Def. HILBERT SPACE

H is a vectorial space on $\mathbb{R}$ (or on $\mathbb{C}$)

$\|\cdot\|_H$ norm $\quad d_H(x,y) = \|x-y\|_H$

H is Banach with respect to the convergence associated with d_H

On H is defined a SCALAR PRODUCT

$$\langle \cdot , \cdot \rangle : H \times H \longrightarrow \mathbb{R}$$

1) $\langle x,y \rangle = \langle y,x \rangle$ SYMMETRIC

2) LINEAR $\quad \langle \lambda x + \mu z, y \rangle =$ $\quad \lambda, \mu \in \mathbb{R}$

$= \lambda \langle x,y \rangle + \mu \langle z,y \rangle$ $\quad x, z, y \in H$

3) CONTINUOUS if $x_n \to x$ in H

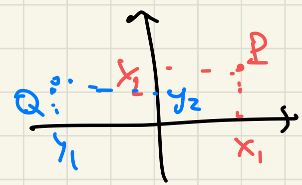
$(d_H(x_n, x) = \|x_n - x\|_H \to 0$ as $n \to +\infty)$

$\forall y \in H$ $\quad \langle x_n, y \rangle \xrightarrow{n \to +\infty} \langle x, y \rangle$

4) CAUCHY-SCHWARTZ $\quad |\langle x, y \rangle| \le \|x\|_H \, \|y\|_H$

5) it is associated with the norm $\quad \langle x, x \rangle = \|x\|_H^2$

Ex $\quad \mathbb{R}^2 = \{ (x_1, x_2) \quad x_i \in \mathbb{R} \}$



$$\langle \underbrace{(x_1, x_2)}_{P}, \underbrace{(y_1, y_2)}_{Q} \rangle = x_1 y_1 + x_2 y_2$$

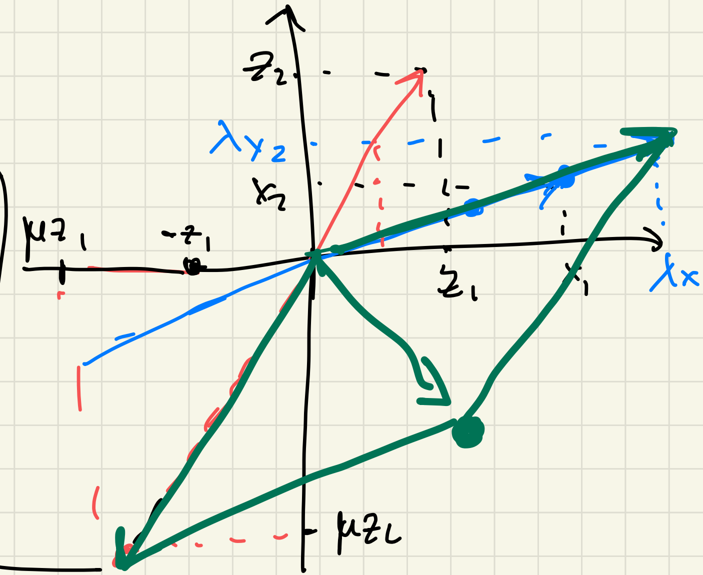
$\mathbb{R}^n$ $\quad \langle (x_1 \dots x_n), (y_1 \dots y_n) \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$

$$\lambda(x_1, x_2) + \mu(z_1, z_2) = (\lambda x_1 + \mu z_1, \lambda x_2 + \mu z_2)$$

$\mu < 0$

$$\langle (\lambda(x_1, x_2) + \mu(z_1, z_2)), (y_1, y_2) \rangle =$$

$$= \lambda x_1 y_1 + \mu z_1 y_1 + \lambda x_2 y_2 + \mu z_2 y_2$$



$$|\langle (x_1, x_2), (y_1, y_2) \rangle| =$$

$$= |x_1 y_1 + x_2 y_2| \leq \underbrace{\sqrt{x_1^2 + x_2^2}}_{|(x_1, x_2)|} \underbrace{\sqrt{y_1^2 + y_2^2}}_{|(y_1, y_2)|}$$

