Quantitative Portfolio Management

Jon Spinney

2018-03-10

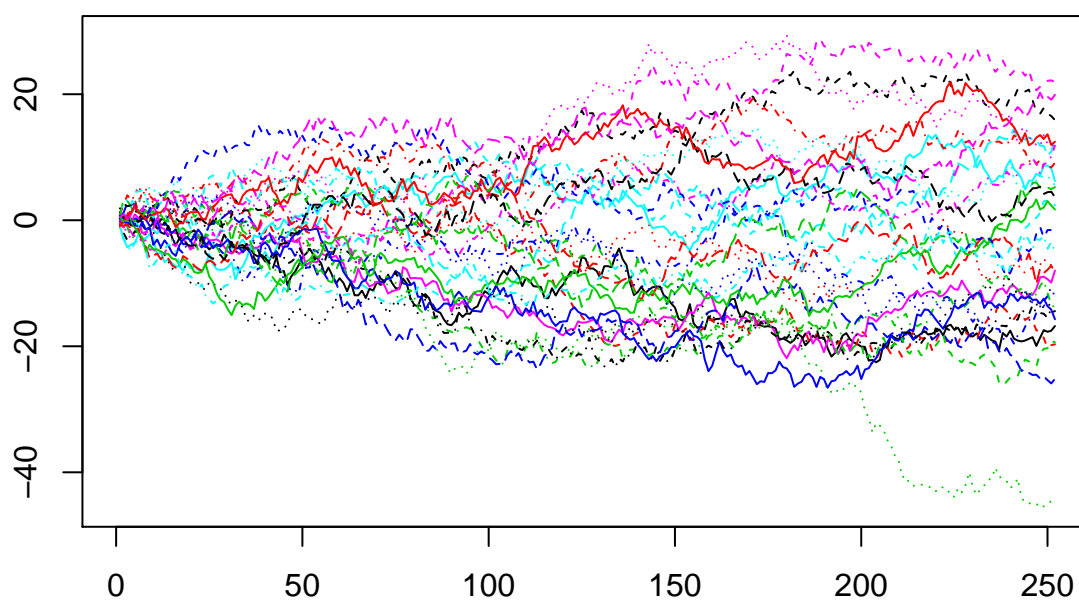# Contents

Chapter 1

Preliminary Material

30 Draws of a Random Walk

This book contains the course notes for the class **MQIM 6604 - Quantitative Portfolio Management** for the Master in Quantitative Investment Management (MQIM) at the University of New Brunswick. All required reading material for the course is contained in this book (there is no other required textbook).

Chapter 2

Introduction

This book contains course notes supporting the class MQIM 6604 - Quantitative Portfolio Management in the Master in Quantitative Investment Management (MQIM) at the University of New Brunswick in Fredericton. This book, in addition to the slides and supplementary reading assigned in lectures, are the only material required for the class.

While there are a plethora of commercially available books related to the field of quantitative equity portfolio management, this book contains a few primary advantages for MQIM students. Namely:

- The book is new, and constantly evolving, allowing it to continually reflect the current practice of quantitative investment management in the field today. While (Grinold and Kahn, 2000) may be considered the longstanding "bible" of active quantitative equity portfolio management and would be an excellent addition to the bookshelf of any practicing quant, it was last updated nearly 20 years ago.
- The book is targeted to be at exactly the correct level for students in the MQIM program, immediately working with appropriate levels of mathematical sophistication for the technically minded students in the program. Conversely, commercially available texts may be at a level that is too low or two high for the MQIM program.
- The R language (which is the primary language of instruction in the UNB MQIM program) is embedded throughout the book, and indeed the book is written in R using markdown and the Bookdown package (Xie, 2015), making this the only book on quantitative equity portfolio management to make such extensive use of R.

2.1 Commercially Available Textbooks Related to Quantitative Equity Portfolio Management

Students that are particularly interested in the field (or are looking for additional companion material to help with 6604) could look at any of the following books:

Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk (Grinold and Kahn, 2000)

- https://www.amazon.ca/Active-Portfolio-Management-Quantitative-Controlling/dp/0070248826
- Generally considered to be the "bible" of quantitative equity portfolio management - contains early references to the authors' work on the Fundamental Law of Active Management. Essential reading for practicing quants.

Risk and Asset Allocation (Meucci, 2005)

- https://www.amazon.ca/Risk-Asset-Allocation-Attilio-Meucci/dp/3642009646/

- Quite technical and idiosyncratic in its approach. The author provides a significant amount of Matlab code related to book examples (partially ported to R in the **meucci** packge) and runs an annual 1-week **Advanced Risk and Portfolio Management ("ARPM")** bootcamp in New York every summer https://www.arpm.co/bootcamp/

Quantitative Equity Portfolio Management: An Active Approach to Portfolio Construction and Management (Chincarini and Kim, 2006)

- https://www.amazon.ca/Quantitative-Equity-Portfolio-Management-Construction/dp/0071459391/
- Targeted at MBA students.

Quantitative Equity Portfolio Management: Modern Techniques and Applications (Qian et al., 2006)

- https://www.amazon.ca/Quantitative-Equity-Portfolio-Management-Applications/dp/1584885580/
- More advanced and more modern than (Grinold and Kahn, 2000), a good read for practicing quants.

Quantitative Equity Investing: Techniques and Strategies (Fabozzi et al., 2010)

- https://www.amazon.ca/Quantitative-Equity-Investing-Techniques-Strategies/dp/0470262478/

Modern Investment Management: An Equilibrium Approach (Litterman, 2003)

- https://www.amazon.ca/Modern-Investment-Management-Equilibrium-Approach/dp/0471124109/
- Written/edited by Robert Litterman and the (at that time) Quantitative Strategies Group of Goldman Sachs Asset Management.

Asset Management: A Systematic Approach to Factor Investing (Ang, 2014)

- https://www.amazon.ca/Asset-Management-Systematic-Approach-Investing/dp/0199959323/
- The most recent of all the commercially available books on this list, this book is non-technical (appropriate for MBA & business school Undergraduate students) but makes an excellent companion book forMQIM 6604 student for extra explanatory material and practical real-world examples of many applications of factor investing and quantitative approaches to portfolio management in general.

2.2   Course Description and Background Information

TBC

test table using kable:

```
knitr::kable(
  head(mtcars[, 1:8], 10), booktabs = TRUE,
  caption = 'A table of the first 10 rows of the mtcars data.'
)
```

2.3   Introduction to R & Exploratory Data Analysis

TBC

Table 2.1: A table of the first 10 rows of the mtcars data.

	mpg	cyl	disp	hp	drat	wt	qsec	vs
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1

Chapter 3

Modern Portfolio Theory

Notes:

- Derive eff front & cal
- equilibrium expected returns

3.1 Markowitz

TBC

3.2 CAPM

Notes:

