Examples using the Ken French data

Bernt Arne Ødegaard

24 November 2021

1 Introduction

On his homepage, Ken French provides a data library very useful for understanding asset returns. In these notes we concentrate on his data for US asset returns.

The folloing contains examples where this data is used. The notes are primarily in the form of exercises, where an exercise involving the Ken French data is given, and then a solution is given, with code in R. The solution is focused on showing the R code, and the generated numbers, they are not complete lectures.

2 Reading Data

The data, as downloaded from Ken French data library page, needs to be split into separate files, since the original data mixes value weigthed, equally weighted, monthly, annual data into one file.

The easiest is to split it into separate files for each dataset one want.

For example, the following R code reads the pricing factors.

```
library(zoo)
Sys.setlocale(category = "LC_ALL", locale = "C")

indir <- "/home/bernt/data/2021/french_us_data/"
filename <- paste0(indir,"F-F_Research_Data_Factors.csv")

FF <- read.table(filename,
                 header=TRUE,
                 skip=3,
                 na.strings=c("-99.99","-999"),
                 sep=",")
dates <- as.yearmon(as.character(FF[,1]),format="%Y%m")
FF <- zoo(coredata(FF), order.by=dates)
RMRF <- FF$Mkt.RF/100
SMB <- FF$SMB/100
HML <- FF$HML/100
RF  <- FF$RF/100
RM <- RF + RMRF
names(RMRF) <- "RMRF"
names(RM) <- "RM"
names(RF) <- "RF"
names(HML) <- "HML"
names(SMB) <- "SMB"
```

The next reads industry portfolios

```
Sys.setlocale(category = "LC_ALL", locale = "C")
library(zoo)
indir <- "/home/bernt/data/2021/french_us_data/"

filename <- paste0(indir,"10_Industry_Portfolios_ew.csv")
FF10IndusEW <- read.table(filename,
                          header=TRUE,skip=11,
                          na.strings=c(" -99.99","-99.99","-999"),
                          sep=",")
dates <- as.yearmon(as.character(FF10IndusEW[,1]),format="%Y%m")
FF10IndusEW <- zoo(coredata(FF10IndusEW[,2:11]), order.by=dates)
FF10IndusEW <- FF10IndusEW / 100

filename <- paste0(indir,"10_Industry_Portfolios_vw.csv")
FF10IndusVW <- read.table(filename,
                          header=TRUE,skip=11,
                          na.strings=c(" -99.99","-99.99","-999"),
                          sep=",")
dates <- as.yearmon(as.character(FF10IndusVW[,1]),format="%Y%m")
FF10IndusVW <- zoo(coredata(FF10IndusVW[,2:11]), order.by=dates)
FF10IndusVW <- FF10IndusVW / 100
```

```
Sys.setlocale(category = "LC_ALL", locale = "C")
library(zoo)
indir <- "/home/bernt/data/2021/french_us_data/"

filename <- paste0(indir,"Portfolios_Formed_on_ME_ew.csv")
data <- read.table(filename,
                   header=TRUE,
                   skip=12,
                   na.strings=c("-99.99","-999"),
                   sep=",")
head(data)
dates <- as.yearmon(as.character(data[,1]),format="%Y%m")
FF3sizeEW <- zoo(coredata(data[,3:5]), order.by=dates)
FF3sizeEW <- FF3sizeEW / 100
FF5sizeEW <- zoo(coredata(data[,6:10]), order.by=dates)
FF5sizeEW <- FF5sizeEW / 100
FF10sizeEW <- zoo(coredata(data[,11:20]), order.by=dates)
FF10sizeEW <- FF10sizeEW / 100
head(FF3sizeEW)
head(FF5sizeEW)
head(FF10sizeEW)

filename <- paste0(indir,"Portfolios_Formed_on_ME_vw.csv")
data <- read.table(filename,
                   header=TRUE,
                   skip=12,
                   na.strings=c(" -99.99","-99.99","-999"),
                   sep=",")
dates <- as.yearmon(as.character(data[,1]),format="%Y%m")
FF3sizeVW <- zoo(coredata(data[,3:5]), order.by=dates)
FF3sizeVW <- FF3sizeVW / 100

FF5sizeVW <- zoo(coredata(data[,6:10]), order.by=dates)
FF5sizeVW <- FF5sizeVW / 100

FF10sizeVW <- zoo(coredata(data[,11:20]), order.by=dates)
FF10sizeVW <- FF10sizeVW / 100
```

3 Descriptive

Exercise 1.

