STA 4702 – Spring 2019 – Homework 1

Part 1: NFL 2014 Combine Data

Download the NFL 2014 Combine data, and select players who are Linebackers. There are 4 variables in the dataset: Height (inches), Weight (pounds), HandLength (inches), and ArmLength (inches).

p.1.a. Obtain a scatterplot matrix of the 4 variables using the **plot()** function.

p.1.b. Obtain the mean vector, the variance-covariance matrix ($\mathbf{S_n}$), and the correlation matrix of the 4 variables, using the **mean()**, **cov()**, and **cor()** functions. What is the angle θ between the Height and Weight vectors?

p.1.c. Obtain the eigenvalues and eigenvectors of the variance-covariance matrix, demonstrate the spectral decomposition of the variance-covariance matrix, and obtain its square-root matrix.

p.1.d. Obtain the following matrices and vectors directly and use them to reconstruct the covariance and correlation matrices.

Let $\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_p \end{bmatrix}$ where $\mathbf{y}_i = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ni} \end{bmatrix}$

Obtain: $\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{bmatrix}$ $\mathbf{d}_i = \mathbf{y}_i - \bar{x}_i \mathbf{1}_n \quad i = 1, \ldots, p$ $\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\mathbf{X} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \cdots & \mathbf{d}_p \end{bmatrix}$

$\mathbf{S} = \frac{1}{n-1}\mathbf{X}'\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\mathbf{X}$ $\mathbf{D}^{1/2}$ $\mathbf{D}^{-1/2}$ $\mathbf{R}$ Generalized Sample Variance: $|\mathbf{S}|$

Show $\left(\bar{x}_1 \mathbf{1}_n\right)'\mathbf{d}_1 = 0$ (to "machine rounding")

p.1.e. Obtain the mean, variance, and standard deviation of the sum of HandLength and ArmLength: directly from the data, and in matrix form based on the mean vector and the variance-covariance matrix.

The following program will read in the data:

```
nfl2014 <- read.csv("http://www.stat.ufl.edu/~winner/sta4702/nfl_combine_2014.csv")
attach(nfl2014); names(nfl2014)

Height_LB <- Height[Position=="LB"]
Weight_LB <- Weight[Position=="LB"]
HandLen_LB <- HandLen[Position=="LB"]
ArmLen_LB <- ArmLen[Position=="LB"]

nfl2014_LB <- data.frame(Height_LB, Weight_LB, HandLen_LB, ArmLen_LB)

detach(nfl2014)
```

The following program was used on the Offensive Linemen (Position=”OL”). Make adjustments to it for the Linebacker data.

```
nfl2014 <- read.csv("http://www.stat.ufl.edu/~winner/sta4702/nfl_combine_2014.csv")
attach(nfl2014); names(nfl2014)

Height_OL <- Height[Position=="OL"]
Weight_OL <- Weight[Position=="OL"]
HandLen_OL <- HandLen[Position=="OL"]
ArmLen_OL <- ArmLen[Position=="OL"]

nfl2014_OL <- data.frame(Height_OL, Weight_OL, HandLen_OL, ArmLen_OL)

detach(nfl2014)
attach(nfl2014_OL)


(n_OL <- length(Height_OL))
one_OL <- rep(1,n_OL)
X_OL <- cbind(Height_OL, Weight_OL, HandLen_OL, ArmLen_OL)
(S_n <- ((n_OL-1)/n_OL) * cov(nfl2014_OL))
(R_n <- cor(nfl2014_OL))

(xbar_OL <- (1/n_OL) * t(X_OL) %*% one_OL)

xbar1_OL <- rep(xbar_OL[1,], n_OL)
xbar2_OL <- rep(xbar_OL[2,], n_OL)
xbar3_OL <- rep(xbar_OL[3,], n_OL)
xbar4_OL <- rep(xbar_OL[4,], n_OL)

d1_OL <- X_OL[,1] - xbar1_OL
d2_OL <- X_OL[,2] - xbar2_OL
d3_OL <- X_OL[,3] - xbar3_OL
d4_OL <- X_OL[,4] - xbar4_OL

t(xbar1_OL) %*% d1_OL
t(xbar2_OL) %*% d2_OL
t(xbar3_OL) %*% d3_OL
t(xbar4_OL) %*% d4_OL

J_n <- matrix(rep(1,n_OL^2),ncol=n_OL)
I_n <- diag(n_OL)

dmat_OL <- (I_n - (1/n_OL)*J_n) %*% X_OL

(SS_dmat_OL <- t(dmat_OL) %*% dmat_OL)
(S_dmat_OL <- (1/(n_OL-1)) * SS_dmat_OL)

(D12_dmat_OL <- diag(sqrt(diag(S_dmat_OL))))

(R_dmat_OL <- solve(D12_dmat_OL) %*% S_dmat_OL %*% solve(D12_dmat_OL))

eigen_OL <- eigen(S_dmat_OL)

(lambda_OL <- eigen_OL$val)
(P_OL <- eigen_OL$vec)

(Lambda_OL <- diag(lambda_OL))
(Lambda12_OL <- diag(sqrt(lambda_OL)))
```

Part 2: Textbook Problems: Chapter 2: 15, 17, 18, 32, 34, 35, 36, 41.