Multivariate Conditional Correlation Models - PCA

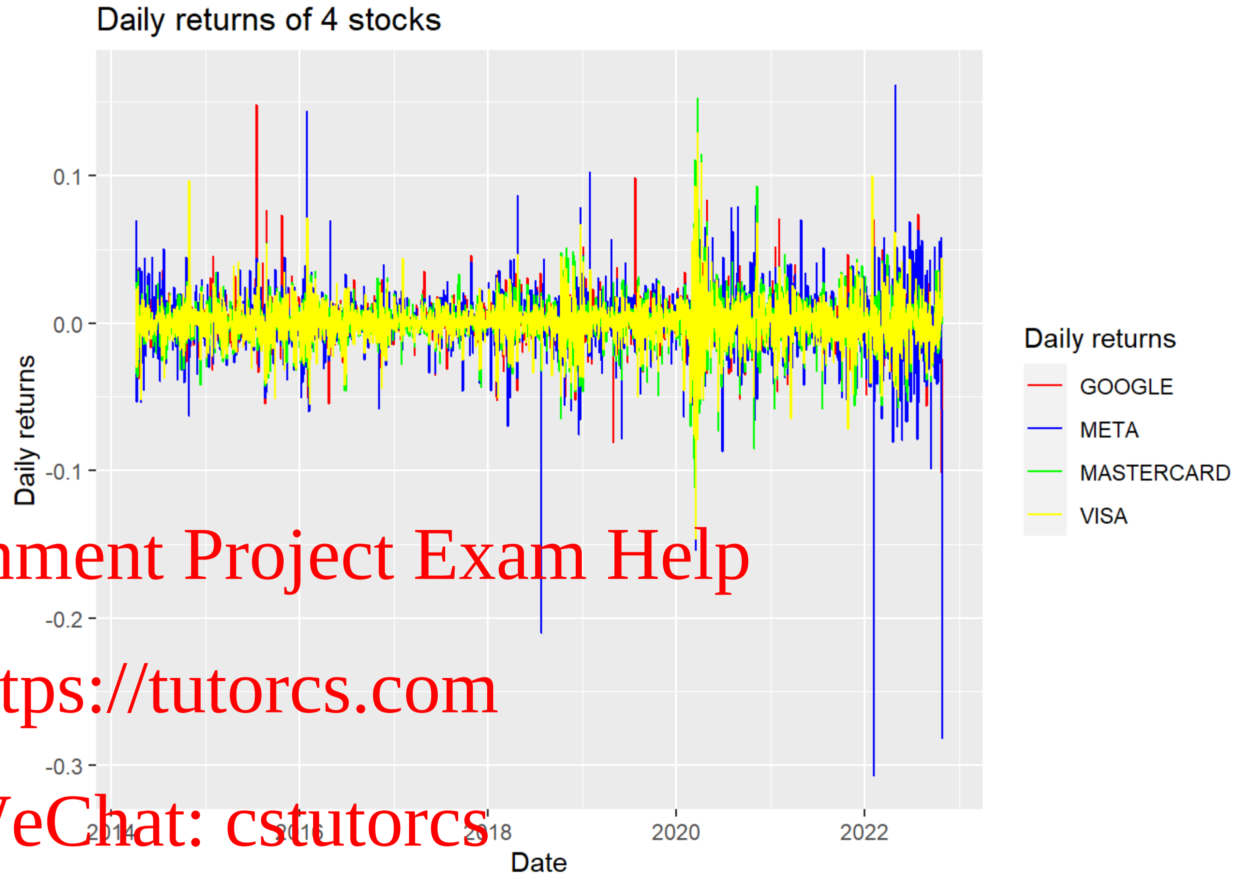
Code

CW6

07 November, 2022

Multivariate Conditional Correlation Models - PCA

One approach to handling the risk arising from groups of highly correlated market variables is PCA. This is a statistical tool with many applications in risk management. It takes historical data on daily stock returns and attempts to identify the main components or factors that explain most of the variation in returns. The aim of the analysis is to replace initial set of variables by a smaller number of *uncorrelated* variables (factors).