Collect the returns data for 10 monthly size based (ew) portfolios from Ken French homepage.

- Describe the return series for the whole period and for the 1970–2013 subperiod.

Solution to Exercise 1.

Reading the data

```
# make sure that the first date do not change, this hardcodes the first date
library(zoo)
FFSizeEW <- read.table("../data/Portfolios_Formed_on_ME_monthly_ew.txt",
                        header=TRUE,skip=11)
FFSize10EW <- zooreg(FFSizeEW[,10:19],start=c(1926,7),frequency=12)
```

A quick and dirty summary is provided by

```
> summary(FFSize10EW)
     Index          Lo10              Dec2              Dec3
 Min.   :1926   Min.   :-30.410   Min.   :-33.280   Min.   :-33.670
 1st Qu.:1948   1st Qu.: -2.743   1st Qu.: -2.792   1st Qu.: -2.473
 Median :1970   Median :  1.125   Median :  1.290   Median :  1.430
 Mean   :1970   Mean   :  1.872   Mean   :  1.377   Mean   :  1.321
 3rd Qu.:1992   3rd Qu.:  5.240   3rd Qu.:  5.015   3rd Qu.:  4.745
 Max.   :2014   Max.   :136.560   Max.   :106.960   Max.   : 85.850
      Dec4              Dec5              Dec6             Dec7
 Min.   :-31.420   Min.   :-30.660   Min.   :-32.64   Min.   :-28.620
 1st Qu.: -2.395   1st Qu.: -2.192   1st Qu.: -2.21   1st Qu.: -2.020
 Median :  1.495   Median :  1.600   Median :  1.52   Median :  1.445
 Mean   :  1.247   Mean   :  1.196   Mean   :  1.20   Mean   :  1.135
 3rd Qu.:  4.883   3rd Qu.:  4.768   3rd Qu.:  4.64   3rd Qu.:  4.480
 Max.   : 70.950   Max.   : 63.720   Max.   : 57.07   Max.   : 55.210
      Dec8              Dec9              Hi10
 Min.   :-31.610   Min.   :-33.010   Min.   :-27.2100
 1st Qu.: -2.132   1st Qu.: -2.033   1st Qu.: -1.7225
 Median :  1.510   Median :  1.410   Median :  1.1500
 Mean   :  1.065   Mean   :  1.029   Mean   :  0.8972
 3rd Qu.:  4.290   3rd Qu.:  4.025   3rd Qu.:  3.9250
 Max.   : 48.160   Max.   : 48.130   Max.   : 36.7000
```

A nicely formatted summary is done by

```
stargazer(FFSize10EW,summary=TRUE,float=FALSE)
```

Statistic	N	Mean	St. Dev.	Min	Max
Lo10	1,052	1.872	10.935	−30.410	136.560
Dec2	1,052	1.377	9.271	−33.280	106.960
Dec3	1,052	1.321	8.368	−33.670	85.850
Dec4	1,052	1.247	7.730	−31.420	70.950
Dec5	1,052	1.196	7.375	−30.660	63.720
Dec6	1,052	1.200	7.043	−32.640	57.070
Dec7	1,052	1.135	6.737	−28.620	55.210
Dec8	1,052	1.065	6.342	−31.610	48.160
Dec9	1,052	1.029	6.062	−33.010	48.130
Hi10	1,052	0.897	5.464	−27.210	36.700

The subperiod data is found by

```
> ffsub <- window(FFSize10EW,start = c(1970,1),end=c(2014,1))
```

The descriptive table:

Statistic	N	Mean	St. Dev.	Min	Max
Lo10	529	1.317	6.625	−27.680	32.940
Dec2	529	1.049	6.650	−30.300	32.390
Dec3	529	1.135	6.420	−28.860	29.880
Dec4	529	1.092	6.177	−29.420	26.410
Dec5	529	1.128	6.076	−28.120	27.540
Dec6	529	1.124	5.706	−25.930	23.020
Dec7	529	1.128	5.599	−25.900	24.810
Dec8	529	1.056	5.438	−23.970	21.170
Dec9	529	1.026	5.077	−22.450	22.510
Hi10	529	0.924	4.785	−20.330	20.970

4 Seasonality

Exercise 2.

Ken French provides return on size sorted portfolios. Using monthly returns, and the 10 size sorted portfolio for the period 1926-2013, test for equality of mean returns across months.

What is the colloquial name for the effect you observe?

Is it stable across decades?

Solution to Exercise 2.