- TBC

3.3 APT & Multifactor Models

TBC

3.4 Modern Approaches

TBC

Chapter 4

Anomalies & Alpha Strategies

TBC

4.1 Early Anomalies Literature

TBC

4.2 Fama French (1990s)

TBC

4.3 Carhart

TBC

4.4 Other Factors

BAB, QMJ, Fama French 2000s

Chapter 5

Forecasting

TBC

5.1 Fractile Zero-Investment Portfolios

TBC

5.2 Factor Mimicking Portfolios

TBC

5.3 Fama-Macbeth Regressions

TBC

5.4 Optimal Predictors

TBC

Chapter 6

Portfolio Optimization

TBC

6.1 Minimum Variance Portfolios

TBC

6.1.1 Case Study: Minimum Variance Portfolios

TBC

6.2 Constrained Optimization

TBC

6.3 Active Portfolio Optimization

TBC

6.4 Modern Optimization

TBC

Chapter 7

Risk Models

TBC

7.1 Risk Measures

- Variance & Volatility
- Semi-deviation
- VaR
- CVaR

7.2 Covariance Matrices and Bayesian Shrinkage Models

TBC

7.3 Principal Components Analysis

TBC

7.4 Fundamental Factor Models

TBC

- Ex 1: Single Index Model
- Ex 2: Industry MOdel
- Ex 3: Barra Style Model

Chapter 8

Post-Modern Portfolio Theory

The original Markowitz spproach to portfolio construction, which we implemented via the Quadratic Utility approach of:

$$max_\omega \quad \omega'\mu - 2^{-1}\lambda^{-1}\omega'\Sigma\omega \quad \text{s.t. } \omega'\mathbf{1} = 1$$

is typically referred to as **Modern Portfolio Theory ("MPT")**. MPT requires the estimation of asset means and co-variances, and trades off return vs. risk in an optimization framework. While the approach is intuitive in its approach, practitioners have found numerous challenges in the implementation of the MPT framework, including:

- unintuitive holdings,
- highly concentrated portfolios, and
- excess turnover.

Each of these issues leads to a common phenomenon of MTP optimal portfolios producing inferior results out of sample relative to more naive approaches. A significant contributor to the challenge of implementing MPT approaches is the combination of the extreme sensitivity of the MPT approach to errors in the mean vector, combined with the real-world challenge of estimating such a noisy parameter. This extreme sensitivity arises as a result of the *full investment constraint* in the optimization process. These challenges have resulted in numerous adaptations of the MPT approach in practice, including:

- Risk Parity Investing,
- Diversification/Concentration Management, and
- Alternative Risk Definitions (e.g. Mean-CVaR) that focus on downside risk as opposed to the symmatric nature of variance/volatility.

Collectively, these and other more recent approaches to portfolio construction can be referred to as **Post-Modern Portfolio Theory ("PMPT")**.

8.1 Risk Parity

Risk parity terminology and meaning often depends on the user, but for the purposes of this book, we define two classes of risk parity approaches to investing - **inverse volatility ("IV")**, which is sometimes also referred to as *naive risk parity*, and **equal risk contribution ("ERC")**. IV as a weighting scheme that weights securities or asset classes proportional to the inverse of their volatility, whereas ERC is a more complicated approach to portfolio construction that results each security's contribution to total portfolio risk being equal. IV strategies therefore consider only the volatilities of the assets in the universe, while ERC strategies consider also the covariance structure between all the assets in the portfolio.

As a visualization, consider a standard 60-40 portfolio of stocks and bonds. Using recent data, the volatility of the equity portion was about 10%, while the volatility of the bond portfolio was closer to 3.5% - the portfolio itself had to total combined volatility of about 6% (the correlation between bonds and stocks during this period was 0). How much of the portfolio total volatility could be attributed to the two components?

Asset Class	Weight (%)	Volatility (%)	% Contrib. to Port Risk
Equities	60%	10.2%	99.4%
Bonds	40%	3.5%	0.6%
Total Portfolio	100%	6.3%	100%

Consider instead if we had used a portfolio approach that sought to equalize the **% Contrib. to Port Risk** at 50% for each asset class:

Asset Class	Weight (%)	Volatility (%)	% Contrib. to Port Risk
Equities	25.5%	10.2%	50%
Bonds	74.5%	3.5%	50%
Total Portfolio	100%	3.7%	100%

The **Equal Risk Contribution** portfolio has just 25% of its assets in the higher risk equity portfolio and 75% of its assets in the relatively lower volatility bond portfolio. The portfolio, with a total risk of 3.7%, derives 50% of its volatility from bonds and 50% from equities.

There are two main quantitative approaches to risk parity - naive risk parity that ignores correlations and weights stocks proportionally to the inverse of their volatility forecasts and equal risk contribution portfolios that take into account the full covariance structure of the market in setting the % Contrib. to Port Risk to be equal for all assets.

8.1.1   Inverse Volatility Weighting (Naive Risk Parity)

For IV, we first estimate the volatility of the assets in the universe and then assign weights to each asset in the resulting portfolio such that:

$$\omega_i = \frac{\sigma_i^{-1}}{\sum_{i=1}^{k} \sigma_i^{-1}}$$

In the IV portfolio construction process, assets with lower volatilities are assigned higher weights, and assets with higher volatilities are assigned lower weights.

```r
library(quantmod)
tickers <- c("F","MMM","MSFT")
getSymbols(tickers)

## [1] "F"    "MMM"  "MSFT"

dat <- merge.xts(F$F.Adjusted, MMM$MMM.Adjusted, MSFT$MSFT.Adjusted)
ret <- (exp(diff(log(dat)))-1)[-1,]
vols <- apply(ret,2,sd)
wgts_rp <- (1/vols)/sum(1/vols)
names(wgts_rp) <- tickers
barplot(wgts_rp,
        main="Weights of Naive Risk Parity Portfolio for F, MMM and MSFT",
```

```
        xlab="Stock",
        ylab="Weight in Portfolio")
```

Weights of Naive Risk Parity Portfolio for F, MMM and MSFT
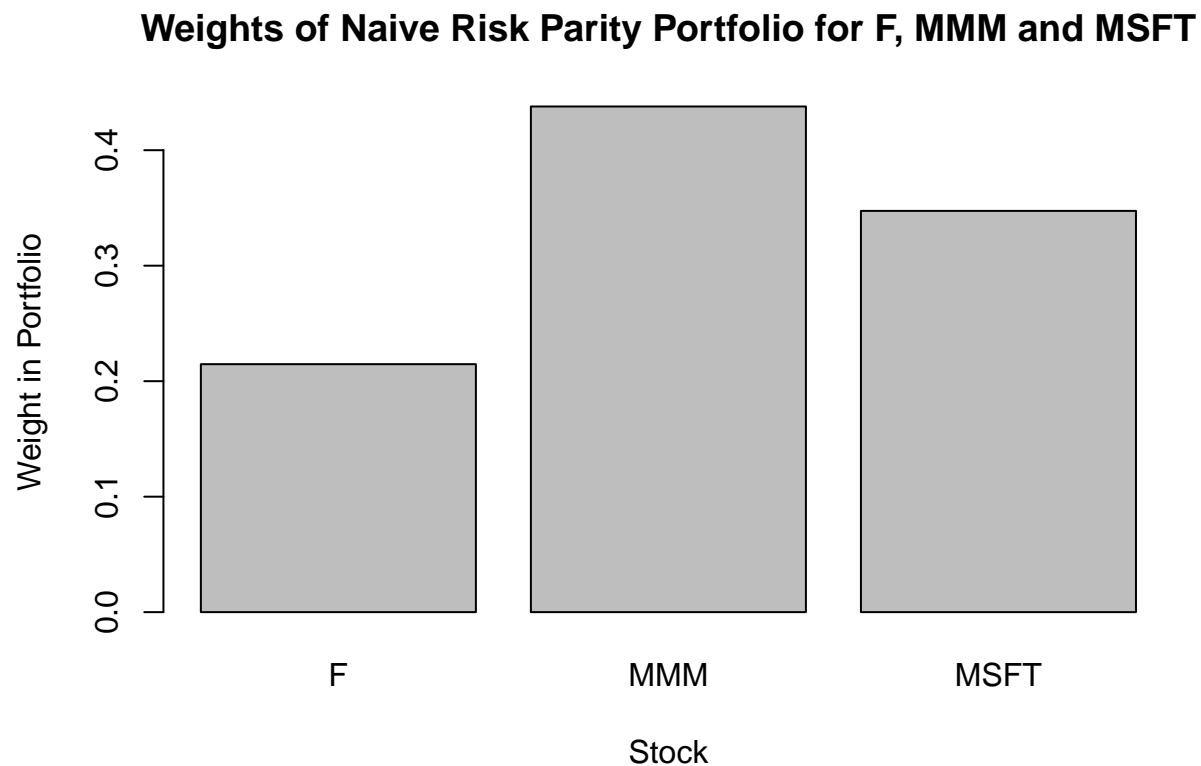
8.1.2 Equal Risk Contributions

A more complicated approach to risk parity is the one that chooses asset weights such that the contribution of each asset to the total portfolio's risk is equal. Like IV approaches, ERC applies relatively higher weights to lower risk assets and vice versa, but unlike IV approaches, the ERC methodology considers the full covariance structure of the market, not just the asset variances.

