Factor Analysis

Blake Pappas

December 17, 2023

Load the Stock Price Data

```r
url <- "http://users.stat.umn.edu/~sandy/courses/8053/Data/Wichern_data/T8-4.DAT"
stock <- read.table(url, sep = "\t", header = F)
colnames(stock) <- c("JP Morgan", "City Bank", "Wells Fargo", "Royal Dutch", "Exxon")
```

Summary Statistics

```r
(Xbar <- colMeans(stock)) # Calculates the mean vector
```

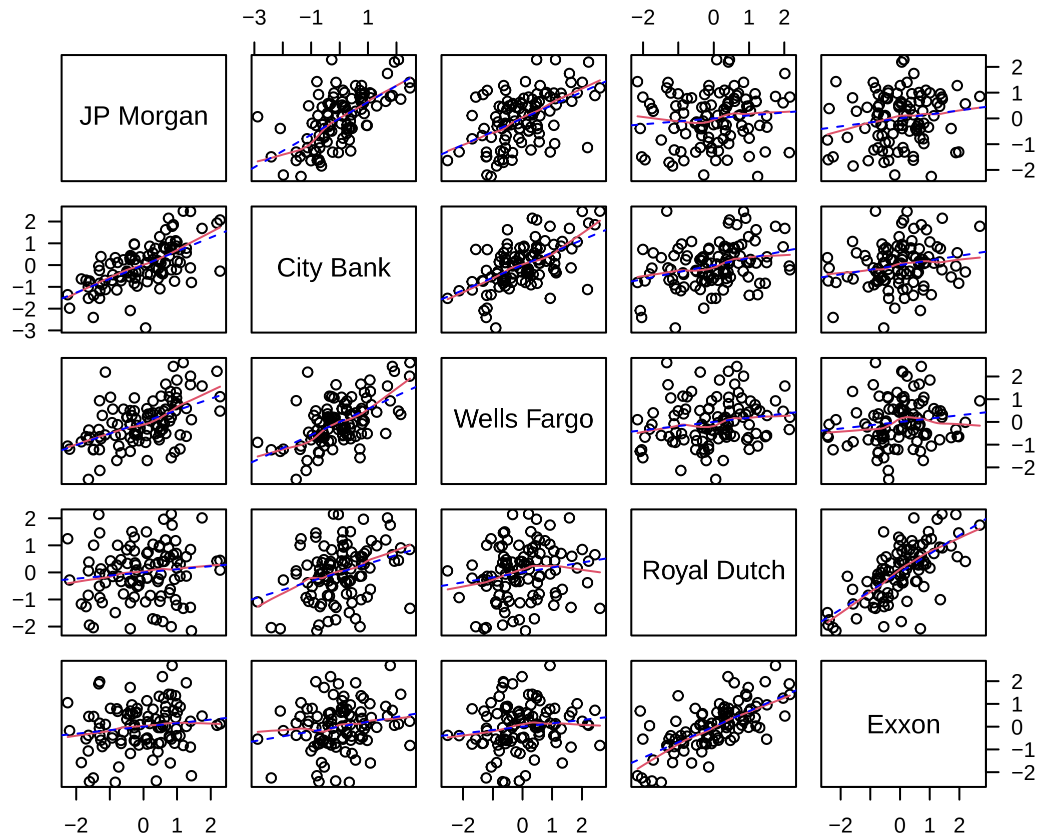
```
##    JP Morgan    City Bank  Wells Fargo  Royal Dutch        Exxon
## 0.0010627806 0.0006554204 0.0016260816 0.0040491252 0.0040386417
```

```r
(S <- cov(stock)) # Calculates the covariance matrix
```

```
##                JP Morgan    City Bank  Wells Fargo  Royal Dutch        Exxon
## JP Morgan   4.332695e-04 0.0002756679 1.590265e-04 6.411929e-05 8.896616e-05
## City Bank   2.756679e-04 0.0004387172 1.799737e-04 1.814512e-04 1.232623e-04
## Wells Fargo 1.590265e-04 0.0001799737 2.239722e-04 7.341348e-05 6.054612e-05
## Royal Dutch 6.411929e-05 0.0001814512 7.341348e-05 7.224964e-04 5.082772e-04
## Exxon       8.896616e-05 0.0001232623 6.054612e-05 5.082772e-04 7.656742e-04
```

