4 Bivariate distributions

In this section we consider general simulation from a bivariate distribution, as well as simulation from the bivariate Normal distribution. In this week's practical we will also look at *transformations* of bivariate random variables.

4.1 Sampling from bivariate distributions

Suppose we have a pair of continuous random variables (X, Y) with joint PDF $f_{X,Y}(x, y)$. Recall that

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X}(y \mid X = x)$$

where $f_X(x)$ is the **marginal** PDF of X and $f_{Y|X}(y \mid X = x)$ is the **conditional** PDF of Y given $X = x$. One way to simulate realizations of the pair (X, Y) is:

1. Simulate a realization x of X from the marginal distribution of X.
2. Simulate a realization y of Y from the conditional distribution of Y given $X = x$.

In the **trivariate case** we can add a third step: simulate Z given $Y = y$ and $X = x$.

Example 4.1: Suppose $X \sim U(0.1, 0.5)$ and

$$f_{Y|X}(Y \mid X = x) = \begin{cases} xe^{-xy}, & \text{when } y \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Write an R function to generate a $2 \times n$ **matrix** of samples from this distribution, and produce a scatter plot.

Note that, given that $X = x$, Y has an exponential distribution with parameter $\lambda = X$:

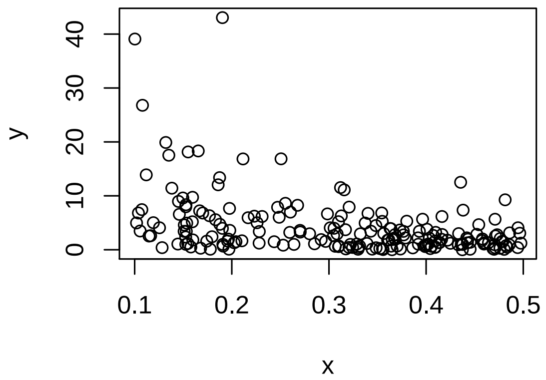
$$Y \mid X = x \sim Exp(x).$$

```r
example4.1 = function(n) {
    output = matrix(0, 2, n)
    x = runif(n, 0.1, 0.5)
    y = rexp(n, rate = x)
    output[1, ] = x
    output[2, ] = y
    output
}
```

Make some notes about the R function.

To obtain a scatterplot:

```
test = example4.1(200)
plot(test[1,], test[2,], xlab = 'x', ylab = 'y')
```



Example 4.2: (Trivariate case.) Suppose that $X \sim U(-1, 1)$ and $Y \sim Exp(\lambda)$ independently of X and that

$$f_{Z|X,Y}(Z \mid X = x, Y = y) = \frac{1}{\sqrt{2\pi}y} e^{-\frac{1}{2}\left(\frac{z-x}{y}\right)^2}.$$

Write an R function to sample realizations (X, Y, Z) in the case $\lambda = 5$ and produce a scatterplot of (X, Z). What would the plot look like if $\lambda = 1$?

Note that, given that $X = x$, $Y = y$ then Z has a normal distribution:

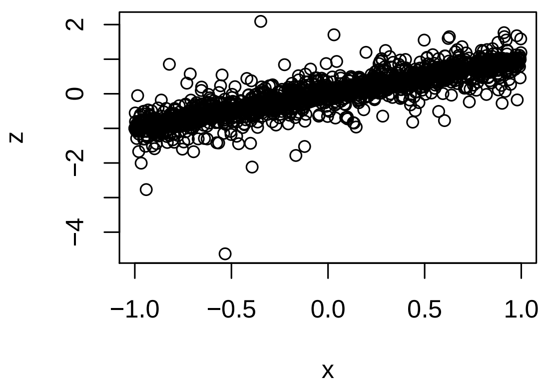
$$Z \mid X = x, Y = y \sim N(x, y^2).$$

```
example4.2 = function(n) {
    output = matrix(0, 3, n)
     x = runif(n, -1, 1)
     y = rexp(n, rate = 5)
     z = rnorm(n, mean = x, sd = y)
     output[1, ] = x
     output[2, ] = y
     output[3, ] = z
     output
}
```

Make some notes about the R function.

