Matrix Norm and Condition Number

- **Norm of a vector x = vector length (size) = $||x||$**
- **How about norm of a <u>*m* by *n*</u> matrix A (<u>any size</u>): $||A||$?**
- **Definition:** $||A|| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- **In other words, $||A||$ is the maximum "amplifying power" of the transformation by A: $||Ax||/||x|| \leq ||A||$ (i.e. $||Ax||/||A|| \leq ||x||$).**
- **Recall:** $A = \begin{bmatrix} 1. & 1. \\ 1. & 1.0001 \end{bmatrix}$ **and** $A' = \begin{bmatrix} .0001. & 1. \\ 1. & 1. \end{bmatrix}$**?**
- **Error equation: $A(x+\delta x)=b+\delta b \rightarrow A(\delta x)=\delta b \rightarrow$ an error δb leads to an error in solution $\delta x=A^{-1}(\delta b)$. Solution is unstable when A^{-1} is large in nature, i.e., A is nearly singular, or when points in the direction that is amplified most by A^{-1}.**
- **Given the norm of $||A^{-1}||$: $||\delta x||/||\delta b||=|| A^{-1}(\delta b) ||/||\delta b|| \leq ||A^{-1}||$**

  $\Rightarrow ||\delta x|| \leq ||A^{-1}||\ ||\delta b|| \Rightarrow ||\delta x||/||x|| \leq ||A^{-1}||||\delta b||/||x|| \leq ||A^{-1}||\ ||\delta b||/(||Ax||/||A||)$

  $\Rightarrow ||\delta x||/||x|| \leq (||A^{-1}||||A||)||\delta b||/||b|| \Rightarrow$ Define condition number $c= ||A||||A^{-1}||$
- **$||\delta x||/||x|| \leq c||\delta b||/||b||$: relative error never exceeds $c\times$relative changes in b**
- **Formula for $||A||$: $||A||^2=\lambda_{\max}$ of A^TA (at least semidefinite!)**

  $\max \frac{\|Ax\|^2}{\|x\|^2} = \max \frac{x^T A^T Ax}{x^T x} = \lambda_{\max}$ **is maximized by the corresponding eigenvector of A^TA. (Recall Rayleigh's quotient) $\Rightarrow ||A||=\sqrt{\lambda_{\max}}$**
- **Formula for condition number: $c^2=\lambda_{\max}/\lambda_{\min}$ ($c = \sqrt{\lambda_{\max} / \lambda_{\min}}$)**

Remaining Questions

- **Recall: to solve $Ax=b$ with no solution $\Rightarrow A^TA\hat{x}=A^Tb$**

  Question: what if columns in A are not independent and A^TA is not invertible? Solution is not unique!
- **Recall: Square matrix diagonalization $S^{-1}AS=\Lambda \Rightarrow$**

  Symmetric: $Q^TAQ=\Lambda$; Schur's lemma: $U^{-1}AU=T$;

  Jordan Form: $M^{-1}AM=J$

  Question: What if A is rectangular?
- **Answer: Singular Value Decomposition (SVD)**
- **Applications**
  - **Too many…..just list a few**
  - **Numerical computation to find eigenvalues and eigenvectors for symmetric matrices**
  - **Numerical computation to find bases for four subspaces**
  - **Data compression for image processing**
  - **Canonical correlation (many-to-many correlation)**
  - **Polar Decomposition for Robotics and Plastic Surgery**
  - **Optimal solution of $Ax=b$**

Basics of SVD: A^TA and AA^T

- **For any A ($m\times n$), A^TA ($n\times n$) is symmetric positive semidefinite and is invertible positive definite when A has independent columns**

- **For any A ($m\times n$), AA^T is symmetric positive semidefinite and is invertible positive definite when A has independent rows**

- **A^TA has the same nullspace as A ($Ax=0 \Leftrightarrow A^TAx=0$); AA^T has the same nullspace as A^T (i.e., the same left-null space of A: $y^TA=0 \Leftrightarrow y^TAA^T=0$)**

