Computational Multilinear Algebra

Charles F. Van Loan

Supported in part by the NSF contract CCR-9901988.
Involves Working With Kronecker Products

\[
\begin{bmatrix}
  b_{11} & b_{12} \\
  b_{21} & b_{22}
\end{bmatrix}
\otimes
\begin{bmatrix}
  c_{11} & c_{12} & c_{13} \\
  c_{21} & c_{22} & c_{23} \\
  c_{31} & c_{32} & c_{33}
\end{bmatrix}
= 
\begin{bmatrix}
  b_{11}c_{11} & b_{11}c_{12} & b_{11}c_{13} & b_{12}c_{11} & b_{12}c_{12} & b_{12}c_{13} \\
  b_{11}c_{21} & b_{11}c_{22} & b_{11}c_{23} & b_{12}c_{21} & b_{12}c_{22} & b_{12}c_{23} \\
  b_{11}c_{31} & b_{11}c_{32} & b_{11}c_{33} & b_{12}c_{31} & b_{12}c_{32} & b_{12}c_{33} \\
  b_{21}c_{11} & b_{21}c_{12} & b_{21}c_{13} & b_{22}c_{11} & b_{22}c_{12} & b_{22}c_{13} \\
  b_{21}c_{21} & b_{21}c_{22} & b_{21}c_{23} & b_{22}c_{21} & b_{22}c_{22} & b_{22}c_{23} \\
  b_{21}c_{31} & b_{21}c_{32} & b_{21}c_{33} & b_{22}c_{31} & b_{22}c_{32} & b_{22}c_{33}
\end{bmatrix}
\]

“Replicated Block Structure”
Properties

Quite predictable:

\[(B \otimes C)^T = B^T \otimes C^T\]
\[(B \otimes C)^{-1} = B^{-1} \otimes C^{-1}\]
\[(B \otimes C)(D \otimes F) = BD \otimes CF\]
\[B \otimes (C \otimes D) = (B \otimes C) \otimes D\]

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 $S_{p,q}$

$S_{3,4} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 8 \\ 1 \\ 5 \\ 9 \\ 2 \\ 6 \\ 10 \\ 3 \\ 7 \\ 11 \end{bmatrix} \equiv \begin{bmatrix} 0 & 4 & 8 \\ 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \end{bmatrix} \rightarrow S_{3,4} \begin{bmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \end{bmatrix}$

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.
\[ C \otimes B = S_{m_1, m_2}(B \otimes C)S_{n_1, n_2}^T \quad B \in \mathbb{R}^{m_1 \times n_1}, \ C \in \mathbb{R}^{m_2 \times n_2} \]

E.g.,

\[
A = B \otimes C = \begin{bmatrix}
  b_{11}c_{11} & b_{11}c_{12} & b_{11}c_{13} & b_{12}c_{11} & b_{12}c_{12} & b_{12}c_{13} \\
  b_{11}c_{21} & b_{11}c_{22} & b_{11}c_{23} & b_{12}c_{21} & b_{12}c_{22} & b_{12}c_{23} \\
  b_{21}c_{11} & b_{21}c_{12} & b_{21}c_{13} & b_{22}c_{11} & b_{22}c_{12} & b_{22}c_{13} \\
  b_{21}c_{21} & b_{21}c_{22} & b_{21}c_{23} & b_{22}c_{21} & b_{22}c_{22} & b_{22}c_{23}
\end{bmatrix}
\]

\[
S_{2, 2}A \ S_{2, 3}^T = = C \otimes B = \begin{bmatrix}
  c_{11}b_{11} & c_{11}b_{12} & c_{12}b_{11} & c_{12}b_{12} & c_{13}b_{11} & c_{13}b_{12} \\
  c_{11}b_{21} & c_{11}b_{22} & c_{12}b_{21} & c_{12}b_{22} & c_{13}b_{21} & c_{13}b_{22} \\
  c_{21}b_{11} & c_{21}b_{12} & c_{22}b_{11} & c_{22}b_{12} & c_{23}b_{11} & c_{23}b_{12} \\
  c_{21}b_{21} & c_{21}b_{22} & c_{22}b_{21} & c_{22}b_{22} & c_{23}b_{21} & c_{23}b_{22}
\end{bmatrix}
\]
Inheriting Structure

