Example...

$$A = \begin{bmatrix} 1 & 10 \\ 2 & 20 \\ 3 & 30 \end{bmatrix} \qquad \Rightarrow \qquad \mathrm{vec}(A) = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 10 \\ 20 \\ 30 \end{bmatrix}$$

- In General...

$$X \in \mathbb{R}^{m\times n} \quad \Rightarrow \quad \mathrm{vec}(X) \ = \begin{bmatrix} X(:,1) \\ X(:,2) \\ \vdots \\ X(:,n) \end{bmatrix}$$

- Big Fact...

$$Y = CXB^T \quad \Rightarrow \quad \mathrm{vec}(Y) = (B \otimes C)\mathrm{vec}(X)$$

Turning Matrix Equations into Vector Equations

Sylvester:

$$FX + XG^T = C \qquad \equiv \qquad (I_n \otimes F \; + \; G \otimes I_m)\,\mathrm{vec}(X) \; = \; \mathrm{vec}(C)$$

Generalized Sylvester:

$$FXH^T + KXG^T = C \qquad \equiv \qquad (H \otimes F \; + \; G \otimes K)\,\mathrm{vec}(X) \; = \; \mathrm{vec}(C)$$

Lyapunov:

$$FX + XF^T = C \qquad \equiv \qquad (I_n \otimes F \; + \; F \otimes I_n)\,\mathrm{vec}(X) \; = \; \mathrm{vec}(C)$$

“Fast” Factoring means “Fast” Solving

If $B, C \in {\rm I\!R}^{m \times m}$, then the m^2-by-m^2 system

$$(B \otimes C)x = f \qquad \equiv \qquad CXB^T = F \qquad x = \text{vec}(X),\;\; f = \text{vec}(F)$$

can be solved in $O(m^3)$ flops:

$$\begin{aligned} CY &= F \\ XB^T &= Y \end{aligned}$$

via factorizations of B and C.

More Dramatic

If $B_1, \ldots, B_d \in \mathbb{R}^{m \times m}$, then the m^d-by-m^d system

$$(B_1 \otimes B_2 \otimes \cdots \otimes B_d)x = f$$

can be solved in $O(m^{d+1})$ flops (instead of $O((m^d)^3)$ flops.)

The Growing Importance
# of the
Kronecker Product

Tensoring Low Dimensional Ideas

Quadrature in one dimension:

$$\begin{aligned} \int_a^b f(x)\,dx &\approx \sum_{i=1}^{n} w_i\, f(x_i) \\ &= w^T f(x) \end{aligned}$$

Quadrature in three dimensions:

$$\begin{aligned} \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3} f(x,y,z)\,dx\,dy\,dz &\approx \sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sum_{k=1}^{n_z} w_i^{(x)} w_j^{(y)} w_k^{(z)} f(x_i, y_j, z_k) \\ &= (w^{(x)} \otimes w^{(y)} \otimes w^{(z)})^T f(x \otimes y \otimes z) \end{aligned}$$

Notes on Tensoring Vectors

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1y_1 \\ x_1y_2 \\ x_1y_3 \\ x_2y_1 \\ x_2y_2 \\ x_2y_3 \end{bmatrix}$$

$$\mathrm{vec}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^{\square} y_1 \;\; y_2 \;\; y_3 \, \Big] \right) = \mathrm{vec}\left(\begin{bmatrix} x_1y_1 & x_1y_2 & x_1y_3 \\ x_2y_1 & x_2y_2 & x_2y_3 \end{bmatrix} \right) = y \otimes x$$

Generalizations of the Familiar

- Higher-order statistics. Instead of looking at the expected value of xx^T, look at the expected value of the *cumulant*

  $$e = x \otimes x \otimes \ \cdots \ \otimes x.$$

  Note that $\text{vec}(xx^T) = x \otimes x$.

- Multidimensional Arrays. Instead of looking for patterns in 2-dimensional arrays via (for example) the SVD, look for patterns in d-dimensional arrays using generalized notions of the SVD. (More later.)

