Lecture 7: Geometric Viewpoint

Definition - For a given probability space $(\Omega, \mathcal{F}, P)$ the function space $L^p(\Omega, \mathcal{F}, P)$ (or L^p for short) consits of all random variables defined on $(\Omega, \mathcal{F}, P)$ for which

$$\mathbb{E}[|X|^p] < \infty, \quad (0 \leq p < \infty)$$

Example:

X ~ has the following density:

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

$\Rightarrow f \notin L^1$ since

$$\mathbb{E}[|X|] = \int_{-\infty}^{\infty} \frac{|x|}{1+x^2}\,dx$$

$$= 2\int_0^{\infty} \frac{x}{1+x^2}\,dx$$

$$= \int_0^{\infty} \frac{1}{1+u}\,du$$

$$= \ln(1+u)\Big|_0^{\infty}$$

$$= \infty.$$

* $L^2(\Omega, \mathcal{F}, P)$ is often called the set of square integrable random variables, or random variables with finite variance.

Theorem - L^2 is a linear subspace of random variables.

proof:

1. Let $X, Y \in L^2$. We need to show $X+Y \in L^2$, that is
$$\mathbb{E}[(X+Y)^2] < \infty.$$
We first prove a Lemma.

Lemma - $(X+Y)^2 \leq X^2 + Y^2$

proof
$$(X-Y)^2 \geq 0$$
$$\Rightarrow X^2 - 2XY + Y^2 \geq 0$$
$$\Rightarrow X^2 + Y^2 \geq 2XY$$
$$\Rightarrow 2X^2 + 2Y^2 \geq X^2 + 2XY + Y^2$$
$$\Rightarrow (X+Y)^2 \leq 2X^2 + 2Y^3.$$

Applying the lemma we have that
$$\mathbb{E}[(X+Y)^2] \leq \mathbb{E}[2X^3 + 2Y^2]$$
$$= 2\mathbb{E}[X^2] + 2\mathbb{E}[Y^2]$$
$$< \infty.$$

2. Let $K \in \mathbb{R}$. Then,
$$\mathbb{E}[K^2X^2] = K^2\mathbb{E}[X^2] < \infty$$
and thus $KX \in L^2$.

Definition - L^2 has a norm defined by:

$$\|X\| = \mathbb{E}[X^2]^{1/2},$$

for all $X \in L^2$. The distance between two random variables $X, Y \in L^2$ is defined by

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

(length of their difference).

The norm satisfies a triangle inequality

$$\|X+Y\| \leq \|X\| + \|Y\|.$$

Theorem - L^2 is an inner product space with inner product

$$\langle X, Y \rangle = \mathbb{E}[XY]$$

proof:

1. Well defined: We need to show that $|\mathbb{E}[XY]| < \infty$. We first prove a lemma:

   Lemma - $|XY| \leq \frac{1}{2}X^2 + \frac{1}{2}Y^2$.

   proof:

   $$(|X| - |Y|)^2 = X^2 - 2|X||Y| + Y^2 \geq 0$$

   $$\Rightarrow X^2 + Y^2 \geq 2|X| \cdot |Y|.$$

   Therefore,

   $$|\mathbb{E}[XY]| \leq \mathbb{E}[|X \cdot Y|] \leq \mathbb{E}[\tfrac{1}{2}X^2 + \tfrac{1}{2}Y^2] < \infty.$$

2. Symmetric: $\mathbb{E}[XY] = \mathbb{E}[YX]$.
3. Linear: $\mathbb{E}[(aX+bY)Z] = a\mathbb{E}[XZ] + b\mathbb{E}[YZ]$.
4. Positive definite: $\mathbb{E}[X^2] = \langle X, X \rangle \geq 0$ and equals 0 if and only if $X = 0$.

Theorem - If $X, Y \in L^2$ then

$$|\mathbb{E}[XY]| \leq \mathbb{E}[X^2]^{1/2}\,\mathbb{E}[Y^2]^{1/2} = \|X\|\cdot\|Y\|.$$

proof:

Let $t \in \mathbb{R}$. Then

$$0 \leq \mathbb{E}[(X-tY)^2] = \mathbb{E}[X^2] - 2t\,\mathbb{E}[XY] + t^2\,\mathbb{E}[Y^2]$$

We make the right hand side as small as possible when t satisfies

$$-2\,\mathbb{E}[XY] + 2t\,\mathbb{E}[Y^2] = 0$$

$$\Rightarrow t = \frac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]}$$

Therefore,

$$0 \leq \mathbb{E}[X^2] - 2\frac{\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]} + \frac{\mathbb{E}[XY]^2}{\mathbb{E}[Y^2]}$$

$$\Rightarrow \mathbb{E}[XY]^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[Y^2]$$

$$\Rightarrow |\mathbb{E}[XY]| \leq \mathbb{E}[X^2]^{1/2}\,\mathbb{E}[Y^2]^{1/2}.$$

Consequences

1. $|Cov(X,Y)| \leq \sqrt{Var(X)}\sqrt{Var(Y)}$

proof:

$$|Cov(X,Y)| = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$$
$$\leq \mathbb{E}[(X-\mathbb{E}[X])^2]^{1/2}\,\mathbb{E}[(Y-\mathbb{E}[Y])^2]^{1/2}$$
$$= \sqrt{Var(X)}\sqrt{Var(Y)}.$$

2. The correlation coefficient

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

satisfies: $-1 \leq \rho(X,Y) \leq 1$.