$$(x_1^n, x_2^n) \to (x_1, x_2)$$

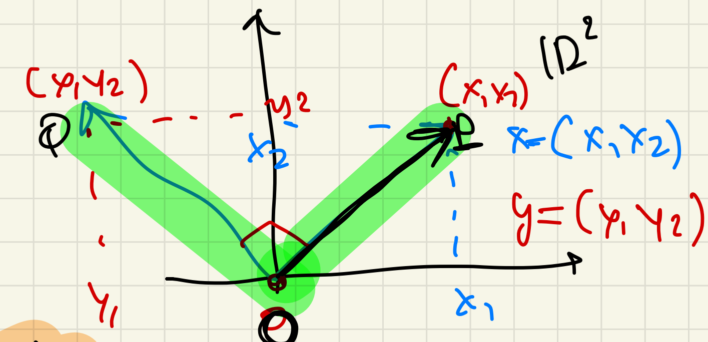
$$x_1^n \to x_1$$

$$x_1^n \to x_2$$

The notion of scalar product permits to introduce the DEFINITION of ORTHOGONALITY

<u>Def</u>: $x \perp y$ ($x, y \in H$) x is orthogonal to y

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



$\langle (x_1, x_2), (y_1, y_2) \rangle = 0 = x_1 y_1 + x_2 y_2$

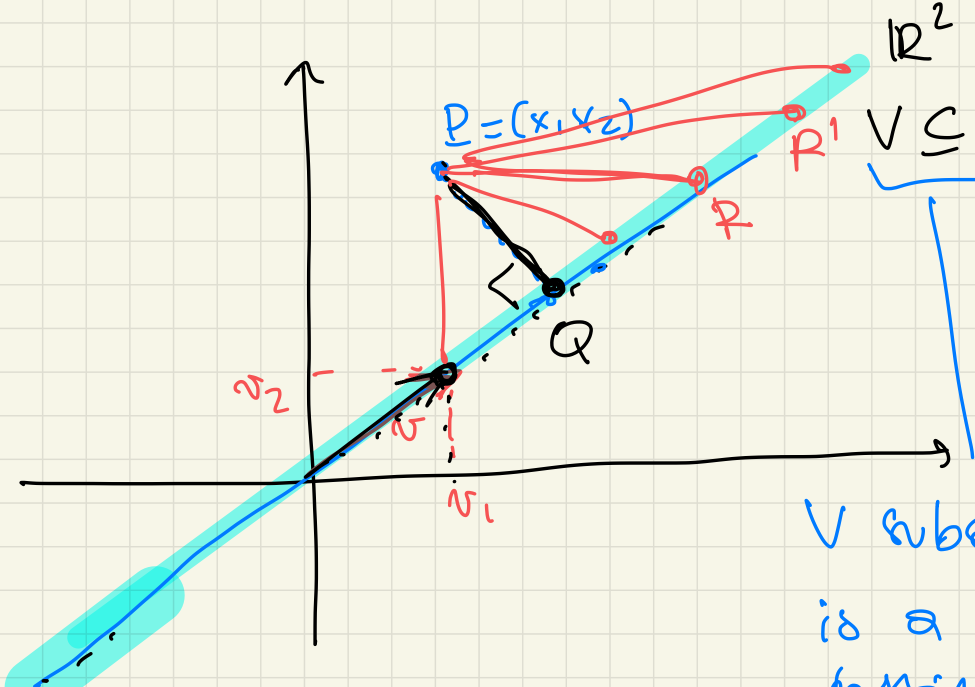
$\iff \overline{OP} \perp \overline{OQ}$

$P = (x_1, x_2)$ $Q = (y_1, y_2)$

$O = (0, 0)$

The main theorem in Hilbert spaces is the

ORTHOGONAL PROJECTION THEOREM



$V \subseteq \mathbb{R}^2$ V subspace

$e_1 = (1,0)$
$e_2 = (0,1)$

V subspace of dim 1 is a straight line passing through 0.

$V = \{ \lambda (v_1, v_2) \, , \, \lambda \in \mathbb{R} \}$

Q: $\exists \, Q \in V$ such that

$|Q-P| = \min_{R \in V} |R-P|$

ORTHOGONAL PROJECTION THEOREM

H Hilbert $\qquad V \subseteq H$ SUBSPACE, CLOSED

(V is a vectorial space CONTAINED in H)

$v_1, v_2 \in V \quad \lambda v_1 + \mu v_2 \in V \quad \forall \lambda, \mu \in \mathbb{R}$

CLOSED $\left[\begin{array}{l} v_n \in V \quad v_n \to x \text{ in } H \quad v_n \text{ is a converging} \\ \text{sequence in } H \implies x \in V \end{array} \right.$

(V contains all the limits of its converging sequences)

then the following holds:

1) $\forall h \in H$ there exists a UNIQUE $v \in V$

such that $\boxed{\|v - h\| = \min_{w \in V} \|w - h\|}$

(there exists a UNIQUE $v \in V$ which has MINIMAL DISTANCE from h, AMONG ALL ELEMENTS IN V)

(2) $\underline{\underline{h - v \in V^{\perp}}}$ (so $\langle h-v, w\rangle = 0$, $\langle v-h, w\rangle = 0$ $\forall w \in V$)

Therefore $h = \underbrace{(h-v)}_{\in V^{\perp}} + \underbrace{v}_{V}$

$\downarrow$

Every element of H can be written in a UNIQUE WAY as the sum of an element of V and an element of $V^{\perp}$.

No proof.

Just prove of the fact that (1) $\Longrightarrow$ (2)

if $v \in V$ is the element at minimal distance from h $\quad \|h-v\| = \min_{w \in V} \|h-w\|$

$\Rightarrow h - v \perp V \qquad (h-v) \in V^{\perp}$.

$r \in \mathbb{R} \qquad w \in V \qquad \underline{\underline{v}} + r w \in V$

$$\|h-v\|^2 \leq \|h-(v+rw)\|^2 \quad \forall r$$

$$\shortparallel$$

$$\langle h-v, h-v \rangle \leq \langle h-(v+rw), h-(v+rw) \rangle \quad \forall r \in \mathbb{R}$$

$$\phi(r) = \langle h - (v + rw), h - (v + rw) \rangle$$

$$\phi(0) \leq \phi(r) \quad \forall r$$

$$\downarrow$$

$$\phi'(0) = 0$$ (since $r = 0$ is a MINIMUM POINT)

$$\phi(r) = \langle h - v - rw, h - v - rw \rangle =$$

$$= \langle h - v, h - v \rangle + \langle -rw, h - v \rangle +$$

$$+ \langle h - v, -rw \rangle + \langle -rw, -rw \rangle =$$

$$= \|h - v\|^2 - 2r \langle h - v, w \rangle + r^2 \|w\|^2$$

$$\phi'(r) = -2 \langle h - v, w \rangle + 2r \|w\|^2$$

$$\phi'(0) = -2 \langle h - v, w \rangle = 0$$ $w \in V \Rightarrow h - v \perp V$

Let us fix a PROBABILITY SPACE

Ω set

$\mathcal{F}$ σ-algebra on Ω (filtration)

$\mathbb{P}$ probability measure

$$M^2 = \{ X : \Omega \to \mathbb{R} \text{ random variable such that } \underbrace{\mathbb{E}(|X|^2)}_{\parallel} < +\infty \}$$

$$\mathbb{E}(|X|^2) = \int_{\mathbb{R}} x^2 \, d\mathcal{L}_X(x)$$

$$\mathcal{L}_X = X \# \underline{\underline{\mathbb{P}}}$$

$$\mathcal{L}_X(A) = \mathbb{P}(\{\omega \mid X(\omega) \in A\})$$

$A \in \mathcal{B}(\mathbb{R})$

$X \in M^2$ $\quad \|X\|_2 = \text{norm} = \sqrt{\mathbb{E}(|X|^2)} = \sqrt{\int_{\mathbb{R}} x^2 \, d\mathcal{L}_X(x)}$

$\downarrow$

distance X, Y

$$\underline{\|X-Y\|_2} = \sqrt{\mathbb{E}(|X-Y|^2)} = \sqrt{\int_{\mathbb{R}\times\mathbb{R}} |x-y|^2 \, d\mathcal{L}_{(X,Y)}(x,y)}$$

$X, Y \in M^2$ $\qquad X : \Omega \to \mathbb{R}$ $\qquad (X,Y) : \Omega \longrightarrow \mathbb{R}\times\mathbb{R} = \mathbb{R}^2$

$\qquad\qquad Y : \Omega \to \mathbb{R}$ $\qquad\qquad \omega \longmapsto (X(\omega), Y(\omega))$