The ERC approac of (Maillard et al., 2010) requires a solution to the following:

$$TRC_i = TRC_j \quad \forall_{i,j}$$

Where

$$TRC_i = \omega_i \frac{\partial \sigma_p}{\partial \omega_i}$$

This ERC problem can be solved via a newton approach (Algorithm 1 in (Chaves et al., 2012)). The R function **SolveRiskParity()** computes ERC portfolio weights for a given covariance matrix (function code not shown).

```
library(tseries)
library(quantmod)
```

```
tickers <- c("F","MMM","MSFT")
getSymbols(tickers)
```

```
## [1] "F"     "MMM"  "MSFT"
```

```
dat <- merge.xts(F$F.Adjusted, MMM$MMM.Adjusted, MSFT$MSFT.Adjusted)
ret <- (exp(diff(log(dat)))-1)[-1,]
sigma <- cov(ret)
erc_soln <- SolveRiskParity(sigma)
wgts <- erc_soln$port_wgt[,1]
names(wgts) <- tickers
barplot(wgts,
        main="Weights of ERC Portfolio for F, MMM and MSFT",
        xlab="Stock",
        ylab="Weight in Portfolio")
```
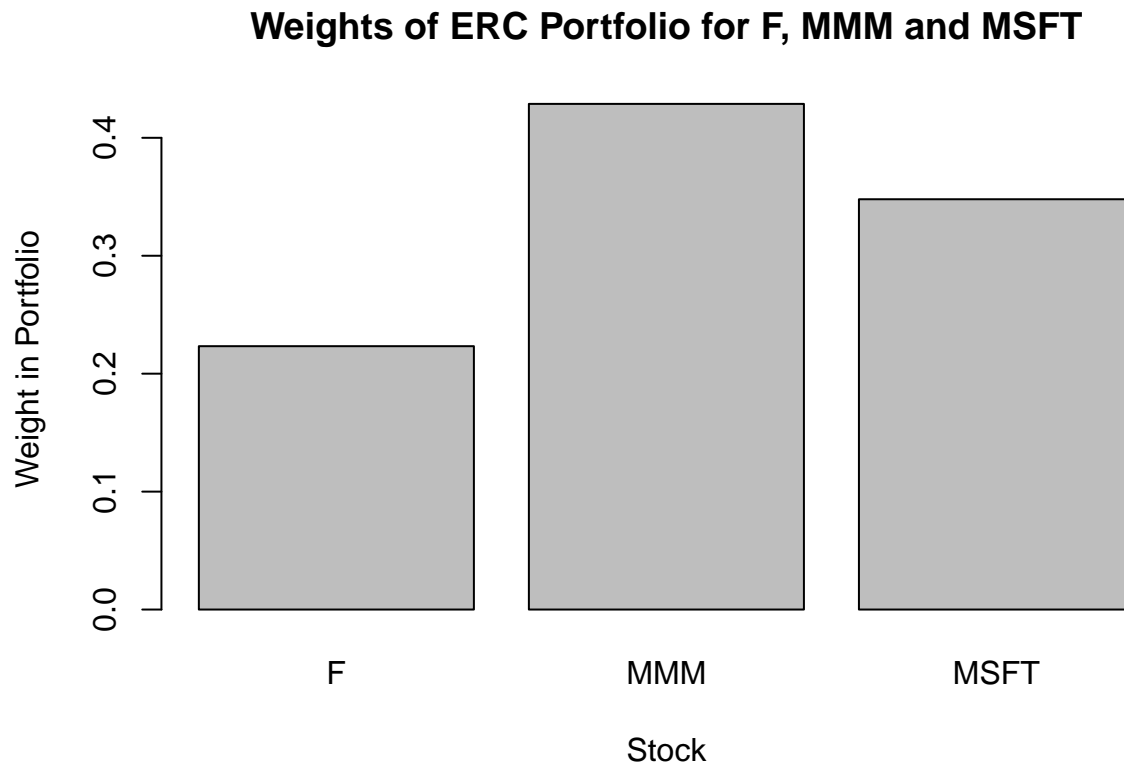

We can confirm that the risk contribution is equal for each asset by verifying that $\omega_i \times \frac{\partial \sigma_p}{\partial \omega_i} = \omega_j \times \frac{\partial \sigma_p}{\partial \omega_j} \quad \forall\, i, j$:

```
contribs <- erc_soln$pctr[,1]
names(contribs) <- tickers
barplot(contribs,
        main="Pct. Contrib to Port. Risk of ERC Portfolio for F, MMM and MSFT",
        xlab="Stock",
        ylab="PCTR")
```

Pct. Contrib to Port. Risk of ERC Portfolio for F, MMM and MSFT
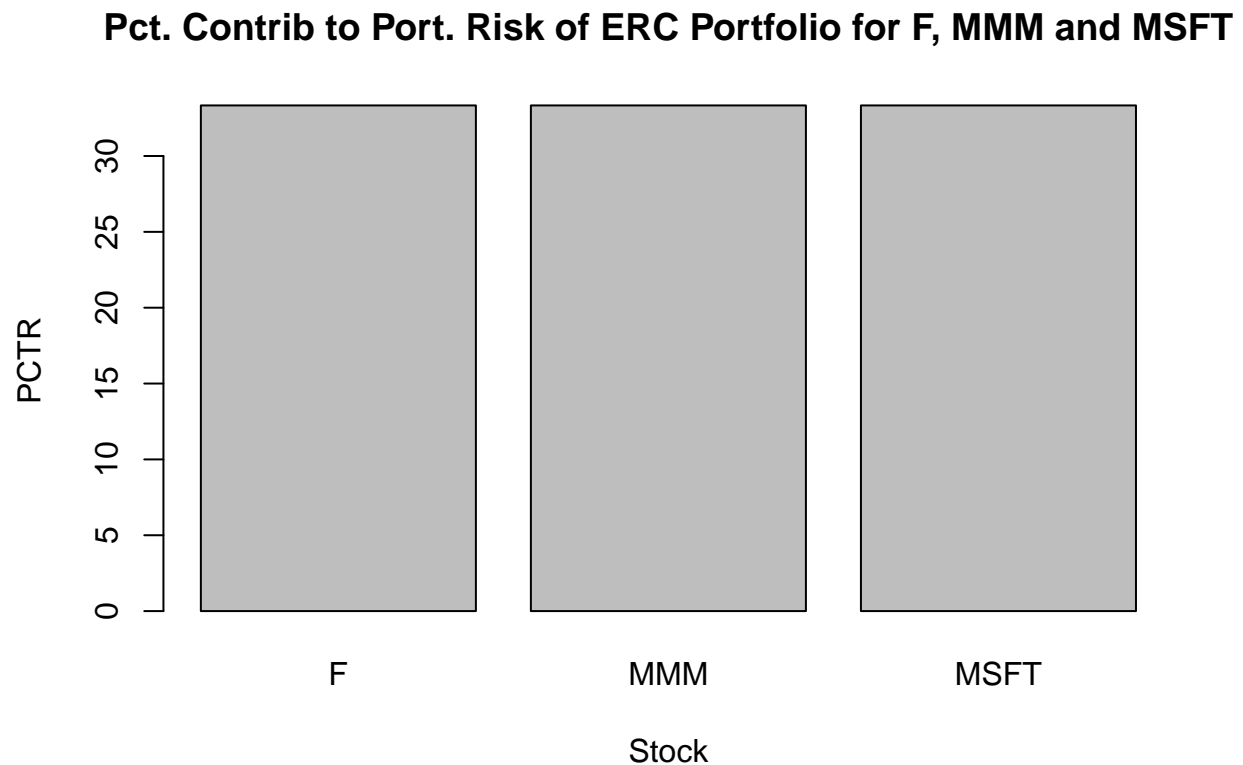
In general, ERC portfolios have reduced risk relative to a market capitalization weighted or equally weighted benchmark due to the relative overweighting of assets with lower volatility and relative underweighting of assets with relatively higher volatility - there is on average a negative relationship between an asset's volatility and its weight in the ERC portfolio:

```r
library(calibrate)
vols <- apply(ret,2,sd)*sqrt(252)
plot(vols,wgts,
     main="Stock Volatility vs. Weight in ERC Portfolio",
     xlab="Volatility",
     ylab="Weight in Portfolio",
     pch=19,
     col="blue"
     )
textxy(vols,wgts, labs=tickers, cx = 1, dcol = "black", m = c(-1, -1))
```

