Singular Value Decomposition (SVD)

Def. An $m \times n$ matrix Σ is <u>rectangular diagonal</u> if $\Sigma_{ij} = 0$ for $i \neq j$.

Examples

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 0 \\ 0 & -6 \\ 0 & 0 \end{bmatrix}$$

The diagonal of Σ is

$$\text{diag}(\Sigma) = (\Sigma_{11}, \Sigma_{22}, \ldots, \Sigma_{kk})$$

where $k = \min(m, n)$.

Example

$\text{diag} = [3 \ 4 \ -1]$, or $[1, -6]$

<u>SVD Theorem</u> Any $m \times n$ real matrix A can be factored as

$$A = Q_1 \Sigma Q_2^T$$

where $Q_1 = m \times m$ orthogonal matrix

$Q_2 = n \times n$ orthogonal matrix

$\Sigma = m \times n$ rectangular diagonal matrix, and

$\operatorname{diag} \Sigma = [\sigma_1, \sigma_2, \ldots, \sigma_k]$ satisfies

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k \geq 0$$

where $k = \min(m, n)$. Moreover, the ~~matrices~~ columns of Q_1 are e-vectors of $A \cdot A^T$, the columns of Q_2 are e-vectors of $A^T A$, and the $(\sigma_i)^2$ are e-values of both $A \cdot A^T$ and $A^T A$. $\boxtimes$

Remark: The entries of $\mathrm{diag}(\Sigma)$ are called singular values of A.

Proof of the theorem

A^TA is $n\times n$ real and symmetric. Hence there exist orthonormal e-vectors $\{v^1, \dots, v^n\}$ such that $A^TAv^j = \lambda_j v^j$. W.L.O.G., we can assume that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

(if not, simply re-order the v^i's).

For $\lambda_j > 0$, say $1 \leq j \leq r$, we define

$$\sigma_j = \sqrt{\lambda_j}$$

and $$q^j = \frac{1}{\sigma_j} A v^j \in \mathbb{R}^m$$

<u>Claim</u> $(q^i)^T q^j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

for $1 \le i,j \le r$.

<u>Pf.</u> $$(q^i)^T q^j = \frac{1}{\sigma_i}\frac{1}{\sigma_j}(v^i)^T A^T A v^j$$

$$= \frac{\lambda_j}{\sigma_i \sigma_j}(v^i)^T v^j$$

$$= \begin{cases} \frac{\lambda_i}{(\sigma_i)^2} & i=j \\ 0 & i \neq j \end{cases} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

$\square$

If $r < m$, we can extend the q^i's to an orthonormal basis for $\mathbb{R}^m$.

Define

$$Q_1 = [q^1 \mid q^2 \mid \cdots \mid q^m]$$

$$Q_2 = [v^1 \mid v^2 \mid \cdots \mid v^n]$$

Define $\Sigma = m \times n$ by

$$\Sigma_{ij} = \begin{cases} \sigma_i \delta_{ij} & 1 \le i, j \le r \\ 0 & \text{otherwise.} \end{cases}$$

Then Σ is rectangular diagonal with

$\mathrm{diag}(\Sigma) = [\sigma_1, \sigma_2, .., \sigma_r, 0 .., 0]$.

Proof of the Theorem

It is enough to show that

$$Q_1^T A Q_2 = \Sigma$$

$$(Q_1^T A Q_2)_{ij} = q_i^T A v^j$$

If $j > r$, then $Av^j = 0$, and thus

$$q_i^T A v^j = 0$$

If $i > r$, then q^i selected to be orthogonal to $\{q^1, \dots, q^r\} = \{\frac{1}{\sigma_1} A v^1, \frac{1}{\sigma_2} A v^2, \dots, \frac{1}{\sigma_r} A v^r\}$, and thus $(q^i)^T A v^i = 0$.

Consider $1 \le i, j \le r$

$$(Q_1^T A Q_2)_{ij} = \frac{1}{\sigma_i} (v^i)^T A^T A v^j$$

$$= \frac{\lambda_i}{\sigma_i} v^{i\,T} v^j$$

$$= \sigma_i \, \delta_{ij}$$

as required.

$\square$