$\downarrow$ I define the joint law $\mathcal{L}_{(X,Y)}$ a measure in $\mathbb{R}^2$

$$\mathcal{L}_{(X,Y)} = (X,Y)_\# \mathbb{P} \quad , \; A, B \in \mathcal{B}(\mathbb{R}) \qquad A \times B \in \mathcal{B}(\mathbb{R}^2)$$

$$\mathcal{L}_{(X,Y)}(A \times B) = \mathbb{P}\{\omega \in \Omega \; , \; X(\omega) \in A \; , \; Y(\omega) \in B\}$$

if X and Y are independent

$$\mathcal{L}_{(X,Y)}(x,y) = \mathcal{L}_X(x)\,\mathcal{L}_Y(y).$$

$$\sim \cdot \sim$$

In M^2 I define the scalar product

$$\langle X, Y\rangle = \mathbb{E}(XY) = \int_{\mathbb{R}\times\mathbb{R}} xy \, d\mathcal{L}_{(X,Y)}(x,y)$$

Note that it satisfies all the properties listed at the beginning

$$E(X\,X) = \mathbb{E}(X)^2 = \|X\|^2$$

Recall $X \perp Y \iff \mathbb{E}(XY)=0$

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

Ex $V = \{ Y \in M^2 \quad \mathbb{E}(Y) = 0 \}$ $\quad V^{\perp} = \{$constant random variables$\}$

Fix $X \in M^2$.

What is the orthogonal projection of X in V?

What is the orthogonal projection of X in $V^{\perp}$?

$$X = X - \mathbb{E}(X) + \mathbb{E}(X)$$

Note that $\mathbb{E}(X)$ is a constant $\Rightarrow \mathbb{E}(X) \in V^{\perp}$

$X - \mathbb{E}(X) \in V$ since $\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(X) = 0$

By the orthogonal projection theorem

$X - \mathbb{E}(X)$ is the orthogonal projection of X in V

(it is the random variable with 0 mean that is at minimal distance from X,

that is: it is the best possible approximation of X by a random variable with 0 mean

$\mathbb{E}(X)$ is the orthogonal projection of X in $V^{\perp}$

($\mathbb{E}(X)$ is the constant which is at minimal distance from X among all other constants

$$\min_{c \in \mathbb{R}} \mathbb{E}(X-c)^2 = \mathbb{E}(X-\mathbb{E}(X))^2.$$

Es $M^2 = \{ X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ random variable
$\mathbb{E}(|X|^2) < +\infty \}$

Let $\mathcal{G} \subsetneq \mathcal{F}$ a σ-algebra contained in $\mathcal{F}$

$M^2_{\mathcal{G}} = \{ X : (\Omega, \mathcal{G}, \mathbb{P}) \to \mathbb{R}$, random variable
$\mathbb{E}(|X|^2) < +\infty \}$

$M^2_{\mathcal{G}}$ are random variables $\mathcal{G}$-measurables

$X \in M^2_{\mathcal{G}} \iff \forall x \in \mathbb{R} \quad \{\omega \in \Omega : X(\omega) \le x\} \in \mathcal{G} \subseteq \mathcal{F} \Rightarrow$
$X \in M^2 \qquad (M^2_{\mathcal{G}} \subseteq M^2)$.

$M^2_{\mathcal{G}} \subsetneq M^2$: let $A \in \mathcal{F} \setminus \mathcal{G}$ (then also $\Omega \setminus A \in \mathcal{F} \setminus \mathcal{G}$

$\mathbb{1}_A : \Omega \to \mathbb{R} \quad \mathbb{1}_A(\omega) = \begin{cases} 0 & \omega \notin A \\ 1 & \omega \in A \end{cases}$ then $\mathbb{1}_A \in M^2 \quad \mathbb{1}_A \notin M^2_{\mathcal{G}}$

$X \in M^2 \qquad X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow \mathbb{R}$

$\downarrow$

orthogonal projection of X in $M^2_{\mathcal{G}}$ is a

RANDOM VARIABLE in $M^2_{\mathcal{G}}$ (measurable w.r.t. $\mathcal{G}$)

(it is at MINIMAL DISTANCE)

which is the best approximation of X

among $\mathcal{G}$-measurable random variables

$\downarrow$

$\underbrace{\mathbb{E}(X \mid \mathcal{G})}$ conditional expectation of X on $\mathcal{G}$.

(the "least square" estimator of X among $\mathcal{G}$-measurable random variables).

$\mathbb{E}(X|\mathcal{G}) \in M^2_{\mathcal{G}}$ and

$$\mathbb{E}(|X-\mathbb{E}(X|\mathcal{G})|^2) = \min_{Y \in M^2_{\mathcal{G}}} \mathbb{E}(|X-Y|^2)$$

So if $X \in M^2_{\mathcal{G}} \Rightarrow \mathbb{E}(X|\mathcal{G}) = X$

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

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

$\Rightarrow \quad \mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})Z) \quad \forall Z \in M^2_{\mathcal{G}}$

$Z = c$ constant $\in M^2_{\mathcal{G}}$ since $\{\omega \mid \cancel{Z}(\omega) \le x\} = \begin{cases} \Omega & x \ge c \\ \emptyset & x < c \end{cases}$

take $c = 1$

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|\mathcal{G}))$$

Ex Let us fix $Y \in M^2$.

