MARKOWITZ'S PORTFOLIO SELECTION MODEL AND RELATED PROBLEMS

by

ABHIJIT RAVIPATI

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Operations Research

written under the direction of

Prof. Dr. ANDRAS PREKOPA

and approved by

New Brunswick, New Jersey

January 2012

ABSTRACT OF THE THESIS

MARKOWITZ'S PORTFOLIO SELECTION MODEL AND RELATED PROBLEMS

By ABHIJIT RAVIPATI

Thesis Director: Prof Dr. ANDRAS PREKOPA

Markowitz's portfolio selection theory is one of the pillars of theoretical finance. This formulation has an inherent instability once the mean and variance are replaced by their sample counterparts. The problem is amplified when the number of assets is large and the sample covariance is singular or nearly singular. This poses a fundamental problem, because solutions that are not stable under sample fluctuations may look optimal for a given sample, but are, in effect, very far from optimal with respect to the average risk. The paper starts with a general introduction to Markowitz's portfolio theory and then discusses further developments and a few notable works in the area and later moves on to discuss the need for regularization and points out a few existing methods for regularization. After which a formulation of the optimal portfolio selection is presented and ends with a few numerical examples.

Table of Contents

Markowitz's Portfolio Theory ……………………………………………………….1

1.1 Introduction……………………………………………………………………..1

1.2 Assumptions……………………………………………………………………..2

1.3 Optimal Portfolio Selection Model…………………………………………….4

1.4 Later Developments…………………………………………………………….7

1.4.1 Konno – Yamazaki, 1991……………………………………………………..10

1.4.1.1Introduction………………………………………………………………..10

1.4.1.2Model…………………………………………………………………11

1.4.2 Young, 1998…………………………………………………………………….14

1.4.2.1 Introduction………………………………………………………………14

1.4.2.2 Model…………………………………………………………………..14

1.4.3 Black and Litterman, 1992…………………………………………………...17

1.4.3.1 Introduction………………………………………………………………17

1.4.3.2 Market equilibrium returns………………………………………………18

1.4.3.3 Investor views…………………………………………………………….19

1.4.3.4 Black- Litterman equation………………………………………………..21

Regularization……………………………………………………………………..22

2.1 Introduction…………………………………………………………………….22

2.2 Regularization applied to the covariance matrix……………………………24

2.2.1 Ridge Regularization……………………………………………………….24

2.2.2 Spectral cut-off regularization……………………………………………..24

2.2.3 Landweber – Fridman regularization……………………………………25

2.3 Regularization scheme as penalized least-square……………………………25

2.3.1 Bridge method…………………………………………………………...26

2.3.2 Least Absolute Shrinkage and Selection Operator (LASSO)…………………27

Mathematical Formulation……………………………………………………….29

Example…………………………………………………………………………31

Appendix……………………………………………………………………….....33

References……………………………………………………………………….39

Markowitz's Portfolio Theory

1.1 Introduction

A little over forty years ago, a University of Chicago graduate student in economics, while in search of a dissertation topic, ran into a stockbroker who suggested that he study the stock market. Harry Markowitz took that advice and developed a theory that became a foundation of financial economics and revolutionized investment practice. His work earned him a share of 1990 Nobel Prize in Economics. A basic premise of economics is that, due to the scarcity of resources, all economic decisions are made in the face of trade-offs. Markowitz identified the trade-off facing the investor: risk versus expected return. The investment decision is not merely which securities to own, but how to divide the investor's wealth amongst securities. This is the problem of "Portfolio Selection;" hence the title of Markowitz's seminal article published in the March 1952 issue of the Journal of Finance. In that article and subsequent works, Markowitz extends the techniques of linear programming to develop the critical line algorithm.

