Lesson 8 — Supplement

Expectation As ~~an~~ Inner Product & Norm

Def (. Vector Space): A vector space is a set $\mathcal{V}$ over a "field" $\mathbb{F}$ (in this discussion, $\mathbb{F}=\mathbb{R}$) with two operations

$\oplus_{\mathcal{V}} : \mathcal{V} \times \mathcal{V} \longrightarrow \mathcal{V}$ and

$\odot_{\mathcal{V}} : \mathbb{F} \times \mathcal{V} \longrightarrow \mathcal{V}$

such that:

If $u, v, w \in V$

1) $u \oplus_V v = v \oplus_V u$ Commutativity

2) $(u \oplus_V v) \oplus_V w = u \oplus_V (v \oplus_V w)$

Associativity

3) $\exists\, 0_V$ s.t.

$u \oplus_V 0_V = u$ (Existence of identity element)

4) $\forall v \in V, \exists (-v) \in V$ s.t

$v \oplus_V (-v) = 0_V$

5) $(\alpha \odot_F \beta) \odot_V u = \alpha \odot_V (\beta \odot_V u)$

$\alpha, \beta \in F (= R)$

(where $\odot_F$ is the scalar multiplication)

6) $(\alpha \oplus_{\mathbb{F}} \beta) \odot_V u = \alpha \odot_V u \oplus_V \beta \odot_V u$

$\forall \alpha, \beta \in \mathbb{F}$

($\oplus_{\mathbb{F}}$ is the scalar sum. It disappears in (6) and $\oplus_V$ appears) (Distributivity of $\oplus_{\mathbb{F}}$ over $\odot_V$)

7) $\alpha \odot_V (u \oplus_V v) = \alpha \odot_V v \oplus_V \alpha \odot_V u$

$\forall \alpha \in \mathbb{F}$ (Distributivity of $\odot_V$ over $\oplus_V$)

8) $1_{\mathbb{F}} \odot_V u = u$

Theorem: The set of all random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ is a vector space with regular sum of random variables and regular scaler product (aX, where $a \in \mathbb{R}$) $\mathbb{F} = \mathbb{R}$

Theorem: The set of all random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with finite variance is a vector space with regular sum and scalar product.

Inner Product:

Def: Let V be a vector space over the field $\mathbb{F}=\mathbb{R}$.

$\langle \cdot,\cdot \rangle : V \times V \longrightarrow \mathbb{R}$ is called a real-valued inner product

iff:

1) $\langle x, x \rangle \geq 0 \quad \forall x \in V$

and $\langle x, x \rangle = 0 \iff x = 0_V$.

(positive definiteness for inner products)

2) $\langle y, x \rangle = \langle x, y \rangle \quad \forall x, y \in V$

3) $\langle x, \alpha y_1 + \beta y_2 \rangle = \alpha \langle x, y_1 \rangle + \beta \langle x, y_2 \rangle \quad \forall \alpha, \beta \in \mathbb{R}$

Examples: $V = \mathbb{R}^n$

$\langle x, y \rangle = x^T y$ is the Euclidean inner product

Def (Norm) Let $(\mathcal{V}, \mathbb{F})$ be a vector space. We call $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ a vector norm iff:

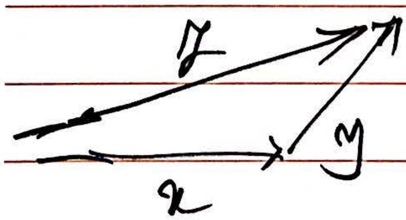
1) $\|x\| \geq 0 \quad \forall x \in \mathcal{V}$

and $\|x\| = 0 \Leftrightarrow x = 0_{\mathcal{V}}$

2) $\|\alpha \odot_{\mathcal{V}} x\| = |\alpha| \, \|x\| \quad \forall x \in \mathcal{V}$
$\alpha \in \mathbb{F} = \mathbb{R}$

3. $\|x \oplus_{\mathcal{V}} y\| \leq \|x\| + \|y\|$

Triangle Inequality



Intuitively, the concept of a vector norm is a generalization of length in Euclidean Space.

$x \in \mathbb{R}^n$

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

p-norm:

$x \in \mathbb{R}^n$

$$\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}$$

Norm induced by an inner product

Let $\langle \cdot , \cdot \rangle$ be an inner product on V. Then $\sqrt{\langle v, v \rangle}$ is a norm on V and $\|v\| = \sqrt{\langle v, v \rangle}$ is called the norm induced by the inner product $\langle \cdot , \cdot \rangle$.

