27 Sept 2018

Review • $(\mathcal{X}, \mathcal{F}, \|\cdot\|)$ a normed space if $(\mathcal{X}, \mathcal{F})$ a vector space, $\mathcal{F} = \mathbb{R}$ or $\mathbb{C}$, and $\|\cdot\| : \mathcal{X} \to \mathbb{R}$ satisfies:

(1) $\forall\, x \in \mathcal{X}$, $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$

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

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

- $d(x,y) := \|x-y\|$
- $d(x,S) := \inf_{y \in S} d(x,y) = \inf_{y \in S} \|x-y\|$
- $x^* = \arg\min_{y \in S} d(x,y) \iff x^* \in S$, $\|x - x^*\| = d(x,S)$ and x^* is the unique element $s \in S$ satisfying $\|x-s\| = d(x,S)$.

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

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

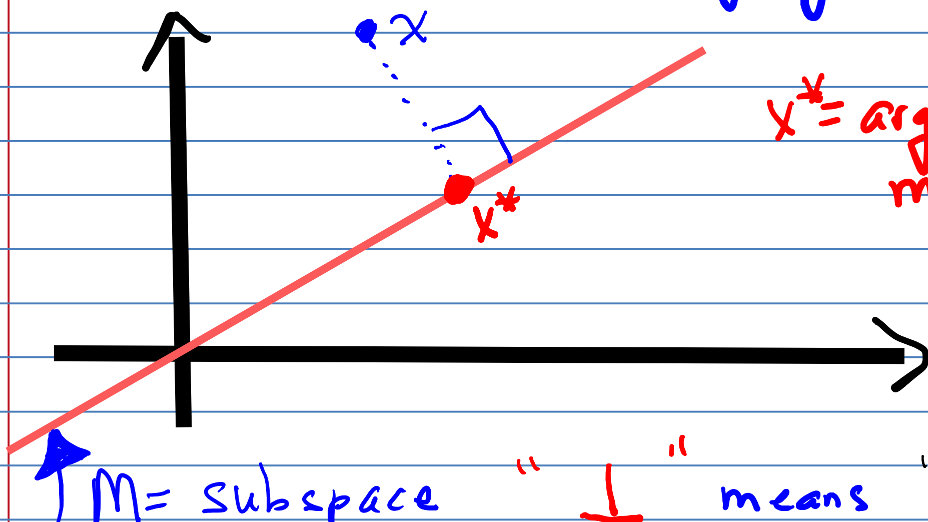
- $a, b \in \mathbb{R}$, $a < b$, $D = [a,b]$, $\mathcal{F} = \mathbb{R}$, $\mathcal{X} = \{f : D \to \mathbb{R}\ \text{cont.}\}$

$$\|f\|_p := \left(\int_a^b |f(\tau)|^p d\tau\right)^{1/p} \quad 1 \leq p < \infty$$

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

ROB 501: $x^* = \arg\min_{y \in S} d(x,S)$ when $S =$ subspace and $\|\cdot\|^2 =$ "sum of squares"

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



$x^* = \underset{m \in M}{\arg\min} \|x - m\|$

$\Updownarrow$

$x - x^* \perp M$

M = subspace

"$\perp$" means "perpedicular or orthogonal"

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

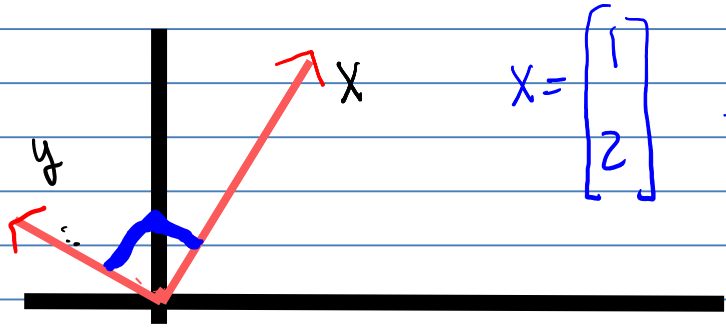
Please read.

