Econometrics Homework 2

b10901069 Chinying Lin

September 2024

Problem 1

Example 1

$Y = (Y_1, Y_2)^\top$
Y_1 is a random variable that takes values 1 and -1 with equal probability.
$Y_2 = Y_1^2 - 1$, that is to say, $Y_2 = 0$.
The mean vector μ is $(0, 0)^\top$.
The covariance matrix Σ is diagonal since Y_2 is always equal to zero, so the covariance between Y_1 and $Y_2 = 0$. However, Y is not a vector of independent variables because Y_2 is determined by Y_1.

Example 2

$Y = (Y_1, Y_2)^\top$
Y_1 and Y_2 are conditionally independent given a latent variable Z, Z is a Bernoulli random variable which takes value 0 or 1 with probability.

- If $Z = 0$, then (Y_1, Y_2) are independent normal variables with mean 0 and variance 1.
- If $Z = 1$, then (Y_1, Y_2) are independent normal variables with mean 3 and variance 1.

The covariance matrix Σ for Y_1 and Y_2 will be diagonal since they are independent given Z. However, marginally Y_1 and Y_2 are not independent, because their distributions are influenced by the common variable Z. The values of Y_1 and Y_2 will be correlated through Z, even though the covariance matrix Σ is diagonal.

Problem 2

if $\Sigma = diag(\sigma_1^2, \sigma_2^2, ..., \sigma_m^2)$ is a diagonal matrix and $Y \sim N(\mu, \Sigma)$, then $cov(Y_i, Y_j) = 0$ for $i \neq j$, that is, the different component of Y are unrelated, and for normal distribution, unrelated implies independent. Therefore, it is proved that random variables $Y_1, Y_2, ..., Y_m$ are independent , indicating Y is a vector of independent variables.

Problem 3

Now, $Y \sim N(0, \sigma^2 I_m)$, then $f(y) = \frac{1}{(2\pi\sigma^2)^{m/2}} \exp\left(-\frac{1}{2\sigma^2} y^T y\right)$.

$Y'Y = \Sigma_{i=1}^m Y_i^2$
$\implies E[Y'Y] = E[\Sigma_{i=1}^m Y_i^2] = \Sigma_{i=1}^m E[Y_i^2] = m\sigma^2$.

$(Y'Y)^2 = (\Sigma_{i=1}^m Y_i^2)^2 = \Sigma_{i=1}^m Y_i^4 + 2\Sigma_{1\le i<j\le m} Y_i^2 Y_j^2$.
For $E[Y_i^4] = 3\sigma^4$, $\Sigma_{i=1}^m E[Y_i^4] = 3m\sigma^4$.
For $E[Y_i^2 Y_j^2], i \neq j$, we can simplify by $E[Y_i^2 Y_j^2] = E[Y_i^2]E[Y_j^2] = \sigma^4$,
then $2\Sigma_{1\le i<j\le m} E[Y_i^2 Y_j^2] = 2\Sigma_{1\le i<j\le m} E[Y_i^2]E[Y_j^2] = m(m-1)\sigma^4$.
$\implies E[(Y'Y)^2] = 3m\sigma^4 + m(m-1)\sigma^4 = (m^2+2m)\sigma^4 = m(m+2)\sigma^4$.

Problem 4

Environmental Variables Setup

```
Data <- read.csv("Equity_Premium.csv")
Matrix <- as.matrix(Data)
TimeMatrix <- Matrix[, 1, drop = FALSE]
Y <- Matrix[, 2, drop = FALSE]
X1 <- Matrix[, 3, drop = FALSE]
X2 <- Matrix[, 6, drop = FALSE]
X3 <- Matrix[, 9, drop = FALSE]
```

Alpha

```
SA_Y <- mean(Y)
indices1 <- X1 > 0.015
indices2 <- X1 <= 0.015
Y_a1 <- Y[indices1, , drop = FALSE]
Y_a2 <- Y[indices2, drop = FALSE]
SA_Y1 <- mean(Y_a1)
SA_Y2 <- mean(Y_a2)

a1 <- (SA_Y - SA_Y2) / (SA_Y1 - SA_Y2)
a2 <- 1 - a1
```

```
> print(a1)
[1] 0.09126984
> print(a2)
[1] 0.9087302
```