Should find that january has higher returns, which is the january effect.

```
library(zoo)
library(stargazer)
Sys.setlocale(category = "LC_ALL", locale = "C")
outdir <- "../../results/2016_04_seasonality/"

source ("../2016_04_read_data/read_size_portfolios.R")
source ("../2016_04_read_data/read_pricing_factors.R")

data <- merge(FFSize10EW,RM,all=FALSE)
data <- data*100.0
head(data)
filename <- paste0(outdir,"descriptive_10_size_ew_1926_2015.tex")
stargazer(data,summary=TRUE,float=FALSE,out=filename)
head(index(data))
mnames <- c("January","February","March", "April","May","June",
          "July","August","September","October","November","December" )
mnths<-months(as.POSIXct(yearmon(index(data))))
head(mnths)
jan <- (mnths==mnames[1])
feb <- (mnths==mnames[2])
mar <- (mnths==mnames[3])
apr <- (mnths==mnames[4])
may <- (mnths==mnames[5])
jun <- (mnths==mnames[6])
jul <- (mnths==mnames[7])
aug <- (mnths==mnames[8])
sep <- (mnths==mnames[9])
oct <- (mnths==mnames[10])
nov <- (mnths==mnames[11])
dec <- (mnths==mnames[12])
mmeans <- colMeans(data[jan])
mmeans <- rbind(mmeans,colMeans(data[feb]))
mmeans <- rbind(mmeans,colMeans(data[mar]))
mmeans <- rbind(mmeans,colMeans(data[apr]))
mmeans <- rbind(mmeans,colMeans(data[may]))
mmeans <- rbind(mmeans,colMeans(data[jun]))
mmeans <- rbind(mmeans,colMeans(data[jul]))
mmeans <- rbind(mmeans,colMeans(data[aug]))
mmeans <- rbind(mmeans,colMeans(data[sep]))
mmeans <- rbind(mmeans,colMeans(data[oct]))
mmeans <- rbind(mmeans,colMeans(data[nov]))
mmeans <- rbind(mmeans,colMeans(data[dec]))
mmeans <- rbind(mmeans,colMeans(data))
rownames(mmeans)<-c(mnames,"All")
filename <- paste0(outdir,"descriptive_10_size_means_per_month_1926_2015.tex")
stargazer(mmeans, float=FALSE, out=filename )

regr <- lm(data$RM~as.numeric(jan))
summary(regr)
```

```
summary(regr)$pvalues
regr$pvalues
regr <- lm(data~as.numeric(jan))
summary(regr)
```

Table 1 Describing size sorted portfolios. US 1926–2015.

Descriptive statistics for monthly returns (in percent) of 10 size sorted portfolios.

Statistic	N	Mean	St. Dev.	Min	Max
Lo.10	1,074	1.814	10.764	−29.140	134.290
Dec.2	1,074	1.355	9.210	−33.020	106.160
Dec.3	1,074	1.286	8.274	−31.170	83.080
Dec.4	1,074	1.224	7.796	−31.770	80.210
Dec.5	1,074	1.162	7.283	−31.440	58.000
Dec.6	1,074	1.183	7.001	−32.830	57.180
Dec.7	1,074	1.096	6.661	−29.110	51.900
Dec.8	1,074	1.061	6.327	−30.730	49.460
Dec.9	1,074	1.010	6.019	−32.310	47.530
Hi.10	1,074	0.890	5.449	−28.430	37.790
RM	1,074	0.928	5.379	−29.100	38.950

Table 2 Return averages split size and month. US 1926–2013.

Return averages, grouped by firm size and month of the year.

	Lo.10	Dec.2	Dec.3	Dec.4	Dec.5	Dec.6	Dec.7	Dec.8	Dec.9	Hi.10	RM
January	10.511	6.124	4.978	4.140	3.534	3.206	2.678	2.200	2.047	1.274	1.472
February	2.244	1.456	1.268	1.005	1.109	1.037	1.031	0.920	0.866	0.339	0.663
March	0.962	0.922	1.023	0.900	0.813	0.911	0.855	0.954	0.790	0.752	0.769
April	1.664	1.317	1.366	1.312	1.513	1.536	1.406	1.263	1.357	1.424	1.331
May	1.550	1.210	0.906	1.054	0.729	0.702	0.699	0.514	0.527	0.632	0.488
June	0.917	0.854	0.646	0.740	0.742	0.654	0.710	0.674	0.790	0.713	0.846
July	2.006	1.551	1.515	1.475	1.290	1.422	1.450	1.434	1.479	1.502	1.504
August	1.563	1.340	1.449	1.439	1.544	1.552	1.389	1.595	1.273	1.134	1.152
September	0.298	-0.607	-0.595	-0.450	-0.643	-0.676	-0.748	-0.866	-0.778	-0.939	-0.734
October	-0.793	-0.366	-0.241	-0.238	-0.187	0.027	0.042	0.324	0.401	0.691	0.420
November	0.936	1.620	1.620	1.615	1.802	1.785	1.715	1.877	1.513	1.497	1.492
December	-0.016	0.885	1.518	1.720	1.711	2.047	1.935	1.849	1.862	1.656	1.737
All	1.814	1.355	1.286	1.224	1.162	1.183	1.096	1.061	1.010	0.890	0.928

