$$\kappa(A) = \|A\| \cdot \|A^{-1}\|$$

Condition # of matrix

$$A = \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & & \ddots & & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}$$

$\in \mathbb{R}^{n \times n}$

A commonly inverted tridiagonal matrix.

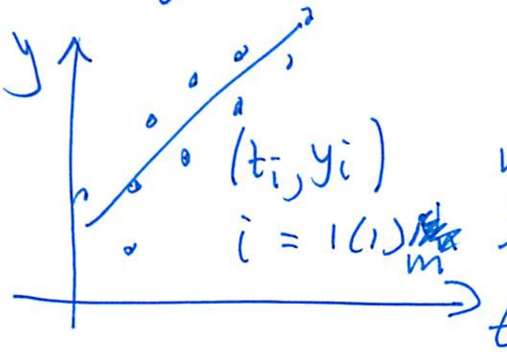
i.e. $u_{xx} = f$ on $[0,1]$

$u(0) = a$, $u(0) = b$

$$Ax = b$$

n	$\kappa_2(A) \sim n^2$
10	60
20	220
100	5100
1000	5.1×10^5

Least Squares problems



Problem: Given ~~measure~~ measurements, find "best-fit" line $y = ct + d$ to the data.

These is no sigle line going through all the data for $m > 2$. This problem is "over-determined", i.e. more constraints (m pairs (t_i, y_i)) than unknowns $[2 - (c, d)]$

$$
\begin{aligned}
y_1 &= a t_1 + b \\
&\vdots \\
y_m &= a t_m + b
\end{aligned}
\iff
\underbrace{\begin{bmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{bmatrix}}_{A}
\underbrace{\begin{bmatrix} c \\ d \end{bmatrix}}_{x}
=
\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{b}
$$

rure is generally no solution to this problem. Instead, ~~pose~~ look for a "best-fit" pair (c, d) that satisfies a "least-squares" problem.

Given $A \in \mathbb{R}^{m \times n}$, $m > n$, $b \in \mathbb{R}^{m}$

~~$\min_{x \in \mathbb{R}}$~~

Find $x \in \mathbb{R}^{n}$ satisfying $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$

or equivalently

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2 \longleftarrow \text{least squares.}$$

For the line fit problem:

$$\min_{(c,d)} \sum_{k=1}^{n} (ct_i + d - y_i)^2$$

$$n = 2$$

$\cancel{f : \mathbb{R}^n \to \mathbb{R}^+}$ Recall $\|y\|_2^2 = y^T y = y \cdot y$

$$\cancel{f(x)} \quad \Rightarrow f(x) = \|Ax - b\|_2^2 : \mathbb{R}^m \to \mathbb{R}^+$$

$$= (Ax - b)^T (Ax - b)$$

$$= x^T A^T A x - (x^T A^T b + \underbrace{b^T A x}_{(Ax)^T}) + b^T b$$

$B \in \mathbb{R}^{n \times n}$

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

Find $x_* \in \mathbb{R}^n$ s.t. $\nabla f(x_*) = 0$ (i.e. a critical pt)

$$\cancel{f(x) = \sum_{\substack{i,j \\ 1 \le i,j \le n}} x_i A_{ij} x_j - 2 \sum_{\substack{i,k \\ 1 \le i \le n \\ 1 \le j \le m}} x_i (A^T)_{ij}}$$

$$f = (Ax)^T(Ax) - 2(Ax)^T b + b^T b$$

$$g_i = \sum_j a_{ij} x_j$$

$$\frac{\partial g_i}{\partial x_p} = a_{ip}$$

$$f = \sum_i g_i(x)^2 - 2\sum_i g_i(x) b_i + b^T b$$

$$(Df)_p = \frac{\partial f}{\partial x_p} = 2\sum_i g_i(x) a_{ip} - 2\sum_i a_{ip} b_i$$

$$= 2 \sum_i a_{ip}(g_i(x) - b_i)$$

$$= 2 \sum_i (A^T)_{pi}\left(\sum_j (A)_{ij} x_j - b_i\right)$$

<u>or</u> $Df(x) = 2(A^T A x - A^T b)$ ✓

The <u>normal equations</u>: Critical point x_* satisfies

$$\boxed{A^T A x = A^T b}$$

$A^T A \in \mathbb{R}^{n \times n}$, $A^T b \in \mathbb{R}^n$

Solving these equations solves the least squares problem, but the normal equations can be poorly conditioned.

Example $A = \begin{pmatrix} \varepsilon & 0 \\ 0 & 1 \end{pmatrix}$; $A^{-1} = \begin{pmatrix} \varepsilon^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ $0 < \varepsilon \ll 1$

$$B = A^T A = \begin{pmatrix} \varepsilon^2 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\kappa_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = 1 \cdot \varepsilon^{-1} = 1/\varepsilon$$

$$\kappa_2(A^T A) = 1/\varepsilon^2$$

Essentially, you squared the condition # of A.

