Lecture 8 cond exp

Monday, October 2, 2017 10:11 AM

Previous lecture

μ measure

$\mu \prec \lambda$ $\quad \lambda(A)=0 \Rightarrow \mu(A)=0$

$$\mu(A) = \int_A f \, d\lambda$$

Radon - Nikodým derivative

Lebesgue integrals

X_n

simple

$\sigma(A_1, A_2, \dots, A_n)$

disjoint subsets partitioning Ω

$X_n \uparrow X$ as $n \to \infty$, X_n simple

$X \geq 0$

$$\int_A X \, dP = \lim_{n \to \infty} \int_A X_n \, dP$$

Extended to arbitrary X

$$\mathbb{E}(X; A) = \int_A X \, dP$$

$$X = X^+ - X^-$$

$X^\pm$ $\max(0, \pm X)$

DF $\mathcal{F}$ is called a σ-algebra generated by r.v. X if it consists of $X^{-1}(B)$, $B \in \mathcal{B}$ Borel subsets

$\exists\ P(B)>0$

$$\mathbb{E}(X \mid B) := \frac{\mathbb{E}(X;B)}{P(B)} = $$

$$= \underbrace{\frac{\mathbb{E}(X\cdot I_B)}{\mathbb{E}(I_B)}} = \frac{\int_B X\,dP}{\int_B dP}$$

My notation $\mathbb{E}(X \,\|\, I_B)$

f, g random functions

Relative expectation $\mathbb{E}(f \,\|\, g) := \frac{\mathbb{E}(f\cdot g)}{\mathbb{E}(g)}$

provided expectations exist

---

r.v. $X \perp Y$

deterministic function

makes sense $\mathbb{E}(f(X,Y) \mid Y=y) = \varphi(y)$

When Y is continuous $\Rightarrow P(Y=y)=0$

$$\mathbb{E}(f(X,Y) \mid Y=y)\Big|_{y=Y} = \varphi(Y) \text{ r.v.}$$

Want to define general conditional expectations wrt a σ-algebra or a random variable.

$\exists\ Y$ is measurable wrt σ-algebra $\mathcal{U}$

$\varphi(Y)$ is measurable wrt $\mathcal{U}$

DF Notation:

$\mathbb{E}(X \mid \mathcal{U})$ or $\mathbb{E}(X \mid Y)$

meaning $\mathbb{E}(X \mid \mathcal{U})$, where $\mathcal{U}$ is a σ-algebra generated by Y

DF Conditional expectation

$(\Omega, \mathcal{F}, P)$

$\exists\ \mathcal{U} \subset \mathcal{F}$

↳ another σ-algebra

Then

r.v. $Y = \mathbb{E}(X | \mathcal{U})$ when

1) Y is $\mathcal{U}$-measurable

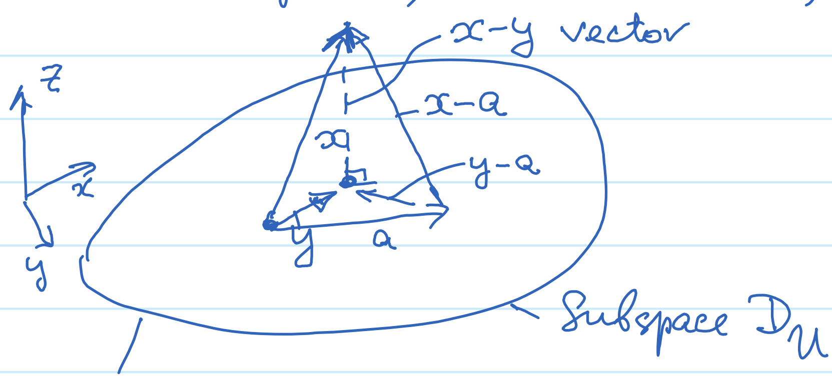
2) $\mathbb{E}(Y; A) = \mathbb{E}(X; A)$, $\forall A \in \mathcal{U}$

Phylosophical intuition

$\mathbb{E}(\cdot | \mathcal{U})$ is an operator that "eats" all the "randomness" in $\mathcal{F}$ except the protected "randomness" $\mathcal{U}$