Unconditional sample correlation between our 4 stocks.

```
# Sample Correlations - all stocks
stocks_corr <- cor(log_returns_demean)
knitr::kable(stocks_corr, digits=4, align = 'c')
```

	GOOGLE	META	MASTERCARD	VISA
GOOGLE	1.0000	0.6440	0.6117	0.6033
META	0.6440	1.0000	0.5051	0.4856

	GOOGLE	META	MASTERCARD	VISA
MASTERCARD	0.6117	0.5051	1.0000	0.8986
VISA	0.6033	0.4856	0.8986	1.0000

PCA analysis on 4 stocks

Original data *r* dimension *T by N*. Use function *prcomp* to identify the directions (eigenvectors) that explain most of the stock return variation. The output of the function:

1. Rotation - matrix whose columns contain the eigenvectors, we will call these the weights *w* dimension *K by K*.
2. x - daily returns of the factors *f* dimension *T by K*, constructed *f=r.w*


```
PCA <- prcomp(x = log_returns_demean)
Table_PCA <- rbind(PCA$rotation, summary(PCA)$importance)

# rotation output - matrix whose columns contain the eigenvectors

knitr::kable(Table_PCA, digits=4, align = 'c')
```

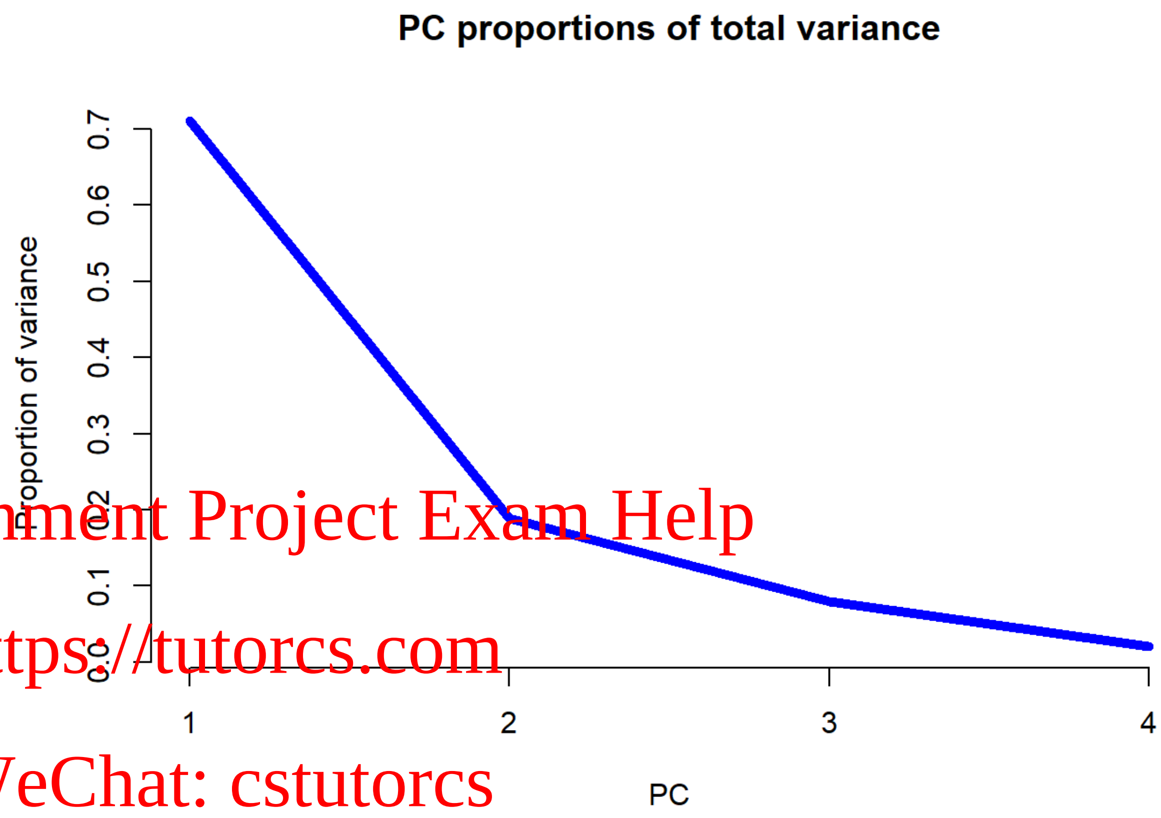
	PC1	PC2	PC3	PC4
GOOGLE	0.4536	-0.0286	-0.8907	0.0091
META	0.6323	-0.6940	0.3441	-0.0139
MASTERCARD	0.4685	0.5268	0.2285	0.6714
VISA	0.4183	0.4900	0.1897	-0.7409
Standard deviation	0.0317	0.0164	0.0106	0.0054
Proportion of Variance	0.7106	0.1890	0.0801	0.0203

