

Session 4: Norms, Inner Products, Orthogonality

1 Norms

Euclidian Norm

$$||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$$

Definition of Norms:(let V be a vector space)

1. Homogeneity: $||\alpha v|| = |\alpha| \times ||v||$ for all $\alpha \in \mathbb{R}^n$ and $v \in V$
2. Positive Definiteness: if $||v|| = 0$ for some v then $v = 0$
3. Triangular Inequality: $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in V$

2 Inner Products

Definition of Inner Products (let V be a vector space)

1. Symmetry: $\langle u, v \rangle = \langle v, u \rangle$ for all $u, v \in V$
2. Linearity: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ and $\langle \alpha v, w \rangle = \alpha \langle v, w \rangle$ for all $v, u, w \in V$ and $\alpha \in \mathbb{R}$
3. Positive Definiteness: $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$

Proposition: If $\langle \cdot, \cdot \rangle$ is an inner product on V then

$$||v|| = \sqrt{\langle v, v \rangle}$$

is a norm on V . We say that the norm $||\cdot||$ is induced by the inner product $\langle \cdot, \cdot \rangle$

Cauchy-Schwartz Inequality Let $||\cdot||$ be the norm induced by the inner product $\langle \cdot, \cdot \rangle$ on the vector space V . Then for all $x, y \in V$:

$$|\langle x, y \rangle| \leq ||x|| \times ||y||$$

Moreover, there is equality if and only if x and y are linearly dependent (i.e. $x = \alpha y$ or $y = \alpha x$ for some $\alpha \in \mathbb{R}$)

3 Orthogonality

Definition of Orthogonality: Let V be a vector space and $\langle \cdot, \cdot \rangle$ be an inner product on V

- We say that vectors x, y are orthogonal if $\langle x, y \rangle = 0$. We write $x \perp y$
- We say that vector x is orthogonal to the set of vectors A if x is orthogonal to all of the vectors in A . We write $x \perp A$

For a family of vectors $\{v_1, \dots, v_n\}$:

- The family is orthogonal if $\langle v_i, v_j \rangle = 0$ for all $i \neq j$
- The family is orthonormal if all the vectors are orthogonal and all of the v_i have unit norm $\|v_1\| = \dots = \|v_k\| = 1$

Proposition: A vector space of finite dimension admits an orthonormal basis

Proposition: Assume that $\dim(V) = n$ and let v_1, \dots, v_n be an orthonormal basis of V . Then the coordinates of a vector $x \in V$ in the basis v_1, \dots, v_n are

$$x = \langle x, v_1 \rangle v_1 + \dots + \langle x, v_n \rangle v_n$$

Pythagorean Theorem: Let $\|\cdot\|$ be the norm induced by $\langle \cdot, \cdot \rangle$ for all $x, y \in V$ we have:

$$x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2$$

Orthogonal Projection Let S be the subspace of \mathbb{R}^n . The orthogonal projection of a vector x onto S is defined as the vector $P_S(x)$ in S that minimizes the distance to x :

$$P_S(x) = \operatorname{argmin}_{y \in S} \|x - y\| \text{ for } y \in S$$

The distance from x to the subspace S is defined by:

$$d(x, S) = \min \|x - y\| = \|x - P_S(x)\|$$

Proposition

Let S be a subspace of \mathbb{R}^n and let (v_1, \dots, v_k) be an **orthonormal basis** of S . Then for all $x \in \mathbb{R}^n$,

$$P_S(x) = \langle v_1, x \rangle v_1 + \dots + \langle v_k, x \rangle v_k.$$

Consequences

Let $V = \begin{pmatrix} | & & | \\ v_1 & \dots & v_k \\ | & & | \end{pmatrix} \overset{\mathbb{R}^{n \times k}}{\text{gather the orthonormal basis-vectors of the subspace } S}$

Proposition

The orthogonal projection is given by $P_S(x) = VV^\top x$.

✚ P_S is a linear transform.

