X Banach space (with a norm $\|\cdot\|$)

$(\cdot,\cdot)$ scalar product associated to the norm if

- $X \times X \longrightarrow \mathbb{R}$

$x, y \longrightarrow (x,y) = x \cdot y = \langle x, y \rangle \in \mathbb{R}$

1) symmetric $(x,y) = (y,x)$

2) linear $(x+z, y) = (x,y) + (z,y)$, $(\lambda x, y) = \lambda (x,y)$ $\lambda \in \mathbb{R}$

3) continuous if $x_n \to x$ in X $(\|x_n - x\| \to 0)$

$(x_n, y) \longrightarrow (x,y)$

4) $|(x,y)| \le \|x\| \, \|y\|$ (Cauchy - Schwartz inequality

it is associated with the norm if

$(x, x) = \|x\|^2$ $\qquad \|x\| = \sqrt{(x,x)}$

Example:

$I \subseteq \mathbb{R}$)

$(L^2(I), \|\cdot\|_2)$ is an Hilbert space

$f, g \in L^2(I)$ by Hölder inequality

$$(f,g) = \int_I f(x)g(x)\,dx < +\infty$$

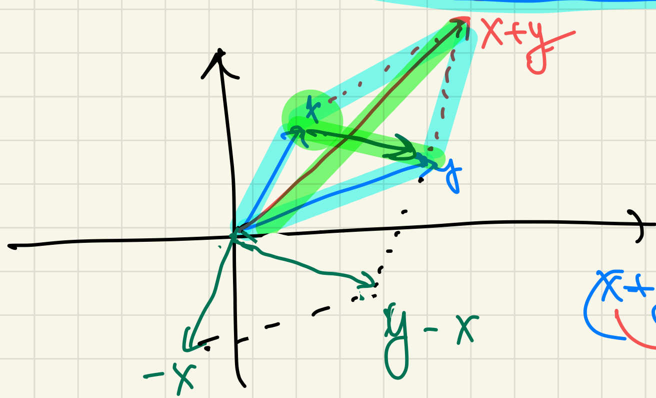
$$|(f,g)| = \left|\int_I f(x)g(x)\,dx\right| \leq \int_I |f(x)||g(x)|\,dx \overset{\text{Hölder}}{\leq} \|f\|_2 \|g\|_2$$

Hölder inequality coincides with the Cauchy-Schwartz inequality

Prop Let $(X, \|\cdot\|)$ be a Banach space.

Then the norm $\|\cdot\|$ is associated to a scalar product if and only if it satisfies the PARALLELOGRAM IDENTITY

$$\forall\, x, y \in X \quad \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$



$X = \mathbb{R}^2$

$$(x+y, x+y) + (x-y, x-y) = \underbrace{2(x,x)} + \underbrace{2(y,y)}$$

Examples of Banach spaces

$I \subseteq \mathbb{R}$

- $L^p(I), \|\cdot\|_p$ $\qquad f \in L^p(I) \Leftrightarrow \int_I |f(x)|^p dx < +\infty$

$$\|f\|_p = \left[\int_I |f(x)|^p dx\right]^{1/p}$$

- $(L^\infty(I), \|\cdot\|_\infty)$ $\qquad f \in L^\infty(I)$ $\quad \exists\, C$ $\quad \sup_{x \in I} |f(x)| \leq C$

$$\|f\|_\infty = \sup_{x \in I} |f(x)|$$

- $(\mathcal{C}(\bar{I}), \|\cdot\|_\infty)$ $\qquad f \in \mathcal{C}(\bar{I})$ if f is continuous in $\bar{I}$, $\|f\|_\infty = \max_{x \in \bar{I}} |f(x)|$

THE UNIQUE HILBERT SPACE is $(L^2(I), \|\cdot\|_2)$

Example of Banach spaces.

