Chapter VI

Inner product space.

VI 1st definitions and properties

1) From Chapter III we know

$\langle , \rangle : \mathbb{R}^n \to \mathbb{R}$

$\langle v, w \rangle := v \cdot w.$ (dot product)

$= \sum_{i=1}^{n} v_i w_i$

Here we had a canonical basis $\vec{e}_1, \dots, \vec{e}_n$.

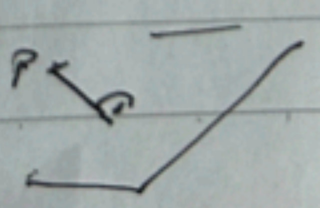
But how do we describe

$\langle , \rangle|_{W \times W}$ for $W \leq \mathbb{R}^n$?

This leads to the theory of inner product spaces of finite dimension

2) Further: In $(\mathbb{R}^n, \langle , \rangle)$ we were able to compute a distance to a set using orthogonality.

Ex.: $dist(P, H) = dist(P, proj_H(P))$ for a hyperplane.



We want to use this idea for ∞-dimensional spaces, like in $C^2([a,b])$.

Def_251 : A tuple $(V, <,>)$ is called an inner product space if

- V is an $\mathbb{R}$-vector space and
- $<,> : V \times V \longrightarrow \mathbb{R}$ is an inner _product_, i.e.

(IP1) $<,>$ is bilinear, i.e.

$$<v_1 + v_2, v> = <v_1, v> + <v_2, v>$$
$$<\lambda w, v> = \lambda <w, v>$$
$$<v, v_1 + v_2> = <v, v_1> + <v, v_2>$$
$$<v, \lambda w> = \lambda <v, w>$$
$$\forall_{v_1, v_2, v, w \in V} \forall \lambda \in \mathbb{R}$$

(IP2) $<,>$ is symmetric

$$<v, w> = <w, v>$$

(IP3) $<,>$ is positive definite, i.e.

$$<v, v> \gneq 0 \quad \forall v \in V \quad v \neq 0.$$

Example 252: 1) $A := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

$\langle v, w \rangle_A := v^T A w \quad v, w \in \mathbb{R}^{2\times 1}$.

$\langle , \rangle_A$ is an inner product on $\mathbb{R}^{2\times 1}$.

Proof: (IP1): $\langle v_1 + v_2, w \rangle_A = (v_1 + v_2)^T A w$

$= (v_1^T + v_2^T) A w = \langle v_1, w \rangle_A + \langle v_2, w \rangle_A$

$\langle \lambda v, w \rangle_A = (\lambda v)^T A w = \lambda (v^T A w) = \lambda \langle v, w \rangle$

(IP2) $\langle v, w \rangle_A = v^T A w \underset{A = A^T}{=} v^T A^T w$

$= w^T A v = \langle w, v \rangle_A$.

(IP3) Take $v \in \mathbb{R}^{2\times 1}$ non-zero.

Then $\langle v, v \rangle_A = 2v_1^2 + 2v_1v_2 + 2v_2^2$

$\underset{(1)}{\geq} 2(v_1^2 - 2|v_1||v_2| + v_2^2) \geq 2(|v_1| - |v_2$

No. 301

Date

2) $A = I : \langle \, , \rangle_I : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$
is the standard inner product

$$\langle v, w \rangle_I = (v_1, \dots, v_n) \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$$= (v_1, \dots, v_n) \circ \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$$= \sum v_i w_i = v \circ w.$$

3) $A = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$

$\langle \, , \rangle_A$ is not an inner product because it is not ~~pol~~ positive definite: $\langle \vec{e}_2, \vec{e}_2 \rangle_A = -1 < 0.$

4) $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 2 & 1 \end{pmatrix}$

$$P_A(\lambda) = \lambda^3 - 5\lambda^2 - 10\lambda + 8$$
$$= (\lambda + 2)(\lambda^2 - 7\lambda + 4)$$

$$= (\lambda + 2)\left(\lambda - \left(\frac{7}{2} + \frac{\sqrt{49 - 16}}{2}\right)\right)\left(\lambda - \left(\frac{7}{2} - \frac{\sqrt{33}}{2}\right)\right)$$

$\langle \, , \rangle_A$ is not an inner product

Idea 253: $A \in \mathbb{R}^{n \times n}$ symmetric.

Then $\langle , \rangle_A$ is an inner product iff all eigenvalues of A are positive.

Proof: Exercise. $\square$

Def 254: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called

(i) positive definite

if $v^T A v > 0 \ \forall \ v \in \mathbb{R}^{n \times 1} \setminus \{\underline{0}\}$

(all eigenvalues positive)

(ii) negative definite

if $v^T A v < 0 \ \forall \ v \in \mathbb{R}^{n \times 1} \setminus \{\underline{0}\}$

(all eigenvalues negative)

(iii) indefinite

if $\exists \ v, w \in \mathbb{R}^{n \times 1} \setminus \{\underline{0}\}$:

$v^T A v > 0 > w^T A w$.

(some eigenvalue positive and some " negative)

Example 255:

	$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$	$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 2 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & & \\ & -2 & \\ & & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$
pos. def.	✓	✗	✓	✗	✗	✗
neg. def.	✗	✗	✗	✗	✓	✗
indef.	✗	✓	✗	✓	✗	✗

Date

$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$: $p_A(\lambda) = \lambda^3 - 3\lambda^2 = \lambda^2(\lambda - 3)$

$\Rightarrow \operatorname{Spec}(A) = \{0, 3\}$

$\Rightarrow$ all eigenvalues ≥ 0
"positive semidefinite."

Analog we say A is negative semidefinite if all eigenvalues ≤ 0.

Example 2.56: (infinite dimensional example)

a) $C([0,1]) = \{ f : [0,1] \to \mathbb{R} \mid f \text{ is continuous} \}$

$\langle f, g \rangle := \int_0^1 f(x) g(x) \, dx.$

b) $\ell_{\text{finite}} = \{ (a_n)_{n \in \mathbb{N}} \mid a_n \in \mathbb{R}$ and only finitely many terms are non-zero$\}$

$(0, 0, 1, -2, 0, -1, 0, 0, \dots) \in \ell_f.$

$$\langle (a_n)_{\mathbb{N}}, (b_n)_{\mathbb{N}} \rangle := \sum_{n=1}^{\infty} a_n b_n$$

c) $\ell^2 = \{ (a_n)_{\mathbb{N}} \mid a_n \in \mathbb{R}, \sum_{n=1}^{\infty} |a_n|^2 < \infty \}$

$$\langle (a_n)_{\mathbb{N}}, (b_n)_{\mathbb{N}} \rangle := \sum_{n=1}^{\infty} a_n b_n$$

Why is $\langle , \rangle$ well-defined, i.e.

$\langle a, b \rangle \in \mathbb{R}$?, i.e. $\sum_{n=1}^{\infty} a_n b_n$

converges?