To obtain a scatterplot:

```
output = example4.2(2000)
plot(output[1,], output[3,], xlab = 'x', ylab = 'z')
```



4.2 The multivariate normal distribution

We met the bivariate normal distribution in MAS1604. The bivariate normal distribution is a generalization of the normal distribution to pairs of random variables, or equivalently, to a distribution on vectors in $\mathbb{R}^2$. The multivariate normal distribution is the analogous distribution on vectors in $\mathbb{R}^n$.

4.2.1 The bivariate normal distribution

Suppose we have constants μ_x, μ_y, σ_x, σ_y and ρ such that $\sigma_x \geq 0$, $\sigma_y \geq 0$ and $-1 \leq \rho \leq 1$. Define the 2×2 matrix Σ by

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}.$$

Then define a joint probability density function by

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{\det\Sigma}} \exp\left(-\frac{1}{2}Q(x,y)\right) \tag{1}$$

where

$$Q(x,y) = (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu}) \tag{2}$$

and the column vectors $\underline{x}, \underline{\mu}$ are defined by $\underline{x} = (x,y)^T$, $\underline{\mu} = (\mu_x, \mu_y)^T$. If random variables (X,Y) have joint probability density given by $f_{X,Y}$ above, then we say that (X,Y) have a **bivariate normal distribution** and write

$$(X,Y)^T \sim N_2(\underline{\mu}, \Sigma).$$

Some of the terms in the definition can be expanded:

$$\sqrt{\det\Sigma} = \sigma_x\sigma_y\sqrt{1-\rho^2}$$

and

$$Q(x,y) = \frac{1}{\sigma_x^2\sigma_y^2(1-\rho^2)}\begin{pmatrix} x-\mu_x & y-\mu_y \end{pmatrix}\begin{pmatrix} \sigma_y^2 & -\rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & \sigma_x^2 \end{pmatrix}\begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}$$
$$= \frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right].$$

However, equations (1) and (2) are **much** more important than the expansions just given.

It can be proved that the function $f_{X,Y}(x,y)$ in equation (1) integrates to 1 and therefore defines a valid joint PDF. The proof is omitted.

Remarks:

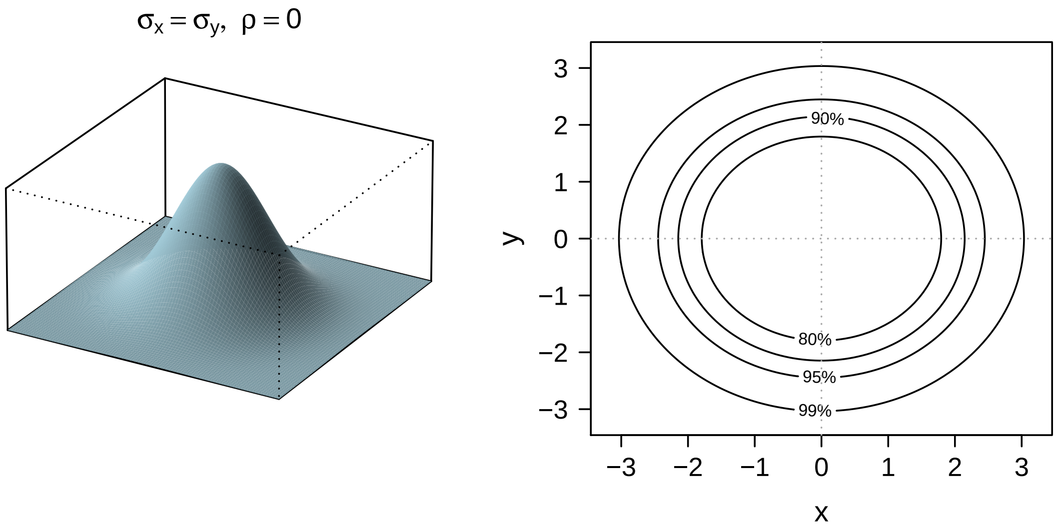
1. The vector $\underline{\mu} = (\mu_x, \mu_y)^T$ is called the **mean vector** and the matrix Σ is called the **covariance** matrix (or sometimes **variance-covariance** matrix).
2. Functions of the form $F(\underline{x}) = \underline{x}^T\Sigma^{-1}\underline{x}$ are called **quadratic forms**. Quadratic forms are functions $\mathbb{R}^n \to \mathbb{R}$ which satisfy certain properties. They crop up in several areas of mathematics and statistics.
3. The matrix Σ and its inverse Σ^{-1} are **positive definite**. A matrix A is positive definite if