- **A^TA and AA^T share the same eigenvalues; if x is the eigenvector of A^TA, the Ax is the eigenvector of AA^T corresponding to the same eigenvalue.**

Proof: $A^TAx=\lambda x \Rightarrow AA^TAx=A\lambda x=\lambda(Ax)$

Diagonalization of A^TA and AA^T

- **Diagonalization: $Q_1^T AA^T Q_1 = \Lambda_1$ and $Q_2^T A^T A Q_2 = \Lambda_2$ where**

$$\Lambda_1 = \begin{array}{cc} & m \\ m & \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & \end{bmatrix} \end{array} \quad \text{and} \quad \Lambda_2 = \begin{array}{cc} & n \\ n & \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & \end{bmatrix} \end{array}$$

with r (=*rank* of $A \le \min(n, m)$) none-zero eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$

Let v_i be the *unit* eigenvector of A^TA corresponding to the nonzero eigenvalues λ_i, then Av_i is the eigenvector of AA^T corresponding to the same eigenvalue.

Then,

$$\|Av_i\|^2 = (Av_i)^T Av_i = v_i^T A^T Av_i = v_i^T \lambda_i v_i = \lambda_i \|v_i\|^2 = \lambda_i \quad i = 1,\ldots,r$$

$$\Rightarrow \|Av_i\| = \sqrt{\lambda_i} \qquad i = 1,\ldots,r$$

The *unit* eigenvector of AA^T : $u_i = \dfrac{Av_i}{\|Av_i\|} = Av_i / \sqrt{\lambda_i} \qquad i = 1,\ldots,r$

$$\Rightarrow Av_i = \sqrt{\lambda_i}\, u_i \quad \text{for } i=1,\ldots,r; \quad Av_i = 0 = \sqrt{\lambda_i}\, u_i \quad \text{for } i > r$$

$$\Rightarrow \underset{m\times n}{A}\ \underset{n\times n}{Q_2} = \underset{m\times m}{Q_1}\ \underset{m\times n}{\Sigma}$$

where $$\Sigma = \begin{array}{cc} & n \\ m & \begin{bmatrix} \sqrt{\lambda_1} & & & & \\ & \ddots & & & \\ & & \sqrt{\lambda_r} & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} \end{array}$$

Singular Value Decomposition

Any *m* by *n* matrix *A* can be factored into

$$A = Q_1 \Sigma Q_2^T = (orthogonal)(diagonal)(orthogonal)$$

where $$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix} = \begin{array}{cc} & n \\ m & \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & \sqrt{\lambda_r} & \\ & & & \end{bmatrix} \end{array}$$

and <u>$\sigma_1,\ldots,\sigma_r$ are positive</u> and called singular values and are square roots of eigenvalues of AA^T and A^TA.

Proof:

(⇒) $AA^T = (Q_1 \Sigma Q_2^T)(Q_2 \Sigma^T Q_1^T) = Q_1 \Sigma \Sigma^T Q_1^T$ **and similarly**

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

(⇐) Since $Q_2^T A^T A Q_2 = \Lambda_2$; $Q_1^T A A^T Q_1 = \Lambda_1$ **and**

$$\underset{m \times n}{A} \; \underset{n \times n}{Q_2} = \underset{m \times m}{Q_1} \; \underset{m \times n}{\Sigma} \Rightarrow Q_1^T A Q_2 = Q_1^T Q_1 \Sigma = \Sigma$$

- **Positive (semi)definite matrices:** $A = Q_1 \Sigma Q_2^T = Q \Lambda Q^T$
- **Indefinite metrics: where** $\Sigma = |\Lambda|$
- **For complex matrices, Σ remains real, Q_1 and Q_2 are Unitary**
- **Reduced form of SVD: Compact SVD**