5 Crossectional asset pricing

Concerns testing of asset pricing models on a cross section of asset returns. Typically performed on portfolios of assets.

5.1 CAPM

Exercise 3.

Suppose the CAPM holds. What restriction does this impose on the following relationship:

$$E[r_i] - r_f = \alpha_i + \beta_i(E[r_m] - r_f)$$

Test this restriction using Ken French's 5 size sorted portfolios. Use data for 1980-2012. Do the tests for each portfolio separately.

Solution to Exercise 3.

Reading the data

```
# make sure that the first date do not change, this hardcodes the first date
library(zoo)
FFSizeEW <- read.table("../data/Portfolios_Formed_on_ME_monthly_EW.txt",
                       header=TRUE,skip=11)
FFSize3EW <- zooreg(FFSizeEW[,2:4],start=c(1926,7),frequency=12)
FFSize5EW <- zooreg(FFSizeEW[,5:9],start=c(1926,7),frequency=12)
FFSize10EW <- zooreg(FFSizeEW[,10:19],start=c(1926,7),frequency=12)
```

Doing the analysis, taking the subperiod 1980-2012.

```
> FFSize5EW <- window(FFSize5EW,start=c(1980,1),end=c(2012,12))
> data <- merge.zoo(FFSize5EW,RMRF,RF,all=FALSE)
> summary(data)
     Index           Lo20               Qnt2              Qnt3
 Min.   :1980   Min.   :-28.120   Min.   :-29.11   Min.   :-27.20
 1st Qu.:1988   1st Qu.: -2.340   1st Qu.: -2.48   1st Qu.: -2.29
 Median :1996   Median :  1.260   Median :  1.89   Median :  1.85
 Mean   :1996   Mean   :  1.215   Mean   :  1.11   Mean   :  1.15
 3rd Qu.:2004   3rd Qu.:  4.580   3rd Qu.:  4.61   3rd Qu.:  4.76
 Max.   :2013   Max.   : 31.780   Max.   : 24.60   Max.   : 23.91
      Qnt4            Hi20              RMRF              RF
 Min.   :-25.00   Min.   :-21.320   Min.   :-23.240   Min.   :0.0000
 1st Qu.: -2.04   1st Qu.: -1.860   1st Qu.: -2.100   1st Qu.:0.2100
 Median :  1.60   Median :  1.460   Median :  1.090   Median :0.4100
 Mean   :  1.14   Mean   :  1.057   Mean   :  0.586   Mean   :0.4101
 3rd Qu.:  4.43   3rd Qu.:  3.880   3rd Qu.:  3.520   3rd Qu.:0.5600
 Max.   : 20.67   Max.   : 17.640   Max.   : 12.460   Max.   :1.3500
```

Looking first at the analysis for the low quintile:

```
> eri <- data$Lo20-data$RF
> erm <- data$RMRF
> reg <- lm(as.matrix(eri)~as.matrix(erm))
> summary(reg)

Call:
lm(formula = as.matrix(eri) ~ as.matrix(erm))

Residuals:
```

```
     Min       1Q   Median       3Q      Max 
-10.9582  -2.3328  -0.3355   1.9856  27.8145 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)     0.19914    0.20927   0.952    0.342    
as.matrix(erm)  1.03408    0.04514  22.907   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 4.115 on 391 degrees of freedom
Multiple R-squared: 0.573,Adjusted R-squared: 0.5719 
F-statistic: 524.7 on 1 and 391 DF,  p-value: < 2.2e-16 
```

Summarizing the analysis for the five size portfolios:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.1991	0.2093	0.95	0.3419
as.matrix(erm)	1.0341	0.0451	22.91	0.0000

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0009	0.1433	0.01	0.9948
as.matrix(erm)	1.1926	0.0309	38.57	0.0000

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0539	0.1149	0.47	0.6391
as.matrix(erm)	1.1700	0.0248	47.19	0.0000

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0703	0.0838	0.84	0.4020
as.matrix(erm)	1.1256	0.0181	62.23	0.0000

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0426	0.0585	0.73	0.4671
as.matrix(erm)	1.0317	0.0126	81.70	0.0000

Using this time period, none of the individual alphas are significant.

