

# ECEn 671: Mathematics of Signals and Systems

## Moon: Chapter 2.

Randal W. Beard

Brigham Young University

September 24, 2020

## Table of Contents

# Section 1

## Metric Spaces

# Spaces

- ▶ One of the objectives of this course is to develop tools that work in a wide variety of settings.
- ▶ We will mostly focus on finite dimensional Hilbert spaces, which include:
  - ▶  $\mathbb{R}^n$ ,  $\mathbb{C}^n$ ,  $\mathbb{C}^{m \times n}$ ,
  - ▶ the set of all functions with finite integral,
  - ▶ the set of all finitely summable sequences,
  - ▶ binary vectors, binary sequences.
- ▶ But does not include important objects like
  - ▶ rotations matrices, quaternions, homogeneous transformations.
- ▶ To make things clear, we will develop the theory systematically in the following order:
  1. Metric space
  2. Norm space / Banach space
  3. Inner product space / Hilbert space

# Metric Spaces

## Definition (Metric Space)

A metric space is a pair  $(\mathbb{X}, d)$  where  $\mathbb{X}$  is a set and

$$d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$$

is a metric defined over  $\mathbb{X}$ .

A metric is a measure of distance between elements in a set.

# Metric Spaces

## Definition (Metric)

Let  $\mathbb{X}$  be a set. Then  $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}$  is a metric if:

$$(M1) \quad d(x, y) = d(y, x), \quad \forall x, y \in \mathbb{X}$$

$$(M2) \quad d(x, y) \geq 0, \quad \forall x, y \in \mathbb{X}$$

$$(M3) \quad d(x, y) = 0, \quad \Longleftrightarrow \quad x = y$$

$$(M4) \quad d(x, z) \leq d(x, y) + d(y, z), \quad \forall x, y, z \in \mathbb{X}$$

(M4) is called the Triangle inequality.

# Examples of Metric Spaces

## Example (E1)

$(\mathbb{R}, d)$  where  $d(x, y) \triangleq |x - y|$  is a metric space.

Note that

- ▶ (M1)  $|x - y| = |y - x|, \forall x, y \in \mathbb{R}.$
- ▶ (M2)  $|x - y| \geq 0, \forall x, y \in \mathbb{R}.$
- ▶ (M3)  $|x - y| = 0$ , if  $x = y$ .
- ▶ (M4)  $|x - z| \leq |x - y| + |y - z| \forall x, y, z \in \mathbb{R}.$

To convince yourself (M4), draw a picture. Note, a picture is not a proof.

# Examples of Metric Spaces

## Example (E2)

$(\mathbb{R}^n, d)$  where

$$d(x, y) \triangleq \left( \sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}$$

where  $x = (x_1, \dots, x_n)^\top$  and  $y = (y_1, \dots, y_n)^\top$ .

Verify that  $d(\cdot, \cdot)$  satisfies (M1)-(M4).



# Examples of Metric Spaces

## Example (E3)

$(\mathbb{R}^n, d)$  where

$$d(x, y) \triangleq \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

where  $p \geq 1$ .

For general  $p \geq 1$ , the triangle inequality is a nontrivial and famous results.

# Examples of Metric Spaces

## Example (E4 bounded sequence space)

Let  $\ell^\infty$  be the set of all sequences of complex numbers where each number is bounded, i.e.,

$$x = (x_1, x_2, x_3, \dots) \in \ell$$

if  $x_i \in \mathbb{C}$  and  $|x_i| < \infty$ .

$(\ell, d)$  is a metric space where

$$d(x, y) = \sup_{j \in \mathbb{N}} |x_j - y_j|.$$

Verify (M1)-(M4).

# Examples of Metric Spaces

## Example (E5 continuous function space)

- ▶ Let  $C[a, b]$  be the set of all continuous functions on  $[a, b]$ , i.e., i.e.  $x \in C[a, b] \Rightarrow x(t)$  is continuous on  $[a, b]$ .

Let

$$d(x, y) = \max_{t \in [a, b]} |x(t) - y(t)|$$

then  $(C[a, b], d)$  is a metric space.

- ▶ This is a different perspective than calculus. In calculus you consider one function at a time. In this class, a function is one point in a larger metric space.

# Examples of Metric Spaces

## Example (E6 discrete metric space)

Let  $\mathbb{X}$  be any set, e.g., the set of three legged dogs, and let

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases}.$$

Then  $(\mathbb{X}, d)$  is a metric space since

- ▶ (M1)  $d(x, y) = d(y, x)$ ,  $\forall x, y \in \mathbb{X}$ .
- ▶ (M2)  $d(x, y) \geq 0$ ,  $\forall x, y \in \mathbb{X}$ .
- ▶ (M3)  $d(x, y) = 0$ , if  $x = y$ .
- ▶ (M4)  $d(x, z) \leq d(x, y) + d(y, z)$   $\forall x, y, z \in \mathbb{X}$ .

# Examples of Metric Spaces

## Example (E7 binary vector space)

Let  $\mathbb{X} = \{0, 1\}^n$  be the set of binary vectors, i.e  $x \in \mathbb{X} \Rightarrow x = (x_1, x_2, \dots, x_n)$  where  $x_i \in \{0, 1\}$ . Let

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

where

$$h(w) = \begin{cases} 1 & \text{if } w \neq 0 \\ 0 & \text{otherwise} \end{cases}.$$

$h$  is called the hamming distance, and simply counts the number of elements in  $x$  and  $y$  that are different.

# Metric Spaces / Norm Spaces / Inner Product Spaces

- ▶ Later in the chapter, we will later introduce the concepts of a norm and a norm space, and an inner product and inner product spaces.
- ▶ Many of the metric spaces introduced above are also norm spaces and inner product spaces, but not all.
- ▶ Metric spaces are the most general of the three.
- ▶ Before introducing the concept of a norm and a normed space, we develop general tools that also work for metric spaces.

## Section 2

## Topology

# Topology

- ▶ In this next section, we develop a set of tools that fall under that category of topology.
- ▶ These tools hold for metric spaces (including norm and inner product spaces).
- ▶ WARNING: There are a lot of definitions. These definitions will help talk formally about things in the future.



# Topology: Open and Closed Sets

## Definition (Ball)

Given a metric space  $(\mathbb{X}, d)$  a  $\delta$ -ball around  $x_0$  is defined to be  $B(x_0, \delta) = \{x \in \mathbb{X} : d(x, x_0) < \delta\}$

## Definition (Interior Point)

A point  $x_o \in \mathbb{X}$  is interior to  $S \subset \mathbb{X}$  if  $\exists \delta > 0$  such that  $B(x_o, \delta) \subset S$ .

## Definition (Open Set)

A set  $\mathbb{X}$  is open if all points in  $\mathbb{X}$  are interior.

## Definition (Closed Set)

A set  $S$  is closed in  $\mathbb{X}$  if  $\mathbb{X} \setminus S$  is open.

# Topology: Convergence

Let  $(\mathbb{X}, d)$  be a metric space.