<u>Proof of well-def</u>:

$m \in \mathbb{N}$. $\quad \sum_{n=1}^{m} |a_n b_n| \leq \sqrt{\sum_{n=1}^{m} |a_n|^2} \sqrt{\sum_{n=1}^{m} |b_n|^2}$

Cauchy-Schwarz

$$\leq \sqrt{\sum^{\infty} |a_n|^2} \sqrt{\sum^{\infty} |b_n|^2} < \infty$$

So $\left(\sum_{n=1}^{m} |a_n b_n| \right)_m$ is bounded and increasing $\Rightarrow$ convergent. $\blacksquare$

of finite dim

Given an inner product space (V, <,>) we can find a symmetric matrix too:

<u>Def 257</u>: Let $B = \{v_1, v_2, \ldots, v_n\}$ be a basis of V. We call

$$\mathrm{Gram}_B(\langle , \rangle) := \left(\langle v_i, v_j \rangle \right)_{\substack{1 \le i \le n \\ 1 \le j \le n}}$$

<u>the Gram matrix</u> of $\langle , \rangle$ w.r.t. B.

It is positive definite.

<u>Example 258</u>:

1) $\langle , \rangle_A$, A pos. def.

$$\mathrm{Gram}_{\{\vec{e}_1, \ldots, \vec{e}_n\}}(\langle , \rangle_A) = A$$

because $\langle \vec{e}_i, \vec{e}_j \rangle_A = \vec{e}_i^{\,T} A \vec{e}_j$

$= a_{ij}$.

2) $\langle , \rangle = \langle , \rangle_{I_2}$ standard inner product on $\mathbb{R}^2$.

$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} =: \{v_1, v_2\}$

$$\mathrm{Gram}_{\mathcal{B}}(\langle , \rangle) = \begin{pmatrix} \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle, & \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle \\ \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle, & \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

(This way one can produce a lot of positive definite matrices)

End of Lecture 20th of Dec 23

3) What happens under basis transformation?

Ex: $\langle , \rangle = \langle , \rangle_{I_2}$

$\mathcal{B}$ as in 2)

$C := \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$

$$\mathrm{Gram}_C(\langle , \rangle) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

No. 307

Date

Take v, w with $[v]_B = \begin{pmatrix} v_1 \\ \vdots \\ v_2 \end{pmatrix}$, $[w]_B = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$

i.e. $v = v_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + v_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$\Rightarrow \langle v, w \rangle = \sum_{\substack{i,j \\ 1 \le i \le 2 \\ 1 \le j \le 2}} v_i w_j \langle v_i, v_j \rangle$$

$$= [v]_B^T \circ \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{pmatrix} \circ [w]_B$$

$$= [v]_B^T \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} [w]_B$$

So $\mathrm{Gram}_C(\langle , \rangle)$

$$= \left(\left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right]_B, \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]_B \right)^T \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \left(\left[\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right]_B \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]_B \right)$$

$$= P_{C \to B}^T \, \mathrm{Gram}_B(\langle , \rangle) \, P_{C \to B}.$$

<u>check</u>: $\text{Gram}_C(\langle, \rangle) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$

$$P_{C \to B} = \left(\left[\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right]_B, \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]_B \right) = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \checkmark$$

<u>Def 259</u>: Given $(V, \langle, \rangle)$ an inner product space, we say $v, w \in V$ are perpendicular (orthogonal to each other) if $\langle v, w \rangle = 0_{\mathbb{R}}$.

Write $v \perp w$.

Let $S \subseteq V$.

$$S^{\perp} := \{ v \in V \mid v \perp w \; \forall w \in S \}$$

"orthogonal complement of S"

Example 260: (a) $\langle , \rangle = \langle , \rangle_{I_3}$

$S = \{\vec{e}_1\} \Rightarrow S^{\perp} = \{v \in \mathbb{R}^3 \mid \langle \vec{e}_1, v\rangle = 0\}$
$= \mathbb{R}\vec{e}_2 + \mathbb{R}\vec{e}_3$

(b) $\langle , \rangle = \langle , \rangle_{I_3}$ $\quad S = \left\{ \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix} \right\}$

$S^{\perp} = \left\{ v \in \mathbb{R}^3 \mid v \perp \begin{pmatrix}1\\1\\0\end{pmatrix} \text{ and } v \perp \begin{pmatrix}1\\0\\1\end{pmatrix} \right\}$

$= \text{null}\left(\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}\right) = \mathbb{R}\begin{pmatrix}-1\\1\\1\end{pmatrix}$

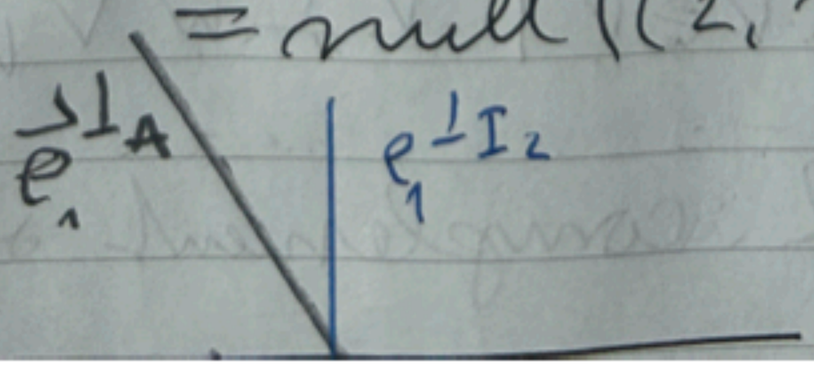
$\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix} \to \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & -1\end{pmatrix}$

(c) Take $\vec{e}_1 \in \mathbb{R}^2$. Consider $\langle , \rangle_{I_2}$ and $\langle , \rangle_A$, $A = \begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}$.

$\vec{e}_1^{\,\perp I_2} = \mathbb{R}\vec{e}_2$

$\vec{e}_1^{\,\perp A} = \text{null}\left(\vec{e}^{\,T}\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}\right)$

$= \text{null}((2, 1)) = \mathbb{R}\begin{pmatrix}-1\\2\end{pmatrix}$



(d) $S = \left\{ \begin{pmatrix} 1\\1\\1\\1 \end{pmatrix}, \begin{pmatrix} 1\\2\\0\\1 \end{pmatrix} \right\} \subseteq \mathbb{R}^{4\times 1}$

$\langle , \rangle = \langle , \rangle_{I_4}$

$S^\perp = \operatorname{null}\left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} \right)$

$= \mathbb{R}\begin{pmatrix} -2\\1\\1\\0 \end{pmatrix} + \mathbb{R}\begin{pmatrix} -1\\0\\0\\1 \end{pmatrix}$

$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2 & 1 \\ & 1 & -1 & 0 \end{pmatrix}$

<u>Prop 260:</u> Let $(V, \langle , \rangle)$ be a finite dimensional inner product space. (i.p.s.)

(a) Let $S \subseteq V$. Then $S^\perp \leq V$.

(b) Let $W \leq V$. Then

(b1) $(W^\perp)^\perp = W$

(b2) $W + W^\perp = V$ direct,

i.e. $W \cap W^\perp = \{0\}$

(b3) $\dim W^\perp = \dim V - \dim W$.