Multivariate Conditional Correlation Models - PCA

	PC1	PC2	PC3	PC4
Cumulative Proportion	0.7106	0.8996	0.9797	1.0000

Try to identify, visually, how many factors to use….

```
par(mfrow=c(1,1))
plot(Table_PCA['Proportion of Variance',], type = 'l', lwd = 5, col
        = 'blue', xlim = c(1,4), main = 'PC proportions of total v
        ariance', xlab = 'PC', ylab = 'Proportion of variance', ax
        es = FALSE)
axis(1, 1:4)
axis(2)
```



Double check how factors are constructed

```
head(log_returns_demean)
```

Multivariate Conditional Correlation Models - PCA

```
##                   GOOGLE         META   MASTERCARD        VISA
## 2014-04-07 -0.009799743  0.003265061 -0.011035780 -0.02153295
## 2014-04-08  0.030080631  0.021286804  0.001945401 -0.00470136
## 2014-04-09  0.015944572  0.069759027  0.027409753  0.02347812
## 2014-04-10 -0.042545576 -0.053732871 -0.033744823 -0.02994868
## 2014-04-11 -0.019888365 -0.010959202 -0.036327408 -0.02537560
## 2014-04-14  0.003042089  0.005878864  0.035039050  0.02136900
```

```
w <- PCA$rotation
w
```



```
##                  PC1         PC2        PC3          PC4
## GOOGLE     0.4535801 -0.02861056 -0.8907097  0.009096991
## META       0.6322937 -0.69396166  0.3441345 -0.013902012
## MASTERCARD 0.4685439  0.52678789  0.2285345  0.671366648
## VISA       0.4182540  0.48999303  0.1896829 -0.740939136
```

```
f <- PCA$x
head(f)
```

```
##                    PC1          PC2           PC3           PC4
## 2014-04-07 -0.01655748 -0.018349960  0.0032458591  0.0084110087
## 2014-04-08  0.02604863 -0.016911670 -0.0199147624  0.0047672132
## 2014-04-09  0.07400282 -0.022923033  0.0205219769  0.0001813933
## 2014-04-10 -0.08160987  0.006054797  0.0060118140 -0.0001050419
## 2014-04-11 -0.04358483 -0.023396424  0.0008279376 -0.0056158028
## 2014-04-14  0.03045200  0.024762066  0.0113744655  0.0076368678
```

Run Univariate GARCH(1,1) on Selected Number of factors

Multivariate Conditional Correlation Models - PCA

```
nf <- 2 # we will only use the first 2 factors which explain 90% of variation
GARCH_1_1 <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
                        mean.model = list(armaOrder = c(0, 0), include.mean = FALSE))
uspec <- multispec(replicate(nf, GARCH_1_1)) # Univariate Garch on 2 factors
GARCH_multifit <- multifit(multispec = uspec, data = f[,1:nf], solver = 'hybrid')

sigma <- sigma(GARCH_multifit)
htMat <- xts(sigma^2, order.by = index(log_returns_demean))
```


