

27 Sept 2018

Review •  $(X, \mathcal{F}, \|\cdot\|)$  a normed space if  $(X, \mathcal{F})$

a vector space,  $\mathcal{F} = \mathbb{R}$  or  $\mathbb{C}$ , and  $\|\cdot\|: X \rightarrow \mathbb{R}$

satisfies: (1)  $\forall x \in X, \|x\| \geq 0$  and  $\|x\| = 0 \Leftrightarrow x = 0$

(2)  $\forall x, y \in X, \|x+y\| \leq \|x\| + \|y\|$

(3)  $\forall \alpha \in \mathcal{F}, \forall x \in X, \|\alpha x\| = |\alpha| \cdot \|x\|$

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

$$\bullet d(x, S) := \inf_{y \in S} d(x, y) = \inf_{y \in S} \|x-y\|$$

$$\bullet x^* = \arg \min_{y \in S} d(x, y) \Leftrightarrow x^* \in S, \|x-x^*\| = d(x, S)$$

and  $x^*$  is the unique element  
 $\in S$  satisfying  $\|x-S\| = d(x, S)$

$$\bullet (\mathcal{F}^n, \mathcal{F}), \quad \|x\|_p := \left( \sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \quad 1 \leq p < \infty$$

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

$$\bullet a, b \in \mathbb{R}, a < b, D = [a, b], \mathcal{F} = \mathbb{R}, X = \{f: D \rightarrow \mathbb{R} \text{ cont.}\}$$

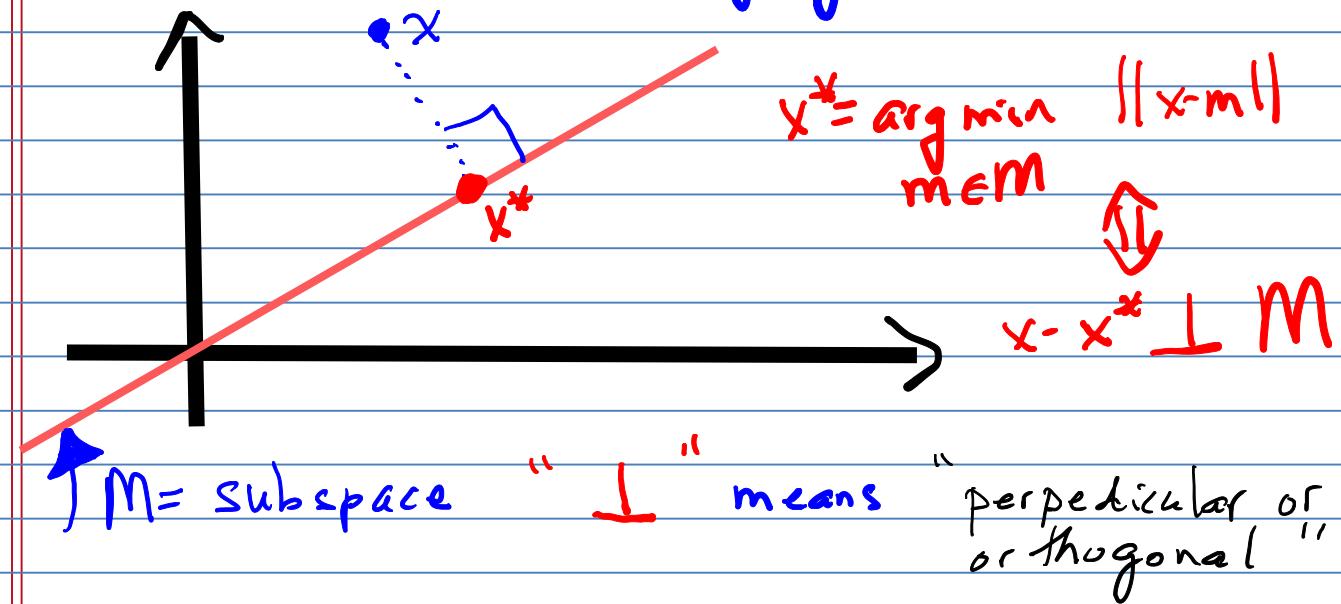
$$\|f\|_p := \left( \int_a^b |f(t)|^p dt \right)^{\frac{1}{p}} \quad 1 \leq p < \infty$$

$$\|f\|_\infty := \sup_{a \leq t \leq b} |f(t)| = \max_{a \leq t \leq b} |f(t)|$$