$(\Omega, \mathcal{F}, \mathbb{P})$ a probability space

$\mathcal{F}$ filtration (a σ-algebra collecting all events)

Ω is a set

$\mathbb{P}$ is a probability measure on Ω

$$\mathbb{P} : \mathcal{F} \longrightarrow [0, +\infty]$$
$$A \longmapsto \mathbb{P}(A)$$

$X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow \mathbb{R}$ measurable

(RANDOM VARIABLE)

X is "transporting" $\mathbb{P}$ measure to a measure on $\mathbb{R}$. $B \in \mathcal{B}(\mathbb{R})$

$$\underbrace{\mathbb{P}_X}_{\text{Borel measure on } \mathbb{R}.}(B) = \mathbb{P}\{\omega \mid X(\omega) \in B\} = \mathbb{P}(X^{-1}(B))$$

$\underline{M}^1 = \{ X$ random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}(X) < +\infty \}$ (it is a vectorial space on $\mathbb{R}$)

$$\mathbb{E}(X) = \int_{\mathbb{R}} x \, d\mathbb{P}_X(x)$$

what is $\mathbb{E}(X)$?

→ X is discrete (with a finite number of values)

$$\mathbb{P}(\omega \mid X(\omega) = i) = \underbrace{\mathbb{P}_X\{i\}}_{} \qquad i = 1 \ldots k$$

$$\mathbb{E}(X) = \sum_{i=1}^{k} \mathbb{P}_X\{i\} \cdot i$$

$$\mathbb{P}_X = \sum_{i=1}^{k} \mathbb{P}(\omega \mid X(\omega)=i) \cdot \delta_{\{i\}}$$

→ X is absolutely continuous $\quad \mathbb{P}_X \ll \mathcal{L}$ (Lebesgue m.)

$\mathbb{P}_X$ has a density f_X $\qquad \mathbb{E}(X) = \int_{\mathbb{R}} x f_X(x)\, dx$

$M^1 = \{X$ random variables, $\mathbb{E}(X) < +\infty\}$

is a vectorial space

$X_1, X_2 \qquad \lambda X_1 + \mu X_2 \in M^1$

$\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2) \ldots$

$\|X\|_1 = \mathbb{E}(|X|)$ this is a norm on M^1

1) $E|X| = 0 \iff |X| = 0$ with probability 1.

2) $\mathbb{E}|\lambda X| = |\lambda|\, E(|X|) \quad \forall \lambda \in \mathbb{R}$.

3) $\mathbb{E}(|X+Y|) \leq \mathbb{E}(|X| + |Y|) = \mathbb{E}|X| + \mathbb{E}(Y)$.

Def. $X_n \to X$ in mean if $X_n \to X$ in M^1, that is $\mathbb{E}(|X_n - X|) \to 0$

$(M^1, \|\cdot\|_1)$ is a Banach space.

Obs $\|\cdot\|_1$ is NOT ASSOCIATED TO any scalar product

It is sufficient to show that the parallelogram identity is not satisfied for some $X, Y \in M^1$.

Example

X Bernoulli of parameter $p \in (0,1)$, $X : \Omega \to \{0, 2\} \subseteq \mathbb{R}$

$\{\omega \mid X(\omega) = 0\} = p \qquad \{\omega \mid X(\omega) = 2\} = 1-p$

$E|X| = E(X) = p \cdot 0 + (1-p) \cdot 2 = 2(1-p) = 2 - 2p$

Y Bernoulli of parameter $p \in (0,1)$, $Y : \Omega \to \{1, 3\} \subseteq \mathbb{R}$

$$\{\omega \mid Y(\omega) = 1\} = p \qquad \{\omega \mid Y(\omega) = 3\} = 1-p$$

$$E(|Y|) = E(Y) = p \cdot 1 + (1-p) \cdot 3 = p + 3 - 3p = 3 - 2p$$

$$X + Y : \Omega \to \{1, 3, 5\}$$

$$Y - X : \Omega \to \{1, -1, 3\} \qquad |Y-X| : \Omega \to \{1, 2\}$$

$$\mathbb{P}(\omega \mid X+Y = 1\} = P(X(\omega)=0)\, P(Y(\omega) = 1) = p^2$$

$$P(\omega \mid X+Y = 3\} = P(X(\omega)=0)\, P(Y(\omega) = 3) + P(X(\omega)=2) P(Y(\omega)=1) =$$

