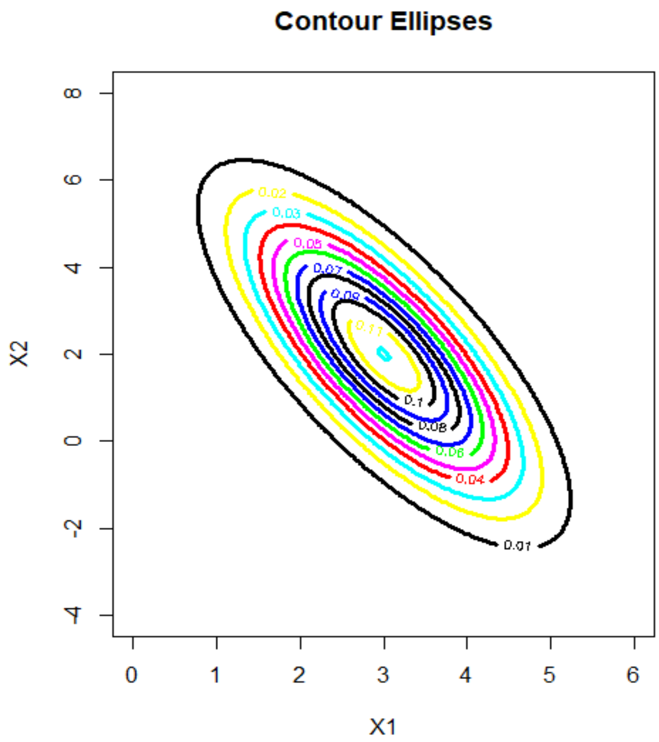
Exercise 4.1 Assume that the random vector Y has the following normal distribution: $Y \sim N_P(0, I)$. Transform it according to (4.49) to create $X \sim N(\mu, \Sigma)$ with mean $\mu = (3, 2)^T$ and $\Sigma = \begin{pmatrix} 1 & -1.5 \\ -1.5 & 4 \end{pmatrix}$. How would you implement the resulting formula on a computer?

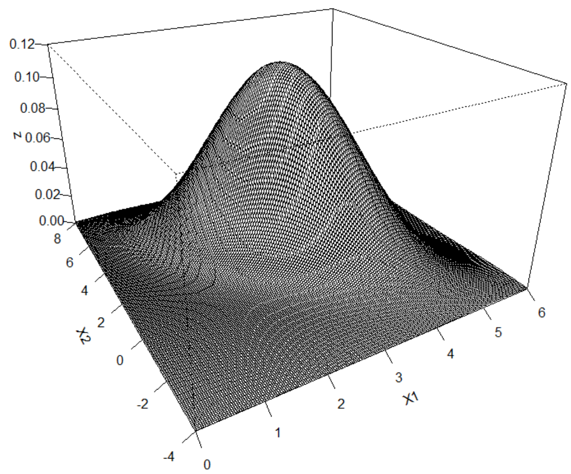
Sol.

$Y \sim N_P(0, I)$ 用 $X = \Sigma^{\frac{1}{2}} Y + \mu$ 做線性轉換得到$X \sim N(\mu, \Sigma)$，多變量常態母體 $X \sim N(\mu, \Sigma)$ 意味它的變數 $(x - \mu)'\Sigma^{-1}(x - \mu) \sim \chi^2(p)$ 服從卡方分配，自由度為 p。



等高線圖中，圖形中心點為平均數 $\mu = (3, 2)^T$，橢圓內的每個值與μ的距離都小於$\chi^2(p)$，數學式可表示為$(x - \mu)'\Sigma^{-1}(x - \mu) \leq \chi^2(p)$，當$x$越靠近橢圓中心$\mu$，表示它的 probability 也越高。

另外，搭配 3D surface plot 可以觀察出 X 只有一個高峰，而且等高線圖的每條等高線間距沒有特別離很遠或很近的，整體呈現對稱狀態，可以推測為常態分布。



Exercise 5.1 Consider $X \sim N_2(\mu, \Sigma)$ with $\mu = (2,2)^T$ and $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and the matrices $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}^T$, $B = \begin{pmatrix} 1 \\ -1 \end{pmatrix}^T$. Show that AX and BX are independent.

