Machine Learning, STOR 565

Random Vectors and the Multivariate Normal

Andrew Nobel

February, 2020

Review of the Univariate Case

Variance and Covariance

Recall: The variance of a random variable X is

$$\mathrm{Var}(X) \;=\; \mathbb{E}(X-\mathbb{E}X)^2 \;=\; \mathbb{E}X^2-(\mathbb{E}X)^2$$

and the covariance of random variables X, Y is

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

Basic Properties

- $\mathrm{Cov}(aX+b, cY+d) = ac\,\mathrm{Cov}(X,Y)$
- $\mathrm{Var}(X) = \mathrm{Cov}(X,X)$
- $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + 2\,\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)$

Univariate Normal

Recall: Given $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ the $\mathcal{N}(\mu, \sigma^2)$ distribution has density

$$f(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \qquad -\infty < x < \infty$$

- $\mu \in \mathbb{R}$ and $\sigma > 0$ called *parameters*, fully determine density f
- *standard normal* is special case $\mu = 0$ and $\sigma^2 = 1$

Notation: If X has density f above, write $X \sim \mathcal{N}(\mu, \sigma^2)$

Univariate Normal

Basic Properties: If $X \sim \mathcal{N}(\mu, \sigma^2)$ then

- $\mathbb{E}X = \mu$ and $\mathrm{Var}(X) = \sigma^2$
- $X \overset{\mathsf{d}}{=} \sigma Z + \mu$ where $Z \sim \mathcal{N}(0, 1)$
- $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Fact: If $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y \sim \mathcal{N}(\eta, \tau^2)$ are independent then

$$X + Y \sim \mathcal{N}(\mu + \eta, \sigma^2 + \tau^2)$$

Random Vectors

Random Vectors

Definition: A k-dimensional *random vector* is a vector of k random variables

$$X = (X_1, \cdots, X_k)^t \in \mathbb{R}^k$$

The *expected value* of X is

$$\mathbb{E}X = (\mathbb{E}X_1, \cdots, \mathbb{E}X_k)^t \in \mathbb{R}^k$$

Basic Properties: Let $a \in \mathbb{R}$, $v \in \mathbb{R}^k$, and $A \in \mathbb{R}^{r \times k}$

- $\mathbb{E}(aX + v) = a\,\mathbb{E}X + v$
- $\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y$
- $\mathbb{E}(AX) = A\,\mathbb{E}X$

Variance Matrix of a Random Vector

Definition: The variance matrix of a k-dimensional random vector $\mathbf{X}$ is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}X)(X - \mathbb{E}X)^t] \in \mathbb{R}^{k \times k}$$

Basic Properties

- $\mathrm{Var}(X)$ is symmetric and non-negative definite
- $\mathrm{Var}(X)_{ij} = \mathrm{Cov}(X_i, X_j)$
- $\mathrm{Var}(AX + v) = A\,\mathrm{Var}(X)\,A^t$

The Multivariate Normal

Multivariate Normal

Definition: A random vector $X \in \mathbb{R}^k$ is *multinormal* if for each $v \in \mathbb{R}^k$ the random variable $\langle X, v \rangle$ is univariate normal.

Notation: If $X \in \mathbb{R}^k$ is multinormal with $\mathbb{E}X = \mu$ and $\mathrm{Var}(X) = \Sigma$ write

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

Fact: Let $Z = (Z_1, \ldots, Z_n)^t$ be a vector of independent standard normals. If $X \sim \mathcal{N}_k(\mu, \Sigma)$ then

$$X \stackrel{\mathsf{d}}{=} \Sigma^{1/2} Z + \mu$$

Multivariate Normal Density

Note: Density of $\mathcal{N}(\mu, \sigma^2)$ can be written in the form

$$f(v) \;=\; \frac{1}{(2\pi)^{1/2}\,\sigma}\exp\left\{-\frac{1}{2}(v-\mu)(\sigma^2)^{-1}(v-\mu)\right\}$$

Fact: If $X \sim \mathcal{N}_k(\mu, \Sigma)$ with $\Sigma > 0$ then X has density

$$f(x) \;=\; \frac{1}{(2\pi)^{k/2}\det(\Sigma)^{1/2}}\exp\left\{-\frac{1}{2}(x-\mu)^t\,\Sigma^{-1}\,(x-\mu)\right\}$$

Standard Multinormal

Ex1: Standard multinormal vector $Z \sim \mathcal{N}_k(0, I)$ has density

$$f(x) \;=\; \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{1}{2} x^t x\right\} \;=\; \prod_{i=1}^{k} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{x_i^2}{2}\right\}$$

Components $Z_1, \ldots, Z_k$ of Z are independent standard normal r.v.

Bivariate Normal

EX 2: Random vector $(X, Y)^t \sim \mathcal{N}_2$ with $\mathrm{Corr}(X, Y) = \rho$ has joint density

$$f(x, y) \;=\; \frac{1}{2\pi\sigma_1\sigma_Y\sqrt{1-\rho^2}} \times$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right\}$$

Note:

- Density is defined only if $-1 < \rho < 1$
- X and Y are independent if and only if $\rho = 0$

Basic Properties of Multivariate Normal

Fact: Let $X \sim \mathcal{N}_k(\mu, \Sigma)$ be multinormal

- If $A \in \mathbb{R}^{l \times k}$, $b \in \mathbb{R}^l$ then $Y = AX + b \ \sim \ \mathcal{N}_l(A\mu + b, A\Sigma A^t)$
- $X_i \perp\!\!\!\perp X_j$ iff $\mathrm{Cov}(X_i, X_j) = 0$
- If $Y \sim \mathcal{N}_k(\mu', \Sigma')$ is independent of X then

$$X + Y \ \sim \ \mathcal{N}_k(\mu + \mu', \Sigma + \Sigma')$$