ROB 501:  $x^* = \arg \min_{y \in S} d(x, y)$  when

$S = \text{subspace}$  and  $\|\cdot\| = \text{"sum of squares"}$

Goal today: tools to make sense of  
this picture for very general vector spaces:



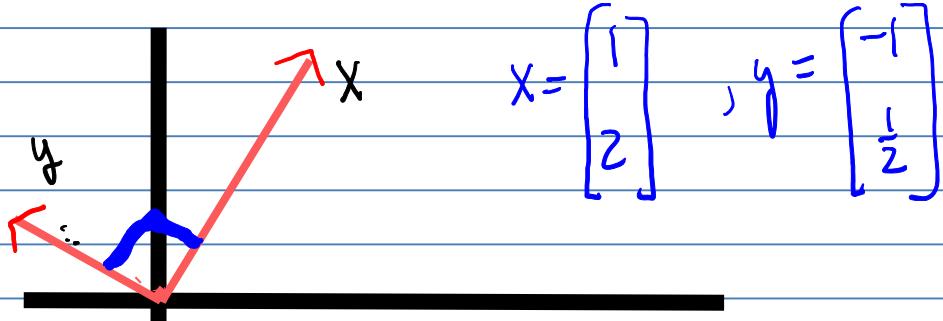
**Please read.** Review Complex Numbers: Let  $z = z_R + j z_I \in \mathbb{C}$ , where  $z_R, z_I \in \mathbb{R}$ . We note that:

- $\bar{z} := z_R - j z_I$  is the complex conjugate of  $z$
- $z \in \mathbb{R} \Leftrightarrow z = \bar{z}$
- $z \cdot \bar{z} = |z|^2$ , and thus,  $|z| = \sqrt{z \cdot \bar{z}}$

$$\begin{cases} z_R = \operatorname{Re}\{z\} \\ z_I = \operatorname{Im}\{z\} \end{cases} \quad \begin{matrix} \nearrow & \searrow \\ c \in \mathbb{R} \end{matrix}$$

$$z = \operatorname{Re}\{z\} + j \operatorname{Im}\{z\}$$

 "Recall":  $x, y \in \mathbb{R}^n$ ,  $x^T y = [x_1 \dots x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^n x_i y_i$



$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, y = \begin{bmatrix} -1 \\ \frac{1}{2} \end{bmatrix}$$

$$x^T y = (1)(-1) + (2)\left(\frac{1}{2}\right) = 0$$

Def. Let  $(X, \mathbb{R})$  be a vector space. A function  $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$

is an inner product if

a)  $\forall x, y \in X, \quad \langle x, y \rangle = \langle y, x \rangle$

b)  $\forall \alpha_1, \alpha_2 \in \mathbb{R}, \quad \forall x_1, x_2 \in X, \quad \forall y \in X$

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle$$

c)  $\forall x \in X, \quad \langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ .

Remark: (a) + (b)  $\Rightarrow$

$$\langle x, \beta_1 y_1 + \beta_2 y_2 \rangle = \beta_1 \langle x, y_1 \rangle + \beta_2 \langle x, y_2 \rangle$$

## Examples

a)  $(\mathbb{R}^n, \mathbb{R})$ ,  $\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i$

b)  $(X, \mathbb{R})$ ,  $X = \mathbb{R}^{n \times m} = \{\text{matrices } n \times m, \text{ coeff. in } \mathbb{R}\}$

$$\langle A, B \rangle := \text{tr}(A^T B)$$

c)  $(X, \mathbb{R})$ ,  $X = \{f: D \rightarrow \mathbb{R}, f \text{ cont.}\}$

$$\langle f, g \rangle := \int_a^b f(t) g(t) dt$$

$D = [a, b]$ ,  $a, b \in \mathbb{R}$ ,  $a < b$ .

