Math 2270-1
Friday 21 Oct.
Meet in the lab room, LCB 115 on Monday... honest!

HW for Fri 10/28
§5.4 1, 2, 3, 5, 21, 22, 23, 25
31, 32, 39

"Least squares solutions" to inconsistent systems

- pages 4, 5 Wed notes

<u>Theorem</u>: The least squares solution $\vec{x}$ to $A^T A \vec{x} = A^T \vec{b}$ is unique, provided the ~~ct~~ columns of $A$ are linearly independent.

<u>proof</u>: If $A_{m \times n}$ then $A^T_{n\times m} A_{m \times n}$ is an $n \times n$ square matrix, and least-squares sol's will be unique iff this matrix is invertible, iff rref ( ) = I, iff the only homogeneous solution

$$(A^T A)\vec{x} = \vec{0}$$

is $\vec{x} = \vec{0}$.

But if $A^T A \vec{x} = \vec{0}$

then $(x^T A^T)(A\vec{x}) = x^T \vec{0} = 0$

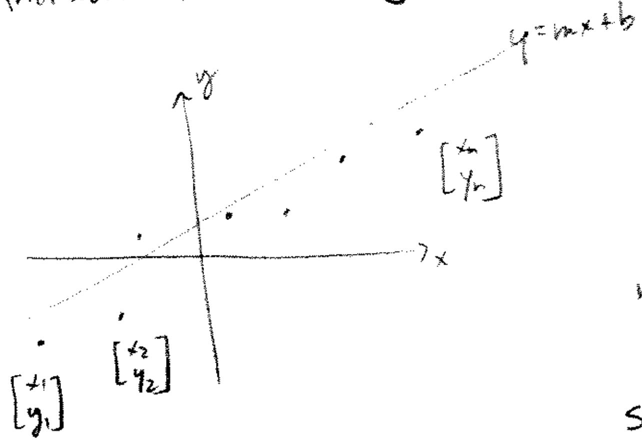
$\Leftrightarrow (A\vec{x})\cdot(A\vec{x}) = 0$

$\Leftrightarrow A\vec{x} = \vec{0}$

$\Leftrightarrow \vec{x} = \vec{0}$ since the columns of $A$ are independent!

<u>example</u>: see pages 4-5 Wed.

And now for something completely different (yet completely the same)




seek $m, b$ to fit $n$ data pts $\left\{ \begin{bmatrix} x_i \\ y_i \end{bmatrix} \right\}$

Ideally $y_i = m x_i + b$

$$m \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

System (for $m$ & $b$) is probably inconsistent

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

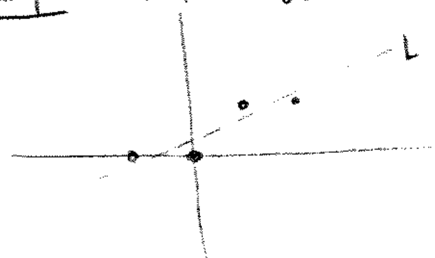
The least squares soltn solves

$$A^T A \begin{bmatrix} m \\ b \end{bmatrix} = A^T \vec{y}$$

$$** \quad \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

and minimizes $\left\| m \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right\|^2 = \sum_{i=1}^{n} (m x_i + b - y_i)^2$

the sum of the squared vertical deviations example 4 pts $\left\{ \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$




$$\begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} m \\ b \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2/5 \\ 3/10 \end{bmatrix}$$

$$\boxed{y = \frac{2}{5} x + \frac{3}{10}}$$

Why stop at straight-line approx?

example : Find the best parabolic fit, $p(t) = c_0 + c_1 t + c_2 t^2$ to the following data points

$\{ [-2,33], [-1,13], [1,3], [2,1], [4,20] \}$

We want

$$\begin{aligned} c_0 - 2c_1 + 4c_2 &= 33 \\ c_0 - c_1 + c_2 &= 13 \\ c_0 + c_1 + c_2 &= 3 \\ c_0 + 2c_1 + 4c_2 &= 1 \\ c_0 + 4c_1 + 16c_2 &= 20 \end{aligned}$$

$$\begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 33 \\ 13 \\ 3 \\ 1 \\ 20 \end{bmatrix}$$

```
> with(linalg):
Warning, the protected names norm and trace have been redefined and unprotected
> A:=matrix(5,3,[1,-2,4,1,-1,1,1,1,1,1,2,4,1,4,16]);
```

$$A := \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 16 \end{bmatrix}$$

```
> b:=vector([33,13,3,1,20]);
```

$$b := [33, 13, 3, 1, 20]$$

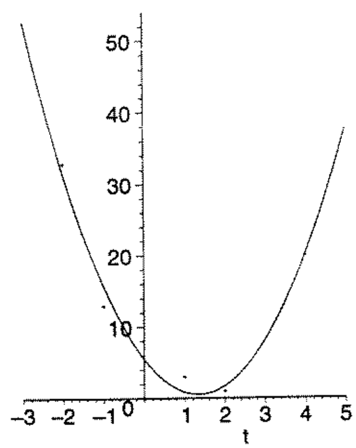
```
> linsolve(transpose(A)&*A,transpose(A)&*b);
```

$$\left[\frac{182}{33}, \frac{-1723}{231}, \frac{214}{77}\right]$$

```
> evalf(%);
```

$$[5.515151515, -7.458874459, 2.779220779]$$

```
> with(plots):
Warning, the name changecoords has been redefined
> points:=pointplot({[-2,33],[-1,13],[1,3],[2,1],[4,20]}):
  curve:=plot(5.515151515 -7.458874459*t + 2.779220779*t^2,
    t=-3..5,color=black):
  display({points,curve});
```




```
>
```

