PCA dual problem (cont.)

Thus:

- could do SVD of $\tilde{X}$, to compute Eigen-analysis
- i.e. replace $d \times d$ analysis by $d \times n$
- Singular Values are $\pm\sqrt{\ }$ of cov. matrix eigenvalues
- (usually taken as + square-root)
- Columns of U can be used for PCA projections
- since they *are* the eigenvectors (i.e. $B = U$)
- so PCA is *both*:
  - eigen-decomposition of covariance matrix
  - singular value decomposition of data matrix

PCA dual problem (cont.)

Improve to $n \times n$ analysis?

Can make U and V “change places” by considering $\tilde{X}^t$

Singular Value Decomposition is:

$$\tilde{X}^t = \left(USV^t\right)^t = VS^tU^t$$

Sizes are useful:

$$n \times d = n \times n \leftrightarrow n \times d \leftrightarrow d \times d$$

Return to an eigen representation as:

$$\tilde{X}^t \tilde{X} = VS^tU^t\left(USV^t\right) = VS^tSV^t$$

PCA dual problem (cont.)

Again for $n < d$:

$$S = \begin{pmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}_{d \times n}, \quad \text{so define} \quad D^* = S^t S = \begin{pmatrix} s_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_n^2 \end{pmatrix}_{n \times n}$$

and define $B^* = V$, to give

$$\widetilde{X}^t \widetilde{X} = B^* D^* \left(B^*\right)^t$$

the "dual" eigen represntation

PCA dual problem (cont.)

Have "dual eigen problem", where

- Treat $\tilde{X}^t$ as "data matrix"
- i.e. rows of data matrix $\tilde{X}$ are replaced by columns
- Based on "dual covariance matrix" $\hat{\Sigma}^*_{n \times n} = \tilde{X}^t \tilde{X}$
- "inner product of data matrix"
- compared to "outer product" for calculation of $\hat{\Sigma}_{d \times d} = \tilde{X}\tilde{X}^t$
- for $n < d$, have faster $n \times n$ eigen-calculation

[go to next page]