Scatterplot Matrix

```r
stock_std <- scale(stock, center = T, scale = T)
stock_std <- as.data.frame(stock_std)
par(las = 1, mgp = c(2.4, 1, 0), mar = c(3.5, 3.5, 1, 0.6))
pairs(stock_std, labels = names(stock),
      panel = function(x, y){panel.smooth(x, y)
abline(lsfit(x, y), lty = 2, col = "blue")})
```



Compute PCs from the Correlation Matrix

```r
(s.cor <- var(stock_std)) # Calculates the correlation matrix
```

```
##             JP Morgan City Bank Wells Fargo Royal Dutch     Exxon
## JP Morgan   1.0000000 0.6322878   0.5104973   0.1146019 0.1544628
## City Bank   0.6322878 1.0000000   0.5741424   0.3222921 0.2126747
## Wells Fargo 0.5104973 0.5741424   1.0000000   0.1824992 0.1462067
## Royal Dutch 0.1146019 0.3222921   0.1824992   1.0000000 0.6833777
## Exxon       0.1544628 0.2126747   0.1462067   0.6833777 1.0000000
```

```r
s.pca <- prcomp(stock, scale = T, center = T) # Principal Component Analysis

s.pca$rotation
```

```
##                    PC1        PC2         PC3        PC4         PC5
## JP Morgan   -0.4690832  0.3680070 -0.60431522  0.3630228  0.38412160
## City Bank   -0.5324055  0.2364624 -0.13610618 -0.6292079 -0.49618794
## Wells Fargo -0.4651633  0.3151795  0.77182810  0.2889658  0.07116948
## Royal Dutch -0.3873459 -0.5850373  0.09336192 -0.3812515  0.59466408
## Exxon       -0.3606821 -0.6058463 -0.10882629  0.4934145 -0.49755167
```

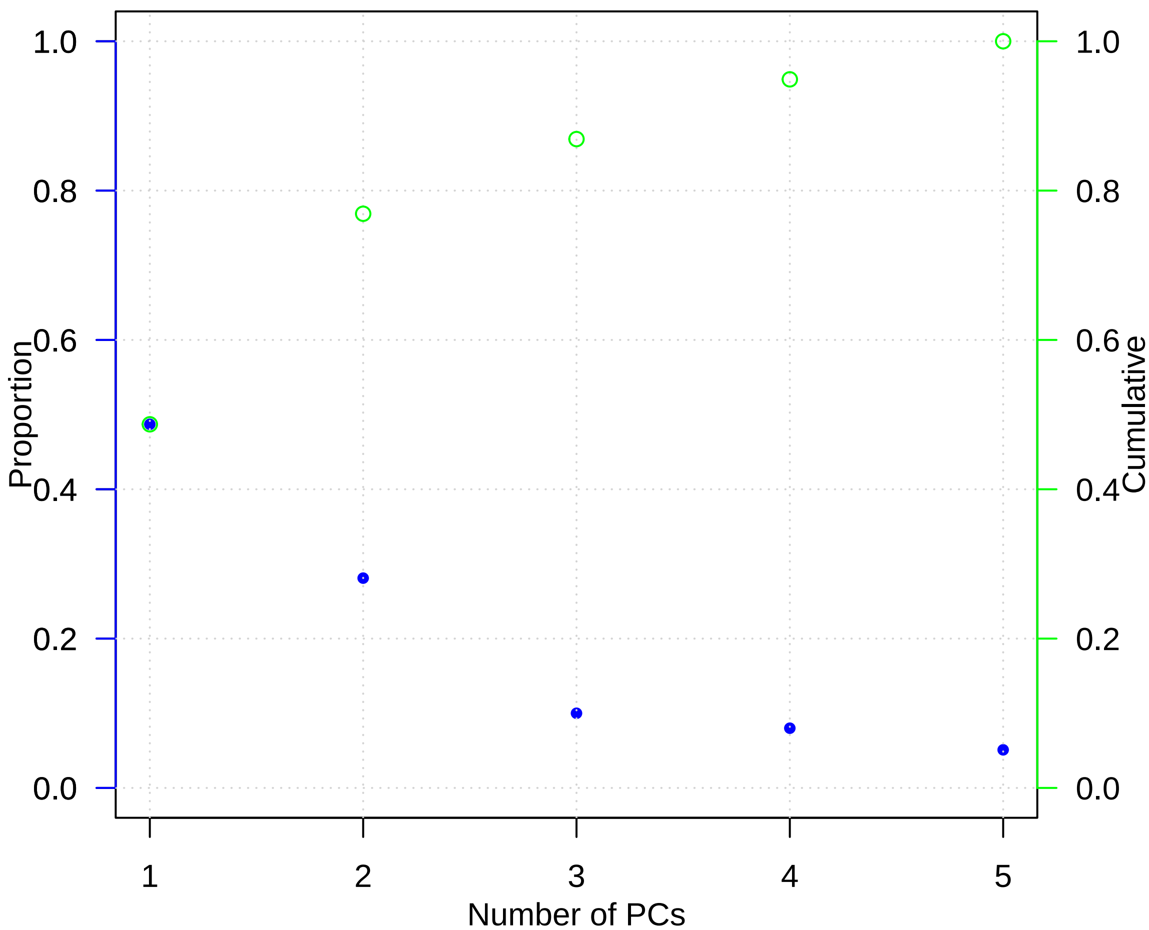
```r
s <- var(s.pca$x)
(Proportion.std <- round(diag(s) / sum(diag(s)), 3)) # Proportion of variability explained by each comp
```

```
##   PC1   PC2   PC3   PC4   PC5
## 0.487 0.281 0.100 0.080 0.051
```

```r
(Cumulative.std <- round(cumsum(diag(s)) / sum(diag(s)), 3)) # Cumulative variability explained by the
```

```
##   PC1   PC2   PC3   PC4   PC5
## 0.487 0.769 0.869 0.949 1.000
```

```r
# Screen Plot
p <- 5
par(las = 1, mgp = c(2, 1, 0), mar = c(3, 3, 1, 3))
plot(1:p, Proportion.std, xlab = "Number of PCs", ylim = c(0, 1),
ylab = "Proportion", pch = 16, cex = 0.8, xaxt = "n", col = "blue")
axis(1, at = 1:p)
mtext("Cumulative", 4, las = 0, line = 2)
axis(4, col = "green"); axis(2, col = "blue")
grid()
points(1:p, Cumulative.std, cex = 1, col = "green")
```



