Norms and inner product
Pre-projection theorem

ROB 501

Necmiye Ozay

- **Last time: Least squares (high level)**

$$\hat{\alpha} = \underset{\alpha \in \mathbb{R}^2}{\operatorname{argmin}} \, ||Y - A\alpha||^2 \qquad \hat{\alpha} = (A^T A)^{-1} A^T Y$$

  - **We will build the proof but we need a few new concepts**

① $(\mathcal{F}^n, \mathcal{F})$ $\quad \|\cdot\|_p$ $\quad \|\cdot\|_\infty \ \ldots$ (last time)

More examples of norms

$\|\cdot\| : \mathcal{X} \longrightarrow \mathbb{R}$ is a norm if

a) $\forall x \in \mathcal{X}$, $\|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$ (positive definiteness)

b) [Triangle Inequality] $\forall x, y \in \mathcal{X}$
$\|x+y\| \leq \|x\| + \|y\|$

c) [positive homogeneity] $\forall \alpha \in \mathcal{F}$, $\forall x \in \mathcal{X}$,
$\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$

where $|\alpha| = \begin{cases} \text{absolute value} & \text{if } \alpha \in \mathbb{R} \\ \text{magnitude} & \text{if } \alpha \in \mathbb{C} \end{cases}$

②

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

$\mathcal{X} = \{ f : [a,b] \longrightarrow \mathbb{R} \mid f \text{ is continuous} \}$, $\mathcal{F} = \mathbb{R}$

$$\|f\|_2 := \sqrt{\int_a^b |f(z)|^2 \, dz}$$ ← integral is well-defined for continuous functions

$$\|f\|_p := \left(\int_a^b |f(z)|^p \, dz \right)^{1/p} \qquad 1 \leq p < \infty$$

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

for continuous functions whose domain is a closed interval

$(\mathcal{X}, \mathbb{R}, \|\cdot\|_2)$

$(\mathcal{X}, \mathbb{R}, \|\cdot\|_p)$

$(\mathcal{X}, \mathbb{R}, \|\cdot\|_\infty)$

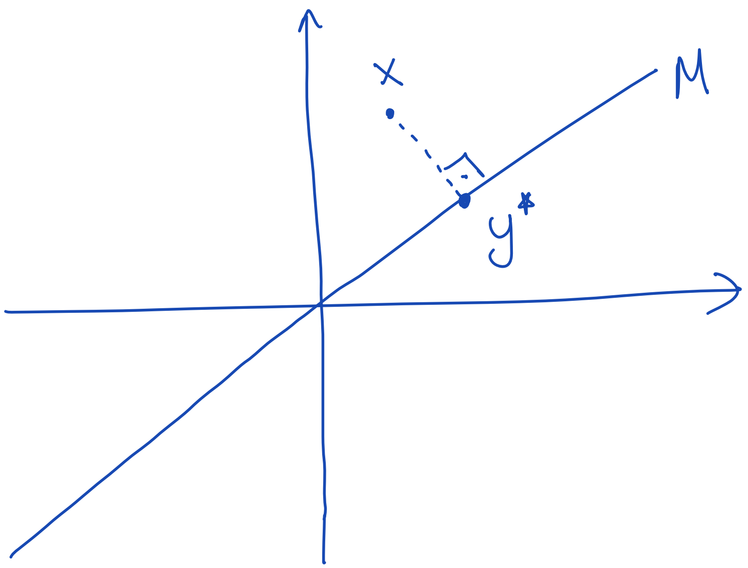
Normed spaces.

More examples of norms

3) Let $\mathcal{X} = \mathbb{R}^{n \times m}$ $\mathcal{F} = \mathbb{R}$. Let $A \in \mathcal{X}$.

$\|A\| = \sqrt{tr(A^T A)}$ is a norm.