If $B$ and $C$ are
\begin{align*}
\text{nonsingular} & \\
\text{lower(upper) triangular} & \\
\text{banded} & \\
\text{symmetric} & \\
\text{positive definite} & \\
\text{stochastic} & \\
\text{Toeplitz} & \\
\text{permutations} & \\
\text{orthogonal} & 
\end{align*}

then $B \otimes C$ is
\begin{align*}
\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{align*}
Factorizations of $B \otimes C$

Obvious...

\[ B \otimes C = (P_B^T L_B U_B) \otimes (P_C^T L_C U_C) = (P_B \otimes P_C)^T (L_B \otimes L_C)(U_B \otimes U_C) \]
\[ B \otimes C = (G_B G_B^T) \otimes (G_C G_C^T) = (G_B \otimes G_C) (G_B \otimes G_C)^T \]
\[ B \otimes C = (Q_B R_B) \otimes (Q_C R_C) = (Q_B \otimes Q_C) (R_B \otimes R_C) \]

Sort of...

\[ B \otimes C = (Q_B \Lambda_B Q_B^T) \otimes (Q_C \Lambda_C Q_C^T) = (Q_B \otimes Q_C) (\Lambda_B \otimes \Lambda_C)(Q_B \otimes Q_C)^T \]
\[ B \otimes C = (U_B \Sigma_B V_B^T) \otimes (U_C \Sigma_C V_C^T) = (U_B \otimes U_C)(\Sigma_B \otimes \Sigma_C)(V_B \otimes V_C)^T \]

Problematic...

\[ B \otimes C = (Q_B R_B P_B^T) \otimes (Q_C R_C P_C^T) = (Q_B \otimes Q_C)(R_B \otimes R_C)(P_B \otimes P_C)^T \]
“Sort of”

\[
\begin{bmatrix}
\times & 0 & 0 \\
0 & \times & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
\otimes
\begin{bmatrix}
\times & 0 & 0 \\
0 & \times & 0 \\
0 & 0 & \times \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
= 
\begin{bmatrix}
\times & 0 & 0 & 0 & 0 & 0 \\
0 & \times & 0 & 0 & 0 & 0 \\
0 & 0 & \times & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \times & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]
\[ P \begin{pmatrix} \begin{bmatrix} x & 0 \\ 0 & x \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{bmatrix} \end{pmatrix} = \begin{bmatrix} x & 0 & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 & 0 \\ 0 & 0 & x & 0 & 0 & 0 \\ 0 & 0 & 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 & 0 & x \end{bmatrix} \]
“Fast” Factoring

Suppose $B \in \mathbb{IR}^{m \times m}$ and $C \in \mathbb{IR}^{n \times n}$ are positive definite.

$A = B \otimes C$ is an $mn$-by-$mn$ symmetric positive definite matrix.

Ordinarily, a Cholesky factorization $GG^T$ this size would cost $O(m^3n^3)$ flops.

But $G_A = G_B \otimes G_C$ and $G_B$ and $G_C$ require $O(m^3 + n^3)$ flops.
“Fast” Solving

If $B, C \in \mathbb{IR}^{m \times m}$, then the $m^2$-by-$m^2$ system

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

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

$$CY = F$$

$$XB^T = Y$$

via factorizations of $B$ and $C$. 
The vec Operation

\[ X \in \mathbb{R}^{m \times n} \implies \text{vec}(X) = \begin{bmatrix} X(:, 1) \\ X(:, 2) \\ \vdots \\ X(:, n) \end{bmatrix} \]

\[ Y = CXB^T \implies \text{vec}(Y) = (B \otimes C)\text{vec}(X) \]
Matrix Equations

Sylvester:
\[ FX + XG^T = C \quad \equiv \quad (I_n \otimes F + G \otimes I_m) \text{vec}(X) = \text{vec}(C) \]

Lyapunov:
\[ FX + XF^T = C \quad \equiv \quad (I_n \otimes F + F \otimes I_n) \text{vec}(X) = \text{vec}(C) \]

Generalized Sylvester:
\[ F_1XG_1^T + F_2XG_2^T = C \quad \equiv \quad (G_1 \otimes F_1 + G_2 \otimes F_2) \text{vec}(X) = \text{vec}(C) \]

These can be solved in \(O(n^3)\) time via the (generalized) Schur decomposition. See Bartels and Stewart, Golub, Nash, and Van Loan, and J. Gardiner, M.R. Wette, A.J. Laub, J.J. Amato, and C.B. Moler.
Talk Outline

