Correlation

$$\mathbb{E}[XY] = \sum_{x_i}\sum_{y_j} x_i y_j p_{X,Y}(x_i, y_j)$$ (integral for continuous rv)

Covariance

$$\mathrm{Cov}(X,Y) \triangleq \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right]$$
$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

If $\mathbb{E}[XY] = 0$, we say X is orthogonal to Y.
If $\mathrm{Cov}(X,Y) = 0$, we say X is uncorrelated to Y.

Properties

Independence of r.v's X, Y

ONE DIRECTIONAL STATEMENT

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \Rightarrow \text{uncorrelatedness}$$

X, Y, Z r.v.'s

$$\mathrm{Cov}(X, aY + bZ) = a\,\mathrm{Cov}(X,Y) + b\,\mathrm{Cov}(X,Z)$$

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

Cauchy-Schwartz Inequality

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

Correlation Coefficient

normalization comes from Cauchy-Schwarz inequality

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}} \qquad \left(|\rho_{X,Y}| \le 1 \right)$$

If $X_1, \dots, X_m$ are <u>pairwise uncorrelated</u>, then

$$Var(X_1 + X_2 + \cdots + X_m) = Var(X_1) + Var(X_2) + \cdots + Var(X_m)$$

i.e

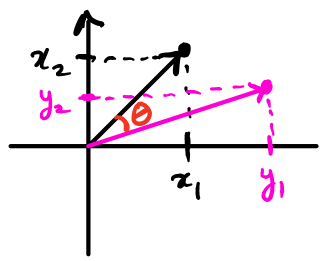
$$Var\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} Var(X_i)$$

Hilbert Space

Example: 2-dimensional Euclidean space.

$$\vec{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\vec{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$



$\langle \vec{x}, \vec{Y} \rangle = x_1 y_1 + x_2 y_2$ (inner product)

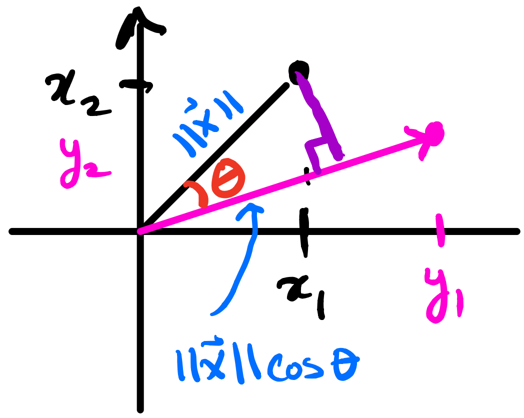
length/norm $(\vec{x}) = \|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$

angle b/t $(\vec{X}, \vec{Y})$ $\quad \cos\theta = \dfrac{\langle \vec{X}, \vec{Y} \rangle}{\|\vec{X}\| \|\vec{Y}\|}$ (via law of cosines)

Cauchy-Schwartz Inequality

magnitude　　　　　norm

$$|\langle \vec{x}, \vec{y} \rangle| \le \|\vec{x}\|\|\vec{y}\|$$



Projection of $\vec{x}$ onto $\vec{y}$

$$\Pi_{\vec{y}}(\vec{x}) = \|\vec{x}\|\cos(\theta)\underbrace{\frac{\vec{y}}{\|\vec{y}\|}}_{\text{along } \vec{y}\text{, normalized to unit length}}$$

notation for projection

$$= \frac{\langle x, y \rangle}{\|y\|^2}\vec{y}$$

Orthogonality

If $\vec{x} \perp \vec{y}$ 　 $\cos\theta = 0$

then $(\|\vec{x} + \vec{y}\|)^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$

$$\Pi_{\vec{y}}(\vec{x}) = 0$$

Vector Space (Linear Space)

* addition defined
* scalar multiplication by $a \in \mathbb{R}$ defined

Inner Product: An inner product on a vector space V is a mapping $V \times V \to \mathbb{R}$ such that $\forall \vec{x}, \vec{y} \in V$ and $a, b \in \mathbb{R}$

