15.7. $A \in \mathcal{M}_{m,n}(\mathbb{R})$

setting $u = [A]^1$ (1st column of A), (wlog $[A]^1 \neq 0$) $\exists$ $rk(A) = 1$,

$[A]^j (j > 1) = [A]^1 \cdot \dfrac{a_{kj}}{a_{k1}}$. ($a_{k1}$: kth comp of $[A]^1$, which is not zero).

setting $v = \left[1 \quad \dfrac{a_{k2}}{a_{k1}} \quad \cdots \quad \dfrac{a_{kn}}{a_{k1}} \right]$, $A = u^t v$ ■

15.9. $\exists$ column vector, $A \in \mathcal{M}_{m,1}(\mathbb{R})$

$\hookrightarrow$ $[u_1 \cdots u_m] \cdot \begin{bmatrix} \sqrt{A^tA} \\ 0 \end{bmatrix} \cdot [1]$, with

$u_1 = \dfrac{A}{\|A\|_2}$, $\{u_1, \dots u_m\}$: ONB of $\mathbb{R}^m$.

$\exists$ row vector, $A \in \mathcal{M}_{1,n}(\mathbb{R})$

$\hookrightarrow$ $[1] \cdot [\sqrt{AA^t} \mid 0] \cdot [v_1 \cdots v_n]^t$. with

$v_1 = \dfrac{A}{\|A\|_2}$, $\{v_1, \dots v_n\}$: ONB of $\mathbb{R}^n$ ■

15.13.

$Ax = 0 \rightarrow A^tAx = 0 \leadsto \ker(A) \subseteq \ker(A^tA)$

$A^tAx = 0 \rightarrow x^tA^tAx = 0 \leadsto \|Ax\|_2^2 = 0 \rightarrow Ax = 0 \leadsto \ker(A^tA) \subset \ker A$

$\therefore \ker A = \ker A^tA$

■

15.14 $A \in M_{n,n}(\mathbb{R})$. $A = U\tilde{\Sigma}V^t$

$\leadsto \|A^2\|_2 \leadsto A^t \cdot A^t \cdot A \cdot A = V\tilde{\Sigma}U^t V\tilde{\Sigma}U^t U\tilde{\Sigma}V^t U\tilde{\Sigma}V^t$

$= V\tilde{\Sigma}U^t V\tilde{\Sigma}^2 V^t U\tilde{\Sigma}V^t := B$

$\|\tilde{\Sigma}V^t U\tilde{\Sigma}\|_2 \leadsto \tilde{\Sigma}U^t V\tilde{\Sigma} \cdot \tilde{\Sigma}V^t U\tilde{\Sigma}$

$= \tilde{\Sigma}U^t V\tilde{\Sigma}^2 V^t U\tilde{\Sigma} := A$

$\rightarrow B = VAV^t$.

$\rightarrow \det(\lambda I - A) = \det[V \cdot (\lambda I - A) \cdot V^t] = \det(\lambda \cdot VV^t - VAV^t)$

$= \det(\lambda I - B)$, two matrices has same EV. $\rightarrow$ same

$\therefore$ same 2-norm ■

15.15

(a) $A = U\tilde{\Sigma}V^t \leadsto A^tA = V\tilde{\Sigma}^tU^tU\tilde{\Sigma}V^t = V\tilde{\Sigma}^t\tilde{\Sigma}V^t$

$AA^t = U\tilde{\Sigma}V^tV\tilde{\Sigma}^tU^t = U\tilde{\Sigma}\tilde{\Sigma}^tU^t$

$\hookrightarrow \det(A^tA - I) = \det[V(\tilde{\Sigma}^t\tilde{\Sigma} - \lambda I)V^t]$

$\det(AA^t - I) = \det[U(\tilde{\Sigma}\tilde{\Sigma}^t - \lambda I)U^t]$

$\leadsto$ same non-zero eigvals.