How do we solve

(1) \[ \min \| W((B \otimes C)x - d) \| \] (weighted least squares)

(2) \[ U^T [B \otimes C \mid d] V = \Sigma \] (total least squares)

(3) \[ (A_p \otimes \cdots \otimes A_1 - \lambda I)x = b \] (shifted linear systems)

given that these problems

(1') \[ \min \| (B \otimes C)x - d \| \]

(2') \[ U^T(B \otimes C)V = \Sigma \]

(3') \[ (A_p \otimes \cdots \otimes A_1)x = b \]

are easy
The Rise of the Kronecker Product

- Thriving Application Areas
- Sparse Factorizations for fast linear transforms
- Tensoring Low-Dimension Techniques
Thriving application areas such as Semidefinite Programming...

Some sample problems...

\[(X \otimes X + A^TDA)u = f.\]

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
Z \otimes I & 0 & X \otimes I
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta z
\end{bmatrix} =
\begin{bmatrix}
rd \\
rp \\
rc
\end{bmatrix}.
\]

See Alizadeh, Haeberly, and Overton (1998).
Symmetric Kronecker Products

For symmetric $X \in \mathbb{IR}^{n \times n}$ and arbitrary $B, C \in \mathbb{IR}^{n \times n}$ this operation is defined by

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

where the “svec” operation is a normalized stacking of $X$’s subdiagonal columns, e.g.,

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

$svec$ stacks the subdiagonal portion of $X$’s columns.
The Rise of the Kronecker Product Cont’d

Sparse Factorizations

Kronecker products are proving to be a very effective way to look at fast linear transforms.
A Sparse Factorization of the DFT Matrix

\[ n = 2^t \]

\[ F_n = A_t \cdots A_1 P_n \]

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

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

\[ \Omega_{L/2} = \text{diag}(1, \omega_L, \ldots, \omega_L^{L/2-1}) \quad \omega_L = \exp(-2\pi i/L) \]
Different FFTs Correspond to Different Factorizations of $F_n$

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

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

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

$$
\text{for } k = t : -1 : 1 \\
\quad x \leftarrow A_q^T x \\
\text{end} \\
y \leftarrow P_n^T x$$
Matrix Transpose

$A \in \mathbb{IR}^{m \times n}$ and $B = A^T$, then $\text{vec}(B) = S_{n,m} \cdot \text{vec}(A)$.

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

\[
\begin{bmatrix}
  a_{11} & a_{21} \\
  a_{12} & a_{22} \\
  a_{13} & a_{23}
\end{bmatrix}
= 
\begin{bmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23}
\end{bmatrix}^T
\]
Multiple Pass Transpose

To compute $B = A^T$ where $A \in \mathbb{IR}^{m \times n}$ factor

$$S_{n,m} = \Gamma_t \cdots \Gamma_1$$

and then execute

\begin{verbatim}
  a = vec(A)
  for k = 1:t
    a ← Γ_k a
  end
  Define $B \in \mathbb{IR}^{n \times m}$ by vec($B$) = a.
\end{verbatim}

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:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} \rightarrow B^{(1)} = \begin{bmatrix} A_1 | A_2 | A_3 | A_4 \end{bmatrix}.$$ 

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

$$B^{(1)} = \begin{bmatrix} A_1 | A_2 | A_3 | A_4 \end{bmatrix} \rightarrow B^{(2)} = \begin{bmatrix} A_1^T | A_2^T | A_3^T | A_4^T \end{bmatrix}.$$ 

Note that $B^{(2)} = A^T$. 
The Rise of the Kronecker Product Cont’d

Tensoring Low-Dimension Techniques

Greater willingness to entertain problems of high dimension.
Tensoring Low Dimensional Ideas

Quadrature in one and three dimensions:

\[
\int_{a}^{b} f(x) \, dx \approx \sum_{i=1}^{n} w_i \, f(x_i)
\]

\[
= \mathbf{w}^T f(x)
\]

\[
\int_{a_1}^{b_1} \int_{a_2}^{b_2} \int_{a_3}^{b_3} f(x, y, z) \, dx \, dy \, dz \approx \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} w_i^{(x)} w_j^{(y)} w_k^{(z)} f(x_i, y_j, z_k)
\]

\[
= (\mathbf{w}^{(x)} \otimes \mathbf{w}^{(y)} \otimes \mathbf{w}^{(z)})^T f(x \otimes y \otimes z)
\]
Higher Dimensional KP Problems


   \textit{In higher-order statistics the calculations involve the “cumulants’} \( x \otimes x \otimes \cdots \otimes x \). \textit{(Note that vec}(xx^T) = x \otimes x\).

