

Overdetermined System

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

$$Ax \cong b$$

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

instead

$$\min_x \|b - Ax\|$$

residual

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

Gauss

"regression analysis"

- Soln is unique iff A has full column rank.

§ 3.2.1 Normal Equation

min $\phi(x)$

$$\min \|r\|^2 = (b - Ax)^T (b - Ax)$$

$$= b^T b - b^T A x - x^T A^T b + x^T A^T A x$$

$$\phi(x) = x^T A^T A x - 2b^T A x + b^T b$$

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

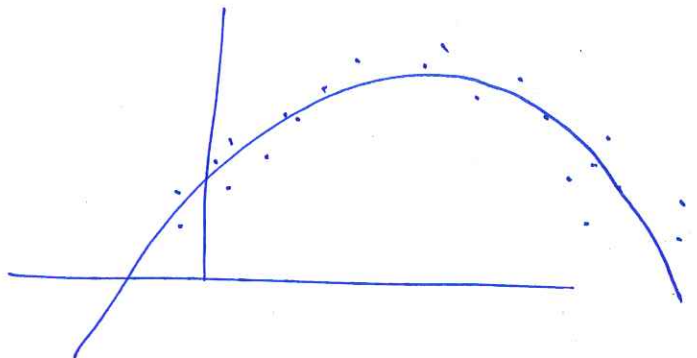
$$\frac{\partial}{\partial x_k} [x_i (A^T A)_{ij} x_j - 2(b^T A)_i x_i + b^T b]$$

$$= \delta_{ik} (A^T A)_{ij} x_j - 2(b^T A)_i \delta_{ik}$$

$$+ x_i (A^T A)_{ij} \delta_{jk}$$

$$= (A^T A)_{kj} x_j - 2(b^T A)_k$$

Example: data fitting



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

$$(1, 3)$$

$$(2, 4)$$

$$(3, 3)$$

⋮

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

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

$$= 2 A^T A x - 2 A^T b = 0$$

$$\Rightarrow \boxed{A^T A x = A^T b} \quad \text{Normal Equations.}$$

min if

$$\frac{\partial^2 \phi}{\partial x^2} > 0$$

$$\frac{\partial^2 \phi}{\partial x^2} = A^T A$$

> 0 if A full col rank.

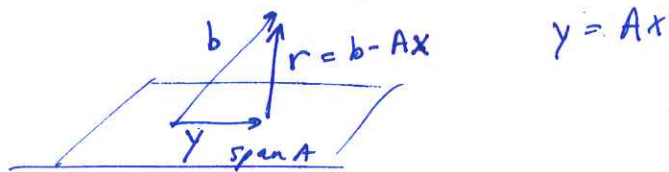
§ 3.2.2. Orthogonality + Orthogonal Projectors.

$$x^T y = 0$$

x, y orthogonal.

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

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



want $r \perp A$.

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

$$\underline{A^T A x = A^T b} \quad \text{Normal equation}$$

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

$$P^2 = P$$

projector

$$+ P = P^T$$

orthogonal projector.

$$P^\perp = I - P$$

is orthogonal projector onto $\text{span}(P^\perp)^\perp$

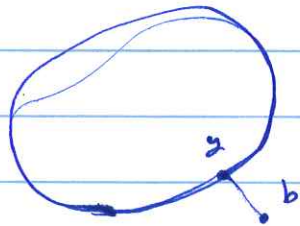
$$v = [P + (I - P)]v = Pv + P^\perp v$$

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 $\nRightarrow x$ unique
 $y = Ax$
 $= A(x+z)$, 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^+ A $m \times n$, $m > n$

$A = U \Sigma V^T$

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

check:

$AA^+ = U \Sigma V^T V \Sigma^+ U^T$
 $= U \Sigma \Sigma^+ U^T$

$(\Sigma \Sigma^+)_{ij} = \sum_{k=1}^n (\Sigma_{ik} \Sigma^+_{kj}) = \sum_{k=1}^n \begin{cases} \sigma_i \sigma_k^{-1} \sigma_k & \text{if } i=k \\ 0 & \text{if } i \neq k \end{cases}$

$= \begin{pmatrix} 1 & & & \\ & \dots & & \\ & & 1 & \\ & & & \dots & \\ & & & & 0 & \dots & \\ & & & & & & 0 \end{pmatrix}$

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

$$\begin{aligned}
 A^+ A &= (V \Sigma^+ U^T)(U \Sigma V^T) \\
 &= V \Sigma^+ \Sigma V^T \\
 &= V I_n V^T \\
 &= I_n \quad \checkmark
 \end{aligned}$$

$$\begin{aligned}
 \Sigma^+ \Sigma &= \begin{pmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_r^{-1} \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & & 0 \end{pmatrix} \\
 &= I_n
 \end{aligned}$$

Goal: Transform

$$Ax \approx b$$

to a square triangular system

§ 3.4

③ ① Normal Equations

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

$$A^T A = L L^T$$

$$\text{Solve } L L^T x = A^T b$$

Bad:

$$\begin{aligned}
 \bullet A &= \begin{pmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} 1 & \epsilon \\ \epsilon & \epsilon^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \epsilon & \epsilon \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \epsilon & \epsilon \end{pmatrix} \\
 &+ \begin{pmatrix} \epsilon^2 & \epsilon^2 \\ \epsilon^2 & \epsilon^2 \end{pmatrix}
 \end{aligned}$$

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

④ § 3.4.2 Augmented System

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

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

Symmetric
indefinite

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

(m*n) x (m*n)

(MATLAB)

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

rule-of-thumb

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

6 §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 \Rightarrow don't amplify error \rightarrow don't need pivoting)

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

5 §3.4.4 Triangular LS.

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

$$Rx \hat{=} c_1$$

$$0 \cdot x \hat{=} c_2$$

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

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

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