$A = Q_1^r \sum^r Q_2^{rT}$ **where** $$\Sigma^r = \begin{array}{cc} & r \\ r & \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_r} \end{bmatrix} \end{array}$$ **and only first *r* columns of Q_1**

and Q_2 corresponding to nonzero singular values are kept in Q_1^r **and** Q_2^r

SVD and Four Fundamental Subspaces of *A*

- **Bases for four fundamental subspaces:**
  - **The first *r* columns of Q_1 (the first *r* eigenvectors of AA^T): column space of *A***
  - **The rest $m-r$ columns of Q_1 (the $m-r$ eigenvectors corresponding to eigenvalue 0 of AA^T): left nullspace of *A***
  - **The first *r* columns of Q_2 (the first *r* eigenvectors of A^TA): row space of *A***
  - **The rest $n-r$ columns of Q_2 (the $n-r$ eigenvectors corresponding to eigenvalue 0 of A^TA): nullspace of *A***

Proof:

$AQ_2 = Q_1\Sigma \Rightarrow$

$$\begin{bmatrix} | & | & & | & | & & | \\ Av_1 & Av_2 & \cdots & Av_r & Av_{r+1} & \cdots & Av_n \\ | & | & & | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | & | & & | \\ \sigma_1 u_1 & \sigma_2 u_2 & \cdots & \sigma_r u_r & 0 & \cdots & 0 \\ | & | & & | & | & & | \end{bmatrix}$$

⇒ **$v_{r+1},\ldots, v_n$ are vectors in null space of *A***

⇒ **$u_1,\ldots, u_r$ are vectors in column space of *A***

And $A = Q_1 \sum Q_2^T \Rightarrow A^T Q_1 = Q_2 \Sigma$

Examples of SVD

- **Example 1 (*A* is diagonal)**

$$\begin{bmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- **Example 2 (*A* has only one column)**

$$A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1]$$

$A^T A$ (1 by 1) and AA^T (3 by 3) both have eigenvalues 9 (always work on the smallest one: in this case the 1-by-1 $A^T A$ is the smallest)

- **Example 3 (*A* is already orthogonal)**

Either $A=QII$ or $A=IIQ$ or even $A=(QQ_2)IQ_2^T$ (for any orthogonal matrix Q_2) but certainly $\Sigma=I$

- **Example 4 (*A* is an incidence matrix)**

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}_{/\sqrt{2} \quad /\sqrt{2}} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{matrix} /\sqrt{6} \\ /\sqrt{2} \\ /\sqrt{3} \end{matrix}$$

where $AA^T = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ **with $\lambda=3, 1$**

More on SVD

- **Transformation by SVD:** $Ax = Q_1 \Sigma (Q_2^T x)$**:**
  - $Q_2^T x = [v_1^T x, \quad v_2^T x, \quad \cdots \quad v_n^T x]^T$ **: Expressing x in R^n as the linear combination of the orthogonal basis of row space and null space (v's)**
  - $\Sigma(Q_2^T x) = [\sigma_1 v_1^T x, \quad \sigma_2 v_2^T x, \quad \cdots \quad \sigma_r v_r^T x]^T$ **: Only the row space part is transformed by multiplying singular values ($\sigma_1, \ldots, \sigma_r$)**
  - $Q_1 \Sigma (Q_2^T x) = u_1 \sigma_1 (v_1^T x) + u_2 \sigma_2 (v_2^T x) + \cdots + u_r \sigma_r (v_r^T x)$ **: transformed result is expressed as the linear combination of orthogonal basis of the column space (u's)**
- **PCA and SVD:** $AA^T = Q_1 \Sigma \Sigma^T Q_1^T$ **and** $A^T A = Q_2 \Sigma^T \Sigma Q_2^T$
  - **v's are coefficients of linear combination of A columns (measurements) to maximize the sum of squares among rows (persons) of A and the $Av_i=\sigma_i u_i$ is the ith PC of $A^T A$**
  - **u's are coefficients of linear combination of A rows to maximize the sum of squares among columns of A and the $A^T u_i=\sigma_i v_i$ is the ith PC of AA^T**
  - $\sigma_1^2 \ge \sigma_2^2 \ge \cdots \ge \sigma_r^2$ **are variances of the linear combination of columns (rows) from the largest to the smallest.**

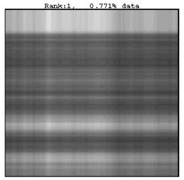
Application: Image Processing

- **Satellite takes a picture containing 1000 by 1000 pixels represented by matrix A; each pixels with a color number $\Rightarrow$ to send 1,000,000 numbers**
- **SVD of Image A:**

$$Q_1 \Sigma Q_2^T = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_r\sigma_r v_r^T$$

where u's are columns of Q_1 and v's are columns of Q_2.