5.2 CAPM test using GMM

Exercise 4.

Consider the CAPM relationship

$$E[r_i] - r_f = \alpha_i + \beta_i(E[r_m] - r_f)$$

If the CAPM holds, $\alpha_i = 0$. One way of testing the CAPM is to test whether this restriction holds equation by equation (the Black Jensen Scholes analysis). However, a more efficient test is to test the joint hypothesis that $\alpha_i = 0$ for all stocks/portfolios in a cross section. One way this can be tested is to use the GMM framework of MacKinlay. Use GMM to estimate the joint system, and test this restriction.

The assets to be used are Ken French's 5 size sorted portfolios. Use data for 1980-2012.

Solution to Exercise 4.

Reading the data

```
# make sure that the first date do not change, this hardcodes the first date
library(zoo)
FFSizeEW <- read.table("../data/Portfolios_Formed_on_ME_monthly_EW.txt",
                        header=TRUE,skip=11)
FFSize3EW <- zooreg(FFSizeEW[,2:4],start=c(1926,7),frequency=12)
FFSize5EW <- zooreg(FFSizeEW[,5:9],start=c(1926,7),frequency=12)
FFSize10EW <- zooreg(FFSizeEW[,10:19],start=c(1926,7),frequency=12)
```

Doing the analysis

```
eRFFSize5EW <- FFSize5EW - RF
eRi <- window(eRFFSize5EW,start=c(1980,1),end=c(2012,12))
data <- merge(eRi,RMRF,all=FALSE)
eRi <- as.matrix(data[,1:5])
eRm <- as.matrix(data[,6])
res   <- gmm(eRi~eRm,x=eRm)
summary(res)
R  <- cbind(diag(5),matrix(0,5,5))
c  <- rep(0,5)
linearHypothesis(res,R,c,test="F")
```

Results.

Take subperiod

```
> eRi <- window(eRFFSize5EW,start=c(1980,1),end=c(2012,12))
> data <- merge(eRi,RMRF,all=FALSE)
> eRi <- as.matrix(data[,1:5])
> eRm <- as.matrix(data[,6])
```

First do the estimation as equation by equation OLS:

```
> summary(lm(eRi~eRm))
Response Lo20 :

Call:
lm(formula = Lo20 ~ eRm)

Residuals:
     Min       1Q   Median       3Q      Max
-10.9582  -2.3328  -0.3355   1.9856  27.8145
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.19914    0.20927   0.952    0.342
eRm          1.03408    0.04514  22.907   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.115 on 391 degrees of freedom
Multiple R-squared: 0.573,Adjusted R-squared: 0.5719
F-statistic: 524.7 on 1 and 391 DF,  p-value: < 2.2e-16


Response Qnt2 :

Call:
lm(formula = Qnt2 ~ eRm)

Residuals:
    Min      1Q  Median      3Q     Max
-8.3814 -1.5499 -0.1874  1.5103 18.8482

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.0009405  0.1433456   0.007    0.995
eRm         1.1925765  0.0309223  38.567   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.819 on 391 degrees of freedom
Multiple R-squared: 0.7918,Adjusted R-squared: 0.7913
F-statistic:  1487 on 1 and 391 DF,  p-value: < 2.2e-16


Response Qnt3 :

Call:
lm(formula = Qnt3 ~ eRm)

Residuals:
    Min      1Q  Median      3Q     Max
-6.5661 -1.2221 -0.1794  1.1986 15.7156

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.05394    0.11493   0.469    0.639
eRm          1.17001    0.02479  47.194   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.26 on 391 degrees of freedom
Multiple R-squared: 0.8507,Adjusted R-squared: 0.8503
F-statistic:  2227 on 1 and 391 DF,  p-value: < 2.2e-16
```

```
Response Qnt4 :

Call:
lm(formula = Qnt4 ~ eRm)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2501 -0.9573 -0.0575  0.7323  9.1194

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.07034    0.08385   0.839    0.402
eRm          1.12564    0.01809  62.231   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.649 on 391 degrees of freedom
Multiple R-squared: 0.9083,Adjusted R-squared: 0.9081
F-statistic:  3873 on 1 and 391 DF,  p-value: < 2.2e-16


Response Hi20 :

Call:
lm(formula = Hi20 ~ eRm)

Residuals:
    Min      1Q  Median      3Q     Max
-6.1496 -0.4402 -0.0131  0.3847  7.0747

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.04261    0.05854   0.728    0.467
eRm          1.03167    0.01263  81.697   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.151 on 391 degrees of freedom
Multiple R-squared: 0.9447,Adjusted R-squared: 0.9445
F-statistic:  6674 on 1 and 391 DF,  p-value: < 2.2e-16
```