Sparse Factorizations

Kronecker products are proving to be a very effective way to look at fast linear transforms such as the FFT:

$$y = F_n x = \begin{bmatrix} \omega_8^0 & \omega_8^0 & \omega_8^0 & \omega_8^0 & \omega_8^0 & \omega_8^0 & \omega_8^0 & \omega_8^0 \\ \omega_8^0 & \omega_8^1 & \omega_8^2 & \omega_8^3 & \omega_8^4 & \omega_8^5 & \omega_8^6 & \omega_8^7 \\ \omega_8^0 & \omega_8^2 & \omega_8^4 & \omega_8^6 & \omega_8^8 & \omega_8^{10} & \omega_8^{12} & \omega_8^{14} \\ \omega_8^0 & \omega_8^3 & \omega_8^6 & \omega_8^9 & \omega_8^{12} & \omega_8^{15} & \omega_8^{18} & \omega_8^{21} \\ \omega_8^0 & \omega_8^4 & \omega_8^8 & \omega_8^{12} & \omega_8^{16} & \omega_8^{20} & \omega_8^{24} & \omega_8^{28} \\ \omega_8^0 & \omega_8^5 & \omega_8^{10} & \omega_8^{15} & \omega_8^{20} & \omega_8^{25} & \omega_8^{30} & \omega_8^{35} \\ \omega_8^0 & \omega_8^6 & \omega_8^{12} & \omega_8^{18} & \omega_8^{24} & \omega_8^{30} & \omega_8^{36} & \omega_8^{42} \\ \omega_8^0 & \omega_8^7 & \omega_8^{14} & \omega_8^{21} & \omega_8^{28} & \omega_8^{35} & \omega_8^{42} & \omega_8^{49} \end{bmatrix} x$$

$\omega_8 = \cos(2\pi/8) + i \cdot \sin(2\pi/8)$.

Recursive Block Structure

$$F_8 S_{2,4} = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \omega_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \omega_8^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & \omega_8^3 \\ \hline 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -\omega_8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -\omega_8^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -\omega_8^3 \end{array}\right] (I_2 \otimes F_4)$$

$F_{n/2}$ “shows up” when you permute the columns of F_n so that the odd-indexed columns come first.

A Sparse Factorization of the DFT Matrix

$$n = 2^t$$

$$F_n = A_t \cdots A_1 P_n$$

$$P_n = S_{2,n/2}(I_2 \otimes S_{2,n/4}) \cdots (I_{n/4} \otimes S_{2,2})$$

$$A_q = I_r \otimes \begin{bmatrix} I_{L/2} & \Omega_{L/2} \\ I_{L/2} & -\Omega_{L/2} \end{bmatrix} \qquad L = 2^q,\ r = n/L$$

$$\Omega_{L/2} = \mathrm{diag}(1, \omega_L, \ldots, \omega_L^{L/2-1}) \qquad \omega_L = \exp(-2\pi i/L)$$

Different FFTs/Different Factorizations of F_n

The Cooley-Tukey FFT is based on $y = F_n x = A_t \cdots A_1 P_n x$

```
$x \leftarrow P_n x$
for $k = 1{:}t$
    $x \leftarrow A_q x$
end
$y \leftarrow x$
```

The Gentleman-Sande FFT is based on $y = F_n x = F_n^T x = P_n^T A_1^T \cdots A_t^T x$

```
for $k = t{:} - 1{:}1$
    $x \leftarrow A_q^T x$
end
$y \leftarrow P_n^T x$
```

Matrix Transpose

$B = A^T$, corresponds to $\mathrm{vec}(B) = S_{n,m} \cdot \mathrm{vec}(A)$.

$$\begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \\ a_{13} \\ a_{23} \end{bmatrix}$$

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

Multiple Pass Transpose

To compute $B = A^T$ where $A \in \mathrm{I\!R}^{m \times n}$ factor $S_{n,m} = \Gamma_t \cdots \Gamma_1$ and then execute

```
$a \; = \mathrm{vec}(A)$
for $k = 1{:}t$
    $a \leftarrow \Gamma_k a$
end
Define $B \in \mathrm{I\!R}^{n \times m}$ by $\mathrm{vec}(B) = a$.
```

Different transpose algorithms correspond to different factorizations of $S_{n,m}$.