The critical line algorithm identifies all feasible portfolios that minimize risk (as measured by variance or standard deviation) for a given level of expected return and maximize expected return for a given level of risk. When graphed in standard deviation versus expected return space, these portfolios form the efficient frontier. The efficient frontier represents the trade-off between risk and expected return faced by an investor when forming his portfolio. Most of the efficient frontier represents well diversified portfolios. This is because diversification is a powerful means of achieving risk reduction.

Markowitz developed mean-variance analysis in the context of selecting a portfolio of common stocks. Over the last two decades, mean-variance analysis has been increasingly applied to asset allocation. Asset allocation is the selection of a portfolio of investments where each component is an asset class rather than an individual security. In many respects, asset allocation is a more suitable application of mean-variance analysis than is stock portfolio selection. Mean-variance analysis requires not only knowledge of the expected return and standard deviation on each asset, but also the correlation of returns for each and every pair of assets. Whereas a stock portfolio selection problem might involve hundred of stocks (and hence thousands of correlations), an asset allocation problem typically involves a handful of asset classes (for example stocks, bonds, cash, real estate, and gold). Furthermore, the opportunity to reduce total portfolio risk comes from the lack of correlation across assets. Since stocks generally move together, the benefits of diversification within a stock portfolio are limited. In contrast, the correlation across asset classes is usually low and in some cases negative. Hence, mean-variance is a powerful tool in asset allocation for uncovering large risk reduction opportunities through diversification.

1.2 Assumptions

As with any model, it is important to understand the assumptions of mean-variance analysis in order to use it effectively. First of all, mean-variance analysis is based on a single period model of investment. At the beginning of the period, the investor allocates his wealth among various asset classes, assigning a nonnegative weight to each asset.

During the period, each asset generates a random rate of return so that at the end of the period, his wealth has been changed by the weighted average of the returns. In selecting asset weights, the investor faces a set of linear constraints, one of which is that the weights must sum to one. Based on the game theory work of Von Neumann and Morgenstern, economic theory postulates that individuals make decisions under uncertainty by maximizing the expected value of an increasing concave utility function of consumption. In a one period model, consumption is end of period wealth. In general, maximizing expected utility of ending period wealth by choosing portfolio weights is a complicated stochastic nonlinear programming problem.

To summarize the assumptions:

1. Investors seek to maximize the expected return of total wealth.
2. All investors have the same expected single period investment horizon.
3. All investors are risk-adverse, that is they will only accept greater risk if they are compensated with a higher expected return.
4. Investors base their investment decisions on the expected return and risk.
5. All markets are perfectly efficient (e.g. no taxes and no transaction costs).

The utility function is assumed to be increasing and concave. In terms of the approximating utility function, this translates into expected utility being increasing in expected return (more is better than less) and decreasing in variance (the less risk the better). Hence, of all feasible portfolios, the investor should only consider those that maximize expected return for a given level of variance, or minimize variance for a given level of expected return. These portfolios form the mean-variance efficient set.

1.3 Optimal Portfolio Selection Model

Assuming the portfolio has N assets with returns R_i, i= 1.. N.

Let,

R_p = Return on the portfolio

R_i = Return on asset i

w_i = Weight of component asset i (that is, the share of asset *i* in the portfolio).

σ_i = Standard deviation of asset i

Portfolio return:

$$R_p = \sum_i w_i E(R_i)$$

Portfolio return variance:

$$\sigma_p^2 = \sum_i \sum_j w_i w_j \sigma_i \sigma_j \rho_{ij}$$

where ρ_{ij} is the correlation coefficient between the returns on assets *i* and *j*.