$$= p \cdot (1-p) + (1-p)\, p = 2p - 2p^2$$

$$P(\omega \mid X+Y = 5) = P(X(\omega) = 2)\, P(X(\omega) = 3) = (1-p)(1-p)$$

$$= (1-p)^2$$

$$\mathbb{E}(X+Y) = \mathbb{E}|X+Y| = 1 \cdot p^2 + 3 \cdot (2p - 2p^2) + 5(1-p)^2 =$$

$$= \cancel{p^2} + 6p - \cancel{6p^2} + 5 + \cancel{5p^2} - 10p = 5 - 4p$$

$$P(Y - X = 1) = P(Y(\omega) = 1)P(X(\omega) = 0) + P(Y(\omega) = 3)P(X(\omega) = 2)$$

$$= p \cdot p + (1-p)(1-p) = p^2 + 1 + p^2 - 2p =$$

$$= 2p^2 + 1 - 2p$$

$$P(Y - X = -1) = P(Y(\omega) = 1)P(X(\omega) = 2) = p \cdot (1-p) = p - p^2$$

$$P(|Y - X| = 1) = P(Y - X = 1) + P(Y - X = -1) =$$

$$= 2p^2 + 1 - 2p + p - p^2 = p^2 - p + 1$$

$$P(|Y - X| = 3) = P(Y = 3)P(X = 0) = (1-p) \cdot p = p - p^2$$

$$E(|Y - X|) = 1 \cdot (p^2 - p + 1) + 3(p - p^2) =$$

$$= p^2 - p + 1 + 3p - 3p^2 = 2p - 2p^2 + 1$$

$$[\mathbb{E}(|X+Y|)]^2 + [\mathbb{E}(|X-Y|)]^2 = \|X+Y\|_1^2 + \|X-Y\|_1^2 =$$

$$= (5-4p)^2 + (2p-2p^2+1)^2$$

$$\neq 2(2-2p)^2 + 2(3-2p)^2 = 2\|X\|_1^2 + 2\|Y\|_1^2$$

$$= 2[\mathbb{E}(|X|)]^2 + 2(\mathbb{E}|Y|)^2$$

So the parallelogram identity IS NOT VERIFIED

$\|\cdot\|_1$ is NOT ASSOCIATED TO a SCALAR PRODUCT $\Rightarrow$ $(M^1, \|\cdot\|_1)$ is NOT Hilbert!

We introduce the following space:

$$M^2 = \{X \text{ random variables } \mathbb{E}(X^2) < +\infty\}$$
$$\mathbb{E}(X) < +\infty$$

$$\|X\|_2 = \left[\mathbb{E}(|X|^2)\right]^{1/2}$$

$\boxed{\text{this is a NORM on the space } M^2.}$

Def: We say that $X_n \to X$ in MEAN SQUARE if $\|X_n - X\|_2 \to 0$ that is $\mathbb{E}\left[|X_n - X|^2\right] \to 0$ as $n \to +\infty$

Note that:

$\mathbb{E}(X \cdot X) = \|X\|_2^2$.

We define the scalar product $M^2 \times M^2 \to \mathbb{R}$

$$X, Y \to \mathbb{E}(X \cdot Y)$$

We need to prove that $\forall X, Y \in M^2$, $\mathbb{E}(X \cdot Y) < +\infty$.

proof :

$$a = \frac{|X|}{\left[\mathbb{E}|X|^2\right]^{1/2}} \qquad b = \frac{|Y|}{\left[\mathbb{E}(|Y|^2)\right]^{1/2}}$$

YOUNG $\quad ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2 \qquad \frac{|X||Y|}{(\mathbb{E}|X|^2)^{1/2}(\mathbb{E}(|Y|^2))^{1/2}} \leq \frac{1}{2}\frac{|X|^2}{\mathbb{E}|X|^2} + \frac{1}{2}\frac{|Y|^2}{\mathbb{E}|Y|^2}$