Stock Volatility vs. Weight in ERC Portfolio
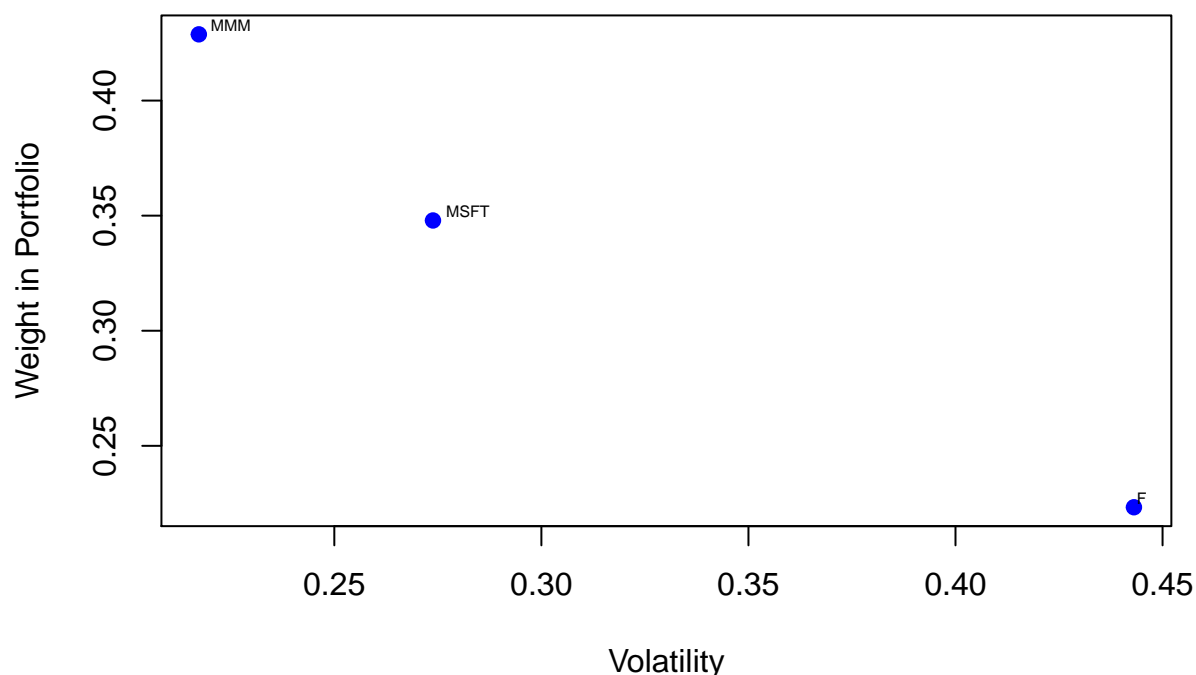
8.2   Concentration/Diversification

Mean-Variance optimal portfolios are often highly concentrated, with potentially very large positions in the most attractive assets:

```r
library(quantmod)
library(quadprog)
tickers <- c("MMM","AXP","AAPL","BA","CAT","CVX","CSCO",
             "KO","DIS","DWDP","XOM","GE","GS","HD","IBM",
             "INTC","JNJ","JPM","MCD","MRK","MSFT","NKE",
             "PFE","PG","TRV","UTX","UNH","VZ","V","WMT")
getSymbols(tickers)
```

```
##  [1] "MMM"  "AXP"  "AAPL" "BA"   "CAT"  "CVX"  "CSCO" "KO"   "DIS"  "DWDP"
## [11] "XOM"  "GE"   "GS"   "HD"   "IBM"  "INTC" "JNJ"  "JPM"  "MCD"  "MRK"
## [21] "MSFT" "NKE"  "PFE"  "PG"   "TRV"  "UTX"  "UNH"  "VZ"   "V"    "WMT"
```

```r
dat <- MMM$MMM.Adjusted
for (i in 2:length(tickers)) {
  obj <- paste(tickers[i],"$",tickers[i],".Adjusted",sep="")
  dat <- merge(dat,eval(parse(text=obj)))
  }
colnames(dat) <- tickers
dat <- dat['2008-03-19/']
```

```r
ret <- (exp(diff(log(dat)))-1)[-1,]

sigma <- cov(ret)

exp_ret <- array(0,dim=c(30,1))
const <- array(1,dim=c(30,1))
Amat <- cbind(const, diag(nrow(sigma)))
bvec <- c(1, rep(0, nrow(sigma)))
port_wgts <- solve.QP(sigma,exp_ret,Amat,bvec=bvec,meq=1)$solution
names(port_wgts) <- tickers

barplot(port_wgts,
        main="Unconstrained Minimum Variance Portfolio Weights",
        ylab="Weight in MVP",
        las=2)
```