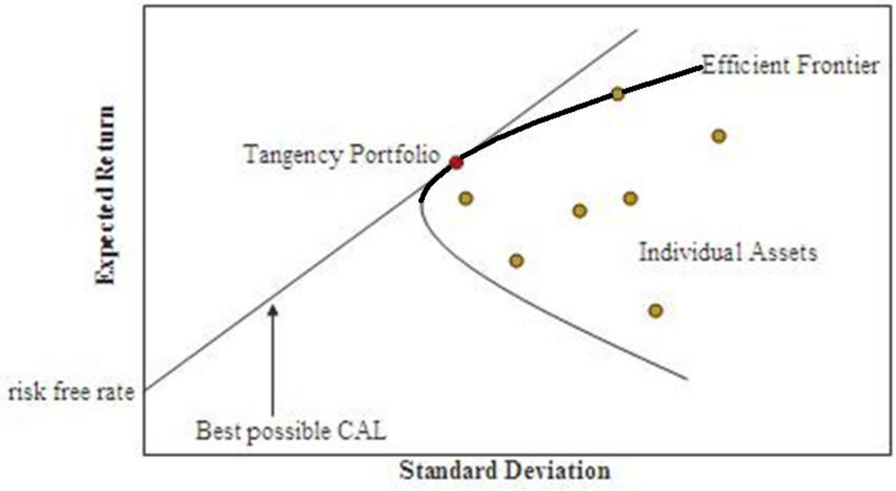
Efficient Frontier
Tangency Portfolio
Expected Return
Individual Assets
risk free rate
Best possible CAL
Standard Deviation

Fig 1.

Markowitz showed that the risk of the portfolio of assets depends on the asset weights and the standard deviations of the asset's returns, and crucially on the correlation (covariance) of the asset returns.

The efficient frontier describes the relationship between the return that can be expected from a portfolio and the riskiness (volatility) of the portfolio. It can be drawn as a curve on a graph of risk against expected return of a portfolio. The efficient frontier gives the best return that can be expected for a given level of risk or the lowest level of risk needed to achieve a given expected rate of return. The efficient frontier is extremely important to the theory of portfolio construction and valuation. The concept of an efficient frontier can be used to illustrate the benefits of diversification. An undiversified portfolio can be moved closer to the efficient frontier by diversifying it. Diversification can, therefore, increase returns without increasing risk, or reduce risk without reducing expected returns.

If the investor has access to risk free investment, the risk of the portfolio can further be reduced, which is shown in Fig.1 by the CAL. The risk-free asset is the (hypothetical) asset which pays a risk-free rate. In practice, short-term government securities (such as US treasury bills) are used as a risk-free asset, because they pay a fixed rate of interest and have exceptionally low default risk. The risk-free asset has zero variance in returns (hence is risk-free); it is also uncorrelated with any other asset (by definition, since its variance is zero).

Within the application of portfolio theory the following two quantities will need to use the corresponding units of measurement throughout the computation:

1. Historical Values: This is the source data which is given in absolute or relative terms.
2. Expected Returns: The expected return of the investment over the period considered which should be given and will be returned in the units used (i.e. absolute or relative) by the historical values.

The values of the expected return which are either evaluated or given will be or will need to be in accordance with the units used within the historical values.

" To use the E-V rule in the selection securities we must have procedures for finding reasonable μ_i and ρ_{ij}. These procedures I believe, should combine statistical techniques and the judgment of practical knowledge………… One suggestion as to tentative μ_i, ρ_{ij} is to use observed for some period of the past. I believe, better methods, which take into account more information can be found. I believe that what is needed is essentially a “probabilistic” reformulation of security analysis.”

- Markowitz, H.M. (March 1952). "Portfolio Selection". *The Journal of Finance*

1.4 Later Developments

Markowitz’s selection model is fundamental to the foundation of the current theory of asset allocation. Since Markowitz proposed his model, numerous portfolio selection models have been developed to advance the former and portfolio theory has been improved and completed in several directions. Some models have been developed to minimize semivariance in different cases such as Huang [15] and Markowitz [16], while other researchers like Konno and Suzuki [17], Liu, Wang and Qiu [18] and Pornchai, Krushnan, Shatid and Arun [19] added the skewness in consideration for portfolio selection.

