Modele Liniowe
Raport nr 1

Michał Kubica

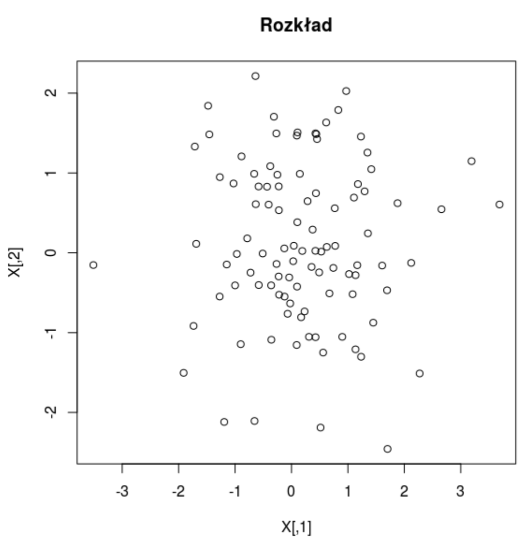
3 grudnia 2018

1 Zadanie 1

Wygenerowano 100 wektorów losowych z rozkładu dwuwymiarowego normalnego $N(0, I)$ poniższym kodem w R:

```
X = matrix(0,100,2)
for (i in 1:100){
  X[i,] = rnorm(2)
}
```

A następnie zaznaczono je na płaszczyźnie.



Rysunek 1

2 Zadanie 2

W tym zadaniu należało przekształcić liniowo otrzymane wektory losowe tak, aby miały one rozkład o średniej $\mu = \binom{4}{2}$ i macierzy kowiariancji Σ. Korzystając z własności i wzorów na liście zadań zadanie sprowadziło się do wyznaczenia takiej macierzy A, aby spełnione było równanie macierzone: $\Sigma = AA^T$. Macierz A wyznaczono na dwa sposoby: metodą Choleskiego(funkcji *chol*) oraz rachunkiem algebraicznym jak w poniższym przykładzie.

$$\Sigma = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}$$

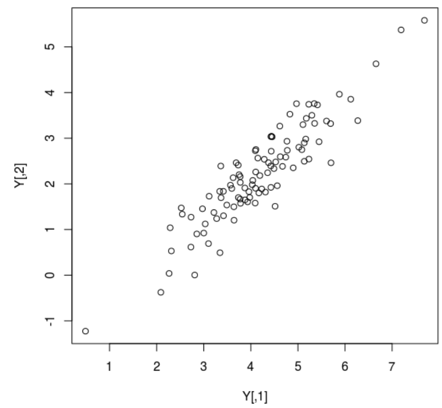
$$\begin{cases} a^2+b^2 = 1 \\ ac+bd = 0,9 \\ c^2+d^2 = 1 \end{cases}$$

przyjmując b = 0 otrzymujemy

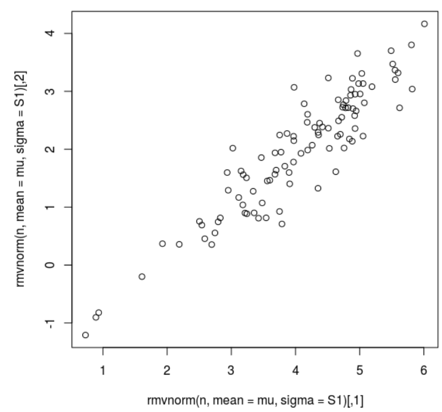
$$\text{A} = \begin{pmatrix} 1 & 0 \\ 0,9 & \sqrt{0.19} \end{pmatrix}$$

Dla każdej z trzech róznych macierzy kowariancji zaznaczono na płaszczyźnie przekształconą chmurę wektorów oraz wygenerowane wektory za pomocą funcji *rmvnorm*.