Then run a GMM esimation

```
> res   <- gmm(eRi~eRm,x=eRm)
> summary(res)

Call:
gmm(g = eRi ~ eRm, x = eRm)


Method:  twoStep

Kernel:  Quadratic Spectral

Coefficients:
                  Estimate     Std. Error   t value      Pr(>|t|)
```

```
Lo20_(Intercept)   1.9914e-01   2.3014e-01   8.6529e-01   3.8688e-01
Qnt2_(Intercept)   9.4053e-04   1.3994e-01   6.7209e-03   9.9464e-01
Qnt3_(Intercept)   5.3937e-02   1.1090e-01   4.8636e-01   6.2671e-01
Qnt4_(Intercept)   7.0342e-02   8.9496e-02   7.8598e-01   4.3188e-01
Hi20_(Intercept)   4.2611e-02   5.8895e-02   7.2351e-01   4.6937e-01
Lo20_eRm           1.0341e+00   4.8468e-02   2.1335e+01  5.3366e-101
Qnt2_eRm           1.1926e+00   3.5770e-02   3.3341e+01  1.0000e-243
Qnt3_eRm           1.1700e+00   3.1389e-02   3.7274e+01  4.3185e-304
Qnt4_eRm           1.1256e+00   2.4935e-02   4.5142e+01   0.0000e+00
Hi20_eRm           1.0317e+00   2.1909e-02   4.7090e+01   0.0000e+00

J-Test: degrees of freedom is 0
                J-test                P-value
Test E(g)=0:    1.04856204379519e-24  *******
```

before finally testing the hypothesis that the intercepts are jointly zero

```
> R  <- cbind(diag(5),matrix(0,5,5))
> c  <- rep(0,5)
> linearHypothesis(res,R,c,test="F")
Linear hypothesis test

Hypothesis:
Lo20_((Intercept) = 0
Qnt2_((Intercept) = 0
Qnt3_((Intercept) = 0
Qnt4_((Intercept) = 0
Hi20_((Intercept) = 0

Model 1: restricted model
Model 2: eRi ~ eRm

  Df  Chisq Pr(>Chisq)
1
2  5 3.9298     0.5596
```

We are not able to reject the null, the CAPM is apparently acceptable for this time period and cross section.

5.3 Estimating m directly

Exercise 5.

The Stochastic Discount factor approach to asset pricing results in the following expression for pricing any excess return:

$$E[m_t er_{it}] = 0$$

Consider an empirical implementation of this where we write the pricing variable m as a function of a set of prespecified factors f:

$$m_t = 1 + bf_t$$

Consider the case of the one factor model $f = 1 + ber_m$, where the only explanatory factor is the return on a broad based market index.

Implement this approach on the set of 5 industry portfolios provided by Ken French. Use data 1980–2012.

Solution to Exercise 5.

Code for doing the analysis

```
source("read_industries.R")
source("read_pricing_factors.R")

                                        # estimate m=1+b*f in crossection
eR <- FFIndusEW-RF
eRi <- window(eR,first=c(1980,1),last=c(2012,12))
# take intersection to align data
data  <- merge(eRi,RMRF,all=FALSE)
eRi    <- as.matrix(data[,1:5])
eRm   <- as.vector(data[,6])
X     <- cbind(eRi,eRm)
g     <- function (parms,X) {
  b <- parms[1];
  f <- as.vector(X[,6])
  m <- 1 + b * f
  e <- m * X[,1:5]
  return (e);
}
library(gmm)
t0   <- c(0.1);
res <-  gmm(g,X,t0,method="Brent",lower=-10,upper=10)
summary(res)
```