Sol.

根據多變量常態分配定理，若$X \sim N_2(\mu, \Sigma)$，則X的任意線性組合皆為常態分配。

$\rightarrow$ *假設*$X_1 = AX = \sum_{i=1}^{2} A_i X_i$, $\quad X_2 = BX = \sum_{i=2}^{2} B_i X_i$

$\rightarrow AX \sim N(A^T\mu, A^T\Sigma A), \quad BX \sim N(B^T\mu, B^T\Sigma B)$

↪ AX 和 BX 皆呈現常態分配

$\rightarrow Cov(X_1, X_2) = A\Sigma B^T = (1 \quad 1)\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\rightarrow A\Sigma = (I_r \quad 0)\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = (\Sigma_{11} \quad \Sigma_{12})$

$\rightarrow A\Sigma B^T = (\Sigma_{11} \quad \Sigma_{12})\begin{pmatrix} 1 \\ -1 \end{pmatrix} = (\Sigma_{11} - \Sigma_{12})$

$\rightarrow A\Sigma B^T \equiv 0$

↪ $Cov(X_1, X_2) = 0$

$\rightarrow \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} X \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}\right)$

$\rightarrow \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} X \sim N_2\left(\begin{pmatrix} A\mu \\ B\mu \end{pmatrix}, \begin{pmatrix} A\Sigma A^T & 0 \\ 0 & B\Sigma B^T \end{pmatrix}\right)$

$\rightarrow \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} X \sim N_2\left(\begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\right)$

$\Rightarrow Because\ \Sigma_{12}, \Sigma_{21} = 0, \quad AX \quad and\ BX \quad are\ independent.$

R Code:

```
# clear all variables
rm(list = ls(all = TRUE))
graphics.off()

# 4.1
# install and load packages
libraries = c("MASS", "mnormt")
lapply(libraries, function(x) if (!(x %in% installed.packages())) {
  install.packages(x)
})
lapply(libraries, library, quietly = TRUE, character.only = TRUE)
# parameter settings
n1     = 200    # number of draws
mu1    = c(3, 2)    # mean vector
sig1 = matrix(c(1, -1.5, -1.5, 4), ncol = 2)    # covariance matrix
# bivariate normal sample
set.seed(1024)
x = mvrnorm(n1, mu1, sig1, 2)
# bivariate normal density
xgrid = seq(from = (mu1[1] - 3 * sqrt(sig1[1, 1])), to = (mu1[1] + 3 * sqrt(sig1[1, 1])),
            length.out = 200)
ygrid = seq(from = (mu1[2] - 3 * sqrt(sig1[2, 2])), to = (mu1[2] + 3 * sqrt(sig1[2, 2])),
            length.out = 200)
z       = outer(xgrid, ygrid, FUN = function(xgrid, ygrid) {
  dmnorm(cbind(xgrid, ygrid), mean = mu1, varcov = sig1)
})
# contour ellipses
contour(xgrid, ygrid, z, xlim = range(xgrid), ylim = range(ygrid), nlevels = 10, col =
c("blue",

"black", "yellow", "cyan", "red", "magenta", "green", "blue", "black"), lwd = 3,
        cex.axis = 1, xlab = "X1", ylab = "X2")
title("Contour Ellipses")
# surface plot
persp(xgrid, ygrid, z, theta=-30, phi=25, expand=0.6, ticktype='detailed', xlab="X1",
ylab = "X2" )
```

```
# 5.1
# parameter settings
n2     = 200   # number of draws
mu2    = c(2, 2)   # mean vector
sig2 = matrix(c(1, 0, 0, 1), ncol = 2)   # covariance matrix
A <- t(matrix(c(1, 1)))
B <- t(matrix(c(1, -1)))
# bivariate normal sample
set.seed(1024)
x2 = mvrnorm(n2, mu2, sig2, 2)
#  先證 AX,BX 常態，再做共變異數=0
m1 <- A%*%mu2
m2 <- B%*%mu2
s11 <- A%*%sig2%*%t(A)
s12 <- A%*%sig2%*%t(B)
s21 <- B%*%sig2%*%t(A)
s22 <- B%*%sig2%*%t(B)
S <- matrix(c(s11, s21, s12, s22), ncol = 2)
```