Beta

```
indices11 <- X1 > 0.015 & X2 > 0.02
indices12 <- X1 > 0.015 & X2 <= 0.02
indices21 <- X1 <= 0.015 & X2 > 0.02
indices22 <- X1 <= 0.015 & X2 <= 0.02
Y_b11 <- Y[indices11, , drop = FALSE]
Y_b12 <- Y[indices12, , drop = FALSE]
Y_b21 <- Y[indices21, , drop = FALSE]
Y_b22 <- Y[indices22, , drop = FALSE]
b11 <- length(Y_b11) / length(Y)
b12 <- length(Y_b12) / length(Y)
b21 <- length(Y_b21) / length(Y)
b22 <- length(Y_b22) / length(Y)
```

```
> print(b11)
[1] 0.05555556
> print(b12)
[1] 0.03571429
> print(b21)
[1] 0.4781746
> print(b22)
[1] 0.4305556
```

Gamma

```
indices31 <- X1 > 0.015 & X3 > -4
indices32 <- X1 <= 0.015 & X3 > -4
indices3 <- X3 > -4
Y_g <- Y[indices3, , drop = FALSE]
Y_g1 <- Y[indices31, , drop = FALSE]
Y_g2 <- Y[indices32, , drop = FALSE]
g1 <- length(Y_g1) / length(Y_g)
g2 <- length(Y_g2) / length(Y_g)
```

```
> print(g1)
[1] 0.1348974
> print(g2)
[1] 0.8651026
```

Rounding to 3 Decimal Place

```
> print(round(a1, 3))
[1] 0.091
> print(round(a2, 3))
[1] 0.909
> print(round(b11, 3))
[1] 0.056
> print(round(b12, 3))
[1] 0.036
> print(round(b21, 3))
[1] 0.478
> print(round(b22, 3))
[1] 0.431
> print(round(g1, 3))
[1] 0.135
> print(round(g2, 3))
[1] 0.865
```

Problem 5

According to $E[(m(X) - X'b)^2]$, $m(X) : 1 * k; b : scalar$.
$E[(m(X) - X^T b)^2]$
$= E[m(X)^2] - 2bE[m(X)X] + b^2 E[XX^T]$
To find its minimum, $\frac{\partial E[(m(X)-X'b)^2]}{\partial b} = 0$ should be satisfied,
that is, $\frac{\partial E[(m(X)-X'b)^2]}{\partial b} = -2E[m(X)X] + 2E[XX^T]b = 0$. Hence, $E[XX^T]b = E[m(X)X]$ and the condition of $E[XX']$ is positive definite is given, so $E[XX']$ has an inverse.
$\implies b = E[XX^T]^{-1}E[E[Y|X]X] = E[XX^T]^{-1}E[E[XY|X]] = E[XX^T]^{-1}E[XY]$

Problem 6

The degrees-of-freedom (DF) setting is essential to define the linear projection coefficient β because it is determined by the covariance between X and Y, as well as the variance of X. For these reasons, X must have a distribution with finite variance.
X is a random variable which is t-distribution with $\nu = 3$, so the probability density function (PDF) of X is given by $f(t) = \frac{\Gamma(2)}{\sqrt{3\pi}\Gamma(\frac{3}{2})}(1 + \frac{t^2}{3})^{-2} = \frac{2}{\sqrt{3\pi}}(1 + \frac{t^2}{3})^{-2}$, with Var(X) $= \frac{\nu}{\nu-2} = 3$.
In simple linear regression, the linear projection coefficient is given by $\beta = \frac{Cov(X,Y)}{Var(X)} = \frac{Cov(X, \frac{1}{1+X^4})}{3}$, apply the function in R script :

```
library(MASS)
n <- 10000
tX <- rt(n, df = 3)
tY <- 1 / (1 + tX^4)
beta <- cov(tX, tY)/3
```

```
> print(beta)
[1] -0.0003361182
```

GitHub Repo

EconometricMethods-homework2-b10901069