There are better ways:

Orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ satisfies

$$Q^T Q = I \in \mathbb{R}^{m \times m} \quad \left(\text{i.e. } Q^T = Q^{-1} \Rightarrow Q Q^T = I \right)$$

i.e. all columns of Q are ~~mutually~~ unit vectors and mutually orthogonal.

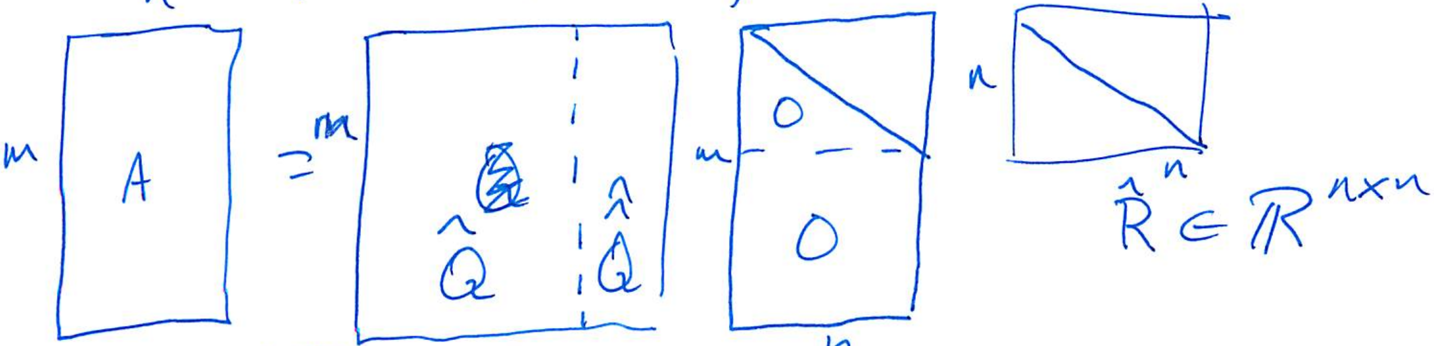
Note
$$\|Qx\|_2 = \left((Qx)^T (Qx) \right)^{1/2} = (x^T \underbrace{Q^T Q}_{= I} x)^{1/2} = \|x\|_2$$

$$\Rightarrow \|Q\|_2 = 1.$$

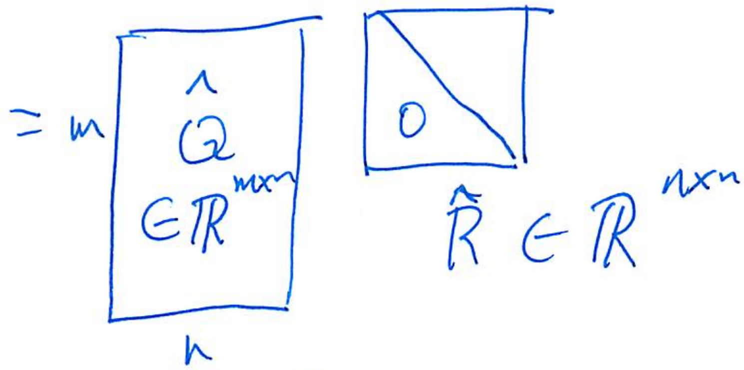
Thm Given $A \in \mathbb{R}^{m \times n}$, $m \geq n$.

Then $\exists$ ~~Q~~ orthogonal $Q \in \mathbb{R}^{m \times m}$ and upper triangular $R \in \mathbb{R}^{m \times n}$ s.t.

$A = QR$; called the QR-decomp



$A \in \mathbb{R}^{m \times n}$ $\qquad Q \in \mathbb{R}^{m \times m}$ $\qquad R \in \mathbb{R}^{m \times n}$



The colums of $\hat{Q}$ remain unit vectors that are mutually orthogonal. Hence

$$\underset{n \times m}{\hat{Q}^T} \underset{m \times n}{\hat{Q}} = I \in \mathbb{R}^{n \times n} \quad (\text{recall } n < m)$$

Thm The solution x_* for the least-squares problem $\min_{x \in \mathbb{R}^n} \|Ax-b\|_2^2$ satisfies

$$\hat{R}x = \hat{Q}^T b \leftarrow$$

requires matrix decomp (still $\sim m^3$) + backsolve.

Pf. $\|Ax-b\|_2^2 = (Ax-b)^T(Ax-b)$

$= (Ax-b)^T Q Q^T (Ax-b) = \|Q^T(Ax-b)\|_2^2$

$\downarrow \quad A = QR, \quad Q^TQ = I$

$= \|Rx - Q^Tb\|_2^2 = \left\| \begin{pmatrix} \hat{R}x \\ \hline 0 \end{pmatrix} - \begin{pmatrix} \hat{Q}^T b \\ \hline c \end{pmatrix} \right\|_2^2$

$= \underbrace{\|\hat{R}x - \hat{Q}^T b\|_2^2} + \|c\|_2^2$ ← indep. of x.

1st term minimized by setting $\boxed{\hat{R}x = \hat{Q}^T b}$

This is well-conditioned ~~to. A setting~~ ~~The condition #.~~ (who sets the condition #? $(\kappa(A^TA))^{1/2}$?)