An Example

If $m = pn$, then $S_{n,m} = \Gamma_2\Gamma_1 \;=\; (I_p \otimes S_{n,n})(S_{n,p} \otimes I_n)$

The first pass $b^{(1)} = \Gamma_1 \text{vec}(A)$ corresponds to a block transposition:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} \qquad \rightarrow \qquad B^{(1)} = \left[\, A_1 \,\middle|\, A_2 \,\middle|\, A_3 \,\middle|\, A_4 \,\right].$$

The second pass $b^{(2)} = \Gamma_2 b^{(1)}$ carries out the transposition of the blocks.

$$B^{(1)} = \left[\, A_1 \,\middle|\, A_2 \,\middle|\, A_3 \,\middle|\, A_4 \,\right] \qquad \rightarrow \qquad B^{(2)} = \left[\, A_1^T \,\middle|\, A_2^T \,\middle|\, A_3^T \,\middle|\, A_4^T \,\right].$$

Note that $B^{(2)} = A^T$.

Semidefinite Programming

Some sample problems...

$$\left(X \otimes X + A^T D A\right) u = f.$$

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ Z \otimes I & 0 & X \otimes I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} r_d \\ r_p \\ r_c \end{bmatrix}.$$

See Alizadeh, Haeberly, and Overton (1998).

Symmetric Kronecker Products

For symmetric $X \in \mathrm{I\!R}^{n \times n}$ and arbitrary $B, C \in \mathrm{I\!R}^{n \times n}$ this operation is defined by

$$(B \circledast C)\mathrm{svec}(X) = \mathrm{svec}\left(\frac{1}{2}\left(CXB^T + BXC^T\right)\right)$$

where the "svec" operation is a normalized stacking of X's subdiagonal columns, e.g.,

$$X \; = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \Rightarrow \; \mathrm{svec}(X) = \Big[\, x_{11} \;\; \sqrt{2}x_{21} \;\; \sqrt{2}x_{31} \;\; x_{22} \;\; \sqrt{2}x_{32} \;\; x_{33} \,\Big]^T .$$

svec stacks the subdiagonal portion of X's columns.

Some Nicely Solved Kronecker Product Problems

The Nearest Kronecker Product Problem

Given $A \in \mathbb{R}^{m \times n}$ with $m = m_1 m_2$ and $n = n_1 n_2$.

Find $B \in \mathbb{R}^{m_1 \times n_1}$ and $C \in \mathbb{R}^{m_2 \times n_2}$ so

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

A bilinear least squares problem. But we can do better...

The NKP is a Nearest Rank-1 problem

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

$$= \left\| \begin{bmatrix} a_{11} & a_{21} & a_{12} & a_{22} \\ a_{31} & a_{41} & a_{32} & a_{42} \\ a_{51} & a_{61} & a_{52} & a_{62} \\ a_{13} & a_{23} & a_{14} & a_{24} \\ a_{33} & a_{43} & a_{34} & a_{44} \\ a_{53} & a_{63} & a_{54} & a_{64} \end{bmatrix} - \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \\ b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} {}^{\square} c_{11} \;\; c_{21} \;\; c_{12} \;\; c_{22} \,\big] \right\|_F$$

Solution Procedure

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

$$\tilde{A} \ = \ \begin{bmatrix} \text{vec}(A_{11})^T \\ \text{vec}(A_{21})^T \\ \text{vec}(A_{31})^T \\ \text{vec}(A_{12})^T \\ \text{vec}(A_{22})^T \\ \text{vec}(A_{32})^T \end{bmatrix} .$$

An SVD solution...

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

$$\text{vec}(B_{opt}) \ = \ \sqrt{\sigma_1}\, U(:,1) \qquad \text{vec}(C_{opt}) \ = \ \sqrt{\sigma_1}\, V(:,1).$$

The Kronecker Product SVD

Ordinary SVD:

$$U^T A V = \Sigma \qquad \Rightarrow \qquad A \;=\; \sum_{k=1}^{r=rank(A)} \sigma_k u_k v_k^T$$

where u_k is the k-th column of U and v_k is the k-th column of V.

KP SVD:

$$U^T \tilde{A} V = \Sigma \qquad \Rightarrow \qquad \sum_{k=1}^{rank(\tilde{A})} \sigma_k\, U_k \otimes V_k$$

where vec(U_k) is the k-th column of U and vec(V_k) is the k-th column of V.