Goal for rest of today and Monday



M is a subspace

$$y^* = \underset{y \in M}{\operatorname{argmin}} \|x - y\|_2$$

$$\iff \quad x - y^* \perp M$$

↳ perpendicular (orthognal)

<u>Recall</u>: $x, y \in \mathbb{R}^2$ $\quad x^T y = [x_1 \ x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \sum_{i=1}^{2} x_i y_i$

$$x \perp y \iff x^T y = 0$$

↳ inner product

Think about least squares: "$A\alpha = Y$" want to fit

$$\min_{\alpha \in \mathbb{R}^n} \| Y - A\alpha \|_2 \qquad A \in \mathbb{R}^{N \times n}$$

$$A = [A_1 \mid A_2 \mid \dots \mid A_n]$$

$$\text{col space}(A) := \{ z = A\alpha \mid \alpha \in \mathbb{R}^n \}$$

$$= \text{span}\{ A_1, A_2, \dots, A_n \}$$

is a subspace of $\mathbb{R}^N$

Inner products

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

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

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

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

c) $\forall x \in \mathcal{X}$, $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = 0$

Remark: (a) + (b) $\Rightarrow \langle x, \beta_1 y_1 + \beta_2 y_2 \rangle = \beta_1 \langle x, y_1 \rangle + \beta_2 \langle x, y_2 \rangle$

$\rightarrow$ See notes at the end of the slide deck for the complex case ($\mathcal{F} = \mathbb{C}$).

Examples:

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

b) $(\mathbb{R}^{n \times m}, \mathbb{R})$, $\langle A, B \rangle = tr(A^T B)$

c) $(X, \mathbb{R})$, $X = \{ f : [a,b] \rightarrow \mathbb{R} \mid f \text{ continuous} \}$

$$\langle f, g \rangle = \int_a^b f(z) g(z) \, dz$$

Def: Given $(\mathcal{X}, \mathcal{F})$ a vector space and $\langle \cdot, \cdot \rangle$ an inner product on it, $(\mathcal{X}, \mathcal{F}, \langle \cdot, \cdot \rangle)$ is an <u>inner product space</u>.

— o —

We will mostly focus on the case $(\mathcal{X}, \mathbb{R}, \langle \cdot, \cdot \rangle)$.

<u>Cauchy-Schwartz inequality:</u>

Let $(\mathcal{X}, \mathbb{R}, \langle \cdot, \cdot \rangle)$ be an inner product space. Then,

$$\forall x, y \in \mathcal{X}, \quad |\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}$$

$$\boxed{|\langle x, y \rangle| \leq \|x\| \cdot \|y\|}$$

— o —

Fact: All inner products induce a norm (but converse is not true).

$\|x\| := \langle x, x \rangle^{1/2}$ is always a norm on $(\mathcal{X}, \mathbb{R})$.

Proof: We want to show $\langle x, x \rangle^{1/2}$ satisfies (a) positive definiteness (b) triangular inequality (c) positive homogeneity.

(a) is satisfied by prop. (c) of inner products

(b) We need to show

$\|x+y\| \leq \|x\| + \|y\|$ for $\|x\| \doteq \langle x,x\rangle^{1/2}$

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

$$\langle x+y, x+y\rangle \overset{\text{prop. (b) of inner products}}{=} \underbrace{\langle x, x+y\rangle} + \underbrace{\langle y, x+y\rangle}$$

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

$=$ by prop (a) of inner products

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

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

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

Cauchy-Schwartz

$$\|x+y\|^2 = \langle x+y, x+y\rangle \leq \underbrace{\|x\|^2 + 2\|x\|\|y\| + \|y\|^2}_{(\|x\|+\|y\|)^2}$$

$$\Rightarrow \quad \|x+y\| \leq \|x\| + \|y\| \quad \text{(triangular inequality)}$$

(c) Follows from prop. (b) of inner products

$$\|\alpha x\| = \langle \alpha x, \alpha x\rangle^{1/2} = (\alpha\langle x, \alpha x\rangle)^{1/2}$$

$$= (\alpha^2\langle x,x\rangle)^{1/2} = |\alpha|\langle x,x\rangle^{1/2}$$

$$= |\alpha| \cdot \|x\|$$

Relation of inner products and norms

Inner products

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

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

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

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

c) $\forall x \in \mathcal{X}$, $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = 0$

