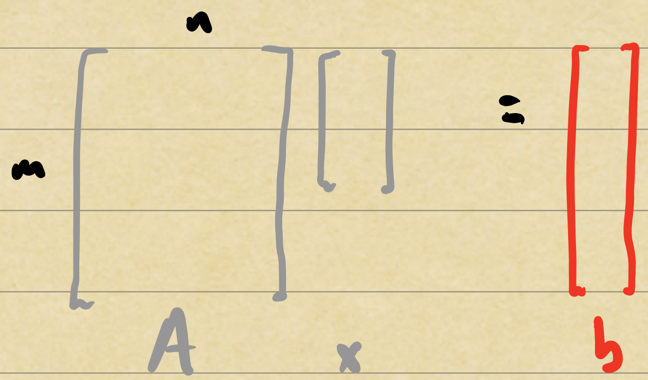
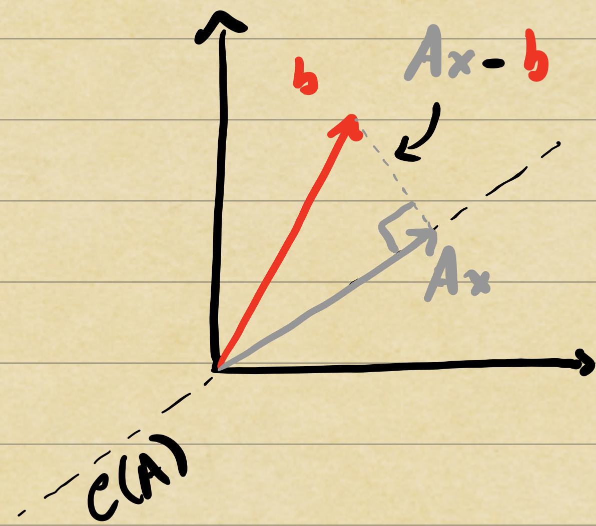
Overdetermined Least-Squares

m equations > n unknowns





(P1) $$x_* = \operatorname{argmin} \|Ax - b\|_2$$

"Least-squares solution"

To find x_*, we have 2 steps:

1) Project $d = \underbrace{P}_{}b$ orthogonal projection onto $C(A)$

2) Solve $Ax_* = d$

Three ways to solve (P1)

Normal Equations: $\underbrace{Ax_*-b}_{r\,=\,\text{residual}} \perp \mathcal{C}(A)$

(18.06) Which of 4 fundamental subspaces does r live in? $\Rightarrow$ nullspace(A^*)!

$$A^*r = 0 \iff \underbrace{A^*A}_{n\times n \text{ matrix}}x_* = \underbrace{A^*b}_{\in\,\mathcal{C}(A)}$$

$$x_* = \underbrace{(A^*A)}_{\substack{\text{nonsingular IFF} \\ \text{rank}(A)=n \text{ (full col. rank)}}}{}^{-1}A^*b$$

Numerically tricky, usually avoid this formulation

$$A^+ = (A^*A)^{-1}A^*b$$

"Pseudoinverse"

Singular Value Decomposition: $A = U\Sigma V^*$

diagonalize $y = V^*x$ $d = U^*b$

$$\|U\Sigma V^* x - b\|_2 = \|\Sigma y - d\|_2$$

minimize over y instead of x (equiv.)

$$\left[\begin{array}{ccc} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{array}\right] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \left[\begin{array}{c} d_1 \\ \vdots \\ d_n \\ \hline d_{n+1} \\ \vdots \\ d_m \end{array}\right]$$

Solve $\Sigma y_* = d$

$\} = \Sigma y_* - d$

So $\quad x_* = V y_* = V \Sigma^{-1}_{1:n,1:n} d_{1:n}$

$$= \underbrace{\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_n^{-1} \end{bmatrix}}_{\tilde{\Sigma}^{-1} = \Sigma^{-1}_{1:n,1:n}} \underbrace{\begin{bmatrix} - & u_1^* & - \\ & \vdots & \\ - & u_n^* & - \end{bmatrix}}_{\tilde{U}^* = U^*_{1:n}} \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}}_{b}$$

pseudoinverse again, much better numerically

$$A^+ = V \tilde{\Sigma}^{-1} \tilde{U}^*$$

$$x_* = \underbrace{V \tilde{\Sigma}^{-1} \tilde{U}^*}_{\substack{A^+ \\ \text{(step 2 solve } Ax = Pb)}} \overbrace{\phantom{\tilde{U}^*}}^{= I_{n \times n} \text{ b/c orthogonal cols of } \tilde{U}} \underbrace{\tilde{U}\tilde{U}^* b}_{\substack{Pb \\ \text{(step 1 orthogonal project)}}}$$

(New!) QR Decomposition:

Diagonalization (SVD) is expensive, can we work with orthogonal transforms but cut cost? => Triangularize A.

$$A = \overset{m \times n}{Q} \, \overset{n \times n}{R}$$

orthogonal basis for $C(A)$ ↑ ↑ upper triangular matrix

Step 1: Project $d = QQ^* b$

Step 2: Solve $Ax_* = d$

$$(QR)\, x_* = QQ^* b$$

solve => $Rx_* = Q^* b$

Pseudoinverse again

$$x_* = \underbrace{R^{-1} Q^*}_{A^+} b$$

Conditioning & Sensitivity

$$x_* = A^+ b$$

A^+ ← linear map

Sensitivity w/respect to perturbed $b + \delta b$

$$x_* + \delta x_* = A^+(b + \delta b)$$

$$\delta x_* = \underbrace{A^+}_{\text{Jacobian}} \delta b$$

Condition #
(relative)

$$\kappa_A(b) = \frac{\|A^+\|}{\|x_*\| / \|b\|} = \|A^+\| \frac{\|b\|}{\|Ax\|} \frac{\|Ax\|}{\|x_*\|}$$

$$= \frac{\|A^+\| \|A\|}{\eta \cos\theta} = \frac{\sigma_1 / \sigma_n}{\eta \cos\theta}$$

$$\cos\theta = \frac{\|Ax\|}{\|b\|} \quad \text{and} \quad \eta = \frac{\|A\| \|x\|}{\|Ax\|}$$

θ ← angle between Ax and b

$$0 \le \theta \le \frac{\pi}{2} \qquad 1 \le \eta \le \frac{\sigma_1}{\sigma_n} \qquad 1 \le \sigma_1/\sigma_n < \infty$$

Roughly corresponds to 2-step solution:

1) Orthogonal projection $d = Pb$

2) Solve reduced $n \times n$ system for x_*

$\theta \approx \frac{\pi}{2}$, Step 1 is ill-conditioned

$\sigma_1/\sigma_n \gg 1$, Step 2 is ill-conditioned.