Results

```
>                                       # estimate m=1+b*f in crossection
> eR <- FFIndusEW-RF
> eRi <- window(eR,first=c(1980,1),last=c(2012,12))
> # take intersection to align data
> data  <- merge(eRi,RMRF,all=FALSE)
> eRi    <- as.matrix(data[,1:5])
> eRm   <- as.vector(data[,6])
> X     <- cbind(eRi,eRm)
> g     <- function (parms,X) {
+   b <- parms[1];
+   f <- as.vector(X[,6])
```

```
+   m <- 1 + b * f
+   e <- m * X[,1:5]
+   return (e);
+ }
> library(gmm)
Loading required package: sandwich
> t0   <- c(0.1);
> res <-  gmm(g,X,t0,method="Brent",lower=-10,upper=10)
> summary(res)

Call:
gmm(g = g, x = X, t0 = t0, method = "Brent", lower = -10, upper = 10)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  3.9418 )

Coefficients:
          Estimate     Std. Error   t value      Pr(>|t|)
Theta[1]  -2.6981e-02   6.8114e-03  -3.9611e+00   7.4595e-05

J-Test: degrees of freedom is 4
                J-test   P-value
Test E(g)=0:    5.02662  0.28458

Initial values of the coefficients
   Theta[1]
-0.02777592

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  NA
Gradian eval. =  NA
>
> proc.time()
   user  system elapsed
  1.344   0.088   1.433
```

Exercise 6.

The Stochastic Discount factor approach to asset pricing results in the following expression for pricing any excess return:

$$E[m_t er_{it}] = 0$$

Consider an empirical implementation of this where we write the pricing variable m as a function of a set of prespecified factors f:

$$m_t = 1 + bf_t$$

Consider the case of the one factor model $f = 1 + ber_m$, where the only explanatory factor is the return on a broad based market index.

Implement this approach on the set of 5 size sorted portfolios provided by Ken French. Use data 1926–2012.

Is the market a relevant pricing factor?

Solution to Exercise 6.

Reading data

```
source("read_size_portfolios.R")
source("read_pricing_factors.R")
eRi <- FFSize5EW - RF
data  <- merge(eRi,RMRF,all=FALSE)
summary(data)
eRi    <- as.matrix(data[,1:5])
eRm   <- as.vector(data[,6])
```

The specification of the GMM estimation:

```
X      <- cbind(eRi,eRm)
g1     <- function (parms,X) {
  b <- parms[1];
  f <- as.vector(X[,6])
  m <- 1 + b * f
  e <- m * X[,1:5]
  return (e);
}
```

Running the GMM analysis

```
t0   <- c(0.1);
res <-  gmm(g1,X,t0,method="Brent",lower=-10,upper=10)
summary(res)
```

Results

```
> summary(data)
     Index           Lo20                Qnt2                 Qnt3
 Min.   :1926   Min.   :-32.010   Min.   :-31.9600   Min.   :-31.3100
 1st Qu.:1948   1st Qu.: -3.110   1st Qu.: -2.8650   1st Qu.: -2.5550
 Median :1970   Median :  0.960   Median :  1.1700   Median :  1.2000
 Mean   :1970   Mean   :  1.373   Mean   :  0.9704   Mean   :  0.8869
 3rd Qu.:1991   3rd Qu.:  4.695   3rd Qu.:  4.5350   3rd Qu.:  4.4050
 Max.   :2013   Max.   :110.670   Max.   : 81.1900   Max.   : 56.8400
      Qnt4             Hi20               RMRF
 Min.   :-29.760   Min.   :-30.100   Min.   :-28.980
 1st Qu.: -2.470   1st Qu.: -2.195   1st Qu.: -2.105
 Median :  1.160   Median :  0.930   Median :  1.010
 Mean   :  0.787   Mean   :  0.655   Mean   :  0.628
 3rd Qu.:  4.125   3rd Qu.:  3.640   3rd Qu.:  3.655
 Max.   : 50.010   Max.   : 41.790   Max.   : 37.770
```

GMM results

```
Call:
gmm(g = g1, x = X, t0 = t0, method = "Brent", lower = -10, upper = 10)

Method:  twoStep
Kernel:  Quadratic Spectral(with bw =  3.56894 )

Coefficients:
          Estimate    Std. Error  t value     Pr(>|t|)
Theta[1]  -0.0199775   0.0060763  -3.2877763   0.0010098

J-Test: degrees of freedom is 4
                J-test     P-value
Test E(g)=0:    14.606060   0.005592

Initial values of the coefficients
  Theta[1]
-0.0253752

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  NA
Gradian eval. =  NA
```

	Model 1
Theta[1]	-0.02^{***}
	(0.01)
Criterion function	1411.21
Num. obs.	1035

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Exercise 7.