$\|\cdot\| : \mathcal{X} \to \mathbb{R}$ is a norm if

a) $\forall x \in \mathcal{X}$, $\|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$ (positive definiteness)

b) [Triangle Inequality] $\forall x, y \in \mathcal{X}$ $\|x + y\| \leq \|x\| + \|y\|$

c) [positive homogeneity] $\forall \alpha \in \mathcal{F}$, $\forall x \in \mathcal{X}$, $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$

where $|\alpha| = \begin{cases} \text{absolute value if } \alpha \in \mathbb{R} \\ \text{magnitude if } \alpha \in \mathbb{C} \end{cases}$

Claim: For any inner product <.,.>, the function $(\langle x,x \rangle)^{1/2}$ is a norm (we call it the norm induced by the inner product), i.e., $||x|| = (\langle x,x \rangle)^{1/2}$

Proof:

- ”a) positive definiteness” of the norm follows from property “c)” of the inner product
- We showed ||.|| defined this way satisfies the triangle inequality last time
- “c) positive homogeneity” of the norm follows from:

$$||\alpha x|| = (\langle \alpha x, \alpha x \rangle)^{1/2} = (\alpha \langle x, \alpha x \rangle)^{1/2} = (\alpha^2 \langle x,x \rangle)^{1/2} = |\alpha|(\langle x, x \rangle)^{1/2}$$

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

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

(c) A vector $x \in \mathcal{X}$ is orthagonal to a set S if $\forall y \in S$, we have $x \perp y$

(d) A set of vectors S is <u>orthanormal</u> if $\forall x \in S$, $\|x\| = 1$ and set S is <u>orthagonal</u>.

<u>Remark</u> $\| \cdot \|$ is always the norm $\langle x, x \rangle^{1/2}$ whenever $(\mathcal{X}, \mathcal{F}, \langle \cdot, \cdot \rangle)$ is coming from mentioned.

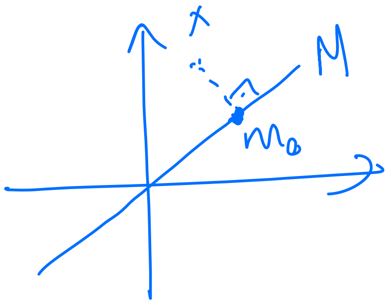
Pythagorean Theorem

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

Proof: for $\mathbb{F} = \mathbb{R}$:

$$\|x+y\|^2 = \langle x+y, x+y \rangle$$
$$= \langle x, x+y \rangle + \langle y, x+y \rangle$$
$$= \langle x, x \rangle + \underbrace{\cancel{\langle x, y \rangle}}_{=0} + \underbrace{\cancel{\langle y, x \rangle}}_{=0} + \langle y, y \rangle$$
$$= \|x\|^2 + \|y\|^2$$

Pre-projection Theorem



$(X, \mathbb{R}, \langle \cdot , \cdot \rangle)$ be an inner product space.

Let $M \subset X$ be a subspace, and $x \in X$.

Then,

(a) If $m_0 \in M$ s.t. $\|x - m_0\| \leq \|x - m\| \ \forall m \in M$, then m_0 is unique.

(b) A necessary and sufficient condition for $m_0 \in M$ to be a minimizer of $\min_{m \in M} d(x, M)$ is that the error vector

$(x-m_0)$ is orthogonal to M, i.e. $(x-m_0) \perp M$

Proof: $(\Longrightarrow)$

Claim b) Let $m_0 \in M$. If $\underbrace{\|x-m_0\| = d(x,M)}_{p}$,

then $\underbrace{(x-m_0) \perp M}_{q}$. $\quad p \Rightarrow q$

Contrapositive: $\neg q \Rightarrow \neg p$

$$x-m_0 \not\perp M \Longrightarrow \|x-m_0\| > d(x,M)$$

$$x-m_0 \not\perp M \Longrightarrow \exists\, \bar{m} \in M \text{ s.t. } x-m_0 \not\perp \bar{m}$$

$$\langle x-m_0, \bar{m} \rangle \neq 0$$

$$\Rightarrow (\bar{m} \text{ is non-zero; i.e, } \|\bar{m}\| \neq 0)$$

We can write:

