Q: what is the matrix P corresponding to projection?

$$P: x \longmapsto a\,\frac{a\cdot x}{a\cdot a} = \frac{a\,a^T}{a\cdot a}\,x \quad \text{so} \quad P = \frac{1}{a\cdot a}\,aa^T \qquad \underbrace{\underbrace{a}_{n\times 1}\,\underbrace{a^T}_{1\times n}}_{n\times n}$$

Example projection onto $\begin{bmatrix}1\\1\\1\end{bmatrix}$ $\quad P = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1&1&1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$

Properties of projections ① P is symmetric: $P^T = \left(\frac{aa^T}{a\cdot a}\right)^T = \frac{(a^T)^T a^T}{a\cdot a} = \frac{aa^T}{a\cdot a} = P.$

② $P = P^2$.

⌈ Remark Q: is there a matrix B s.t. $(Ax)\cdot y = x\cdot(By)$ for all x, y?

A: yes $B = A^T$: $\quad (Ax)^T y = x^T A^T y = x^T(A^T y) = x\cdot(Ay)$. ⌋

§ Projections and least squares

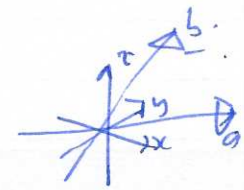
Lots of equations in one variable: $ax = b$, e.g. $\begin{matrix}2x = b_1\\ 3x = b_2\\ 4x = b_3\end{matrix}$ $\quad Ax = b.$

probably inconsistent, but can look for a solution $\hat{x}$ which minimizes the error squared $\quad$ ~~[2 3 4]~~ $\begin{bmatrix}2\\3\\4\end{bmatrix}x = \begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$ write $A = a$

$$E^2 = (2x-b_1)^2 + (3x-b_2)^2 + (4x-b_3)^2 = \|ax - b\|^2$$

$$\frac{d}{dx}(E^2) = 2(2x-b_1).2 + 2(3x-b_2).3 + 2(4x-b_3).4 = 0$$

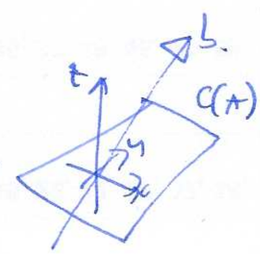
$$\hat{x} = \frac{2b_1 + 3b_2 + 4b_3}{2^2+3^2+4^2} = \frac{a.b}{a.a}$$



General case (1 var) $ax = b$ has least squares solution $\hat{x} = \frac{a.b}{a.a}$.

note error vector is $b - \hat{x}a$ is perpendicular to a: $a.(b - \hat{x}a)$

$= a.b - \frac{a.b}{a.a}\,a.a = 0.$



Several variables $\quad Ax = b \quad$ with #equations $\gg$ #variables.
$\qquad\qquad\qquad\qquad\qquad\qquad m \gg n$.

(probably inconsistent)

find $\hat{x}$ to minimize least squares error $E^2 = \|Ax - b\|^2$ $\quad = \mathrm{col}(A)$

Fact: this is minimized when error ($b - A\hat{x}$) is perpendicular to $C(A)$

<u>recall</u> $C(A)^\perp = N(A^T)$ so $A^T(b - A\hat{x}) = 0$ ~ $\boxed{A^TAx = A^Tb}$.
normal equations.

<u>check</u> $E^2 = \|Ax-b\|^2 = (Ax-b)\cdot(Ax-b) = (Ax-b)^T(Ax-b)$

$$E^2 = \cancel{Ax}\ x^TA^TAx - b^TAx - \cancel{b^TAx} + b^Tb \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$
x^TAb $\quad - b^TA \frac{\partial}{\partial x_i}(x) -$

$$\frac{\partial}{\partial x_i}(E^2) = \frac{\partial}{\partial x_i}(x^T)A^TAx + x^TA^TA\frac{\partial}{\partial x_i}(x) - \cancel{b^TA}\frac{\partial}{\partial x_i}(x^T)Ab = 0.$$

$\frac{\partial}{\partial x_i}(x) = e_i$ $\frac{\partial}{\partial x_i}(x^T) = e_i^T$ so get:
$$\begin{matrix} e_i^T A^TAx + x^TA^TAe_i \\ e_i^T Ab \quad - b^TAe_i \end{matrix} = 0$$

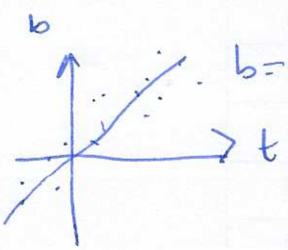
so $e_i^T(A^TAx - A^Tb) + (x^TA^TA - b^TA)e_i = 0$

$\uparrow$ $A^TAx = A^Tb \Rightarrow \frac{\partial}{\partial x_i}(E^2) = 0$ for all x_i $\Rightarrow$ critical point. $\square$.