🌀 VV^\top is its matrix.

4 Proofs:

PROBLEM 2.2:

Recall that to prove that two sets are equal ($A=B$) can prove that $A \subseteq B$ and $B \subseteq A$.

- ⊙ let $x \in \text{Im}(A)$, by definition there exists $v \in \mathbb{R}^n$ such that $x = Av$

$$= \begin{pmatrix} | & & | \\ c_1 & \dots & c_n \\ | & & | \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$$= v_1 c_1 + v_2 c_2 + \dots + v_n c_n$$

So that $x \in \text{Span}(c_1, \dots, c_n)$.

This shows that $\text{Im}(A) \subseteq \text{Span}(c_1, \dots, c_n)$

- ⊙ let $x \in \text{Span}(c_1, \dots, c_n)$, so x is a linear combination of c_1, \dots, c_n : there exists $\alpha_1, \dots, \alpha_n$ in \mathbb{R} such that

$$x = \alpha_1 c_1 + \dots + \alpha_n c_n$$

$$= \begin{pmatrix} | & & | \\ c_1 & \dots & c_n \\ | & & | \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

\nearrow call the vector $v \in \mathbb{R}^n$

$$= Av \Rightarrow x \in \text{Im}(A)$$

This proves $\text{Span}(c_1, \dots, c_n) \subseteq \text{Im}(A)$

- ⊙ Overall conclusion: Since $\text{Span}(c_1, \dots, c_n) \subseteq \text{Im}(A)$ and $\text{Im}(A) \subseteq \text{Span}(c_1, \dots, c_n)$ we can conclude that $\text{Im}(A) = \text{Span}(c_1, \dots, c_n)$

Figure 1: $\text{Im}(A) = \text{span of columns}$ proof

PROBLEM 1.5

On side obvious $\textcircled{A} \quad V = G \Rightarrow \begin{cases} \dim(V) = \dim(G) \\ V \subset G \end{cases} \textcircled{B}$

Other side:

\textcircled{B}
Assume $\begin{cases} \dim(V) = \dim(G) = n \\ V \subset G \end{cases}$

There exists a basis (v_1, \dots, v_n) of V .

Since $V \subset G$, (v_1, \dots, v_n) is also a family of linearly independent vectors of G .

$\Rightarrow (v_1, \dots, v_n)$ is a basis of G since $\dim G = n$

$\Rightarrow G = \text{Span}(v_1, \dots, v_n) = V \quad \textcircled{A}$

Conclusion $\textcircled{A} \Rightarrow \textcircled{B}$ and $\textcircled{B} \Rightarrow \textcircled{A} \Rightarrow A \Leftrightarrow B$

Figure 2: subspaces are the same example 2

An inequality

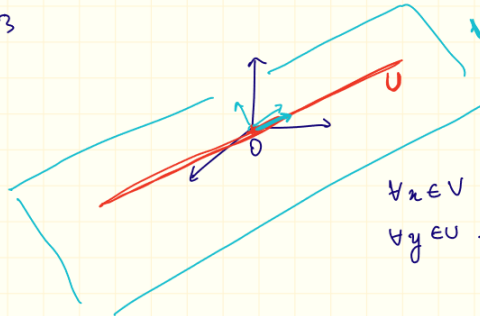
Proposition

Let U and V be two subspaces of \mathbb{R}^n . Assume that $U \subset V$. Then

$$\dim(U) \leq \dim(V) \leq n.$$

If moreover $\dim(U) = \dim(V)$, then $U = V$. PROOF (*)

$$\dim \mathbb{R}^3 = 3$$



$$\begin{aligned} \forall x \in V &\Rightarrow x \in \mathbb{R}^3 \Rightarrow V \subset \mathbb{R}^3 \\ \forall y \in U &\Rightarrow y \in V \Rightarrow U \subset V \end{aligned}$$