<u>Proof:</u> (a) • $v_1, v_2 \in S^\perp \Rightarrow \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$ $= 0 + 0$ for $w \in S$. $\Rightarrow v_1 + v_2 \in S^\perp$

• $0 \in S^\perp$

• $\lambda \in \mathbb{R}, v \in S^\perp \Rightarrow \langle \lambda v, w \rangle = \lambda \langle v, w \rangle = 0$ for $w \in S \Rightarrow \lambda v \in S^\perp$.

Subspace criterion $\Rightarrow S^\perp \leq V$

Date

(b) Let $B := \{v_1, \dots, v_n\}$ be a basis of V and $A := \mathrm{Gram}_B(<,>)$.

Consider $f : V \longrightarrow \mathbb{R}^n$ given by $f(v_i) := \vec{e}_i$, $i = 1, \dots, n$ and linear extension
$f(\sum_{i=1}^n \lambda_i v_i) := \sum \lambda_i f(v_i) = \sum \lambda_i \vec{e}_i$

Then f is bijective and
$$< f(v), f(v') >_A = < v, v' >$$
for all $v, v' \in V$.
("f is an isometry from $(V, <,>)$ to $(\mathbb{R}^n, <,>_A)$").

So we can replace $(V, <,>)$ by $(\mathbb{R}^n, <,>_A)$.
(That is how we use an isometry

(e3) Let $\{w_1, \dots, w_m\}$ be a basis of W
Then
$\{Aw_1, \dots, Aw_m\}$ ~~has~~
is linearly independent as A is invertible (Why?).

$$\Longrightarrow \text{nullity}\left(\begin{pmatrix} w_1^T \\ \vdots \\ w_m^T \end{pmatrix} A^T\right) \underset{\substack{\Uparrow \\ A = A^T}}{=} \text{nullity}\left(\begin{pmatrix} w_1^T \\ \vdots \\ w_m^T \end{pmatrix} A\right)$$

$= \dim V - m$
↑
rk-nullity thm

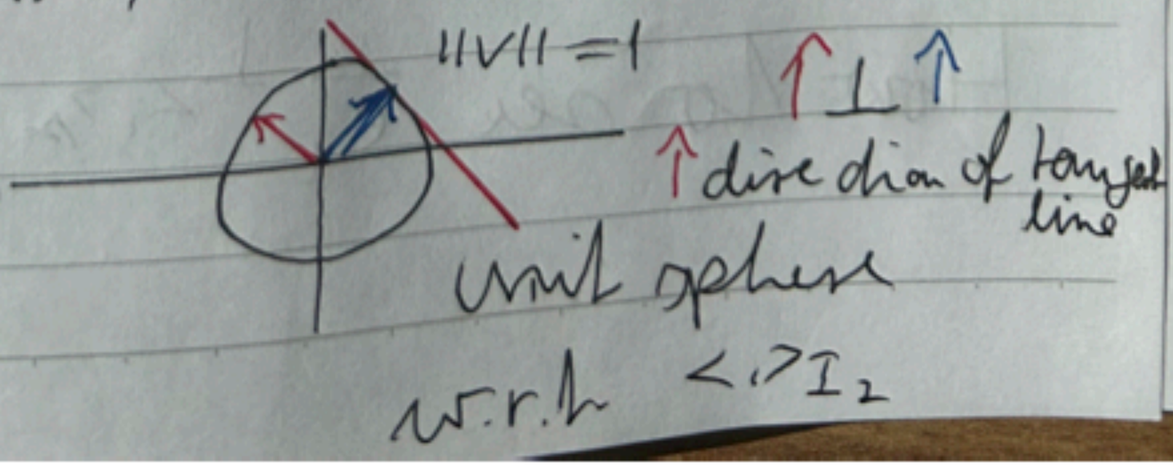
$\Rightarrow \dim W^\perp = \dim V - \dim W.$
$\Rightarrow$ (b-3)
For (b-2) $w \in W^\perp \cap W \Rightarrow \langle w, w \rangle = 0 \Rightarrow w = 0.$
$W + W^\perp$ is direct
$\Rightarrow \dim(W + W^\perp) = \dim W + \dim W^\perp$
$= \dim V$
$\Rightarrow V = W + W^\perp$
↑
$W + W^\perp, V$ finite dim.

For (b-1): $W \subseteq (W^\perp)^\perp$, both have same finite dimension.
$\Rightarrow W = (W^\perp)^\perp$ □

Example 262: (Unit sphere)

Consider $(\mathbb{R}^2, \langle , \rangle_A)$, A positive def.

$A = I_2$:

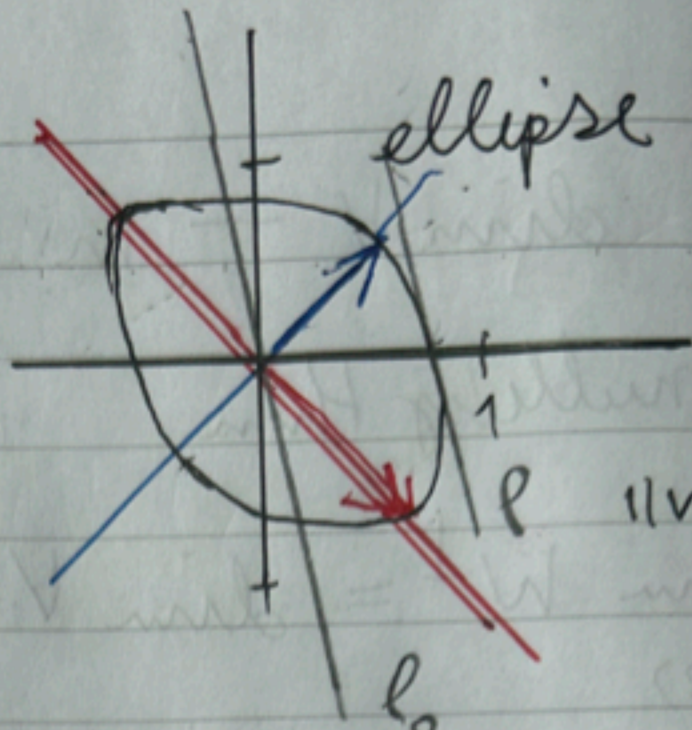


No. 3|3

Date

$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

$\|v\|_A := \sqrt{\langle v, v\rangle_A}$



ellipse $2x^2+2xy+2y^2=1$

$\|v\|_A = 1$

eigenbasis $\left\{ \overset{3}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}, \overset{1}{\begin{pmatrix} 1 \\ -1 \end{pmatrix}} \right\}$

normalize $v := \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{\langle \binom{1}{1}, \binom{1}{1} \rangle_A}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{\sqrt{6}}$

$w = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\|\binom{1}{-1}\|_A} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{1}{\sqrt{2}}$

$v \perp_{\langle,\rangle_{I_2}} w$ and $v \perp_{\langle,\rangle_A} w$

$\Big(\; v^T A w = \frac{1}{\sqrt{12}} (1,1) \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$= \frac{1}{\sqrt{12}} (3,3) \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0$

How to see $\vec{e}_1^{\perp_{\langle,\rangle_A}}$?

Write S^1_A or $S^1_{<,>_A}$ for $\{v \in \mathbb{R}^2 \mid \|v\|_A = 1\}$

the unit sphere wrt. A (or $<,>_A$).