The Stochastic Discount factor approach to asset pricing results in the following expression for pricing any excess return:

$$E[m_t er_{it}] = 0$$

Consider an empirical implementation of this where we write the pricing variable m as a function of a set of prespecified factors f:

$$m_t = 1 + bf_t$$

Consider the case of the three factor model $f = 1+b_1 er_m+b_2 SMB+b_3 HML$, where the explanatory factors are the return on a broad based market index, and the two Fama French factors SMB and HML.

Implement this approach on the set of 5 size sorted portfolios provided by Ken French. Use data 1926–2012.

Which are the relevant pricing factors?

Solution to Exercise 7.

Organizing the data

```
library(gmm)
source("read_size_portfolios.R")
source("read_pricing_factors.R")

eRi <- FFSize5EW - RF
data  <- merge(eRi,RMRF,SMB,HML,all=FALSE)
summary(data)
eRi    <- as.matrix(data[,1:5])
eRm   <- as.vector(data$RMRF)
smb <- as.vector(data$SMB)
hml <- as.vector(data$HML)
```

The GMM specification

```
X      <- cbind(eRi,eRm,smb,hml)
g3     <- function (parms,X) {
  b1 <- parms[1];
  b2 <- parms[2];
  b3 <- parms[3];
  erm <- as.vector(X[,6])
  smb <- as.vector(X[,7])
  hml <- as.vector(X[,8])
  m <- 1 + b1 * erm + b2*smb + b3*hml
  e <- m * X[,1:5]
  return (e);
}
```

Data, overview

```
> summary(data)
     Index           Lo20               Qnt2                Qnt3
 Min.   :1926   Min.   :-32.010   Min.   :-31.9600   Min.   :-31.3100
 1st Qu.:1948   1st Qu.: -3.110   1st Qu.: -2.8650   1st Qu.: -2.5550
 Median :1970   Median :  0.960   Median :  1.1700   Median :  1.2000
 Mean   :1970   Mean   :  1.373   Mean   :  0.9704   Mean   :  0.8869
 3rd Qu.:1991   3rd Qu.:  4.695   3rd Qu.:  4.5350   3rd Qu.:  4.4050
 Max.   :2013   Max.   :110.670   Max.   : 81.1900   Max.   : 56.8400
```

```
      Qnt4              Hi20              RMRF              SMB
 Min.   :-29.760   Min.   :-30.100   Min.   :-28.980   Min.   :-16.3900
 1st Qu.: -2.470   1st Qu.: -2.195   1st Qu.: -2.105   1st Qu.: -1.5200
 Median :  1.160   Median :  0.930   Median :  1.010   Median :  0.0500
 Mean   :  0.787   Mean   :  0.655   Mean   :  0.628   Mean   :  0.2352
 3rd Qu.:  4.125   3rd Qu.:  3.640   3rd Qu.:  3.655   3rd Qu.:  1.7750
 Max.   : 50.010   Max.   : 41.790   Max.   : 37.770   Max.   : 39.0400
      HML
 Min.   :-13.450
 1st Qu.: -1.295
 Median :  0.220
 Mean   :  0.382
 3rd Qu.:  1.745
 Max.   : 35.480
```

Running the GMM analysis

```
> t0   <- c(1.0,0,0);
> res <-  gmm(g3,X,t0)
> summary(res)

Call:
gmm(g = g3, x = X, t0 = t0)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  4.83032 )

Coefficients:
          Estimate    Std. Error  t value     Pr(>|t|)
Theta[1]  -0.0133913   0.0086600  -1.5463407   0.1220223
Theta[2]   0.0131771   0.0145060   0.9083904   0.3636720
Theta[3]  -0.0778065   0.0264944  -2.9367123   0.0033171

J-Test: degrees of freedom is 2
                J-test   P-value
Test E(g)=0:    3.51806  0.17221

Initial values of the coefficients
    Theta[1]     Theta[2]     Theta[3]
-0.009775731  0.011717385 -0.076275222

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  100
Gradian eval. =  NA
```

	Model 1
Theta[1]	-0.01
	(0.01)
Theta[2]	0.01
	(0.01)
Theta[3]	-0.08^{***}
	(0.03)
Criterion function	339.91
Num. obs.	1035

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

5.4 Estimating factors in system

Exercise 8.

Consider the moment condition

$$E[er - (a + \beta f)] = 0$$

$$E[f'(r - (a + \beta f))] = 0$$

$$E[er - \beta\lambda] = 0$$

where f is the excess return on the market. Estimate this using the five industry portfolios on the US.

Solution to Exercise 8.