Week9Assignment

John Navarro

December 3, 2016

Part 1 Data

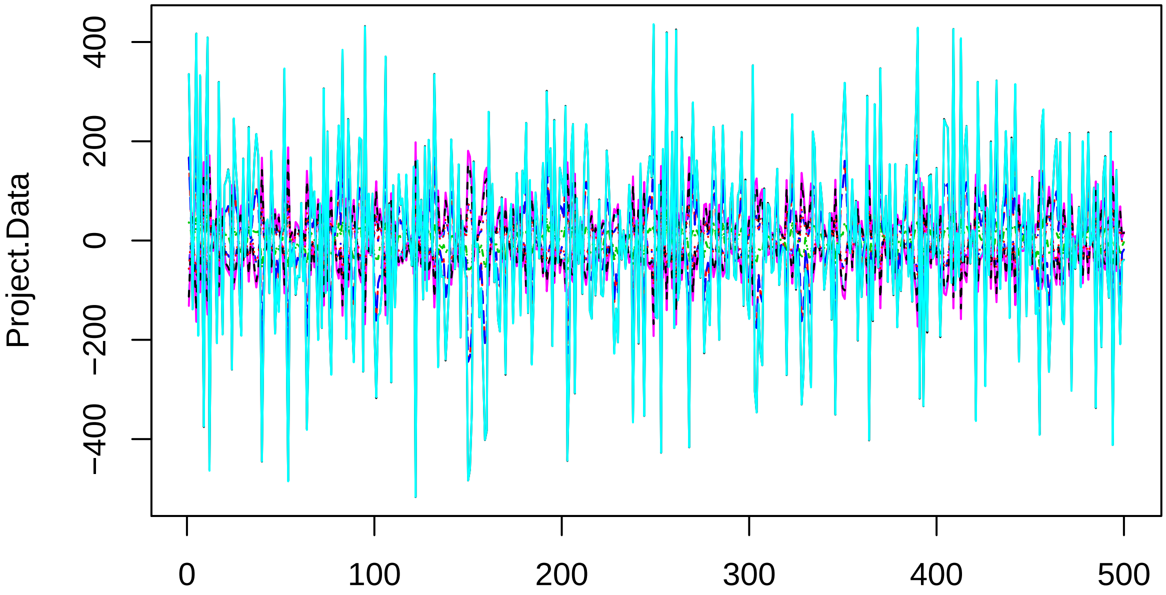
```
Project.Data<-read.csv(file="C:/Users/JohntheGreat/Documents/MSCA/StatisticalAnalysis/Week9/Assignments/
 Project.Data[1:10,]
```

```
##            V1         V2        V3         V4         V5         V6
## 1   -10.00000   -8.00000  -6.00000  -4.000000   -2.00000    0.00000
## 2   335.26043  167.30996  36.42212 -57.023949 -112.88559 -131.66066
## 3   106.04431   61.94413  22.68583  -4.491896  -16.86543  -23.94030
## 4  -138.25926  -69.13931 -14.48162  24.403158   47.02190   55.09567
## 5   139.51863   66.76133  11.15244 -28.738979  -53.45132  -61.10555
## 6   416.99904  207.59255  45.27888 -70.756695 -140.82071 -163.84332
## 7  -190.35076 -104.67105 -34.61287  14.833058   41.78893   52.80835
## 8   331.03604  177.06454  52.22067 -36.066936  -85.00319 -104.34305
## 9    18.52366   16.95822  13.74827  11.802604   12.21568   11.16780
## 10 -375.72268 -182.05494 -34.09990  72.048443  137.85975  158.20484
##            V7         V8        V9        V10        V11
## 1     2.00000   4.000000   6.00000    8.00000   10.00000
## 2  -113.07746 -56.864257  36.48108  167.10119  335.37523
## 3   -20.52873  -1.442949  23.81151   57.95828  108.23602
## 4    47.68518  23.851115 -14.68544  -68.41763 -138.65609
## 5   -52.72716 -29.341695  10.92991   67.54925  139.08538
## 6  -140.40840 -71.099858  45.15218  208.04116  416.75236
## 7    44.31460  12.730956 -35.38898 -101.92300 -191.86184
## 8   -88.76262 -32.937981  53.37590  172.97410  333.28527
## 9    10.74362  13.027791  14.20061   15.35655   19.40438
## 10  135.88303  73.693656 -33.49248 -184.20571 -374.54004
```

```
#Separate the first row of Project.Data
Data.Levels<-as.numeric(Project.Data[1,])
Project.Data<-Project.Data[-1,]
head(Project.Data)
```

```
##          V1         V2        V3         V4         V5         V6
## 2  335.2604  167.30996  36.42212 -57.023949 -112.88559 -131.66066
## 3  106.0443   61.94413  22.68583  -4.491896  -16.86543  -23.94030
## 4 -138.2593  -69.13931 -14.48162  24.403158   47.02190   55.09567
## 5  139.5186   66.76133  11.15244 -28.738979  -53.45132  -61.10555
## 6  416.9990  207.59255  45.27888 -70.756695 -140.82071 -163.84332
## 7 -190.3508 -104.67105 -34.61287  14.833058   41.78893   52.80835
##           V7         V8        V9        V10       V11
## 2 -113.07746 -56.864257  36.48108  167.10119  335.3752
## 3  -20.52873  -1.442949  23.81151   57.95828  108.2360
## 4   47.68518  23.851115 -14.68544  -68.41763 -138.6561
## 5  -52.72716 -29.341695  10.92991   67.54925  139.0854
## 6 -140.40840 -71.099858  45.15218  208.04116  416.7524
## 7   44.31460  12.730956 -35.38898 -101.92300 -191.8618
```

```
#Plot the data
matplot(Project.Data,type="l")
```