## Definition (Convergence)

Given a sequence  $\{x_n\}_{n=1}^{\infty}$ , where  $x_n \in \mathbb{X}$ , the following are equivalent

- ▶  $\lim_{n \rightarrow \infty} x_n = x^*$
- ▶  $x_n \rightarrow x^*$
- ▶  $\forall \epsilon > 0, \exists N(\epsilon)$  such that  $n \geq N \Rightarrow d(x_n, x^*) < \epsilon$

A sequence  $\{x_n\}_{n=1}^{\infty}$  in  $\mathbb{X}$  with a limit  $x^* \in \mathbb{X}$  is said to converge.

# Topology: Convergence

Note that a limit may not always exist (similar to min, max)  
For example,  $\lim_{t \rightarrow \infty} \sin(t)$  does not exist.

## Definition (lim sup)

Define lim sup as the largest limit (possibly infinity) of any subsequence.

## Definition (lim inf)

Define lim inf is the smallest limit of all possible subsequences.

## Example

- ▶  $\limsup_{t \rightarrow \infty} \sin(t) = 1$  since the subsequence  $t_n = \frac{k\pi}{2}, k = 1, 5, 9, \dots$  converges to 1
- ▶  $\liminf_{t \rightarrow \infty} \sin(t) = -1$  since the subsequence  $t_n = \frac{k\pi}{2}, k = 3, 7, 11, \dots$  converges to -1

# Topology: Cauchy Sequence

## Definition (Cauchy Sequence)

A sequence  $\{x_n\}_{n=1}^{\infty}$  in a metric space  $(\mathbb{X}, d)$  is said to be a Cauchy sequence if  $\forall \epsilon > 0, \exists N(\epsilon) > 0$  such that  $n, m > N \Rightarrow d(x_n, x_m) < \epsilon$

A sequence is Cauchy if elements in its tail get increasingly closer together. Note that we have not said anything about an element of convergence.

## Theorem

*If a sequence  $\{x_n\}_{n=1}^{\infty}$  in  $\mathbb{X}$  converges to an element  $x^* \in \mathbb{X}$  then it is a Cauchy sequence.*

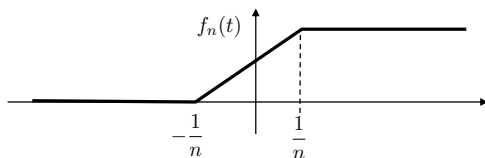
The converse is not true!! I.e., not all Cauchy sequences converge.

# Topology: Cauchy Sequence

Example (from book)

Let  $\mathbb{X} = C[-1, 1]$  and  $d(f, g) = \left( \int_{-1}^1 (f(t) - g(t))^2 dt \right)^{\frac{1}{2}}$

let  $f_n :$



By integration we get:

$$d(f_n, f_m) = \frac{1}{6m^3n} (m^3 + 4m^2n + mn^2 + 2n^3)$$

$$\rightarrow 0 \text{ for } n, m \text{ large } (m > n)$$

but  $f_n$  converges to a discontinuous function which is not in  $\mathbb{X}$ .

This is undesirable

# Topology: Complete Metric Space

## Definition (Complete metric space)

A metric space  $(\mathbb{X}, d)$  is complete if every Cauchy sequence in  $\mathbb{X}$  converges to a value in  $\mathbb{X}$ .

## Implication

$C[a, b]$  with metric  $(\int_a^b |f - g|^2 dt)^{1/2}$  is not complete.

- ▶ Banach spaces are complete normed spaces (discussed later).
- ▶ Hilbert spaces (extremely important in signal processing and control) are complete inner product spaces (discussed later).
- ▶ The importance of  $L_p$  and  $\ell_p$  are that they are complete spaces.

## Section 3

# Vector Spaces

# Vector Spaces

A field is a set of scalars with well defined addition and multiplication operations.

## **Example of fields:**

- ▶  $\mathbb{R}$  with normal addition and multiplication operations
- ▶  $\mathbb{C}$  with complex addition and complex multiplication
- ▶ The set of quaternions, with addition and quaternion multiplication
- ▶ Binary numbers  $\{0, 1\}$  where addition is the “or” operator and multiplication is the “and” operator.



# Vector Spaces

## Definition (Linear Vector Space)

A linear vector space is a pair  $(\mathbb{X}, \mathbb{F})$ , where  $\mathbb{X}$  is a set of objects, and  $\mathbb{F}$  is a field, this is closed under addition and scalar multiplication. i.e.,

- ▶  $x \in \mathbb{X}, \alpha \in \mathbb{F} \Rightarrow \alpha x \in \mathbb{X}$
- ▶  $x, y \in \mathbb{X} \Rightarrow x + y \in \mathbb{X}$ .

By implication

- ▶  $x \in \mathbb{X}, \alpha, \beta \in \mathbb{F} \Rightarrow (\alpha + \beta)x = \alpha x + \beta x \in \mathbb{X}$
- ▶  $x, y \in \mathbb{X}, \alpha \in \mathbb{F} \Rightarrow \alpha(x + y) = \alpha x + \alpha y \in \mathbb{X}$
- ▶  $x, y \in \mathbb{X}, \alpha, \beta \in \mathbb{F} \Rightarrow \alpha x + \beta y \in \mathbb{X}$ .

# Vector Spaces: Subspace

## Definition (Subspace)

A subspace  $V \subset \mathbb{X}$  is a subset of  $\mathbb{X}$  that is also a linear vector space, in particular it contains zero.

**Important property:** A vector space contains a zero element.

# Vector Spaces: Examples

The following are vector spaces:

- ▶  $(\mathbb{R}^n, \mathbb{R})$ ,  $(\mathbb{C}^n, \mathbb{C})$ ,  $(\mathbb{R}^{m \times n}, \mathbb{R})$ ,  $(C[a, b], \mathbb{R})$ ,  $(\ell^\infty, \mathbb{R})$ ,  $(L^\infty, \mathbb{R})$ .

The following are NOT vector spaces:

- ▶ The set  $\mathbb{X} = \mathbb{R} \times [0, 2\pi]$ , (a cylinder) is not a vector space for any field  $\mathbb{F}$ . This is the state space for an inverted pendulum.
- ▶ The set of rotation matrices is not a vector space for any field  $\mathbb{F}$ . This is in the configuration space for robots and satellites.
- ▶ The set of unit quaternions is not a vector space for any field. Quaternions are used extensively in robotics, quantum mechanics, and computer graphics.
- ▶ There are many useful spaces that are NOT linear vector spaces.

# Vector Spaces: Linear Independence

Let  $S$  be a vector space and let  $T \subset S$ . ( $T$  may have uncountable infinite members).  $T$  is linearly independent if for each finite nonempty subset of  $T$ . i.e.,  $\{p_1, \dots, p_n\}$  where  $p_i \in T$ , we have that

$$c_1 p_1 + \dots + c_n p_n = 0 \quad \Longleftrightarrow \quad c_1 = c_2 = \dots = c_n = 0.$$

Otherwise  $T$  is linearly dependent.

# Vector Spaces: Linear Independence

## Example

Let  $S = \mathbb{R}^3$  then the set  $T = \{(1, 0, 0)^\top, (0, 1, 0)^\top\} \subset \mathbb{R}^3$  is linearly independent since

$$c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

if and only if  $c_1 = c_2 = 0$ .