<u>Remarks</u> • if columns of A are linearly independent, then A^{-1} exists and $\hat{x} = (A^TA)^{-1}A^Tb$

• $A\hat{x}$ = projection of b onto $\mathrm{col}(A)$, i.e. $A\hat{x} = A(A^TA)^{-1}Ab$.

<u>Example</u> Best fit straight line. data points (t_i, b_i)



$b = C + Dt$

$$\begin{matrix} C + Dt_1 = b_1 \\ C + Dt_2 = b_2 \\ \vdots \\ C + Dt_m = b_m \end{matrix} \qquad \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \\ 1 & t_n \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{bmatrix}$$

best solution $\hat{x} = \begin{pmatrix} \hat{C} \\ \hat{D} \end{pmatrix}$ minimizes $E^2 = \|b - Ax\|^2$ $\quad Ax = b$

$= \sum_{i=1}^{n} (b_i - C - Dt_i)^2$, i.e. $A^TA\begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = A^Tb$

$$\sim \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_n \end{bmatrix}\begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}\begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_n \end{bmatrix}\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$\begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix}\begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

Eigenvalues

recall: eigenvalues are solutions to $\det(A-\lambda I)=0$

algebraic multiplicity = # of times $(\lambda-\lambda_i)$ appears in $\det(A-\lambda I)$.

characteristic equation

geometric multiplicity = dim of solutions to $(A-\lambda_i I)x=0$
(eigenspace)

Example $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ $\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$

$\lambda=1$ has alg. multiplicity $=2$

solve $(A-I)x=0$ $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $x_1=t$, $x_2=0$ solution $\left\{t\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$.

free var

dim of eigenspace $= 1 < 2$.

Example $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ $\det(A-\lambda I) = (1-\lambda)^2(2-\lambda)$

$\lambda = 1, 1, 2$.

eigenvectors for $\lambda=1$ $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ ← two dimensional eigenspace.

$\lambda=2$ $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Special case: triangular matrices $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ $\det(A-\lambda I) = (1-\lambda)(4-\lambda)(6-\lambda)$

eigenvalues = diagonal entries.

useful facts. $\det(A) =$ product of eigenvalues $\lambda_1\lambda_2\cdots\lambda_n$

$\operatorname{trace}(A) = a_{11}+a_{22}+\cdots+a_{nn} = \lambda_1+\lambda_2+\cdots+\lambda_n$ sum of eigenvalues.

Warning: row operations do not preserve eigenvalues!

$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ ← different eigenvalues!

Example rotation by $\pi/2$: $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ $\det(A-\lambda) = \lambda^2+1$ (58)

$\lambda = \pm i$

find eigenvectors: $\lambda = i$: $\begin{bmatrix} -i & -1 \\ 1 & i \end{bmatrix} \rightsquigarrow \begin{bmatrix} -i & -1 \\ 0 & 0 \end{bmatrix}$ $v = t\begin{bmatrix} 1 \\ -i \end{bmatrix}$

t

check: $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix}$. ✓.

Example: projection matrix $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ eigenvalues: $(\frac{1}{2}-\lambda)^2 - \frac{1}{4} = 0$

$\lambda = 1 \quad \lambda = 0$ $\qquad \lambda^2 - \lambda = 0$

eigenvectors: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $\qquad \lambda(\lambda-1) = 0$

$\lambda = 0, 1$

Matrix powers $A^k = \underbrace{A.A.\cdots A}_{k \text{ times}}$.

Prop^n^ If v is an eigenvector for A with eigenvalue λ then v A^k λ^k.

check $Av = \lambda v$ $\rightsquigarrow$ $A^k v = A^{k-1}Av = A^{k-1}\lambda v = \lambda A^{k-2}Av = \lambda^2 A^{k-2} v$
$= \cdots = \lambda^k v$.

Prop^n^ if $S^{-1}AS = D$ then $S^{-1}A^kS = D^k$

$A = SDS^{-1}$ $\qquad A^k = SD^kS^{-1}$

Proof $A^k = (SDS^{-1})^k = \underbrace{SD\overset{I}{\cancel{S^{-1}S}}D\overset{I}{(S^{-1}\cdot)}\cdot\overset{I}{(\cdot S)}DS^{-1}}_{k-\text{tim}} = SD^kS^{-1}$. □.

Prop^n^ If λ is an eigenvalue of A, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.

Proof $Av = \lambda v$ $\quad v = \lambda A^{-1}v$ $\quad \frac{1}{\lambda}v = A^{-1}v$ ✓ □.