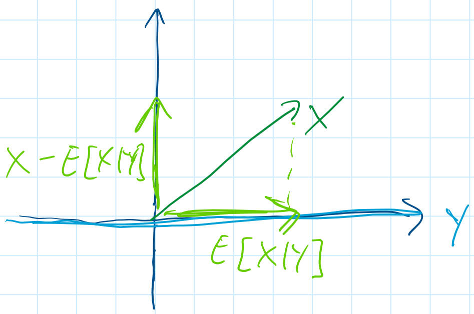
CONDITIONAL EXPECTATION given σ-algebra

Conditional expectation as "best approximation" of X given Y



approximation of vector X in vector subspace generated by Y
"is orthogonal projection" on $\langle Y \rangle$

- $E[X|Y] \in \langle Y \rangle$
- $(X - E[X|Y]) \cdot E[X|Y] = 0$

But random variables are functions $\Omega \to \mathbb{R}$ that are infinite-dimensional objects

Given $(\Omega, \mathcal{A}, P)$ probability space
Consider $\mathcal{F} \subseteq \mathcal{A}$ SUB σ-ALGEBRA
$\mathcal{F}$ represents smaller information
A random variable $X : \Omega \to \mathbb{R}$ is called
$\mathcal{F}$-MEASURABLE if $\{X \in A\} \in \mathcal{F}$ $\forall A \subseteq \mathbb{R}$ interval

Given $X : \Omega \to \mathbb{R}$ $\mathcal{A}$-measurable
We want to construct the best approximation of X as a $\mathcal{F}$-measurable random variable
(Knowing only the information carried by $\mathcal{F}$)

Def $(\Omega, \mathcal{A}, P)$, $X \in L^1(\Omega, \mathcal{A}, P)$ X is $\mathcal{A}$-meas.
$\mathcal{F} \subseteq \mathcal{A}$ sub σ-algebra.
The CONDITIONAL EXPECTATION of X given $\mathcal{F}$ is a random variable, denoted $E[X|\mathcal{F}]$, such that
i) $E[X|\mathcal{F}]$ is $\mathcal{F}$-measurable,

i) $E[X|\mathcal{F}]$ is $\mathcal{F}$-measurable,

ii) $E[X|\mathcal{F}] \in L^1(\Omega, \mathcal{F}, P)$ $\qquad E[X|\mathcal{F}] : \Omega \to \mathbb{R}$

iii) $E[X \mathbb{1}_A] = E[\ E[X|\mathcal{F}]\ \mathbb{1}_A] \quad \forall A \in \mathcal{F}$.

Properties

1) $E[X|\mathcal{F}]$ exists and is unique almost surely.

i.e. if Y and Y' are both $E[X|\mathcal{F}]$ then

$$P(Y = Y') = 1$$

2) if $X \in L^2(\Omega, \mathcal{A}, P)$.

L^2 has a SCALAR PRODUCT

$X, Y \in L^2(\Omega) \longrightarrow \langle X, Y \rangle = E[XY]$

Consider $L^2(\Omega, \mathcal{F}, P)$ subspace of $L^2(\Omega, \mathcal{A}, P)$

$\hookrightarrow$ space of $\mathcal{F}$-measurable square integrable r.v.

Conditional expectation $E[X|\mathcal{F}]$ is the ORTHOGONAL PROJECTION of X on the subspace $L^2(\Omega,\mathcal{F},P)$

that is, $E[X|\mathcal{F}] \in L^2(\Omega, \mathcal{F}, P)$ and

$X - E[X|\mathcal{F}]$ is orthogonal to $L^2(\Omega, \mathcal{F}, P)$, i.e.

$\langle X - E[X|\mathcal{F}], Y \rangle = 0 \quad \forall Y \in L^2(\Omega, \mathcal{F}, P)$

3) if $\mathcal{F} = \sigma(Y)$ then $E[X|\mathcal{F}] = E[X|Y]$

4) LINEARITY: $X, Y \in L^1(\Omega)$, $a, b \in \mathbb{R}$

$E[aX + bY|\mathcal{F}] = a\,E[X|\mathcal{F}] + b\,E[Y|\mathcal{F}]$

5) MONOTONICITY: $P(X \leq Y) = 1$ then

$E[X|\mathcal{F}] \leq E[Y|\mathcal{F}]$ almost surely

6) $|E[X|\mathcal{F}]| \leq E[|X|\,|\mathcal{F}]$

7) product $XY \in L^1(\Omega)$ and Y is $\mathcal{F}$-measurable then
$E[XY|\mathcal{F}] = Y\,E[X|\mathcal{F}]$