Factor Loadings and Specific Variances

```
# m = 2, Factor Loadings
lambda <- s.pca$sdev^2 # Lambda is essentially the eigenvalues
e <- s.pca$rotation
sqrt(lambda[1]) * e[, 1]
```

```
##   JP Morgan   City Bank Wells Fargo Royal Dutch       Exxon
##  -0.7323218  -0.8311791  -0.7262022  -0.6047155  -0.5630885
```

```
sqrt(lambda[2]) * e[, 2]
```

```
##   JP Morgan   City Bank Wells Fargo Royal Dutch       Exxon
##   0.4365209   0.2804859   0.3738582  -0.6939569  -0.7186401
```

```
# Specific Variances
sVar <- diag(s.cor - (lambda[1] * e[, 1] %*% t(e[, 1]) + lambda[2] * e[, 2] %*% t(e[, 2]))) # Using PC

# Residual Matrix
round(s.cor - (lambda[1] * e[, 1] %*% t(e[, 1]) + lambda[2] * e[, 2] %*% t(e[, 2]) + diag(sVar)), 2) #
```

```
##             JP Morgan City Bank Wells Fargo Royal Dutch Exxon
## JP Morgan        0.00     -0.10       -0.18       -0.03  0.06
## City Bank       -0.10      0.00       -0.13        0.01 -0.05
## Wells Fargo     -0.18     -0.13        0.00        0.00  0.01
## Royal Dutch     -0.03      0.01        0.00        0.00 -0.16
## Exxon            0.06     -0.05        0.01       -0.16  0.00
```