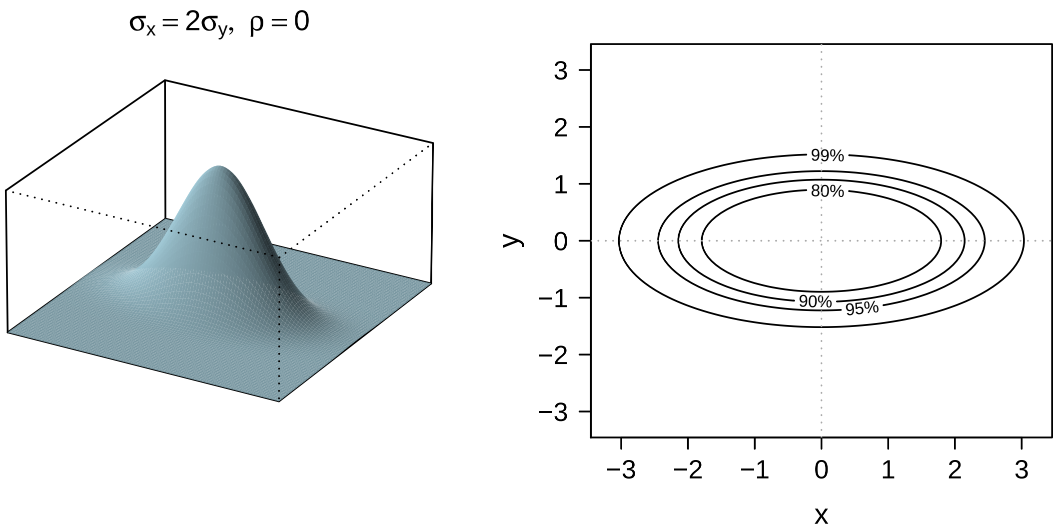
$$\underline{x}^T A \underline{x} \geq 0$$

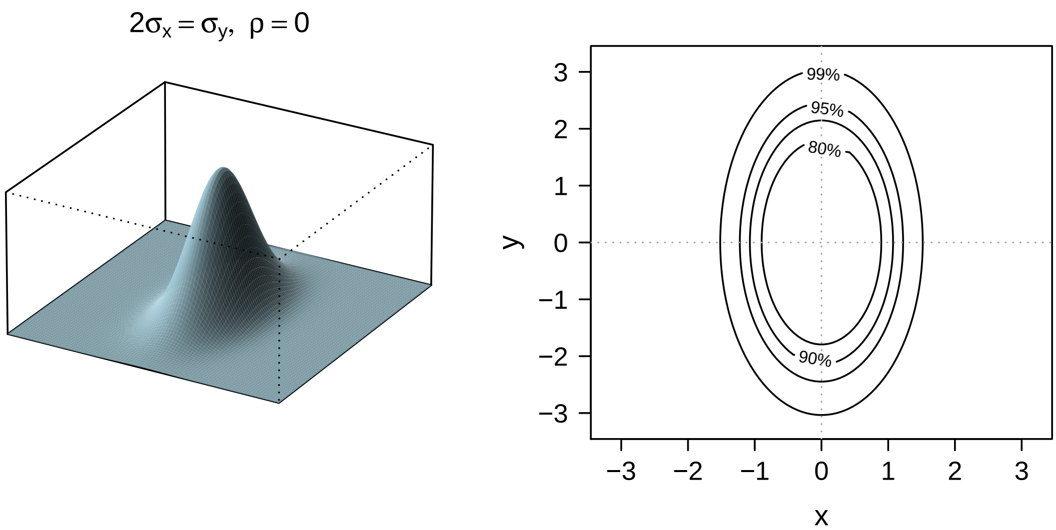
for all non-zero vectors $\underline{x}$.

