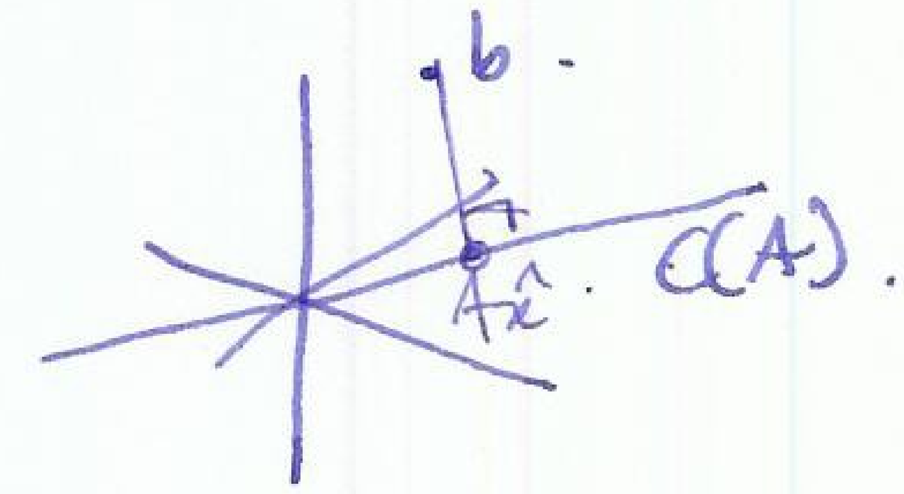
3.3 Projections and least squares

Lots of equations in one variable $ax=b$, e.g.

$$2x=b_1 \qquad ax=b.$$
$$3x=b_2$$
$$4x=b_3.$$

probably inconsistent, but can look for solution $\hat{x}$ which minimizes the error squared:

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



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

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

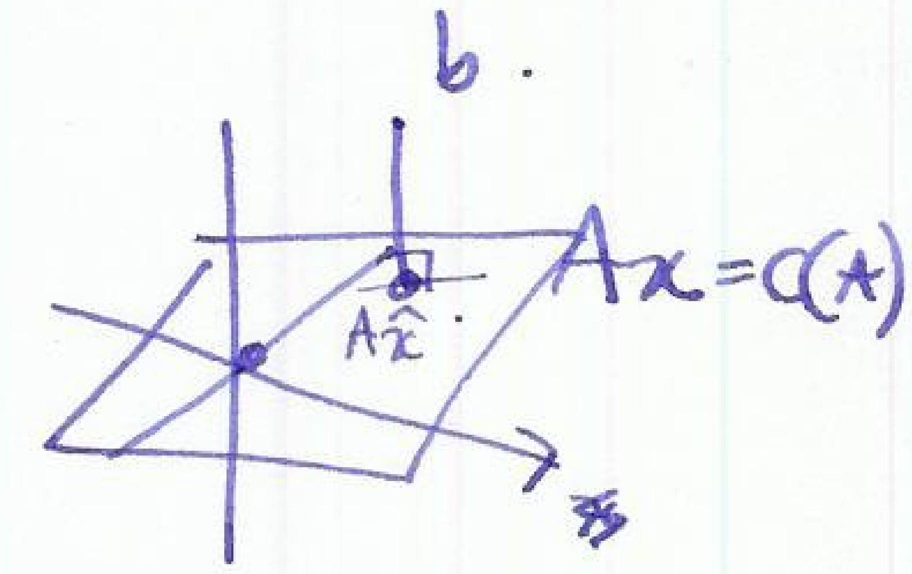
General case (1 var) $ax=b$ has least-squares solution $\hat{x}=\dfrac{a^Tb}{a^Ta}$.

note: $a^T(b-\hat{x}a) = a^Tb - \dfrac{a^Tb}{a^Ta}\cdot a^Ta = 0$

↑ error vector perpendicular to a.

Several variables

$\underset{m\times n}{A}x=b$ with $m>>n$ (probably inconsistent)



find $\hat{x}$ to minimize least squares error $E^2=\|Ax-b\|^2$

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

$\mathbb{R}^m$.