$\sigma(Y) = \sigma$-algebra contained in $\mathcal{F}$ generated by $\{\omega, Y(\omega) \le x\}$ $\forall x \in \mathbb{R}$

= minimal σ-algebra which makes Y measur.

$M^2_{\sigma(Y)} = \{$ random variables in M^2 which are measurable with respect to $\sigma(Y) \}$

(IT IS POSSIBLE TO PROVE)

$= \{ h(Y)$ $h : \mathbb{R} \to \mathbb{R}$ measurable such that $h(Y) : \Omega \longrightarrow \mathbb{R}$

$$\omega \xrightarrow{Y} Y(\omega) \xrightarrow{h} h(Y(\omega))$$

$$\Omega \xrightarrow{Y} \mathbb{R} \xrightarrow{h} \mathbb{R}$$

is a random variable and $\mathbb{E}((h(Y))^2) < +\infty \}$

Example take Y : tossing a coin

$$Y(\omega) = \begin{cases} 1 & \text{if } \omega \text{ is head} \\ 0 & \text{if } \omega \text{ is tail} \end{cases}$$

$$\sigma(Y) = \{ \Omega, \phi, \{\omega \mid Y(\omega)=1\}, \{\omega \mid Y(\omega)=0\}\}$$

For $Y \in M^2$ we define

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

$= \mathbb{E}(X \mid \sigma(Y))$ orthogonal projection of X in $M^2_{\sigma(Y)}$

$\mathbb{E}(X \mid Y)$ is the random variable $h(Y)$ which best approximate X.

$\mathbb{E}(X|Y) \in M^2_{\sigma(Y)}$ so $\mathbb{E}(X|Y) = h(Y)$

$h(y) = \mathbb{E}(X|Y=y)$

$\downarrow$ ...

how to compute $\mathbb{E}(X|Y)$?

$$\min_{Z \in M^2_{\sigma(Y)}} \mathbb{E}(X-Z)^2 = \min_{\substack{\{f:\mathbb{R}\to\mathbb{R}\} \\ \text{meas.}}} \mathbb{E}(X-f(Y))^2$$

$\Rightarrow$ since $Z \in M^2_{\sigma(Y)} \Leftrightarrow Z = f(Y)$ $f:\mathbb{R}\to\mathbb{R}$

$h:\mathbb{R}\to\mathbb{R}$ such that $\mathbb{E}(X-h(Y))^2 = \min\limits_{f:\mathbb{R}\to\mathbb{R}} \mathbb{E}(X-f(Y))^2$

then $h(Y) = \mathbb{E}(X|Y)$ LEAST SQUARE ESTIMATOR of X given Y.

solving the minimization problem is quite difficult!

In some case it is easy:

take X independent of $Y \Rightarrow$ so $\mathbb{E}(X \cdot f(Y)) = \mathbb{E}(X)\,\mathbb{E}(f(Y))$
$\forall f : \mathbb{R} \to \mathbb{R}$

Observe that $X - \mathbb{E}(X) \perp M^2_{\sigma(Y)}$ since

$\forall Z \in M^2_{\sigma(Y)} \quad Z = f(Y) \quad f : \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}\left((X - \mathbb{E}(X))\,f(Y)\right) = \mathbb{E}\left(X f(Y) - \mathbb{E}(X) f(Y)\right) = \left(\begin{matrix}\text{by} \\ \text{independ.}\end{matrix}\right.$$

$$= \mathbb{E}(X)\,\mathbb{E}(f(Y)) - \mathbb{E}(X)\,\mathbb{E}(f(Y)) = 0$$

Moreover $\mathbb{E}(X) \in M^2_{\sigma(Y)}$ since it is a constant

$$X = \underbrace{\mathbb{E}(X)}_{\in\, M^2_{\sigma(Y)}} + \underbrace{X - \mathbb{E}(X)}_{\in\, (M^2_{\sigma(Y)})^\perp} \Rightarrow \mathbb{E}(X \mid Y) = \mathbb{E}(X)$$

constant!