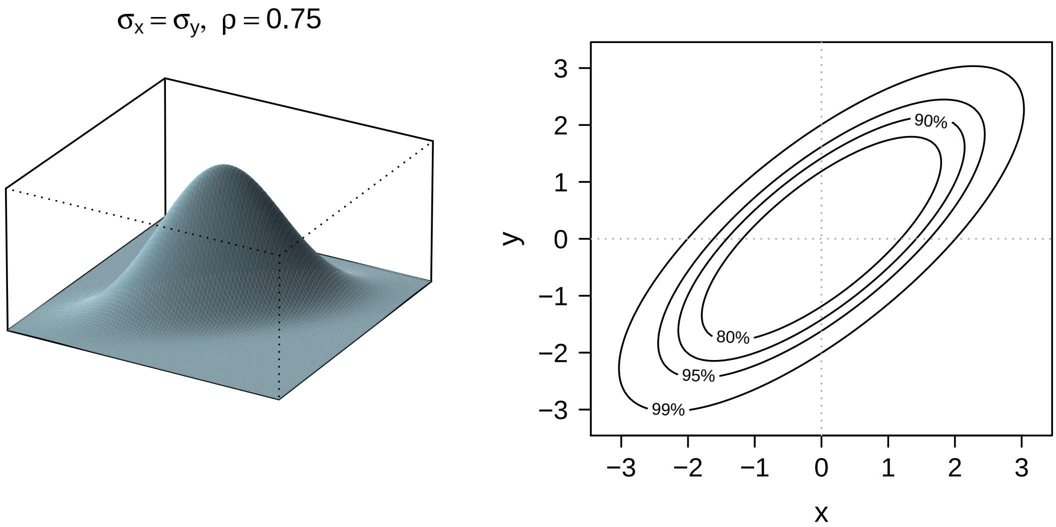
4. It follows that when $\mu_x = \mu_y = 0$, $Q(x,y)$ is of positive definite quadratic form.

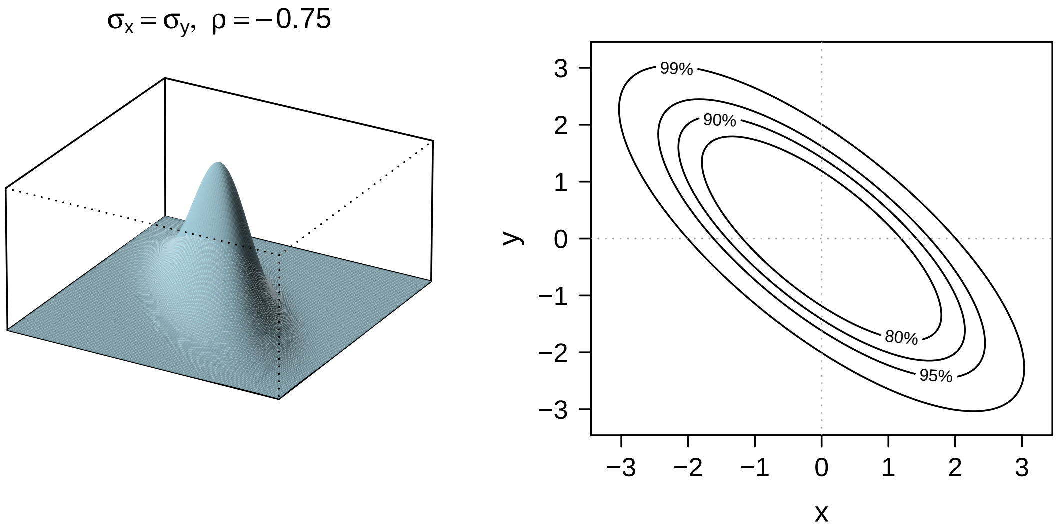
Over the next two pages we will visualize the bivariate Normal distribution using 3-D plots and contour plots.