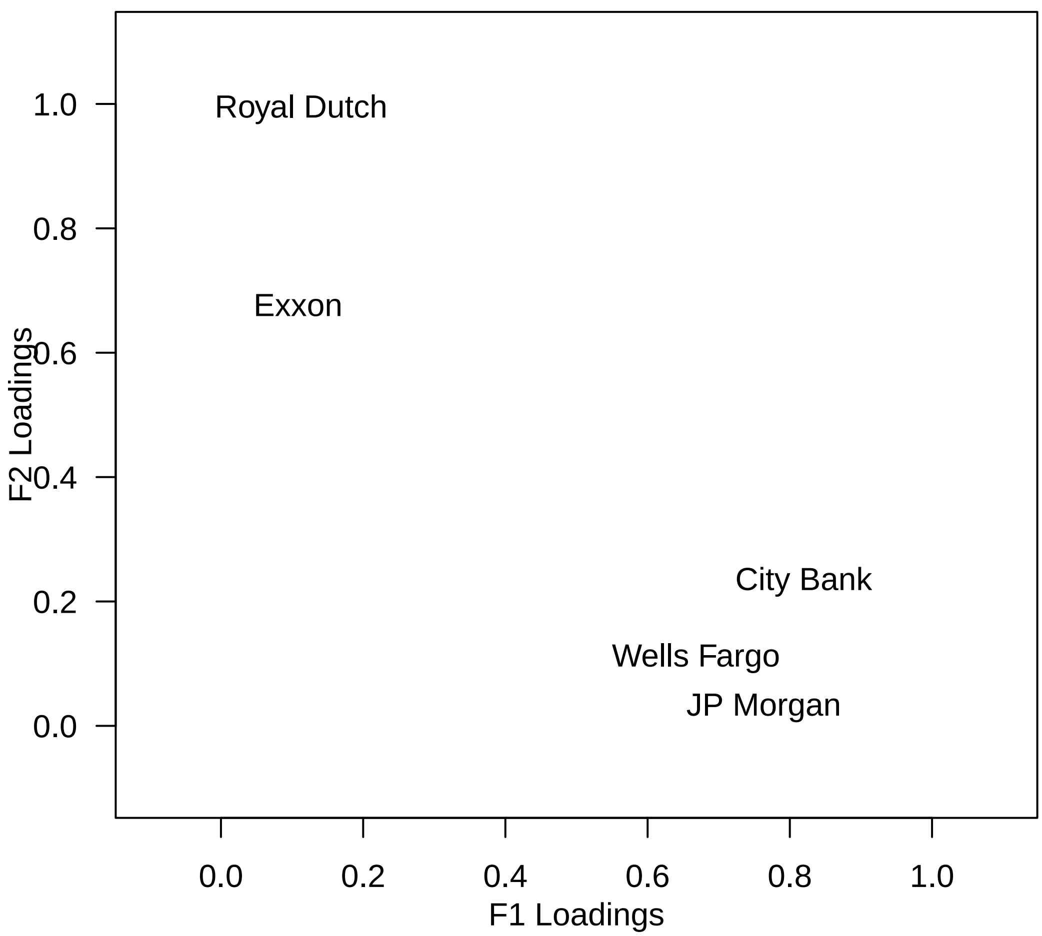
MLE

```
(stock.fac <- factanal(stock, factors = 2,
method = "mle", scale = T, center = T))
```

```
##
## Call:
## factanal(x = stock, factors = 2, method = "mle", scale = T, center = T)
##
## Uniquenesses:
##   JP Morgan   City Bank Wells Fargo Royal Dutch       Exxon
##       0.417       0.275       0.542       0.005       0.530
##
## Loadings:
##             Factor1 Factor2
## JP Morgan   0.763
## City Bank   0.819   0.232
## Wells Fargo 0.668   0.108
## Royal Dutch 0.113   0.991
## Exxon       0.108   0.677
```

```
##
##                 Factor1 Factor2
## SS loadings       1.725   1.507
## Proportion Var    0.345   0.301
## Cumulative Var    0.345   0.646
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 1.97 on 1 degree of freedom.
## The p-value is 0.16
```

```
par(las = 1, mgp = c(2, 1, 0), mar = c(3, 3, 1, 3))
plot(stock.fac$loadings, xlab = "F1 Loadings", ylab = "F2 Loadings",
     type = "n", xlim = c(-0.1, 1.1), ylim = c(-0.1, 1.1))
text(stock.fac$loadings, labels = colnames(stock))
```




```
# Residual Matrix
pred <- (stock.fac$loadings %*% t(stock.fac$loadings)) + diag(stock.fac$uniqueness)
(resid <- s.cor - pred)
```

```
##                 JP Morgan     City Bank   Wells Fargo   Royal Dutch
## JP Morgan    1.055860e-07  7.496780e-06 -2.564223e-03 -3.325561e-04
## City Bank    7.496780e-06  3.255673e-08  1.608871e-03  2.116218e-04
```

```
## Wells Fargo -2.564223e-03  1.608871e-03  5.157368e-08 -9.518792e-06
## Royal Dutch -3.325561e-04  2.116218e-04 -9.518792e-06 -1.559500e-06
## Exxon        5.198222e-02 -3.307885e-02  5.547153e-04  1.218853e-04
##                     Exxon
## JP Morgan    5.198222e-02
## City Bank   -3.307885e-02
## Wells Fargo  5.547153e-04
## Royal Dutch  1.218853e-04
## Exxon        2.670491e-07
```