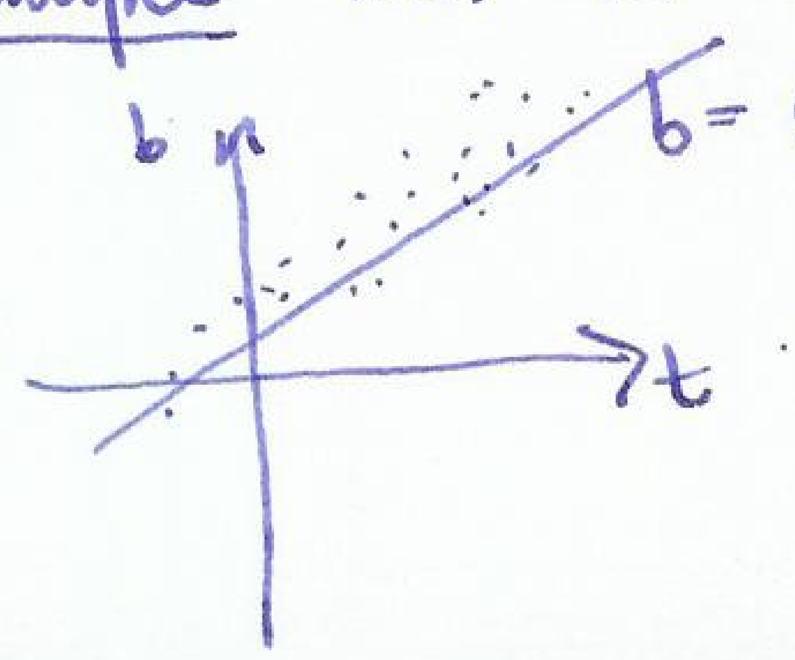
recall $C(A)^\perp = N(A^T)$ so $\boxed{A^T(b-A\hat{x})=0}$ or $\boxed{A^TA\hat{x}=A^Tb.}$

Note (3L) • normal equations: $A^TA\hat{x}=A^Tb$

• If cols of A are linearly independent, then $(A^TA)^{-1}$ exists and $\hat{x} = (A^TA)^{-1}A^Tb$.

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

Example Best fit straight line



data points (t_i, b_i)

$$C+Dt_1=b_1$$
$$C+Dt_2=b_2$$
$$\vdots$$
$$C+Dt_m=b_m$$

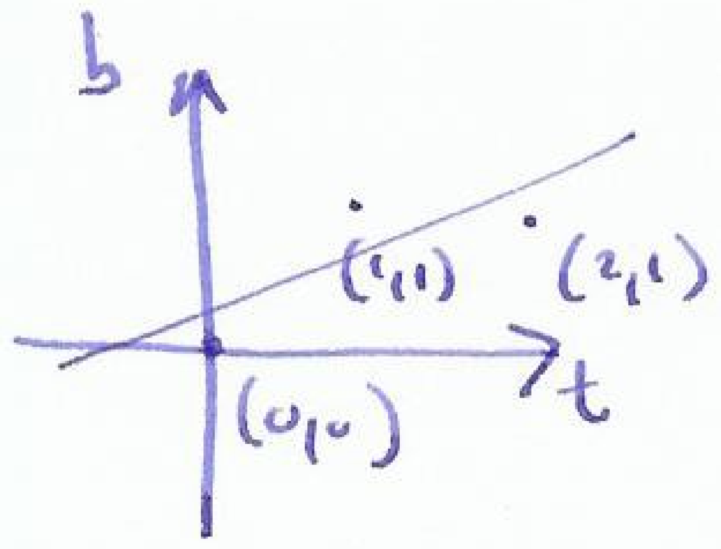
$$\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \qquad Ax = b$$

best solution $\hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$ minimizes $E^2 = \|b - Ax\|^2 = \sum_{i=1}^{m} (b_i - C - Dt_i)^2$

i.e. $A^T A \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = A^T b \quad \sim \quad \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_m \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$.