2.1 $\Sigma = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$



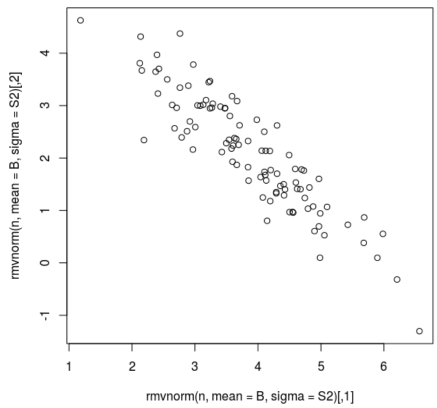
Rysunek 2: Przekształcona chmura wektorów



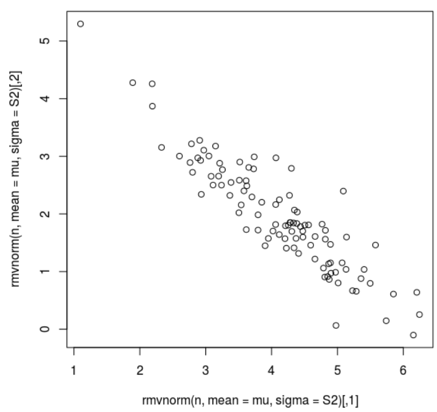
Rysunek 3: Wylosowane 100 wektorów losowych za pomocą rvrnorm

Na obu wykresach widać wyraźną zależność liniową między wektorami losowymi. Wyniki się zgadzają z przewidywaniami związanymi z postacią macierzy kowariancji.

2.2 $\Sigma = \begin{pmatrix} 1 & -0.9 \\ -0.9 & 1 \end{pmatrix}$



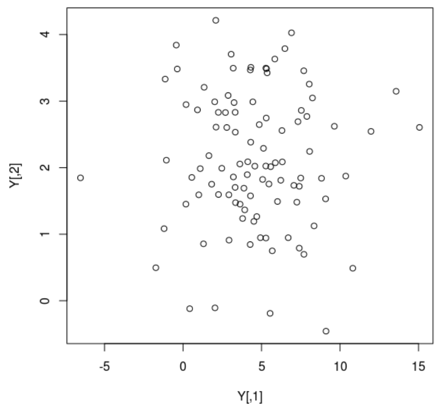
Rysunek 4: Przekształcona chmura wektorów



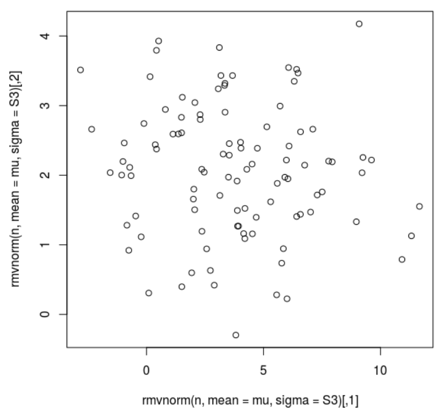
Rysunek 5: Wylosowane 100 wektorów losowych za pomocą rvrnorm

Na obu wykresach widać wyraźną odwrotną zależność liniową między wektorami losowymi. Wyniki się zgadzają z przewidywaniami związanymi z postacią macierzy kowariancji.

2.3 $\Sigma = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$



Rysunek 6: Przekształcona chmura wektorów



Rysunek 7: Wylosowane 100 wektorów losowych za pomocą rvrnorm

W tym przypadku nie ma jakiejkolwiek zależności liniowej między wektorami, ale to jednak nie wyklucza istnienia zależności nieliniowej.

2.4 Kod R

```
S1 = matrix(c(1, 0.9, 0.9, 1), 2, 2)
S2 = matrix(c(1, -0.9, -0.9, 1), 2, 2)
S3 = matrix(c(9, 0, 0 ,1), 2, 2)

sqrt(0.19)
A1 = t(chol(S1))
A2 = t(chol(S2))
A3 = t(chol(S3))

mu = c(4,2)

Y = t(A1%*%t(X)+mu)
plot(Y)
plot(rmvnorm(n, mean=mu, sigma=S1 ))

Y = t(A2%*%t(X)+mu)
plot(Y)
plot(rmvnorm(n, mean=mu, sigma=S2 ))

Y = t(A3%*%t(X)+mu)
plot(Y)
plot(rmvnorm(n, mean=mu, sigma=S3 ))
```