Some Modified Least Squares Problems

How do we solve

$$(1) \qquad \min \| \, W((B \otimes C)x - d) \, \| \qquad \text{(weighted least quares)}$$

$$(2) \qquad U^T \, [B \otimes C \mid d] \, V = \Sigma \qquad \text{(total least squares)}$$

given that these problems

$$(1') \qquad \min \| \, (B \otimes C)x - d \, \|$$

$$(2') \qquad U^T (B \otimes C) V = \Sigma$$

are *easy*

Weighted Least Squares Problems

$$\min \parallel W^{-1/2}((B \otimes C)x - b) \parallel_2 \qquad \equiv \qquad \left[\begin{array}{cc} W & B \otimes C \\ B^T \otimes C^T & 0 \end{array} \right] \left[\begin{array}{c} r \\ x \end{array} \right] = \left[\begin{array}{c} b \\ 0 \end{array} \right]$$

Compute the QR factorizations

$$B = Q_B \left[\begin{array}{c} R_B \\ 0 \end{array} \right] \qquad C = Q_C \left[\begin{array}{c} R_C \\ 0 \end{array} \right]$$

The augmented system transforms to

$$\begin{bmatrix} E_{11} & E_{12} & R_B \otimes R_C \\ E_{21} & E_{22} & 0 \\ R_B^T \otimes R_C^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{r}_1 \\ \tilde{r}_2 \\ x \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ 0 \end{bmatrix}$$

where

$$\begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}$$

is a simple permutation of $(Q_B \otimes Q_C)^T W (Q_B \otimes Q_C)$. Solve the E_{22} system via congugate gradients exploiting structure.

Total Least Squares

LS: $$\min_{b+r \in \operatorname{ran}(A)} \| r \|_2^2 \qquad\qquad Ax_{\mathrm{LS}} \; = \; b + r_{\mathrm{opt}}$$

TLS: $$\min_{b+r \in \operatorname{ran}(A+E)} \| E \|_F^2 + \| r \|_2^2 \qquad\qquad (A + E_{\mathrm{opt}})x_{\mathrm{TLS}} \; = \; b + r_{\mathrm{opt}}$$

“Errors in Variables”

Total Least Squares Solution

To solve

$$\min_{b + r \,\in\, \mathrm{ran}(A+E)} \quad \| \, E \, \|_F^2 + \| \, r \, \|_2^2 \qquad\qquad (A + E_{\mathrm{opt}})x_{\mathrm{TLS}} \;=\; b + r_{\mathrm{opt}}$$

compute the SVD of $[A \mid b] \in \mathrm{I\!R}^{m \times n+1}$:

$$U^T \;\square\; A \big| b \,\big] V = \Sigma$$

and set

$$x_{\mathrm{TLS}} = -V(1{:}n, n+1)/V(n+1, n+1)$$

TLS When A is a Kronecker Product

We need the last column of V in $U^TFV = \Sigma$ where

$$F = \left[\; B \otimes C \,\middle|\, b \;\right]$$

First compute the SVDs of B and C:

$$U_B^T B V_B = \Sigma_B \qquad U_C^T C V_C = \Sigma_C$$

If

$$\tilde{U} = U_B \otimes U_C \qquad\qquad \tilde{V} = \left[\begin{array}{c|c} V_B \otimes V_C & 0 \\ \hline 0 & 1 \end{array}\right]$$

then

$$\tilde{U}^T F \tilde{V} \;=\; \left[\; \Sigma_B \otimes \Sigma_C \;\middle|\; g \;\right] \;\equiv\; \tilde{F} \qquad \text{where} \quad g = \tilde{U}^T b$$

We need the smallest right singular vector of $\tilde{F}$.