Figure 3: subspaces are the same

2.2 Inequalities

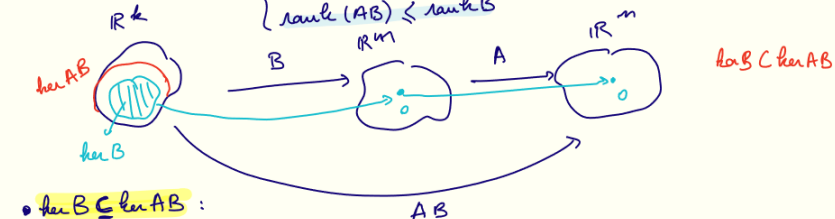
Proposition

$C = \subseteq$

Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times k}$. Then the following holds

1. $\text{rank}(A) \leq \min(n, m)$.
2. $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$.

Proof. Show that $\begin{cases} \text{rank}(AB) \leq \text{rank}(A) \\ \text{rank}(AB) \leq \text{rank}(B) \end{cases} \longrightarrow \text{HW Q}^\circ$



• $\text{ker } B \subseteq \text{ker } AB$:

for $x \in \text{ker } B$, $Bx = 0 \Rightarrow ABx = 0 \Rightarrow x \in \text{ker } AB$.

• $\dim \text{ker } AB + \text{rank } AB = \dim \text{ker } B + \text{rank } B = k$ By the rank nullity theorem.

$\dim \text{ker } B \leq \dim \text{ker } AB$
 $k - \text{rank } B \leq k - \text{rank } AB \longrightarrow \text{rank } B \geq \text{rank } (AB) \quad \square$

2. The rank-nullity theorem 2.2 Inequalities

17/27

Figure 4: rank nullity

PROBLEM 3.4

$$(a) \begin{cases} \text{rank}(A) = \dim \text{Im}(A) \\ \text{rank}(AB) = \dim \text{Im}(AB) \end{cases}$$

$$\text{yet } \text{Im}(AB) \subset \text{Im}(A) \Rightarrow \text{rank}(AB) \leq \text{rank } A$$

$$(b) \quad \pm \text{ trivial to show } \ker(L) \subset \ker(L^T L)$$

$$\begin{aligned} \pm \text{ now for any } x \in \ker(L^T L) \quad & L^T L x = 0 \\ & x^T L^T L x = 0 \\ \Rightarrow \quad & \|Lx\|^2 = 0 \\ \Rightarrow \quad & Lx = 0 \\ \Rightarrow \quad & x \in \ker(L) \end{aligned}$$

$$\text{so } \ker(L^T L) \subset \ker L$$

$$\text{Conclusion: } \ker(L^T L) = \ker L$$

$$(c) \quad \text{rank}(L^T L) \leq \text{rank}(L^T)$$

$$\Rightarrow m - \dim \ker(L^T L) \leq \text{rank}(L^T) \quad \text{rank nullity theorem}$$

$$\Rightarrow m - \dim \ker(L) \leq \text{rank}(L^T)$$

$$\Rightarrow \text{rank}(L) \leq \text{rank}(L^T)$$

⊕ apply the same inequality to L^T :

$$\text{rank}(L^T) \leq \text{rank}(\underbrace{(L^T)^T}_L)$$

$$\text{conclusion } \text{rank}(L) \leq \text{rank}(L^T) \leq \text{rank}(L)$$

$$\hookrightarrow \text{rank}(L) = \text{rank}(L^T)$$

Figure 5: rank L = rank L^T proof

Why?

x_1, \dots, x_k linearly dependent \Leftrightarrow one of them is a linear combination of the others - B

① Assume x_1, \dots, x_k linearly dependent: $A \Rightarrow B$

exists $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ such that $\alpha_1 x_1 + \dots + \alpha_k x_k = 0$
not all 0

$\alpha_i \neq 0 \Rightarrow \alpha_i x_i = -\alpha_1 x_1 - \dots - \alpha_k x_k$ (no term in x_i)