Building Conditional Variance Covariance Matrix


```
# Set residual unexplained variation constant
errors <- log_returns_demean - f[, 1:nf] %*% t(PCA$rotation[,1:nf])
omega <- diag(colMeans(errors^2))
ht <- array(dim = c(length(Stocks), length(Stocks), dim(log_returns_demean)[1]))
for (i in 1:dim(log_returns_demean)[1]) {
  ht[, , i] <- PCA$rotation[,1:nf] %*% diag(as.numeric(htMat[i, ])) %*% t(PCA$rotation[,1:nf]) + omega
}
```


Conditional versus Unconditional Correlations

```
# 1. Low volatility date: 2017-04-12
# 2. High volatility date: 2020-03-30

ind1 <- match(x = as.Date('2017-04-12'), index(log_returns_demean))
corr1 <- cov2cor(ht[, , ind1])

ind2 <- match(x = as.Date('2020-03-30'), index(log_returns_demean))
corr2 <- cov2cor(ht[, , ind2])

corr1
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.0000000 0.4315436 0.3720299 0.3551491
## [2,] 0.4315436 1.0000000 0.1929858 0.1714308
## [3,] 0.3720299 0.1929858 1.0000000 0.7674956
## [4,] 0.3551491 0.1714308 0.7674956 1.0000000
```



```
corr2
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,] 1.0000000 0.9055692 0.8796905 0.8720304
## [2,] 0.9055692 1.0000000 0.6596087 0.6471912
## [3,] 0.8796905 0.6596087 1.0000000 0.9928305
## [4,] 0.8720304 0.6471912 0.9928305 1.0000000
```

```
stocks_corr
```

Multivariate Conditional Correlation Models - PCA

```
##                 GOOGLE      META  MASTERCARD      VISA
## GOOGLE     1.0000000 0.6440230   0.6116893 0.6033374
## META       0.6440230 1.0000000   0.5051198 0.4856127
## MASTERCARD 0.6116893 0.5051198   1.0000000 0.8986232
## VISA       0.6033374 0.4856127   0.8986232 1.0000000
```

Extra Code - relation between OLS and PCA

```
b <- apply(X = log_returns_demean, MARGIN = 2, FUN = function(x) lm
        (x ~ f))

bhat <- rbind(b$GOOGLE$coefficients, b$META$coefficients, b$MASTERC
        ARD$coefficients, b$VISA$coefficients)[,2:5]

rownames(bhat) <- c('GOOGLE', 'META', 'MASTERCARD', 'VISA')

knitr::kable(bhat, digits = 4, align = 'c')
```


	fPC1	fPC2	fPC3	fPC4
GOOGLE	0.4536	-0.0286	-0.8907	0.0091
META	0.6323	-0.6940	0.3441	-0.0139
MASTERCARD	0.4685	0.5268	0.2285	0.6714
VISA	0.4183	0.4900	0.1897	-0.7409

```
# Factor betas in columns - r = f*b

knitr::kable(PCA$rotation, digits = 4)
```

	PC1	PC2	PC3	PC4
GOOGLE	0.4536	-0.0286	-0.8907	0.0091
META	0.6323	-0.6940	0.3441	-0.0139
MASTERCARD	0.4685	0.5268	0.2285	0.6714
VISA	0.4183	0.4900	0.1897	-0.7409

PCA - with Yields Data

In order for this to work without specifying any path, save the csv file in the same folder as the codes and go to 'Session' -> 'Set Working Directory' -> 'To Source File Location'


```
Yields <- read.csv('Treasury Yields.csv')[, 2:11]
names(Yields) <- gsub('X', '', names(Yields))
l <- names(Yields)
```