Lemma (Cauchy-Schwartz-Bunyakovsky)

For a vector space V with inner product $\langle \cdot , \cdot \rangle$

$$|\langle u,v\rangle|^2 \leq \langle u,u\rangle \cdot \langle v,v\rangle$$

or

$$|\langle u,v\rangle| \leq \|u\| \, \|v\|$$

u, v linearly dependent

Equality occurs when $u = av$, $a \in \mathbb{R}$

Proof: $0 \leq \left\| u - \frac{\langle u,v\rangle}{\|v\|^2} v \right\|^2$

One can see that on the vector space of random variables with finite variance, the following inner product can be defined:

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

The norm induced by the inner product is

$$\|X\| = \sqrt{\mathbb{E}[X^2]}$$

(Show that they are indeed a norm and an inner product)

Also, the Cauchy-Schwartz-Bunyakovsky inequality holds in this vector space, i.e.

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$$

or

$$\left(\mathbb{E}[XY]\right)^2 \leq \mathbb{E}[X^2]\,\mathbb{E}[Y^2]$$

Exercise: (Rewrite the proof using $\mathbb{E}[XY]$ as the inner product)

Remark: Correlation Coefficient

is a "measure" of (linear)

dependence of X and Y.

In a Euclidean space,

such a measure is called

the directional cosine

between two vectors

$$\cos(\varphi) = \frac{\langle u, v \rangle}{\|u\|\|v\|}$$

$\cos\varphi = 1$ when $u = av$, $a \geq 0$

(i.e. u and v are collinear

and have the same direction)

and $\cos\varphi = -1$ iff u and v

are collinear and have

opposite directions.

Similarly, we saw that

$\rho(X, Y) = 1$ iff $\tilde{Y} = a\tilde{X}$, $a > 0$

and

$\rho(X, Y) = -1$ iff $\tilde{Y} = a\tilde{X}$, $a < 0$

Remark: Our previous discussion shows the relationship between linear algebra and probability theory (obviously, they are related in many other ways)

The set of all random variables can be viewed as a vector space. If one is interested in defining "length" and "direction" for random variables, one needs to

consider the subspace of

a random variables with finite variance.

(subset of a vector space that

is a vector space itself is

called a subspace)

The square root of the second

moment of a random variable

in this vector space can be

seen as its length and

$E[XY]$ can be seen as

the inner product between X and Y. Also, $\frac{E[XY]}{\sqrt{E[X^2]}\sqrt{E[Y^2]}}$

$= \frac{\langle X, Y \rangle}{\|X\| \|Y\|}$ can be viewed as

a measure of collinearity

between X and Y.

$\mathbb{E}[XY]$ is sometimes called the correlation between X and Y, $\mathrm{Corr}(X,Y)$

Obviously, $\mathrm{Cor}(X,Y) = \mathrm{Corr}(\tilde{X},\tilde{Y})$.

Orthogonality

In vector spaces that are equipped with an inner product, orthogonality is defined as having "zero collinearity".

This is an extension of the concept of orthogonality in the Euclidean Space.

Def: Assume V is a vector space equipped with inner product $\langle u,v\rangle$.

u,v are called "orthogonal" iff $\langle u,v\rangle = 0$.

Therefore, the concept of having zero correlation is analogous to the concept of orthognality.

Def: Assume that X, Y are $\mathcal{L}^2$ r.v.'s on $(\Omega, \mathcal{F}, \mathbb{P})$. They are

called orthogonal if $\mathbb{E}[XY]=0$ (they have zero correlation)

Remark: By $\mathcal{L}^r$ r.v.'s, we mean r.v.'s whose r^{th} moments are finite.

Remark: All of the above discussions about norm, inner product, Cauchy-Schwartz inequality etc can be stated for random variables or "centralized random variables." In the latter case, $\|\tilde{X}\| = \sqrt{\mathrm{Var}(X)} = \sigma_X$

$\langle \tilde{X}, \tilde{Y} \rangle = \mathrm{Cov}(X, Y)$,

$$\frac{\langle \tilde{X}, \tilde{Y} \rangle}{\|\tilde{X}\| \, \|\tilde{Y}\|} = \rho(X, Y)$$