$\rightarrow$ take the expected value

$$\frac{\mathbb{E}(|X||Y|)}{(\mathbb{E}|X|^2)^{1/2}(\mathbb{E}(|Y|^2))^{1/2}} \leq \frac{1}{2}\frac{\mathbb{E}|X|^2}{\mathbb{E}|X|^2} + \frac{1}{2}\frac{\mathbb{E}|Y|^2}{\mathbb{E}|Y|^2} = \frac{1}{2} + \frac{1}{2} = 1$$

$|\mathbb{E}(XY)| \leq \mathbb{E}|XY| \leq \left[\mathbb{E}|X|^2\right]^{1/2}\left[\mathbb{E}|Y|^2\right]^{1/2} < +\infty$ since $X, Y \in M^2$

So $(M^2, \|\cdot\|_2)$ is an HILBERT SPACE

Def $X, Y \in M^2$ X is orthogonal to Y if

$$\underbrace{\mathbb{E}}_{\sim}(X \cdot \underbrace{Y}_{\sim}) = 0$$

$C \subseteq M^2$ $C^\perp = \{ Y \in M^2$ such that $\mathbb{E}(X \cdot Y) = 0$
$\forall X \in C \}$

Example $C = \{ X \in M^2 \ \ \mathbb{E}(X) = 0 \}$

$C^\perp = ?$ We prove that $C^\perp = \{$ constant random variables $\}$.

1) if $Y \equiv c$ constant then
$E(X \cdot Y) = c\, \mathbb{E}(X) = 0 \quad \forall X \in C$.

2) if Y is such that $E(X \cdot Y) = 0 \ \forall X \in C \implies$

$$\Rightarrow \mathbb{E}(X\cdot(Y-\mathbb{E}(Y))) = \underbrace{\mathbb{E}(X\cdot Y)}_{=0} - \mathbb{E}(Y)\cdot \underbrace{\mathbb{E}(X)}_{=0} = 0$$

$$\forall X \in C$$

$$\Downarrow$$

but $Y-\mathbb{E}(Y) \in C$ since $\mathbb{E}(Y-\mathbb{E}(Y)) = 0$!

so $\mathbb{E}((Y-\mathbb{E}(Y))(Y-\mathbb{E}(Y))) = 0$

$$\mathbb{E}(|Y-\mathbb{E}(Y)|^2) = 0 \Rightarrow |Y-\mathbb{E}(Y)|^2 = 0 \text{ with probability } 1$$

$\Rightarrow Y = \mathbb{E}(Y)$ with probability one

$\Rightarrow Y$ is constant.

$(\Omega, \mathcal{F}, \mathbb{P})$ $\quad$ $\mathcal{F}$ filtration

$\mathcal{G} \subseteq \mathcal{F}$ $\quad$ $\mathcal{G}$ a sub filtration

$\mathcal{G}$ is a σ-algebra

$M^2_{\mathcal{G}} = \{ X \in M^2$ such that X is meas. with resp to $\mathcal{G}$

$X : (\Omega, \underline{\underline{\mathcal{G}}}, \mathbb{P}) \to \mathbb{R}$ is a measurable function:

$\forall B \in \mathcal{B}(\mathbb{R})$ $\quad$ $X^{-1}(B) = \{\omega \mid X(\omega) \in B\} \in \mathcal{G} \subseteq \mathcal{F} \} \subseteq M^2$

($M^2_{\mathcal{G}}$ is a closed subspace of M^2).

ORTHOGONAL PROJECTION THEOREM.

$M^2 = (M^2_{\mathcal{F}}, \|\cdot\|_2)$ $\quad$ $\mathcal{G} \subseteq \mathcal{F}$ ($\mathcal{G}$ is a σ-algebra contained in $\mathcal{F}$).

$M^2_{\mathcal{G}}$ is a closed subspace of M^2.

