Section 6.10.8
Principal Component Analysis
PCA

PCA gives us:
- Dimensionality Reduction
- PCA Space
- What things are similar and different?

Applications:
- Identification
- Classification
- Recommender Service

$\underline{X_{m \times n}}$

Features/Measurements

$$\begin{array}{cccc} & x_1 & x_2 & x_3 \cdots \\ s_1 & | & | & | \\ s_2 & | & | & | \\ s_3 & | & | & | \\ \vdots & & & \end{array}$$

Samples

$$\begin{array}{cc} & S_1 \; S_2 \; S_3 \cdots \\ x_1 & \text{———} \\ x_2 & \text{———} \\ x_3 & \text{———} \\ \vdots & \end{array}$$

Samples

* zero mean for each feature
  repmat → matrix of means

$$D = \begin{bmatrix} | & | & \\ (x_1-\mu_1) & (x_2-\mu_2) & \cdots \\ | & | & \end{bmatrix} \quad D = \begin{bmatrix} - (x_1-\mu_1) - \\ - (x_2-\mu_2) - \\ \vdots \end{bmatrix}$$

$$\mathrm{cov}_X = \begin{bmatrix} \sigma_{x_1}^2 & \sigma_{x_2x_1} & \sigma_{x_3x_1} \cdots \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 & \cdots \\ \vdots & & \ddots \end{bmatrix}$$

$$\mathrm{cov}_X = \frac{D^T D}{m-1} \qquad \mathrm{cov}_X = \frac{D D^T}{n-1}$$

* [V, L] = eig(covx);
* sort V by eigenvalues
* Limit W to top k columns of V based on the eigenvalues

$$W = [\, v_1 \;\; v_2 \;\; \cdots \;\; v_k \,]$$

$$\boxed{Y = DW} \qquad \boxed{Y = W^T D}$$

— The PCA Space —

- Y has smaller dimension than X.
- Y is a linear combination of D data.
- Y shows the Principal component of the data.

To reconstruct X in reduced dimension:

[usually not needed]

* Reverse the steps

$\tilde{X} = D + \mu$

$Y = DW$

$D = YW^T$

$$\boxed{\hat{X} = YW^T + \mu}$$

$\tilde{x} = D + \mu$

$Y = W^T D$

$D = WY$

$$\boxed{\hat{x} = WY + \mu}$$