Unconstrained Minimum Variance Portfolio Weights
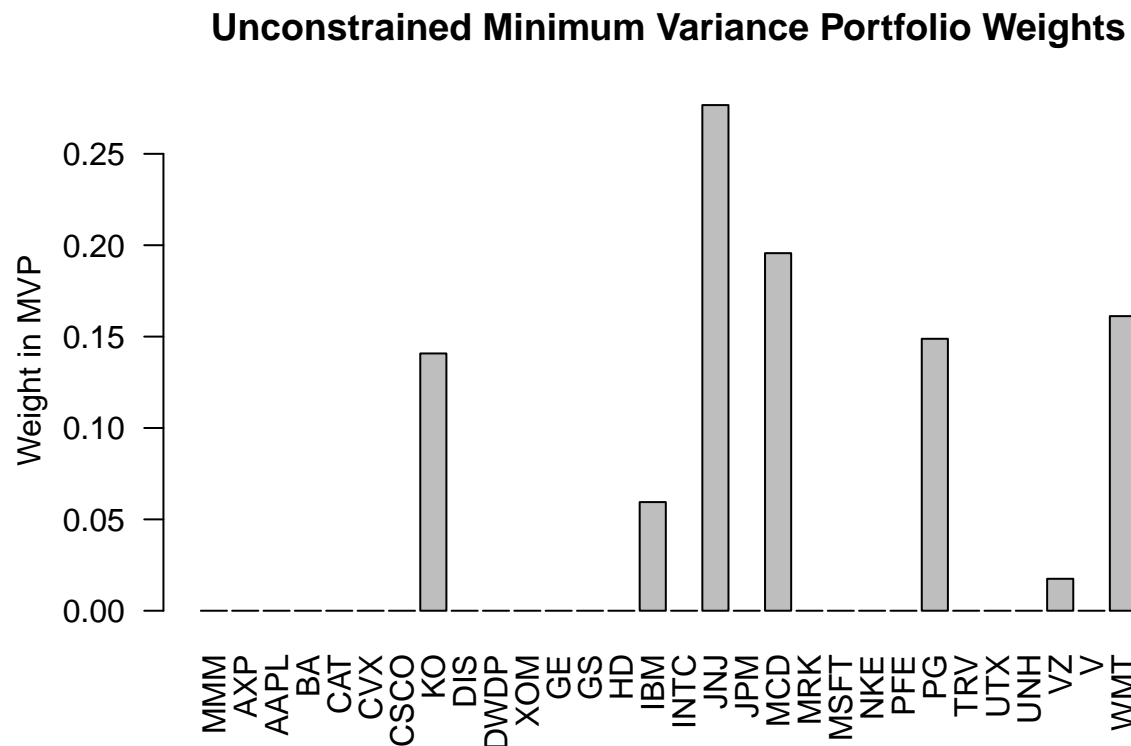
Using a representative sample of 30 large cap U.S. stocks, an unconstrained MVP results in the portfolio holding as much as 20-30% in certain stocks.

Two methods to prioritize diversification/deconcentration in the optimization process include the **Most Diversified Portfolio ("MDP")** approach of (Choueifaty and Coignard, 2008) as well as the direct modification of the objective function using the **Herfindahl-Hirschman Index ("HHI")**.

8.2.1 Most Diversified Portfolio

The MDP is the portfolio that maximizes the diversification ratio of the portfolio, defined as the ratio of the weighted average volatility of the portfolio constituents to the volatility of the portfolio:

$$DR = \frac{\omega' v}{\sqrt{\omega' \Sigma \omega}}$$

Where v is a vector of individual asset volatilities. Maximization of the diversification ratio is equivalent to the following optimization objective function:

$$min_\omega \quad \omega' C \omega \quad \text{s.t. } \omega' \mathbf{1} = 1$$

with C the correlation matrix of assets in the investment universe.

```r
library(quantmod)
tickers <- c("F","MMM","MSFT")
getSymbols(tickers)
```

```
## [1] "F"    "MMM"  "MSFT"
```

```r
dat <- merge.xts(F$F.Adjusted, MMM$MMM.Adjusted, MSFT$MSFT.Adjusted)
ret <- (exp(diff(log(dat)))-1)[-1,]
C <- cor(ret)
mdp <- as.numeric((solve(C)%*%rep(1,3))/as.numeric(rep(1,3)%*%solve(C)%*%rep(1,3)))
names(mdp) <- tickers
barplot(mdp,
        main="Weights of Most Diversified Portfolio",
        xlab="Stock",
        ylab="Weight")
```

Weights of Most Diversified Portfolio
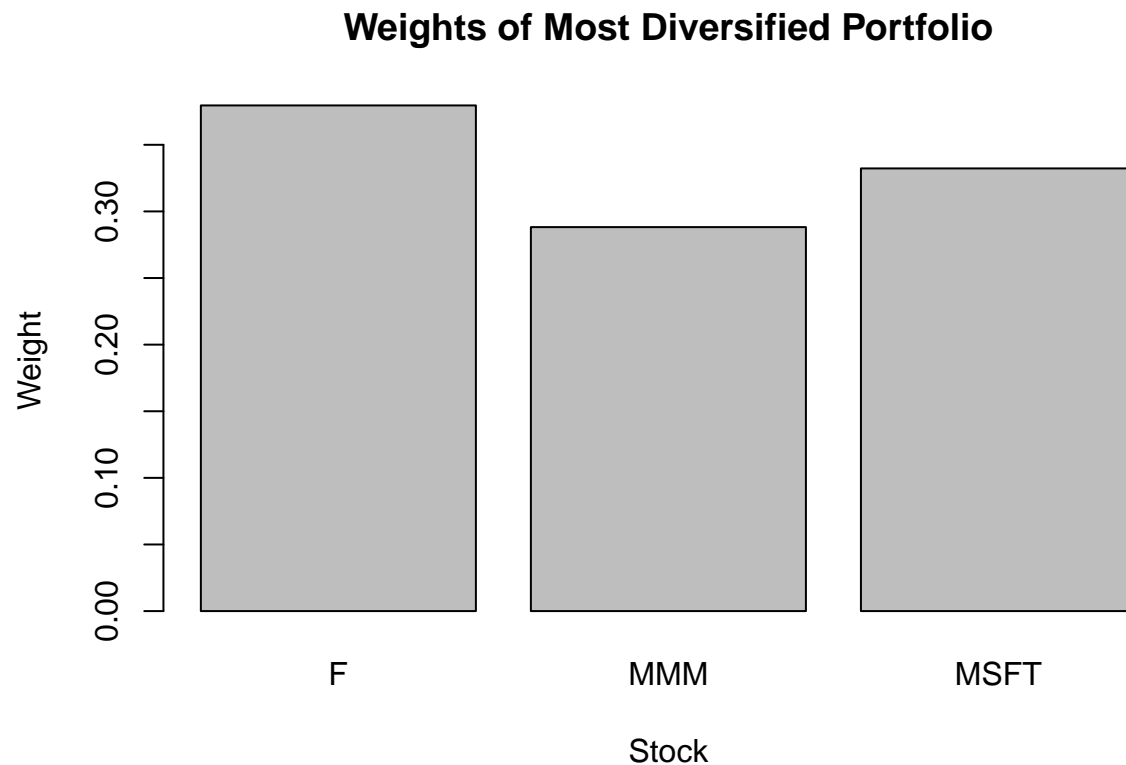
8.2.2 HHI Based Deconcentration

The Herfindahl-Hirschman Index (defined as $HHI(\omega) = \sum_{i=1}^{N} \omega_i^2$) is a measure of market concentration and was applied in a portfolio optimization context in (King, 2007).The approach generalizes to that of the use of a modified covariance matrix in the optimization objective function. For a minimum variance portfolio with a concentration penalty:

$$\min_{\omega} \quad \omega'(\Sigma + \lambda_c \mathbf{I})\omega \quad \text{s.t. } \omega'\mathbf{1} = 1$$

Where:

- $\mathbf{I}$ is an identity matrix
- λ_c is a *concentration risk aversion* parameter

Larger values of λ_c produce portfolios that are less concentrated in nature, while a value of $\lambda_c = 0$ results in the unconstrained minimum variance formulation (no weight given to avoidance of concentration risk).