$\forall X \in M^2 \quad \exists! \, Y \in M^2_{\mathcal{G}}$ such that

$$d(X,Y) = \|X-Y\|_2 = \left[\mathbb{E}(X-Y)^2\right]^{1/2} =$$

$$= \min_{Z \in M^2_{\mathcal{G}}} \left[\mathbb{E}(X-Z)^2\right]^{1/2} = \min_{Z \in M^2_{\mathcal{G}}} d(X,Z).$$

$X - Y$ is orthogonal to $M^2_{\mathcal{G}}$

$$\left(X = Y + W \qquad W \in \left(M^2_{\mathcal{G}}\right)^{\perp} \right)$$

$Y = \mathbb{E}(X \mid \mathcal{G})$ = conditional expectation of X given $\mathcal{G}$

= best estimator of X, given the information contained in the filtration $\mathcal{G}$.

Y is MEASURABLE with respect to $\mathcal{G}$.

$X - \mathbb{E}(X|\mathcal{G})$ is orthogonal to every element of $M^2_{\mathcal{G}}$

$$\left[\begin{array}{l} X - \mathbb{E}(X|\mathcal{G}) \perp \underline{\mathbb{E}(X|\mathcal{G})}^{\in M^2_{\mathcal{G}}} \\ \Rightarrow \mathbb{E}\big((X - \mathbb{E}(X|\mathcal{G})) \cdot \mathbb{E}(X|\mathcal{G})\big) = 0 \\ \mathbb{E}(X\,\mathbb{E}(X|\mathcal{G})) = \mathbb{E}([\mathbb{E}(X|\mathcal{G})]^2) \end{array}\right.$$

$\forall Z \in M^2_{\mathcal{G}}$ $\boxed{\mathbb{E}((X - \mathbb{E}(X|\mathcal{G}))Z) = 0}$

---

observe that every constant random variable is in $M^2_{\mathcal{G}}$ for every $\mathcal{G}$ σ-algebra take Z constant r.v.

$$\boxed{Z(\omega) \equiv c \quad \forall \omega}$$

$\forall B \in \mathcal{B}(\mathbb{R})$ $\quad Z^{-1}(B) = \{\omega \mid \overbrace{Z(\omega) \in B}^{c}\} = \{\omega \mid c \in B\}$

$$= \begin{cases} \emptyset & \text{if } c \notin B \\ \Omega & \text{if } c \in B \end{cases}$$

if Z is constant $\Rightarrow Z \in M^2_{\mathcal{G}}$ $\forall \mathcal{G}$ σ-algebra

because $\forall \mathcal{G}$ σ-algebra, $\emptyset, \Omega \in \mathcal{G}$

$$\overbrace{\mathbb{E}\left(\left(X - \mathbb{E}(X \mid \mathcal{G})\right) \cdot c\right) = 0} \Longrightarrow \mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{G}))$$

(apply the orthogonality cond to $Z(\omega) \equiv c$)

A particular case is the following.

$\boxed{\text{Fix } Z \in M^2}$ $\boxed{Z^{-1}(B) = \{\omega \mid Z(\omega) \in B\} \in \mathcal{Y}}$

$\mathcal{G}(Z)$ = filtration generated by Z =

= smallest σ-algebra which contains all the elements $\boxed{Z^{-1}(B)}$ for $\underline{B \in \mathcal{B}(\mathbb{R})}$

Example:

if $Z \equiv c$ (Z constant) $\mathcal{G}(Z) = \{\emptyset, \Omega\}$

So for $X \in M^2$ $\exists!$ $\mathbb{E}(X \mid \mathcal{G}(Z)) = \mathbb{E}(X \mid Z)$ CONDITIONAL EXPECTATION

the orthogonal projection of X on $M^2_{\mathcal{G}(Z)}$.

$$\begin{tikzcd}
(\Omega, \mathbb{P}, \mathcal{F}) \arrow[r, "Z"] \arrow[rr, bend right] & \mathbb{R} \arrow[r, "h"] & \mathbb{R} \arrow[l, bend right]
\end{tikzcd}$$

$M^2_{\mathcal{G}(Z)} = \{ Y \in M^2$ such that Y is measurable w.r.to $\mathcal{G}(Z)$.

$Y^{-1}(B) = \{\omega \mid Y(\omega) \in B\} \in \underbrace{\mathcal{G}(Z)} \} =$

$= \{ h(Z)$, for $h : \mathbb{R} \to \mathbb{R}$ Borel measurable$\}$

$E(h(Z)^2) < +\infty$.