Case of  $(X, \mathbb{C}, \langle \cdot, \cdot \rangle)$  is given on  
the next page. We will only need  
it in ROB 501 on rate occasions.

**Definition:** Let  $(X, \mathbb{C})$  be a vector space. A function

$$\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{C}$$

is an **inner product** if

(a)  $\forall x, y \in X, \langle x, y \rangle = \overline{\langle y, x \rangle}$

(b)  $\forall x_1, x_2, y \in X$  and  $\forall \alpha_1, \alpha_2 \in \mathbb{C}$ , (i.e., linear in the left argument)

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle$$

(c)  $\forall x \in X, \langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ .

**Remarks:**

- In the case of a real vector space  $(X, \mathbb{R})$ , replace (a) with

(a'):  $\langle x, y \rangle = \langle y, x \rangle$ . It is easy to show that we then have linearity in both the left and right sides.

- Going back to the complex case,  $(X, \mathbb{C})$ , (a) and (b) together imply that  $\forall x, y_1, y_2 \in X$  and  $\forall \alpha_1, \alpha_2 \in \mathbb{C}$ ,

$$\begin{aligned}\langle x, \alpha_1 y_1 + \alpha_2 y_2 \rangle &= \overline{\langle \alpha_1 y_1 + \alpha_2 y_2, x \rangle} \\ &= \overline{\alpha_1} \overline{\langle y_1, x \rangle} + \overline{\alpha_2} \overline{\langle y_2, x \rangle} \\ &= \overline{\alpha_1} \langle y_1, x \rangle + \overline{\alpha_2} \langle y_2, x \rangle \\ &= \overline{\alpha_1} \langle x, y_1 \rangle + \overline{\alpha_2} \langle x, y_2 \rangle\end{aligned}$$

## Cauchy-Schwarz Inequality

Let  $(X, \mathbb{C}, \langle \cdot, \cdot \rangle)$  be an inner product

space. Then,  $\forall x, y \in X$ ,

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}$$

To show:  $\|x\| := \langle x, x \rangle^{1/2}$  is in

fact a norm on  $(X, \mathcal{F})$ .

K  
N  
O  
W

**Theorem:** [Cauchy-Schwarz Inequality] Suppose that  $\mathcal{F} = \mathbb{R}$  or  $\mathbb{C}$ . Let  $(X, \mathcal{F}, \langle \cdot, \cdot \rangle)$  be an **inner product space** (i.e.  $(X, \mathcal{F})$  is a vector space and  $\langle \cdot, \cdot \rangle$  is an inner product on  $X$ ). Then, for all  $x, y \in X$ ,

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}.$$

skip

**Proof.** If  $y = 0$ , the result is obviously true. Hence, assume  $y \neq 0$ . For all scalars  $\lambda$  we have that

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \lambda \langle x, y \rangle - \bar{\lambda} \langle y, x \rangle + |\lambda|^2 \langle y, y \rangle,$$

because the inner product of a vector with itself is a non-negative real number.

For the particular choice  $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$ , direct calculation shows

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

which gives

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle} = \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}.$$

Triangle Inequality  $\mathcal{F}=\mathbb{R}$

$$\|x+y\| \leq \|x\| + \|y\| \quad \text{True?}$$

$$\|x+y\|^2 \leq (\|x\| + \|y\|)^2$$

$$\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2$$

Let's now expand the left hand side!

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

Because of C.S.

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

Missing Step

$$2 \langle x, y \rangle \leq 2 |\langle x, y \rangle| \leq 2 \|x\| \cdot \|y\|$$

$$\therefore \|x+y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2$$

$$\|x+y\| \stackrel{\uparrow}{\leq} \|x\| + \|y\|$$

$\Delta$ -Inequality.