Take tangent line l on S^1_A at

$\mathbb{R}^{>0}\vec{e}_1 \cap S^1_A$. The parallel line l_0 through 0 is $\vec{e}_1^{\perp_A}$.

Here: ① $<v, \vec{e}_1>_A = 0 \iff v^T\circ\binom{2}{1} = 0$

$\iff v \in \mathbb{R}\binom{-1}{2}$

② Take derivative of

$$1 = 2x^2 + 2xy + 2y^2 \text{ at } \vec{e}_1$$

$$\Rightarrow \quad 0 = 2^2 x_0 dx + 2x_0 dy + 2y_0 dx + 2^2 y_0 dy$$
$$= 2(2x_0 + y_0)dx + (4y_0 + 2x_0)dy$$
for $(x_0, y_0) = (1, 0)$

$\Rightarrow$ l_0 is given by $4x + 2y = 0$

$\Rightarrow$ $l_0 = \mathbb{R}\binom{-1}{2}$.

Date

Def 263: $\|v\|_{2,<>} := \sqrt{<v,v>}$

Theorem 264: Let $(V, <,>)$ be a i.p.s

Then

(i) $|<v,w>|^2 \leq \|v\|^2 \|w\|^2$

$\forall v,w \in V$

Cauchy-Schwarz inequality

(ii) $\|\cdot\|_{<,>}$ is a norm on V,
(i.e. ≥ 0 and satisfies
(N1), (N2), (N3) of Theorem 123)

Proof: (i) $v,w \in V$. For $t \in \mathbb{R}$:

$0 \leq <tv+w, tv+w>$
$= t^2 <v,v> + 2t<v,w> + <w,w>$

$= (t\|v\| + \|w\|)^2 + 2t(<v,w> - \|v\|\|w\|)$

If $v=0$ then $|<v,w>| = |<0,w>| = 0$
$= \|0\| \cdot \|w\|$

If $v \neq 0$ plug in $t := -\frac{\|w\|}{\|v\|}$. $\square$ (i)

(ii) Exercise $\square$

End of Lecture 22nd of Dec 23

<u>Question 265:</u> $(V, \langle , \rangle)$ finite dim i.p.s

$W \leq V$.

$$V = W \oplus W^{\perp} \xrightarrow{\text{proj}_W} W$$
$$w_1 + w_2 \longmapsto w_2$$

1) How to compute $\text{proj}_W(v)$?

2) Given ~~an~~ a basis of W,

How to get an ON basis?

<u>Def 266:</u> (ON-basis) Let $(V, \langle , \rangle)$ be an i.p.s., $B = \{v_1, \ldots, v_n\}$ a basis of V.

(i) B is called <u>orthogonal</u> if $v_i \perp v_j \; \forall_{i \neq j}$.
(O-basis)

(ii) — " — <u>orthonormal</u> if B is orthogonal and $\|v_i\| = 1 \; \forall \; i = 1, \ldots, n$.
(ON-basis)

Similar orthogonal / orthonormal set.

(ii) Exercise $\square$

End of Lecture 22nd of Dec 23.

Question 265: $(V, \langle , \rangle)$ finite dim i.p.s.

$W \leq V$.

$$V = W \oplus W^{\perp} \xrightarrow{\mathrm{proj}_W} W$$
$$w_1 + w_2 \longmapsto w_2$$

1) How to compute $\mathrm{proj}_W(v)$?

2) Given ~~an~~ a basis of W,

How to get an ON basis?

Def 266: (ON-basis) Let $(V, \langle , \rangle)$ be an i.p.s., $B = \{v_1, \dots, v_n\}$ a basis of V.

(i) B is called orthogonal if $v_i \perp v_j \ \forall i \neq j$.
(O-basis)

(ii) — " — orthonormal if B is orthogonal and $\|v_i\| = 1 \ \forall i = 1, \dots, n$.
(ON-basis)

Similar orthogonal / orthonormal set.

Date

Example 267:

(a) $\{\vec{e}_1, \vec{e}_2\}$ in $\mathbb{R}^2$ is an ON-basis w.r.t. $\langle\,|\,\rangle_{I_2}$.

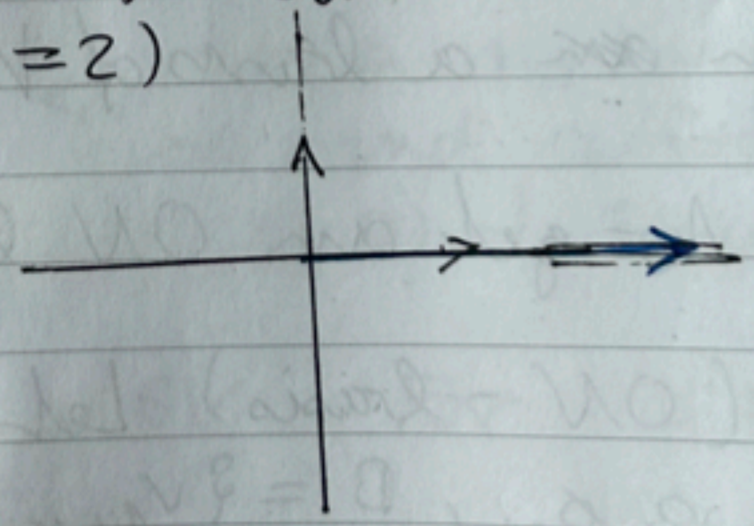
Proof: • basis ✓

• $\langle \vec{e}_1, \vec{e}_2 \rangle = \vec{e}_1 \cdot \vec{e}_2 = 0$

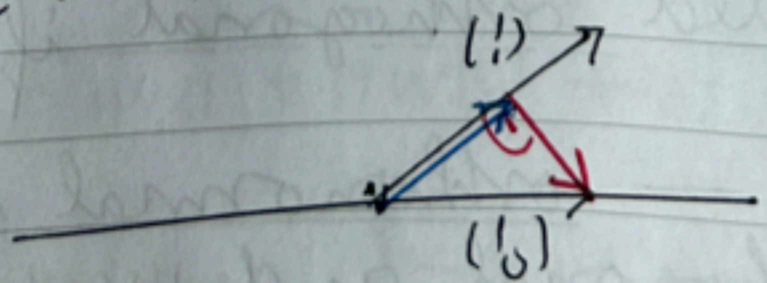
• $\|\vec{e}_1\|_{I_2} = 1 = \|\vec{e}_2\|_{I_2}$. $\square$

$\{2\vec{e}_1, \vec{e}_2\}$ is an O-basis, not an ON basis

$(\|2\vec{e}_1\| = 2)$



(b) $\left\{ \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix} \right\}$ is a basis of $\mathbb{R}^2$



How to get an O-basis containing $\begin{pmatrix}1\\1\end{pmatrix}$?

Replace $\begin{pmatrix}1\\0\end{pmatrix}$ by

$$\begin{pmatrix}1\\0\end{pmatrix} - \mathrm{proj}_{\begin{pmatrix}1\\1\end{pmatrix}}\begin{pmatrix}1\\0\end{pmatrix}$$

$$= \begin{pmatrix}1\\0\end{pmatrix} - \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\right\rangle \begin{pmatrix}1\\1\end{pmatrix}\frac{1}{\sqrt{2}}$$

$$= \begin{pmatrix}1\\0\end{pmatrix} - \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}$$

$\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}\right\}$ is an O-basis.

