Solving Linear Least Squares with SVD

What we have learned so far…

$\boldsymbol{A}$ is a $m \times n$ matrix where $m > n$
(more points to fit than coefficient to be determined)

Normal Equations: $\boldsymbol{A}^T \boldsymbol{A}\, \boldsymbol{x} = \boldsymbol{A}^T\, \boldsymbol{b}$

- The solution $\boldsymbol{A}\, \boldsymbol{x} \cong \boldsymbol{b}$ is unique if and only if $rank(\mathbf{A}) = n$
($\boldsymbol{A}$ is full column rank)

- $rank(\mathbf{A}) = n$ → columns of $\boldsymbol{A}$ are ***linearly independent*** → n non-zero singular values → $\boldsymbol{A}^T\ \boldsymbol{A}$ has only positive eigenvalues → $\boldsymbol{A}^T \boldsymbol{A}$ is a symmetric and positive definite matrix → $\boldsymbol{A}^T \boldsymbol{A}$ is invertible

$$\boldsymbol{x} = (\boldsymbol{A}^T \boldsymbol{A})^{-1} \boldsymbol{A}^T\, \boldsymbol{b}$$

- If $rank(\mathbf{A}) < n$, then $\boldsymbol{A}$ is rank-deficient, and solution of linear least squares problem is ***not unique***.

Condition number for Normal Equations

Finding the least square solution of $\boldsymbol{A}\,\boldsymbol{x} \cong \boldsymbol{b}$ (where $\boldsymbol{A}$ is full rank matrix) using the Normal Equations

$$\boldsymbol{A}^T\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{A}^T\,\boldsymbol{b}$$

has some advantages, since we are solving a square system of linear equations with a symmetric matrix (and hence it is possible to use decompositions such as Cholesky Factorization)

However, the normal equations tend to worsen the conditioning of the matrix.

$$cond(\boldsymbol{A}^T\boldsymbol{A}) = (cond(\boldsymbol{A}))^2$$

How can we solve the least square problem without squaring the condition of the matrix?

SVD to solve linear least squares problems

$\boldsymbol{A}$ is a $m \times n$ rectangular matrix where $m > n$, and hence the SVD decomposition is given by:

$$\boldsymbol{A} = \begin{pmatrix} \vdots & \cdots & \vdots \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & 0 \\ & & \vdots \\ & & 0 \end{pmatrix} \begin{pmatrix} \cdots & \mathbf{v}_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \mathbf{v}_n^T & \cdots \end{pmatrix}$$

We want to find the least square solution of $\boldsymbol{A}\,\boldsymbol{x} \cong \boldsymbol{b}$, where $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}}$