2. de Lathauwer, de Moor, and Vandewalle (2000)

   \textit{multilinear SVD, low rank approximation of tensors}
Least Squares Problems that Involve Kronecker Products
Least squares problems of the form

\[ \min \| (B \otimes C)x - b \| \]

can be efficiently solved by computing the QR factorizations (or SVDs) of \( B \) and \( C \). See Fausett and Fulton (1994) and Fausett, Fulton, and Hashish (1997).

Barrlund (1998) on surface fitting with splines...

\[ \| (A_1 \otimes A_2)x - f \| \quad \text{such that} \quad (B_1 \otimes B_2)x = g \]
Weighted Least Squares

\[
\min \| W^{-1/2}(B \otimes C)x - b \|_2 \equiv \begin{bmatrix}
W & B \otimes C \\
B^T \otimes C^T & 0
\end{bmatrix} \begin{bmatrix}
r \\
x
\end{bmatrix} = \begin{bmatrix}
b \\
0
\end{bmatrix}
\]

Compute the QR factorizations

\[
B = Q_B \begin{bmatrix}
R_B \\
0
\end{bmatrix} \quad C = Q_C \begin{bmatrix}
R_C \\
0
\end{bmatrix}
\]
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 conjugate gradients exploiting structure.
When the blocks are Kronecker products...

\[
\min \left\| \begin{bmatrix} B_1 \otimes C_1 \\ B_2 \otimes C_2 \end{bmatrix} x - b \right\|_2
\]

Compute a pair of GSVD’s:

\[
B_1 = U_1 B D_1 B X_B^T, \quad B_2 = U_2 B D_2 B X_B^T
\]

\[
C_1 = U_1 C D_1 C X_C^T, \quad C_2 = U_2 C D_2 C X_C^T.
\]

\[
\begin{bmatrix} B_1 \otimes C_1 \\ B_2 \otimes C_2 \end{bmatrix} = \begin{bmatrix} U_1 B \otimes U_2 B & 0 \\ 0 & U_1 C \otimes U_2 C \end{bmatrix} \begin{bmatrix} D_1 B \otimes D_2 B \\ D_1 C \otimes D_2 C \end{bmatrix} X_B^T \otimes X_C^T.
\]
Total Least Squares

\[
\text{LS: } \min \| r \|_2^2 \quad \text{subject to } b + r \in \text{ran}(A)
\]
\[
Ax_{LS} = b + r_{opt}
\]

\[
\text{TLS: } \min \| E \|_F^2 + \| r \|_2^2 \quad \text{subject to } b + r \in \text{ran}(A + E)
\]
\[
(A + E_{opt})x_{TLS} = b + r_{opt}
\]

“Errors in Variables”
Total Least Squares Solution

To solve

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

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

$$U^T [A \mid b] V = \Sigma$$

and set

$$x_{\text{TLS}} = -V(1:n, n+1)/V(n+1, n+1)$$
If $A$ is a Kronecker Product $B \otimes C$...

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

$$F = \begin{bmatrix} B \otimes C \mid b \end{bmatrix}$$

First compute the SVDs of $B$ and $C$:

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

If

$$\tilde{U} = U_B \otimes U_C \quad \tilde{V} = \begin{bmatrix} V_B \otimes V_C \mid 0 \end{bmatrix}$$

then

$$\tilde{U}^T F \tilde{V} = \begin{bmatrix} \Sigma_B \otimes \Sigma_C \mid g \end{bmatrix} \equiv \tilde{F} \quad \text{where} \quad g = \tilde{U}^T b$$

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

The right singular vectors of $\tilde{F}$ are the eigenvectors of

$$C = \tilde{F}^T \tilde{F} = \left[ \sum_B \otimes \sum_C \mid g \right]^T \left[ \sum_B \otimes \sum_C \mid g \right] = \begin{bmatrix} D & z \\ z^T & \alpha \end{bmatrix}$$

where

$$D = \sum_B^T \sum_B \otimes \sum_C^T \sum_C \quad \text{(square and diagonal)}$$

$$z = (\sum_B \otimes \sum_C)^T g$$

$$\alpha = g^T g$$
Eigenvectors/Eigenvalues of Bordered Matrices

\[ C = \begin{bmatrix}
  d_1 & 0 & 0 & 0 & z_1 \\
  0 & d_2 & 0 & 0 & z_2 \\
  0 & 0 & d_3 & 0 & z_3 \\
  0 & 0 & 0 & d_4 & z_4 \\
  z_1 & z_2 & z_3 & z_4 & \alpha
\end{bmatrix} \]