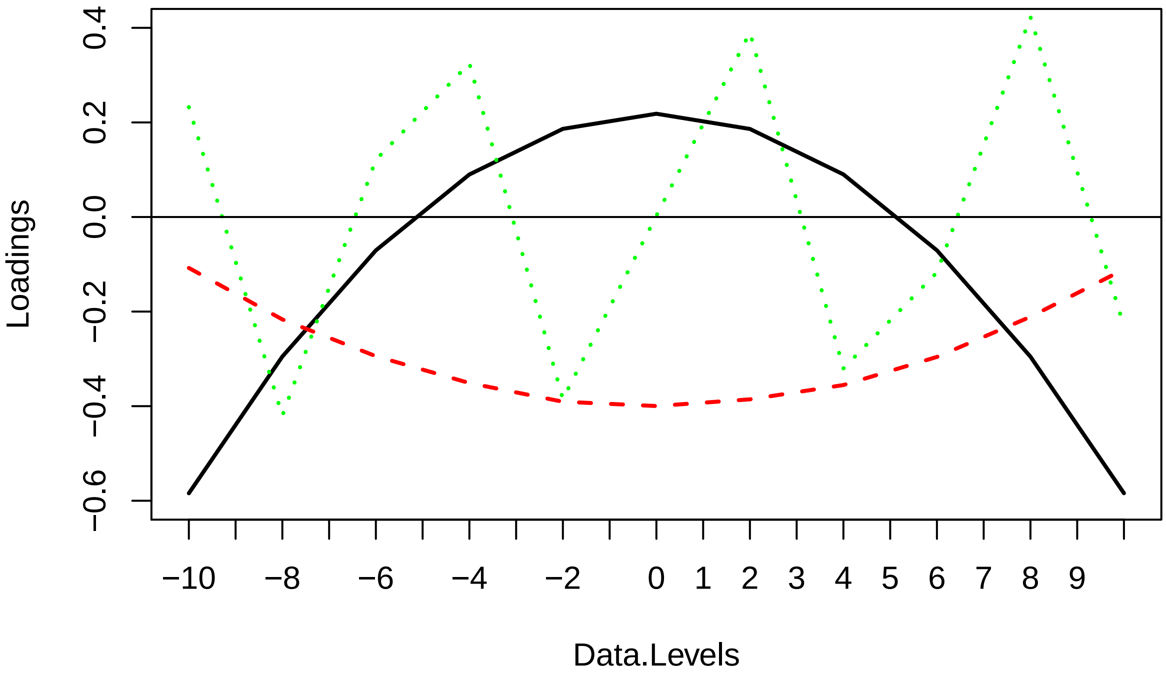


#Part 2 PCA

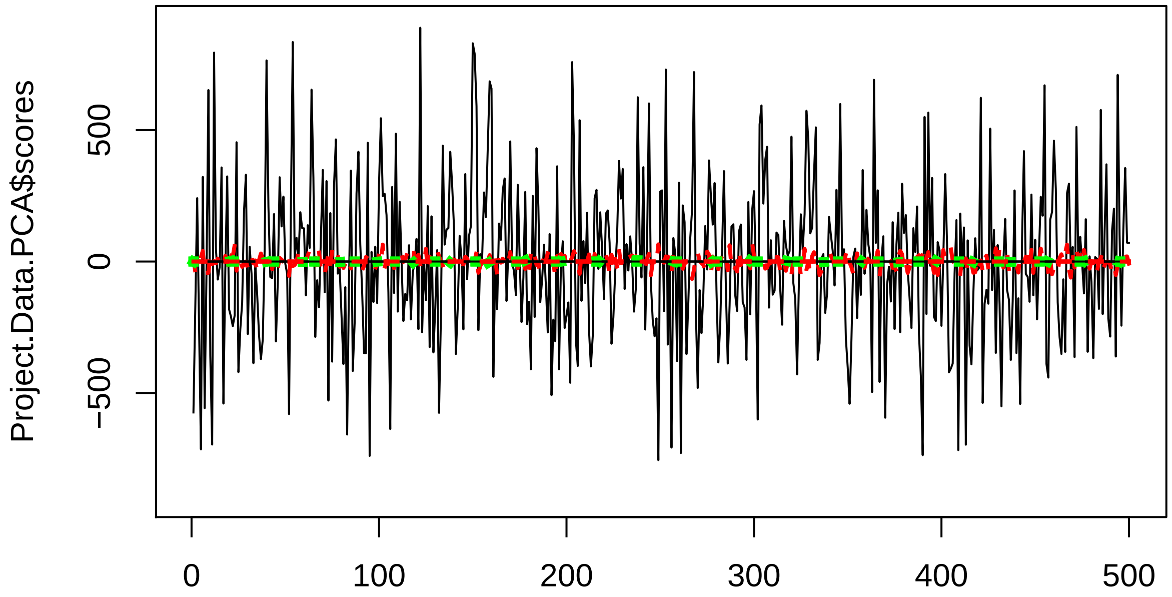
```
Project.Data.PCA <- princomp(Project.Data)
names(Project.Data.PCA)
```

```
## [1] "sdev"     "loadings" "center"   "scale"    "n.obs"    "scores"
## [7] "call"
```