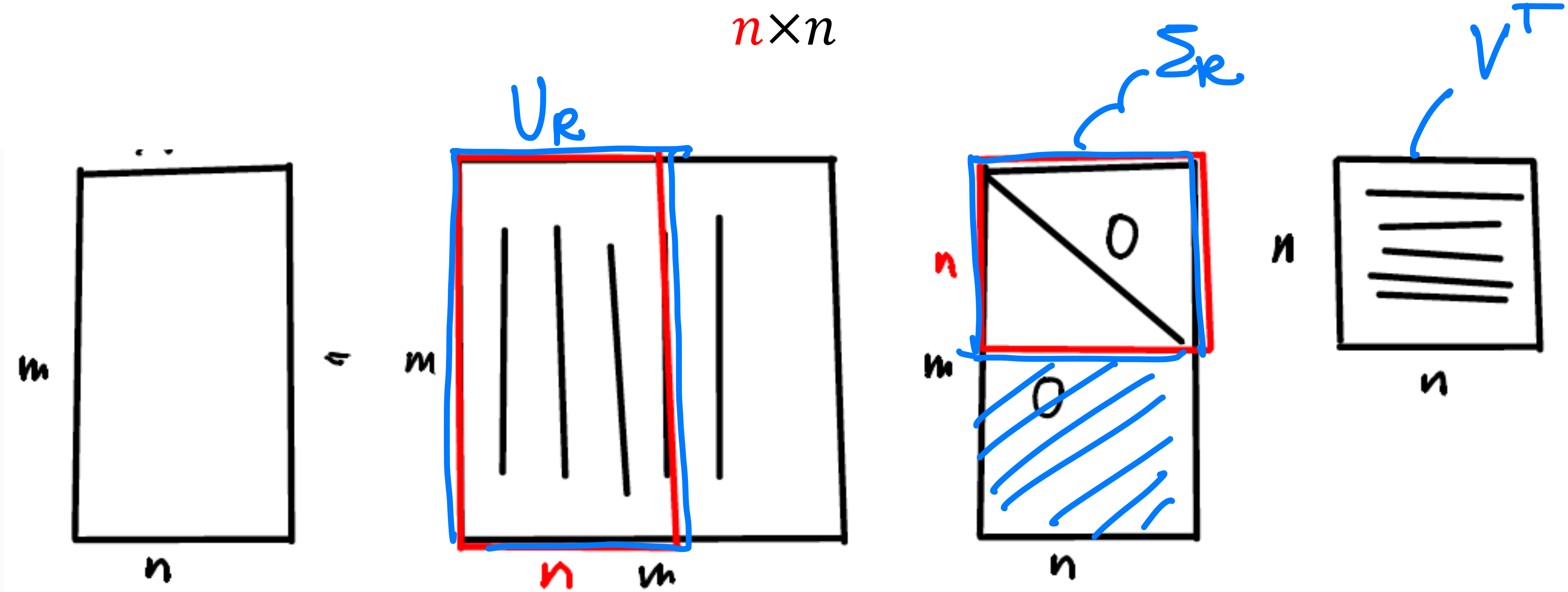
or better expressed in reduced form: $\boldsymbol{A} = \boldsymbol{U}_R\ \boldsymbol{\Sigma}_{\boldsymbol{R}}\ \boldsymbol{V}^{\boldsymbol{T}}$

Recall Reduced SVD

$m > n$

$$\boldsymbol{A} = \boldsymbol{U}_R \, \boldsymbol{\Sigma}_R \, \boldsymbol{V}^T$$

A: $m \times n$, U_R: $m \times n$, Σ_R: $n \times n$, V^T: $n \times n$



Shapes of the Reduced SVD

Suppose you compute a reduced SVD $A = U\Sigma V^T$ of a 10×14 matrix A. What will the shapes of U, Σ, and V be?

Hint: Remember the transpose on V!

The shape of U will be ____ $\times$ ____ .

The shape of Σ will be ____ $\times$ ____ .

The shape of V will be ____ $\times$ ____ .

SVD to solve linear least squares problems

$$\boldsymbol{A} = \boldsymbol{U_R}\ \boldsymbol{\Sigma_R}\ \boldsymbol{V}^T$$

$$\boldsymbol{A} = \begin{pmatrix} \vdots & \cdots & \vdots \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \begin{pmatrix} \cdots & \mathbf{v}_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \mathbf{v}_n^T & \cdots \end{pmatrix}$$

$$A x = b \longrightarrow A^T A x = A^T b$$

$$(U_R \Sigma_R V^T)^T (U_R \Sigma_R V^T) x = (U_R \Sigma_R V^T)^T b$$

$$(V^T)^T \Sigma_R^T \underbrace{U_R^T U_R} \Sigma_R V^T x = (V^T)^T \Sigma_R^T U_R^T b$$

$$\Sigma_R^{-1} = ?$$

$$V \Sigma_R^T \Sigma_R V^T x = V \Sigma_R^T U_R^T b$$

$$V \Sigma_R^2 V^T x = V \Sigma_R U_R^T b \Longrightarrow \boxed{\Sigma_R^2 V^T x = \Sigma_R U_R^T b}$$

① Full rank A $A_{m\times n}$: $\text{rank}(A) = n$

$$\Sigma_R^2 V^T x = \Sigma_R U_R^T b \implies V^T x = \Sigma_R^{-1} U_R^T b$$

$$\boxed{x = V \Sigma_R^{-1} U_R^T b}$$

x: unique solution, $n\times 1$; V: $n\times n$; Σ_R^{-1}: $n\times n$; U_R^T: $n\times m$; b: $m\times 1$ ✓