How to get an ON-basis?:

$$\left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\left(\frac{1}{\sqrt{2}}\right)}\begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}\right\} = \left\{\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\right\}$$

Gram-Schmidt process 268:
(to orthonormalize a linear independent set)

Given $\{v_1, \dots, v_m\}$ linearly independent in V.

Step 1: ~~$u_1 := \frac{1}{\|v_1\|}v_1$~~. $w_1 := v_1$, $u_1 := \frac{1}{\|w_1\|}w_1$

No. 319

Date

<u>Step 2:</u> $w_2 := v_2 - \langle v_2, u_1 \rangle u_1$

$u_2 := \frac{1}{\|w_2\|} w_2$

<u>Step 3:</u> $w_3 := v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$

$u_3 := \frac{1}{\|w_3\|} w_3$

$\vdots$

<u>Step i:</u> $w_i := v_i - \langle v_i, u_1 \rangle u_1 - \dots - \langle v_i, u_{i-1} \rangle u_{i-1}$

$u_i := \frac{1}{\|w_i\|} w_i.$

Then $\{w_1, \dots, w_n\}$ is an O-set
and $\{u_1, \dots, u_n\}$ is an ON-set.

<u>Proof:</u> • ~~w~~ is welldefined, because
$w_i \neq 0$, because $v_i \notin \operatorname{span}\{v_1, \dots, v_{i-1}\}$

• $\|u_i\| = 1$

• $\operatorname{span}\{v_1, \dots, v_i\} = \operatorname{span}\{w_1, \dots, w_i\}$

Date

- For $i < j$: $\langle w_i, w_j \rangle = 0$, because $w_i \in \operatorname{span}\{u_1, \dots, u_{j-1}\}$ and for $1 \le k \le j-1$

$$\langle w_j, u_k \rangle = \langle v_j - \sum_{\ell=1}^{j-1} \langle v_j, u_\ell \rangle u_\ell, u_k \rangle$$

$$\underset{\uparrow}{=} \langle v_j, u_k \rangle - \langle v_j, u_k \rangle \langle u_k, u_k \rangle$$

(induction j, so by induction hypothesis: $\{u_1, \dots, u_{j-1}\}$ is ON)

$$\underset{\uparrow}{=} 0$$

$\|u_k\| = 1$.

The base case is given by $\{u_1\}$ being ON. □

Example 269:

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Step 1: $w_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

Step 2: $w_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, u_1 \right\rangle u_1$

$$= \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{2}} \right\rangle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{2}}$$

$$= \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

$$u_2 = \frac{1}{\|w_2\|} w_2 = \frac{1}{\sqrt{\frac{3}{2}}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \sqrt{\frac{2}{3}} \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

Step 3: $w_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{2}} \right\rangle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{2}}$

$$- \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \frac{1}{\sqrt{6}} \right\rangle \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

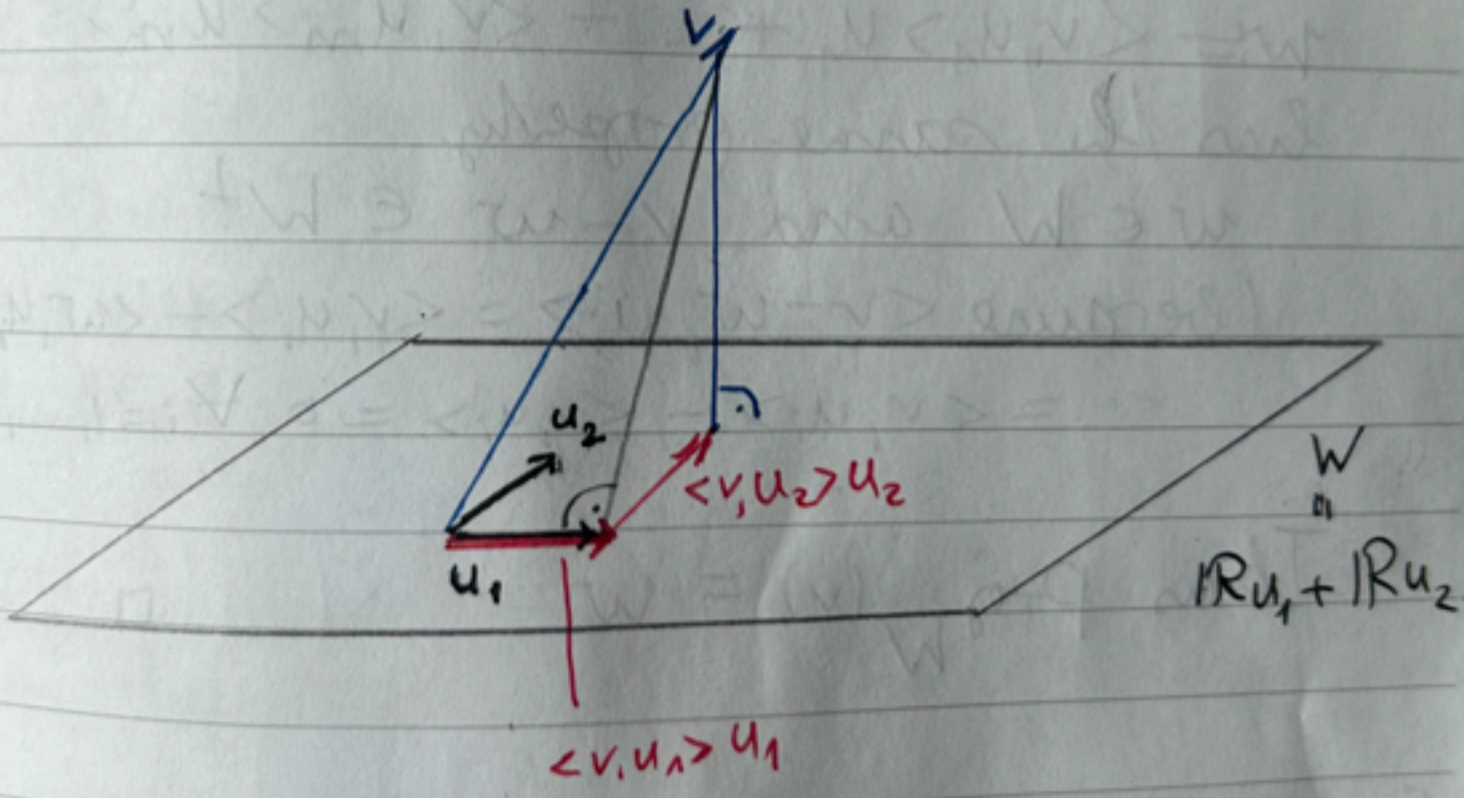
$$= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} \cancel{-\frac{1}{3}} \\ \cancel{\frac{1}{3}} \\ \cancel{\frac{4}{3}} \end{pmatrix} = \frac{1}{3}$$

$$= \begin{pmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} \qquad u_3 = \frac{1}{\|w_3\|} w_3 = \frac{1}{\frac{2}{3}\sqrt{3}} w_3 = \begin{pmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$$

Thus $\{u_1, u_2, u_3\} = \left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$

is an ON set, in fact an ON basis of $\mathbb{R}^{3\times1}$!