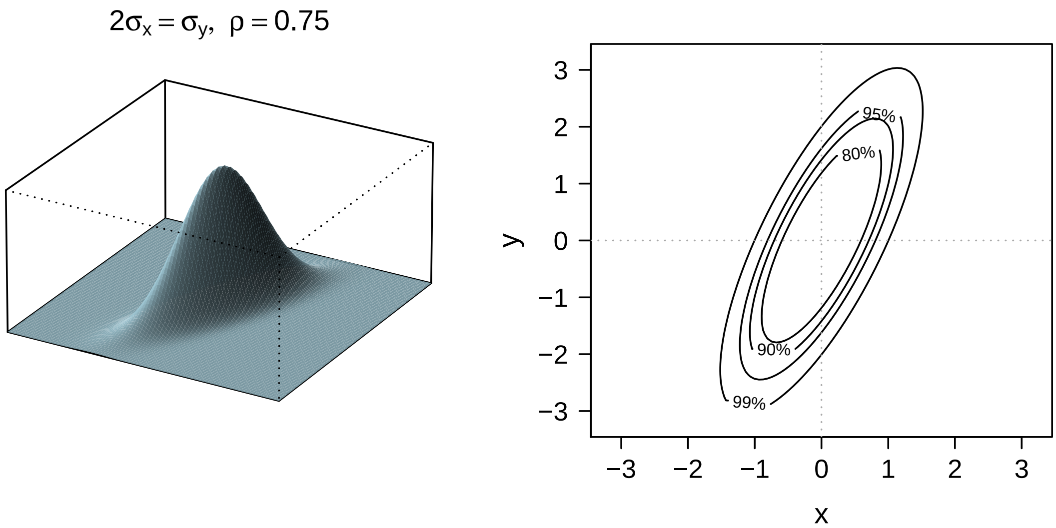

$\sigma_x = \sigma_y,\ \rho = 0$
90%
80%
95%
99%
y
x
−3 −2 −1 0 1 2 3


$\sigma_x = 2\sigma_y,\ \rho = 0$
99%
80%
90%
95%
y
x
−3 −2 −1 0 1 2 3


$2\sigma_x = \sigma_y,\ \rho = 0$
99%
95%
80%
90%
y
x
−3 −2 −1 0 1 2 3


$\sigma_x = \sigma_y, \ \rho = 0.75$
99%
95%
90%
80%
y
x


$\sigma_x = \sigma_y, \ \rho = -0.75$
99%
95%
90%
80%
y
x


$2\sigma_x = \sigma_y, \ \rho = 0.75$
99%
95%
90%
80%
y
x

The previous two pages show contour and surface plots for the density function on the xy-plane for different parameter values. Comments:

1. $Q(x, y) \geq 0$ with equality only when $\underline{x} = \underline{\mu}$. It follows that the density function has its mode at $\underline{x} = \underline{\mu}$. The above plots all have $\mu_x = \mu_y = 0$.
2. Changing the values of μ_x, μ_y does not change the shape of the plots, but corresponds to a translation of the xy-plane i.e. changing μ_x, μ_y just shifts the contours / surface to a new mode position.
3. The contours of equal density are **circular** when $\sigma_x = \sigma_y$ and $\rho = 0$ and **elliptical** when $\sigma_x \neq \sigma_y$ or $\rho \neq 0$.
4. σ_x and σ_y control the extent to which the distribution is **dispersed**.
5. The parameter ρ is the **correlation** of X, Y i.e. $\text{Cor}(X, Y) = \rho$. (The proof is omitted.) Thus for non-zero ρ, the contours are at an angle to the axes.

When (X, Y) has a bivariate normal distribution then given a general set $A \subseteq \mathbb{R}^2$, calculation of the probability $\Pr((X, Y) \in A)$ requires a computer. However, the marginal and conditional distributions of the bivariate normal distribution are particularly simple – one of the reasons the distribution is so useful in statistics.

Suppose $(X, Y)^T \sim N_2(\underline{\mu}, \Sigma)$. The following results hold.

1. The marginal distributions are normal:
$$X \sim N(\mu_x, \sigma_x^2) \quad \text{and} \quad Y \sim N(\mu_y, \sigma_y^2).$$
2. The conditional distributions are normal:
$$\begin{aligned} X|Y = y &\sim N(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y), \sigma_x^2(1 - \rho^2)) \quad \text{and} \\ Y|X = x &\sim N(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1 - \rho^2)). \end{aligned}$$
3. When $\rho = 0$, X and Y are independent.
4. Linear combinations of X and Y are also normally distributed:
$$aX + bY \sim N(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\rho\sigma_x\sigma_y)$$
where a, b are constants. Thus a **linear combination** of normal random variables is also normally distributed itself.

Proofs are omitted.

4.2.2 Simulating bivariate normal random variables

R has built-in functions to simulate bivariate and multivariate normal random variables. However, we can simulate using the strategy introduced in section 4.1: first simulate from the marginal for X (or marginal for Y if preferred), and then simulate Y from the conditional $Y|X = x$.

