LEAST SQUARES METHOD

Solving $Ax=b$ when $b\notin C(A)=W$

$\Leftrightarrow$ Solve $Ax=p_W(b)$ instead

How to avoid the projection?

$b\notin W$

Theorem: Let $S=\{x\in\mathbb{R}^n \mid Ax=p_W(b)\}$. Then $S=S'$.

$$S'=\{x\in\mathbb{R}^n \mid A^TAx=A^Tb\}.$$

Proof: we already know: $b-p_W(b)\in Null(A^T)$

(as $(b-p_W(b))\perp W$, and $W=C(A)=R(A^T)$)

$\subseteq$: $x\in S$: $Ax=p_W(b)\Rightarrow A^T(Ax-p_W(b))=0\Leftrightarrow Ax-p_W(b)\in Null(A^T)\Leftrightarrow$

$\Leftrightarrow \underbrace{Ax-p_W(b)+(p_W(b)-b)}_{=Ax-b}\in Null(A^T)\Leftrightarrow A^T(Ax-b)=0\Leftrightarrow x\in S'$.

$\supseteq$: note that $\overbrace{Ax}^{\in C(A)}-\overbrace{p_W(b)}^{\in C(A)}\in C(A)$, and $\overset{\in C(A)^\perp}{Ax-p_W(b)}\perp C(A)$ $\xRightarrow{\text{thm on }\perp}$ $Ax-p_W(b)=0$, i.e., $x\in S$.

ORTHOGONAL MATRICES

Consider VS $\mathbb{R}^n$ with standard inner product, $\overbrace{\text{and } \|\ \|}^{\text{the induced norm}}$.

Def: A matrix $Q\in\mathbb{R}^{n\times n}$ is orthogonal if $Q^TQ=I$.

Theorem: The following claims are equivalent:

a) Q is orthogonal,
b) Q is invertible and $Q^{-1}=Q^T$,
c) $Q\cdot Q^T=I$
d) Q^T is orthogonal.

} follows from properties of matrix inverse

E.g. In $\mathbb{R}^2$, matrices of rotation, reflection.

GRAM-SCHMIDT ORTHOGONALIZATION

given a basis $(b_1, \dots, b_n)$, constructs an ON basis $(v_1, \dots, v_n)$

for $i = 1, \dots, n$ do

$(*)$ $\left[\begin{array}{l} y_i := b_i - \sum_{j=1}^{i-1} \langle b_i | v_j \rangle \cdot v_j \\ v_i = \dfrac{y_i}{\|y_i\|} \end{array}\right.$

projection of b_i onto $\mathrm{span}(v_1, \dots, v_{i-1})$

Theorem (QR decomposition of invertible matrices):
For every invertible $B \in \mathbb{R}^{n \times n}$, there exists an orthogonal matrix Q and an upper triangular matrix R with positive diagonal s.t. $B = Q \cdot R$.

Proof: Consider Gram-Schmidt ortogonalization of $b_1, \dots, b_n$:

note: $\forall i$, v_i is a linear combination of $b_1, \dots, b_i$, with positive coefficient of b_i.

$\Rightarrow$ For $Q = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$ and $B = \begin{pmatrix} | & & | \\ b_1 & \cdots & b_n \\ | & & | \end{pmatrix}$, $Q = B \cdot P$

for an upper triangular P with positive diagonal!

Recall (?): An inverse of an upper Δ is upper Δ.

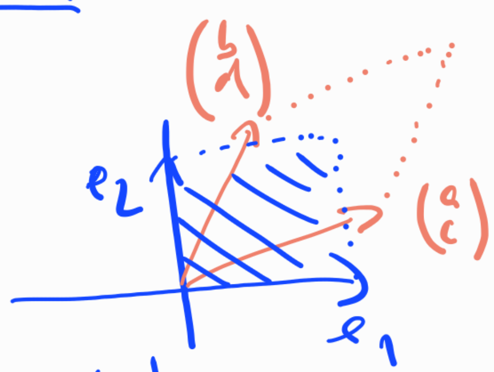
Let $R = P^{-1}$. Then $B = Q \cdot R$.