However, the set  $T = \{(1, 1, 0)^\top, (2, 2, 0)^\top\} \subset \mathbb{R}^3$  is linearly dependent since

$$c_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 + 2c_2 \\ c_1 + 2c_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

when  $c_1 = -2$  and  $c_2 = 1$  (as only on example).

# Vector Spaces: Span

## Definition (Span)

Let  $S$  be a vector space, then  $\text{span}(T)$  is the set of all linear combinations of  $T \subseteq S$ .

## Example

$$\text{span} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}$$

## Example

$$\text{span} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} : \alpha, \beta \in \mathbb{R} \right\} = \mathbb{R}^2.$$

# Vector Spaces: Basis

## Definition (Basis)

$T$  is a basis for the vector space  $S$  if  $T$  is linearly independent and  $\text{span}(T) = S$ .

## Definition (Dimension)

The dimension of the vector space  $S$  is the smallest number of linearly independent vectors needed to span  $S$ .

## Example

One possible basis for  $\mathbb{R}^n$  is given by

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

Therefore  $\dim(\mathbb{R}^n) = n$ .

# Vector Spaces: Basis

## Example

One possible basis for  $\ell^\infty$  is given by

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

Therefore  $\dim(\ell^\infty) = \infty$ .



# Vector Spaces: Basis

## Example

The set of all polynomials  $P$  is a vector space with basis

$$\{1, t, t^2, \dots\}$$

Therefore  $\dim(P) = \infty$ .

## Example

The set of all polynomials of degree  $\leq q$   $P^q$  is a vector space with basis

$$\{1, t, t^2, \dots, t^q\}$$

Therefore  $\dim(P^q) = q$ .

## Section 4

### Normed Spaces

# Norms and Normed Spaces

## Definition (Norm)

Let  $S$  be a vector space,  $\|x\|$  is a norm if:

$$(N1) \quad \|x\| \geq 0 \quad \forall x \in S$$

$$(N2) \quad \|x\| = 0 \quad \Leftrightarrow x = 0$$

$$(N3) \quad \|\alpha x\| = \alpha \|x\|$$

$$(N4) \quad \|x + y\| \leq \|x\| + \|y\| \quad (\text{triangle inequality})$$

## Differences between norms and metrics:

- ▶ Norms only have one argument (the length of a vector), where metrics are distances between elements of a set.
- ▶ Norms are only defined for vector spaces!  
(i.e. there is no norm for rotation matrices, but there are metrics!)
- ▶ Norms scale with the vector (N3)  
(there are metrics that don't scale), e.g.

$$d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

- ▶ Every norm is also a metric

$$\|x - y\| = d(x, y)$$

$$\|x\| = d(x, 0)$$

# Definition: Normed Space

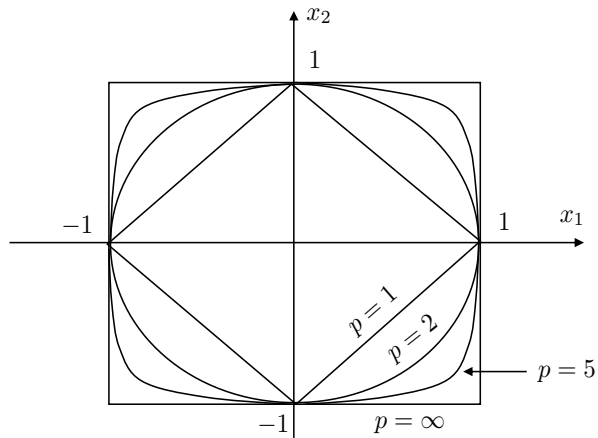
A normed space is a pair  $(\mathbb{X}, \|\cdot\|)$  where  $\mathbb{X}$  is a vector space and  $\|\cdot\|$  is a norm.

## Example (Normed Spaces)

$\mathbb{R}^n$  is a vector space. All of the following norms are valid:

- ▶ one-norm  $\|x\|_1 = \sum_{i=1}^n |x_i|$  (power vectors)
- ▶ two-norm  $\|x\|_2 = \left( \sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}$  (energy vectors)
- ▶ infinity-norm  $\|x\|_\infty = \max_{i=1, \dots, n} |x_i|$  (bounded vectors)
- ▶ p-norm  $\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$

# Unit Circle in $\mathbb{R}^2$



# Normed Space Example: Sequence Spaces

Let  $\ell$  be the set of sequences:  $x = (x_1, x_2, x_3, \dots)$ . The following normed vector spaces can be defined:

- ▶  $\ell_1$ : (power sequences) If  $\|x\|_{\ell_1} = \sum_{i=1}^{\infty} |x_i|$  then  
 $\ell_1 \triangleq \{x \in \ell : \|x\|_{\ell_1} < \infty\}$
- ▶  $\ell_2$ : (energy sequences) If  $\|x\|_{\ell_2} = (\sum_{i=1}^{\infty} |x_i|^2)^{1/2}$  then  
 $\ell_2 \triangleq \{x \in \ell : \|x\|_{\ell_2} < \infty\}$
- ▶  $\ell_{\infty}$ : (bounded sequences) If  $\|x\|_{\ell_{\infty}} = \sup_{j \in \mathbb{N}} |x_j|$  then  
 $\ell_{\infty} \triangleq \{x \in \ell : \|x\|_{\ell_{\infty}} < \infty\}$
- ▶  $\ell_p$ : If  $\|x\|_{\ell_p} = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$  then  $\ell_p \triangleq \{x \in \ell : \|x\|_{\ell_p} < \infty\}$   
for  $1 \leq p \leq \infty$

# Normed Space Examples

## Example

Consider the sequence  $x = (1, 1, 1, \dots)$ :

- ▶  $x \in \ell_\infty$ , but
- ▶  $x \notin \ell_p$  for  $1 \leq p < \infty$ .

## Example

Consider the sequence  $x = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots)$

- ▶  $x \notin \ell_1$  (prove this), but
- ▶  $x \in \ell_p$   $p > 1$  (prove this)



# Normed Space Example: Function Spaces

Let  $L^n(\Omega)$  be the set of functions on  $\Omega$ .  $x \in L^n(\Omega)$  is an equivalent class of functions, i.e. equal except on a set of measure zero. (picture) The following norms are valid:

- ▶  $L_1^n(\Omega)$  (power signals). If  $\|x\|_{L_1^n(\Omega)} = \int_{\Omega} \|x(t)\| dt$ , then  $L_1^n(\Omega) = \{x \in L^n(\Omega) \mid \|x\|_{L_1^n(\Omega)} < \infty\}$ .
- ▶  $L_2^n(\Omega)$  (energy signals). If  $\|x\|_{L_2^n(\Omega)} = \left(\int_{\Omega} \|x(t)\|^2 dt\right)^{1/2}$ , then  $L_2^n(\Omega) = \{x \in L^n(\Omega) \mid \|x\|_{L_2^n(\Omega)} < \infty\}$ .
- ▶  $L_p^n(\Omega)$ . If  $\|x\|_{L_p^n(\Omega)} = \left(\int_{\Omega} \|x(t)\|^p dt\right)^{1/p}$ , then  $L_p^n(\Omega) = \{x \in L^n(\Omega) \mid \|x\|_{L_p^n(\Omega)} < \infty\}$ ,  $1 \leq p \leq \infty$ .
- ▶  $L_{\infty}^n(\Omega)$  (bounded signals). If  $\|x\|_{L_{\infty}^n(\Omega)} = \sup_{t \in \Omega} \|x(t)\|$ , then  $L_{\infty}^n(\Omega) = \{x \in L^n(\Omega) \mid \|x\|_{L_{\infty}^n(\Omega)} < \infty\}$ .