The common assumptions are that the investor has enough historical data and that the situation of asset markets in future can be correctly predicted by the historical data. Since sometimes this is not, practical problems arise. For example, when new stocks are listed in the stock market, there is no historical information for these securities. Random, fuzzy and random fuzzy optimization models proved some useful methods for investors to tackle the uncertainty. A number of researchers have shown that mean-variance efficient portfolios, based on estimates, are highly sensitive to perturbations of these estimates. Jobson, Korkie and Ratti (20) and Jobson and Korkie (21) detail these problems and suggest the use of shrinkage estimators.

Some authors like Carlsson, Fuller and Majlender [22], Leon, Liern and Vercher [23] and Vercher, Bermudez and Segura [24] use fuzzy numbers to replace uncertain returns of the securities. Tanaka and Guo [25] and Tanaka, Guo and TÄurksen [26] used possibilistic distributions to model uncertainty in returns. Arenas-Parra, Bilbao-Terol and Rodrguez-Ura [27] introduced vague goals for return rate, risk and liquidity based on expected intervals. A measure of downside risk is incorporated by Feiring, Wong, Poon, and Chan (28), and Konno, Shirakawa, and Yamazaki (29) who use an approximation to the lower semi-third moment in their Mean-Absolute Deviation-Skewness portfolio model. Konno and Yamazaki proposed the mean absolute deviation (MAD) model as an alternative to the mean variance (MV) model claiming that it retains all the positive features of the mean variance model , not only saves computing time but also does not require the covariance matrix.

Addressing the issue of estimation risk, Frost and Savarino (30) show that constraining portfolio weights, by restricting the action space during the optimization, reduces estimation error**.** Jorion (31) proposes a resampling method aimed at estimation error. In an attempt to maintain the decision simplicity associated with the efficient frontier and still accommodate parameter uncertainty, Michaud (32) proposes a sampling based method for estimating a resampled efficient frontier. Polson and Tew (33) argue for the use of posterior predictive moments instead of point estimates for mean and variance of an assumed sampling model.

Using a Bayesian approach**,** Britten-Jones (34) proposes placing informative prior densities directly on the portfolio weights. Chopra and Ziemba (35) showed that errors in means are about ten times as important as errors in variances, and errors in variances are about twice as important as errors in covariances. Best and Grauer (36) showed that optimal portfolios are very sensitive to the level of expected returns. Jorion (37) use a shrinkage approach while Treynor and Black (38) advocate the use of investors' views in combination with historical data. Kandel and Stambaugh (39) examine predictability of stock returns when allocating between stocks and cash by a risk-averse Bayesian investor. Zellner and Chetty (40), Klein and Bawa (41) and Brown (42) emphasize using a predictive probability model. P´astor and Stambaugh (43) study the implications of different pricing models on optimal portfolios, updating prior beliefs based on sample evidence. P´astor (44) and Black and Litterman (45) propose using asset pricing models to provide informative prior distributions for future returns.

A number of researchers are targeting their efforts in modeling time variations in the conditional dependence of asset returns in terms of conditional covariances and correlations (Bollerslev et al. (46) or Engle (47) to cite a few). Based on data from the last 150 years**,** Goetzmann, Li and Rouwenhorst (48) found that correlations between equity returns vary substantially over time and achieve their highest levels during periods characterized by highly integrated financial markets. Longin and Solnik (49) studied shifts in global equity markets correlation structure and rejected the hypothesis of constant correlations among international stock markets. Further**,** Ang and Chen (50) confirmed this for the US market for correlations between stock returns and an aggregate market index. Others connect the variability of stock return correlations to the phase of the business cycle**.** Ledoit, Santa-Clara and Wolf (51) and Erb, Harvey and Viskanta (52) show that correlations are time-varying and depend on the state of the economy, tending to be higher during periods of recession. Similar evidence is brought forward by Moskowitz (53) who links time variation of volatilities and covariances to NBER recessions.