(b) $A^tAV = V\tilde{\Sigma}^t\tilde{\Sigma}V^t \cdot V = V \cdot \tilde{\Sigma}^t\tilde{\Sigma} = [\sigma_1^2 v_1, \cdots \sigma_n^2 v_n]$

$AA^tU = U\tilde{\Sigma}\tilde{\Sigma}^tU^t \cdot U = U\tilde{\Sigma} \cdot \tilde{\Sigma}^t = [\sigma_1^2 u_1, \cdots, \sigma_m^2 u_m]$

($\sigma_1^2 \cdots \sigma_r^2$: non-zero diagonal comp. of $\tilde{\Sigma}$,

$\sigma_{r+1}^2, \cdots = 0$)

(c) Find orthonormal set of eigvec of A^tA, AA^t and

sign it as V, U. Setting $\sigma_1, \cdots \sigma_r$: square root of

non-zero eigval of A^tA. $A = U\tilde{\Sigma}V^t$.

(d)

$$A^tA = \begin{bmatrix} 5 & 10 \\ 10 & 25 \end{bmatrix} \leadsto \lambda_1 = 15+10\sqrt{2},\ \lambda_2 = 15-10\sqrt{2}$$

$$\therefore\ \sigma_1 = \sqrt{15+10\sqrt{2}},\quad \sigma_2 = \sqrt{15-10\sqrt{2}}$$

$$\begin{bmatrix} 5 & 10 \\ 10 & 25 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = (15+10\sqrt{2})\cdot\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1+\sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} 5 & 10 \\ 10 & 25 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = (15-10\sqrt{2})\cdot\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1-\sqrt{2} \end{bmatrix}$$

$$\therefore\ U_1 = \frac{1}{\sqrt{4+2\sqrt{2}}}\begin{bmatrix} 1 \\ 1+\sqrt{2} \end{bmatrix},\quad U_2 = \frac{1}{\sqrt{4-2\sqrt{2}}}\begin{bmatrix} 1 \\ 1-\sqrt{2} \end{bmatrix}$$

$$AA^t = \begin{bmatrix} 17 & 14 \\ 14 & 13 \end{bmatrix} \leadsto \lambda_1 = 15+10\sqrt{2},\ \lambda_2 = 15-10\sqrt{2}$$

$$\begin{bmatrix} 17 & 14 \\ 14 & 13 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = (15+10\sqrt{2})\cdot\begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} 17 & 14 \\ 14 & 13 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = (15-10\sqrt{2})\cdot\begin{bmatrix} x \\ y \end{bmatrix}$$

$$\hookrightarrow \begin{bmatrix} 2-10\sqrt{2} & 14 \\ 14 & -2-10\sqrt{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \leadsto \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1+5\sqrt{2} \\ 7 \end{bmatrix}$$

$$\begin{bmatrix} 2+10\sqrt{2} & 14 \\ 14 & -2+10\sqrt{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \leadsto \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1-5\sqrt{2} \\ 7 \end{bmatrix}$$

$$\therefore\ u_1 = \frac{1}{\sqrt{100+10\sqrt{2}}}\begin{bmatrix} 1+5\sqrt{2} \\ 7 \end{bmatrix},\quad u_2 = \frac{1}{\sqrt{100-10\sqrt{2}}}\begin{bmatrix} 1-5\sqrt{2} \\ 7 \end{bmatrix}$$

$$\therefore\ A = [u_1\ u_2]\cdot \operatorname{diag}\left[\sqrt{15+10\sqrt{2}},\ \sqrt{15-10\sqrt{2}}\right]\cdot [U_1, U_2]^t \quad \blacksquare$$

15.16

(a) $x = \sum_i x_i v_i$

so $Ax = [U \tilde{\Sigma} V^t] (\sum_i x_i v_i)$

$$= [U \tilde{\Sigma}] \cdot [x_1, \cdots, x_n]^t = \sum_{i=1}^{n} \sigma_i x_i [U]^i$$

$$\|Ax\|_2^2 = \sum_{i=1}^{n} \sigma_i^2 x_i^2$$

$$\Rightarrow \sigma_n^2 \|x\|_2^2 \leq \sum_{i=1}^{n} \sigma_i^2 x_i^2 \leq \sigma_1^2 \|x\|_2^2 \quad (\because \sigma_n^2 \leq \cdots \leq \sigma_2^2 \leq \sigma_1^2)$$ //

(b) $\frac{\|Ax\|_2}{\|x\|_2} \leq \sigma_1$ (from (a)).

