Overdetermined System

$A \quad m \times n \qquad m > n$

$$Ax \cong b$$

usually no solution in the usual sense. ($b \notin \text{span}(A)$)

instead

$$\min_x \| b - Ax \| \qquad \underline{\text{residual}}$$

"least squares" = min sum of squares $\sum_i (b_i - (Ax)_i)^2$

Gauss

"regression analysis"

- Sol'n is unique iff A has full column rank.

§ 3.2.1 Normal Equation

$\min \phi(x)$

$$\min \|r\|^2 = (b-Ax)^T(b-Ax)$$
$$= b^Tb - b^TAx - x^TA^Tb + x^TA^TAx$$

$$\phi(x) = x^TA^TAx - 2b^TAx + b^Tb$$

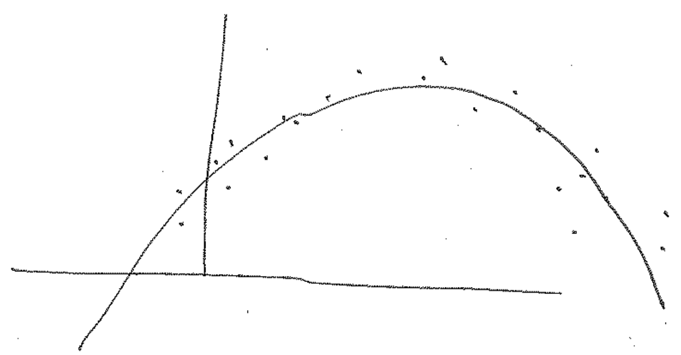
$$\frac{d\phi}{dx} = 0 \Rightarrow$$

$$\frac{\partial}{\partial x_k}\left[x_i (A^TA)_{ij} x_j - 2(b^TA)_i x_i + b^Tb \right]$$

$$= \delta_{ik}(A^TA)_{ij}x_j - 2(b^TA)_i \delta_{ik} + x_i (A^TA)_{ij}\delta_{jk}$$

$$= (A^TA)_{kj}x_j - 2(b^TA)_k + x_i(A^TA)_{ik}$$

__Example: data fitting__



$$f(x) = ax^2 + bx + c$$

(1, 3)

(2, 4)

(3, 3)

$\vdots$

$$\begin{pmatrix} 1^2 & 1 & 1 \\ 2^2 & 2 & 1 \\ 3^2 & 3 & 1 \\ \vdots & & \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} \cong \begin{pmatrix} 3 \\ 4 \\ 3 \\ \vdots \end{pmatrix}$$

$$= A^TAx + (x^TA^TA)^T - 2(b^TA)^T$$

$$= 2A^TAx - 2A^Tb = 0$$

$$\Rightarrow \boxed{A^TAx = A^Tb}$$ <u>Normal Equations</u>.

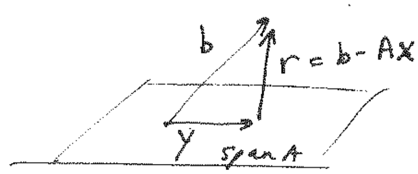
min if $\frac{\partial^2\phi}{\partial x^2} > 0$ $\qquad \frac{\partial^2\phi}{\partial x^2} = A^TA \succ 0$ if A full col rank.

$\boxed{\S 3.2.2.}$ <u>Orthogonality & Orthogonal Projectors.</u>

$x^Ty = 0 \qquad x, y$ orthogonal.

usually, $b \notin \operatorname{span}(A)$

closest $y \in \operatorname{span}(A)$



$y = Ax$

want $r \perp A$.

$$A^Tr = 0 \quad \Rightarrow \quad A^T(b - Ax) = 0$$

$\underline{A^TAx = A^Tb}$ Normal equations!

y is "orthogonal projection" of b onto $\operatorname{span}(A)$

$P^2 = P$ projector

$+ \ P = P^T$ orthogonal projector.

$P^\perp = I - P$ is orthogonal projector onto $\operatorname{span}(P)^\perp$

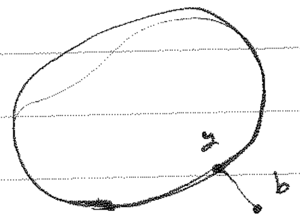
$$v = [P + (I - P)]v = Pv + P^\perp v$$

LECTURE 6

Thursday, 10/18

2 § 3.2. Existence + Uniqueness of LS. ~~problem~~ solution

$\phi(\vec{y}) = \|\vec{b} - \vec{y}\|_2$ continuous + coercive
$\therefore$ sol'n exists



$y \in \text{span}(A)$ unique $\not\Rightarrow$ x unique

$$y = Ax = A(x+z), \quad \text{if } Az = 0.$$

x unique iff A has full col. rank

✓ 3.2.1. Normal Eqs.

