

## Chapter 5

# Inner Product Spaces

- 10/6:
- We define

$$\ell^2(\mathbb{R}) = \left\{ \{a_n\}_{n \geq 1} \subset \mathbb{R} : \sum_1^\infty |a_n|^2 < \infty \right\}$$

- **Inner product:** A map  $V \times V \rightarrow \mathbb{F}$  that takes  $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \cdot \mathbf{y}$ . Denoted by  $\cdot, (\cdot, \cdot), \langle \cdot, \cdot \rangle$ .
- Properties of the inner product:
  - $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$  (symmetry).
  - $(\alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z}) = \alpha(\mathbf{x}, \mathbf{z}) + \beta(\mathbf{y}, \mathbf{z})$  (linearity).
  - $(\mathbf{x}, \mathbf{x}) \geq 0$ .
  - $(\mathbf{x}, \mathbf{x}) = 0$  iff  $\mathbf{x} = 0$ .
- If  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , then

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i$$

- If  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ , then

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i \bar{y}_i$$

- If  $f, g \in \mathbb{P}_n(t)$ , then

$$(f, g) = \int_{-1}^1 f \bar{g} dt$$

- The conjugate of a polynomial is the polynomial with the conjugate of the coefficients of the original polynomial. Symbolically, if  $f = \sum_{i=0}^n \alpha_i t^i$  is a polynomial, then  $\bar{f} = \sum_{i=0}^n \bar{\alpha}_i t^i$ .
- It is a fact that

$$\left| \sum_{n=1}^{\infty} a_n \bar{b}_n \right| \leq \|(a_n)_{n \geq 1}\| \|(b_n)_{n \geq 1}\|$$

- Suppose we want to define the inner product between two matrices.
  - A common one is

$$(A, B) = \text{tr}(B^* A)$$

where  $B^* = \bar{B}^T = \overline{B^T}$  is the conjugate transpose.

- We define the norm as a function  $V \rightarrow [0, \infty)$  given by

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$$

- Properties of the norm.

- $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ .
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ .
- $\|\mathbf{x}\| = 0$  iff  $\mathbf{x} = 0$ .

- In  $\mathbb{R}^n$ ,

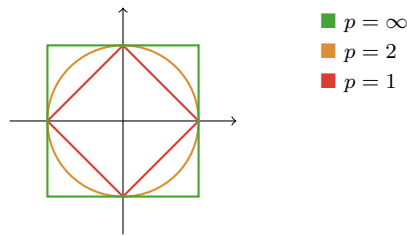


Figure 5.1: The unit ball of norms corresponding to  $p = 1, 2, \infty$ .

- The standard norm is

$$\|\mathbf{x}\| = \sqrt{\sum |x_i|^2}$$

- We can also define

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum |x_i|^p}$$

- We can even define

$$\|\mathbf{x}\|_\infty = \max |x_i|$$

- And we can prove that all of these are valid norms.
- Only the norm corresponding to  $\ell^2$  is given by an inner product, but all the other quantities are still norms as defined by the properties (see Treil (2017)).
- Figure 5.1 shows the unit ball of each norm, i.e., the set of all points which have norm 1.

- The parallelogram rule:

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$$

- Orthogonality: Given  $\mathbf{v}, \mathbf{w}$ , if  $\mathbf{v} \perp \mathbf{w}$ , then  $(\mathbf{v}, \mathbf{w}) = 0$ .

- In particular, if  $\mathbf{v} \perp \mathbf{w}$ , then

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

- Let  $E$  be a subspace of  $V$ . If  $\mathbf{v} \perp E$ , then  $\mathbf{v} \perp \mathbf{e}$  for all  $\mathbf{e} \in E$ , i.e.,  $\mathbf{v} \perp$  a set of vectors spanning  $E$ .
- Any set of orthogonal vectors is linearly independent. Thus, if  $V$  is  $n$  dimensional, then  $\mathbf{v}_1, \dots, \mathbf{v}_n$  orthogonal is a basis.
- Let  $E$  be a subspace of  $V$ . Take  $\mathbf{v} \in V$ . We want to define the projection  $P_E \mathbf{v}$  of  $\mathbf{v}$  onto  $E$ .
  - We have that  $P_E \mathbf{v} \in E$  and  $\mathbf{v} - P_E \mathbf{v} \perp E$ .
  - Additionally, we have that

$$\|\mathbf{v} - P_E \mathbf{v}\| \leq \|\mathbf{v} - \mathbf{e}\|$$

for all  $\mathbf{e} \in E$ .

