LECTURE 42

Singular Value Decomposition (SVD) of a rectangular matrix.
Consider a $K \times N$ data matrix

$$A : \mathbb{R}^N = \mathbb{R}(A^\top) \oplus \mathcal{N}(A) \rightarrow \mathbb{R}_K = \mathbb{R}(A) \oplus \mathcal{N}(A^\top)$$

The problem in data analysis is:

**Given:** The data vector $d$

**The data matrix $A$**

**Find:** The "best" solution $w$ to $Aw = d$ (x)

**Discussion:**

**CASE 1:** If the columns of $A$ are independent (but the rows are dependent, in general) then even though $(*)$ has no solution,

$$\|Ab^* - d\| = \min \Rightarrow A^\top Ab^* = A^\top d \Rightarrow b^* = (A^\top A)^{-1} A^\top d$$

is unique and given by

$$b^* = (A^\top A)^{-1} A^\top d$$

whenever $\dim \mathcal{N}(A) = 0$.

This is because

$$\mathcal{N}(A) = \mathcal{N}(A^\top A)$$

so that $\mathcal{N}(A^\top A)$ is also trivial, which guarantees that $(A^\top A)^{-1}$ exists.

Very useful side comment:

$$\dim \mathcal{R}(A) = \dim \mathcal{R}(A^\top) \quad \mathcal{R}(A^\top) = \mathcal{R}(A)$$
Case 2. If the columns of $A$ are dependent, then $\mathcal{N}(A) = \mathcal{N}(A^T)$ is non-trivial.

In that case there exist infinitely many least squares solutions to $(\ast)$. However, among them is a unique one which is "best".

The "best", i.e. optimal solution is the one that has minimum length.

$$b^* = (A^T A)^{-1} A^T$$
The singular value decomposition of the $K \times M$ data matrix $A$ has the following form.

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

$$Q_1^T \cdot A \cdot Q_2 = \Sigma$$
Step I: Diagonalize $A^T A$ via $Q_2$

$A^T A$ is positive semi-definite.

$Q^{-1} A Q = A^T A$

Step II: Construct $Q_1$ from $A Q_2^\top = Q_1 \Sigma$

(2)

a) Construct $\{q_i\}_{i=1}^r \subset K \rightarrow \Sigma$

From $A^T A x_i = \lambda_i x_i \Rightarrow \lambda_i = \frac{x_i^\top A^T A x_i}{x_i^\top x_i}$ we have $\lambda_i = \lambda_1, \ldots, \lambda_r$

and $A: R^m \rightarrow R^k$

obtain $q_i = \frac{A x_i}{\sigma_i}$ for $i = 1, \ldots, r$

They satisfy

$q_i^\top q_j = \frac{x_i^\top A^T A x_j}{\sigma_i \sigma_j} = \begin{cases} 0 & i \neq j \\ 1 + \epsilon & i = j \end{cases}$

b) Using G-S extend $q_1, \ldots, q_r$ to an o.n. basis $q_1, \ldots, q_r, q_{r+1}, \ldots, q_k$ for $R^k$. 

Thus we have

$$Q_1 = \begin{bmatrix} q_1 & \cdots & q_r & q_{r+1} & \cdots & q_k \end{bmatrix}$$

$$R^k \rightarrow R^k$$

from $\Pi^A$ via $G-S$.

**Step III**

$$Q_1^T \times \text{Eq.}(2) \Rightarrow Q_1^T \ A \ Q_2 = \Sigma$$

$$\text{Eq.}(2) \times Q_2^T \Rightarrow \ A = Q_1 \Sigma \ Q_2^T$$