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

$z_R = Re\{z\}$, $z_I = Im\{z\}$ $\in \mathbb{R}$

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

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



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

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

Def. Let $(\mathcal{X}, \mathbb{R})$ be a vector space. A function $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is an inner product if

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

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

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

c) $\forall x \in \mathcal{X}$, $\langle x, x \rangle \geq 0$ 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) $(\mathbb{R}^n, \mathbb{R})$, $\quad \langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i$

b) $(\mathcal{X}, \mathbb{R})$, $\quad \mathcal{X} = \mathbb{R}^{n \times m} = \{$Matrices $n \times m$, coeff in $\mathbb{R}\}$

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

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

$$\langle f, g \rangle := \int_a^b f(\tau)\, g(\tau)\, d\tau$$

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

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

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

$$< \cdot, \cdot >: X \times X \to \mathbb{C}$$

is an **inner product** if

(a) $\forall x, y \in X$, $< x, y > = \overline{< y, x >}$

(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)

$$< \alpha_1 x_1 + \alpha_2 x_2, y > = \alpha_1 < x_1, y > + \alpha_2 < x_2, y >$$

(c) $\forall x \in X$, $< x, x > \geqslant 0$ and $< x, x > = 0 \Leftrightarrow x = 0$.

Remarks:

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

(a′): $< x, y > = < y, x >$. 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 $\forall \alpha_1, \alpha_2 \in \mathbb{C}$,

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

Cauchy-Schwarz Inequality

Let $(X, \mathbb{R}, \langle \cdot, \cdot \rangle)$ 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 "n fact a norm on $(\mathcal{X}, \mathbb{R})$.

KNOW

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

$$|< x, y >| \leqslant < x, x >^{1/2} \cdot < y, y >^{1/2} .$$

skip

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

$$0 \leq \; < x - \lambda y, x - \lambda y > = < x, x > - \lambda < y, x > - \bar{\lambda} < x, y > + |\lambda|^2 < y, y >,$$

because the inner product of a vector with itself is a non-negative real number. For the particular choice $\lambda = \frac{<x,y>}{<y,y>}$, direct calculation shows

$$0 \leq \; < x, x > - \frac{| < x, y > |^2}{< y, y >},$$

which gives

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

$\blacksquare$

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

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

$$\Updownarrow$$

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

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

Let's now expand the left hand side!

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

$$= \langle x, x+y\rangle + \langle y, x+y\rangle$$

$$= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$$

$$= \|x\|^2 + \underline{2\langle x, y\rangle} + \|y\|^2$$

Because of C.S.

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

Missing step

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

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

$$\Updownarrow$$

$$\|x+y\| \leq \|x\| + \|y\|$$ Δ-Inequality!!

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

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

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

$$\|x\| := <x,x>^{1/2}$$

is a norm on X.

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

$$\|x+y\| \leqslant \|x\| + \|y\|.$$

This is equivalent to showing:

$$\|x+y\|^2 \leqslant \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2.$$

Brute force computation:

$$\begin{aligned}
\|x+y\|^2 &= <x+y, x+y> \\
&= <x, x+y> + <y, x+y> \\
&= \overline{<x+y, x>} + \overline{<x+y, y>} \\
&= \overline{<x,x> + <y,x>} + \overline{<x,y> + <y,y>} \\
&= <x,x> + <x,y> + \overline{<x,y>} + <y,y> \\
&= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\{<\mathrm{x},\mathrm{y}>\}
\end{aligned}$$

where $\mathrm{Re}\{<\mathrm{x},\mathrm{y}>\}$ denotes the real part of the complex number $<x,y>$. However, for any complex number α, $\mathrm{Re}\{\alpha\} \leq |\alpha|$, and thus we have

$$\begin{aligned}
\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\{<\mathrm{x},\mathrm{y}>\} \\
&\leqslant \|x\|^2 + \|y\|^2 + 2|<x,y>| \\
&\leqslant \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|,
\end{aligned}$$

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

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

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

b) A set of vectors S is <u>orthogonal</u> 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 <u>orthonormal</u>.

