March 15, 2021 ① b not close to R(A) ② A ill-conditioned ✓

data points: (a_i^T, b_i) $i = 1, \dots, m$

$$\hat{x} = \arg\min_x \|Ax - b\|_2^2$$

$$\hat{x} = R^{-1}Q^T b = V\Sigma^{-1}U^T b = A^+ b$$

- $A\hat{x} = b + r$ $\quad A^T r = 0$

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad r = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

- $\hat{x}_1 = A_1^+ b_1$ how good on A?

$$\|A\hat{x}_1 - b\|^2$$
$$= \|A\hat{x}_1 + r - A\hat{x}\|^2$$
$$= \|A(\hat{x}_1 - \hat{x}) + r\|^2$$
$$= \|A(\hat{x}_1 - x)\|^2 + \|r\|^2 \quad (A^T r = 0)$$

bias: best we can hope for

Variance: sensitivity to data perturbations

Claim: $\hat{x} = A_1^+(b_1 + r_1)$

Proof: $= \arg\min_{x_1} \|A_1 x_1 - (b_1 + r_1)\|$ $\quad \geq 0$

$$\|AA_1^+(b_1 - b_1 - r_1)\|$$
$$\leq \|A\|\,\|A_1^+\|\,\|r\|$$

$\|A\| \to \sigma_1(A)$, $\quad \|A_1^+\| \sim 1/\sigma_n(A)$

$$= \kappa_2(A) \cdot \|r\|$$

Idea: better conditioned problems
(↘ variance)

with a little more
(↗ bias)

$A^+b = V_{n\times n}\,\Sigma_{n\times n}\,U^T_{n\times m}\,b_m$ $\quad m \times n$

$A_1^+b_1 = V_{1\,(n\times n)}\,\Sigma_{n\times n}\,U^T_{1\,(n\times m_1)}\,b_{1\,(m_1)}$ $\quad m_1 \times n$