## Section 5

### Inner Product Spaces

# Inner Product Spaces

## Definition (Inner Product)

Let  $S$  be a vector space over  $\mathbb{R}$ . An inner product  $\langle \cdot, \cdot \rangle: S \times S \rightarrow \mathbb{R}$  has the following properties:

$$(IP1) \quad \langle x, y \rangle = \overline{\langle y, x \rangle}$$

$$(IP2) \quad \langle \alpha x, y \rangle = \alpha \langle x, y \rangle$$

$$(IP3) \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$$

$$(IP4) \quad \langle x, x \rangle > 0 \quad \text{if } x \neq 0, \langle x, x \rangle = 0 \Leftrightarrow x = 0$$

## Definition (Inner Product Space)

A vector space with an inner product defined is called an inner-product space.

## Definition (Hilbert Space)

A complete inner-product space is called a Hilbert space.

# Inner Product Spaces: Examples

- ▶  $\mathbb{R}^n$ :  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^T y$  is called the Euclidean inner product.
- ▶  $\mathbb{C}^n$ :  $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y_i} = y^H x$
- ▶  $\mathbb{R}^n$  with the Euclidean inner product is a Hilbert space .
- ▶  $\mathbb{C}^n$  with the Euclidean inner product is a Hilbert space.
- ▶ All finite-dimensional inner-product spaces are Hilbert spaces.

# Inner Product Spaces: Examples

- ▶ Real sequences  $\ell_2$ :  $\langle x, y \rangle_{\ell_2} = \sum_{i=1}^{\infty} x_i y_i$
- ▶ Complex sequences  $\ell_2$ :  $\langle x, y \rangle_{\ell_2} = \sum_{i=1}^{\infty} x_i \overline{y_i}$
- ▶ Both of these examples are Hilbert spaces.

# Inner Product Spaces: Examples

- ▶ Complex function space  $L_2^n(\Omega)$  with inner product:

$$\langle x, y \rangle = \int_{-\infty}^{\infty} y^H(t)x(t) dt$$

is a Hilbert space, but

- ▶ Continuous function  $C[a, b]$  with the same inner product is NOT a Hilbert space.

# Norms vs Inner Products

Every inner product defines a norm (but not vice-versa)

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}}$$

where  $\|\cdot\|$  is called the norm induced by the inner product  $\langle \cdot, \cdot \rangle$ .

## Examples of induced norms

$$\|\cdot\|_2: \langle x, x \rangle^{1/2} = \left( \sum_{i=1}^n x_i^2 \right)^{1/2} = \|x\|_2$$

$$\|\cdot\|_{\ell_2}: \langle x, x \rangle^{1/2} = \left( \sum_{i=1}^{\infty} x_i^2 \right)^{1/2} = \|x\|_{\ell_2}$$

$$\|\cdot\|_{L_2}: \langle x, x \rangle^{1/2} = \left( \int_{\Omega} x^T(t)x(t)dt \right)^{1/2} = \left( \int_{\Omega} \|x(t)\|_2^2 dt \right)^{1/2} = \|x\|_{L_2}$$

Note that induced norms are all 2-norms.



# Cauchy-Schwartz Inequality

## Theorem (Cauchy-Schwartz)

*Let  $S$  be any inner product space (doesn't need to be Hilbert) and*

*let  $\|\cdot\| = \langle \cdot, \cdot \rangle^{1/2}$*

*then  $\forall x, y \in S$*

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

*with equality iff  $y = \alpha x$  where  $\alpha \in \mathbb{F}$  is any scalar in the field  $\mathbb{F}$ .*

# Cauchy-Schwartz Inequality: Proof

The inequality clearly holds if either  $x = 0$  or  $y = 0$ . Therefore assume that  $x \neq 0$  and  $y \neq 0$ . Then

$$\begin{aligned}\|x - \alpha y\|^2 &= \langle x - \alpha y, x - \alpha y \rangle \\&= \langle x, x \rangle - \alpha \langle y, x \rangle - \langle x, \alpha y \rangle + \langle \alpha y, \alpha y \rangle \\&= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\overline{\langle x, \alpha y \rangle}} + \overline{\overline{\langle \alpha y, \alpha y \rangle}} \\&= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\langle \alpha y, x \rangle} + \alpha \overline{\langle \alpha y, y \rangle} \\&= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\alpha} \overline{\langle y, x \rangle} + \alpha \overline{\alpha} \overline{\langle y, y \rangle} \\&= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\alpha} \langle x, y \rangle + |\alpha|^2 \langle y, y \rangle \\&= \|x\|^2 - \alpha \langle y, x \rangle - \overline{\alpha} \langle x, y \rangle + |\alpha|^2 \|y\|^2\end{aligned}$$

# Cauchy-Schwartz Inequality: Proof

Recall the technique of completing the square:

$$\begin{aligned} ax^2 + bx + c &= a\left(x^2 + \frac{b}{a}x\right) + c \\ &= a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c. \end{aligned}$$

Complete the square in  $\alpha$ :

$$\begin{aligned} \|x - \alpha y\|^2 &= \|y\|^2 \left( \alpha \bar{\alpha} - \alpha \frac{\overline{\langle x, y \rangle}}{\|y\|^2} - \bar{\alpha} \frac{\langle x, y \rangle}{\|y\|^2} \right) + \|x\|^2 \\ &= \|y\|^2 \left( \alpha - \frac{\langle x, y \rangle}{\|y\|^2} \right) \left( \bar{\alpha} - \frac{\overline{\langle x, y \rangle}}{\|y\|^2} \right) - \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \|x\|^2 \end{aligned}$$

# Cauchy-Schwartz Inequality: Proof

Let  $\alpha^* = \frac{\langle x, y \rangle}{\|y\|^2}$  to get

$$0 \leq \|x - \alpha^* y\|^2 = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}$$

$$\Rightarrow |\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2$$

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

## Section 6

### Notions of Convergence

# Notions of Convergence

## Definition (Strong Convergence/ Convergence in norm)

$x_n$  converges strongly to  $x$ , i.e.  $x_n \xrightarrow{s} x$  iff

$$\|x_n - x\| \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

## Definition (Weak Convergence / Convergence in inner product)

$x_n$  converges weakly to  $x$ , i.e.  $x_n \xrightarrow{w} x$  iff

$$\langle x_n, y \rangle \rightarrow \langle x, y \rangle, \forall y \in S,$$

Note that this must hold for all  $y \in S$ , therefore Example 2.4.4 in the book is bogus!

# Notions of Convergence (cont.)

## Theorem (Strong vs. Weak Convergence)

Let  $(x_n)$  be a sequence in a normed space  $\mathbb{X}$ . Then

- A. Strong convergence  $\Rightarrow$  weak convergence with the same limit
- B. The converse of (A.) is not generally true
- C. If  $\dim \mathbb{X} < \infty$ , then weak convergence  $\Rightarrow$  strong convergence.

