PCA on Swap Rates

Code

Casey Tirshfield

The file `m_swap.txt` contains the monthly swap rates r_{kt} for eight maturities $T_k = 1, 2, 3, 4, 5, 7, 10,$ and 30 years from July 2000 to June 2007.

Hide

```
# import .txt file as data frame
df <- read.table('m_swap.txt', skip=1, header=TRUE)
# we deal with monthly change
swap.diff <- apply(df[,-1], 2, diff)
```

(a) Perform a principal component analysis (PCA) of the data using the sample covariance matrix.

Hide

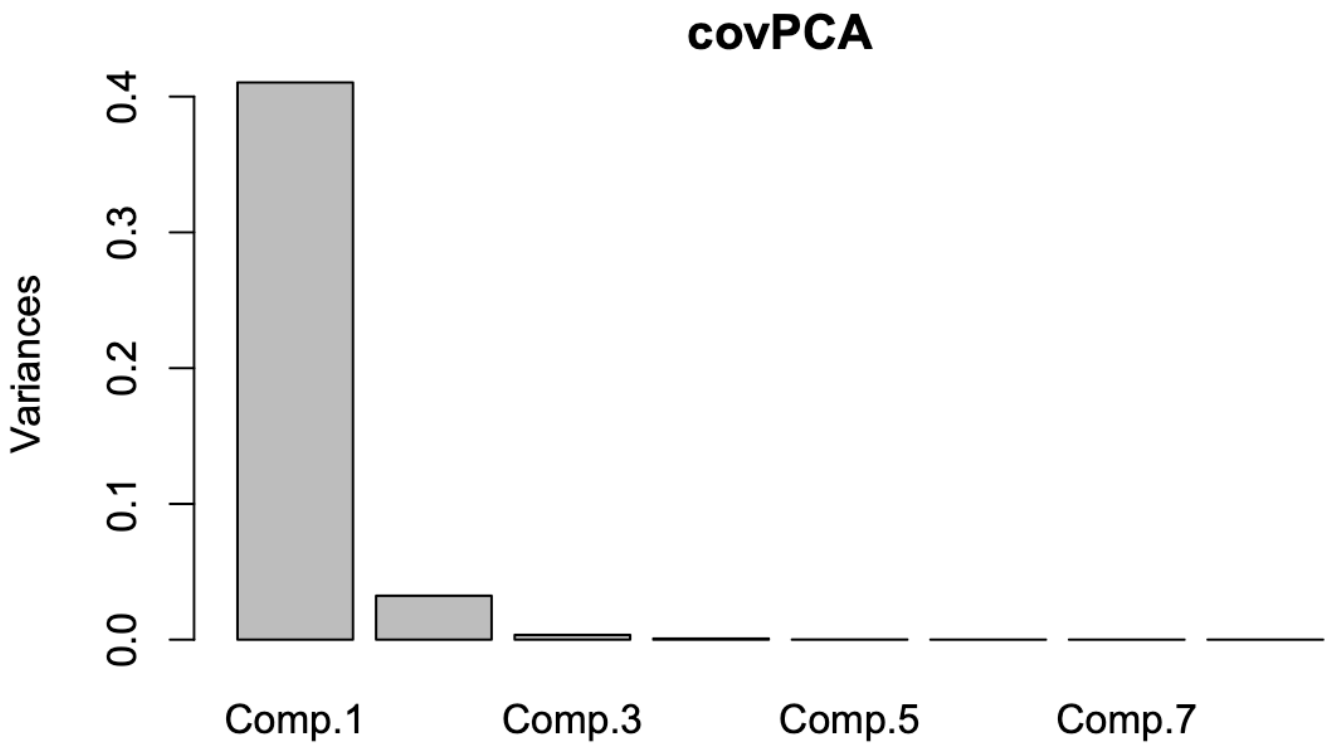
```
covPCA <- princomp(swap.diff, cor=FALSE)
summary(covPCA)
```

```
Importance of components:
                          Comp.1     Comp.2      Comp.3      Comp.4
Standard deviation     0.6406272 0.17987516 0.059185563 0.028120172
Proportion of Variance 0.9176055 0.07234153 0.007832076 0.001767996
Cumulative Proportion  0.9176055 0.98994707 0.997779143 0.999547139
                             Comp.5       Comp.6       Comp.7       Comp.8
Standard deviation     0.0111330377 6.299070e-03 4.914828e-03 3.842626e-03
Proportion of Variance 0.0002771231 8.871522e-05 5.400848e-05 3.301425e-05
Cumulative Proportion  0.9998242620 9.999130e-01 9.999670e-01 1.000000e+00
```

This plots $\frac{\lambda_i}{\text{tr}(\mathbf{V})}$ for each i:

Hide

```
screeplot(covPCA)
```



(b) Perform a PCA of the data using the sample correlation matrix.