3. There exists $\theta \in [0, \pi]$ such that

$$\mathbb{E}[XY] = \cos\theta \, \|X\| \cdot \|Y\|.$$

proof:

By Cauchy-Schwarz inequality

$$|\mathbb{E}[XY]| \le \|X\|^{1/2} \cdot \|Y\|^{1/2}.$$

$$\Rightarrow -1 \le \frac{\mathbb{E}[XY]}{\|X\| \cdot \|Y\|} \le 1$$

There exists $\theta \in [0, \pi]$ such that

$$\cos\theta = \frac{\mathbb{E}[XY]}{\|X\| \cdot \|Y\|}$$

$$\Rightarrow \mathbb{E}[XY] = \cos\theta \, \|X\| \cdot \|Y\|.$$

4. The angle between random variables $X, Y \in L^2$ is given by

$$\cos\theta = \frac{\mathbb{E}[XY]}{\|X\| \, \|Y\|}.$$

5. If X, Y are uncorrelated, then $\theta = \pi/2$

proof:

If X, Y are uncorrelated then $\mathbb{E}[XY] = 0$ implying that $\theta = \pi/2$.

Definition - We say a sequence of random variables X_n converges in L^2 to some X in L^2 if

$$\lim_{n\to\infty} \|X_n - X\| = 0.$$

Theorem (Weak Law of Large Numbers) - Let $X_i \in L^2$ be I.I.D. $\mathbb{E}[X_i] = \mu$, $\mathbb{E}[X_i^2] = \sigma^2$, $\mathbb{E}[X_i X_j] = \mu^2$. Then the empirical mean

$$\frac{1}{n} S_n = \frac{1}{n}(X_1 + \dots + X_n)$$

converges to μ in L^2.

proof:

$$\|S_n - \mu\|^2 = \left\| \frac{1}{n}(X_1 + \dots + X_n) - \mu \right\|^2$$

$$= \frac{1}{n^2} \|(X_1 - \mu) + (X_2 - \mu) + \dots + (X_n - \mu)\|^2$$

$$= \frac{1}{n^2} \mathbb{E}[[(X_1-\mu) + (X_2-\mu) + \dots + (X_n-\mu)][(X_1-\mu) + \dots + (X_n-\mu)]]$$

$$= \frac{1}{n^2}\left(\mathbb{E}[(X_1-\mu)^2] + \dots + \mathbb{E}[(X_n-\mu)^2]\right)$$

$$= \frac{\sigma^2}{n}.$$

Therefore,

$$\lim_{n\to\infty} S_n = 0.$$

Corollary:

$$\|S_n - \mu\| = \frac{\sigma}{\sqrt{n}} \quad \text{(slow convergence).}$$

Projections:

If $X, Y \in L^2$, what is the closest random variable to X lying in span$\{Y\}$.

If $Z_1 \in$ span$\{Y\}$ then there exists $t \in \mathbb{R}$ such that

$$Z_1 = tY.$$

We want to minimize:

$$\| X - tY \|$$

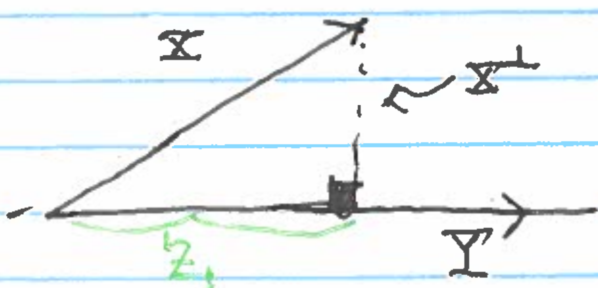
which is equivalent to minimizing

$$f(t) = \| X - tY \|^2 = \mathbb{E}[(X - tY)^2]$$
$$= \mathbb{E}[X^2] - 2t\,\mathbb{E}[XY] + t^2\,\mathbb{E}[Y^2]$$
$$\Rightarrow f'(t) = 2t\,\mathbb{E}[Y^2] - 2\,\mathbb{E}[XY]$$
$$\Rightarrow t = \frac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]}.$$

Therefore,

$$Z_1 = \frac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]} Y = \text{Proj}_Y(X).$$

is called the orthogonal projection onto Y.



$$X = X^{\perp} + Z_1$$
$$\Rightarrow X^{\perp} = X - \frac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]} Y.$$