1.4.1 Konno – Yamazaki, 1991

1.4.1.1 Introduction

Konno and Yamazaki (1991) proposed a new model using mean absolute deviation (MAD) as risk measure to overcome the weaknesses of the mean-variance model proposed by Markowitz. One of the most significant reasons problems being the computational difficulty associated with solving a large scale quadratic problem associated with a dense covariance matrix. They stated that equilibrium models have to impose several unrealistic assumptions to derive a relation between rate of return and on assets and market portfolio but, data from Tokyo stock exchange showed that this relation is very unstable and that the information provided by CAPM can best serve as a first order approximation.

Konno and Yamazaki employed L1 –mean absolute deviation as a risk measure instead of variance, so they could overcome most of the problems of Markowitz's model while maintaining its advantages over equilibrium models. Some of the problems that are rarely practically solved are as follows:

Computational burden: Solving large scale dense quadratic problems can prove difficult.

Investor’s perception: A large number of investors were not fully convinced of the validity of the standard deviation as a measure of risk.

Transaction costs and cut-off effect: This means that the investor who invests in many different stocks in small costs will be inconvenienced as he will have to bear the burden of transaction costs. Also, since the investor cannot buy stocks in fractions and he/she will have to round it off to integers.

1.4.1.2 Model

They introduced the **L1** risk function

$$w(x) = E[\left|\sum_{j=1}^{n} R_j x_j - E\left[\sum_{j=1}^{n} R_j x_j\right]\right|]$$

Where,

R_j = Random variable representing the rate of return on asset S_j

x_j = Amount invested in S_j

M_0 = Total fund amount

E [.] = Expected value of random variable in bracket

They then go on to state and prove the following theorem:

If (R_1,……..R_n) are multivariate normally distributed , then

$$w(x) = \sqrt{\frac{2}{\pi}}\,\sigma(x)$$

Where $\sigma(x)$ = Standard deviation

They proved that these two measures (w(x) and R_i) are the same if ($R_1 \ldots R_n$) are multivariate normally distributed.

So the Model becomes the following;

$$\text{Min} \quad w(x)\ E\left[\left|\sum_{j=1}^{n} R_j x_j - E\left[\sum_{j=1}^{n} R_j x_j\right]\right|\right]$$

$$\text{ST} \quad \left|\ \begin{array}{l} \sum_{j=1}^{n} E[R_j]x_j \geq \rho M_0\ , \\ \sum_{j=1}^{n} x_j = M_0\ , \\ 0 \leq x_j \leq u_j\ , \quad j = 1,\ldots\ldots,n\ . \end{array} \right. \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 12.1$$

Konno and Yamazaki assumed that the expected value of the random variable can be approximated by the average from the data.

Therefore,

$$r_j = E[R_j] = \sum_{t=1}^{T} r_{jt}/T$$

Now,

$$E\left[\left|\sum_{j=1}^{n} R_j x_j - E\left[\sum_{j=1}^{n} R_j x_j\right]\right|\right] = \frac{1}{T}\sum_{t=1}^{T}\left|\sum_{j=1}^{n}(r_{jt} - r_j)x_j\right|$$

Let

$$a_{jt} = r_{jt} - r_j\ , \quad j = 1,\ldots\ldots,n; \quad t = 1,\ldots..,T\ .$$

Model in 12.1 can be stated as,

$$\text{Min} \quad \sum_{t=1}^{T}\left|\sum_{j=1}^{n} a_{jt}x_j\right| / T$$

$$\text{ST} \quad \begin{cases} \sum_{j=1}^{n} r_j x_j \geq \rho M_0, \\ \sum_{j=1}^{n} x_j = M_0, \\ 0 \leq x_j \leq u_j, \quad j = 1,\ldots\ldots,n \end{cases}$$