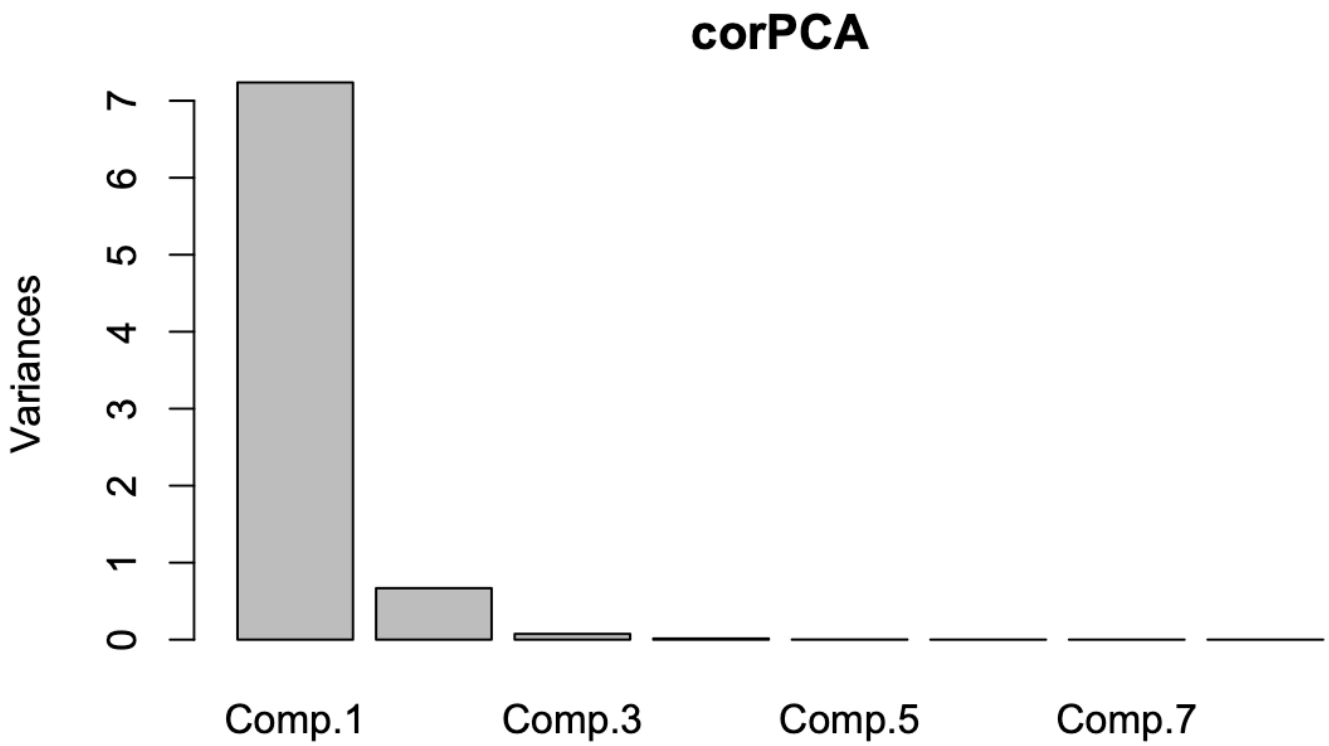
Hide

```
corPCA <- princomp(swap.diff, cor=TRUE)
summary(corPCA)
```

```
Importance of components:
                          Comp.1     Comp.2      Comp.3      Comp.4
Standard deviation     2.6903256 0.81796806 0.273625853 0.121072253
Proportion of Variance 0.9047315 0.08363397 0.009358888 0.001832311
Cumulative Proportion  0.9047315 0.98836542 0.997724309 0.999556621
                            Comp.5       Comp.6       Comp.7       Comp.8
Standard deviation     0.0473501672 2.595412e-02 1.991026e-02 1.532842e-02
Proportion of Variance 0.0002802548 8.420207e-05 4.955233e-05 2.937005e-05
Cumulative Proportion  0.9998368756 9.999211e-01 9.999706e-01 1.000000e+00
```

