The Quotient Singular Value Decomposition
(Van Loan, Paige and Saunders)

Let $A, B \in \mathbf{R}^{m \times n}, m > n^*$, $C^T[A^T \;\; B^T]$, and $k = \text{rank}(C)$. Then there exists unitary $U, V \in \mathbf{R}^{m \times m}$ and diagonal matrices $\mathbf{\Sigma}_A, \mathbf{\Sigma}_B$ so that:

$$A = U\mathbf{\Sigma}_A P^T, \quad B = V\mathbf{\Sigma}_B P^T$$

and

$$\mathbf{\Sigma}_A = \begin{bmatrix} I_A^r & & \\ & S_A^s & \\ & & O_A^{k-(r+s)} \end{bmatrix}, \mathbf{\Sigma}_B = \begin{bmatrix} O_B^r & & \\ & S_B^s & \\ & & I_B^{k-(r+s)} \end{bmatrix}$$

First $r$ cols of $P$: Row$(A) \bigcap$ Null$(B)$
Next $s$ cols of $P$: Row$(A) \bigcap$ Row$(B)$
Remainder: Null$(A) \bigcap$ Row$(B)$

Stationary pts to $\dfrac{x^T A^T A x}{x^T B^T B x}$

*Only require $A \in \mathbf{R}^{m \times p}, B \in \mathbf{R}^{n \times p}$

## Joint Diagonalization and Generalized Eigenvectors

If $A$ full rank- generalized evecs= 2 evec comps (commonly done in practice):

Joint Diagonalization of $A^T A$, $B^T B$:

Let $A^T A = \Phi \Theta \Phi^T$
Take $\hat{B} = B \Phi \Theta^{-1/2}$ (Whitening)
Take $\hat{B}^T \hat{B} = \Psi \Lambda \Psi^T$

Joint Diagonalizer: $\Phi \Theta^{-\frac{1}{2}} \Psi^T$, since:

$$\Psi \Theta^{-\frac{1}{2}} \Phi^T A^T A \Phi \Theta^{-\frac{1}{2}} \Psi^T = I$$

$$\Psi \Theta^{-\frac{1}{2}} \Phi^T B^T B \Phi \Theta^{-\frac{1}{2}} \Psi^T = \Lambda$$

## Application 6:
## Max Noise Fraction for Time Series*

Let $X^{t\times n}$ be a multidimensional time series contaminated with noise:

$$X = S + N$$

Assumption: $X^T X = S^T S + N^T N$

Goal: Find a splitting set, $\{\psi_i\}_{i=1^n}$ so if

$$\Psi_1 = \{\psi_1, \ldots, \psi_r\}, \Psi_2 = \left\{\psi_{r+1}, \ldots, \psi_n\right\}$$

then:

$$\begin{array}{ll} S\Psi_1 = \widehat{S} & S\Psi_2 = 0 \\ N\Psi_1 = 0 & N\Psi_2 = \widehat{N} \end{array}$$

Solution: $\Psi_1$ are gen evecs of $(S, N)$ at $\infty$, $\Psi_2$ are gen evecs of $(S, N)$ at 0.

PROBLEM: We don't have $N$!

*Kirby and Anderle

Solve by estimating $N^T N$...

Define $\Delta\cdot = \cdot(n+1,:) - \cdot(n,:)$

Assumption: $N^T N \approx \Delta X^T \Delta X$.

Now we clean the data by maximizing the noise fraction:

$$\max_{\Delta X w \neq 0} \frac{w^T X^T X w}{w^T \Delta X^T \Delta X w}$$

which is performed by the generalized eigenvector routine...

Example: Data consisting of $\cos(t), \sin(t)$ and noise, with a random projection to $\mathbf{R}^3$.