In general though computing $\mathbb{E}(X|Y)$ is difficult

$$\min_{f:\mathbb{R}\to\mathbb{R}} \mathbb{E}(X-f(Y))^2 = \mathbb{E}(X-\mathbb{E}(X|Y))^2$$

$$\mathbb{E}(X-h(Y))^2 \qquad h(Y)=\mathbb{E}(X|Y)$$

I restrict the set where I compute the minimum

I consider just $f:\mathbb{R}\to\mathbb{R}$ $\quad f(x)=ax+b$

f LINEAR.

$$V=\{\underbrace{aY+b}, a,b\in\mathbb{R}\} \subsetneq M^2\sigma(Y)$$

the orthogonal projection of X in V

is <u>not</u> $\mathbb{E}(X \mid Y)$ but it is just the

best LINEAR APPROXIMATION of X given Y

(the LINEAR FUNCTION of Y which is at MINIMAL distance from X)

$\bar{a}, \bar{b} \in \mathbb{R}$

$$\mathbb{E}\,|X - \bar{a}Y - \bar{b}|^2 = \min_{(a,b)\in\mathbb{R}^2} \mathbb{E}(|X - aY - b|^2)$$

Ex : let us find $\bar{a}, \bar{b}$

$$\min_{a,b} \mathbb{E}((X - aY - b)^2) = \min_{a,b} \big(\mathbb{E}(X^2) + a^2\mathbb{E}(Y)^2 + b^2 +$$

$$- 2a\,\mathbb{E}(XY) - 2b\,\mathbb{E}(X) + 2ab\,\mathbb{E}(Y)\big).$$

$$\phi(a,b) := \mathbb{E}(X^2) + a^2\,\mathbb{E}(Y^2) + b^2 - 2a\,\mathbb{E}(XY) - 2b\,\mathbb{E}(X) + 2ab\,\mathbb{E}(Y)$$

$\downarrow$

Compute $\min_{a,b} \phi(a,b)$

$$\begin{cases} \dfrac{\partial \phi}{\partial a}(a,b) = \cancel{2}a\,\mathbb{E}(Y^2) - \cancel{2}\,\mathbb{E}(XY) + \cancel{2}b\,\mathbb{E}(Y) = 0 \\[2ex] \dfrac{\partial \phi}{\partial b}(a,b) = \cancel{2}b - \cancel{2}\,\mathbb{E}(X) + \cancel{2}a\,\mathbb{E}(Y) = 0 \end{cases}$$

$$\begin{cases} a = \dfrac{\mathbb{E}(XY)}{\mathbb{E}(Y^2)} - \dfrac{b\,\mathbb{E}(Y)}{\mathbb{E}(Y^2)} = \dfrac{\mathbb{E}(XY)}{\mathbb{E}(Y^2)} - \dfrac{\mathbb{E}(X)\mathbb{E}(Y)}{\mathbb{E}(Y^2)} + a\dfrac{(\mathbb{E}(Y))^2}{\mathbb{E}(Y^2)} \\[2ex] b = \mathbb{E}(X) - a\,\mathbb{E}(Y) \end{cases}$$

$$\Rightarrow \begin{cases} a\left(1 - \dfrac{(\mathbb{E}(Y))^2}{\mathbb{E}(Y^2)}\right) = \dfrac{\mathrm{Cov}(X,Y)}{\mathbb{E}(Y^2)} \\ b = \mathbb{E}(X) - a\,\mathbb{E}(Y) \end{cases}$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

$$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2$$

$$\Rightarrow a\left(\frac{\mathbb{E}(Y^2) - (\mathbb{E}(Y))^2}{\cancel{\mathbb{E}(Y^2)}}\right) = \frac{\mathrm{Cov}(X,Y)}{\cancel{\mathbb{E}(Y^2)}} \to a = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}\, Y}$$

$$b = \mathbb{E}(X) - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}\, Y} \cdot \mathbb{E}(Y)$$

$L(X|Y)$ = the linear least square estimator of X given Y

$$= \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}\, Y} Y + \mathbb{E}(X) - \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}\, Y} \mathbb{E}(Y).$$

$$a = \frac{\cancel{E(Y^2)}\,[E(XY) - E(X)E(Y)]}{\cancel{E(Y^2)}\,[E(Y^2) - [E(Y)]^2]} = \frac{Cov(X,Y)}{Var(Y)}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$Var\, Y = E(Y^2) - E(Y)^2$$

$$b = E(X)E(Y^2) - E(XY)E(Y)$$