$$\frac{\|Av_1\|_2}{\|v_1\|_2} = \|Av_1\|_2 = \| U \tilde{\Sigma} V^t \cdot v_1 \|_2 = \| U \tilde{\Sigma} e_1 \|_2$$

$$= \| U \cdot \sigma_1 e_1 \|_2 = \| \sigma_1 \cdot [U]^1 \|_2 = \sigma_1$$

(c) $\exists\ A^{-1} = V \Sigma^{-1} U^t = \sum_{j=1}^{n} [V]^j \cdot \frac{1}{\sigma_j} \cdot [U]_j$

$$= \sum_{j=1}^{n} [V]^{n+1-j} \cdot \frac{1}{\sigma_{n+1-j}} \cdot [U]_{n+1-j}.$$

so setting $V' = ([V]^n, \cdots [V]^1)$, $\tilde{\Sigma}' = \mathrm{diag}(\frac{1}{\sigma_n}, \cdots, \frac{1}{\sigma_1})$,

$U' = ([U]^n, \cdots [U]^1)$, $A^{-1} = V' \tilde{\Sigma}' U'^t$ (SVD)

$\frac{\|Ax\|_2}{\|x\|_2} \leq \frac{1}{\sigma_n}$ (from (a)).

$$\frac{\|A^{-1}u_n\|_2}{\|u_n\|_2} = \|A^{-1}u_n\|_2 = \| V \Sigma^{-1} U^t u_n \|_2 = \| V \Sigma^{-1} e_n \|_2$$

$$= \| V \cdot \frac{1}{\sigma_n} e_n \|_2 = \| \frac{1}{\sigma_n} [V]^n \|_2 = \frac{1}{\sigma_n}$$ ■

15.18. 15.20. 15.25 → code problem

15.29.

(a) $A = U\Sigma V^t$, with $A, U, V \in M_{n,n}(\mathbb{R})$

∵ U, V, A is non-singular, $\det A = \det U \cdot \det \Sigma \cdot \det V^t$

→ $\det \Sigma \neq 0$ ⟹ Σ is non-singular

Inverse of Σ is $\Sigma^{-1} = \mathrm{diag}\left(\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_n}\right)$.

(b) $Ax = b$ ⟹ $(U\Sigma V^t)x = b$ ⟹ $x = V\Sigma^{-1}U^t b$.

(c), (d): code

(e) Gaussian Elimination ⟹ $\frac{2}{3}n^3 - \frac{1}{2}n^2 - \frac{1}{6}n$ flops

∵ SVD, it is approximately $2n^3$

∴ GE is better !!

∎

16.2. Yes. ∵ A is inv. ∃! $\underset{x}{\mathrm{argmin}} \|A\bar{x} - b\|_2$

Also, ∃! $\bar{x}$ from G-E. s.t. $A\bar{x} = b$ ⟹ $\underset{x}{\mathrm{argmin}} \|A\bar{x} - b\|_2 = \bar{x}$ ∎

16.3. $b \in \langle [A]^1, \cdots [A]^n \rangle^{\perp}$ ⟹ $b \in (\mathrm{Im} A)^{\perp}$.

↳ ∵ $A^t b = 0$ ⟹ $([A]^i)^t b = 0$ ($\forall i$) ∎

16.6. A: full-rk mtx with $A \in M_{m,n}(\mathbb{R})$, $m < n$.

(a) $A^t A x = A^t b$

⟹ $A^t A \left(A^t (AA^t)^{-1} b + (I - A^t (AA^t)^{-1} A) y \right)$

$= A^t b + A^t A y - A^t A A^t (AA^t)^{-1} y = A^t b$.

∴ $x = A^t (AA^t)^{-1} b + (I - A^t (AA^t)^{-1} A) y$ satisfies normal eq.

$$A = U\tilde{\Sigma}V^t \rightsquigarrow A^tA = \underline{V\tilde{\Sigma}^t\tilde{\Sigma}V^t.}$$

$$V\tilde{\Sigma}^t\tilde{\Sigma}V^t x = V\tilde{\Sigma}^t U^t b$$

$$\rightsquigarrow \underline{(\tilde{\Sigma}^t\cdot\tilde{\Sigma})}V^t x = \underline{\tilde{\Sigma}^t U^t b}$$

$$\downarrow \begin{bmatrix} \sigma_1^2 v_1^t \\ \sigma_m^2 v_m^t \\ \hline 0 \end{bmatrix} \qquad \hookrightarrow \begin{bmatrix} \sigma_1 u_1^t b \\ \vdots \\ \sigma_m u_m^t b \\ \hline 0 \end{bmatrix}$$