Projection intuition

Vector algebra, Linear algebra, Analytic geometry



Projection of x on $D_{\mathcal{U}}$:= y vector

$$y = \arg\min_{a \in \mathcal{U}} (x-a)^T \cdot (x-a)$$

$\forall\ x, y$ vectors Scalar product

$$\langle x, y \rangle := x^T \cdot y$$

$$d(x,y) = \sqrt{(x-y)^T (x-y)}$$

$\mathbb{E}(X)$

$$\arg\min_a \mathbb{E}(X-a)^2 = \arg\min_a \mathbb{E}X^2 - 2\mathbb{E}X \cdot a + a^2$$

$$a_{min} = \frac{-(-2\mathbb{E}X)}{2\cdot 1} = \mathbb{E}X$$

$X \in L_2$, a space of functions $X(\omega)$: $\mathbb{E}X^2 < \infty$

$X, Y \qquad \langle X, Y \rangle = \mathbb{E}(X \cdot Y)$

$$\arg\min_a \mathbb{E}(X-a)^2 = \arg\min_a \langle X-a, X-a \rangle =$$

$$= \arg\min_a d^2(X, a)$$

$D_{\mathcal{U}}$ will be the space of r.v. measurable wrt $\mathcal{U}$

Consider a "discrete" σ-algebra $\mathcal{U}$ as an example that will help develop the intuition

$X \in L_2$

Y – discrete

$\Omega = \bigcup_j A_j$ – disjoint

$$\mathcal{U} = \sigma(A_1, A_2, \dots)$$

Z measurable wrt $\mathcal{U} \Rightarrow$

$$Z(\omega) = \sum_k z_k \cdot I_{A_k}$$

Set of functions

$\{ I_{A_k} \}$ can be considered basis functions

$Z(\omega)$ is a linear form

The following condition

$$\underset{\in L_2}{X - Y} \perp Y - a, \quad \forall a \in D_{\mathcal{U}}$$

defines the projection of X on $D_{\mathcal{U}}$ in the sense of minimal distance from X to $D_{\mathcal{U}}$

Note: This condition is equivalent to $X - Y \perp a$ $\forall a \in D_{\mathcal{U}}$

Proof:

We have:

$$\mathbb{E}\left((X-a)^2\right) = \mathbb{E}\left(X - Y + Y - a\right)^2 \overset{X-Y \perp Y-a}{=}$$

$$= \mathbb{E}\left((X-Y)^2\right) + \mathbb{E}\left((Y-a)^2\right) \underbrace{\geq}_{= \text{ iff. } a = Y} \mathbb{E}\left((X-Y)^2\right)$$

$$d^2(X, a) = d^2(X, Y) + d^2(Y, a)$$

We have just shown that

$a = Y$ minimizes $\mathbb{E}\left((X-a)^2\right) \iff X - Y \perp Y - a$ or $X - Y \perp D_{\mathcal{U}}$

So that the condition $X - Y \perp Y - a, \ \forall a \in D_{\mathcal{U}}$ defines Y as a projection of X onto $D_{\mathcal{U}}$ just as the condition $Y = \underset{a \in D_{\mathcal{U}}}{\arg\min}\, d^2(X, a)$ $\square$

Because $Y \in D \Rightarrow Y = \sum y \cdot I$

$a \in D_{\mathcal{U}}$ $\square$

Because $Y \in D_{\mathcal{U}} \Rightarrow Y = \sum_k y_k \cdot I_{A_k}$

$\uparrow$ are some coefficients

This condition $X - Y \perp a$, $\forall a \in D_{\mathcal{U}}$ uniquely defines the Y and its coefficients y_k

Let's find them.

On the one hand $\mathbb{E}(Y; A_k) = \sum_j y_j \underbrace{I_{A_j} \cdot I_{A_k}}_{=0\ j \neq k} = y_k \cdot \mathbb{P}(A_k)$

So that

$$y_k = \frac{\mathbb{E}(Y; A_k)}{P(A_k)} := \mathbb{E}(Y \mid A_k), \quad P(A_k) > 0$$

and

$$Y = \sum_k \underbrace{\mathbb{E}(Y \mid A_k)}_{y_k} \cdot I(A_k)$$

Now $X - Y \perp D_{\mathcal{U}} \iff$

$$\mathbb{E}\left((X - Y) \cdot I_{A_k}\right) = 0$$

defines a system of linear eqns for y_k

$$\mathbb{E}(X - Y; A_k) = \mathbb{E}(X; A_k) - \underbrace{\mathbb{E}(Y; A_k)}_{y_k P(A_k)} = 0$$

$\Rightarrow$

$$\mathbb{E}(X; A_k) = y_k \cdot \mathbb{P}(A_k) \Rightarrow$$

$$\Rightarrow y_k = \frac{\mathbb{E}(X; A_k)}{P(A_k)}$$

Also, $\forall A \in \mathcal{U}$

$$\mathbb{E}(X; A) = \sum \mathbb{E}(X; A_k) =$$

$$= \sum_{k}^{k} y_k P(A_k) = \mathbb{E}(Y; A)$$

Summarizing:

Properties of the above construction

1) Y is $\mathcal{U}$-measurable (i.e. $\in D_{\mathcal{U}}$)

2) $\mathbb{E}(Y; A) = \mathbb{E}(X; A)$

These are the properties that went into the definition of $\mathbb{E}(X|\mathcal{U})$ that is general and does not rely on the discrete assumption for the Y.