LECTURE 42

Singular Value Decomposition (SVD) of a rectangular matrix

Consider a $K \times M$ <u>data matrix</u>

Using Haykin's notation

$$A : \underset{\text{domain space}}{R^M = \mathcal{R}(A^T) \oplus \mathcal{N}(A)} \longrightarrow \underset{\text{target space}}{R^K = \mathcal{R}(A) \oplus \mathcal{N}(A^T)}$$

$$\boxed{\text{cause}} \; w \rightsquigarrow A w = \vec{d} \; \boxed{\text{observed effect}}$$

The problem in data analysis is:

Given: The data vector $\vec{d}$

The data matrix A

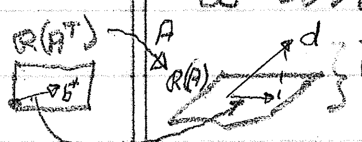
Find: The "best" solution w to $A\vec{w} = \vec{d}$ (*)

Disussion:

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

$\|Ab^* - \vec{d}\| = \min \Rightarrow A^TAb^* = A^Td \Rightarrow b^* = (A^TA)^{-1}A^Td$ is unique

furnishes the "best" solution in the form of the <u>least squares</u> solution which is unique and given by



$$b^* = (A^TA)^{-1}A^Td$$

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

This is because

$$\boxed{\mathcal{N}(A) = \mathcal{N}(A^TA)}$$

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

Very useful side comment:

$\dim \mathcal{R}(A) = \dim \mathcal{R}(A^T)$

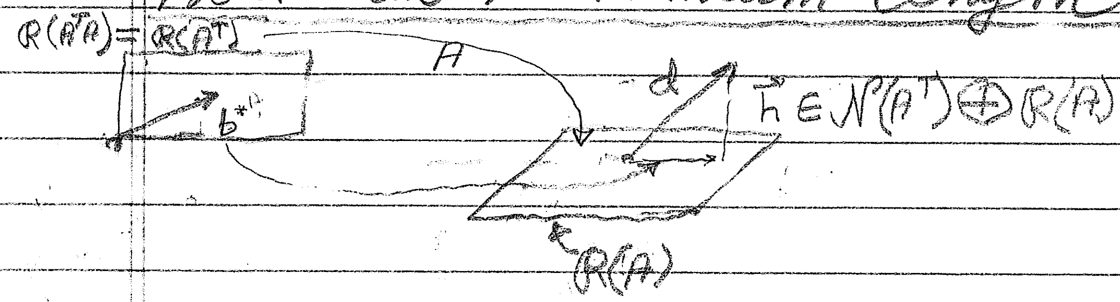
$$\boxed{\mathcal{R}(A^T) = \mathcal{R}(A^TA)}$$

CASE 2. If the columns of A are dependent then $\mathcal{N}(A) = \mathcal{N}(A^TA)$ is non-trivial

In that case there exist infinitely many least squares solutions to (*)

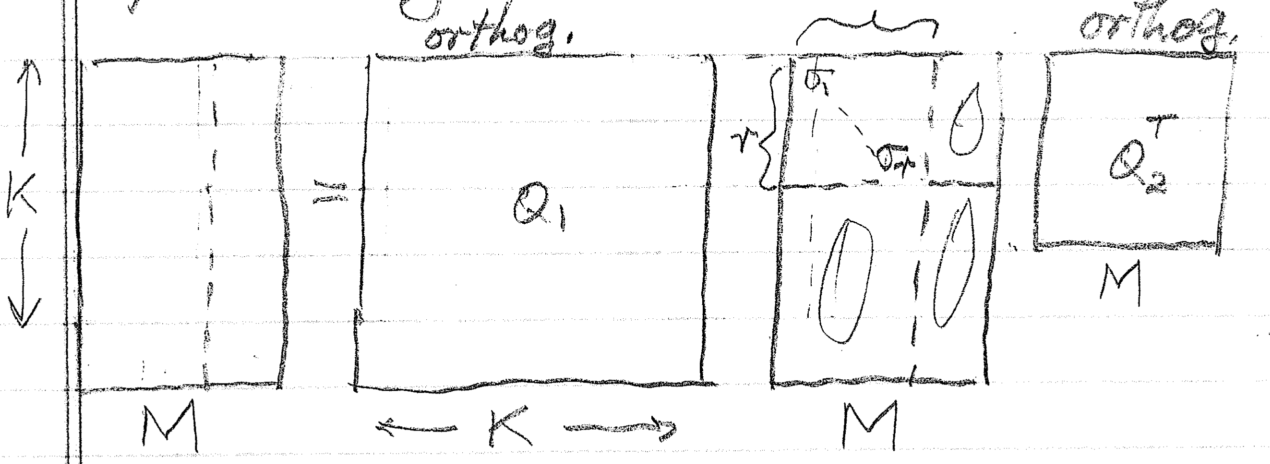
However, among them there is a unique one which is "best":

The "best", i.e optimal solution to Eq. (*) is the one that has minimum length.

