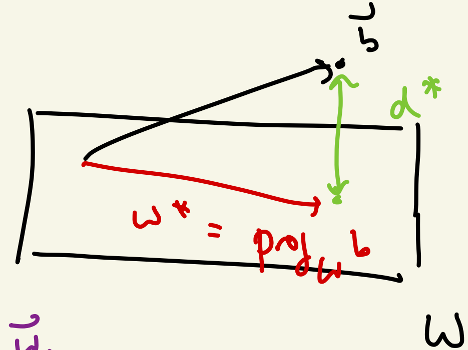
Last time

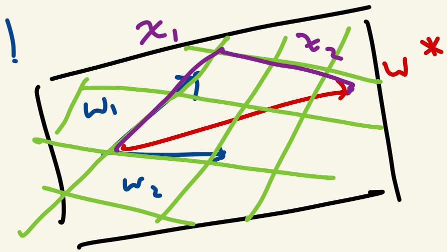


1. Find a basis $w_1 \dots w_k$
2. $A = (w_1 \dots w_k)$
3. $x^* = (A^TA)^{-1}A^Tb$

not closest pt!

$w^* = x_1\vec{w}_1 + x_2\vec{w}_2$

4. $w^* = Ax^* = A(A^TA)^{-1}A^Tb$ Closest point!



5. $d^* = \sqrt{\|b\|^2 - x^{*T}f} = \sqrt{\|b\|^2 - x^{*T}(A^Tb)}$

actual distance

Thm All along $w^* = \text{proj}_W b$

$P = A(A^TA)^{-1}A^T$ is called a projection matrix.

Today Suppose we have a system of equations

$$A\vec{x} = \vec{b}.$$

Maybe this system is inconsistent, i.e. no solution!

What vector is closest to being a solution?

Def A **least squares** solution of $A\vec{x} = \vec{b}$ is a vector x^* such that

$$\|Ax^* - b\|^2 = \min_{\vec{x} \in \mathbb{R}^n} \|Ax - b\|^2.$$

Thm Suppose $\ker(A) = 0$. Then there is a unique least squares solution

$$x^* = (A^TA)^{-1}A^Tb.$$

Pf Consider the subspace $W = \operatorname{img}(A) = \{ w = Ax \mid \tilde{x} \}$

$\min \|Ax - b\|^2 = \min \|w - b\|^2$ where $w \in \operatorname{img}(A)$.

Therefore $Ax^* = w^*$ is the closest pt on $\operatorname{img}(A)$ to $\tilde{b}$ and so x^* would be the least squares solution.

Note that if $\ker(A) \neq 0$, suppose $0 \neq z \in \ker(A)$.

Then $\tilde{x} = x^* + z$ is also a least squares solution

$Az = 0 \iff z \in \ker(A)$

$$\|A\tilde{x} - b\|^2 = \|A(x^* + z) - b\|^2$$

$$= \|Ax^* + \cancel{A}z - b\|^2$$

$$= \|Ax^* - b\|^2 = \min \|Ax - b\|^2$$

$\tilde{x}, x^*$ are both LS sol's.

So $\ker(A) = 0$ is necessary for uniqueness of LS soln.

<u>Ex</u> $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. This is an inconsistent system!

$$1x + 0y = 1 \longrightarrow x = 1$$

$$0x + 1y = 2 \longrightarrow y = 2$$

$$x + 3y = 1 \longrightarrow 1 + 2 \cdot 3 = 7 \neq 1$$ inconsistent!

Which vector $\begin{pmatrix} x \\ y \end{pmatrix}$ is closest by distance to being a solution to this system? Which $\begin{pmatrix} x \\ y \end{pmatrix}$ minimizes this expression?

$$\min \left\| \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\|^2$$

By least squares formula

$$x^* = (A^TA)^{-1}A^T\vec{b} \qquad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$$A^TA = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 10 \end{pmatrix} \qquad (A^TA)^{-1} = \frac{1}{11}\begin{pmatrix} 10 & -3 \\ -3 & 2 \end{pmatrix}$$

$$A^Tb = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

$$x^* = \frac{1}{11}\begin{pmatrix} 10 & -3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 5 \end{pmatrix} = \frac{1}{11}\begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

$x^* = \begin{pmatrix} 5/11 \\ 4/11 \end{pmatrix}$ is the least square solution to

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 5/11 \\ 4/11 \end{pmatrix} = \begin{pmatrix} 5/11 \\ 4/11 \\ 17/11 \end{pmatrix}$ is closest to $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ that we could get!

Application!