① $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$

② $\langle \vec{x}, \vec{x} \rangle \geq 0$ w/ equality iff $\vec{x} = \vec{0}$

③ $\langle x, a\vec{y} + b\vec{z} \rangle = a\langle \vec{x}, \vec{y} \rangle + b\langle \vec{x}, \vec{z} \rangle$

Hilbert Space

A Hilbert space is a vector space that

① Has an inner product $\langle \cdot, \cdot \rangle$ defined

② Is complete w/ respect to the norm $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$ induced by $\langle \cdot, \cdot \rangle$

→ complete in the sense every Cauchy sequence converges

Hilbert Space of Random Variables

All random variables w/ finite 2nd moments form a Hilbert space w/ inner product

r.v's NOT vectors

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

Inner Product $\Rightarrow$ norm/length $\Rightarrow$ distance metric

length/norm of $X = \sqrt{\langle X, X \rangle} = \sqrt{\mathbb{E}[X^2]}$

Cauchy-Schwarz ⇒ this ≤ 1

Angle b/t X, Y; θ_{XY}: $\cos(\theta_{XY}) = \dfrac{\langle X, Y \rangle}{\|X\|\|Y\|} = \dfrac{\mathbb{E}[XY]}{\sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}}$

Projection of X onto Y

a random variable

$$\Pi_Y(X) = \frac{\langle X, Y \rangle}{\|Y\|^2} Y = \frac{\mathbb{E}[XY]}{\mathbb{E}[Y^2]} Y$$

Orthogonality

$$\langle X, Y \rangle \triangleq \mathbb{E}[XY] = 0$$

$$\Pi_Y(X) = 0$$

Pythagorean

all X_i pairwise orthogonal

$$\mathbb{E}\left[\left(\sum_{i=1}^{m} X_i\right)^2\right] = \sum_{i=1}^{m} \mathbb{E}[X_i^2]$$

Mean Squared Error (MSE)

Have r.v. X which can't be directly observed.
Want to estimate $\hat{X}$.

Note: we are in a Hilbert space of random variables

The estimation error

$$W = X - \hat{X}$$

The MSE of $\hat{X}$ is

$$\mathbb{E}[(X-\hat{X})^2] = \mathbb{E}[W^2] = \langle W, W \rangle = \|W\|^2 = (d(X, \hat{X}))^2$$

Minimum Mean Squared Error

$$\hat{X}_{MMSE} = \underset{x}{\text{argmin}}\ \mathbb{E}[(X-\hat{x})^2]$$

MMSE of X using a constant

$$\hat{X}_{MMSE} = a^*$$

where

$$a^* = \underset{a \in \mathbb{R}}{\text{argmin}}\ \mathbb{E}[(X-a)^2]$$

it's $\mathbb{E}[X]$!

$$MSE = \mathbb{E}[(X-a)^2]$$

$$= \mathbb{E}[(X - \mathbb{E}[X] + \mathbb{E}[X] - a)^2]$$

$$= \mathbb{E}[(X-\mathbb{E}[X])^2] + \cancel{2\mathbb{E}[(X-\mathbb{E}[X])(\mathbb{E}[X]-a)]} \rightarrow 0 + \mathbb{E}[(\mathbb{E}[X]-a)^2]$$

Want to minimize

$$\mathbb{E}[(X-\mathbb{E}[X])^2] + \mathbb{E}[(\mathbb{E}[X]-a)^2]$$

$$\rightarrow \text{Var}(X) + (\mathbb{E}[X]-a)^2$$

Take $a = \mathbb{E}[X]$

MMSE of X given $Y=y$

similar proof

$$\hat{X}_{MMSE} = \mathbb{E}[X|Y=y]$$

$$MSE(\hat{X}_{MMSE}) = \text{Var}(X|Y=y)$$