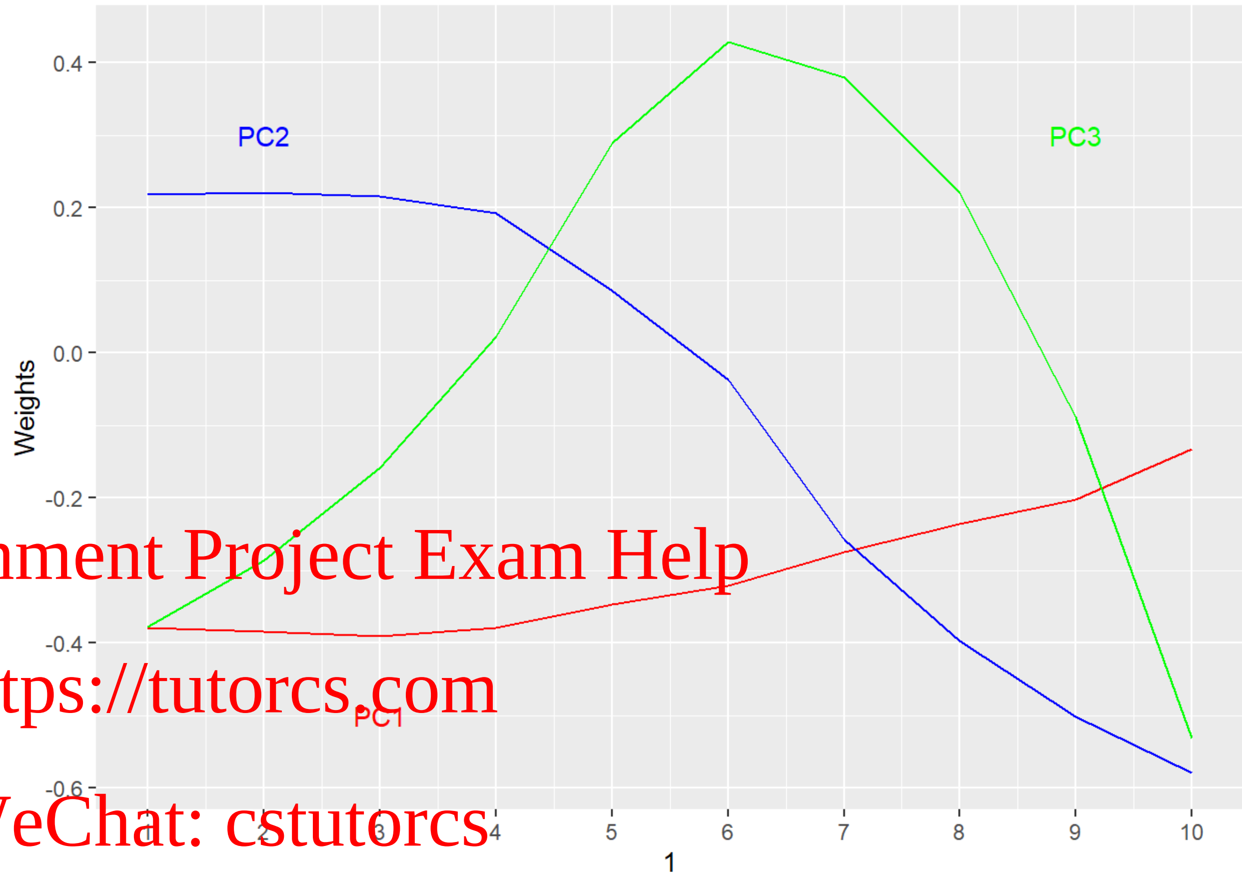
```
PCA <- prcomp(x = Yields)
Table_PCA <- rbind(PCA$rotation, summary(PCA)$importance)
knitr::kable(Table_PCA, digits = 3, align = 'c')
```

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
1M	-0.379	0.219	-0.378	0.551	0.222	-0.497	-0.252	0.007	-0.013	-0.057
3M	-0.384	0.221	-0.286	0.215	-0.047	0.520	0.543	-0.160	0.198	0.206
6M	-0.391	0.216	-0.158	-0.253	-0.393	0.202	-0.153	0.395	-0.358	-0.457
1Yr	-0.378	0.193	0.023	-0.484	-0.215	-0.191	-0.312	-0.244	0.120	0.575
2Yr	-0.346	0.085	0.289	-0.314	0.400	-0.106	0.134	-0.329	0.304	-0.548

