Chapter 5. Multiple Random Variables

5.9: The Multivariate Normal Distribution

Slides (Google Drive) Alex Tsun Video (YouTube)

In this section, we will generalize the Normal random variable, the most important continuous distribution! We were able to find the joint PMF for the Multinomial random vector using a counting argument, but how can we find the Multivariate Normal density function? We'll start with the simplest case, and work from there.

5.9.1 The Special Case of Independent Normals

Suppose $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are independent normal RVs.
Then by independence, their joint PDF is (recall that $\exp(z)$ is just another way to write e^z):

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_X^2}(x-\mu_X)^2\right) \cdot \frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_Y^2}(y-\mu_Y)^2\right), \qquad x, y \in \mathbb{R}$$

The mean vector μ is given by:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}$$

And the covariance matrix Σ is given by:

$$\Sigma = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}$$

Then, we say that (X, Y) has a bivariate Normal distribution, which we will denote:

$$(X, Y) \sim \mathcal{N}_2(\boldsymbol{\mu}, \Sigma)$$

This is nice and all, if we have two independent Normals. But what if they aren't independent?

5.9.2 The Bivariate Normal Distribution

We'll now see how we can construct the joint PDF of two (possibly dependent) Normal RVs, to get the Bivariate Normal PDF.

Definition 5.9.1: The Bivariate Normal Distribution

Let $Z_1, Z_2 \sim \mathcal{N}(0,1)$ be iid standard Normals, and $\mu_X, \mu_Y, \sigma_X^2 > 0, \sigma_Y^2 > 0$ and $-1 \le \rho \le 1$ be scalar parameters. We construct from these two RVs a random vector (X, Y) by the transformations:

1. We construct X by taking Z_1, multiplying it by σ_X, and adding μ_X:

$$X = \sigma_X Z_1 + \mu_X$$

2. We construct Y from both Z_1 and Z_2, as shown below:

$$Y = \sigma_Y(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_Y$$

From this transformation, we get that marginally (show this by computing the mean and variance of X, Y and closure properties of Normal RVs),

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2) \qquad Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

Additionally,

$$\rho(X,Y) = \rho = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\,\mathsf{Var}(Y)}} = \frac{\mathsf{Cov}(X,Y)}{\sigma_X \sigma_Y} \Rightarrow \mathsf{Cov}(X,Y) = \rho\sigma_X\sigma_Y$$

That is, for the the RVTR (X, Y),

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$$

By using the multivariate change-of-variables formula from 4.4, we can turn the "simple" product of standard normal PDFs into the PDF of the bivariate Normal:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{z}{2(1-\rho^2)}\right), \qquad x, y \in \mathbb{R}$$