Using a minimum variance portfolio as an example, we can observe the impact of a concentration penalty on the resulting weight of a 3 asset portfolio:

```
library(tseries)
library(quantmod)
tickers <- c("F","MMM","MSFT")
getSymbols(tickers)

## [1] "F"    "MMM"  "MSFT"
```

```r
dat <- merge.xts(F$F.Adjusted, MMM$MMM.Adjusted, MSFT$MSFT.Adjusted)
ret <- (exp(diff(log(dat)))-1)[-1,]
sigma <- cov(ret)
lambda_c <- seq(0,0.01,length.out=50)
port_wgts <- array(0,dim=c(3,50))
for (i in 1:dim(port_wgts)[2]) {
  sigma_adj <- sigma + lambda_c[i]*diag(3)
  port_wgts[,i] <- {(solve(sigma_adj)%*%rep(1,3))/
      as.numeric(rep(1,3)%*%solve(sigma_adj)%*%rep(1,3))}
}
colnames(port_wgts) <- as.character(round(lambda_c,digits=3))
par(mar=c(5.1, 4.1, 4.1, 8.1), xpd=TRUE)
barplot(port_wgts,
        main="Effect of Concentration Penalty on \n Minimum Variance Portfolio Weights",
        xlab="Value of Concentration Risk Aversion Parameter",
        ylab="Weight",
        col=c("grey","lightblue","lightgreen")
        )
legend("topright", inset=c(-0.2,0), legend=tickers,col=c("grey","lightblue","lightgreen"),lty=1)
```
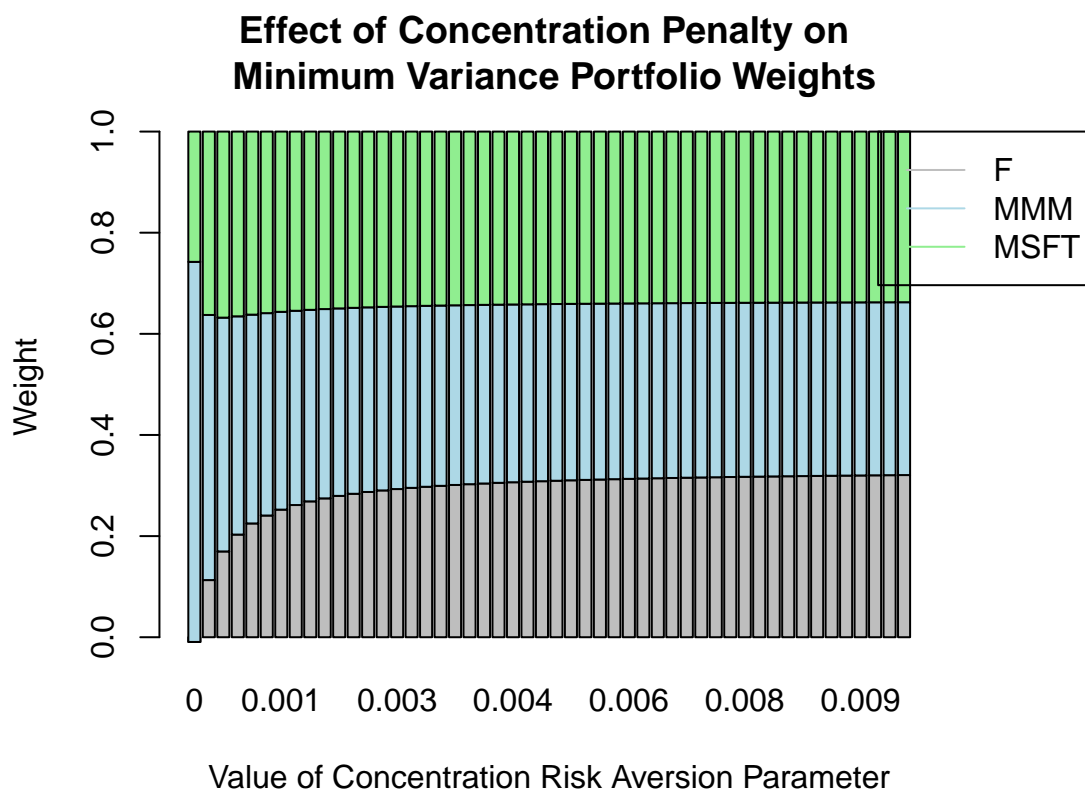

8.3 Mean-CVaR

While the standard MPT optimization approach uses variance as the measure of risk in the objective function:

$$\max_{\omega} \quad \omega'\mu - \frac{1}{2}\lambda\omega'\Sigma\omega$$

(Rockafellar and Uryasev, 2000) replaced variance with the downside risk measure **Conditional Value-at-Risk ("CVaR")** - sometimes referred to as Expected Tail Loss ("ETL"):

$$CVaR_\alpha = -\mathbb{E}(r|r \leq q^r(\alpha))$$

Aplying this in a quadratic utility framework results in the objective function

$$\max_{\omega} \quad \omega'\mu - \lambda \cdot \text{CVaR}_\alpha(\omega) \quad \text{s.t. } \omega'\mathbf{1} = 1$$

(TODO): Finish CVaR Section with example

8.4   Synthesis of PMPT Approaches

(TODO): Results from paper showing optimality, etc., reasons why for each approach.

Appendix 1 - Special Topics

TBC

Hedge Fund Replication

TBC

Robustness

TBC

Appendix 2 - Background Information

Matrix Algebra Review

Basic portfolio mathematics is much easier to understand (and write clearly) if the reader has understanding of some simple matrix algebra techniques. For example, the variance σ_p^2 of portfolio of two assets written out in fully expanded form is often expressed as the sum of three separate terms:

$$\sigma_p^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 \sigma_2 \sigma_2 \rho_{1,2} \tag{8.1}$$

For a three asset portfolio, we need to take the sum of six separate terms (in fact, for n assets, we would have to take the sum of $\frac{n(n+1)}{2}$ terms). Conversely, in matrix language this would be written as simply $\sigma_p^2 = \omega' \Sigma \omega$ irrespective of the number of assets in the portfolio.

This appendix covers the very basics of matrix mathematics required to understand the above and succeeed in this class - including basic matrix computational techniques and an introduction to finding inverse matrices. Further coverage is available in Dr. Eric Zivot's *Matrix Algebra Review* available at http://faculty.washington. edu/ezivot/econ424/matrixreview.pdf, or alternatively for a textbook treatment, the first few chapters of Chiang et al *Fundamental Methods of Mathematical Economics* contains most of what would be necessary for this course.

Basic Concepts

Matrix addition is defined for same-dimension matrices:

$$\begin{bmatrix} 1 & 4 \\ 3 & 7 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 4 & 8 \end{bmatrix}$$

A matrix can be multiplied by a scalar by multiplying that scalar by each element of the matrix:

$$2 \times \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 2 & 2 \end{bmatrix}$$

The dot-product of two conformable vectors is the sum of the corresponding elements multiplied together:

$$\begin{bmatrix} 1 & 4 \end{bmatrix} \times \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 1 \times 2 + 4 \times 1 = 6$$

Matrix multiplication is defined in a similar way, in that matrices are conformable when the number of columns in the first matrix is the same as the number of rows in the second:\

$$\begin{bmatrix} 1 & 4 \\ 3 & 7 \end{bmatrix} \times \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \times 2 + 4 \times 1 & 1 \times 3 + 4 \times 1 \\ 3 \times 2 + 7 \times 1 & 3 \times 3 + 7 \times 1 \end{bmatrix} = \begin{bmatrix} 6 & 7 \\ 13 & 16 \end{bmatrix}$$

Division of matrices is not concept that is defined, although a similar idea is contained within the concept of inverse matrices.

Inverse Matrices

The Identity Matrix I is simply a matrix with 1's in the diagonal elements and 0's elsewhere:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

If you check, you can see that any matrix X multiplied with the Identity matrix results in the original matrix X - that is, $XI = IX = X$ and $II = I$. Another interesting matrix is that of the Inverse matrix. For any matrix X, its inverse is denoted X^{-1}, and an any matrix multiplied by its inverse results in the Identity Matrix - that is, $XX^{-1} = X^{-1}X = I$. This is therefore analogous to division for matrices, and will prove to be useful.

