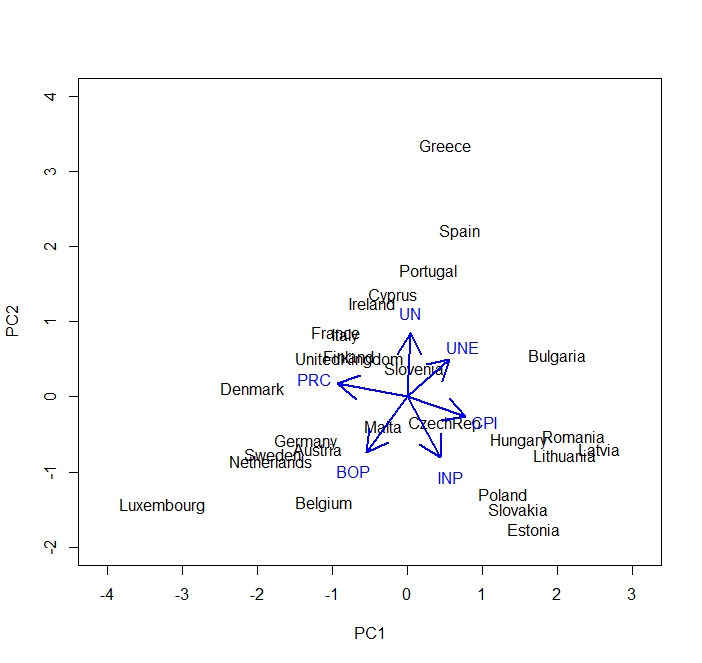
### Exercise 2  
  
remove(list = ls())  
  
euro = read.table(file = "Europe.txt", header = T, dec =".")  
attach(euro)  
europe = data.frame(euro)  
europe = na.omit(europe)  
ls(europe)

## [1] "BOP" "CPI" "ï..Country" "INP" "PRC"   
## [6] "UN" "UNE"

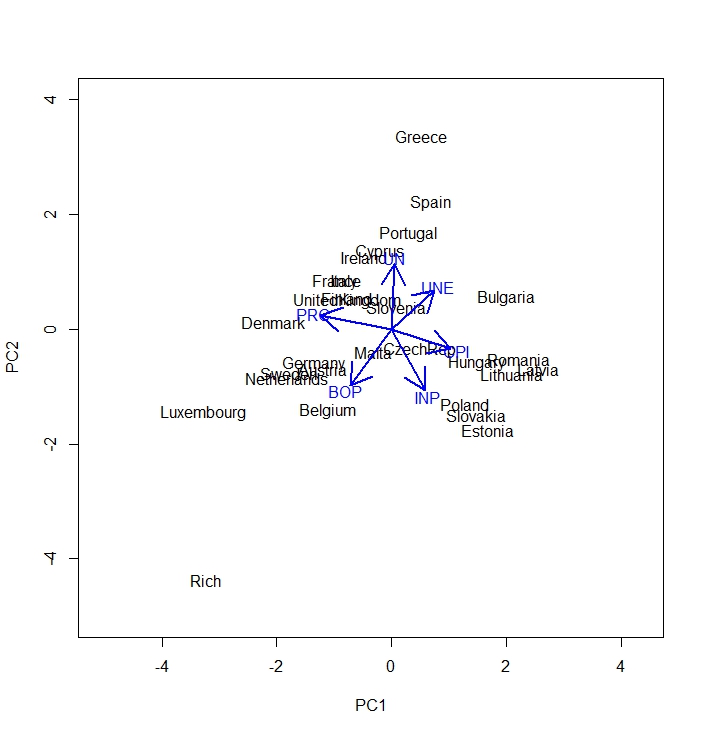
Country = ï..Country  
numbers = cbind(CPI, UNE, INP, BOP, PRC, UN)  
  
# a)   
R = cor(numbers)  
E = eigen(R)$vectors  
  
X = scale(numbers, center = T, scale = T)  
Yhat = X%\*%E  
y1hat = Yhat[,1]\*(-1)  
y2hat = Yhat[,2]  
  
plot(y1hat, y2hat, type = "n", asp = 1, xlab = "PC1", ylab = "PC2", xlim = c(-4,3), ylim = c(-2, 4))  
text(y1hat, y2hat, Country)  
  
arrows(0,0,-E[,1]\*1.5,E[,2]\*1.5, lwd = 2, col = "blue")  
text(-E[,1]\*2, E[,2]\*2, c("CPI","UNE","INP","BOP","PRC","UN"), col = "blue")