<u>geometric picture 270:</u>



<u>Corollary 271:</u> Let $(V, \langle\,\rangle)$ be a finite dim. i.p.s. and $\{u_1, \dots, u_m\}$ be an ON basis of $W \leq V$. Then

$$\operatorname{proj}_W(v) = \langle v, u_1 \rangle u_1 + \dots + \langle v, u_m \rangle u_m$$

for all $v \in V$.

<u>Proof:</u> $\mathrm{proj}_W(v)$ is the unique element of W s.t. $v - \mathrm{proj}_W(v) \perp$ ~~$\mathrm{proj}_W(v)$~~ W,

i.e. we have

$$v = \underbrace{(v - \mathrm{proj}_W(v))}_{\in W^{\perp}} + \underbrace{\mathrm{proj}_W(v)}_{\in W}$$

(by definition of proj_W.)

$w := \langle v, u_1 \rangle u_1 + \dots + \langle v, u_m \rangle u_m$
has the same property.

$w \in W$ and $v - w \in W^{\perp}$

(because $\langle v - w, u_i \rangle = \langle v, u_i \rangle - \langle w, u_i \rangle$
$= \langle v, u_i \rangle - \langle v, u_i \rangle = 0 \ \forall i = 1, \dots, m$)

Thus $\mathrm{proj}_W(v) = w$ $\square$

<u>Example 272:</u> $(\mathbb{R}^4, \langle \, , \rangle_{I_4})$

$$W = \mathbb{R} \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \frac{1}{\sqrt{3}}}_{u_1} + \mathbb{R} \underbrace{\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \frac{1}{\sqrt{3}}}_{u_2}$$

$$\mathrm{proj}_W \left(\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} \right) = \langle v, u_1 \rangle u_1 + \langle v, u_2 \rangle u_2$$

$$= \frac{1}{3}(v_1 + v_2 + v_4) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{3}(v_1 - v_2 + v_3)$$

$$= \frac{1}{3}\begin{pmatrix} 2v_1 + v_3 + v_4 \\ 2v_2 - v_3 + v_4 \\ v_1 - v_2 + v_3 \\ v_1 + v_2 + v_4 \end{pmatrix}$$

Ex: $\mathrm{proj}_W(\vec{e}_1) = \frac{1}{3}\begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \end{pmatrix}$.

What about the orthogonal projection onto a column space of a matrix A?

Prop. 273: Let $A \in \mathbb{R}^{m\times n}$ of rank n.
Then there exist
$Q \in \mathbb{R}^{m\times n}$ with $\{c_1(Q), \dots, c_n(Q)\}$ ON
and $R \in \mathbb{R}^{n\times n}$ upper triangular invertible
such that $A = QR$

Def 274: Such a decomposition $A = QR$ as in Prop 273 is called Q-R decomposition.

Proof of Prop 273: Let v_j be the jth column of A.
Gram-Schmidt $\Rightarrow$ We get $\{u_1, \dots, u_n\}$ ON-basis of $\mathrm{col}(A)$.

$$v_j = \langle v_j, u_1\rangle u_1 + \dots + \langle v_j, u_j\rangle u_j$$

So $A = QR$ with $Q = (u_1, \dots, u_m)$

and $R = \begin{pmatrix} \langle v_1, u_1 \rangle & \langle v_2, u_1 \rangle & \langle v_3, u_1 \rangle & \cdots & \langle v_m, u_1 \rangle \\ & \langle v_2, u_2 \rangle & \langle v_3, u_2 \rangle & & \langle v_m, u_2 \rangle \\ & & \langle v_3, u_3 \rangle & & \langle v_m, u_3 \rangle \\ & 0 & & \ddots & \vdots \\ & & & & \langle v_m, u_m \rangle \end{pmatrix}$

□

Example 275:

$A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} \qquad v_1 := \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \; v_2 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$

G-S: Step 1: $u_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$

Step 2: $w_2 = v_2 - \langle v_2, u_1 \rangle u_1$

$$= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{6} \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$u_2 = \frac{1}{\|w_2\|} w_2 = \sqrt{2}\, w_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Q: $Q = (u_1, u_2) = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ -1/\sqrt{6} & 1/\sqrt{2} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix}$

No. 326

Date

R:
$$R = \begin{pmatrix} \langle v_1, u_1 \rangle & \langle v_2, u_1 \rangle \\ & \langle v_2, u_2 \rangle \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{6} & \frac{3}{\sqrt{6}} \\ & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$A = QR$

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} \sqrt{6} & \frac{3}{\sqrt{6}} \\ & \frac{1}{\sqrt{2}} \end{pmatrix}$$

IV 2. Least square method

We are given a set of data

$$(b_1, a_{11}, \dots, a_{1n}) = (b_1, \vec{a}_1)$$
$$(b_2, a_{21}, \dots, a_{2n}) = (b_2, \vec{a}_2)$$
$$\vdots$$
$$(b_m, \vec{a}_m)$$

We want a linear map $\ell : \mathbb{R}^{1\times n} \to \mathbb{R}$ such

$$\begin{pmatrix} \ell(\vec{a}_1) \\ \ell(\vec{a}_2) \\ \vdots \\ \ell(\vec{a}_m) \end{pmatrix} \text{ is nearest to } \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} =: \vec{b}$$

i.e. minimize

$$\left\| \vec{b} - \begin{pmatrix} \ell(\vec{a}_1) \\ \vdots \\ \ell(\vec{a}_m) \end{pmatrix} \right\| \text{ amoung all}$$

linear $\ell : \mathbb{R}^{1\times n} \to \mathbb{R}$

Date

Reformulate: $\ell(\vec{a}) = \vec{a} \circ \vec{m}$, $\vec{m} \in \mathbb{R}^{n \times 1}$
normal f.l.

$$\begin{pmatrix} \ell(\vec{a}_1) \\ \vdots \\ \ell(\vec{a}_m) \end{pmatrix} = A\vec{m}, \quad A := \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix} = (a_{ij}) \in \mathbb{R}^{m \times n}.$$

So minimize $\|b - A\vec{m}\|$

among all $\vec{m} \in \mathbb{R}^{n \times 1}$, i.e.