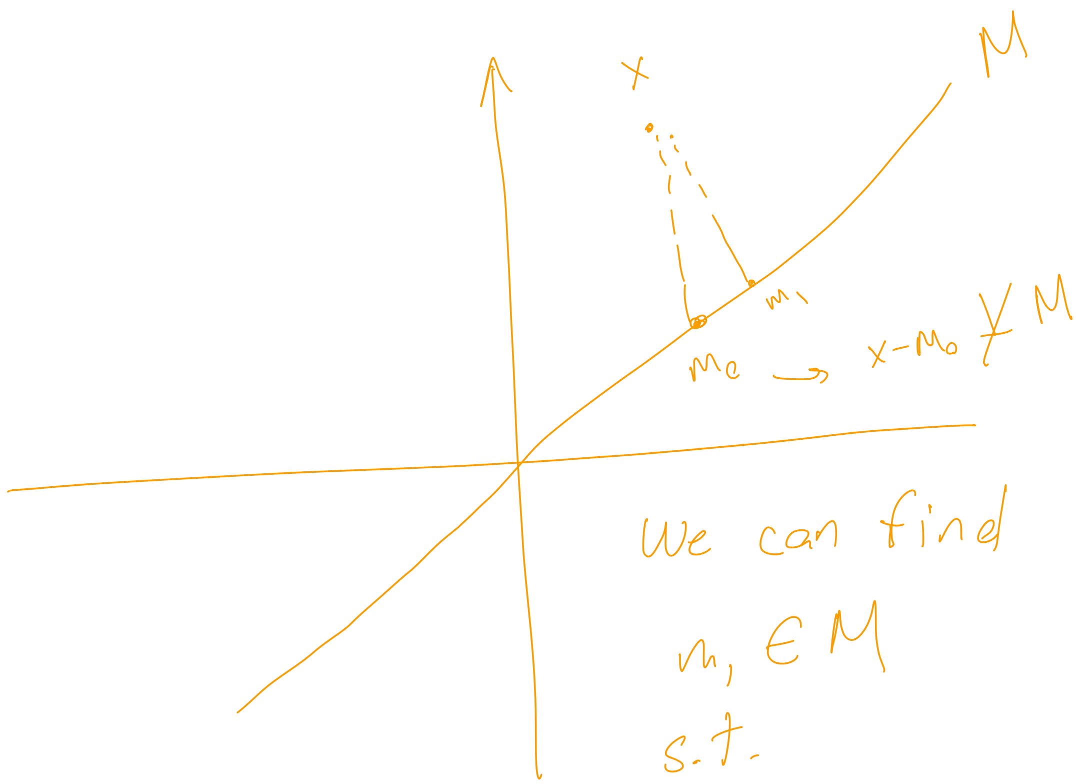
$$\left\langle x-m_0, \frac{\bar{m}}{\|\bar{m}\|} \right\rangle = \underbrace{\frac{1}{\|\bar{m}\|}}_{\neq 0} \underbrace{\langle x-m_0, \bar{m} \rangle}_{\neq 0} \neq 0$$

Thus, without loss of generality, we can assume $\|\bar{m}\| = 1$.

