6. Multivariate Normal Distributions

Thursday, January 20, 2022 4:30 PM

Random Variables (i.e., scalars)

Notation: $\mu_X = \mathbb{E}[X]$, $\sigma_X = \sqrt{Var[X]}$

Def Covariance

$$Cov(X,Y) = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \in \mathbb{R}$$
$$= \mathbb{E}[XY] - \mu_X \mu_Y$$

Def Correlation

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} =: \rho_{XY} \in [-1,1]$$

(unitless)

Facts

- $Var(X) = Cov(X,X)$
- $Cov(aX+b, Y) = a \cdot Cov(X,Y)$
- $Cov(X,Y) = Cov(Y,X)$ symmetry
- $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$ distributive

Random Vectors

Let $\vec{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ $\vec{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$

Def (vector) covariance $\vec{X} \in \mathbb{R}^m$, $\vec{Y} \in \mathbb{R}^n$

$Cov(\vec{X}, \vec{Y})$ is an $m \times n$ matrix,

(i,j) entry is $Cov(X_i, Y_j)$

Facts

- $Var(\vec{X}) := Cov(\vec{X}, \vec{X})$, often denoted Σ or Σ_X

  Σ is a symmetric matrix and non-negative definite

- $Cov(A\vec{X} + \vec{\mu}, B\vec{Y} + \vec{\nu}) = A \cdot Cov(\vec{X}, \vec{Y}) \cdot B^T$

  matrix vector

- $Cov(\vec{X}, \vec{Y}) = Cov(\vec{Y}, \vec{X})^T$

Def (vector) correlation

$$[Cor(\vec{X}, \vec{Y})]_{ij} = \frac{[Cov(\vec{X}, \vec{Y})]_{ij}}{\sigma_{X_i} \sigma_{Y_j}} = Cor(X_i, Y_j)$$

Normal Distribution (aka Gaussian)

If $X \sim N(\mu, \sigma^2)$ then it's probability density function, f,

is $f(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \exp\left(-\frac{1}{2} (x-\mu)^2 \right)$

Multivariate Normal Distribution

$\vec{X} \in \mathbb{R}^n$ has a multivariate normal distribution if its pdf f

is $$f(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n \, |\Sigma|}} \cdot \exp\Bigg(\underbrace{-\frac{1}{2} \underbrace{(\vec{x}-\vec{\mu})}_{\text{vector}}^T \overset{\text{matrix}}{\Sigma^{-1}} \underbrace{(\vec{x}-\vec{\mu})}_{\text{vector}}}_{\text{scalar}} \Bigg)$$

$|\Sigma| := \det(\Sigma)$

written $\vec{X} \sim N(\vec{\mu}, \Sigma)$ or $\sim N_n(\vec{\mu}, \Sigma)$

or $\sim MVN(\vec{\mu}, \Sigma)$

N.B. Books/papers usually use **bold** to denote vectors

also, $\vec{X} \sim N(\mu, \Sigma)$ then:

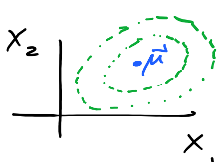
$X_1 \sim N(\mu_1, \Sigma_{11})$

$X_i \sim N(\mu_i, \Sigma_{ii})$

and in fact all marginals are (multi-variate) normal

2D case, $\vec{x} \in \mathbb{R}^2$

$f(\vec{x})$



level sets are ellipsis, determined by properties of Σ

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Generally:

$\vec{X} \sim N(\vec{\mu}, \Sigma)$

If Σ is diagonal (e.g., $\Sigma = \sigma^2 \cdot \overset{\text{identity}}{I}$)

this means X_i and X_j are ($i \neq j$) uncorrelated

but in fact for multivariate normal, X_i and X_j are actually independent

So $\vec{X} \sim N(\vec{0}, \sigma^2 I)$

$\Rightarrow X_i$ are iid $N(0, \sigma^2)$