Suppose we've taken some data

$$(t_1, y_1), \ldots, (t_n, y_n). \qquad n \gg 0$$

Suppose we expect a linear relationship

$$y = \alpha t + \beta.$$

Error: $e_i = y_i - (\alpha t_i + \beta)$

we find α, β which minimizes

$$e_1^2 + e_2^2 + \cdots + e_n^2 \,?$$

actual

expected

$$\vec{e} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} y_1 - (\alpha t_1 + \beta) \\ \vdots \\ y_n - (\alpha t_n + \beta) \end{pmatrix} \qquad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} \alpha t_1 + \beta \\ \vdots \\ \alpha t_n + \beta \end{pmatrix}$$

$$= \vec{y} - \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \vec{y} - A\vec{x}$$

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \qquad x = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$$

Can we find α, β which minimizes

$$e_1^2 + e_2^2 + \dots + e_n^2 = \|\vec{e}\|^2$$

$$= \|\vec{y} - A\vec{x}\|^2$$

So $x = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$ which minimizes error satisfies

$$\min \|\vec{y} - A\vec{x}\|^2 \;!$$

$x = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$ is the least squares solution to $\underline{A\begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \vec{y}}$

$(t_1, y_1) \;\ldots\; (t_n\ y_n)$ $\quad \underline{\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}, \quad \underline{A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}}$

So least squares line can be computed by

$$\underline{x^* = (A^T A)^{-1} A^T \vec{y}}.$$

Ex $(0,2), (1,3), (3,7), (6,12)$

$t = 0, 1, 3, 6$ time

$y = 2, 3, 7, 12$ height

Expected $y = \alpha t + \beta$

Which line $y = \alpha t + \beta$ fits the data w/ least squared error?

$$\begin{pmatrix} \beta \\ \alpha \end{pmatrix}^* = (A^T A)^{-1} A^T \vec{y} \quad \text{where} \quad A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \\ 1 & t_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{pmatrix}$$

$$\vec{y} = \begin{pmatrix} 2 \\ 3 \\ 7 \\ 12 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 6 \end{pmatrix}$$

$$= \begin{pmatrix} 4 & 10 \\ 10 & 46 \end{pmatrix} \qquad (A^T A)^{-1} = \frac{1}{84} \begin{pmatrix} 46 & -10 \\ -10 & 4 \end{pmatrix}$$

$$A^T y = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 6 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 7 \\ 12 \end{pmatrix} = \begin{pmatrix} 24 \\ 96 \end{pmatrix}$$

$$\begin{pmatrix} \beta \\ \alpha \end{pmatrix}^* = (A^TA)^{-1} A^T y = \frac{1}{84}\begin{pmatrix} 46 & -10 \\ -10 & 4 \end{pmatrix}\begin{pmatrix} 24 \\ 96 \end{pmatrix} = \frac{1}{84}\begin{pmatrix} 144 \\ 144 \end{pmatrix}$$

$$= \begin{pmatrix} 12/7 \\ 12/7 \end{pmatrix}$$

The best fit line is $y = \frac{12}{7} t + \frac{12}{7}$!

$$y = \underline{(\alpha_0 + \alpha_1 t)} + \alpha_2 t^2 + \cdots + \alpha_k t^k$$

numbers!

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & & t_1^k \\ 1 & t_2 & t_2^2 & & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & t_n & t_n^2 & & t_n^k \end{pmatrix} \qquad \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_k \end{pmatrix} = (A^TA)^{-1} A^T y$$

$(x, y)(z)$

Input (x, y) 2D pos

measured z height $z = \alpha x + \beta y + \gamma$

$$e_i = z_i - (\alpha x_i + \beta y_i + \gamma)$$

$z = x^\alpha$

$\ln(z) = \alpha \ln(x)$

$$\vec{e} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} - \begin{pmatrix} 1 & x_i & y_i \\ \vdots & & \\ 1 & x_n & y_n \end{pmatrix} \begin{pmatrix} \gamma \\ \alpha \\ \beta \end{pmatrix}$$

Suppose we wanted to compute $(A^TA)^{-1}A^T$ quickly.

$A = QR$ A is not square! $A = QR$ isn't quite the same

$$A = \begin{pmatrix} u_1 \cdots u_k \end{pmatrix} \begin{pmatrix} r_{11} & & * \\ & \ddots & r_{kk} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(A^TA)^{-1}A^T = \left((QR)^T QR\right)^{-1} A^T$$

$$= \left(R^T \overset{I}{\cancel{Q^T}} Q R\right) A^T$$

$$= (R^TR)^{-1} A^T \qquad R = \begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ much easier!

Can we compute QR quickly? ✓ $\square$

HW 8 due tonight!

HW 9 due 12/4 but I'll still post it later today.

$$x + 2y - z = 0$$

$$\begin{pmatrix} 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

$$\text{plane} = \ker \begin{array}{ccc} & \text{free} & \text{free} \\ \left(1 \right. & 2 & \left. -1 \right) \\ \text{leading } 1 & & \end{array}$$

x determined y, z free

$$x = -2y + z$$

$$\text{plane} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2y + z \\ y \\ z \end{pmatrix} = \underline{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}} y + \underline{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}} z$$

basis

$$W \subseteq \mathbb{R}^3 \qquad W = \mathbb{R}^3$$