## Proof:

(A) By definition of strong convergence,

$$x_n \xrightarrow{s} x^* \Rightarrow \|x_n - x^*\| \rightarrow 0$$

so let  $y$  be any element in  $\mathbb{X}$  then

$$|\langle x_n, y \rangle - \langle x^*, y \rangle| = |\langle x_n - x^*, y \rangle| \leq \|x_n - x^*\| \|y\|$$

but the RHS  $\rightarrow 0$  which implies that the LHS  $\rightarrow 0$  which implies weak convergence.



## Proof:

(B) Before proving part (B) let's first understand what is wrong with Example 2.4.4 in the book.

$$x_n = (0, 0, 0, \dots, 1, 0, \dots)$$

$$y = (1, 1/2, 1/4, 1/8, \dots)$$

Then  $\langle x_n, y \rangle \rightarrow 0$  but this does not imply weak convergence since it must hold for all  $y \in \mathbb{X}$ .

## Proof:

To prove part (B) we need a counter example. Again let  $x_n = (0, 0, \dots, 0, 1, 0, \dots)$  and let  $\mathbb{X} = \ell_2$  i.e.

$$\begin{aligned} y \in \mathbb{X} &\Rightarrow \left( \sum_{i=1}^{\infty} |y_i|^2 \right)^{\frac{1}{2}} < \infty \\ &\Rightarrow y_i \rightarrow 0 \quad \text{as } i \rightarrow \infty \end{aligned}$$

so

$$\begin{aligned} \langle x_n, y \rangle &= y_n \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad \forall y \in \mathbb{X} \\ &\Rightarrow \{x_n\} \xrightarrow{w} 0 \end{aligned}$$

but there is no  $x^*$  such that  $\|x_n - x^*\| \rightarrow 0$ .

## Proof:

(C) Suppose that  $x_n \xrightarrow{w} x$  and  $\dim(\mathbb{X}) = k$  then

$$\forall y \in \mathbb{X} \quad \langle x_n, y \rangle \rightarrow \langle x, y \rangle.$$

Let  $\{e_1, \dots, e_k\}$  be an orthonormal basis for  $\mathbb{X}$ , i.e.  $\langle e_i, e_j \rangle = \delta_{ij}$ , then

$$\begin{aligned} x_n &= a_1^{(n)} e_1 + \dots + a_k^{(n)} e_k \\ x &= a_1 e_1 + \dots + a_k e_k. \end{aligned}$$

## Proof:

Then since  $\langle x_n, y \rangle \rightarrow \langle x, y \rangle \quad \forall y$ , let  $y = e_j$

$$\Rightarrow \langle a_1^{(n)} e_1 + \cdots + a_k^{(n)} e_k, e_j \rangle = a_j^{(n)}$$

and

$$\langle a_1 e_1 + \cdots + a_k e_k, e_j \rangle = a_j$$

so

$$\langle x_n, e_j \rangle \rightarrow \langle x, e_j \rangle \Rightarrow a_j^{(n)} \rightarrow a_j \quad \forall j = 1, \dots, k$$

Also,

$$\begin{aligned} \|x_n - x\| &= \left\| \sum_{j=1}^k a_j^{(n)} e_j - \sum_{j=1}^k a_j e_j \right\| = \left\| \sum_{j=1}^k (a_j^{(n)} - a_j) e_j \right\| \\ &\leq \sum_{j=1}^k |a_j^{(n)} - a_j| \|e_j\| \rightarrow 0 \end{aligned}$$

$\Rightarrow$  strong convergence

# Equivalence of Norms

## Theorem

*Let  $\dim(\mathbb{X}) = k$  and let  $\|\cdot\|$  and  $\|\cdot\|_0$  be two different norms on  $\mathbb{X}$  then  $\exists a, b$  such that*

$$a \|x\|_0 \leq \|x\| \leq b \|x\|_0$$

## Proof.

(in book page 96)



**Implication:** For convergence proofs, it doesn't matter which norm you use, therefore, use the one that simplifies the proof.

## Section 7

# Orthogonality

# Orthogonality

Let  $x, y \in \mathbb{X}$  where  $\mathbb{X}$  is an inner product space. Then the angle between  $x$  and  $y$  is

$$\theta = \cos^{-1} \left( \frac{\langle x, y \rangle}{\|x\| \|y\|} \right).$$

i.e.

$$\langle x, y \rangle = \|x\| \|y\| \cos \theta$$

# Orthogonality, cont.

## Definition (Colinear)

Two vectors  $x, y \in \mathbb{X}$  are said to be colinear if

$$\theta = 180 * n \quad n = 0, \pm 1, \pm 2, \dots$$

## Definition (Orthogonal)

Two vectors  $x, y \in \mathbb{X}$  are said to be orthogonal if

$$\theta = 90 * n \quad n = \pm 1, \pm 3, \pm 5, \dots$$

i.e.,  $\langle x, y \rangle = 0$ .

If  $\langle x, y \rangle = 0$  we write  $x \perp y$ .



## Orthogonality, cont.

### Example (Vectors in $L_2[0, 2\pi]$ )

The functions  $x = \sin(t)$  and  $y = \cos(t)$  are orthogonal since

$$\langle x, y \rangle = \int_0^{2\pi} \sin(t)\cos(t)dt = 0.$$

### Example (Vectors in $\ell$ )

The sequences

$$x = (1, 1, 1, 1, 0, 0, \dots)$$

$$y = (1, -1, 1, -1, 1, \dots)$$

are orthogonal since

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i = 0.$$

# Other useful inner products and norms: Weighting

## Definition (Positive Definite Matrix)

A matrix  $W : \mathbb{R}^k \rightarrow \mathbb{R}^k$  is positive definite (PD) if

$$\forall x \in \mathbb{R}^k \quad x^T W x > 0$$

- ▶  $W$  is positive semi-definite (PSD) if  $x^T W x \geq 0$
- ▶  $W$  is negative definite (ND) if  $x^T W x < 0 \quad \forall x \in \mathbb{R}^k$
- ▶  $W$  is negative semi-definite (NSD) if  $x^T W x \leq 0 \quad \forall x \in \mathbb{R}^k$
- ▶ Otherwise it is indefinite

## Examples of positive definiteness

- ▶  $W = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  is PD since

$$x^T W x = x_1^2 + x_2^2 > 0 \quad \forall x \in \mathbb{R}^2$$

- ▶  $W = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  is PSD since

$$x^T W x = x_1^2 = 0 \quad \forall x = \begin{pmatrix} 0 \\ \alpha \end{pmatrix} \neq 0$$

- ▶  $W = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$  is indefinite since

$$x^T W x = -x_1^2 + x_2^2$$

which can be positive or negative depending on  $x$ .

# Examples of Inner Products

## Weighted Inner Products and Norms

If  $W > 0$  then  $\langle x, y \rangle_W = x^H W y$  is a valid inner product which induces the weighted norm  $\|x\|_W = (x^H W x)^{\frac{1}{2}}$

We can define weighted inner products for functions:

$$\langle f, g \rangle_W = \int f(t)g(t)w(t)dt$$

where  $w(t) > 0$  except on a set of measure zero.

