Data matrix

- n by p data matrix or design matrix

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix}$$

- Each row consists of a vector of measurements x_i^T from p variables for a particular observation; X is often column centered

National Space Institute

PCA

- R-mode or primal analysis: decompose p by p variance-covariance matrix $S = X^T X/(n-1) = 1/(n-1)\sum_{i=1}^n x_i x_i^T$

$$\frac{1}{n-1} X^T X u_i = \lambda_i u_i$$

- Projections or scores are $x^T u_i$
- Variance of scores maximized

$$u_i^T S u_i = \lambda_i u_i^T u_i = \lambda_i$$

National Space Institute

PCA

- Q-mode or dual analysis: decompose n by n (Gram) matrix $XX^T/(n-1)$, multiply from left with X, $v_i \propto X u_i$

$$\frac{1}{n-1} XX^T (Xu_i) = \lambda_i (Xu_i)$$

$$\frac{1}{n-1} XX^T v_i = \lambda_i v_i$$

National Space Institute

PCA

- Primal and dual problems have same eigenvalues, eigenvectors related by

$$u_i = X^T v_i / \sqrt{(n-1)\lambda_i}$$

$$v_i = X u_i / \sqrt{(n-1)\lambda_i}$$

National Space Institute

Gram matrix

- Consists of inner products only

$$XX^T = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \cdots & x_1^T x_n \\ x_2^T x_1 & x_2^T x_2 & \cdots & x_2^T x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n^T x_1 & x_n^T x_2 & \cdots & x_n^T x_n \end{bmatrix}$$

National Space Institute

SVD

X is n by p cc data matrix

$X = UDV^T$

D is n by p diagonal with nonnegative, non-increasing singular values

U is n by n unitary

V is p by p unitary

National Space Institute

SVD

- X is n by p cc data matrix
- $X = UDV^T$
- D is n by p diagonal with nonnegative, non-increasing singular values
- U is n by n unitary
- V is p by p unitary

National Space Institute

SVD 'economy size'

- X is n by p data matrix, $n > p$
- $X = UDV^T$
- D is p by p diagonal with nonnegative, non-increasing singular values
- U is n by p unitary
- V is p by p unitary

National Space Institute

SVD / PCA / EOF

$$
\begin{aligned}
S &= X^T X/(n-1) \\
Sp_i &= \lambda_i p_i \\
SP &= P\Lambda \\
S &= P\Lambda P^T
\end{aligned}
$$

$$
\begin{aligned}
X^T X &= (UDV^T)^T UDV^T \\
&= VDU^T UDV^T \\
&= VD^2V^T \quad (= VD(VD)^T)
\end{aligned}
$$

National Space Institute

SVD / PCA / EOF

$$
\begin{aligned}
P &= V \\
D^2 &= (n-1)\Lambda
\end{aligned}
$$

Principal component scores

$$
\begin{aligned}
XP &= UDV^T V \\
&= UD
\end{aligned}
$$

National Space Institute