Frequency Response

Suppose we wish to evaluate the following function for many different values of μ:

$$\phi(\mu) \;=\; c^T(A - \mu I)^{-1}d$$

It pays to triangularize A via the real Schur decomposition:

$$Q^TAQ \;=\; T \;=\; \left[\begin{array}{ccccc} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{array}\right]$$

$$\phi(\mu) \;=\; c^TQ(T - \mu I)^{-1}Q^Td \qquad (O(n^2) \text{ per evaluation})$$

Solving $(F \otimes G - \lambda I)\, x = b$.

$$\begin{bmatrix} f_{11}G - \lambda I_m & f_{12}G & f_{13}G & f_{14}G \\ 0 & f_{22}G - \lambda I_m & f_{23}G & f_{24}G \\ 0 & 0 & f_{33}G - \lambda I_m & f_{34}G \\ 0 & 0 & 0 & f_{44}G - \lambda I_m \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

Solve the triangular system

$$(f_{44}G - \lambda I_m)\, y_4 = c_4$$

for y_4. Substituting this into the above system reduces it to

$$\begin{bmatrix} f_{11}G - \lambda I_m & f_{12}G & f_{13}G \\ 0 & f_{22}G - \lambda I_m & f_{23}G \\ 0 & 0 & f_{33}G - \lambda I_m \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \tilde{c}_1 \\ \tilde{c}_2 \\ \tilde{c}_3 \end{bmatrix}$$

where $\tilde{c}_i = c_i - f_{i4}Gy_4$, $i = 1{:}3$.

Orthogonal Decompositions of 3-Dimensional Arrays

Think Matrix $\longleftrightarrow$ Think Vector

Suppose $A = UBV^T$ where A and B are m-by-n and

$$U = [\, u_1,\ u_2,\ \cdots,\ u_m \,] \qquad\qquad V = [\, v_1,\ v_2,\ \cdots,\ v_n \,] \,.$$

Then we can write A as a linear combination of rank-1 matrices

$$A \;=\; \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} \; (\; u_i v_j^T \;)$$

and vec(A) as a linear combination of rank-1 tensors

$$\text{vec}(A) = \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} \; (\; v_j \otimes u_i \;)$$

The SVD in Tensor Notation

If A is m-by-n and has SVD

$$A = U\Sigma V^T$$

with $r = \text{rank}(A)$, then

$$\text{vec}(A) \;=\; \sum_{i=1}^{m}\sum_{j=1}^{n} \sigma_{ij} \; (\, v_j \otimes u_i \,) \;=\; \sum_{k=1}^{r} \sigma_{kk} \; (\, v_k \otimes u_k \,)$$

The SVD represents an mn-by-1 vector as a minimal combination of rank-1 tensors that are mutually orthogonal.

Representing $A = A(1{:}m, 1{:}n, 1{:}p)$

Turn A into a vector

$$\mathrm{vec}(A) \;=\; \begin{bmatrix} \mathrm{vec}(A(:,:,1)) \\ \vdots \\ \mathrm{vec}(A(:,:,p)) \end{bmatrix}$$

and represent it as a sum of rank-1 tensors

$$\mathrm{vec}(A) = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} \sigma_{ijk} \left(w_k \otimes v_j \otimes u_i \right)$$

where each of the following is orthogonal:

$$U = [\, u_1, \ldots, u_m] \qquad V = [\, v_1, \ldots, v_n] \qquad W = [\, w_1, \ldots, w_p]$$

Concentrating Mass

- It can be shown that

$$\mathrm{vec}(A) = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{p} \sigma_{ijk} \, (\, w_k \otimes v_j \otimes u_i \,) \; = (W \otimes V \otimes U)\sigma$$

  where $\sigma = \mathrm{vec}(\Sigma)$, $\Sigma = \Sigma(1{:}m, 1{:}n, 1{:}p)$.

- This representation is "good" if most of σ's mass is concentrated in as few components as possible.

- But you can forget this...

$$\mathrm{vec}(A) \; = \; \sum_{k=1}^{r} \sigma_{kkk} \; (w_k \otimes v_k \otimes u_k)$$

The Kronecker Product SVD (KPSVD)

If

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{bmatrix} \qquad\qquad A_{ij} \in {\rm I\!R}^{s\times t}$$

then

$$A = \sum_{i=1}^{r} \sigma_i \; B_i \otimes C_i \qquad r = \min\{pq, st\}$$

where $B_i \in {\rm I\!R}^{p\times q}$, $C_i \in {\rm I\!R}^{s\times t}$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$.