Now, consider a system of equations - three equations and three unknowns - below:

$$3x_1 + 5x_2 + x_3 = 29$$
$$7x_1 + 2x_2 + 4x_3 = 60$$
$$-6x_1 + 3x_2 + 2x_3 = -4$$

How can we solve this system of equations (that is, find the values of x_1, x_2, x_3 that make the left and right hand sides equal)? The method of Gauss-Jordan elimination is often taught, but this is tedious even for 3×3 systems (there are shortcut methods for 2×2 matrices), and so we ignore the details behind it and skip forward to computational implementations. It turns out that we can re-write the system equations as follows in matrix form, and that through the use of the inverse matrix, we can easily solve such a system. If we write:

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 7 & 2 & 4 \\ -6 & 3 & 2 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} b = \begin{bmatrix} 29 \\ 60 \\ 4 \end{bmatrix}$$

Then the equivalent system in matrix form is written $Ax = b$. The question is how to find the values of x that make $Ax = b$. Because of the properties of the Identity Matrix I and the Inverse Matrix we defined above, we are able to take the following steps:

$$Ax = b$$
$$A^{-1}Ax = A^{-1}b$$
$$Ix = A^{-1}b$$
$$x = A^{-1}b$$

And so we can solve any such system of equations (provided a solution exists, which is essentially equivalent to the existence of the Inverse Matrix for the Matrix A) by simply pre-multiplying the b vector by the matrix A^{-1}.

Note that in R, it is often recommended to use the *solve()* function to find the matrix inverse. You can verify the above in R as follows:

```r
(A = matrix(c(3, 5, 1, 7, 2, 4, -6, 3, 2),nrow=3,byrow=T))
```

```
##      [,1] [,2] [,3]
## [1,]    3    5    1
## [2,]    7    2    4
## [3,]   -6    3    2
```

```r
(b = matrix(c(29,60,-4),nrow=3,byrow=T))
```

```
##      [,1]
## [1,]   29
## [2,]   60
## [3,]   -4
```

```r
(A.inv <- solve(A))
```

```
##             [,1]        [,2]        [,3]
## [1,]  0.0441989  0.03867403 -0.09944751
## [2,]  0.2099448 -0.06629834  0.02762431
## [3,] -0.1823204  0.21546961  0.16022099
```

```r
(x <- A.inv %*% b)
```

```
##      [,1]
## [1,]    4
## [2,]    2
## [3,]    7
```

And in fact we get the proper solution of $x = (4, 2, 7)$, that is, $x_1 = 4, x_2 = 2, x_3 = 7$.

Sums

However, despite the fact that we will use matrices whenever appropriate, there are times when it is useful or more intuitive to use sums instead. For example, the sum of all integers from 1 to 100 could be written in summation notation as:

$$1 + 2 + ... + 100 = \sum_{i=1}^{100} i$$

Often, expressions can be written equivalently suing summation notation or using matrix notation, and we will often find summation notation convenient, particularly when encountering concepts like risk decomposition for the first time. For the purposes of this text, it will be helpful to remember that:

- $\mathbf{c}'\mathbf{x} = \sum_{i=1}^{n} c_i x_i$ when $\mathbf{c}$ and $\mathbf{x}$ are vectors of length n
- $\mathbf{c}'\mathbf{X}\mathbf{c} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j X_{i,j}$ when $\mathbf{c}$ is a vector of length n and X is an $n \times n$ matrix

Probability Review

This section covers only the bare minimum of the very basics of probability and expectation that is needed to complete this course. For a good probability review that does assume some prior knowledge, consult Eric Zivot's *Probability Review* available at http://faculty.washington.edu/ezivot/econ424/probreview.pdf. For more in depth coverage beginning with the basics in a textbook-like format, Grinstead and Snell's *Introduction to Probability* (Grinstead and Snell, 2006) is a good choice and is available free online.

Probability Distributions & Random Variables

A Random Variable X (hereafter: R.V.) is a function that maps each outcome from a sample space into a real number. For example, in the case of tossing a regular 6-sided die, the R.V. could take the values (1, 2, 3, 4, 5, 6).

A Probability Mass Function (hereafter: PMF) is a function $f_X(x)$ that maps the values of a R.V. X on to numbers between 0 and 1, inclusive. That is, the PMF gives the probability of the outcome of the R.V. occurring:

$$f_X(x) = \mathbb{P}(X = x) \tag{8.2}$$

Which can be read as "the probability that the R.V. X takes on the value x" (there are similar definitions for continuous R.V.'s, although we ignore them for this class). Using this definition, we can also define the cumulative probability mass as:

$$F_X(x) = \sum_{i=1}^{n} \mathbb{P}(X = x_i) \tag{8.3}$$

Where the total sum of all possible non-overlapping events in the sample space amounts to $\mathbb{P}(\Omega) = 1$.

Rules of Expectation & Variance

For a discrete random variable, we interpret the Expected Value ($\mathbb{E}$) of that random variable to be its average value in repeated sampling, and it is defined as the probability weighted sum of all possible outcomes. In mathematical notation, the expectation of the R.V. X is:

$$\mathbb{E}X = \sum_{i=1}^{n} x_i * \mathbb{P}(X = x_i) \tag{8.4}$$

For n equally likely events, this can be simplified to:

$$\mathbb{E}X = \frac{1}{n} * (\sum_{i=1}^{n} x_i) \tag{8.5}$$

For example, if we define a random variable X that pays \$1 if a (fair) coin comes up as heads, and -\$1 if the coin comes up as tails, then the expected value of that random variable is:

$$\mathbb{E}(X) = \frac{1}{2} * 1 + \frac{1}{2} * (-1)$$
$$= \frac{1}{2} * (1 + (-1))$$
$$= 0$$

There are various rules of expectations that we should know:

- The expected value of a constant c is that constant: $\mathbb{E}(c) = c$
- The expected value of a random variable X plus a constant c is the expected value of the random variable plus the constant: $\mathbb{E}(X + c) = \mathbb{E}(X) + c$
- The expected value of a constant a multiplied by a random variable X is the constant multiplied by the expectation of the random variable: $\mathbb{E}(aX) = a\mathbb{E}(X)\backslash$

In addition, there are other rules of expectation that are important to know, but not necessarily for this course, and so they are omitted. However, we also need to be aware of some similar rules or properties for variance and co-variance. Recall that the definition of the variance of a random variable is:

$$Var(X) = \mathbb{E}[(X - \mu_X)^2] \tag{8.6}$$

Properties:

- $Var(X + c) = Var(X)$
- $Var(aX) = a^2 Var(X)\backslash$

And likewise, the co-variance of two discrete random variables X and Y is:

$$Cov(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] \tag{8.7}$$

For which the following properties exist:

- $Cov(X + a, Y + b) = Cov(X, Y)$
- $Cov(aX, bY) = abCov(X, Y)\backslash$

And this in turn allows us to define the following rules:

- $Var(X + Y) = Var(X) + Var(Y) + 2 * Cov(X, Y)$
- $Var(\omega_1 X + (1 - \omega_1)Y) = \omega_1^2 Var(X) + (1 - \omega_1)^2 Var(Y) + 2 * \omega_1(1 - \omega_1)Cov(X, Y)$

Rule 2 should look familiar to anyone who has seen the expression for the variance of a 2 asset portfolio.

