Asset pricing data in R

Bernt Arne Ødegaard

21 December 2023

1 Introduction

For purposes of asset pricing, the typical data is monthly returns and interest rates. To do analysis typically calculate the *excess* return (return above the risk free rate).

2 Calculating individual stock returns directly from market data

To do this, the steps are:

- Download asset (stocks, indices, ...) prices.
- For each asset, calculate monthly returns (for example using the `quantmod` package).
- Align the series along the time dimension.

The following is an example that gathers monthly return of three large US IT companies (Microsoft, Apple and Alphabet (Google)) and calculates an eqully weighted portfolio of the three, using data from Yahoo Finance.

```
library(quantmod)
first_date <- as.Date("2010-01-01")
microsoft <- getSymbols("MSFT",auto.assign=FALSE,from=first_date)
apple     <- getSymbols("AAPL",auto.assign=FALSE,from=first_date)
alphabet  <- getSymbols("GOOG",auto.assign=FALSE,from=first_date)

microsoft_daily_prices <- microsoft$MSFT.Adjusted
microsoft_monthly_returns <- monthlyReturn(microsoft_daily_prices)
index(microsoft_monthly_returns) <- as.yearmon(index(microsoft_monthly_returns))
names(microsoft_monthly_returns) <- "MicroSoft"

apple_daily_prices <- apple$AAPL.Adjusted
apple_monthly_returns <- monthlyReturn(apple_daily_prices)
index(apple_monthly_returns) <- as.yearmon(index(apple_monthly_returns))
names(apple_monthly_returns) <- "Apple"

alphabet_daily_prices <- alphabet$GOOG.Adjusted
alphabet_monthly_returns <- monthlyReturn(alphabet_daily_prices)
index(alphabet_monthly_returns) <- as.yearmon(index(alphabet_monthly_returns))
names(alphabet_monthly_returns) <- "Alphabet"

stock_returns <- merge(microsoft_monthly_returns,
```

```
                        apple_monthly_returns,
                        alphabet_monthly_returns,
                        all=FALSE)
ew <- (1/3) * (stock_returns$MicroSoft + stock_returns$Apple + stock_returns$Alphabet)
names(ew) <- "ew"
```

3 Getting the risk free interest rate right

Careful about the risk free rate. If you get risk free rates by looking at current quotes, be aware that is the forward looking interest rate. For example, the current 1 month interest rate quote is the interest rate implied in buying a one-month bond *now*, and getting the principal back in one month. The risk free interest rate over a month is therefore the quoted one month interest rate at the last date of the previous month.

Some sources (such as the Ken French data) has taken care of this, the risk free rate is matched to the other returns. Other sources (such as the Norwegian data below) may give the forward-looking interest rates, in which case lagging the interest rate is necessary.

For the previous example, download the US 1 month constant maturity treasury rate, and use it to construct excess returns.

```
library(quantmod)
source ("read_stocks.R")
                                        # daily observations of 1-month constant maturity treasury rate
treas_1m <- getSymbols("DGS1MO", src="FRED", auto.assign=FALSE)
names(treas_1m) <- "treas_1m"
treas_1m <- na.omit(treas_1m)
                                        #create monthly observations
Rf <- treas_1m[endpoints(treas_1m,"month"),]
names(Rf) <- "Rf"
Rf <- Rf/100                            # don't want percentage interest rates
                                        # pick last observation in a given month
index(Rf) <- as.yearmon(index(Rf))
                                        # that last observation should be matched with returns the next
Rf <- na.omit(lag(Rf,1))
                                        # take intersection, all dates with both returns and Rf
data <- merge(apple_monthly_returns,Rf,all=FALSE)
excess_monthly_returns_apple <- data$Apple-data$Rf
```

4 The Ken French data

Discuss a source which provides ready-made monthly returns:

Ken French provides portfolio returns and "pricing factors" for the US and Globally. (web page)

Note that the data on Ken French web site needs editing, he collects both equally weighted and value weighted versions of return series, and towards the end of the files are annual figures. To use, the easiests is to remove the annual series, and split the files into two, one equally weighted and one value weighted.