- Lastly, we have that  $P_E \mathbf{v}$  is unique.
- If we receive a basis of a vector space, how do we create out of that a basis that is orthogonal? The process of doing this is called **Gram-Schmidt orthogonalization**.
  - We keep  $\mathbf{v}_1$ , subtract  $P_{\mathbf{v}_1} \mathbf{v}_2$  from  $\mathbf{v}_2$ , subtract  $P_{\{\mathbf{v}_1, \mathbf{v}_2\}} \mathbf{v}_3$  from  $\mathbf{v}_3$ , and on and on.
- If we are given a set of orthogonal vectors, we can normalize them by dividing each by its norm. This creates an orthonormal list. The standard basis is orthonormal.
- Let

$$E^\perp = \{v \in V : v \perp E\}$$

- It follows that  $V = E \oplus E^\perp$ .
- How close can we come to solving  $A\mathbf{x} = \mathbf{b}$  if we cannot solve it exactly (i.e., if the columns are not linearly independent)?
  - Let  $A$  be an  $m \times n$  matrix, and let  $\mathbf{b} \in \mathbb{R}^m$ .
  - Then the best solution is given by minimizing  $\|A\mathbf{x} - \mathbf{b}\|$ . We minimize this with projections. A special case of this is least squares regression! More details in Treil (2017).

10/8:

- Soug is gonna send us a hefty amount of reading for the weekend.
- Least square approximation:
  - If we want to minimize  $\|A\mathbf{x} - \mathbf{b}\|$ , the best we can do is project  $\mathbf{b}$  onto the range of  $A$ .
  - Let  $\mathbf{v}_1, \dots, \mathbf{v}_k$  be an orthogonal basis of range  $A$ .
  - Then

$$\text{Proj}_{\text{range } A} \mathbf{b} = \sum_{k=1}^k \frac{(\mathbf{b}, \mathbf{v}_k)}{\|\mathbf{v}_k\|^2} \mathbf{v}_k$$

- Matrix equation form:

$$\text{Projection}_{\text{range } A} = A(A^*A)^{-1}A^*$$

if  $A^*A$  is invertible, where  $A^* = \bar{A}^T$ .

■ Soug never uses this though.

- The minimum is found when  $\mathbf{b} - A\mathbf{x} \perp \text{range } A$ . Implies that  $\mathbf{b} - A\mathbf{x} \perp \mathbf{a}_k$  for all  $k$ . Implies  $(\mathbf{b} - A\mathbf{x}, \mathbf{a}_k) = \bar{\mathbf{a}}_k^T (\mathbf{b} - A\mathbf{x}) = 0$ .
- Note that we're letting  $\bar{\mathbf{a}}_k^T$  be the row vector

$$\bar{\mathbf{a}}_k^T = (\bar{a}_{1,k} \quad \cdots \quad \bar{a}_{n,k})$$

- We also have  $\bar{A}^T (\mathbf{b} - A\mathbf{x}) = 0$ , from which it follows that  $A^*A\mathbf{x} = A^*\mathbf{b}$ , so  $\mathbf{x} = (A^*A)^{-1}A^*\mathbf{b}$ . Thus,  $\text{Proj}_{\text{range } A} = Ax$ , so  $\text{Proj}_{\text{range } A} = A(A^*A)^{-1}A^*\mathbf{b}$ .