- **We may keep only the first few terms with larger singular values, say 60 of them:**

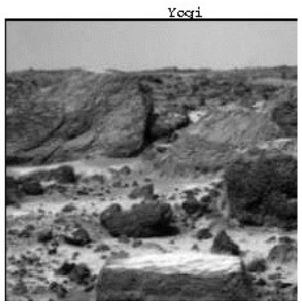


$$\begin{aligned} Q_1 \sum Q_2^T &= u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_{60}\sigma_{60} v_{60}^T \\ &= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_{60} u_{60} v_{60}^T \end{aligned}$$

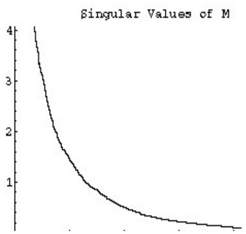
i.e. sum of $u_i v_i^T$ with weights σ_i ($u_i v_i^T$? recall rank-one matrix)

- **Throw away 950 terms and only send back 60+60•(1000+1000)**
- **The image gets more and more lucid as more singular-value terms are added.**
- **Example: Martian image of a rock called "Yogi" by the Sojourner rover: 256*264 (rank 256)**

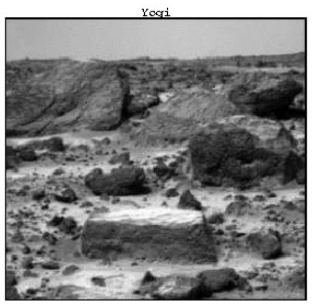
Original



Singular Value



Rank: 81 (62.4%)



Polar Decomposition for Robotics and Plastic Surgery

- **Robot arm: rotate and stretch-out/draw-back**
- **Plastic Surgery: Rotation and Stretching/Compression of a certain part of your body**
- **Material deformation expressed by a matrix A**
- **Every real square matrix can be factored into**

$$A = QS$$

where Q is *orthogonal* and S is *symmetric positive (semi)definite*

Proof:

$$A = Q_1 \Sigma Q_2^T = (Q_1 Q_2^T)(Q_2 \Sigma Q_2^T) \Rightarrow \; Q = Q_1 Q_2^T \,; S = Q_2 \Sigma Q_2^T$$

- **Example of $A=QS$**

$$\begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$$

- **Example of $A=S'Q$** $\quad A = Q_1 \Sigma Q_2^T = (Q_1 \Sigma Q_1^T)(Q_1 Q_2^T)$

$$\begin{bmatrix} 1 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

- **Orthogonal Q: rotation or reflection**
- **Symmetric (semi)definite $S (= Q_2 \Sigma Q_2^T)$: eigenvalues $\sigma_1, \ldots, \sigma_r \Rightarrow$ stretch/compress through directions of columns in Q_2**

Optimal Solution of $Ax=b$

- $Ax=b$

  (1) **Rows of A are dependent $\Rightarrow$ very likely no solution (b is not in the column space of A) $\Rightarrow A\hat{x}=p \Rightarrow A^T A\hat{x}=A^T b$**

  (2) **Columns of A are dependent? $A^{\mathrm{T}}A$ not invertible with null space $\Rightarrow$ No unique Solution!**

- **Optimal solution of $Ax=b$:**

  Solution of $A\hat{x}=p$ with minimum length

Example 1: $A=\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

- **Rows are dependent $\Rightarrow$ project b onto the column space**

$$A\hat{x}=p \text{ is } \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix}=\begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \Rightarrow \hat{x}_1=b_1/\sigma_1;\ \hat{x}_2=b_2/\sigma_2$$

- **Columns are dependent: $\hat{x}_3$ and $\hat{x}_4$ can be randomly chosen**

  Optimal solution: $\hat{x}_3=\hat{x}_4=0$ such the length of $\hat{x}$ is minimum