In the examples below use an example edited file that only contains monthly (ew) returns.

R file:

```
##
## read Ken French data. In this example, read 5 industry portfolios and the Fama-French 3 factors
##
library(xts)
```

```
Sys.setlocale(category = "LC_ALL", locale = "C")
filename <- "5_Industry_Portfolios_ew.csv"
FF5IndusEW <- read.table(filename,
                          header=TRUE,skip=10,
                          na.strings=c(" -99.99","-99.99","-999"),
                         sep=",")
                                        # convert to time series
dates <- as.yearmon(as.character(FF5IndusEW[,1]),format="%Y%m")
FF5IndusEW <- xts(coredata(FF5IndusEW[,2:6]), order.by=dates)
                                        # data is in percent, divide by 100 to get returns
FF5IndusEW <- FF5IndusEW / 100
                                        # now get 3 factors
filename <- "F-F_Research_Data_Factors.csv"

FF <- read.table(filename,
                 header=TRUE,
                 skip=3,
                 na.strings=c("-99.99","-999"),
                 sep=",")
                                        # make time series
dates <- as.yearmon(as.character(FF[,1]),format="%Y%m")
FF <- xts(coredata(FF[,2:5]), order.by=dates)
                                        # make returns from percentage returns
RMRF <- FF$Mkt.RF/100
SMB  <- FF$SMB/100
HML  <- FF$HML/100
RF   <- FF$RF/100
                                        # the market return is the risk free rate plus the market risk p
RM   <- RF + RMRF
names(RMRF) <- "RMRF"
names(RM) <- "RM"
names(RF) <- "RF"
names(HML) <- "HML"
names(SMB) <- "SMB"
```

5 The Norwegian Ken French type asset pricing data

For Norwegians, there is an equivalent series available at ba-odegaard.no

How to prepare this data for use:

```
library(xts)
Sys.setlocale(category = "LC_ALL", locale = "C")
filename <- "industry_portfolios_monthly_vw.txt"
IndusRet <- read.table(filename, header=TRUE, sep=",")
                                        # convert to time series
dates <- as.Date(as.character(IndusRet[,1]),format="%Y%m%d")
months <- as.yearmon(dates)
                                        # note: actual returns, not percentage returns
IndusRet <- xts(coredata(IndusRet[,2:11]),order.by=months)
filename <- "market_portfolios_monthly.txt"
```

```
MarketRet <-read.table(filename,header=TRUE,sep=",")
months <- as.yearmon(as.Date(as.character(MarketRet$date),format="%Y%m%d"))
                                        # equally and value weighted market portfolios
Rm_ew <- xts(MarketRet$EW,order.by=months)
Rm_vw <- xts(MarketRet$VW,order.by=months)
                                        #read monthly risk free rates
filename <- "Rf_monthly.txt"
Rf <-read.table(filename,header=TRUE,sep=",",skip=1)
months <- as.yearmon(as.Date(as.character(Rf$date),format="%Y%m%d"))
Rf <- xts(Rf[,2],order.by=months)
Rf <- lag(Rf,1)               # these are forward looking interest rates, but for asset pricing purposes
names(Rf) <- "Rf"              # match with the month for which this interest rate is the risk free retu
                                        # read pricing factors
filename <- "pricing_factors_monthly.txt"
factors <- read.table(filename,header=TRUE,sep=",")
dates <- as.Date(as.character(factors[,1]),format="%Y%m%d")
months <- as.yearmon(dates)
factors <- xts(coredata(factors[,2:6]),order.by=months)
SMB  <- na.omit(factors$SMB)
HML  <- na.omit(factors$HML)
UMD  <- na.omit(factors$UMD)
LIQ  <- na.omit(factors$LIQ)
```