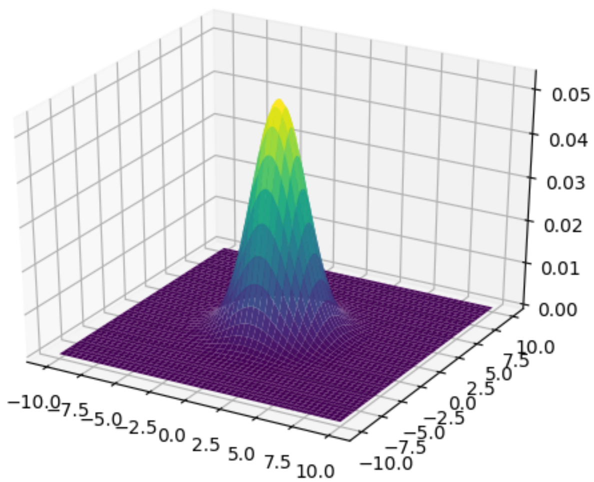
where

$$z = \frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}$$

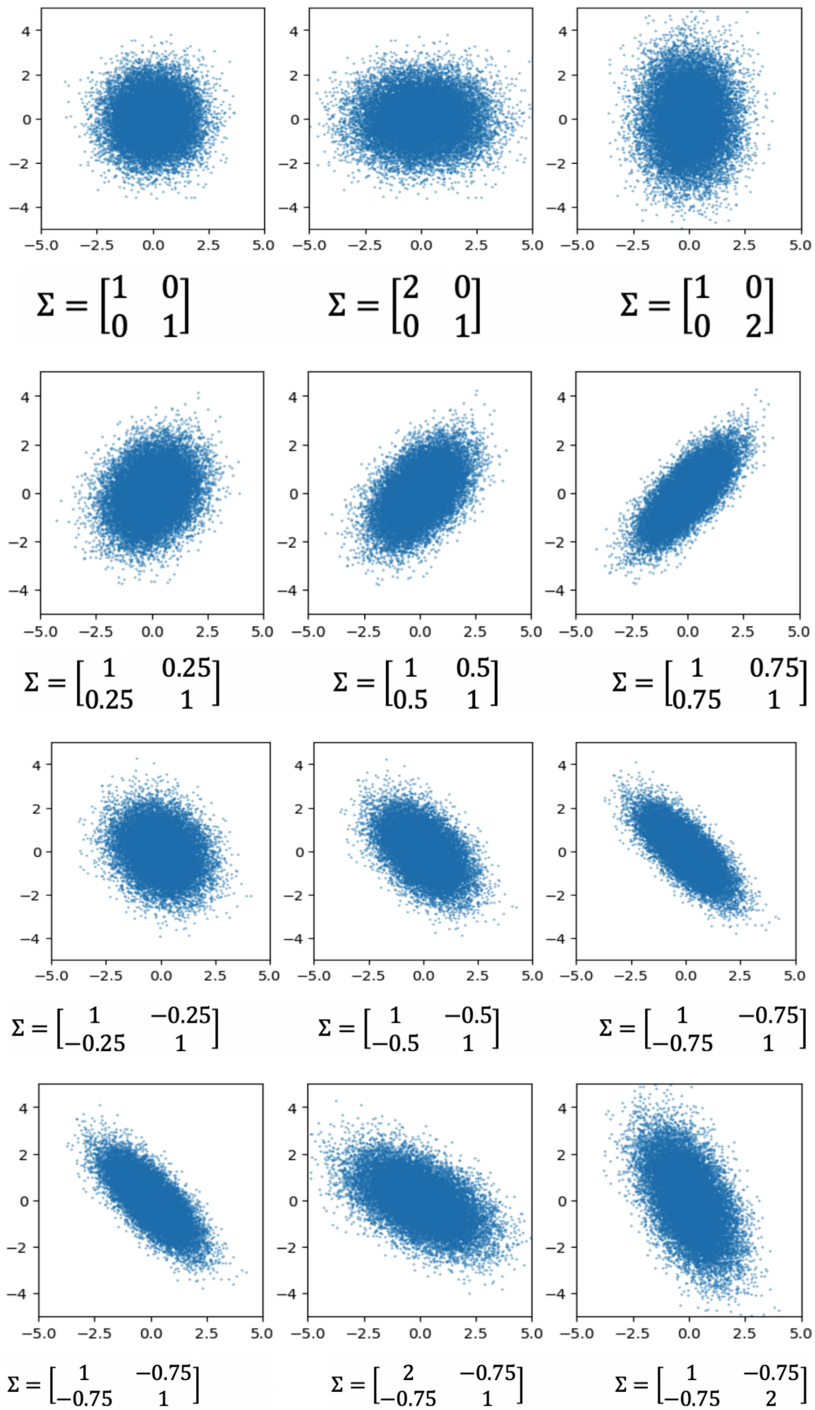
Finally, we write:

$$(X, Y) \sim \mathcal{N}_2(\boldsymbol{\mu}, \Sigma)$$

The visualization below shows the density of a bivariate Normal distribution. On the xy-plane, we have the actual two Normas, and on the z-axis, we have the density. Marginally, both variables are Normals!


0.05
0.04
0.03
0.02
0.01
0.00
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0

Now let's take a look at the effect of different covariance matrices Σ on the distribution for a bivariate normal, all with mean vector $(0, 0)$. Each row below modifies one entry in the covariance matrix; see the pictures graphically to explore how the parameters change the shape!



5.9.3 The Multivariate Normal Distribution

Definition 5.9.2: The Multivariate Normal Distribution

A random vector $\mathbf{X} = (X_1, ..., X_n)$ has a multivariate Normal distribution with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and (symmetric and positive-definite) covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$, written $\mathbf{X} \sim \mathcal{N}_n(\boldsymbol{\mu}, \Sigma)$, if it has the following joint PDF:

$$f_{\mathbf{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right), \qquad \boldsymbol{x} \in \mathbb{R}^n$$

While this PDF may look intimidating, if we recall the PDF of a univariate Normal $W \sim \mathcal{N}(\mu, \sigma^2)$:

$$f_W(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(w - \mu)^2\right)$$

We can note that the two formulae are quite similar; we simply extend scalars to vectors and matrices!

Additionally, let us recall that for any RVs X and Y:

$$X \perp Y \quad \to \quad \mathsf{Cov}(X, Y) = 0$$

If $\mathbf{X} = (X_1, \ldots, X_n)$ is Multivariate Normal, the converse also holds:

$$\mathsf{Cov}(X_i, X_j) = 0 \quad \to \quad X_i \perp X_j$$

Unfortunately, we cannot do example problems as they would require a deeper knowledge of linear algebra, which we do not assume.