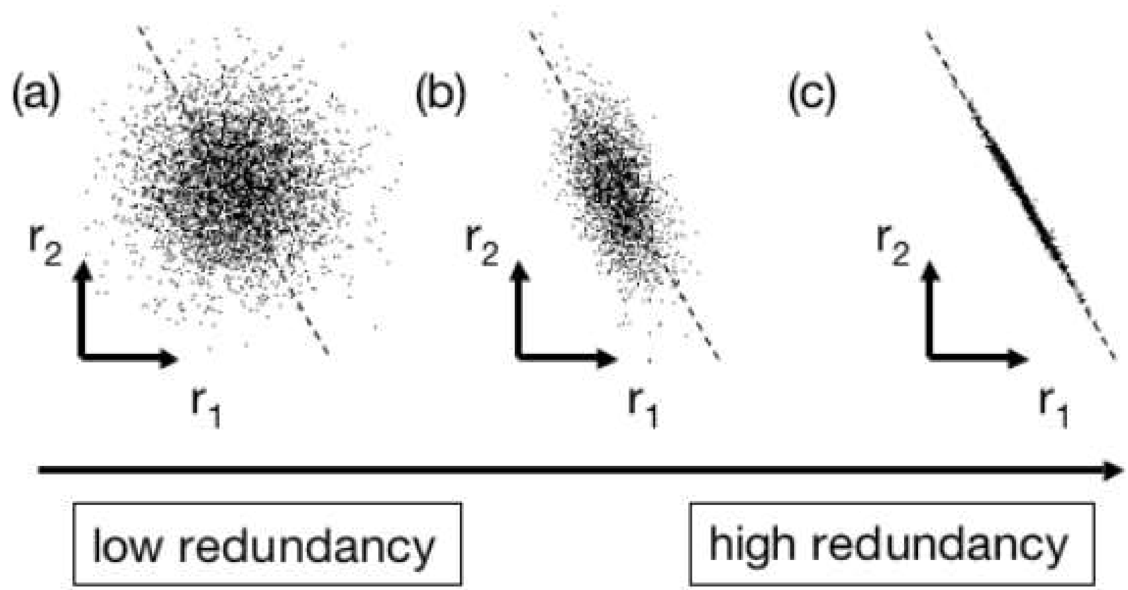
Dimensionality Reduction

December 7, 2012

Principal Component Analysis (PCA) statistical problem formulation



- Covariation matrix for (a) $\begin{pmatrix} \sigma_{1,1} & 0 \\ 0 & \sigma_{2,2} \end{pmatrix}$
- Covariation matrix for (c) $\begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{2,2} \end{pmatrix}$

Covariance

$$X = \begin{pmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{pmatrix}$$

Let each x_i has zero mean

$\sigma_{i,j} = \frac{1}{n-1} x_i^T x_j$

- $i = j$ - variance
- $i <> j$ - covariance
- if $\sigma_{i,j} = 0$ then x_i and x_j tottaly uncorrelated

$S_X = \frac{1}{n-1} X^T X$ - covariance matrix

Change of Basis

- Let P be the $n \times n$ matrix. And each row is basis vector of new subspace

$$PX = \begin{pmatrix} - & p_1 & - \\ - & p_2 & - \\ \cdots & \cdots & \cdots \\ - & p_n & - \end{pmatrix} \begin{pmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_m \\ | & | & \cdots & | \end{pmatrix}$$

where each $||p_i|| = 1$

- Each product $p_i \cdot x_j$ - is the length of progection of each x_j onto vector p_i
- In other words each product $p_i \cdot x_j$ is the coordinate on axis p_i

Desired Covariance Properties

- We would like to find matrix P such that the covariances between transformed feature vectors to be zero.

$$S_y = \frac{1}{n-1} Y^T Y, \text{ where } Y = PX$$

- In this case we eliminate all redundancy.
- The rows of matrix P are the principal components

Eigenvectors of Covariance

- It could be shown that if rows of matrix P will be chosen as eigenvectors of XX^T

$$S_Y = \frac{1}{n-1}YY^T = \frac{1}{n-1}(PX)(PX)^T = \frac{1}{n-1}D$$

- D - some diagonal matrix

Egenvectors and SVD

- SVD - of $m \times n$ matrix M is a factorization of the form

$$M = U\Sigma V^T$$

- U is an $m \times m$ real or complex unitary matrix
- Σ is an $m \times n$ diagonal matrix with nonnegative real numbers on the diagonal
- V^T is an $n \times n$ real or complex unitary matrix.
- It could be easily shown that
  - columns of U is the eigenvectors of MM^T
  - squares of Σ diagonal elements is the eigenvalues of MM^T

Usage of SVD for PCA

We want to map vector $x^{(i)} \in \mathbb{R}^n \to y^{(i)} \in \mathbb{R}^k$

- Extract mean from each column of X
- Estimate covariance matrix S_X
- Using SVD factorize XX^T columns of U is principal components of X

$$U = \begin{pmatrix} | & | & \cdots & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & \cdots & | \end{pmatrix} \to U' = \begin{pmatrix} | & | & \cdots & | \\ u_1 & u_2 & \cdots & u_k \\ | & | & \cdots & | \end{pmatrix}$$

- Convert each $x^{(i)}$ element to $y^{(i)}$ by

$$U'^T \cdot x^{(i)} = y^{(i)}$$