Announcements

1) Last job candidate tomorrow, talk at 3-4 in CB 2090

2) Apply for scholarships (mainly for undergrad)

Least-Squares
or
Solving $Ax = b$

Recall: Vandermonde Matrices and Polynomial Interpolation:

given n points in $\mathbb{R}^2$ with distinct x-coordinates, $\exists$ a degree $(n-1)$ polynomial that passes through (interpolates) the points.

Sometimes the oscillations of the polynomial are a bit too wild!

Example 1: (not a good fit)

$$v = [-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5]$$

$$y = [0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0]$$

$$A = \text{fliplr}(\text{vander}(v))$$

Solve: $A \backslash y'$

Plot the resulting polynomial against the data

Fix: Reduce degree.

Penalty: the polynomial no longer passes exactly through the given points, but the oscillations are tamped down.

We'll be "solving" the system

$Ax = b$ where b is the desired output and is an m-vector, x is an n-vector, $A \in \mathbb{C}^{m \times n}$, $m \geq n$.

If $A \in \mathbb{C}^{m\times m}$ is invertible, the we simply apply A^{-1} to find x.

But if A is not invertible, there may be no solution!

Example 2: $A = \begin{bmatrix} 2 & -6 \\ 1 & -3 \end{bmatrix}$

$$b = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

Find $\begin{bmatrix} x \\ y \end{bmatrix}$ with

$$A \begin{bmatrix} x \\ y \end{bmatrix} = b.$$

However,

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - 6y \\ x - 3y \end{bmatrix}$$

However,

$$A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x-6y \\ x-3y \end{bmatrix}$$

$$= x\begin{bmatrix} 2 \\ 1 \end{bmatrix} - y\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= x\begin{bmatrix} 2 \\ 1 \end{bmatrix} - 3y\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= (x-3y)\begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

So $b \in \operatorname{ran}(A)$ if and only if $b = c\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

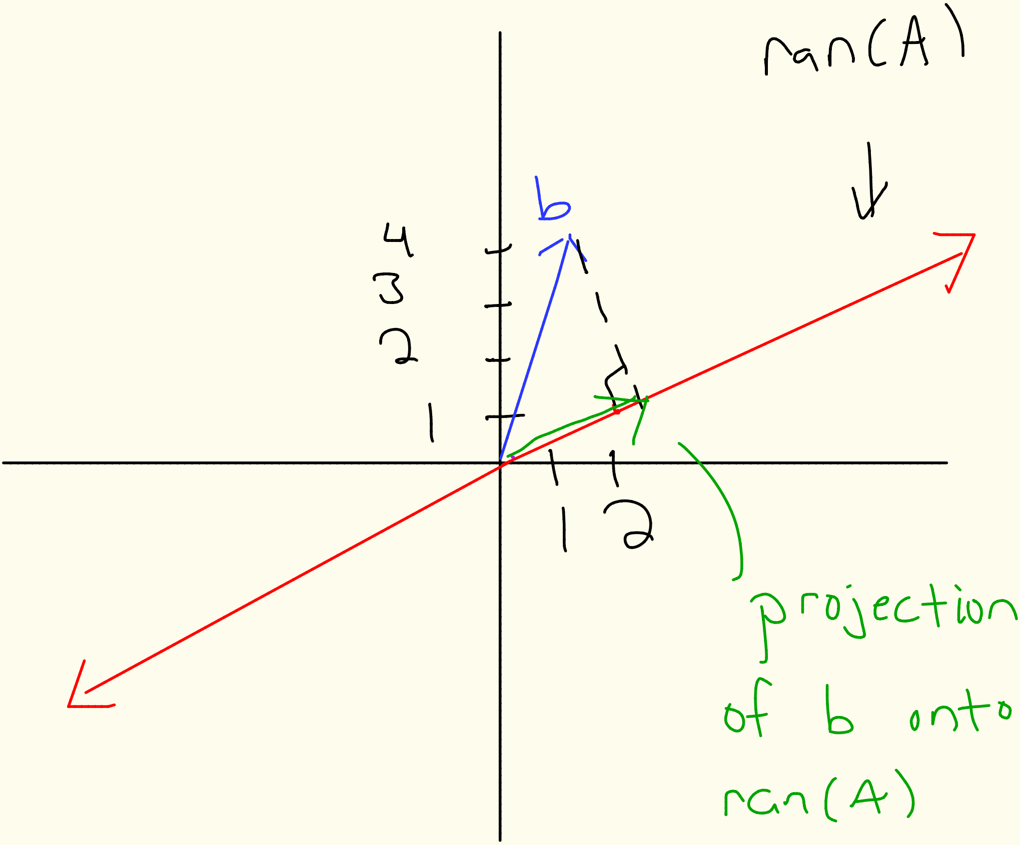
for some constant c.

But our $b = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

which is not a scalar multiple of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$!

Therefore, $Ax = b$ has no solution.

Picture



If v = projection of b onto ran(A), $\|v-b\|_2$ is minimized.

We'll reduce to finding a vector x with $Ax = v$. There will be infinitely many x values, but only one v! Indeed, if $y \in \ker(A)$, $A(x+y) = v$. Here, $\ker(A)$ is one-dimensional.

Lemma: $(\ker(A^*) = (\operatorname{ran}(A))^{\perp})$

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

$$\ker(A^*) = (\operatorname{ran}(A))^{\perp}.$$

Proof: Show inclusion both ways

1) Show $\ker(A^*) \subseteq (\operatorname{ran}(A))^{\perp}$

Let $x \in \ker(A^*)$, $y \in \mathbb{C}^n$.

Then

$$(Ay)^* x$$

$$= (y^* A^*) x$$

$$= y^* (A^* x)$$

$$= 0$$ since $x \in \ker(A^*)$, so $A^* x = 0_{\mathbb{C}^n}$

Therefore, x is orthogonal to $\operatorname{ran}(A)$, and so is in $(\operatorname{ran}(A))^{\perp}$.

2) $\operatorname{ran}(A)^\perp \subseteq \ker(A^*)$

Let $x \in \operatorname{ran}(A)^\perp$.

Then

$$\|A^*x\|_2^2 = (A^*x)^* A^* x$$

$$= (x^* A) A^* x$$

$$= x^* \big(\underbrace{A A^* x}_{\in\, \operatorname{ran}(A)}\big)$$

$$= 0 \quad \text{since } x \in \operatorname{ran}(A)^\perp.$$

This shows that

$x \in \ker(A^*)$ and

so the proof is

complete: we have

$$\operatorname{ran}(A)^\perp \subseteq \ker(A^*)$$

and

$$\ker(A^*) \subseteq \operatorname{ran}(A)^\perp,$$

so $\operatorname{ran}(A)^\perp = \ker(A^*)$

$\square$

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

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

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

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

$$Pb = Ax.$$

If A is rank n,

then A^*A is invertible

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

The system

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

at the normal equations

for the data (A, b).