Not responsible for proof when  $\mathcal{F} = \mathbb{C}$ .

**Corollary:** Let  $(X, \mathcal{F}, \langle \cdot, \cdot \rangle)$  be an inner product space. Then

$\mathcal{F} = \mathbb{R}$  or  $\mathbb{C}$

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

is a norm on  $X$ .

**Proof:** The main thing to establish is the triangle inequality:

$$\|x + y\| \leq \|x\| + \|y\|.$$

This is equivalent to showing:

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2.$$

Brute force computation:

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

where  $\operatorname{Re}\{\langle x, y \rangle\}$  denotes the real part of the complex number  $\langle x, y \rangle$ .

However, for any complex number  $\alpha$ ,  $\operatorname{Re}\{\alpha\} \leq |\alpha|$ , and thus we have

$$\begin{aligned}\|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}\{\langle x, y \rangle\} \\&\leq \|x\|^2 + \|y\|^2 + 2|\langle x, y \rangle| \\&\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|,\end{aligned}$$

where the last inequality is from the Cauchy-Schwarz Inequality. ■

Def. (a) Two vectors  $x, y \in (\mathbb{X}, \mathcal{J}, \langle \cdot, \cdot \rangle)$  are orthogonal if  $\langle x, y \rangle = 0$ .

Notation  $x \perp y \Leftrightarrow \langle x, y \rangle = 0$

b) A set of vectors  $S$  is orthogonal if  $(\forall x, y \in S, x \neq y) \Rightarrow x \perp y$

c) If in addition,  $\forall x \in S \Rightarrow \|x\| = 1$ , then  $S$  is orthonormal.

(orthonormal  $\Rightarrow$  [orthogonal and norm one])

Pythagorean Thm If  $x \perp y$ ,  
then  $\|x+y\|^2 = \|x\|^2 + \|y\|^2$

(Pf for  $\mathbb{X}=\mathbb{R}$ )

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle \quad \text{o because } x \perp y \\ &= \|x\|^2 + \|y\|^2\end{aligned}$$

□

## Pre-Projection Theorem Let

$(X, \mathbb{R}, \langle \cdot, \cdot \rangle)$  be an inner product space. Let  $M \subset X$  be a subspace and  $x \in X$ . Then