- Adjoint of a linear map  $T : V \rightarrow W$  is the  $A^*$  discussed above.
  - First, we'll do this for matrices. And then we'll do it for any finite-dimensional vector space.
  - Let  $A$  be an  $m \times n$  matrix. We claim then that

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$$

for all  $\mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m$ . Proof:

$$\begin{aligned} (A\mathbf{x}, \mathbf{y}) &= \bar{\mathbf{y}}^T A\mathbf{x} \\ &= \mathbf{y}^* A\mathbf{x} \\ &= (A^*\mathbf{y})^* \mathbf{x} \\ &= (\mathbf{x}, A^*\mathbf{y}) \end{aligned}$$

- Properties of the adjoint:

$$(AB)^T = B^T A^T$$

$$(AB)^* = B^* A^*$$

$$(A^*)^* = A$$

- $A^*$  is the unique matrix  $B$  such that  $(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, B\mathbf{y})$ .
- Let  $\mathbf{v}_1, \dots, \mathbf{v}_n$  be a basis of  $V$ , and let  $\mathbf{w}_1, \dots, \mathbf{w}_m$  be a basis of  $W$ .
- Definition of  $A^*$ : If  $(A\mathbf{x}, \mathbf{y}) = (y, A^*\mathbf{x})$  for all  $\mathbf{x} \in V$  and  $\mathbf{y} \in W$ .
- But it's not enough to define something; we have to check that it exists.
- If  $[A]_{AB}$ , then  $[A^*]_{AB}$ .
- More properties (give criteria for solving systems of equations):

$$\ker A^* = (\text{range } A)^\perp$$

$$\ker A = (\text{range } A^*)^\perp$$

$$\text{range } A = (\ker A^*)^\perp$$

$$\text{range } A^* = (\ker A)^\perp$$

■ Soug proves these.

• Isometries and unitary operators.

- $U : X \rightarrow Y$  is an isometry if  $\|\mathbf{x}\| = \|U\mathbf{x}\|$  for all  $\mathbf{x} \in X$ . It is an isometry because it preserves the distance between points.
- It immediately follows that  $\|\mathbf{x}_1 - \mathbf{x}_2\| = \|U\mathbf{x}_1 - U\mathbf{x}_2\| = \|U(\mathbf{x}_1 - \mathbf{x}_2)\|$ .
- This definition is equivalent to an inner product one:  $(\mathbf{x}, \mathbf{y}) = (U\mathbf{x}, U\mathbf{y})$ . This follows from the definition of the norm.
- We have

$$(\mathbf{a}, \mathbf{b}) = \frac{1}{4} \sum_{\alpha=\pm 1, \pm i} \alpha \|\mathbf{a} + \alpha \mathbf{b}\|^2$$

■  $(a+b)^2 - (a-b)^2 = 4ab$  for any  $a, b \in \mathbb{R}$ , so  $ab = \frac{1}{4}[(a+b)^2 - (a-b)^2]$ . Thus, in a real inner product space,

$$(\mathbf{a}, \mathbf{b}) = \frac{1}{4} \left( \|\mathbf{a} + \mathbf{b}\|^2 - \|\mathbf{a} - \mathbf{b}\|^2 \right)$$

■ It follows that isometries preserve inner products.

- $U$  is an isometry if and only if  $U^*U = I$ . Proof:

$$(\mathbf{x}, \mathbf{x}) = (U\mathbf{x}, U\mathbf{y}) = (U^*U\mathbf{x}, \mathbf{x})$$

$$(\mathbf{x}, \mathbf{y}) = (U\mathbf{x}, U\mathbf{y}) = (U^*U\mathbf{x}, \mathbf{y})$$

$$(\mathbf{x}, \mathbf{y}) = (U\mathbf{x}, U\mathbf{y}) = (\mathbf{x}, \mathbf{y})$$

for all  $\mathbf{y}$ .

- An isometry is unitary if it is invertible.

■ Thus,  $U : X \rightarrow Y$  an isometry is unitary iff  $\dim X = \dim Y$ .

- Note that it follows that  $U^* = U^{-1}$  for  $U$  an isometry.
- $U$  unitary implies  $|\det U| = 1$ , so  $\lambda$  an eigenvalue of  $U$  implies that  $|\lambda| = 1$ .
- $A$  is diagonalizable iff it has an orthogonal basis of eigenvectors.