Moreover, vec(B_1),...,vec(B_r) are mutually orthogonal and so are vec(C_1),...,vec(C_r).

Algorithm ($m = n = p$ for clarity)

- Set up

$$\tilde{A} \; = \; [A_1, \ldots, A_n\,] \qquad A_i = A(:,:,i) \in {\rm I\!R}^{n\times n}$$

and note that $\mathrm{vec}(\tilde{A}) = \mathrm{vec}(A)$.

- Compute the KPSVD

$$\tilde{A} = \sum_{k=1}^{n} \tau_k \left(\, w_k^T \otimes B_k \,\right)$$

where $W = [\, w_1, \ldots, w_n\,]$ is orthogonal and $B_1, \ldots, B_n$ are each n-by-n. it follows that

$$\mathrm{vec}(A) = \mathrm{vec}(\tilde{A}) = \sum_{k=1}^{n} \tau_k \left(\, w_k \otimes \mathrm{vec}(B_k) \,\right)$$

- Compute the SVD

$$[\, B_1, \ldots, B_n \,] \;=\; \sum_{i=1}^{n} \mu_i u_i \begin{bmatrix} g_{1i} \\ \vdots \\ g_{ni} \end{bmatrix}^T$$

where $U = [u_1, \ldots, u_n\,]$ and

$$G \;=\; \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} \qquad g_{ik} \in \mathbb{R}^n$$

have orthogonal columns.

- $B_k \;=\; \sum_{i=1}^{n} \mu_i \left(u_i g_{ki}^T\right) \;\Rightarrow\; \text{vec}(B_k) = \sum_{i=1}^{n} \mu_i \left(g_{ki} \otimes u_i\right)$

- $\text{vec}(A) \;=\; \sum_{k=1}^{n} \tau_k \,(\, w_k \otimes \text{vec}(B_k) \,) \;=\; \sum_{k=1}^{n} \sum_{i=1}^{n} \tau_k \mu_i \left(w_k \otimes g_{ki} \otimes u_i\right)$

- Set $\tilde{G} \ = \ [\, g_{11},\ g_{21},\ g_{31},\ g_{12},\ g_{22},\ g_{32},\ g_{13},\ g_{23},\ g_{33}\,]$ and compute its QR factorization

$$\tilde{G} = VR \qquad V = [\, v_1, \ldots, v_n \,] \ \text{(orthogonal)}$$

- Since $g_{ki} = \sum_{j=1}^{n} r_{j,k+(i-1)n} v_j$ it follows that

$$\begin{aligned}
\mathrm{vec}(A) &= \sum_{k=1}^{n} \sum_{i=1}^{n} (\tau_k \cdot \mu_i)\,(w_k \otimes g_{ki} \otimes u_i) \\
&= \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} (\, \tau_k \cdot \mu_i \cdot r_{j,k+(i-1)n} \,)\,(w_k \otimes v_j \otimes u_i) \\
&= \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sigma_{ijk}\,(w_k \otimes v_j \otimes u_i)
\end{aligned}$$

$$\boxed{\sigma_{ijk} = \tau_k \cdot \mu_i \cdot r_{j,k+(i-1)n}}$$

De Lathauwer, De Moor, Vandewalle (2000)

Compute the SVDs:

$$
\begin{aligned}
[\ A(1,:,:),\ \ \ldots\ \ , A(n,:,:)\ ] &= V D_1 G_1^T \\
{}^{\square}\ A(:,1,:)^T, \ldots, A(:,n,:)^T\ ] &= W D_2 G_2^T \\
[\ A(:,:,1),\ \ \ldots\ \ , A(:,:,n)\ ] &= U D_3 G_3^T
\end{aligned}
$$

Set $\Sigma = \texttt{\textbf{reshape}}(D_3^T(W \otimes V))$ and then

$$
\mathrm{vec}(A) \ = \ \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sigma_{ijk} \, (w_k \otimes v_j \otimes u_i)
$$

Research Goals

- To heighten the profile of the Kronecker product and develop an "infrastructure" of methods thereby making it easier for the numerical linear algebra community to spot Kronecker "opportunities".

- To understand better the concept of "rank" as it applies to multidimensional arrays and to develop effective algorithms for its calculation.