Hide

```
screeplot(corPCA)
```

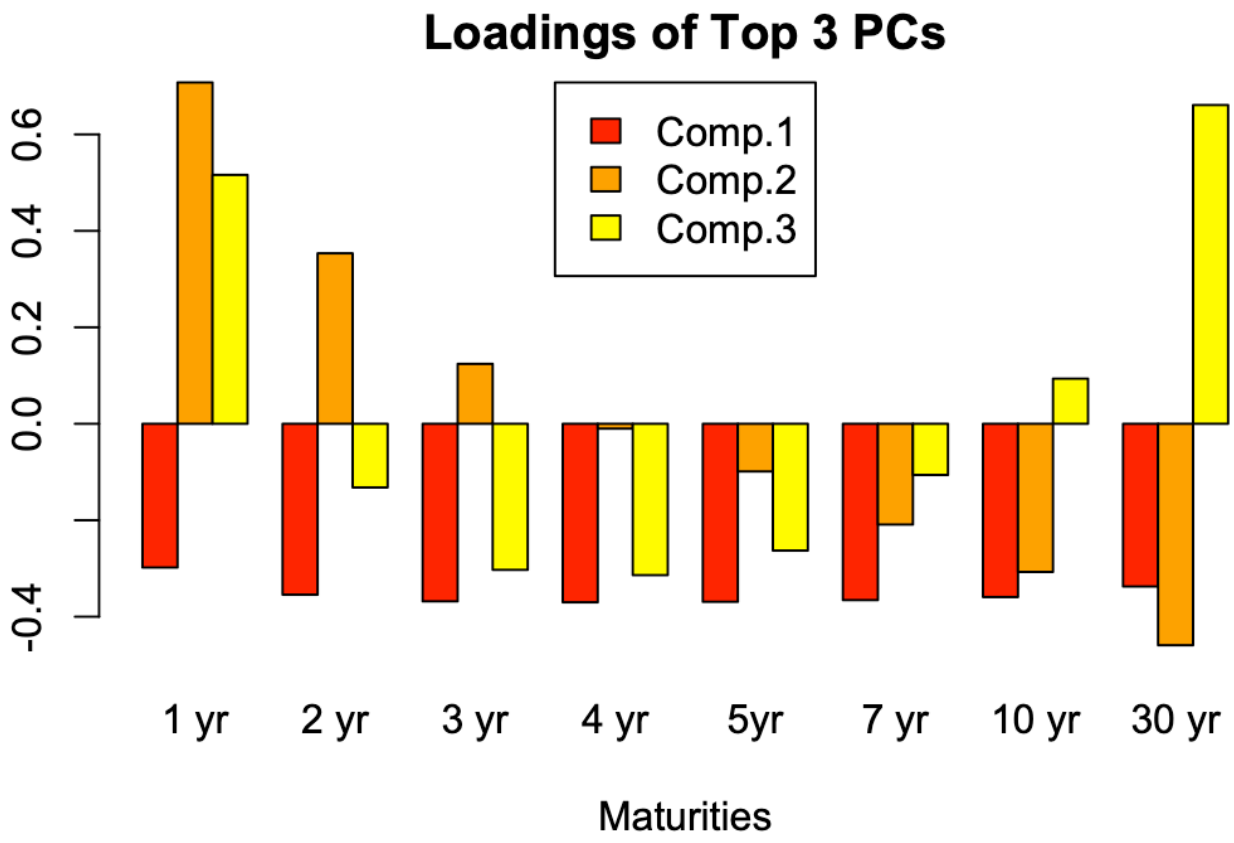


(c) Compare your results with those in Section 2.2.3 of Tze Leung Lai and Haipeng Xing's book, "Statistical Models and Methods for Financial Markets" for the daily data. Discuss the influence of sampling frequency on the result.

Hide

```
loadings <- t(corPCA$loadings[,1:3])
colnames(loadings) <- c('1 yr', '2 yr', '3 yr', '4 yr', '5yr',
                        '7 yr', '10 yr', '30 yr')
barplot(loadings, main = 'Loadings of Top 3 PCs', xlab='Maturities',
        col=c('red', 'orange', 'yellow'),

legend = rownames(loadings), beside=TRUE, args.legend=list(x='top'))
```



Both my results and the results given in Section 2.2.3 of Tze Leung Lai and Haipeng Xing's book, "Statistical Models and Methods for Financial Markets", indicated that greater than 90% of the variance can be explained by the first principal component. This would lead me to believe that the affects of sampling frequency are not significant, and that the rates share a common factor.