$\Rightarrow x_i = -\frac{\alpha_1}{\alpha_i} x_1 - \dots - \frac{\alpha_k}{\alpha_i} x_k$ since $\alpha_i \neq 0$

② Assume $x_i = \beta_1 x_1 + \dots + \beta_k x_k$ (no term i) $\beta_1, \dots, \beta_k \in \mathbb{R}$

$B \Rightarrow A$

$\Rightarrow \beta_1 x_1 + \dots - x_i + \dots + \beta_k x_k = 0 \rightarrow$ linearly dependent -
 \downarrow
 $\beta_i = -1 \neq 0$

$A \Leftrightarrow B$

2. Span & linear dependency 2.2 Linear dependency

20/39

Figure 6: how to prove linear dependence.png

3.2 Pythagorean Theorem

Theorem (Pythagorean theorem)

Let $\|\cdot\|$ be the norm induced by $\langle \cdot, \cdot \rangle$. For all $x, y \in V$ we have

$$x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Proof. $x \perp y \Rightarrow \|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + 2 \underbrace{\langle x, y \rangle}_{=0 \text{ since } x \perp y} + \|y\|^2$
 $= \|x\|^2 + \|y\|^2$

Figure 7: pythagorean theorem

(c) As $P_S(x) - x$ and $P_S(x) \in S$ are orthogonal (previous question) we can use Pythagorean theorem:

$$\|x\|^2 = \|x - P_S(x)\|^2 + \|P_S(x)\|^2$$

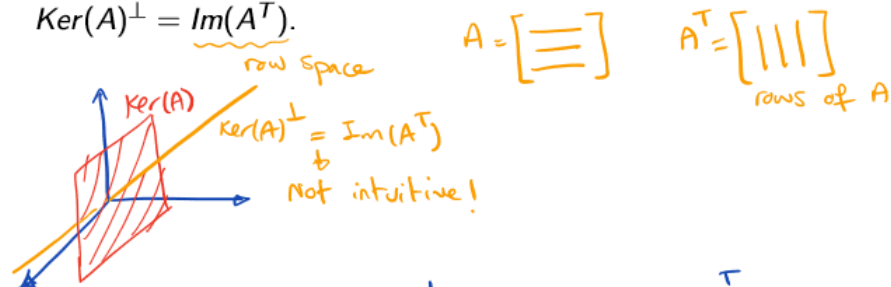
$$\Rightarrow \|x\|^2 \geq \|P_S(x)\|^2$$

$$\Rightarrow \|x\| \geq \|P_S(x)\|$$

steps allowed because $\|x\|$ and $\|P_S(x)\|$ are both positive -

Figure 8: norm of proj x leq norm of x

6. Let $A \in \mathbb{R}^{n \times m}$. Assume the Euclidean inner product. Prove that $\text{Ker}(A)^\perp = \text{Im}(A^T)$.



(i) show $\text{Im}(A^T) \subseteq \text{Ker}(A)^\perp \iff \forall x \in \text{Im}(A^T) \Rightarrow x \in \text{Ker}(A)^\perp$

for $\forall x \in \text{Im}(A^T) \exists y$ s.t. $x = A^T y$

$$\forall z \in \text{Ker}(A), \langle x, z \rangle = x^T z = (A^T y)^T z = y^T \underbrace{A z}_{=0} = 0$$

That is, x is orthogonal to all vectors in $\text{Ker}(A)$

$$\Rightarrow x \in \text{Ker}(A)^\perp$$

(ii) show $\text{Ker}(A)^\perp \subseteq \text{Im}(A^T)$ \leftarrow Not easy to show

So use result from problem (5). The above is equivalent

$$\text{to } \text{Im}(A^T)^\perp \subseteq \text{Ker}(A) \iff \forall x \in \text{Im}(A^T)^\perp \Rightarrow x \in \text{Ker}(A)$$

Show this!