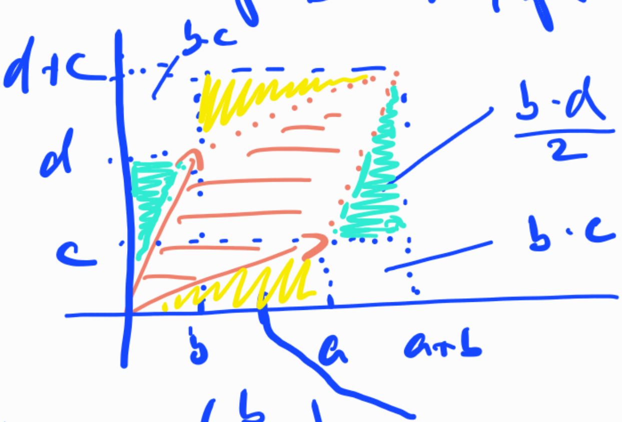
Moreover, by construction $Q^T \cdot Q = I$. ∎

Meaning: Decomposition into a product of special matrices.

DETERMINANTS

Motivation 1: Consider linear transformation $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x) = A\cdot x$.





$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad A\cdot e_1 = \begin{pmatrix} a \\ c \end{pmatrix} \quad Ae_2 = \begin{pmatrix} b \\ d \end{pmatrix}$$

f changes a square into a parallelogram

The area of it: $(a+b)(c+d) - 2\cdot bc - ac - bd =$

$= \cancel{ac} + bc + ad + \cancel{bd} - 2bc - \cancel{ac} - \cancel{bd}$

$= ad - bc$

The area of the square $= 1$ $\Rightarrow$ the area changed $(ad-bc)$-times

---

Motivation 2: Consider system $Ax = h$.

We know: exactly 1 solution $\Leftrightarrow$ A is invertible.

How to recognize it? We have an algorithm (Gauss elim.) but we want a formula

i) Assume $a \neq 0$ $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \sim \begin{pmatrix} a & b \\ 0 & ad-bc \end{pmatrix}$

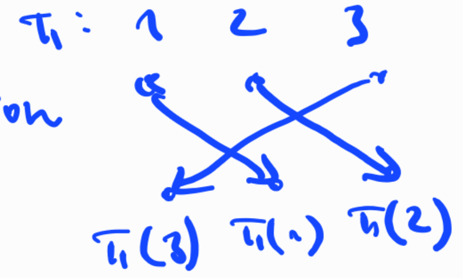
i.e., $ad-bc \neq 0 \iff A$ invertible

ii) For $a = 0$: $bc \neq 0 \iff A$ invertible

$\quad\quad\quad\quad\quad\quad\quad\; \overset{\shortparallel}{ad-bc}$

In both cases, the same number $ad-bc$ reveals something important about the matrix A.

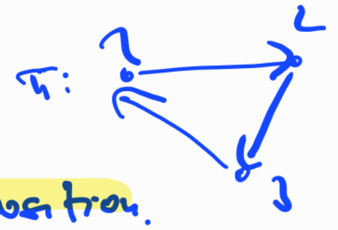
Recall:



permutation of $\{1,\dots,n\}$ is a bijection

$$\pi : \{1,\dots,n\} \to \{1,\dots,n\}$$

sign of a permutation



$$\mathrm{sgn}(\pi) = (-1)^{\#\text{inversions}} = (-1)^{n-\#\text{cycles}}$$
$$= (-1)^{\#\text{transpositions}}$$

transposition

E.g. $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ $\quad t_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ $\quad t_{ij}(k) = \begin{cases} i & k=j \\ j & k=i \\ k & k\neq i,j \end{cases}$

$\pi = t_{12} \circ t_{23}$ $\qquad t_{12} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$

2 inversions $\quad$ 1 cycle $\quad$ 2 transpositions

$$\mathrm{sgn}(\pi) = -1^{2} = -1^{3-1} = -1^{2} = 1$$

$\circ$ ... composition

Lemma: i) $\mathrm{sgn}(p \circ q) = \mathrm{sgn}(p) \cdot \mathrm{sgn}(q)$,

ii) $\mathrm{sgn}(p^{-1}) = \mathrm{sgn}(p)$.

Notation: S_n - set of all permutations of $\{1,\dots,n\}$

For every $\sigma \in S_n$: $\{\pi \circ \sigma : \pi \in S_n\} = S_n$.