✓ 3.2.2 Orthogonality + Orthogonal Projectors

1 Pseudo inverse A^+ $\quad$ A $m \times n$, $m > n$

$$A = U\Sigma V^T$$

$$A^+ = V\Sigma^+ U^T, \quad \Sigma^+_{ij} = \begin{cases} (\Sigma_{ji})^{-1} & \text{for } \Sigma_{ji} \neq 0 \\ 0 & \text{for } \Sigma_{ji} = 0 \end{cases}$$

check:

$$\underset{m\times n\ n\times m}{AA^+} = U\Sigma V^T V \Sigma^+ U^T = U\Sigma\Sigma^+ U^T$$

$$(\Sigma\Sigma^+)_{ij} = \sum_{k=1}^{n} (\Sigma_{ik}\Sigma^+_{kj}) = \cancel{\delta_{\min(i,j)}} \; \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \ \sigma_n \\ & 0 & \end{pmatrix} \begin{pmatrix} \sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & 0 \\ & & \ddots \ \sigma_n^{-1} & \end{pmatrix}$$

$$= \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{n} & \\ & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{pmatrix}$$

$$= \overset{m\times m}{\begin{bmatrix} \tilde{U} & 0 \end{bmatrix} \begin{bmatrix} \tilde{U}^T \\ 0 \end{bmatrix}} = \sum_{i=1}^{n} u_i u_i^T \quad (\neq I_m)$$

$$A^+ A = (V\Sigma^+ U^T)(U\Sigma V^T) = V\Sigma^+\Sigma V^T = V I_n V^T = I_n \checkmark$$

$$\Sigma^+\Sigma = \begin{pmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_n^{-1} \end{pmatrix}\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \end{pmatrix} = I_n$$

Goal: Transform

$$Ax \cong b$$

to a square triangular system

$\boxed{\S 3.4}$

(3) (1) Normal Equations

$$A^T A x = A^T b$$

$$A^T A = LL^T$$

Solve $LL^T x = A^T b$

Bad:

- $A = \begin{pmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} 1 & \varepsilon & \\ 1 & & \varepsilon \end{pmatrix}\begin{pmatrix} 1 & 1 \\ \varepsilon & \\ & \varepsilon \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} \cancel{\varepsilon^2} & \\ & \varepsilon^2 \end{pmatrix} \approx 0.$

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

(4) $\boxed{\S 3.4.2}$ Augmented System

$$A^T(\underbrace{b - Ax}_{r}) = 0$$

$$\Rightarrow \begin{pmatrix} A^T r = 0 \\ r + Ax = b \end{pmatrix} \Rightarrow$$

symmetric indefinite

$$\boxed{\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}}$$

$(m+n) \times (m+n)$

(MATLAB) Scale

$$\begin{pmatrix} \alpha I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} r/\alpha \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

rule-of-thumb

$$\alpha \doteq \frac{\max|a_{ij}|}{1000}$$

(6) $\boxed{\S 3.4.3}$ Orthogonal Transformations

Gaussian transformations do not preserve Euclidean Norm, but orthogonal transformations do!

e.g., Rotation, Reflection

$$\|Ax - b\| = \|QAx - Qb\|$$

for orthogonal Q.

$$\|Qx\|_2^2 = (x^T Q^T)(Qx) = x^T(Q^T Q)x = x^T x = \|x\|_2^2$$

(Orth matrices don't change length ⇒ don't amplify error ⟶ don't need pivoting)

(But ... more expensive that G.E.)

(5) $\boxed{\S 3.4.4}$ Triangular LS.

$$\begin{pmatrix} R \\ 0 \end{pmatrix} x \cong \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

$$Rx \cong c_1$$

$$0 \cdot x \cong c_2$$

$$\|r\|_2^2 = \underbrace{\|c_1 - Rx\|_2^2}_{\text{force this to be } 0.} + \underbrace{\|c_2\|_2^2}_{\text{no control over this term.}}$$

LS. soln : $\boxed{Rx = c_1}$ Solve for x by back-substitution

residual: $\|r\|_2^2 = \|c_2\|_2^2$