Theorem: (least squares)

Consider the equation $Ax = b$ with $A \in \mathbb{C}^{m \times n}$ for $m \geq n$. A **least squares solution** is a vector x with $\|Ax - b\|_2$ minimal.

This is equivalent to:

1) $r = Ax - b \in \mathrm{ran}(A)^\perp$

2) $r \in \ker(A^*)$

3) $A^*Ax = A^*b$

4) If P is the orthogonal projection onto $\mathrm{ran}(A)$, then

$$Pb = Ax.$$

If A is rank n,

then A^*A is invertible

and $\boxed{x = (A^*A)^{-1}A^*b}$.

The system

$A^*Ax = A^*b$ is known

as the normal equations

for the data (A, b).

proof:

1) and 2) are equivalent by the lemma from last class.

2) and 3) are equivalent by basic arithmetic:

If $r = Ax - b \in \ker(A^*)$,

then $0 = A^* r = A^*(Ax - b)$

$= A^* A x - A^* b$

Then $A^*Ax = A^*b$,

so $2) \Rightarrow 3)$.

Conversely, if $A^*Ax = A^*b$,

$0 = A^*Ax - A^*b$

$= A^*(Ax-b) \Rightarrow Ax-b \in \ker(A^*)$.

So $3) \Rightarrow 2)$

If we know

$$Ax = Pb, \quad \text{then}$$

$$\begin{aligned} P(Ax - b) &= PAx - Pb \\ &= Ax - Pb \quad \text{(since } P \text{ projects onto } \operatorname{ran}(A)\text{)} \\ &= 0. \end{aligned}$$

This says

$$Ax - b \in \operatorname{ran}(P)^\perp$$
$$= \operatorname{ran}(A)^\perp,$$

which shows 4) $\Rightarrow$ 1).

Conversely, if $Ax - b \in \operatorname{ran}(A)^\perp$,

then
$$0 = P(Ax - b)$$
$$= PAx - Pb$$
$$= Ax - Pb,$$

so $Ax = Pb$ and 1) $\Rightarrow$ 4).

We know 1) – 4) are equivalent.

Now suppose $Ax = Pb$.

We'll show

$\| Ax - b \|_2$ is minimized.

Suppose $z \in \operatorname{ran}(A)$.

$$\| z - b \|_2^2$$

$$= \| z - Pb + Pb - b \|_2^2$$

$$= \|$$

$$\|z-b\|_2^2$$

$$= \|z - Pb + Pb - b\|_2^2$$

$$= \|Pz - Pb + Pb - b\|_2^2$$

$$= \|P(z-b) + (P-I_n)b\|_2^2$$

$$= \|P(z-b)\|_2^2 + \|(P-I_n)b\|_2^2$$

Since $P(z-b)$ and $(P-I_n)b$ are orthogonal.

$$\geq \|(P-I_n)b\|_2^2$$

$$= \| Pb - b \|_2^2$$

$$= \| Ax - b \|_2^2$$

since $Pb = Ax$

This says

$$\| z - b \|_2^2 \geq \| Ax - b \|_2^2$$

for all $z \in \operatorname{ran}(A)$, and

so $Ax = z$ minimizes

$\| z - b \|_2$. enough

Definition : (pseudo inverse)

Let $A \in \mathbb{C}^{m \times n}$, $m \geq n$.

The pseudo inverse of A is the matrix

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

provided $A^{*}A$ is invertible.

Note again that by least squares, $A^T A$ is invertible if and only if $\operatorname{rank}(A) = n$.

Application

Back to polynomial interpolation: use fewer points.

Instead use a 7^{th} degree approximation obtained by truncating the vandermonde matrix to 8 columns. (multiply by eye(11,8))

Use the normal equations. If B is your truncated Vandermonde matrix, then B has rank n.

Then B^*B is invertible, so the coefficients of the 7^{th} degree polynomial are given by solving

$$B^*Bx = B^*y$$

Then

$$x = (B^* B)^{-1} B y$$

$$= B^+ y$$

Example 1: (best-fit line)

Take

$(1,3),(3,-1),(4,8) \in \mathbb{R}^2.$

We want the least-squares solution in the form of a line.

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

Apply least squares to

$$B \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 8 \end{bmatrix}.$$

Equivalent to solving

the normal equations

$$B^* B \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = B^* \begin{bmatrix} 3 \\ -1 \\ 8 \end{bmatrix}.$$

Solve using Matlab

QR Decomposition and Least-Squares

"modern-classical"

Implement solving

$Ax=b$ on a computer.

Assume A has rank n.

Compute the reduced QR factorization

$$A = \hat{Q}\hat{R}.$$

Then

$$\hat{Q}\hat{R}x = b, \text{ so}$$

Then

$$\hat{Q}\hat{R}x = b \text{, so}$$

multiplyling both sides on the left by $(\hat{Q})^*$,

$$(\hat{Q})^* b = \underbrace{(\hat{Q})^* \hat{Q}}_{\text{projection onto ran}(\hat{R})} \hat{R} x$$

$$= \hat{R} x$$

Compute $(\hat{Q})^* b$, and then, since $\hat{R}$ is upper triangular, solve

$$\hat{R} x = (\hat{Q})^* b$$

by back substitution, starting with the last row.