$$\therefore \bar{x} = \begin{bmatrix} \frac{1}{\sigma_1^2} v_1 & \cdots & \frac{1}{\sigma_m^2} v_m \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 u_1^t b \\ \vdots \\ \sigma_m u_m^t b \end{bmatrix}$$

$$= \sum_{j=1}^{m} \frac{1}{\sigma_j} \cdot v_j u_j^t b = \sum_{j=1}^{n} v_j \sigma_j^{-1} u_j^t \cdot b = \sum_{j=1}^{n} v_j \sigma_j \cdot \sigma_j^{-2} \cdot u_j^t b$$

$$= \sum_{j=1}^{n} v_j \sigma_j \cdot u_j^t u_j \sigma_j^{-2} \cdot u_j^t b = \sum_{i,j}^{n} v_i \sigma_i u_i^t u_j \sigma_j^{-2} u_j^t b$$

$$= [v_1 \cdots v_n] \operatorname{diag}[\sigma_1 \cdots \sigma_n] \begin{bmatrix} u_1^t \\ \vdots \\ u_n^t \end{bmatrix}$$

$$\cdot [u_1 \cdots u_n] \cdot \operatorname{diag}\left[\frac{1}{\sigma_1^2}, \cdots, \frac{1}{\sigma_n^2}\right] \cdot \begin{bmatrix} u_1^t \\ \vdots \\ u_n^t \end{bmatrix} b$$

$$= V\cdot\tilde{\Sigma}\cdot U^t \cdot U \cdot \tilde{\Sigma}^t \tilde{\Sigma} U^t b = \underline{A^t(AA^t)^{-1}b}$$

$\exists\ (\tilde{\Sigma}^t\cdot\tilde{\Sigma})V^t x = 0$ case, $\exists\ \mathrm{rk}(\tilde{\Sigma}^t\cdot\tilde{\Sigma})V^t = m.$

$\dim \ker(\tilde{\Sigma}^t\cdot\tilde{\Sigma})V^t = n-m. \to \ker(\tilde{\Sigma}^t\cdot\tilde{\Sigma})V^t = \langle v_{m+1}, \cdots, v_n \rangle.$

$$(I - A^t(AA^t)^{-1}A)y = (I - V\tilde{\Sigma}^t(\tilde{\Sigma}\tilde{\Sigma}^t)^{-1}\tilde{\Sigma}V^t)y$$

$$= V\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-m} \end{array}\right)V^t y = [0 \mid v_{m+1}, \cdots, v_n]\,\underline{V}^t y$$

Inv. Mtx

$\ker(\tilde{\Sigma}^t\cdot\tilde{\Sigma})V^t = \{(I - A^t(AA^t)^{-1}A)y : \forall y \in \mathbb{R}^n\}$

$\therefore x = A^t(AA^t)^{-1}b + (I - A^t(AA^t)^{-1}A)y$ is all solution of $\operatorname{argmin} \|Ax-b\|_2$

From the proof, $A^t(AA^t)^{-1}b \in \langle v_1, \dots, v_m \rangle$, $(I - A^t(AA^t)^{-1}A)y \in \langle v_{m+1}, \dots, v_n \rangle$ $\Rightarrow$ $A^t(AA^t)^{-1}b \perp I - A^t(AA^t)^{-1}A$

$\therefore$ Least norm solution : $y=0 \Rightarrow$ $\underline{A^t(AA^t)^{-1}b}$

(c) No. too many multiplication, inverse calculations. ∎

16.13. $\exists\, x_1, x_2 \in \{\underset{x}{\operatorname{argmin}} \|Ax-b\|_2\}$, If $x_1 \neq x_2$, $x_1 - x_2 \in \ker A$.

$\therefore \underset{x}{\operatorname{argmin}} \|Ax-b\|_2$ is in form of $\bar{x} + \eta$ ($\eta \in \ker A$).

we need a lemma.

(Thm) $A = U\Sigma V^t$, SVD. setting $U = [u_1, \cdots u_m]$, $V = [v_1, \cdots, v_n]$.