# Examples of Inner Products

## Definition (Expectation)

Expectation is a weighted inner product with weight  $f_{\mathbb{X}\mathbb{Y}}(x, y)$

$$\langle \mathbb{X}, \mathbb{Y} \rangle = \int xy f_{\mathbb{X}\mathbb{Y}}(x, y) dx dy = E[\mathbb{X}\mathbb{Y}]$$

if  $\mathbb{X}$  is a zero mean then

$$\langle x, x \rangle = \text{var}(x)$$

is the norm induced by  $E[\cdot]$

# Examples of Inner Products

- ▶ Let  $\mathbb{I}(m, n)$  be the set of grayscale images with  $m \times n$  pixels, each valued between  $[0, 255]$ .
- ▶ A valid inner on  $\mathbb{I}(m, n)$  is given by

$$\langle I, J \rangle = \sum_{u=1}^m \sum_{v=1}^n I(u, v)J(u, v), \quad \forall I, J \in \mathbb{I}(m, n).$$

# Orthogonal Subspaces

## Definition (Orthogonal Subspaces)

Let  $V, W$  be subspaces of  $S$ .  $V \perp W$  if

$$\forall v \in V \text{ and } \forall w \in W, \quad \langle v, w \rangle = 0$$

## Definition (Orthogonal Complement)

$V^\perp$ , called the orthogonal complement of  $V$ , is the set

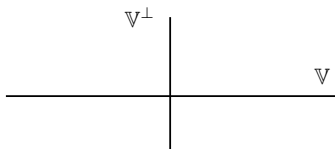
$$V^\perp = \{x \in S : \forall v \in V, \langle x, v \rangle = 0\}$$

# Orthogonal Subspaces, cont.

## Example

Let  $S = \mathbb{R}^2$  and  $V = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, \alpha \in \mathbb{R} \right\}$  then

$$V^\perp = \left\{ \begin{pmatrix} 0 \\ \alpha \end{pmatrix}, \alpha \in \mathbb{R} \right\}$$





# Orthogonal Subspaces, cont.

## Theorem

*Let  $V$  and  $W$  be subspaces of an inner product space  $S$  (not necessarily Hilbert). Then*

1.  $V^\perp$  is a closed subspace of  $S$
2.  $V \subset V^{\perp\perp}$  ( $V = V^{\perp\perp}$  if  $S$  is complete)
3. If  $V \subset W$  then  $W^\perp \subset V^\perp$
4.  $V^{\perp\perp\perp} = V^\perp$
5. If  $x \in V \cap V^\perp$  then  $x = 0$
6.  $\{0\}^\perp = S$  and  $S^\perp = \{0\}$

Prove one in class.

# Inner Sum and Direct Sum

## Definition (Inner Sum)

If  $V$  and  $W$  are linear subspaces then

$$V + W = \{x : x = v + w, v \in V \text{ and } w \in W\}$$

is the inner sum.

## Definition (Orthogonal Sum)

If  $V$  and  $W$  are orthogonal subspaces then the sum

$$V \oplus W = \{x : x = v + w, v \in V \text{ and } w \in W\}$$

is called the orthogonal sum.

## Definition (Disjoint Subspaces)

Two subspaces are said to be disjoint if

$$V \cap W = \{0\}$$

# Inner Sum and Direct Sum, cont.

## Lemma

*Let  $V + W$  be subspaces of  $S$  and let  $x \in V + W$  then the representation  $x = v + w$  is unique iff  $V + W$  are disjoint.*

## Proof.

( $\Leftarrow$ ) Assume  $V, W$  are disjoint but  $x = v + w$  is not unique i.e.  $x = v_1 + w_1 = v_2 + w_2$  where  $v_1 \neq v_2$  and  $w_1 \neq w_2$ . This implies that  $v_1 - v_2 = w_2 - w_1$  but  $v_1 - v_2 \in V$  and  $w_2 - w_1 \in W$  since  $V, W$  are subspaces. Since  $V \cap W = \{0\}$  we must have that  $v_1 - v_2 = w_2 - w_1 = 0$  or  $v_1 = v_2$  and  $w_1 = w_2$  which is a contradiction. □

# Inner Sum and Direct Sum, cont.

## Lemma

*If  $V$  and  $W$  are orthogonal subspaces then the representation of  $x \in V \oplus W$  is unique (i.e.  $x = v + w$ , where  $v \in V$  and  $w \in W$ ).*

## Example

Let  $S = \mathbb{R}^2$ , let  $V = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\}$ , let

$W = \left\{ \begin{pmatrix} 0 \\ \alpha \end{pmatrix} : \alpha \in \mathbb{R} \right\}$  Then

$$\begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

is a unique decomposition.

# Difference between a Hamel basis and a Complete basis.

## Definition

An orthonormal set of basis vectors  $T = \{p_1, p_2, \dots\}$  is said to be a complete basis for a Hilbert space  $S$  if every  $x \in S$  can be represented as

$$x = \sum_{j=1}^{\infty} c_j p_j$$

Examples of complete bases: Fourier functions:  $e^{j\omega t}$

Legendre & Chebyshev polynomials

Difference: A Hamel basis  $\Rightarrow$  every  $x$  can be represented by a finite representation. A complete basis allows functions through a limiting process.

## Section 8

# Linear Operators

# Operators and Transformations

## Definition (Linear Operator)

Let  $\mathcal{L} : \mathbb{X} \rightarrow \mathbb{Y}$  be an operator from  $\mathbb{X}$  to  $\mathbb{Y}$ .  $\mathcal{L}$  is a linear operator if

1.  $\mathcal{L}[\alpha x] = \alpha \mathcal{L}[x] \quad \forall x \in \mathbb{X} \quad \forall \alpha \in \mathbb{F}$
2.  $\mathcal{L}[x_1 + x_2] = \mathcal{L}[x_1] + \mathcal{L}[x_2], \quad \forall x_1, x_2 \in \mathbb{X}$

# Examples of Linear Operators

## Example (Matrices)

Operators from  $\mathbb{C}^n$  to  $\mathbb{C}^m$  are  $m \times n$  matrices.

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay$$

$A$  is a linear operator.

## Example (Differential Equations with no input)

The differential equation  $\dot{x} = Ax$ ;  $x(0) = x_0$  defines a linear operator from  $\mathbb{R}^n$  to  $L_2[0, T]$

$$y(t) = \mathcal{L}[x_0] \text{ where } \mathcal{L}[x_0] = e^{At}x_0$$

$\mathcal{L}$  is linear since

$$e^{At}(\alpha x_{01} + \beta x_{01}) = \alpha e^{At}x_{01} + \beta e^{At}x_{02}$$



# Examples of Linear Operators

## Example (Convolution)

Convolution is a linear operator from  $L_\infty$  to  $L_\infty$  if  $h(t) \in L_1[-\infty, \infty]$ , i.e.

$$y(t) = \mathcal{L}[x(t)] = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau$$

(Recall: for a system to be BIBO stable required that  $\int_{-\infty}^{\infty} |h(\tau)|d\tau < \infty$  i.e.  $h(t) \in L_1[-\infty, \infty]$ )

# Examples of Linear Operators

## Example (Fourier Transform)

(E4) The Fourier transform defines a linear operator from  $L_2[-\infty, \infty]$  to  $L_2[-\infty, \infty]$ .

$$X(j\omega) = \mathcal{L}[x(t)] \triangleq \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt$$

There are many examples of linear operators!