Which is equivalent to the following linear program:

$$\text{Min} \quad \sum_{t=1}^{T} y_t / T$$

$$\text{ST} \quad \begin{cases} y_t + \sum_{j=1}^{n} a_{jt}x_j \geq 0, \quad t = 1,\ldots..,T, \\ y_t - \sum_{j=1}^{n} a_{jt}x_j \geq 0, \quad t = 1,\ldots..,T, \\ \sum_{j=1}^{n} r_j x_j \geq \rho M_0, \\ \sum_{j=1}^{n} x_j = M_0, \\ 0 \leq x_j \leq u_j, \quad j = 1,\ldots\ldots,n \end{cases}$$

Konno-Yamazaki state the following advantages over Markowitz’s model:

1. No need to calculate the covariance matrix.
2. Solving their linear program is much easier compared to solving a quadratic program.
3. The optimal solution size is smaller
4. T can be uses as a control variable to restrict the number of assets in the portfolio

1.4.2 Young, 1998

1.4.2.1 Introduction

Young (1998) proposed a principle for choosing portfolios based on historical returns. He was the first to apply the minimax model to the portfolio selection problem. If each of the two players behaves rationally, then game theory asserts that a solution for every situation can be determined by assuming that the players seek to minimize their maximum expected losses – Minimax criterion.

Young used minimum return rather than variance as a measure of risk. He defined the optimal portfolio as that one that would minimize the maximum loss over all past historical periods, subject to a restriction on the minimum acceptable average return across all observed time periods. He stated that if an investor's utility function is more risk averse than is implied by mean-variance analysis, or if returns data are skewed, or if the portfolio optimization problem involves a large number of decision variables, his model would be advantageous to use.

1.4.2.2 Model

$$\text{MAX}_{M_p,w} \quad M_p$$

$$\text{ST} \qquad \left| \begin{array}{l} \sum_{j=1}^{N} w_j y_{jt} - M_p \geq 0, \quad \text{t= 1,…..,T,} \\[2em] \sum_{j=1}^{N} w_j \bar{y}_j \geq G, \\[2em] \sum_{j=1}^{N} w_j \leq W, \end{array} \right.$$

$$w_j \geq 0, \quad j=1,\ldots,N.$$

Where,

y_{jt} = Return on one dollar invested in security j in time period t

w_j = Portfolio allocation to security j

$\overline{y_j}$ = Average return on security j

$$= {}^{1}/_{T} \sum_{t=1}^{T} y_{jt}$$

y_{pt} = Return on portfolio in time period t

$$= \sum_{j=1}^{N} w_j y_{jt}$$

E_p = Average return on portfolio

$$= \sum_{j=1}^{N} w_j \overline{y_j}$$

M_p = Minimum return on portfolio

$$= \min_t y_{pt}$$

The optimum portfolio maximizes M_p under imposed restrictions,

1. E_p (average return) exceeds a minimum level G

2. Net asset allocations does not exceed total budget allocation W

Thus, M_p represents the portfolio's minimum return at the end of each time period and since M_p is being maximized, the portfolio will take on the maximum value of the minimum returns. According to Young, this model presents logical advantages over other portfolio optimization models if asset prices are not normally distributed and similar results when they are.

He states an equivalent model that seeks to maximize expected return, subject to a restriction that the portfolio return exceeds some threshold *H* in each observation period:

$$\max_w \quad E = \sum_{j=1}^{N} w_j \bar{y}_j$$

$$\text{ST} \quad \left| \begin{array}{l} \sum_{j=1}^{N} w_j y_{jt} \geq H, \quad t = 1,\ldots..,T, \\ \sum_{j=1}^{N} w_j \leq W, \end{array} \right.$$

$$w_j \geq 0, \quad j=1,\ldots.,N.$$

This model has considerable advantage as it is a linear program and it allows the model to treat additional complexities such as:

1. Transaction costs
2. Logical side constraints like
    - Inclusion/exclusion of both assets a and b
    - Holding more than $d worth of asset a

Thus the minimax model is capable of incorporating a large number of modeling complexities and variations.