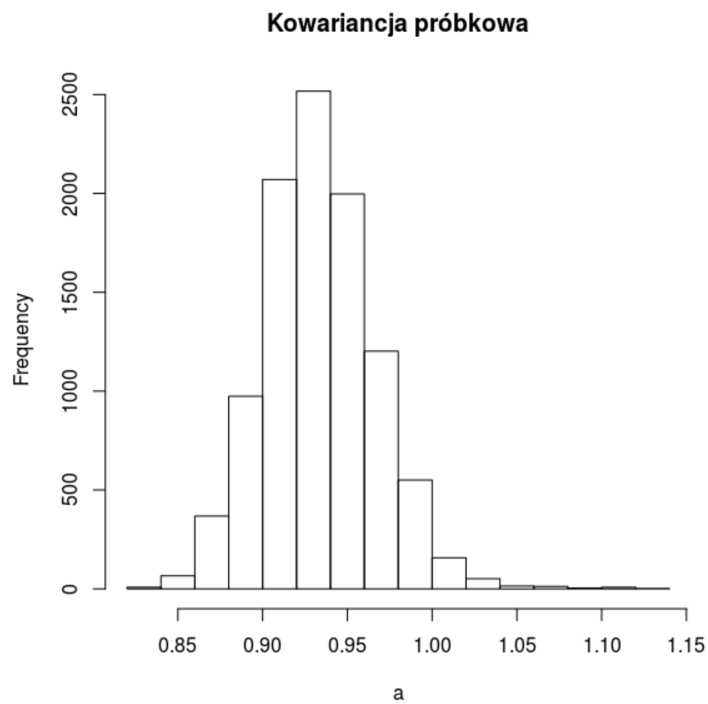
3 Zadanie 3

Za pomocą funkcji *rnorm* wygenerowano 200 wektorów losowych z rozkładu wielowymiarowego normalnego $N(0, I_{100x100})$. Korzystając z własności przekształcenia liniowego i metody Choleskiego wyznaczono macierz A taką, że $\tilde{X} = AX$ oraz $\tilde{X} \sim N(0, \Sigma)$, gdzie

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & 0.9 & \dots & 0.9 \\ 0.9 & 1 & \dots & 0.9 \\ \vdots & \vdots & \ddots & 0.9 \\ 0.9 & 0.9 & 0.9 & 1 \end{pmatrix}$$

Ze względu na trudności, związane z wizualizacją wektorów losowych o wymiarze 100, zweryfikowano poprawność wygenerowanych wektorów losowych wyliczając średnią, rysując histogram próbkowych wariancji współrzędnych wektora $\tilde{X}$ oraz próbkowych kowariancji między różnymi współrzędnymi tego wektora.

Wyliczona średnia wynosi: -0.003316602



Rysunek 8: Histogram próbkowych kowariancji

Kowariancja próbkowa waha się od 0.9 do 1, a więc jest to zgodne z oczekiwaniami i postacia macierzy kowariancji. A więc na podstawie wyliczenia średniej i powyższego wykresu można powiedzieć, że wketory został wygenerowane poprawnie.

3.1 Kod R

```
my_rmvnorm=function(mu,Sigma){
  r = length(mu)
  L = t(chol(Sigma))
  Z = rnorm(r)
  return(L %*% Z + mu)
}

n = 200
m = 100

Sigma = diag(m)
```

```
mu = rep(0,m)

X = matrix(0,m,n)
for(i in 1:n){
 X[,i] = my_rmvnorm(mu, Sigma)
}
X
X = t(X)

x = rep(0,m)
for(i in 1:m){
  x[i] = mean(X[,i])
}
x
mean(x)
mean(X)

sigma = matrix(0.9,m,m) + 0.1*diag(m)

A = t(chol(sigma))
Y = A%*%t(X)

Y
plot(t(Y))

hist(var(Y), main='Histogram prbkowych wariancji')


a = matrix(0, 100, 100)
for (i in 1:100){
  for (j in 1:100){
    a[i,j] = cov(Y[i,], Y[j,])
  }
}
a
hist(a, main='Kowariancja prbkowa')
```