② Rank deficient $A_{m\times n}$: $\text{rank}(A) = r < n$

$\Sigma_R^2 V^T x = \Sigma_R U_R^T b$ ⟶ solution is not unique

① Find x s.t $\min_x \|Ax - b\|_2^2$ ✓

+ $\min_x \|x\|_2$ ✓

$\Sigma_R^2 V^T x = \Sigma_R U_R^T b$

rank(A) = r < n

Change of variable $y = V^T x \rightarrow x = Vy$

Let's solve $\Sigma_R \, y = U_R^T b$

$U_R = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}$ m×n

$$\begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} u_1^T b \\ u_2^T b \\ \vdots \\ u_r^T b \\ \vdots \\ u_n^T b \end{bmatrix}$$

Σ_R n×n $\quad \underline{y} \quad U_R^T b$

$y_1 \sigma_1 = u_1^T b = u_1 \cdot b$

$y_1 = u_1^T b / \sigma_1$

$y_2 = u_2^T b / \sigma_2$

$\vdots$

$y_r = u_r^T b / \sigma_r$

$y_i = \dfrac{u_i^T b}{\sigma_i} \quad i = 1, \ldots, r$

$y_i = ~~0~~ \quad i = r+1, \ldots, n$?

$y_{r+1} = ?$

anything!! → solution not unique

min ‖x‖

min ‖Vy‖₂

set it to zero!

In summary:

$$y_i = \begin{cases} \dfrac{u_i^T b}{\sigma_i}, & \text{if } i = 1, \ldots, r \\ 0, & \text{if } i = r+1, \ldots, n \end{cases} \quad \Big\} \ \underline{y}$$

Compute $\underset{\sim}{x} \longrightarrow \underset{\sim}{x} = \underline{\underline{V}}\, \underline{y} = \sum_{i=1}^{n} y_i \underset{\sim}{v_i}$

$$\underset{\sim}{x} = \sum_{i=1}^{n} \left(\frac{u_i^T b}{\sigma_i} \right) \underset{\sim}{v_i}$$

$\sigma_i \neq 0$

$\underline{\underline{A}}\,\underline{x} = \underline{b}$

$\underline{\underline{A}}$ is rank deficient

$$\underset{\sim}{x} = \sum_{i=1}^{n} \left(\frac{u_i^T b}{\sigma_i}\right) \underset{\sim}{v_i}$$

$\sigma_i \neq 0$

$U_R^T b$ — U_R^T: $n \times m$, b: $m \times 1$ → $O(mn)$

$\Sigma_R^T z$ — Σ_R^T: $n \times n$, z: $n \times 1$ — Σ is diag → $O(n)$

$$\Sigma_R^2 V^T x = \Sigma_R U_R^T b \rightarrow \Sigma_R^+$$

$$V^T x = \Sigma_R^+ U_R^T b$$

$\Sigma_R^+ = \Sigma_R^{-1}$ full rank

$$\underset{\sim}{x} = \underline{\underline{V}}\ \underline{\underline{\Sigma_R^+}}\ \underline{\underline{U_R^T}}\ \underline{b}$$

($z = U_R^T b$, $y = \Sigma_R^+ z$)

$V y$ — V: $n \times n$, y: $n \times 1$ → $O(n^2)$

SVD → $O(mn^2)$

$m \geq n$ | overall $O(mn)$

Example:

Consider solving the least squares problem $\boldsymbol{A\,x} \cong \boldsymbol{b}$, where the singular value decomposition of the matrix $\boldsymbol{A} = \boldsymbol{U\,\Sigma\,V^T x}$ is:

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x} \cong \begin{bmatrix} 12 \\ 9 \\ 9 \\ 10 \end{bmatrix}$$

Determine $\|\boldsymbol{b} - \boldsymbol{A\,x}\|_2$

Example

Suppose you have $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}}\boldsymbol{x}$ calculated. What is the cost of solving

$$\min_{\boldsymbol{x}} \|\boldsymbol{b} - \boldsymbol{A}\,\boldsymbol{x}\|_2^2 \ ?$$

A) $O(n)$
B) $O(n^2)$
C) $O(mn)$
D) $O(m)$
E) $O(m^2)$