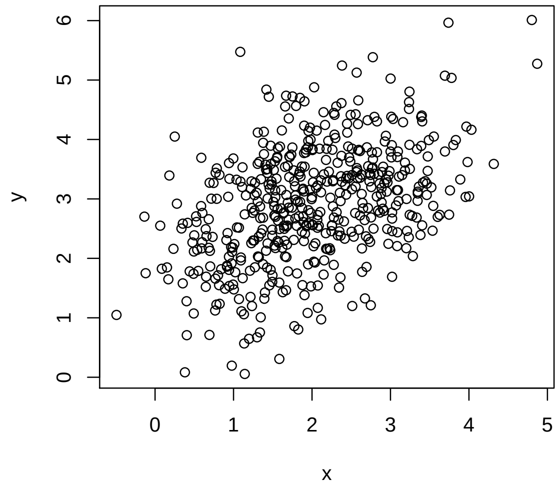
Example 4.3: Suppose $(X, Y)^T \sim N_2(\underline{\mu}, \Sigma)$ where $\mu_x = 2$, $\mu_y = 3$, $\sigma_x = 1$, $\sigma_y = 1$ and $\rho = 0.5$. Simulate a sample of size 500 from this distribution and draw a scatter plot. Use simulation to find $\Pr\left(X^2 + Y^2 < 9\right)$.

The marginal distribution of X is $X \sim N(2, 1^2)$. Using the formula for the conditional

$$\begin{aligned} Y|X = x &\sim N(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1 - \rho^2)) \\ &\sim N(3 + 0.5(x - 2), 0.75). \end{aligned}$$

To sample:

```
npts = 500
x = rnorm(npts, mean = 2, sd = 1)
y = rnorm(npts, mean = 3+0.5*(x-2), sd = sqrt(0.75))
```



To find $\Pr\left(X^2 + Y^2 < 9\right)$ approximately, count the number of points in the region:

```
npts = 10000
x = rnorm(npts, mean = 2, sd = 1)
y = rnorm(npts, mean = 3+0.5*(x-2), sd = sqrt(0.75))
f = x^2+y^2
sum(f<9)/npts
[1] 0.2877
```

4.2.3 The multivariate normal distribution

The multivariate normal distribution is a distribution on vectors in $\mathbb{R}^n$. Suppose that $\underline{X}$ is a random vector with n entries, i.e. $\underline{X} = (X_1, \ldots, X_n)^T$. Then

$$\underline{X} \sim N_n(\underline{\mu}, \Sigma)$$

if $X_1, \ldots, X_n$ have joint PDF given by

$$f_{\underline{X}}(\underline{x}) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}Q(\underline{x})\right)$$

where

$$Q(\underline{x}) = (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu}).$$

This definition makes sense for any column vector $\underline{\mu} \in \mathbb{R}^n$ and any **positive definite** $n \times n$ matrix Σ. Some remarks:

1. The vector $\underline{\mu}$ is the **mean** of the distribution and Σ is called the **covariance** matrix.
2. All the marginal distributions of $\underline{X}$ are normal. (We do not specify their parameters here, however).
3. Similarly, all the conditional distributions of $\underline{X}$ are normal. (Again, we do not specify the parameters of these distributions here).

Example 4.4: Suppose $\underline{X} \sim N_3(0, \Sigma)$ where

$$\Sigma = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

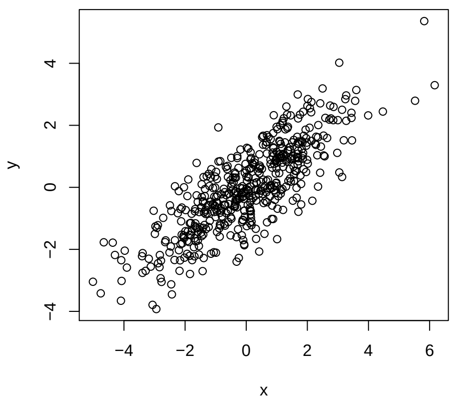
Use R to sample from this distribution. Create a scatterplot of (X, Y) samples and compute the correlation.

In R, you use the `mvrnorm` function to sample:

```
library(MASS)
npts = 500
mu = c(0, 0, 0)
sigma = matrix(data = c(3,2,1,2,2,1,1,1,3), nrow = 3, byrow = TRUE)
x = mvrnorm(npts, mu, sigma)
```

Now `x` has dimension `npts`×3, so column 1 is X and column 2 is Y:

```
plot(x[,1], x[,2], xlab = 'x', ylab = 'y')
```



The sample correlation is:

```
cor(x[,1],x[,2])
[1] 0.812629
```