- **Optimal solution x^+: the minimum-length solution of $A\hat{x}=p$:**

$$x^+=\begin{bmatrix} \frac{b_1}{\sigma_1} \\ \frac{b_2}{\sigma_2} \\ 0 \\ 0 \end{bmatrix}=\begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Optimal Solution x^+ and Pseudo-inverse A^+

- **Recall:**

$$x^+=\begin{bmatrix} \frac{b_1}{\sigma_1} \\ \frac{b_2}{\sigma_2} \\ 0 \\ 0 \end{bmatrix}=\begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Optimal solution $x^+=A^+b \Rightarrow A^+$ is called pseudo-inverse of A; i.e. If

$$A=\begin{array}{c} \\ m \end{array}\begin{array}{c} n \\ \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix} \end{array} \quad \text{then} \quad A^+=\begin{array}{c} \\ n \end{array}\begin{array}{c} m \\ \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & \end{bmatrix} \end{array}$$

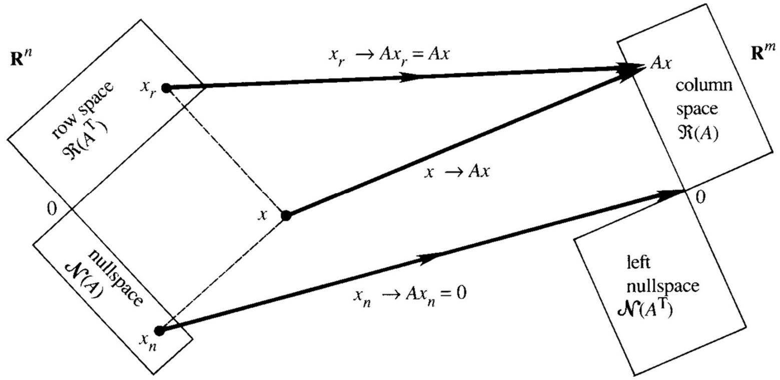
$$\text{and} \quad x^+=A^+b=\begin{bmatrix} b_1/\sigma_1 \\ \vdots \\ b_r/\sigma_r \end{bmatrix}$$

- $(A^+)^+=A$

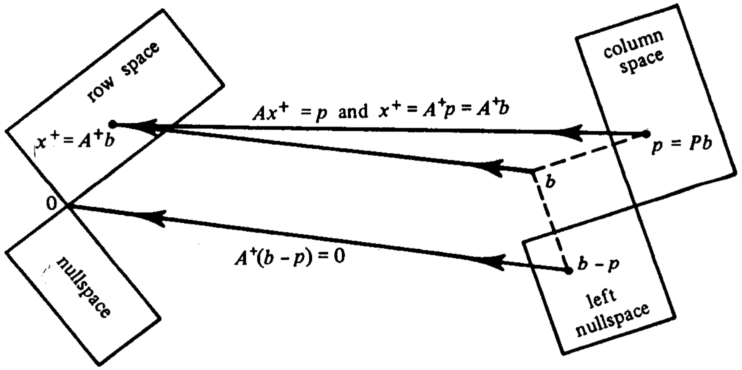
- **A is invertible $\Rightarrow A^{-1}=A^+$**

Row Space component of $\hat{x}$ and x^+

- **Least square solution $A\hat{x}=p \Rightarrow$ split $\hat{x}$ into row space component and null space component: $\hat{x}=\hat{x}_r+\hat{x}_n$**
  - **$A\hat{x}_r=p$ since $A\hat{x}_n=0$**
  - **$\|\hat{x}\|^2=\|\hat{x}_r\|^2+\|\hat{x}_n\|^2$, so $\hat{x}$ is shortest when $\hat{x}_n=0$**
  - **all $\hat{x}$ has the same row space component $\hat{x}_r$ and $x^+=\hat{x}_r$**
- **Recall**



- **Now, reverse direction:**



Finding A^+

Example:

$Ax=b$ is $-x_1+2x_2+2x_3=18$

Solutions are on the whole plane $-x_1+2x_2+2x_3=18$

A=[-1 2 2] $\Rightarrow x^+$=[-2 4 4]=$\hat{x}_r$

Any other solution $\hat{x}=\hat{x}_r+\hat{x}_n$ with $\hat{x}_n \neq 0$, e.g. [-2 5 3], [-2 7 1], [-6 3 3] are longer than [-2 4 4]

$$A^+=\begin{bmatrix}-1 & 2 & 2\end{bmatrix}^+=\begin{bmatrix}-\frac{1}{9}\\ \frac{2}{9}\\ \frac{2}{9}\end{bmatrix} \quad \text{and} \quad A^+\begin{bmatrix}18\end{bmatrix}=\begin{bmatrix}-2\\4\\4\end{bmatrix}$$

- **Formula of A^+: $A=Q_1\Sigma Q_2^T \Rightarrow A^+=Q_2\Sigma^+Q_1^T$ where**

$$\Sigma^+=\begin{array}{c c} & m \\ n & \begin{bmatrix}1/\sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & 1/\sqrt{\lambda_r} & \\ & & & \end{bmatrix}\end{array}=\begin{array}{c c} & m \\ n & \begin{bmatrix}1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & \end{bmatrix}\end{array}$$

$$\text{(Recall: } \Sigma=\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}=\begin{array}{c c} & n \\ m & \begin{bmatrix}\sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & \sqrt{\lambda_r} & \\ & & & \end{bmatrix}\end{array}\text{)}$$

Examples of A^+ Formula

- **Example 1:** $A=\begin{bmatrix}\sigma_1 & 0 & 0 & 0\\ 0 & \sigma_2 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$

  A **is diagonal already** $\Rightarrow Q_1=I_{3\times 3}$; $Q_2=I_{4\times 4}$

$$A^+=\begin{bmatrix}\frac{1}{\sigma_1} & 0 & 0\\ 0 & \frac{1}{\sigma_2} & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$

- **Example 2:** $A=[-1\ 2\ 2]$

  $Q_1=[1]$ **with singular value=3:**

$$\begin{bmatrix}-1 & 2 & 2\end{bmatrix}=\begin{bmatrix}1\end{bmatrix}\begin{bmatrix}3 & 0 & 0\end{bmatrix}\begin{bmatrix}-\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\end{bmatrix}$$

$$\begin{bmatrix}-1 & 2 & 2\end{bmatrix}^+=\begin{bmatrix}-\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\end{bmatrix}\begin{bmatrix}\frac{1}{3}\\ 0\\ 0\end{bmatrix}\begin{bmatrix}1\end{bmatrix}=\begin{bmatrix}-\frac{1}{9}\\ \frac{2}{9}\\ \frac{2}{9}\end{bmatrix}$$

Proof of A^+ Formula

- $A=Q_1\Sigma Q_2^T \Rightarrow A^+=Q_2\Sigma^+ Q_1^T$

Proof:

$$\|Ax-b\|=\|Q_1\Sigma Q_2^T x-b\|=\|Q_1^T(Q_1\Sigma Q_2^T x-b)\|=\|\Sigma Q_2^T x-Q_1^T b\|.$$

Let $y=Q_2^T x=Q_2^{-1}x$ **and** $\|y\|=\|x\|$

$\Rightarrow$ **min** $\|Ax-b\| \equiv$ **min** $\|\Sigma y-Q_1^T b\|$

$\Rightarrow$ **solving the optimal solution for** $\Sigma y=Q_1^T b$

Σ is a diagonal matrix

(like the one in Example 1: $A=\begin{bmatrix}\sigma_1 & 0 & 0 & 0\\ 0 & \sigma_2 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\Rightarrow A^+=\begin{bmatrix}\frac{1}{\sigma_1} & 0 & 0\\ 0 & \frac{1}{\sigma_2} & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$**)**

$\Rightarrow\ y^+=\Sigma^+ Q_1^T b$

$\Rightarrow\ x^+=Q_2 y^+=Q_2\Sigma^+ Q_1^T b$