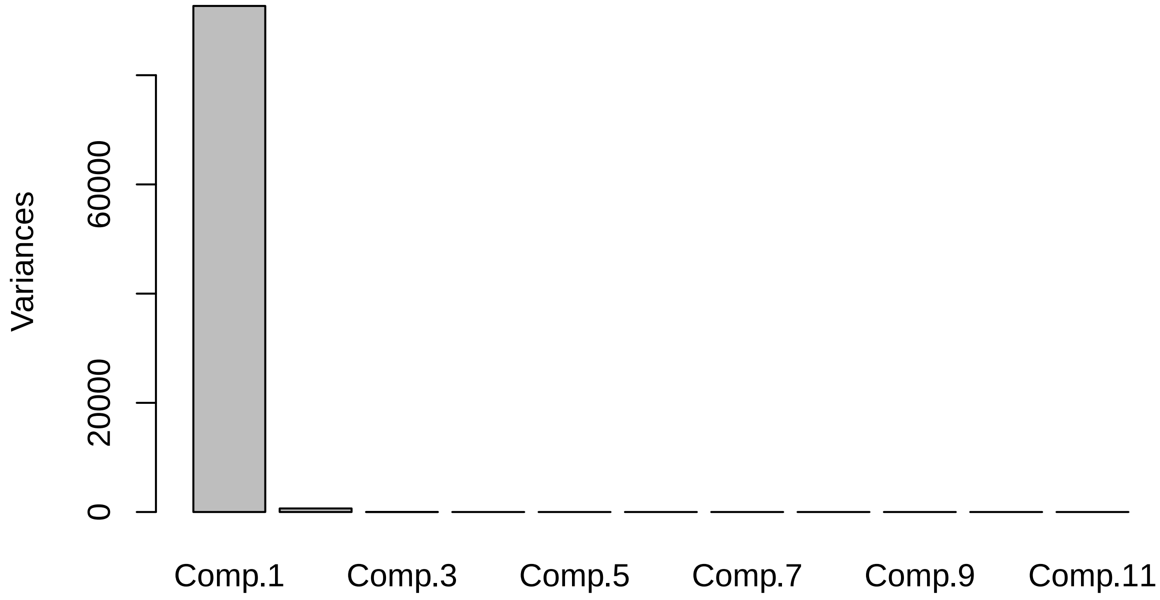
```
#Plot Loadings
matplot(Data.Levels,Project.Data.PCA$loadings[,1:3],type="l",lty=c(1,2,3),lwd=2,xaxt="n", xlab="Data.Le
abline(h=0)
#axis(1, -10:10,labels=colnames(swissPredictors))
axis(1, -10:10)
```



```
#Plot Factors
matplot(Project.Data.PCA$scores,type="l",lty=c(1,2,3),lwd=c(1,2,5),lend=2,ylim=c(-900,900),col=c("black
```



```
#Plot Variances
barplot(Project.Data.PCA$sdev^2, ylab = "Variances")
```



```
Project.Data.PCA$sdev^2
```

```
##       Comp.1       Comp.2       Comp.3       Comp.4       Comp.5 
## 9.268609e+04 6.489602e+02 5.576366e+00 3.053585e-11 3.062401e-12 
##       Comp.6       Comp.7       Comp.8       Comp.9      Comp.10 
## 1.785265e-12 1.371684e-12 0.000000e+00 0.000000e+00 0.000000e+00 
##      Comp.11 
## 0.000000e+00
```

```
Project.Data.PCA$sdev^2/sum(Project.Data.PCA$sdev^2)
```

```
##       Comp.1       Comp.2       Comp.3       Comp.4       Comp.5 
## 9.929877e-01 6.952601e-03 5.974211e-05 3.271442e-16 3.280887e-17 
##       Comp.6       Comp.7       Comp.8       Comp.9      Comp.10 
## 1.912634e-17 1.469547e-17 0.000000e+00 0.000000e+00 0.000000e+00 
##      Comp.11 
## 0.000000e+00
```

Estimate PCA using manual calculation with eigen(). For this recall the steps on slide 16 of the lecture notes.