Since $M^2_{g(Z)} = \{ h(Z) \, , \quad h : \mathbb{R} \to \mathbb{R} \text{ Borel measurable} \}$

$\mathbb{E}(X \mid Z) = h(Z)$ where $h : \mathbb{R} \to \mathbb{R}$ is the function which minimizes:

$$\underbrace{\mathbb{E}\left(|X - h(Z)|^2 \right)}_{= \left[\text{dist}\left(X, h(Z)\right)\right]^2} = \min_{\substack{g : \mathbb{R} \to \mathbb{R} \\ g \text{ measurable}}} \mathbb{E} \underbrace{|X - g(Z)|^2}_{= \left(\text{dist}\left(X, g(Z)\right)\right)^2} ,$$

$h(Z) = \mathbb{E}(X \mid Z)$ is the BEST ESTIMATOR of X given Z: it is a function of Z which has minimal distance (in $\|\cdot\|_2$ sense) from X.

$\mathbb{E}(X|Z) = h(Z)$ where

$$\underbrace{\mathbb{E}(|X - \mathbb{E}(X|Z)|^2)}_{\mathbb{E}|X - h(Z)|^2} = \min_{\substack{g:\mathbb{R}\to\mathbb{R} \\ \text{measurable}}} \mathbb{E}(|X - g(Z)|^2$$

INSTEAD OF MINIMIZING AMONG ALL POSSIBLE MEASURABLE functions g, I RESTRICT the set (IMPOSING a CONSTRAINT).

I impose the constraint that g is LINEAR:

$g: \mathbb{R} \to \mathbb{R}$ linear $\iff$ $g(x) = Ax + B \quad \forall x$

for $A, B \in \mathbb{R}$.

I consider this other problem

$$\min_{g(x)=Ax+B} \mathbb{E}\,|X-g(Z)|^2 = \min_{(A,B)\in\mathbb{R}^2} \mathbb{E}\,|X-AZ-B|^2$$

$$= \mathbb{E}\,|X-\bar{A}Z-\bar{B}|^2 \qquad (\bar{A}\ \bar{B}\ \text{MINIMUM})$$

$\bar{A}Z+\bar{B}$ is the BEST LINEAR ESTIMATOR of X given Z (it is not the best possible estimator of X, which is $\mathbb{E}(X|Z)=h(Z)$)

h in general NOT LINEAR..

$$\bar{A} = \frac{\mathrm{Cov}(X,Z)}{\mathrm{Var}(Z)} \qquad \bar{B} = \mathbb{E}(X)\cdot\left[1-\frac{\mathrm{Cov}(X,Z)}{\mathrm{Var}\,Z}\right]$$

How to find $\bar{A}, \bar{B}$?

$$\mathbb{E}(X - AZ - B)^2 =$$

$$E(X)^2 + A^2\,\mathbb{E}(Z)^2 + B^2 - 2A\,\mathbb{E}(XZ) - 2B\,\mathbb{E}(X) + 2AB\,\mathbb{E}(Z)$$

$$F(A, B) = [\mathbb{E}(X)^2] + A^2\,\mathbb{E}[(Z)^2] + B^2 - 2A\,\mathbb{E}(XZ) - 2B\,\mathbb{E}(Z) + 2AB(\mathbb{E}(Z))$$

$F : \mathbb{R}^2 \to \mathbb{R}$

$$\frac{\partial F}{\partial A} = 0 \qquad \frac{\partial F}{\partial B} = 0$$

$$D^2 F = \begin{pmatrix} \frac{\partial^2 F}{\partial A^2} & \frac{\partial^2 F}{\partial A \partial B} \\ \frac{\partial^2 F}{\partial A \partial B} & \frac{\partial^2 F}{\partial B^2} \end{pmatrix}$$

If $Z \equiv c$ $\quad \underbrace{\mathcal{G}(Z) = \{\Omega, \phi\}}$

$$\mathbb{E}(X \mid \mathcal{G}(Z)) = \mathbb{E}(X).$$

$$M^2_{\mathcal{G}(Z)} = \{\text{constant random variable}\}$$