Least square problem 276: (LSP)

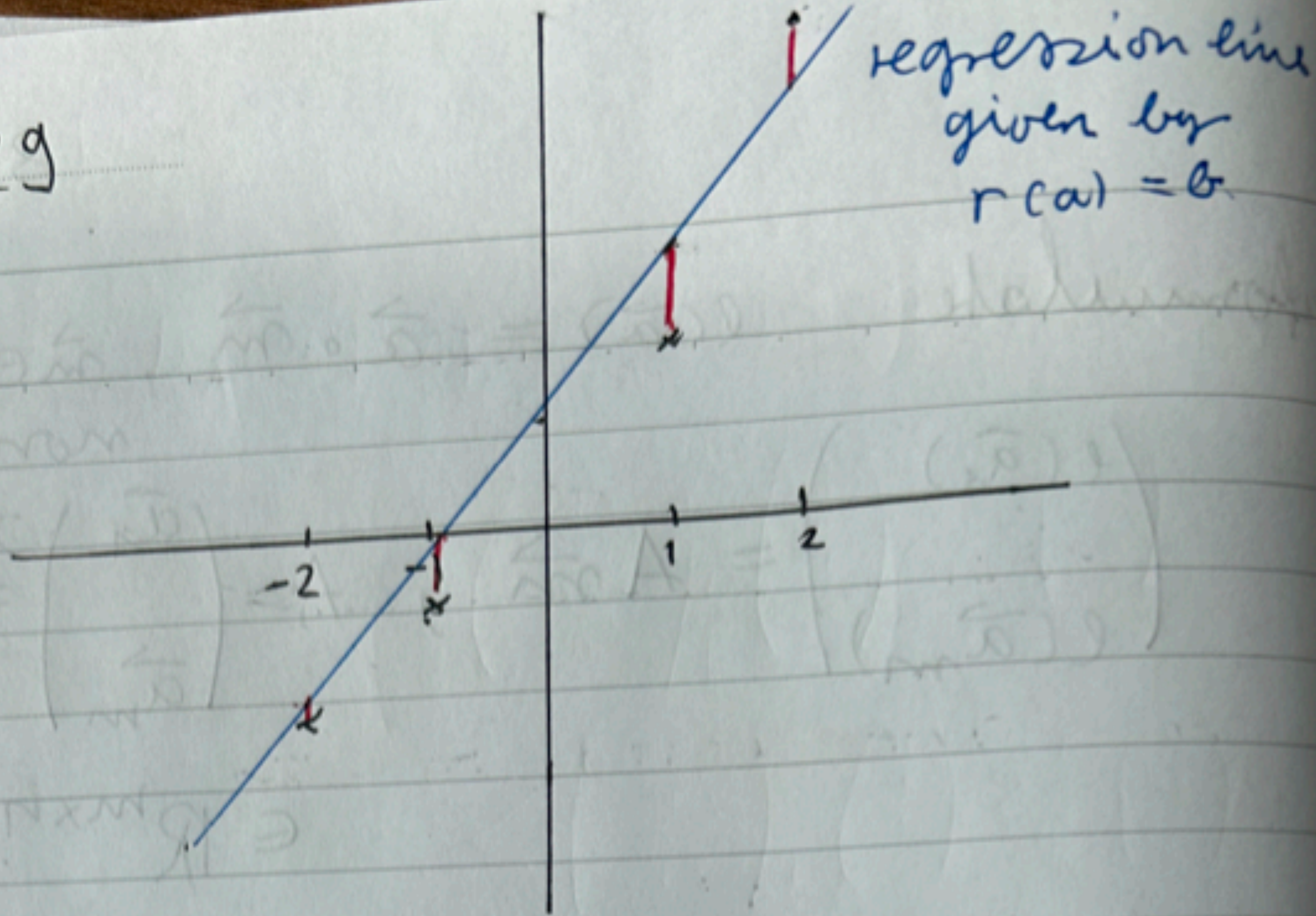
Find $\vec{m}_0 \in \mathbb{R}^{n \times 1}$ such that

$$\|b - A\vec{m}_0\| = \min_{\vec{m} \in \mathbb{R}^{n \times 1}} \|b - A\vec{m}\|.$$

Example 277: (regression)

data

a_i	1	2	-1	-2
b_i	1.5	4	-0.5	-1.5



regression line r: minimize the sum of the squares of the red distances.

$$r(a) = m_1 a + m_0 \quad , \; m_1, m_0 \in \mathbb{R}$$
$$= (a, 1) \circ \begin{pmatrix} m_1 \\ m_0 \end{pmatrix}$$

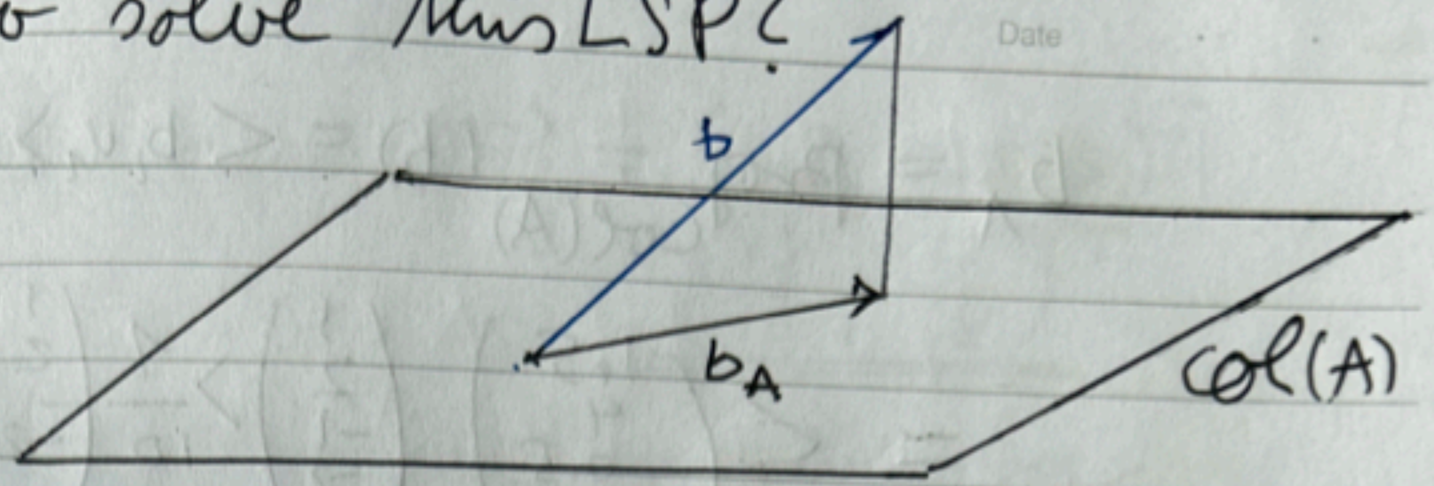
We get LSP: minimize $\|b - A\vec{m}\|$ over $\vec{m} \in \mathbb{R}^2$.

with $A := \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ a_3 & 1 \\ a_4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix}$

and $b := \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 4 \\ -0.5 \\ -1.5 \end{pmatrix}$

How to solve this LSP?