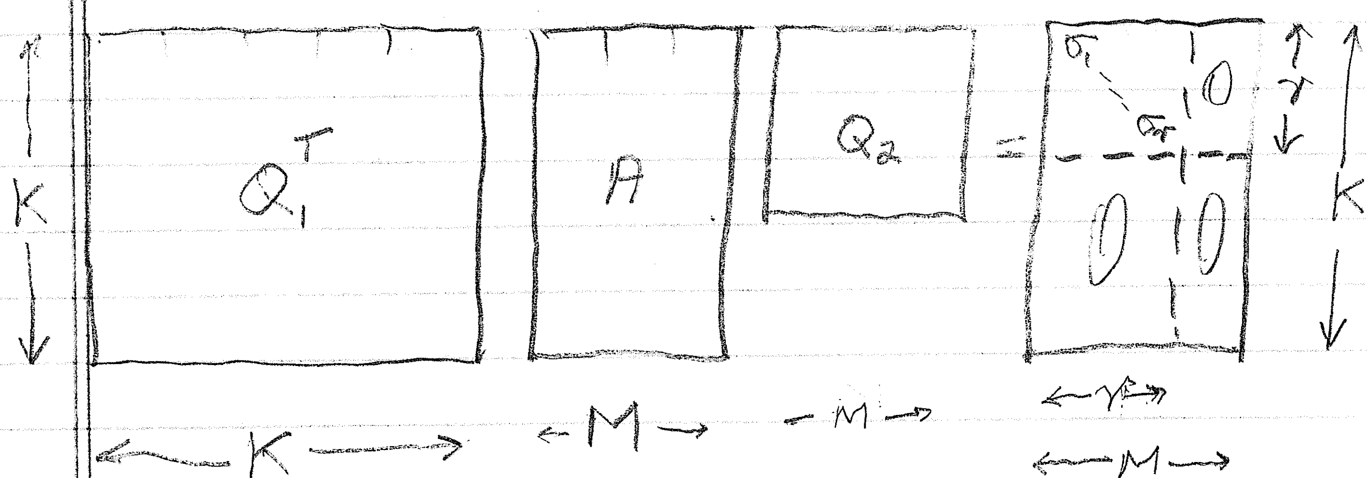


$$b^* = (A^TA)^{-1}A^Td$$

The singular value decomposition of the $K \times M$ data matrix A has the following form



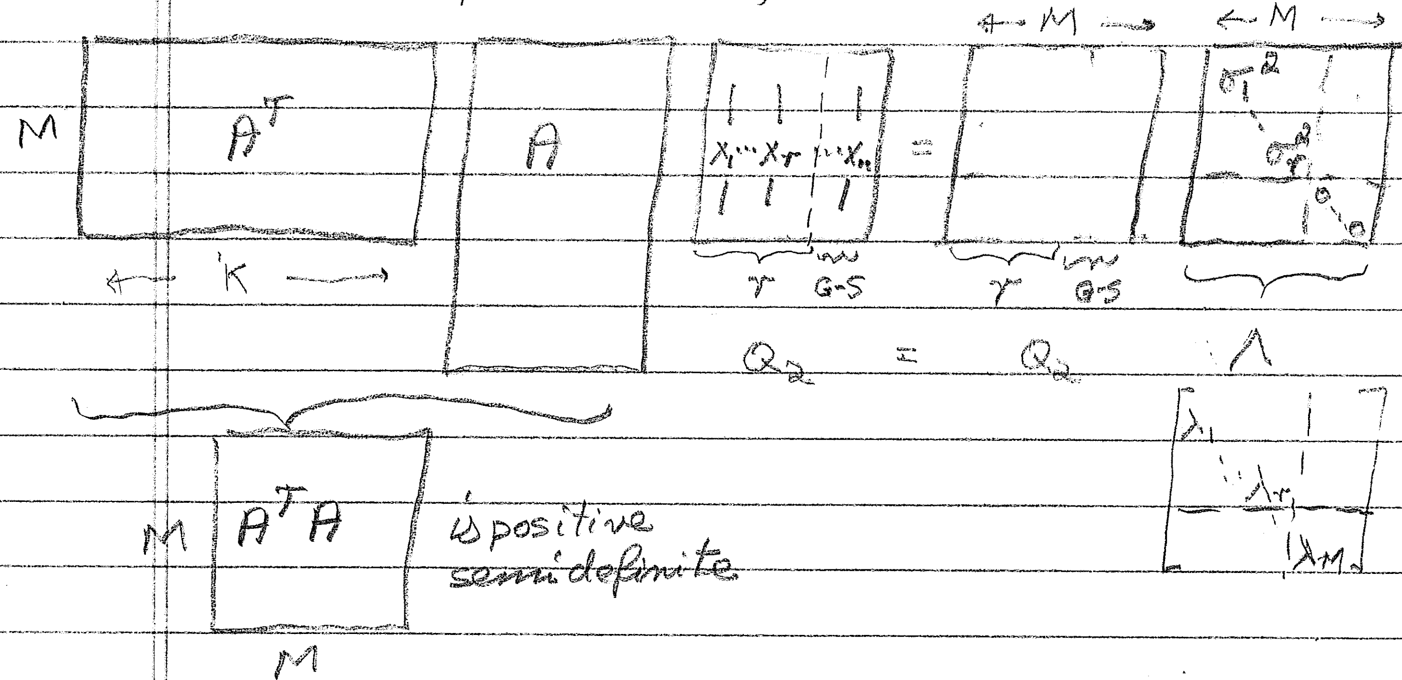
$$A \;=\; Q_1 \quad \Sigma \quad Q_2^T$$