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

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

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

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

$$= \langle x, x \rangle + \langle y, y \rangle + 2\cancel{\langle x, y \rangle}$$ 0 because $x \perp y$

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

$\square$

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

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

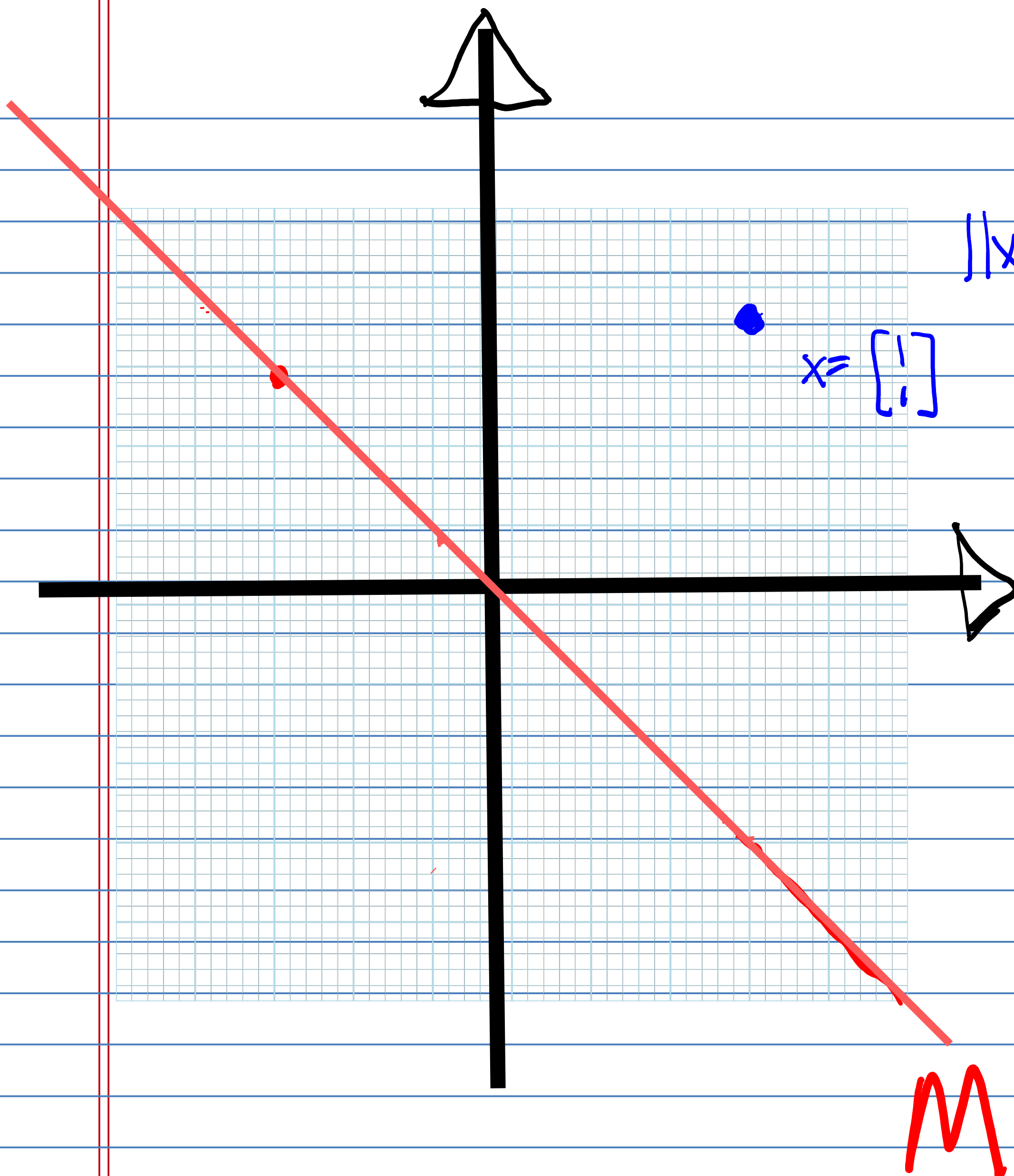
$$\|x - m_0\| \leq \|x - m\| \quad \forall m \in M,$$

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 $x - m_0 \perp M$.

Non-unique minimum is shown next for the 1-norm.

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



The minimizing vector is not unique. Indeed,

$$S = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;\middle|\; x_2 = -x_1 , \; |x_1| \le 1 \right\} = \text{"}\underset{m \in M}{\text{arg min}}\text{"}\, d(x, M)$$

for $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. That is,

$\forall \hat{x} \in S, \quad \|x-\hat{x}\| = 2 = \inf_{m \in M} \|x-m\|$

Exercise $\forall \hat{x} \in S$

$\|x-\hat{x}\|_1 = |1-\hat{x}_1| + |1+\hat{x}_1|, \quad -1 \le \hat{x}_1 \le 1$

$= 1-\hat{x}_1 + 1+\hat{x}_1 = 2$