The eigenvalues of \( C \) are the zeros of

\[ f(\mu) = \mu - \alpha + z^T (D - \mu I)^{-1} z = \mu - \alpha + \sum_{i=1}^{N} \frac{z_i^2}{d_i - \mu} \]

Can assume that the \( d_i \) are distinct and that the \( z_i \) are nonzero.

The zeros of \( f \) are separated by the \( d_i \). Can show that the smallest root of \( f \) (a.k.a. the smallest eigenvalue of \( \tilde{F} \)) is between 0 and the smallest \( d_i \).
From the eigenvalue equation

\[
\begin{bmatrix}
  D & z \\
  z^T & \alpha \\
\end{bmatrix}
\begin{bmatrix}
  x \\
  -1 \\
\end{bmatrix}
= \mu_{\text{min}}
\begin{bmatrix}
  x \\
  -1 \\
\end{bmatrix}
\]

it follows that

\[
x = (D - \mu_{\text{min}}I)^{-1}z
\]

Thus the sought-after singular vector of \( \tilde{F} \) is just a unit vector in the direction

\[
\begin{bmatrix}
  (D - \mu_{\text{min}}I)^{-1}z \\
  -1 \\
\end{bmatrix}
\]
Overall Solution Procedure

1. First compute the SVDs of $B$ and $C$:
   
   $U_B^T B V_B = \Sigma_B \quad U_C^T C V_C = \Sigma_C$

2. Find smallest root of $f(\mu) = \mu - \alpha + z^T (D - \mu I)^{-1} z$.

3. Get smallest singular vector $w$ of
   
   $[ \Sigma_B \otimes \Sigma_C \mid g ]$

4. Get the TLS solution from
   
   $v = \begin{bmatrix} V_B \otimes V_C & 0 \\ 0 & 1 \end{bmatrix} w$
A Shifted Kronecker Product System
Frequency Response

Suppose we wish to evaluate for many different values of $\mu$.

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

It pays to recompute the real Schur decomposition..

$$Q^T A Q = T = \begin{bmatrix}
\times & \times & \times & \times & \times & \times \\
0 & \times & \times & \times & \times & \times \\
0 & 0 & 0 & \times & \times & \times \\
0 & 0 & 0 & 0 & \times & \times \\
0 & 0 & 0 & 0 & 0 & \times \\
\end{bmatrix}$$

$$\phi(\mu) = c^T Q(T - \mu I)^{-1}Q^T d \quad (O(n^2) \text{ per evaluation})$$
Solve \((A_p \otimes \cdots \otimes A_1 - \lambda I) x = b\).

First compute the real Schur decompositions \(A_k = Q_k T_k Q_k^T, \ i = 1:p\)

Set \(Q = Q^{(p)} \otimes \cdots \otimes Q^{(1)}\).

It follows that

\[ Q^T \left( A^{(p)} \otimes \cdots \otimes A^{(1)} \right) Q = T^{(p)} \otimes \cdots \otimes T^{(1)} \equiv T \]

and we get

\[ \left( T^{(p)} \otimes \cdots \otimes T^{(1)} - \lambda I_n \right) y = c \]

where \(y \in \mathbb{IR}^n\) and \(c \in \mathbb{IR}^n\) are defined by

\[ x = \left( Q^{(p)} \otimes \cdots \otimes Q^{(1)} \right) y \quad c = \left( Q^{(p)} \otimes \cdots \otimes Q^{(1)} \right)^T b. \]
Unshifted Case

The standard approach for solving

\[
\left(T^{(p)} \otimes \cdots \otimes T^{(1)}\right) y = c \quad T^{(i)} \in \mathbb{R}^{n_i \times n_i}
\]
is to recognize that

\[
T = \prod_{i=1}^{p} \left( I_{\rho_i} \otimes T^{(i)} \otimes I_{\mu_i} \right)
\]

where \( \rho_i = n_1 \cdots n_{i-1} \) and \( \mu_i = n_{i+1} \cdots n_p \) for \( i = 1:p \).