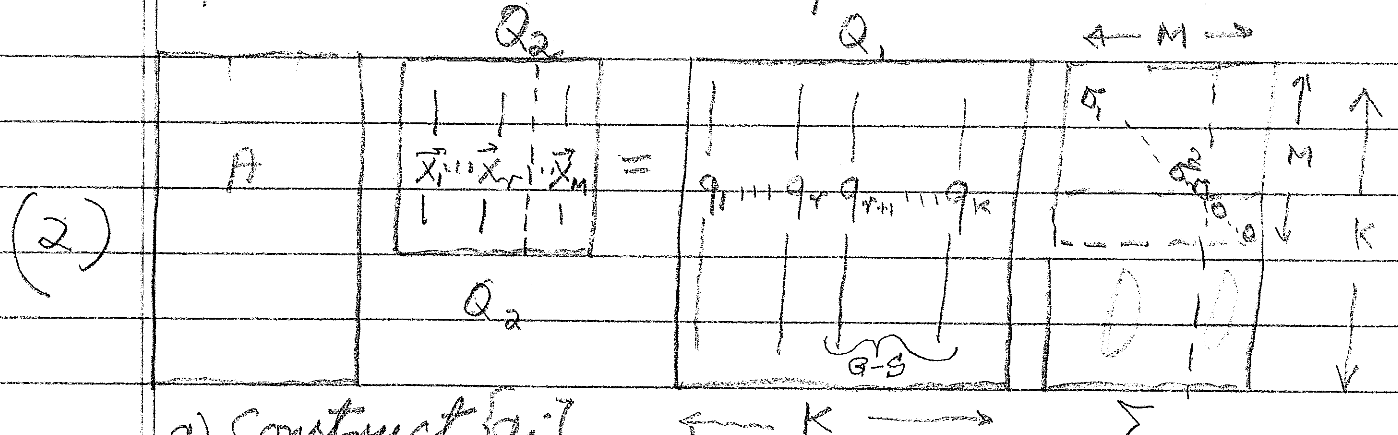
$$Q_1^T \cdot A \quad Q_2 = \Sigma$$

Step I: Diagonalize $A^T A$ via Q_2

$A^T A$ is positive semi definite.



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



a) Construct $\{q_j\}$

From $A^T A x_j = \lambda_j x_j \equiv \sigma_j^2 x_j, \; j = 1, \ldots, r$

and $A : R^M \to R^K$

obtain $q_j = \dfrac{A x_j}{\sigma_j}$ for $j = 1, \ldots, r$

They satisfy

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

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

$$Q_1 = \begin{bmatrix} | & & | & | & & | \\ q_1 & \cdots & q_r & q_{r+1} & \cdots & q_k \\ | & & | & | & & | \end{bmatrix} : R^k \longrightarrow R^k$$

$$\underbrace{q_1 \cdots q_r}_{\text{from } A^T A} \qquad \underbrace{q_{r+1} \cdots q_k}_{\text{via G-S}}$$

Step III

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

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