8) TOWER PROPERTY : $\mathcal{G} \subseteq \mathcal{F} \subseteq \mathcal{A}$
$E[\,E[X|\mathcal{F}]\,|\,\mathcal{G}] = E[X|\mathcal{G}]$

9) if $\mathcal{G} = \{\emptyset, \Omega\}$ then $E[X|\mathcal{G}] = E[X]$
As a consequence
$E[\,E[X|\mathcal{F}]\,] = E[X]$

10) if X and $\mathcal{F}$ are independent
(that is, $\sigma(X)$ and $\mathcal{F}$ are independent) then
$E[X|\mathcal{F}] = E[X]$

11) $X \in L^p(\Omega)$ $p \geq 1$ then $E[X|\mathcal{F}] \in L^p(\Omega)$

12) JENSEN's INEQUALITY
$E[\varphi(X)|\mathcal{F}] \geq \varphi(E[X|\mathcal{F}])$ $\forall$ φ CONVEX

MARTINGALES

$(X_n)_{n\geq 1}$ discrete time stochastic process
We can define $E[X|Y]$ where Y is a random vector
$\rightarrow E[X|Y_1,\dots,Y_n] = h(Y_1,\dots,Y_n)$

Def A stochastic process $(X_n)_{n\geq 0}$ is a MARTINGALE
(for its natural filtration) if

- $X_n \in L^1(\Omega)$ $\forall n$
- $E[X_{n+1}|X_n, X_{n-1},\dots,X_0] = X_n$ a.s.

- $X_n \in L^1(\Omega) \quad \forall n$
- $E[X_{n+1} \mid X_n, X_{n-1}, \dots, X_0] = X_n$ a.s.

Examples

1) RANDOM WALK (symmetric)

let $(Z_n)_{n\geq 1}$ independent sequence of r.v. with $E[Z_n]=0$

$X_n = Z_1 + \dots + Z_n \quad , X_0 = 0$

$\Rightarrow \quad X_{n+1} = X_n + Z_{n+1}$

Note Z_{n+1} is independent of X_n

$(X_n)_{n\geq 0}$ is a MARTINGALE

$E[X_{n+1} \mid X_n, \dots, X_1] \overset{?}{=} X_n$

$= E[X_n + Z_{n+1} \mid X_n, \dots, X_1]$

$\overset{\text{LINEARITY}}{=} \underset{\overset{\|}{X_n}}{E[X_n \mid X_n, \dots, X_1]} + \underbrace{E[Z_{n+1} \mid X_n, \dots, X_1]}_{\overset{\|\ (10) - \text{INDEPENDENCE}}{E[Z_{n+1}]}}$

$= X_n + \overbrace{E[Z_{n+1}]}^{=0} = X_n$

2) $(Z_n)_{n\geq 0}$ independent with $E[Z_n] = 1 \quad \forall n$

Let $X_n = Z_0 \cdot Z_1 \cdot \dots \cdot Z_n = \prod_{j=0}^{n} Z_j$

$(X_n)_{\geq 0}$ is a martingale $\qquad X_{n+1} = X_n \cdot Z_{n+1}$

$E[X_{n+1} \mid X_n, \dots, X_0] = E[X_n Z_{n+1} \mid X_n, \dots, X_0]$

X_n is $(X_1, \dots, X_n)$-measurable

$\overset{(7)}{=} X_n \, E[Z_{n+1} \mid X_n, \dots, X_0]$

$\overset{(10)}{=} X_n \, E[Z_{n+1}] = X_n \cdot 1$

$Z_{n+1} \perp\!\!\!\perp (X_n, \dots, X_0)$

Recall that if $\mathcal{F} = \sigma(Y_1, \dots, Y_n)$ then

X is $\sigma(Y_1, \dots, Y_n)$ measurable if and only if

X is $\sigma(Y_1,..,Y_n)$ measurable if and only if $X = f(Y_1,..,Y_n)$ X is a function of $(X_1,..,X_n)$

- FILTRATION

Given $(\Omega, \mathcal{A}, P)$

filtration is increasing source of information with time

$\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}$ is a FILTRATION if
$\mathcal{F}_n$ is a σ-algebra, $\mathcal{F}_n \subseteq \mathcal{A}$ and $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$

We say that a stochastic process $X = (X_n)_{n \geq 0}$ is ADAPTED to a filtration $\mathbb{F} = (\mathcal{F}_n)_n$ if
X_n is $\mathcal{F}_n$-measurable

The filtration generated by a process $X = (X_n)_{n \geq 0}$, called NATURAL FILTRATION is $\mathbb{F}^X = (\mathcal{F}_n^X)_{n \geq 0}$
$\mathcal{F}_n^X = \sigma(X_0, X_1, .., X_n)$

<u>Def</u> A process $(X_n)_{n \geq 1}$ is a MARTINGALE for the filtration $\mathbb{F}$ (denoted $\mathbb{F}$-martingale) if

- $X_n \in L^1$ $\forall n$
- X is adapted to $\mathbb{F}$
- $E[X_{n+1} \mid \mathcal{F}_n] = X_n$ a.s.