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

Example



$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

$$\left[\begin{array}{cc|c} 3 & 3 & 2 \\ 0 & 2 & 1 \end{array}\right] \qquad \hat{D} = \frac{1}{2} \quad \hat{C} = \frac{1}{6}$$

§3.4 Orthogonal bases and Gram-Schmidt

recall orthonormal basis $\{v_1, \ldots, v_n\}$ $\quad v_i \cdot v_j = 0 \quad i \neq j \qquad \|v_i\| = 1$ for all i.

Example: $\{e_1, e_2, e_3\}$.

Def^n^ We say a square matrix $\underset{n \times n}{Q}$ is orthogonal if the columns of Q are an orthonormal basis for $\mathbb{R}^n$.

Useful fact $Q^T Q = I$

$$\begin{bmatrix} - & q_1^T & - \\ - & q_2^T & - \\ & \vdots & \\ - & q_n^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots 1 \end{bmatrix} = I_n$$

Example ① $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad Q^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$.

② Any permutation matrix.

<u>Fact</u> Q preserves lengths, i.e. $\|Qx\| = \|x\|$ for any x

<u>Proof</u> $\|Qx\|^2 = (Qx)^T Qx = x^T Q^T Q x = x^T I x = x^T x = \|x\|^2$. $\square$.

- How to write a vector b as a sum of the q_i

$$b = c_1 q_1 + c_2 q_2 + \cdots + c_n q_n.$$

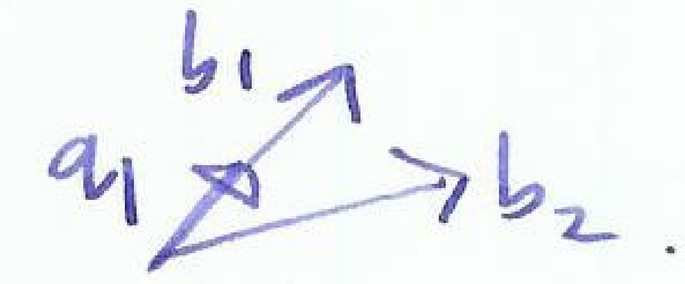
$$q^T_i b = q^T_i(\text{——}) = c_1 \underbrace{q^T_i q_1}_{0} + \cdots + c_i \underbrace{q^T_i q_i}_{1} + \cdots + c_n \underbrace{q^T_i q_n}_{0} = c_i$$

so $\quad c_i = q^T_i b = b \cdot q_i$

i.e. $\quad b = (b\cdot q_1) q_1 + (b \cdot q_2) + \cdots + (b \cdot q_n) q_n.$

<u>Gram-Schmidt</u>: take any basis $\{b_1, b_2, \ldots, b_n\}$ and produce an orthonormal basis $\{q_1 \ldots q_n\}$

① set $\quad q_1 = \dfrac{b_1}{\|b_1\|}$



② set $\quad q_2 = \dfrac{b_2 - q_1^T b\, q_1}{\|\text{length}\|}$.

need to subtract off component of b_2 in direction q_1, i.e. projection of b_2 onto q_1, i.e. $\dfrac{b_2 \cdot q_1 q_1}{q_1 \cdot q_1} = b_2 \cdot q_1 \cdot q_1$

③ set $\quad q_3 = \dfrac{b_3 - \text{components in directions } q_1, q_2}{\|\text{length}\|}$

$$q_3 = \frac{b_3 - (b_3 \cdot q_1) q_1 - (b_3 \cdot q_2) q_2}{\|b_3 - (b_3 \cdot q_1) q_1 - (b_3 \cdot q_2) q_2\|}.$$

ek.

<u>Example</u> $\left\{ \underset{b_1}{\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}}, \underset{b_2}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}, \underset{b_3}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}} \right\}$ $\quad q_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. $\quad q'_2 = \overset{b_2}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} - (b_2 \cdot q_1) q_1$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

$q_2 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. $\quad q'_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \left\{ \dfrac{1}{\sqrt{2}} \dfrac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1/2 \\ 1/2 \end{bmatrix}. \quad q_3 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.\right.$