# Range and Null Space of an Operator

## Definition (Range Space)

Let  $\mathcal{L} : \mathbb{X} \rightarrow \mathbb{Y}$  be a linear operator. The range space (or image) of  $\mathcal{L}$  is

$$\mathcal{R}(\mathcal{L}) = \{y \in \mathbb{Y} : y = \mathcal{L}[x] \text{ and } x \in \mathbb{X}\} \subseteq \mathbb{Y}$$

## Definition (Null Space)

The Null space or kernel of  $\mathcal{L}$  is

$$\mathcal{N}(\mathcal{L}) = \{x \in \mathbb{X} : \mathcal{L}[x] = 0\} \subseteq \mathbb{X}$$

## Example of Range and Null Space

- ▶ Consider the matrix  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  which defines a linear operator from  $\mathbb{R}^3$  to  $\mathbb{R}^2$ .
- ▶ Note that  $y = Ax = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$ .
- ▶ Therefore, the range space is

$$\mathcal{R}(A) = \left\{ \begin{pmatrix} \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\} \subset \mathbb{R}^2.$$

- ▶ Similarly, the null space is

$$\mathcal{N}(A) = \left\{ \begin{pmatrix} 0 \\ \alpha \\ \beta \end{pmatrix} : \alpha, \beta \in \mathbb{R} \right\} \subset \mathbb{R}^3.$$

## Section 9

# Projections

# Projections

- ▶ Suppose that  $V$  and  $W$  are disjoint subspaces of  $S$  such that  $V + W = S$ , i.e.

$$x \in S \Rightarrow x = v + w$$

where  $v \in V$  and  $w \in W$  is a unique decomposition.

- ▶ Define the linear operator  $P : S \rightarrow V \subset S$  as

$$Px = P(v + w) = v$$

- ▶ Note that  $P(Px) = Pv = v$

# Projections, cont.

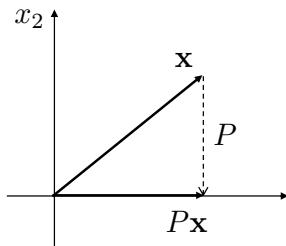
## Definition (Projection Operator)

Let  $P : S \rightarrow S$  such that  $P^2 = P$ , then  $P$  is called a projection operator or idempotent.

## Example

Let  $P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ , then  $P \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$

i.e.  $P$  projects elements of  $P$  onto the  $x_1$  axis:



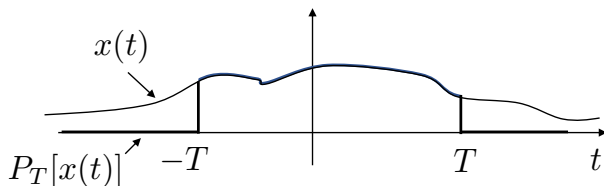
## Projections, cont.

### Example

Truncation: let

$$(P_T x)(t) = \begin{cases} x(t), & -T \leq t \leq T \\ 0, & \text{otherwise} \end{cases}$$

Then  $P_T$  projects  $x(t)$  onto its truncated function:





## Projections, cont.

### Theorem (Moon 2.7)

*Let  $P : S \rightarrow S$  be a projection operator, then*

$$S = R(P) + N(P)$$

**Proof.**

Homework problem.



## Projections, cont.

### Theorem

*If  $P : S \rightarrow S$  is a projection operator then so is  $(I - P) : S \rightarrow S$*

Proof.

$$\begin{aligned}(I - P)^2 &= (I - P)(I - P) = \\ &= I - P - P - P^2 \\ &= I - P - P + P \\ &= I - P\end{aligned}$$



## Projections, cont.

- Note that if  $P : S \rightarrow V$  and  $I - P : S \rightarrow W$  then  $V$  and  $W$  are disjoint and  $S = V + W$  since

$$x = \underbrace{Px}_{\in V} + \underbrace{(I - P)x}_{\in W}.$$

- $V$  and  $W$  are disjoint. If not, then  $\exists x_0 (\neq 0) \in S$  such that

$$\begin{aligned} Px_0 &= (I - P)x_0 = x_0 - Px_0 \\ 2Px_0 &= x_0 \\ \Rightarrow Px_0 &= \frac{1}{2}x_0 \\ \text{and } P^2x_0 &= \frac{1}{4}x_0 = \frac{1}{2}x_0 \Leftrightarrow x_0 = 0 \end{aligned}$$

# Projections, cont.

## Definition (Orthogonal Projection)

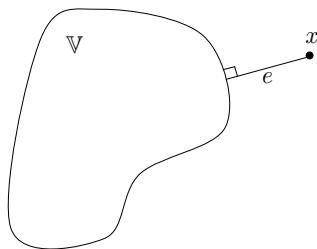
If  $V$  and  $W$  are orthogonal then  $P$  is an orthogonal projection

## Theorem

*$P$  is an orthogonal projection iff  $R(P) \perp N(P)$*

# Applications to Engineering

Given a point  $x \in S$ , suppose that we want to approximate  $x$  by a point in  $\mathbb{V} \subset S$  (assuming  $x \notin \mathbb{V}$ ) then we want to find the point in  $\mathbb{V}$  that is closest to  $x$ .



This is given by the orthogonal projection of  $x$  onto  $\mathbb{V}$ . i.e.

$$\langle e, v \rangle = 0 \quad \forall v \in \mathbb{V}$$

# Applications to Engineering

Let  $\mathbf{n}$  be a unit vector in  $\mathbb{R}^3$  (i.e.,  $\|\mathbf{n}\| = 1$ ), then

$$\Pi_{\mathbf{n}}^{\perp} \triangleq \mathbf{n}\mathbf{n}^{\top}$$

is a projection operation. Geometrically  $\Pi_{\mathbf{n}}^{\perp} \mathbf{x} = \mathbf{n}\mathbf{n}^{\top} \mathbf{x}$  find the projection of  $\mathbf{x}$  along the unit vector  $\mathbf{n}$

Also

$$\Pi_{\mathbf{n}} = I - \mathbf{n}\mathbf{n}^{\top}$$

is a projection operator. Geometrically,  $\Pi_{\mathbf{n}} \mathbf{x}$  projections  $\mathbf{x}$  onto the 2D space that is orthogonal to  $\mathbf{n}$ .

# The Projection Theorem

## Theorem

*Let  $\mathbb{S}$  be a Hilbert space and let  $\mathbb{V}$  be a closed subspace of  $\mathbb{S}$ . For any  $x \in \mathbb{S}$  there exists a unique  $v_0 \in \mathbb{V}$  closest to  $x$ ; i.e.*

$$\|x - v_0\| \leq \|x - v\| \quad \forall v \in \mathbb{V}.$$

*Furthermore  $v_0$  minimizes  $\|x - v_0\|$  iff  $x - v_0$  is orthogonal to  $\mathbb{V}$ .*

# The Projection Theorem, proof

## Step 1. Show that $v_0$ exists.

Assume  $x \notin \mathbb{V}$  and let  $\delta = \inf_{v \in \mathbb{V}} \|x - v\|$ . We need to show that in fact  $\exists v_0 \in \mathbb{V}$  such that  $\|x - v_0\| = \delta$ .