<u>Note</u> that if $\mathbb{F} = \mathbb{F}^X$ then get same definition as before

X is SUBMARTINGALE if $E[X_{n+1} \mid \mathcal{F}_n] \geq X_n$ a.s.
X is SUPERMARTINGALE if $E[X_{n+1} \mid \mathcal{F}_n] \leq X_n$ a.s.

Properties

- X SUPERMART $\Rightarrow -X$ is SUBMART
- X, Y MART $\Rightarrow aX + bY$ MART $\forall\, a, b \in \mathbb{R}$
- X, Y SUPERMART $\Rightarrow aX + bY$ SUPERMART if $a, b \geq 0$
- X, Y SUPERMART $\Rightarrow Z = \min(X, Y)$ SUPERMART

  $\hookrightarrow Z_n = \min(X_n, Y_n)$
- X, Y SUB MART $\Rightarrow Z = \max(X, Y)$ SUBMART
- X MARTINGALE $\Rightarrow E[X_n] = E[X_0] \quad \forall n$

  Indeed, $E[X_{n+1}] \underset{(9)}{=} E[E[X_{n+1} \mid \mathcal{F}_n]] = E[X_n]$
- X SUBMARTINGALE $\Rightarrow E[X_{n+1}] \geq E[X_n]$
- X SUPERMARTINGALE $\Rightarrow E[X_{n+1}] \leq E[X_n]$

Example $(\Omega, \mathcal{A}, P)$ $\quad Y$ $\mathcal{A}$-meas. r.v. $Y \in L^1(\Omega)$

given a filtration $\mathbb{F} = (\mathcal{F}_n)_{n \geq 0}$

Let $X_n = E[Y \mid \mathcal{F}_n]$

Then $(X_n)_{n \geq 0}$ is a MARTINGALE

$$E[X_{n+1} \mid \mathcal{F}_n] = E[E[Y \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n]$$

TOWER PROPERTY $\mathcal{F}_{n+1} \supseteq \mathcal{F}_n$ $\quad = E[Y \mid \mathcal{F}_n] = X_n$

Theorem CONVERGENCE THEOREM for MARTINGALES

if $(X_n)_{n \geq 0}$ is a SUBMARTINGALE and $\sup_n E[X_n^+] < \infty$

then there exists $X \in L^1(\Omega)$ such that

$X_n \xrightarrow[n \to \infty]{} X$ almost surely.

Remark

- Condition $\sup E[X_n^+] < \infty$ means that $E[X_n^+] \leq C$

Remark

- Condition $\sup_n E[X_n^+] < \infty$ means that $E[X_n^+] \leq C$ for a constant C independent of n.
- if $X_n \leq 0 \ \forall n$ then condition is satisfied, as $E[X_n^+] = 0$
- if $(X_n)_{n \geq 0}$ is SUPERMARTINGALE and $\sup_n E[X_n^-] < \infty$ then $\exists X \in L^1(\Omega)$ s.t. $X_n \xrightarrow[n \to \infty]{} X$ a.s.
  In particular, this is true if $X_n \geq 0 \ \forall n$
- Note that a martingale is both a submartingale and a supermartingale.
- For a MARKOV CHAIN, the ergodic theorem establishes convergence in distribution $\mu_{X_n} \xrightarrow[n \to \infty]{} \pi$ to the stationary distribution
  → The result for martingales is much stronger.

→ A MARKOV CHAIN might not be a martingale (for example, the non-symmetric random walk)

$$E[X_{n+1} \mid X_1, \dots, X_n] \underset{\text{MARKOV PROPERTY}}{=} E[X_{n+1} \mid X_n] = h(X_n)$$

MARTINGALE property says that $h(X_n) = X_n$

- non-symmetric random walk: $X_{n+1} = X_n + Z_{n+1}$ with $E[Z_{n+1}] \neq 0$
  then $E[X_{n+1} \mid X_n] = X_n + E[Z_{n+1}] \neq X_n$

→ A MARTINGALE might not be a MARKOV CHAIN