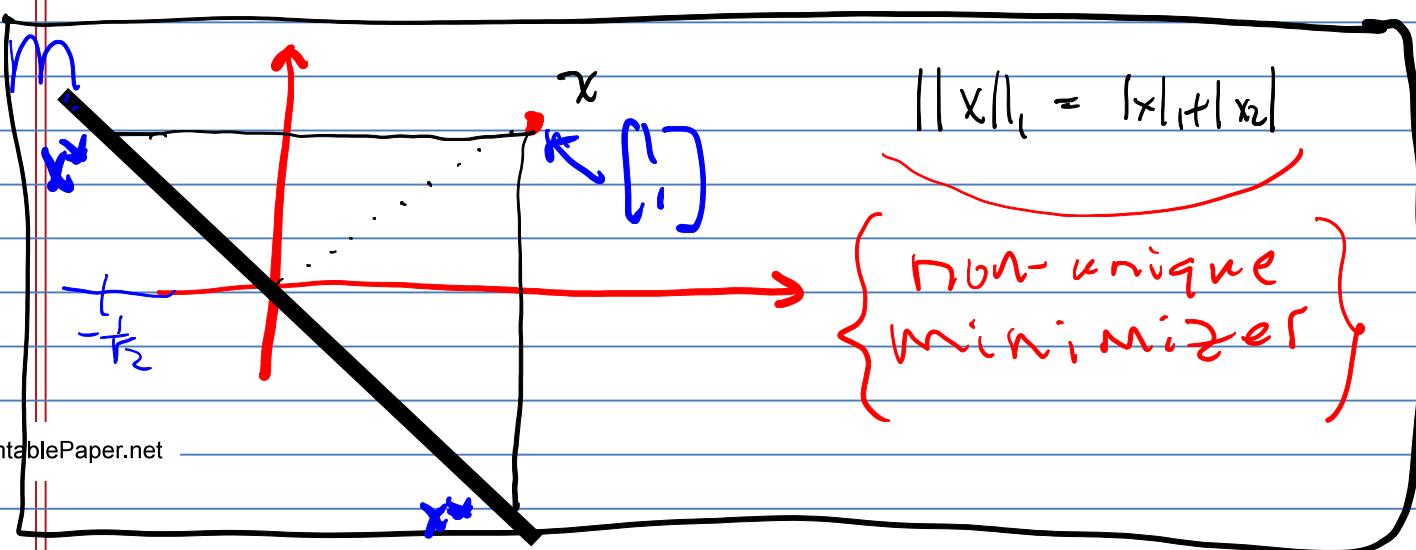
a) If  $\exists m_0 \in M$  such that

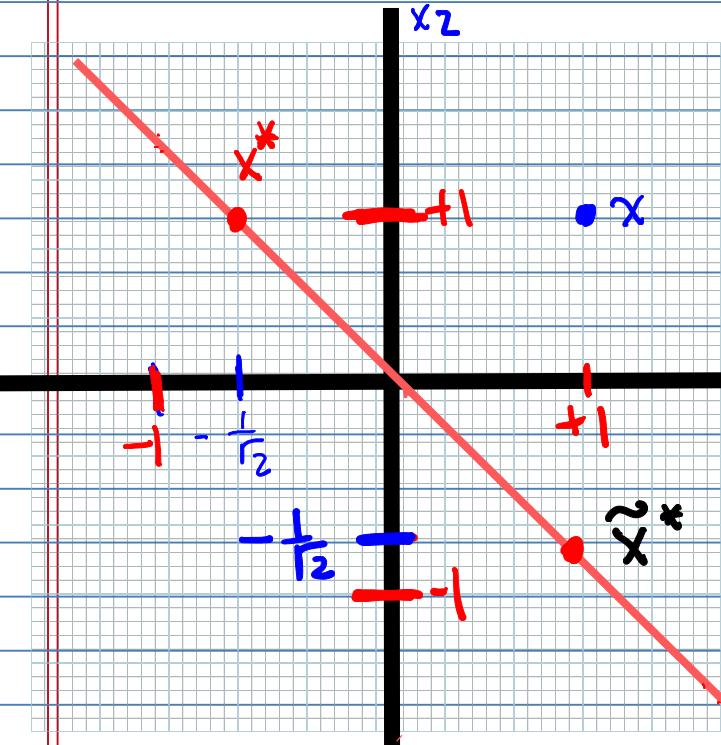
$$\|x - m_0\| \leq \|x - m\| \quad \forall m \in M,$$

then  $m_0$  is unique.

b) A necessary and sufficient condition that  $m_0 \in M$  be a minimizing vector in  $M$  is that the error vector

is orthogonal to  $M$ ; that is  $x - m_0 \perp M$ .





$$\|x\|_1 = |x_1| + |x_2|$$

$$x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Let  $e = x - x^*$ , then  $e = \begin{bmatrix} 1 + \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}$ ,  $\|e\|_1 = 1 + \frac{\sqrt{2}}{2}$

$$\text{Consider } \|x\|_1 = \|x - 0\|_1 =$$

$$\rightarrow = \left\| \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\|_1 = 2 > 1 + \frac{\sqrt{2}}{2}$$

$\therefore 0$  is not a minimizing vector.

Consider  $\tilde{x}^* = \begin{bmatrix} 1 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ , then

$$\|x - \tilde{x}^*\|_1 = \left\| \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\|_1 = \left\| \begin{bmatrix} 0 \\ 1 + \frac{1}{\sqrt{2}} \end{bmatrix} \right\|_1 = 1 + \frac{1}{\sqrt{2}}$$

∴ There are two values that achieve the minimum,