$$\text{Let } y \in \text{Im}(A^T) \text{ \& } x \in \text{Im}(A^T)^\perp \Rightarrow \langle x, y \rangle = x^T y = 0$$

$$\text{and } \exists v \text{ s.t. } y = A^T v \Rightarrow \langle A^T v, x \rangle = (A^T v)^T x = v^T A x = 0$$

not zero \leftarrow

$$\Rightarrow Ax = 0$$

$$\Rightarrow x \in \text{Ker}(A)$$

Figure 9: complement to $\text{ker}(a) = \text{im}(a^t)$

PROBLEM 10

(a) as usual

(b) We can use the rank nullity theorem for the linear transformation corresponding to the orthogonal projection on S .

$$\begin{cases} \text{Im}(P_S) = S \\ \text{Ker}(P_S) = S^\perp \end{cases}$$

(c) For any $u \in \mathbb{R}^n$ $P_S(u) \in S$ and $u - P_S(u) \in S^\perp$

$$\begin{aligned} u &= P_S(u) + (u - P_S(u)) \\ &= x + y \quad \square \end{aligned}$$

Figure 10: dim s complement stuff (not that great tbh)

let y be a vector of S , and (v_1, \dots, v_n) an orthonormal basis of S .

We have,

$$y = \langle v_1, y \rangle v_1 + \dots + \langle v_n, y \rangle v_n$$

$$P_S(x) = \langle v_1, x \rangle v_1 + \dots + \langle v_n, x \rangle v_n$$

$$\text{So that } \langle y, x \rangle = \langle \langle v_1, y \rangle v_1 + \dots + \langle v_n, y \rangle v_n, x \rangle$$

$$= \langle v_1, y \rangle \langle v_1, x \rangle + \dots + \langle v_n, y \rangle \langle v_n, x \rangle$$

$$\text{and } \langle y, P_S(x) \rangle = \langle \langle v_1, y \rangle v_1 + \dots, \langle v_1, x \rangle v_1 + \dots \rangle$$

would have all crossed terms:

$$\langle v_i, y \rangle \langle v_i, v_j \rangle \langle v_j, x \rangle$$

$$= 0 \text{ if } i \neq j$$

$$= 1 \text{ if } i = j$$

so that only remains: because v_1, \dots, v_n orthonormal

$$\begin{aligned} \langle y, P_S(x) \rangle &= \langle v_1, y \rangle \langle v_1, x \rangle + \dots + \langle v_n, y \rangle \langle v_n, x \rangle \\ &= \langle y, x \rangle \end{aligned}$$

Figure 11: dot product of xy

PROOF: $P_S(x) = \arg \min_{y \in S} \|x - y\|$

let $y \in S$, $y = \alpha_1 v_1 + \dots + \alpha_k v_k$ for some $\alpha_1 \dots \alpha_k$ in \mathbb{R}

$$\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$$

$$\odot \langle x, y \rangle = \langle x, \alpha_1 v_1 + \dots + \alpha_k v_k \rangle = \sum_{i=1}^k \alpha_i \langle x, v_i \rangle$$

$$\odot \|y\|^2 = \sum_{i=1}^k \alpha_i^2$$

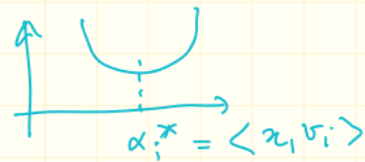
Proof

$$\min_y \|x - y\|^2 \Leftrightarrow \min_y -2\langle x, y \rangle + \|y\|^2$$

$$\Leftrightarrow \min_{\alpha_1 \dots \alpha_k} \sum_{i=1}^k \alpha_i^2 - 2\alpha_i \langle x, v_i \rangle$$

$$\downarrow f_i(\alpha_i) = \alpha_i^2 - 2\alpha_i \langle x, v_i \rangle$$

find α_i^* such that $f_i(\alpha_i^*)$ is minimum



Conclude: minimizer y given by

$$y = \langle x, v_1 \rangle v_1 + \dots + \langle x, v_k \rangle v_k = P_S(x)$$

□

Figure 12: argmin projection proof