Let  $\{v_i\}$  be a sequence in  $\mathbb{V}$  such that  $\|x - v_i\| \rightarrow \delta$  and show that  $\{v_i\}$  is Cauchy.

Need parallelogram law:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Consider

$$\begin{aligned} \|(v_j - x) + (x - v_i)\|^2 + \|(v_j - x) - (x - v_i)\|^2 \\ = 2\|v_j - x\|^2 + 2\|x - v_i\|^2 \end{aligned}$$

$$\Rightarrow \|v_j - v_i\|^2 = 2\|v_j - x\|^2 + 2\|x - v_i\|^2 - 4\left\|\frac{(v_j + v_i)}{2} - x\right\|^2$$



# The Projection Theorem, proof

$$v_i, v_j \in \mathbb{V} \Rightarrow \frac{v_j + v_i}{2} \in \mathbb{V} \Rightarrow \left\| \frac{(v_j - v_i)}{2} - x \right\|^2 \geq \delta^2$$

$$\Rightarrow \|v_j - v_i\|^2 \leq 2 \|v_j - x\|^2 + 2 \|v_i - x\|^2 - 4\delta^2$$

But  $\|v_j - x\| \rightarrow \delta$

$$\Rightarrow \|v_j - v_i\| \rightarrow 0$$

and is therefore Cauchy.

Since  $\mathbb{V}$  is a Hilbert space

$$v_i \rightarrow v_0 \in \mathbb{V}.$$

Note that this proof is not constructive, i.e. it doesn't tell you how to construct the sequence  $\{v_i\}$ .

# The Projection Theorem, proof

**Step 2. Show that**  $v_0 = \arg \min_{v \in \mathbb{V}} \|x - v\| \Rightarrow x - v_0 \perp \mathbb{V}$ .

Proof by contradiction. Suppose that  $x - v_0$  is not perpendicular to  $\mathbb{V}$ . Then there exists a  $v \in \mathbb{V}$  such that

$$\langle x - v_0, v \rangle = \delta \neq 0$$

and w.l.o.g. (why?) let  $\|v\| = 1$

Let  $z = v_0 + \delta v \in \mathbb{V}$  then

$$\begin{aligned}\|x - z\|^2 &= \|x - v_0 - \delta v\|^2 = \|x - v_0\|^2 - 2\operatorname{Re} \langle x - v_0, \delta v \rangle + \|\delta v\|^2 \\ &= \|x - v_0\|^2 - 2\delta^2 + \delta^2 < \|x - v_0\|^2\end{aligned}$$

which is a contradiction since  $v_0$  is the minimizer.

# The Projection Theorem, proof

**Step 3.** Suppose  $(x - v_0) \perp \mathbb{V}$  then  $\forall v \in \mathbb{V}$  such that  $v \neq v_0$

$$\begin{aligned}\|x - v\|^2 &= \|x - v_0 + v_0 - v\|^2 \\&= \|x - v_0\|^2 + 2\operatorname{Re} \langle x - v_0, v_0 - v \rangle + \|v_0 - v\|^2 \\&= \|x - v_0\|^2 + \|v_0 - v\|^2 \\&> \|x - v_0\|^2\end{aligned}$$

**Step 4. Uniqueness** Same as proof on page 25 of notes.

# Closed Subspace

## Theorem (Moon Theorem 2.10)

*Let  $\mathbb{V}$  be a closed subspace of a Hilbert space  $\mathbb{S}$ , then*

$$\mathbb{S} = \mathbb{V} \oplus \mathbb{V}^\perp$$

$$\mathbb{V} = \mathbb{V}^{\perp\perp}$$

**Proof.**

In book.



## Section 10

### Gram Schmidt Orthogonalization

# Application: Gram Schmidt Orthogonalization

Given a set  $T = \{p_1, \dots, p_n\}$

Find a set  $T' = \{q_1, \dots, q_{n'}\}$   $n' \leq n$  such that

$$\text{span}(T') = \text{span}(T) \text{ and } \langle q_i, q_j \rangle = \delta_{ij}$$

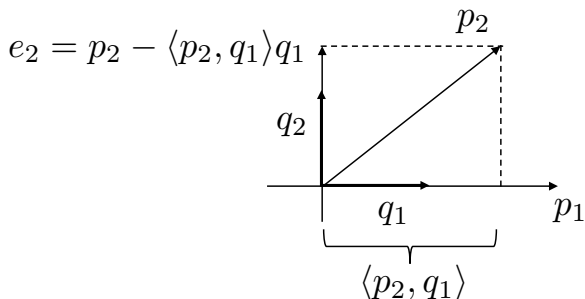
**Step 1.** Normalize the First Vector

$$q_1 = \frac{p_1}{\|p_1\|} \quad (\text{i.e. } \langle q_1, q_1 \rangle = 1)$$

## Application: Gram Schmidt Orthogonalization, cont

Step 2. Let  $e_2$  be  $p_2$  minus the projection of  $p_2$  on  $q_1$  i.e.

$$e_2 = p_2 - \langle p_2, q_1 \rangle q_1$$



Then normalize  $e_2$ :

$$q_2 = \frac{e_2}{\|e_2\|}$$

# Application: Gram Schmidt Orthogonalization, cont

Step 3. Let  $e_k$  be  $p_k$  minus the projection of  $p_k$  on  $q_1, \dots, q_{k-1}$ :

$$e_k = p_k - \sum_{j=1}^{k-1} \langle p_k, q_j \rangle q_j \Rightarrow q_k = \frac{e_k}{\|e_k\|}$$



## Example: Gram Schmidt Orthogonalization

Given the set

$$T = \{p_1, p_2, p_3\} \triangleq \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$$

find a set  $T' = \{q_1, q_2, q_3\}$  where the vectors in  $T'$  are orthonormal and  $\text{span}(T) = \text{span}(T')$ .

$$q_1 = \frac{p_1}{\|p_1\|} = \frac{\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}}{2} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

## Example: Gram Schmidt Orthogonalization, cont.

$$\begin{aligned}e_2 &= p_2 - \langle p_2, q_1 \rangle q_1 \\&= \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}^\top \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\&= \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\&= \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}\end{aligned}$$

$$\text{Therefore } q_2 = \frac{e_2}{\|e_2\|} = \frac{\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}}{2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

## Example: Gram Schmidt Orthogonalization, cont.

$$\begin{aligned}e_3 &= p_3 - \langle p_3, q_1 \rangle q_1 - \langle p_3, q_2 \rangle q_2 \\&= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}^\top \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}^\top \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\&= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\&= \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}\end{aligned}$$

$$\text{Therefore } q_3 = \frac{e_3}{\|e_3\|} = \frac{\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}}{3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

## Example: Gram Schmidt Orthogonalization, cont.

Therefore, the Gram Schmidt orthonormalization of

$$T = \{p_1, p_2, p_3\} \triangleq \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}$$

is

$$T' = \{q_1, q_2, q_3\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Note that  $\text{span}(T) = \text{span}(T')$ .