# b)  
Rich = data.frame(ï..Country = "Rich", CPI = mean(CPI), UNE = mean(UNE), INP = mean(INP), BOP = max(BOP)\*3, PRC = mean(PRC), UN = mean(UN))  
rbind.data.frame(europe, Rich)

## ï..Country CPI UNE INP BOP PRC UN  
## 1 Belgium 116.0300 4.770000 125.5900 908.6 6716.500 -1.6000000  
## 2 Bulgaria 141.2000 7.310000 102.3900 27.8 1094.700 3.5000000  
## 3 CzechRep 116.2000 4.880000 119.0100 -277.9 2616.400 -0.6000000  
## 4 Denmark 114.2000 6.030000 88.2000 1156.4 7992.400 0.5000000  
## 5 Germany 111.6000 4.630000 111.3000 499.4 6774.600 -1.3000000  
## 6 Estonia 135.0800 9.710000 111.5000 153.4 2194.100 -7.7000000  
## 7 Ireland 106.8000 10.200000 111.2000 -166.5 6525.100 2.0000000  
## 8 Greece 122.8300 11.300000 78.2200 -764.1 5620.100 6.4000000  
## 9 Spain 116.9700 15.790000 83.4400 -280.8 4955.800 0.7000000  
## 10 France 111.5500 6.770000 92.6000 -337.1 6828.500 -0.9000000  
## 11 Italy 115.0000 5.050000 87.8000 -366.2 5996.600 -0.5000000  
## 12 Cyprus 116.4400 5.140000 86.9100 -1090.6 5310.300 -0.4000000  
## 13 Latvia 144.4700 12.110000 110.3900 42.3 1968.300 -3.6000000  
## 14 Lithuania 135.0800 11.470000 114.5000 -77.4 2130.600 -4.3000000  
## 15 Luxembourg 118.1900 3.140000 85.5100 2016.5 10051.600 -3.0000000  
## 16 Hungary 134.6600 6.770000 115.1000 156.2 1954.800 -0.1000000  
## 17 Malta 117.6500 4.150000 101.6500 359.4 3378.300 -0.6000000  
## 18 Netherlands 111.1700 3.230000 103.8000 1156.6 6046.000 -0.4000000  
## 19 Austria 114.1000 2.990000 116.8000 87.8 7045.500 -1.5000000  
## 20 Poland 119.9000 6.280000 146.7000 -74.8 2124.200 -1.0000000  
## 21 Portugal 113.0600 9.680000 89.3000 -613.4 4073.600 0.8000000  
## 22 Romania 142.3400 4.760000 131.8000 -128.7 1302.200 3.2000000  
## 23 Slovenia 118.3300 5.560000 105.4000 39.4 3528.300 1.8000000  
## 24 Slovakia 117.1700 9.190000 156.3000 16.0 2515.300 -2.1000000  
## 25 Finland 114.6000 5.920000 101.0000 -503.7 7198.800 -1.3000000  
## 26 Sweden 112.7100 6.100000 100.5000 1079.1 7476.700 -2.3000000  
## 27 UnitedKingdom 120.9000 6.110000 90.3600 -24.3 6843.900 -0.8000000  
## 28 Rich 120.6752 7.001481 106.1952 6049.5 4824.563 -0.5592593

XRich = c(0,0,0,(3\*max(BOP)-mean(BOP))/sd(BOP),0,0)  
YhatRich = XRich%\*%E  
plot(y1hat, y2hat, type = "n", asp = 1, xlab = "PC1", ylab = "PC2",ylim = c(-5,4))  
text(y1hat,y2hat, Country)  
text(-YhatRich[1,1],YhatRich[1,2], "Rich")  
arrows(0,0,-E[,1]\*2,E[,2]\*2, col = "blue", lwd = 2)  
text(-E[,1]\*2.2,E[,2]\*2.2, c("CPI","UNE","INP","BOP","PRC","UN"), col = "blue")



### Exercise 3  
  
remove(list = ls())  
  
euro = read.table(file = "Europe.txt", header = T, dec =".")  
attach(euro)  
europe = data.frame(CPI, UNE, INP, BOP, PRC, UN)  
europe = na.omit(europe)  
ls(europe)

## [1] "BOP" "CPI" "INP" "PRC" "UN" "UNE"

X = scale(europe, center = T, scale = T)  
R = cor(europe)  
E = eigen(R)$vectors  
  