$r = rk(A)$, $x_{LS} := \sum_{i=1}^{r} \left(\frac{u_i^t b}{\sigma_i}\right) v_i$ minimizes $\|Ax-b\|_2$ and has the smallest 2-norm of all minimizers. Also,

$$r_{LS}^2 = \| Ax_{LS} - b \|^2 = \sum_{i=r+1}^{m} (u_i^t b)^2$$

(proof)

$$\|b - Ax\|_2^2 = \|U^t b - U^t A x\|_2^2 = \|U^t b - U^t A V \cdot V^t x\|_2^2 = \|U^t b - \tilde{\Sigma} \alpha\|_2^2,$$

with $\alpha = V^t x = [\alpha_1, \dots, \alpha_n]^t$.

$$\tilde{\Sigma}\alpha = [\sigma_1\alpha_1 \cdots \sigma_r\alpha_r \ 0 \cdots 0]^t$$

$$U^t b = \begin{bmatrix} u_1^t \\ \vdots \\ u_m^t \end{bmatrix} b = \begin{bmatrix} u_1^t b \\ \vdots \\ u_m^t b \end{bmatrix} \rightarrow \|U^t b - \Sigma\alpha\|_2^2 = \sum_{i=1}^{r} (\sigma_i\alpha_i - u_i^t b)^2 + \sum_{i=r+1}^{m} (u_i^t b)^2$$

$$\underset{\alpha}{\operatorname{argmin}} \|U^t b - \Sigma\alpha\|_2^2 = \left[\frac{u_i^t b}{\sigma_i} \,\middle|\, 0\right]$$

$$\Rightarrow x = V\alpha = \sum_{i=1}^{r} \frac{u_i^t b}{\sigma_i} v_i + \sum_{i=r+1}^{n} \alpha_i v_i$$

setting $\alpha_{r+1}, \cdots \alpha_m = 0$.

$\|x\|_2^2$ min !!

□

WTS $\| b - Ax_{LS} \|_2 = \| AA^+ x - b \|_2$

$(A^+ := V \Sigma^+ U^t , \quad \Sigma^+ = \mathrm{diag}(\frac{1}{\sigma_1}, \cdots \frac{1}{\sigma_r}, 0, \cdots, 0))$

$A^+ = V\Sigma^+ U^t$ and $A^+ b = [v_1 \cdots v_n] \ \mathrm{diag}(\frac{1}{\sigma_1}, \cdots \frac{1}{\sigma_r}, \cdots 0) \cdot \begin{bmatrix} u_1^t \\ \vdots \\ u_m^t \end{bmatrix} . b$

$= [\sigma_1 v_1 \cdots \ \sigma_r v_r \ 0 \cdots 0] . \begin{bmatrix} u_1^t b \\ \vdots \\ u_m^t b \end{bmatrix}$

$= \sum_{i=1}^{r} (\frac{u_i^t b}{\sigma_i}) v_i = x_{LS}$ //

$\therefore$ All solutions are a form of $\hat{x} = A^+ b + h \ (h \in \ker(A))$ ■

16.15.

(a) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $A^t A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\sigma_1 = 1, \sigma_2 = 0$

and $\tilde{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \ V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$

$\therefore A^+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\underbrace{A + \delta A}_{A_1} = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \\ 0 & 0 \end{bmatrix}$ and $A_1^t A_1 = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon^2 \end{bmatrix}$

$\therefore (A + \delta A)^+ = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\varepsilon^2} \end{bmatrix} . \begin{bmatrix} 1 & 0 & 0 \\ 0 & \varepsilon & 0 \end{bmatrix}$

$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\varepsilon} & 0 \end{bmatrix}$

$\therefore A^+ - (A + \delta A)^+ = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{\varepsilon} & 0 \end{bmatrix}.$ and $\| A^+ - (A+\delta A)^+ \|_2 = ?$

and $\begin{bmatrix} 0 & 0 \\ 0 & -1/\varepsilon \\ 0 & 0 \end{bmatrix} . \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1/\varepsilon & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\varepsilon^2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\rightarrow \| A^+ - (A + \delta A)^+ \|_2 = \sqrt{1/\varepsilon^2} = \frac{1}{\varepsilon}$

(b) $\lim_{\delta A \to 0} \|A^+ - (A+\delta A)^+\|_F$ could get large...

$\leadsto$ rank-deficient system is

hard to solve numerically. $\blacksquare$