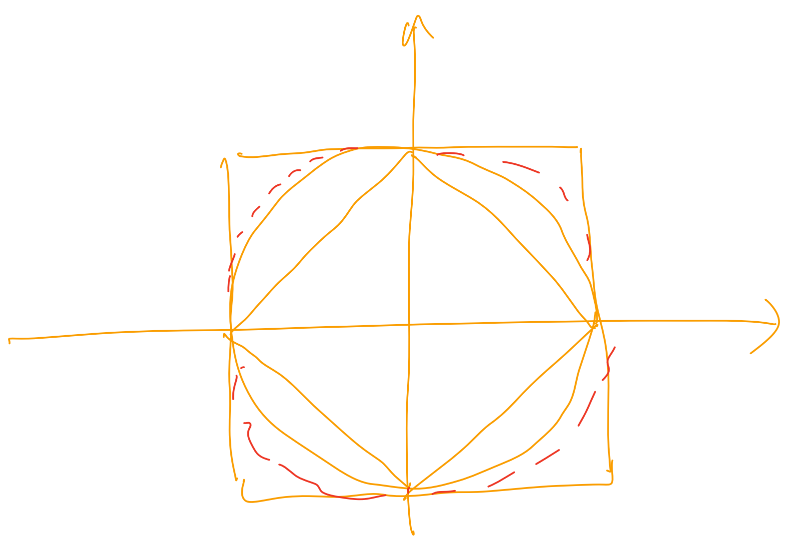
Define: $\beta = \langle x-m_0, \bar{m} \rangle \neq 0$

$m_1 = m_0 + \beta \bar{m} \in M$

"To show $\|x - m_1\| < \|x - m_0\|$"



We can find $m_1 \in M$ s.t. $\|x - m_0\| > \|x - m_1\|$



$\|x\|_2 = \langle x, x \rangle^{1/2}$

inner product $\overset{\not\Leftarrow}{\longrightarrow}$ norms $\overset{\not\Leftarrow}{\longrightarrow}$ distances

please review

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

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

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 $\forall \alpha_1, \alpha_2 \in \mathbb{C}$,

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

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

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

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. ■

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

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

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

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

■

OFFICE HOURS

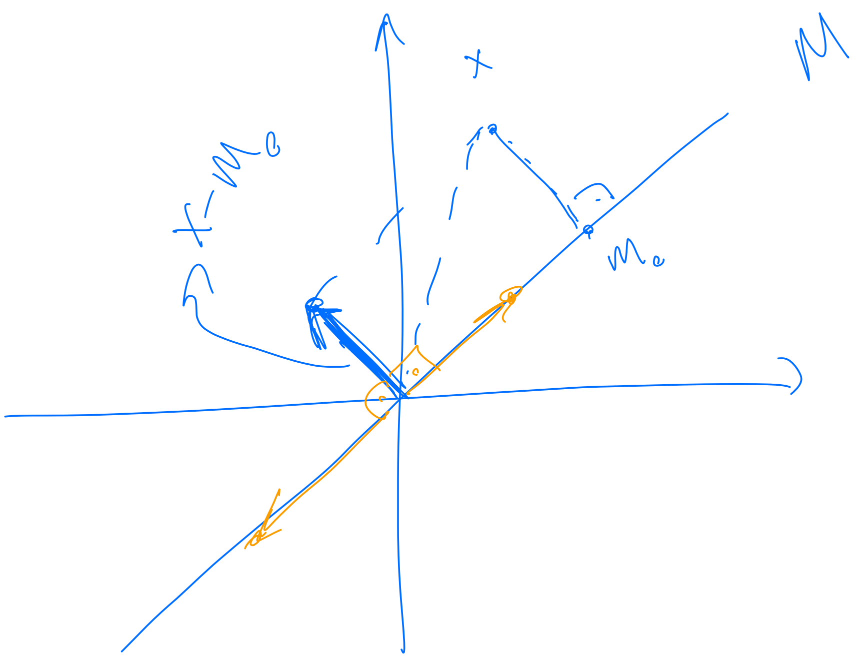
$x_1, x_2, y_1, y_2 \in \mathbb{R}^2$

$$y_1 = P x_1$$

$$y_2 = P x_2$$

$$\underbrace{[y_1 ; y_2]}_{A} = P \underbrace{[x_1 ; x_2]}_{B}$$

$$A = P \cdot B$$



$x - m_0$