Multivariate Conditional Correlation Models - PCA

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
3Yr	-0.321	-0.037	0.429	0.012	0.441	-0.020	0.225	0.329	-0.505	0.320
5Yr	-0.275	-0.257	0.380	0.274	-0.136	0.178	-0.279	0.425	0.573	0.038
7Yr	-0.236	-0.397	0.221	0.308	-0.222	0.209	-0.235	-0.590	-0.374	-0.087
10Yr	-0.203	-0.501	-0.088	-0.075	-0.382	-0.515	0.523	0.078	0.040	-0.030
30Yr	-0.133	-0.579	-0.531	-0.279	0.423	0.231	-0.211	0.086	0.020	0.046
Standard deviation	0.042	0.010	0.004	0.001	0.001	0.000	0.000	0.000	0.000	0.000
Proportion of Variance	0.943	0.048	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000
Cumulative Proportion	0.943	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000


```
ggplot(data = PCA$rotation, mapping = aes(x = c(1:10))) +
  geom_line(aes(y = PCA$rotation[, 'PC1']), color = 'red') +
  geom_line(aes(y = PCA$rotation[, 'PC2']), color = 'blue') +
  geom_line(aes(y = PCA$rotation[, 'PC3']), color = 'green') +
  ylab('Weights') +
  scale_x_continuous(breaks = seq(1, 10),1) +
  annotate(geom="text", x = 3, y = - 0.5, label="PC1", color="red")
        +
  annotate(geom="text", x = 2, y = 0.3, label="PC2", color="blue")
        +
  annotate(geom="text", x = 9, y = 0.3, label="PC3", color="green")
```



PCA - with Currency Data

```
Data_Ccy <- read.csv('Currency Excess Returns.csv')[2:10]
l <- names(Data_Ccy)

CcyLogRet <- log(1 + Data_Ccy)

PCA <- prcomp(x = CcyLogRet)
Table_PCA <- rbind(PCA$rotation, summary(PCA)$importance)

knitr::kable(Table_PCA, digits=2, align = 'c')
```

Multivariate Conditional Correlation Models - PCA

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
GBP	0.22	0.04	0.25	-0.03	0.93	-0.01	-0.12	-0.12	0.04
CAD	0.24	0.27	-0.09	-0.53	0.01	-0.54	0.53	0.04	-0.01
AUD	0.42	0.31	-0.37	-0.13	-0.09	-0.16	-0.71	0.20	0.03
NZD	0.44	0.27	-0.46	0.46	0.05	0.31	0.39	-0.25	-0.03
JPY	0.12	-0.76	-0.52	-0.31	0.14	0.08	0.03	-0.04	-0.05
SEK	0.39	-0.09	0.26	-0.13	-0.08	0.42	0.18	0.58	0.45
CHF	0.32	-0.37	0.18	0.47	-0.13	-0.58	-0.02	-0.12	0.37
NOK	0.38	-0.04	0.37	-0.35	-0.30	0.25	-0.13	-0.66	-0.03
EUR	0.35	-0.17	0.26	0.16	-0.06	-0.06	0.04	0.31	-0.81
Standard deviation	0.07	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.01
Proportion of Variance	0.62	0.12	0.09	0.05	0.04	0.02	0.02	0.02	0.01
Cumulative Proportion	0.62	0.74	0.83	0.88	0.92	0.95	0.97	0.99	1.00

Interest Rates

```
Data_IntRate <- read.csv('Interest Rates.csv')[2:11]
IntRateDiff <- Data_IntRate[,1:dim(Data_IntRate)[2] - 1] - Data_Int
        Rate[, dim(Data_IntRate)[2]]
AvgIRDiff <- colMeans(IntRateDiff)
plot(x = PCA$rotation[,2], y = AvgIRDiff, xlab = 'Weights')
```

Multivariate Conditional Correlation Models - PCA