We then obtain...

\[
y \leftarrow c
\]
for \( i = 1:p \)

\[
y \leftarrow \left( I_{\rho_i} \otimes T^{(i)} \otimes I_{\mu_i} \right)^{-1} y
\]
end
The $p = 2$ case: $(F \otimes G - \lambda I)x = b$

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

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 $\tilde{c}_i = c_i - f_{i4}Gy_4$, $i = 1:3$. 
The General Triangular Case

It is recursive...

\[
\textbf{function } \ y = \text{KPShiftSolve}(T, c, \lambda, n, \alpha) \\
n \text{ is a } p \text{-by-1 integer vector} \\
T \text{ is a } p \text{-by-1 cell array where } T\{i\} \text{ is an} \\
n_i \text{-by-} n_i \text{ upper triangular matrix for } i = 1:p. \\
\text{Assume that } c \in \mathbb{R}^N \text{ where} \\
N = n_1 \cdots n_p \text{ and} \\
\lambda \text{ and } \alpha \text{ are scalars so } A = (\alpha T\{p\} \otimes \cdots \otimes T\{1\} - \lambda I_N) \\
A \text{ is nonsingular. } y \in \mathbb{R}^N \text{ solves } Ay = c.
\]
\[ p \leftarrow \text{length}(n) \]
if \( p == 1 \)
\[ y \leftarrow (\alpha T\{1\} - \lambda I)c \]
else
\[ m \leftarrow n_1 \cdots n_{p-1} \]
for \( i = n_p: -1:1 \)
\[ \text{id}x \leftarrow 1 + (i - 1)m:im \]
\[ \alpha \leftarrow T\{p\}(i, i) \]
\[ y(\text{id}x) \leftarrow \text{KPShiftSolve}(T, b(\text{id}x), \lambda, n(1:p - 1), \alpha) \]
\[ z \leftarrow (T\{p - 1\} \otimes \cdots \otimes T\{1\})y(\text{id}x) \]
for \( j = 1:i - 1 \)
\[ \text{id}x \leftarrow 1 + (j - 1)m:im \]
\[ c(\text{id}x) \leftarrow c(\text{id}x) - T\{p\}(j, i)z \]
end
end
Two-by-Two Bumps

\[
\begin{bmatrix}
  f_{11}G - \lambda I_m & f_{12}G & f_{13}G & f_{14}G \\
  0 & f_{22}G - \lambda I_m & f_{23}G & f_{24}G \\
  0 & 0 & f_{33}G - \lambda I_m & f_{34}G \\
  0 & 0 & f_{43}G & f_{44}G - \lambda I_m
\end{bmatrix}
\begin{bmatrix}
  y_1 \\
  y_2 \\
  y_3 \\
  y_4
\end{bmatrix}
= \begin{bmatrix}
  c_1 \\
  c_2 \\
  c_3 \\
  c_4
\end{bmatrix}
\]

Solve for \(y_3\) and \(y_4\) at the same time.

Revert to localized complex arithmetic.

\[
Z^H \begin{bmatrix}
  f_{33} & f_{34} \\
  f_{43} & f_{44}
\end{bmatrix} Z = \begin{bmatrix}
  t_{33} & t_{34} \\
  0 & t_{44}
\end{bmatrix}
\]
High Order Kronecker Products

\[ y = (F_p \otimes \cdots \otimes F_1) x \quad F_i \in \mathbb{R}^{n_i \times n_i} \]

It is convenient to make use of the factorization

\[ F_p \otimes \cdots \otimes F_1 = M_p \cdots M_1 \]

where

\[ M_i = \prod_{n_i,N/n_i}^T (I_{N/n_i} \otimes F_i) \quad N = n_1 \cdots n_p \]

**MATLAB:**

\[
Z \leftarrow x \\
\text{for } i = 1:p \\
\quad Z \leftarrow (F_i \cdot \text{reshape}(Z, n_i, N/n_i))^T \\
\text{end} \\
y \leftarrow \text{reshape}(Z, N, 1)
\]
Conclusion

Our goal is 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”.

With precedent...

- The development of effective algorithms for the QR and SVD factorizations turned many “$A^T A$” problems into least square/singular value calculations.

- The development of the QZ algorithm (Moler and Stewart (1973)) for the problem $Ax = \lambda Bx$ prompted a rethink of “standard” eigenproblems that have the form $B^{-1} Ax = \lambda x$. 