```
#create centered matrix
centered.Project.Data <- scale(Project.Data, scale = FALSE)
cov.matrix <- cov(centered.Project.Data)
eigen.decomp <- eigen(cov.matrix)
```

Calculate 3 factor loadings using PCA and using manual method based on eigen-decomposition. Combine them in one matrix Project.Data.PCA.by.eigen.Loadings and compare

```
#eigen columns
Project.Data.PCA.Eigen.Loadings1 <- eigen.decomp$vectors[,1:3]

#bind with PCA columns
Project.Data.PCA.Eigen.Loadings1 <- cbind(Project.Data.PCA.Eigen.Loadings1,Project.Data.PCA$loadings[,1

#rename the columns
colnames(Project.Data.PCA.Eigen.Loadings1) <- c("L1.eigen", "L2.eigen", "L3.eigen", "L1.PCA", "L2.PCA",
head(Project.Data.PCA.Eigen.Loadings1)
```

```
##      L1.eigen   L2.eigen     L3.eigen     L1.PCA     L2.PCA       L3.PCA
## V1  0.5842042 -0.1077103 -0.232409041 -0.5842042 -0.1077103  0.232409041
## V2  0.2951056 -0.2165408  0.420358898 -0.2951056 -0.2165408 -0.420358898
## V3  0.0705508 -0.2943480 -0.120991690 -0.0705508 -0.2943480  0.120991690
## V4 -0.0898977 -0.3511162 -0.324880404  0.0898977 -0.3511162  0.324880404
## V5 -0.1864045 -0.3906019  0.384963011  0.1864045 -0.3906019 -0.384963011
## V6 -0.2183951 -0.3996945 -0.002667934  0.2183951 -0.3996945  0.002667934
```