Idea 1:



Project b onto col(A), $b_A := proj_{col(A)}(b)$,

and solve $A\vec{x} = b_A$

End of Lecture 27th Dec 23

Step 1:

Compute b_A:

$A = QR$: $v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

- $w_1 := v_1$, $u_1 := \frac{1}{\|w_1\|} w_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$

- $w_2 = v_2 - \langle v_2, u_1 \rangle u_1$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - 0 \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix} \frac{1}{\sqrt{10}} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$u_2 := \frac{1}{\|w_2\|} w_2 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

We only need $Q = (u_1, u_2) = \begin{pmatrix} 1/\sqrt{10} & 1/2 \\ 2/\sqrt{10} & 1/2 \\ -1/\sqrt{10} & 1/2 \\ -2/\sqrt{10} & 1/2 \end{pmatrix}$

(Check $Q^T Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$)

No. 331

Date

$$b_A = \text{proj}_{Col(A)}(b) = \langle b, u_1 \rangle u_1 + \langle b, u_2 \rangle u_2$$

$$= \left\langle \begin{pmatrix} 1.5 \\ 4 \\ -0.5 \\ -1.5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix} \right\rangle \frac{1}{10} \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$$

$$+ \left\langle \begin{pmatrix} 1.5 \\ 4 \\ -0.5 \\ -1.5 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle \frac{1}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1}{10}(1.5 + 8 + 0.5 + 3) \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}$$

$$+ \frac{1}{4}(1.5 + 4 - 0.5 - 1.5) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{13}{10} \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix} + \frac{3.5}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{40} \begin{pmatrix} 87 \\ 139 \\ -17 \\ -69 \end{pmatrix}$$

Step 2:

Solve $A\vec{m} = b_A$:

$$\left(\begin{array}{cc|c} 1 & 1 & 87/40 \\ 2 & 1 & 139/40 \\ -1 & 1 & -17/40 \\ -2 & 1 & -69/40 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cc|c} 1 & 1 & 87/40 \\ 0 & -1 & -35/40 \\ 0 & 2 & 70/40 \\ 0 & 3 & 105/40 \end{array}\right) \rightarrow \left(\begin{array}{cc|c} 1 & 0 & 52/40 \\ 0 & 1 & 35/40 \end{array}\right)$$

$u_2 > u_2$

The solution is $\vec{m}_* = (1.3, 0.875)^T$
$= \left(\frac{13}{10}, \frac{7}{8}\right)^T$

The regression line is given by

$$r(a) = \frac{13}{10} a + \frac{7}{8}.$$

(Note we have computed $\vec{m}_*$ s.t.

$$\|b - A\vec{m}_*\| = \min_{\vec{m} \in \mathbb{R}^{2\times 1}} \|b - A\vec{m}\|.)$$

To complicated!

Idea 2: Heuristic:

We want $\vec{m}$ s.t. $b \approx A\vec{m}$

Trick: Multiply with A^T from left and solve

$$A^T A \vec{m} = A^T b.$$

If $(A^T A)$ invertible: then
$$\vec{m} = (A^T A)^{-1} A^T b.$$

No. 333

Does this work? Answer: Yes!

Theorem 278: Let $A \in \mathbb{R}^{m \times n}$ of rank n and $b \in \mathbb{R}^{m \times 1}$. Then

(i) $A^T A$ is invertible

(ii) $\vec{m}_* := (A^T A)^{-1} A^T b$ is the unique solution of the LSP

$$\min_{\vec{m} \in \mathbb{R}^{n \times 1}} \| b - A\vec{m} \|.$$

Finish Example 277 first:

$$A^T A = \begin{pmatrix} 1 & 2 & -1 & -2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 10 & 0 \\ 0 & 4 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} 1 & 2 & -1 & -2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1.5 \\ 4 \\ -0.5 \\ -1.5 \end{pmatrix}$$

$$= \begin{pmatrix} 13 \\ 3.5 \end{pmatrix}$$

$$(A^T A | A^T b) = \left(\begin{array}{cc|c} 10 & & 13 \\ & 4 & 3.5 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & & 1.3 \\ & 1 & 7/8 \end{array}\right)$$

We get $\vec{m}_* = (1.3, 0.875)^T$

Proof (Theorem 278): $A = QR$

Q-R decomposition $Q = (u_1, \ldots, u_n)$

$$b_A = \mathrm{proj}_{\mathrm{col}(A)}(b) \in \mathrm{col}(A).$$

So $\exists\, \vec{m} \in \mathbb{R}^{n\times 1}$: $b_A = A\vec{m}$.

For such $\vec{m}$ we have:

$$A\vec{m} = b_A = \sum_{i=1}^{n} \langle b, u_i \rangle u_i$$

$$= Q \begin{pmatrix} \langle b, u_1 \rangle \\ \vdots \\ \langle b, u_n \rangle \end{pmatrix} = Q Q^T b$$

$$\Rightarrow \quad A^T A \vec{m} = R^T \underbrace{(Q^T Q)}_{I_n} Q^T b = A^T b$$

Exercise: $A^T A$ is invertible

$$\Rightarrow \quad \vec{m} = (A^T A)^{-1} A^T b \qquad \square$$

Def 279: Let $A \in \mathbb{R}^{m\times n}$ of rank n.

$A^{+} := \mathrm{Pen}(A) := (A^T A)^{-1} A^T$ is called the "Penrose inverse" of A.

3)
'8

VI.3 Application: QR for solving LSP

Question 280: Does the QR-decomposition simplify solving an LSP?
Answer: Yes.

Theorem 281: (LSP with QR)
Given $b \in \mathbb{R}^{m\times 1}$, $A \in \mathbb{R}^{m\times n}$ of rk n with $A = QR$. (Q-R-decomp.).
Then the LSP has a unique solution which is

$$\vec{x}_* = R^{-1} Q^T b.$$

Proof: Instead of $A\vec{x} \overset{LSP}{=} b$ (I)
consider $Qy \overset{LSP}{=} b$ (II)

$(y = R\vec{x})$

Thm 278 $\Rightarrow$ The unique least square sol^n of (II) is $\underbrace{(Q^TQ)^{-1}Q^Tb}_{= Q^Tb} =: y_*$.

(Note: $Q^TQ = I_n$)

$\Rightarrow$ The unique least square sol^n of (I) is $\vec{x}_* = R^{-1}y_* = R^{-1}Q^Tb$ $\square$

Thm

Example 282: (related to Ex 277)

$$A = \begin{array}{c} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix} \\ \quad v_1 \quad v_2 \end{array} = \underbrace{\begin{array}{c} \quad u_1 \qquad u_2 \\ \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{1}{2} \\ \frac{2}{\sqrt{10}} & \frac{1}{2} \\ -\frac{1}{\sqrt{10}} & \frac{1}{2} \\ -\frac{2}{\sqrt{10}} & \frac{1}{2} \end{pmatrix} \end{array}}_{Q} R$$

with

$$R = \begin{pmatrix} \langle u_1, v_1 \rangle & \langle u_1, v_2 \rangle \\ & \langle u_2, v_2 \rangle \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{10} & 0 \\ & 2 \end{pmatrix}$$

Thm 281 $\Longrightarrow$

$$\vec{m}_x = \begin{pmatrix} \frac{1}{\sqrt{10}} & \\ & 2^{-1} \end{pmatrix} Q^T \begin{pmatrix} 1.5 \\ 4 \\ -0.5 \\ -1.5 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{10}} & \\ & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{10}} & \frac{2}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & -\frac{2}{\sqrt{10}} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1.5 \\ 4 \\ -0.5 \\ -1.5 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{10}(1.5 + 8 + 0.5 + 3) \\ \frac{1}{4}(1.5 + 4 - 0.5 - 1.5) \end{pmatrix} = \begin{pmatrix} 13/10 \\ 7/8 \end{pmatrix}.$$

$\Longrightarrow$ $r(a) = \frac{13}{10} a + \frac{7}{8}$ is the sol^n.