Note: For a 2×2 matrix $(n=2)$ $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$,

we get the formula $a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$, i.e.,

we sum products of the form $\prod_{i=1}^{n} a_{i\pi(i)}$

over all permutations $\pi \in S_n$,

appropriately multiplied by ± 1, depending on $\mathrm{sgn}(\pi)$

Definition: **Determinant** of a square matrix $A \in T^{n\times n}$

is the number $$\det(A) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) \cdot \prod_{i=1}^{n} a_{i\pi(i)}.$$

For a triangular (upper or lower) matrix A,

$\det(A) = \prod_{i=1}^{n} a_{ii}$ - product of diagonal entries

Theorem:

i) If $A = \begin{pmatrix} - v_1 - \\ \vdots \\ - v_n - \end{pmatrix} \in T^{n \times n}$ (matrix), $z \in T^n$ (vector), $\alpha, \beta \in T$ (scalars),

then for each $i \in \{1, \dots, n\}$

$$\det \underbrace{\begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & \alpha v_i + \beta z & - \\ & \vdots & \\ - & v_n & - \end{pmatrix}}_{=B} = \alpha \cdot \det(A) + \beta \cdot \det \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & z & - \\ & \vdots & \\ - & v_n & - \end{pmatrix} \leftarrow i\text{-th row}$$

(i.e., det is a linear function of each its rows)

ii) If A' is obtained from A by a swap of two rows, then $\det(A') = -\det(A)$.

iii) $\det(I) = 1$.

notation: i-th row: $v_i = (a_{i1} \cdots a_{in})$

Proof: i) by def.

$$\det(B) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot a_{1\pi(1)} \cdots (\alpha a_{i\pi(i)} + \beta z_{\pi(i)}) \cdots a_{n\pi(n)} =$$

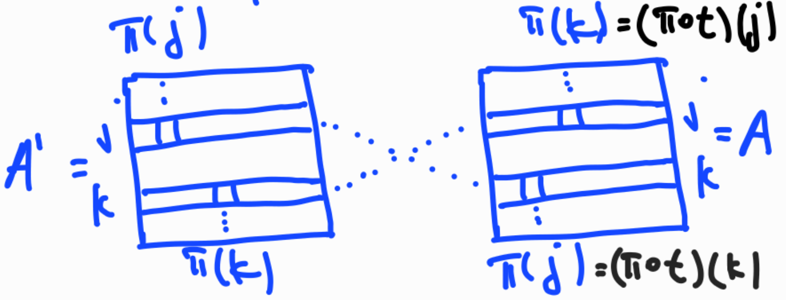
$$= \alpha \underbrace{\boxed{\sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot a_{1\pi(1)} \cdots a_{i\pi(i)} \cdots a_{n\pi(n)}}}_{=\det(A)} +$$

$$\beta \underbrace{\boxed{\sum_{\pi \in S_n} \operatorname{sgn}(\pi)\, a_{1\pi(1)} \cdots z_{\pi(i)} \cdots a_{n\pi(n)}}}_{= \det \begin{pmatrix} - v_1 - \\ \vdots \\ - z - \\ \vdots \\ - v_n - \end{pmatrix}}$$

ii) Let $j \neq k$ be the indices of the swapped rows, and $t \in S_n$ be a transposition $t(j) = k$, $t(k) = j$, $t(i) = i$ for $i \neq j, k$.

$$\det(A') = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \prod_{i=1}^{n} a'_{i\pi(i)} = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \prod_{i=1}^{n} a_{i(\pi \circ t)(i)} =$$

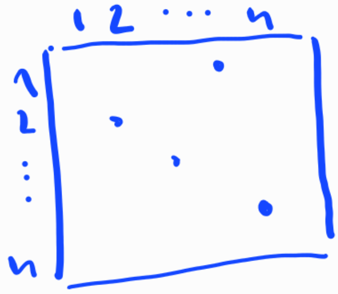
$$= \sum_{(\pi \circ t) \in S_n} -\operatorname{sgn}(\pi \circ t) \prod_{i=1}^{n} a_{i(\pi \circ t)(i)} = -\det(A)$$



iii) we already know - I is an upper triangular.

∎

Theorem: $\det(A) = \det(A^T)$



$j = \pi(i) \iff \pi^{-1}(j) = i$

Proof: Note that for each $\pi \in S_n$:

$$\prod_{i=1}^{n} a_{i\pi(i)} = \prod_{j=1}^{n} a_{\pi^{-1}(j)j}$$

As $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$, we get:

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} = \sum_{\pi^{-1} \in S_n} \operatorname{sgn}(\pi^{-1}) \prod_{j=1}^{n} a_{\pi^{-1}(j)j}$$

$$= \sum_{\pi' \in S_n} \operatorname{sgn}(\pi') \prod_{j=1}^{n} A^T_{j\pi'(j)} = \det(A^T) \qquad \blacksquare$$

denote $\pi' = \pi^{-1}$