Conditioning of least squares problems

Conditioning of linear least squares is a more complicated problem than the one for linear systems.
We won't give a full proof:

Theorem (Trefethen, Bau, Theorem 18.1)

Consider the linear least squares problem $\boxed{\min \|Ax - b\|}$ with $A \in \mathbb{R}^{m \times n}$ with full column rank. Its relative condition number with respect to the input b is

$$\boxed{\kappa_{rel, b \to x} \leq \frac{\kappa(A)}{\cos\theta},}$$

and with respect to A it is

$$\kappa_{rel, A \to x} \leq \boxed{\kappa(A) + \kappa(A)^2 \tan\theta,}$$

where θ is the angle such that $\cos\theta = \frac{\|Ax\|}{\|b\|}$.

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| = \frac{\sigma_1}{\sigma_n} \quad \text{for } A \in \mathbb{R}^{n \times n}$$

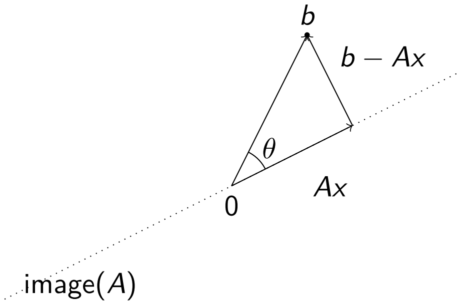
$$\kappa(A) = \cancel{\|A\| \cdot \|A^{-1}\|} = \frac{\sigma_1}{\sigma_n} \quad \text{for } A \in \mathbb{R}^{m \times n}$$

↯

inverse of A does not exist

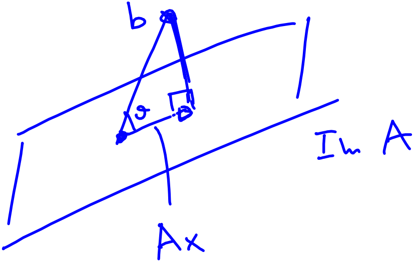
def: $\kappa(A) = \dfrac{\sigma_{max}}{\sigma_{min}}$ also for non-square A

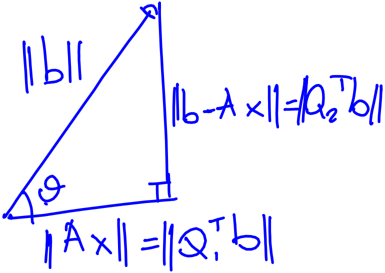
The geometric picture



b 'split' into two orthogonal components: Ax and $b - Ax$.
QR and SVD reveal their norms: if $A = QR, Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ or $A = U\Sigma V^T$, $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ (as in their thing versions) then

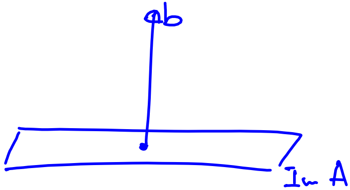
$$\|Ax\| = \|Q_1^T b\| = \|U_1^T b\| = \|b\| \cos\theta,$$
$$\|b - Ax\| = \|Q_2^T b\| = \|U_2^T b\| = \|b\| \sin\theta.$$





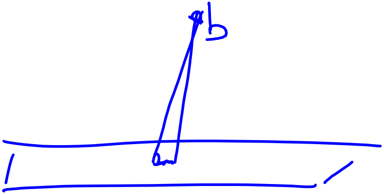
$$\|Q^T(Ax-b)\| = \left\| \begin{bmatrix} A \\ 0 \end{bmatrix} x - \begin{bmatrix} Q_1^T b \\ Q_2^T b \end{bmatrix} \right\| = \left\| \begin{bmatrix} \overbrace{\boxed{Ax - Q_1^T b}}^{\text{we set this to } 0} \\ Q_2^T b \end{bmatrix} \right\|$$

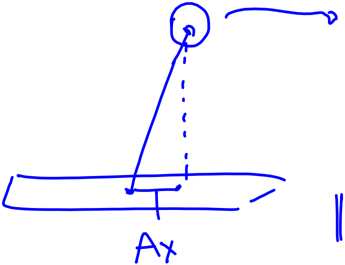
$$\min \|Ax-b\| = \|Q_2^T b\|$$



Case 1: $\mathrm{Im}\,A$, b almost orthogonal

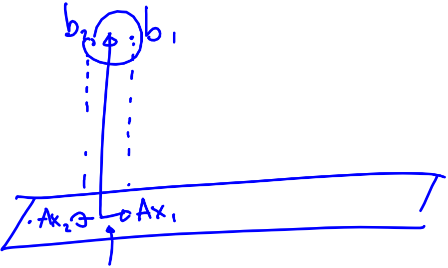
$\Rightarrow \theta \approx 90^\circ$





b perturbed to something in this ball of radius δ with $\dfrac{\delta}{\|b\|}$ small

$\|Ax\| = \|b\| \cdot \cos\theta$



$$\frac{\|b_1 - b_2\|}{\|b_1\|} \quad \text{small}$$

$$\frac{\|x_1 - x_2\|}{\|x_1\|} \quad \text{large}$$

Some intuition

A small $\frac{\|\Delta b\|}{\|b\|}$ may correspond to a large $\frac{\|Q_1^T \Delta b\|}{\|Q_1^T b\|}$ — the denominator differs by a factor $\cos\theta$.

Solving a linear system + this possible error amplification gives $\kappa_{rel,b\to x} \leq \frac{\kappa(A)}{\cos\theta}$.

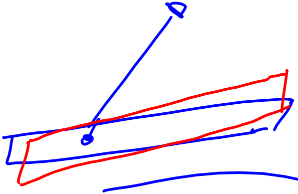
This perturbation is more substantial if b and Ax are almost orthogonal, i.e., $\cos\theta \approx 0$.

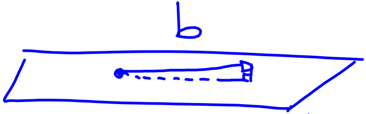
Some intuition

Perturbing ΔA may 'tilt' image(A): this changes both $Q_1^T b$ and the matrix in the linear system.

This perturbation is less substantial if Ax and b are almost aligned, i.e., $\sin\theta \approx 0$.

$\theta \approx 0$



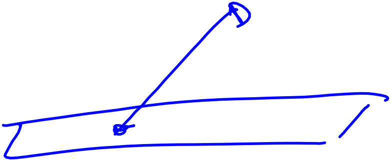


$\|Q_1^T b\|$ large

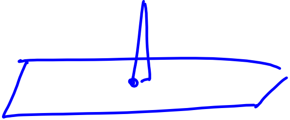
$\|Q_2^T b\|$ small

$\tan\theta \approx 0$

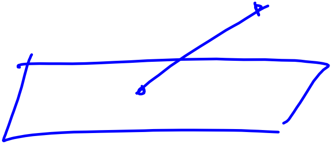
$\kappa_{A\to x} \approx \kappa(A)$



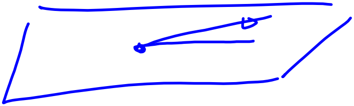
$\rightsquigarrow \kappa_{A\to x} \approx \kappa(A)^2$



Im A, b almost orthogonal $=$ no bounds, could be ∞



general $\approx \kappa^2(A)$



almost parallel $\approx \kappa(A)$