# i)  
A1hat = X%\*%E[,1]%\*%t(E[,1])  
head(A1hat,5)

## [,1] [,2] [,3] [,4] [,5] [,6]  
## 1 -0.5660436 -0.4131536 -0.3217987 0.4032920 0.6883677 -0.02347820  
## 2 1.0154376 0.7411651 0.5772815 -0.7234740 -1.2348774 0.04211804  
## 3 0.2549705 0.1861023 0.1449520 -0.1816601 -0.3100706 0.01057559  
## 4 -1.0526220 -0.7683059 -0.5984211 0.7499670 1.2800975 -0.04366036  
## 5 -0.6885770 -0.5025904 -0.3914596 0.4905940 0.8373810 -0.02856060

AE1 = (dim(X)[1]-1)\*sum(eigen(R)$values[2:6])  
AE1

## [1] 97.11576

# ii)  
A2hat = X%\*%E[,1:2]%\*%t(E[,1:2])  
head(A2hat, 5)

## [,1] [,2] [,3] [,4] [,5] [,6]  
## 1 -0.3242375 -0.89125343 0.43842950 1.10504305 0.5172357 -0.82318795  
## 2 0.9231616 0.92361347 0.28716965 -0.99127032 -1.1695714 0.34729651  
## 3 0.3139868 0.06941505 0.33049678 -0.01038758 -0.3518378 -0.18460519  
## 4 -1.0729105 -0.72819144 -0.66220720 0.69108732 1.2944562 0.02343844  
## 5 -0.5872209 -0.70299193 -0.07280039 0.78474171 0.7656489 -0.36376896

AE2 = (dim(X)[1]-1)\*sum(eigen(R)$values[3:6])  
AE2

## [1] 57.18878

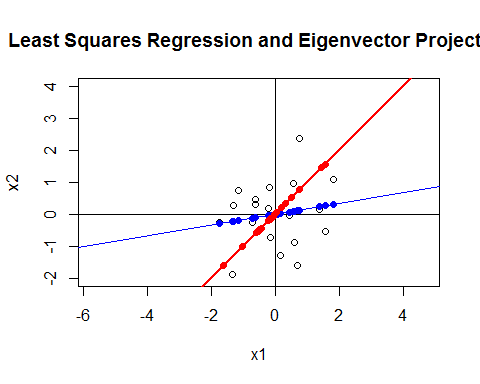
# iii)  
A3hat = X%\*%E[,1:3]%\*%t(E[,1:3])  
head(A3hat, 5)

## [,1] [,2] [,3] [,4] [,5] [,6]  
## 1 -0.3856722 -1.1109001 0.5970611 1.0433610 0.4527950 -0.5999865  
## 2 0.7990121 0.4797432 0.6077384 -1.1159197 -1.2997955 0.7983505  
## 3 0.1447466 -0.5356673 0.7674949 -0.1803092 -0.5293588 0.4302699  
## 4 -1.0107256 -0.5058625 -0.8227760 0.7535226 1.3596837 -0.2024887  
## 5 -0.6380826 -0.8848371 0.0585305 0.7336752 0.7122985 -0.1789808

AE3 = (dim(X)[1]-1)\*sum(eigen(R)$values[4:6])  
AE3

## [1] 33.29049

### Exercise 4  
  
remove(list = ls())  
  
appro = read.table(file = "Approx2dim.txt", header = T, dec =".")  
attach(appro)  
x1 = ï..x1  
Approx2 = data.frame(x1,x2)  
Approx2 = na.omit(Approx2)  
  
plot(x1,x2, asp = 1, xlim = c(-4,3), ylim = c(-2, 4), main = "Least Squares Regression and Eigenvector Projection")  
abline(h = 0, v = 0)  
  
reslm = lm(x2 ~ x1 - 1)  
x2hatlm = reslm$fitted.values  
  
abline(reslm, col= "blue")  
points(x1,x2hatlm, pch = 16, col = "blue")  
  
S = cov(Approx2)  
E = eigen(S)$vectors  
X = scale(Approx2, center = F, scale = T)  
  
Ahat = X%\*%E[,1]%\*%t(E[,1])  
reslm2 = lm(Ahat[,2] ~ Ahat[,1] -1)  
x2hatlm2 = reslm2$fitted.values  
  
abline(reslm2, col = "red", lwd = 2)  
points(Ahat[,1], x2hatlm2, pch = 16, col ="red")



#Approximation Error for OLS = sum of residuals  
sum((reslm$residuals)^2)

## [1] 18.46061

# Approximation Error  
AE = (dim(X)[1]-1)\*sum(eigen(S)$values[2])  
AE

## [1] 15.79869