QUIZ 2 - QUESTION 2

Amanjit Gill

November 8, 2022

```
library(ggm)
```

```
## Warning: package 'ggm' was built under R version 4.2.2
```

```
Sigma <- matrix(
    c(
        177, 40, -14, 4,
        40, 98, -37, -4,
        -14, -37, 314, 5,
        4, -4, 5, 1
    ),
    4, 4
)

# PART A

(r12.34 <- pcor(c(1,2,3,4), Sigma))
```

```
## [1] 0.4788521
```

```
# PART A manually

Sig11 <- Sigma[1:2, 1:2]
Sig12 <- Sigma[1:2, 3:4]
Sig21 <- Sigma[3:4, 1:2]
Sig22 <- Sigma[3:4, 3:4]

Sig12.34 <- Sig11 - Sig12 %*% solve(Sig22) %*% Sig21

(r12.34 <- Sig12.34[1,2] / sqrt(Sig12.34[1,1] * Sig12.34[2,2]))
```

```
## [1] 0.4788521
```

```
# PART B

C <- Sigma[2:4, 2:4]

sig0 <- Sigma[1,2:4]

sigYY <- Sigma[1,1]

(r1.234 <- sqrt(t(sig0) %*% solve(C) %*% sig0 / sigYY))
```

```
##           [,1]
## [1,] 0.5624804
```

```r
# PART C

# perform eigendecomposition
eig <- eigen(Sigma)

# extract variances and PC vectors
p <- nrow(Sigma)
sum_vars <- sum(eig$values)

vars <- c()
PCs <- list()
var_conts <- c()

for (i in 1:p) {
    vars <- c(vars, eig$values[i])
    PCs[[i]] <- eig$vectors[,i]
    var_conts <- c(var_conts, vars[i]/sum_vars*100)
}

PCs
```

```
## [[1]]
## [1]  0.14338438  0.18544166 -0.97201274 -0.01560591
##
## [[2]]
## [1] 0.92368901 0.32727505 0.19840428 0.01803896
##
## [[3]]
## [1]  0.35356243 -0.92514658 -0.12528191  0.05832539
##
## [[4]]
## [1] -0.03511620  0.05105128 -0.01146382  0.99801263
```

```r
# choose k using 90% variance explained
var_conts.c <- cumsum(var_conts)
(min(which(var_conts.c >= 90)))
```

```
## [1] 3
```

```r
# choose k using Kaiser's rule
# since Sigma isn't standardised (yet), use an alternative definition
(max(which(vars > mean(vars))))
```

```
## [1] 2
```

```r
# REDO PCA WITH STANDARDISED SIGMA

# standardised Sigma is same as correlation matrix
Sigma.std <- cov2cor(Sigma)

# now redo PCA
eig.std <- eigen(Sigma.std)

# compute variance contribution for each variable
sum_vars.std <- sum(eig.std$values)
```

```r
vars.std <- c()
var_conts.std <- c()

for (i in 1:p) {
    vars.std <- c(vars.std, eig.std$values[i])
    var_conts.std <- c(var_conts.std, vars.std[i]/sum_vars.std)
}

var_conts.std
```

```
## [1] 0.40166487 0.31773379 0.20167561 0.07892574
```