The variance of a random variable X is often denoted as σ_X^2 in many applications, and we should remember that the standard deviation is just the square root of the variance $\sigma_X = \sqrt{Var(X)}$. Finally, note that the correlation coefficient ρ is defined as:

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} \tag{8.8}$$

Calculus Review

In this course, we use calculus primarily to motivate the concept of portfolio optimization. Consequently we are concerned primarily with the concept of the derivative (of multivariate functions) and of the techniques of optimization. These topics are briefly reviewed here here in a highly informal/non-rigorous way. For a

more complete but easy to follow treatment of some applications that are necessary for this class, please consult a textbook such as Chiang and Wainwright's *Fundamental Methods of Mathematical Economics.*

Differentiation of Vector & Matrix Valued Functions

For functions of one variable, there are standard computational techniques for obtaining derivatives. For example, for powers of x:

$$f(x) = 3x^3 + 2x - 8$$
$$f'(x) = 9x^2 + 2$$
$$f''(x) = 18x$$
$$f^{(3)}(x) = 18$$

More a slightly more complicated function such as $y = 4(x^2 - 2)^3$, we can use the *chain rule*. If we let $u = x^2 - 2$:

$$y(u) = 4u^3$$
$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$
$$= (12u^2)(2x)$$
$$= 24x(x^2 - 2)^2$$

In the case when we have a function of more than one variable, we take *partial* derivatives by treating all variables that are not of interest as constants. Given a function like $f(x_1, x_2, x_3) = 2x_1^2 x_2 x_3 + 4x_1 x_2^2 + 3x_2 + 4x_3^2$, we would denote the first derivative w.r.t. x_1 as $f_{x_1}(x_1, x_2, x_3)$, the first derivative w.r.t x_2 as $f_{x_2}(x_1, x_2, x_3)$, and the first derivative w.r.t x_3 as $f_{x_3}(x_1, x_2, x_3)$. Two successive differentiations with respect to x_1 would be labeled $f_{x_1 x_1}(x_1, x_2, x_3)$. Alternatively, we are likely to use alternative notation such as $\frac{\partial y}{\partial x_1}$ or $\frac{\partial^2 y}{\partial x_1^2}$ to represent the first and second derivatives w.r.t. to x_1, and $\frac{\partial^2 y}{\partial x_1 \partial x_2}$ to be the partial second derivative with respect to x_1 and then x_2. For example:

$$f(x_1, x_2, x_3) = 2x_1^2 x_2 x_3 + 4x_1 x_2^2 + 3x_2 + 4x_3^2$$
$$f_{x_1} = 4x_1 x_2 x_3 + 4x_2^2$$
$$f_{x_2} = 2x_1^2 x_3 + 8x_1 x_2 + 3$$
$$f_{x_3} = 2x_1^2 x_2 + 8x_3$$
$$f_{x_1 x_2} = 4x_1 x_3 + 8x_2$$

And so on.

Given the following:

$$a = \begin{bmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}$$

Then:

$$\frac{\partial}{\partial x}a'x = \begin{bmatrix} \frac{\partial}{\partial x_1}a_1x_1 & \frac{\partial}{\partial x_2}a_2x_2 & ... & \frac{\partial}{\partial x_n}a_nx_n \end{bmatrix} = a' \tag{8.9}$$

Given additionally:

$$A = \begin{bmatrix} a_{1,1} & \sigma_{1,2} = a_{2,1} & ... & a_{1,n} = a_{n,1} \\ a_{2,1} = a_{1,2} & a_{2,2} & ... & a_{2,n} = a_{n,2} \\ ... & ... & ... & ... \\ a_{n,1} = a_{1,n} & a_{n,2} = a_{2,n} & ... & a_{n,n} \end{bmatrix}$$

Then:

$$\frac{\partial}{\partial x}Ax = A \tag{8.10}$$

$$\frac{\partial}{\partial x}x'Ax = 2Ax \tag{8.11}$$

This final rule will prove very useful in aspects of simple portfolio optimization.

Maximization & Minimization

Given a function, such as $f(x) = -2x^2 + 5$, we find the critical points of f by differentiating w.r.t. the decision variable (in this case, x) and setting the first derivative equal to zero and solving for the location:

$$f(x) = -2(x-2)^2 + 5$$
$$f'(x) = -4(x-2)$$
$$0 = -4(x-2)$$
$$x = 2$$

And so the critical point of the original function f occurs when $x = 2$, which would be at the point $(2, 5)$.

We decide whether the critical point $(2, 5)$ is a minimum or maximum by investigating the sign of the second derivative - a negative second derivative indicates that this is a maximum, while a positive second derivative indicates that this is a minimum. In the case of the function f, the critical point is a maximum, since $f''(x) = -4$.

There are similar rules for multivariable functions as well.

Constrained Maximization & Minimization

In most applications of maximization and minimization, we would like to be able to constrain the decision variable in some way. Consider the following:

$$max_\omega \quad \omega'\Sigma\omega$$

Which is the objective function for a global minimum variance portfolio. The unconstrained solution for this problem is found when:

$$\frac{\partial}{\partial \omega} \omega' \Sigma \omega = 2 \Sigma \omega$$

$$F.O.C. \quad 0 = 2 \Sigma \omega$$

$$0 = \omega$$

Or, in english, the portfolio variance is minimized when all weights are 0. This is uninteresting, and usually we with to find the minimum variance portfolio weights in the case of *full investment* wherein the portfolio weights sum to 1, which we can represent in matrix language with $\omega' \mathbf{1} = 1$. We can add this constraint via the lagrangian $1 - \omega' \mathbf{1} = 0$. We now have the objective function:

$$max_\omega \quad \omega' \Sigma \omega \quad \text{s.t. } \omega' \mathbf{1} = 1$$

With f.o.c.:

$$\frac{\partial}{\partial \omega} \left[\frac{1}{2} \omega' \Sigma \omega + (1 - \omega' \mathbf{1}) \right] = 0$$

Solution to this will yield the (more interesting) minimum variance portfolio with full investment:

$$\omega^* = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}$$

Bibliography

Ang, A. (2014). *Asset Management: A Systematic Approach to Factor Investing.* Oxford University Press, New York, 1st edition. ISBN 978-0199959327.

Chaves, D., Hsu, J., Li, F., and Shakernia, O. (2012). Efficient algorithms for computing risk parity portfolio weights. *Journal of Investing*, 21(3):150–163.

Chincarini, L. B. and Kim, D. (2006). *Quantitative Equity Portfolio Management: An Active Approach to Portfolio Construction and Management.* McGraw-Hill Education, New York, 1st edition. ISBN 978-0071459396.

Choueifaty, Y. and Coignard, Y. (2008). Toward maximum diversification. *The Journal of Portfolio Management*, 35(1):40–51.

Fabozzi, F. J., Focardi, S. M., and Kolm, P. N. (2010). *Quantitative Equity Investing: Techniques and Strategies.* Wiley, New York, 1st edition. ISBN 978-0470262474.

Grinold, R. C. and Kahn, R. N. (2000). *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk.* McGraw-Hill Education, New York, 2nd edition. ISBN 978-0070248823.

Grinstead, C. M. and Snell, J. L. (2006). *Grinstead and Snell's Introduction to Probability.* American Mathematical Society, 2nd edition.

King, D. (2007). Portfolio optimization and diversification. *The Journal of Asset Management*, 8(5):296–307.

Litterman, B. (2003). *Modern Investment Management: An Equilibrium Approach.* Wiley, New York, 1st edition. ISBN 978-0471124108.

Maillard, S., Roncalli, T., and Teiletche, J. (2010). The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, 36(4):60–70.

Meucci, A. (2005). *Risk and Asset Allocation.* Springer, New York, 1st edition. ISBN 978-3540279044.

Qian, E. E., Hua, R. H., and Sorensen, E. H. (2006). *Quantitative Equity Portfolio Management: Modern Techniques and Applications.* Chapman and Hall/CRC, New York, 1st edition. ISBN 978-1584885580.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41.

Xie, Y. (2015). *Dynamic Documents with R and knitr.* Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.