Additional §5.4 topic:

The 4 fundamental subpaces for a linear transformation

$L(\vec{x}) = A\vec{x}$, $L: \mathbb{R}^n \to \mathbb{R}^m$.

- $\ker(A) := \{\vec{x} \in \text{domain s.t. } A\vec{x} = \vec{0}\} = \{\vec{x} \in \text{domain s.t. } \vec{x} \perp \text{ each row of } A\}$
  $= (\text{span}\{\text{rows of } A\})^{\perp}$
  $:= (\text{rowspace}(A))^{\perp}$
- $\text{Image}(A) := \{\vec{y} \in \text{codomain s.t. } \exists\, \vec{x} \text{ with } A\vec{x} = \vec{y}\}$
- $\ker(A)^{\perp} = \{\vec{z} \in \text{domain s.t. } \vec{z} \cdot \vec{x} = 0 \ \ \forall \vec{x} \in \ker(A)\}$
  $= ((\text{rowspace}(A))^{\perp})^{\perp} = \text{rowspace}(A)$
- $(\text{Image}(A))^{\perp} = (\text{span}(\text{cols}(A)))^{\perp} = \ker(A^T)$

<u>Example</u> $L : \mathbb{R}^3 \to \mathbb{R}^3$

$$L(\vec{x}) = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & -3 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ -1 & 1 & -3 \end{bmatrix}$$

reduced column echelon form!, $= (\mathrm{rref}(A^T))^T$ !

$$\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathrm{rref}(A^T) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathrm{rcef}(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -2 & 0 \end{bmatrix}$$

$A\vec{x} = \vec{0}$: $x_3 = t$, $x_2 = -t$, $x_1 = 3t$

$A^T\vec{y} = \vec{0}$: $y_3 = t$, $y_2 = 2t$, $y_1 = -t$

$$\ker(A^T) = \mathrm{span}\left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$$

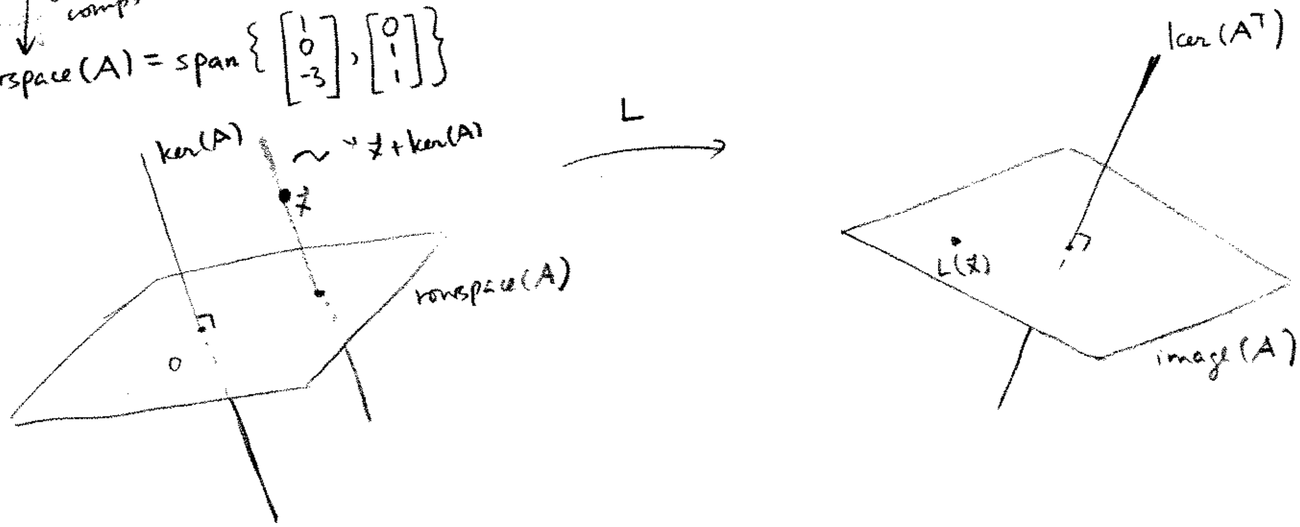
orthog comps!

$$\mathrm{image}(A) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\} = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} \right\}$$

$$\ker(A) = \mathrm{span}\left\{ \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \right\}$$

orthog comps!

$$\mathrm{rowspace}(A) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$




$$L(\vec{x} + \ker(A)) = L(\vec{x})$$