1.3 SINGULAR VALUE DECOMPOSITION

Singular Values and Singular Vectors

Singular values of an $m \times n$ matrix A are the square roots of $\min\{m, n\}$ eigenvalues of A^*A.

$$\sigma(A) = \sqrt{\lambda(A^*A)}$$

Right singular vectors of a matrix A are the eigenvectors of A^*A.

$$\sigma(A)^2 v - A^*Av = 0$$

Left singular vectors of a matrix A are the eigenvectors of AA^*.

$$\sigma(A)^2 u - AA^*u = 0$$

$$\bar{\sigma}(A) = \text{ the largest singular value of } A = \max_{\|x\|=1} \|Ax\| = \|A\|_2$$

The largest possible size change of a vector by A.

$$\underline{\sigma}(A) = \text{ the smallest singular value of } A = \min_{\|x\|=1} \|Ax\|$$

The smallest possible size change of a vector by A.

Singular Values and Singular Vectors (Continued)

Condition number: $c(A) = \frac{\underline{\sigma}(A)}{\bar{\sigma}(A)}$

$$A\bar{v} = \bar{\sigma}\bar{u}$$

$$A\underline{v} = \underline{\sigma}\ \underline{u}$$

$$\Downarrow$$

$\bar{v}$ ($\underline{v}$): highest (lowest) gain input direction

$\bar{u}$ ($\underline{u}$): highest (lowest) gain observing direction

Singular Value Decomposition

Let $A \in \mathbf{R}^{m \times n}$. Suppose σ_i be singular values of A such that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0, \quad p = \min\{m, n\}$$

Let

$$U = [u_1, u_2, \cdots, u_m] \in \mathbf{R}^{m \times m} \qquad V = [v_1, v_2, \cdots, v_n] \in \mathbf{R}^{n \times n}$$

where u_i, v_j denote left and right orthonormal singular vectors of A, respectively. Then

$$A = U \Sigma V^*, \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} = \sum_{i=1}^{p} \sigma_i(A) u_i v_i^*$$

where

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}$$

Consider $y = Ax$. Then Σ is simply the representation of A when x and y are represented in the coordinate systems consisting of right and left singular vectors, respectively.

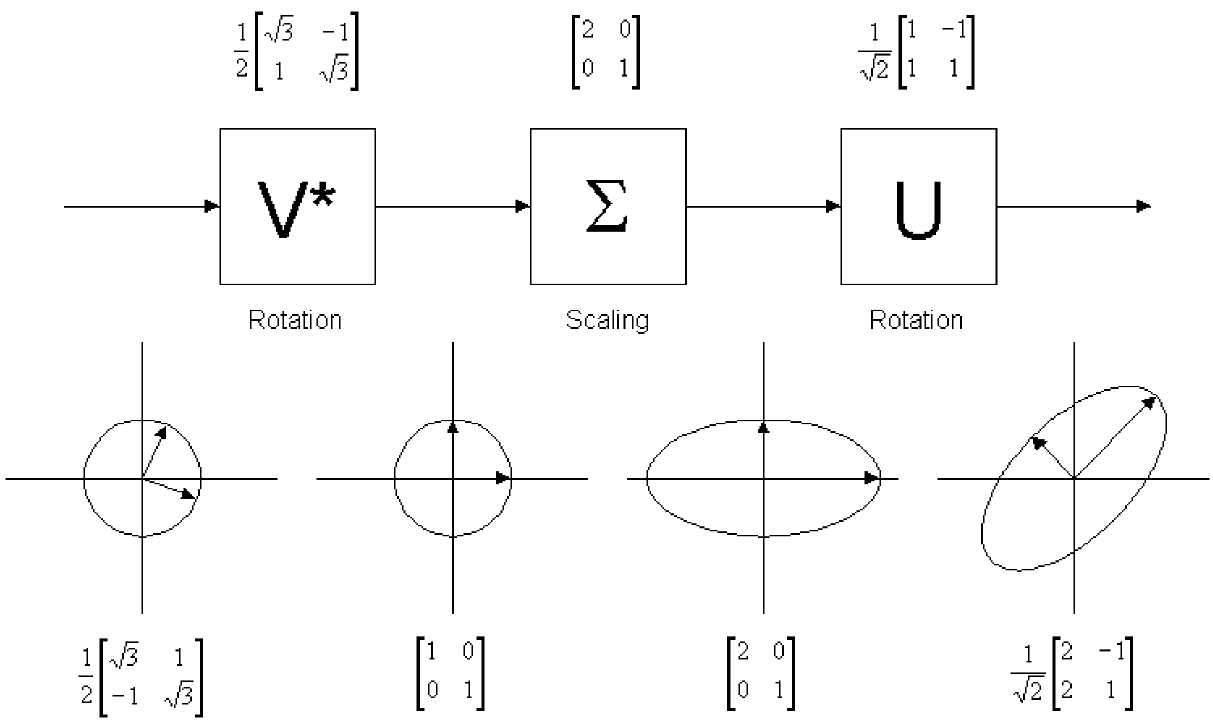
Singular Value Decomposition (Continued)

Example:

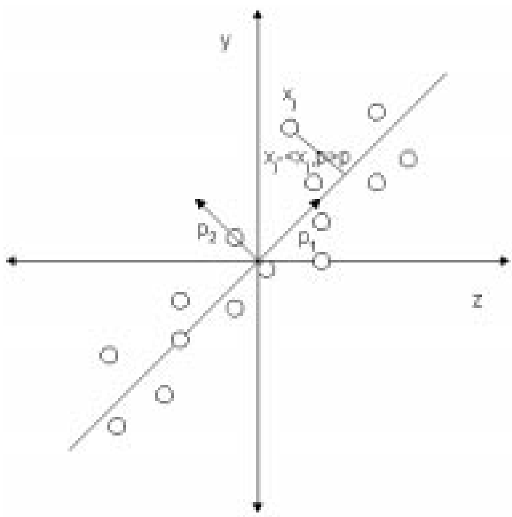
$$A = \begin{bmatrix} 0.8712 & -1.3195 \\ 1.5783 & -0.0947 \end{bmatrix}$$

$$\Downarrow$$

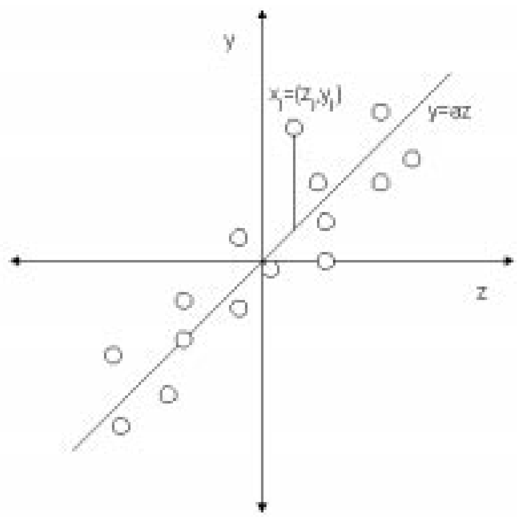
$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad V = \frac{1}{2}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix}$$



Principal Component Analysis



Concept of Principal Vector



Concept of Least Squares
Linear Regression of y on z

Given N n-dimensional vectors $\{x_1, x_2, \cdots, x_N\}$, the principal vector p is

$$p = \arg \min_{\|p\|=1} \sum_{i=1}^{N} \|x_i - \langle x_i, p\rangle p\|^2$$

$$= \arg \min_{\|p\|=1} \sum_{i=1}^{N} \left[\langle x_i, x_i\rangle - 2\langle x_i, p\rangle^2 + \langle x_i, p\rangle^2 \langle p, p\rangle\right]$$

$$= \arg \min_{\|p\|=1} \sum_{i=1}^{N} -\frac{\langle x_i, p\rangle^2}{\langle p, p\rangle} = \arg\max \sum_{i=1}^{N} \frac{\langle x_i, p\rangle^2}{\langle p, p\rangle} = \arg\max \alpha(p)$$

where

$$\alpha(p) = \sum_{i=1}^{N} \frac{x_i^T p p^T x_i}{p^T p}$$

Principal Component Analysis (Continued)

At the extremum,

$$0 = \frac{1}{2}\frac{d\alpha}{dp} = \sum_{i=1}^{N} \frac{x_i x_i^T p}{p^T p} - \sum_{i=1}^{N} \frac{x_i^T p p^T x_i p}{(p^T p)^2}$$

$$\Downarrow$$

$$0 = \sum_{i=1}^{N} x_i x_i^T p - \sum_{i=1}^{N} \frac{x_i^T p p^T x_i}{p^T p} p = X X^T p - \lambda p \quad \text{Singular Value Problem for } X$$

where

$$X = [x_1 \; x_2 \; \cdots \; x_N], \quad \lambda = \sum_{i=1}^{N} \frac{x_i^T p p^T x_i}{(p^T p)^2}$$

The SVD of X is

$$X = P\Lambda^{\frac{1}{2}} V^T = p_1 \lambda_1^{\frac{1}{2}} u_1^T + \cdots + p_n \lambda_n^{\frac{1}{2}} u_n^T$$

where

$$P = [p_1 \; p_2 \; \cdots \; p_n], \quad V = [v_1 \; v_2 \; \cdots \; v_N],$$

$$\Lambda = [diag[\lambda_i^{\frac{1}{2}}] \; 0] \quad 0 = X^T X v - \lambda v$$

$$\lambda_1^{\frac{1}{2}} \geq \cdots \geq \lambda_n^{\frac{1}{2}}$$

The approximation of X using first m significant principal vectors:

$$X \approx \bar{X} = \bar{P}\bar{\Lambda}^{\frac{1}{2}} \bar{U}^T = p_1 \lambda_1^{\frac{1}{2}} u_1^T + \cdots + p_m \lambda_m^{\frac{1}{2}} u_m^T$$

where

$$\bar{P} = [p_1 \; p_2 \; \cdots \; p_m], \quad \Lambda = diag[\lambda_i^{\frac{1}{2}}]_{i=1}^{m} \quad \bar{V} = [v_1 \; v_2 \; \cdots \; v_m]$$

Principal Component Analysis (Continued)

$$p_i^T X = p_i^T (p_1 \lambda_1^{\frac{1}{2}} u_1^T + \cdots + p_n \lambda_n^{\frac{1}{2}} u_n^T) = \lambda_i^{\frac{1}{2}} u_i^T$$

$$\Downarrow$$

$$\bar{P}^T X = \bar{U}^T$$

$$\Downarrow$$

$$\bar{X} = \bar{P}\bar{U}^T = \bar{P}\bar{P}^T X$$

and the residual is

$$\tilde{X} = X - \bar{X} = (I - \bar{P}\bar{P}^T)X$$