1.4.3 Black and Litterman, 1992

1.4.3.1 Introduction

The Black-Litterman asset allocation model, created by Fischer Black and Robert Litterman, is a portfolio construction method that overcomes the problem of highly concentrated portfolios, input-sensitivity, and estimation error maximization. Their model uses a Bayesian approach to combine the subjective views of an investor regarding the expected returns of one or more assets with the market equilibrium vector of expected returns (the prior distribution) to form a new, mixed estimate of expected returns.

The Black-Litterman asset allocation model was introduced in Black and Litterman (1990) and expanded in Black and Litterman (1991, 1992). The Black Litterman model combines the CAPM (Sharpe (1964)), reverse optimization (Sharpe (1974)), mixed estimation (Theil (1971, 1978)), the universal hedge ratio / Black's global CAPM (see Black (1989a, 1989b)), and mean-variance optimization (Markowitz (1952)). The approach works by combing historical information with additional data and forms the updated distribution of expected returns. If the investor has no subjective opinions the weights are based on market equilibrium data, but if he does hold any subjective views then the weights on individual assets are shifted from the market equilibrium weights.

The key inputs to the Black and Litterman model are market equilibrium returns and investor views. This framework incorporates the investor views that helps investors control the magnitude of the tilts caused by views.

1.4.3.2 Market equilibrium returns

The Black and Litterman model uses the market equilibrium weights or capital asset pricing model (CAPM) as the basis. CAPM is developed by forming the efficient frontier of the market portfolios and tracing the capital market line (CML). The CML is tangent to the efficient frontier at the market portfolio therefore, there is no other combination of risky and riskless assets that can provide better returns for a given level of risk.

CAPM:

$$E\,(r_i) = r_j + \frac{\sigma_i}{\sigma_m}\,(r_m - r_j)$$

$$= r_j + \beta_i\,(r_m - r_j)$$

Where,

$E\,(r_i)$ = Expected return on asset i

r_f = Risk free asset return

r_m = Return on market portfolio

σ_i = Standard deviation of returns on asset i

σ_m = Standard deviation of returns on market portfolio

$\beta_i = \frac{\sigma_i}{\sigma_m}$

The model uses CAPM in reverse. It assumes market portfolio is held by mean-variance investors and it uses optimization to back out the optimal expected returns. They define market equilibrium returns as:

$$\pi = \lambda \Sigma \omega \qquad \text{……..19.1}$$

Where,

N = Number of assets

π = Vector of implied excess returns (N,1)

Σ = Covariance matrix of returns (N,N)

ω = Vector of market capitalization weights of the assets (N,1)

λ = Risk aversion coefficient

$$= \frac{(r_m - r_j)}{{\sigma_m}^2}$$

1.4.3.3 Investor views

The views of the investors are incorporated into the model in the following form:

$$Q + \varepsilon = \begin{bmatrix} Q_1 \\ \vdots \\ \\ Q_K \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \\ \varepsilon_K \end{bmatrix}$$

Where,

K= Number of investor views

Q = Vector of investor views

ε = Error term

If $\varepsilon = 0$ that means the investor has 100% confidence in his views. ω denotes the variance of each error term. Assuming that each error term is independent of each other the covariance matrix Ω is a diagonal matrix with the following form:

$$\Omega = \begin{bmatrix} \omega_1 & 0 & \cdots & & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & & & \\ & & & & 0 \\ 0 & \cdots & & 0 & \omega_K \end{bmatrix}$$

Using the above formulation the model incorporates both absolute as well as relative views of the investors.

1.4.3.4 Black- Litterman equation

The Black-Litterman equation combines the equilibrium returns and investor views into one equation to determine expected returns which is used to determine the optimal portfolio weights.

$$E[R] = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q] \qquad \ldots\ldots\ldots\ldots 21.1$$

Where,

E [R] = Vector of combined results (N,1)

τ = Scalar indicating uncertainty of CAPM prior

Σ = Covariance matrix of equilibrium excess returns (N,N)

P = Matrix of investor views (N,1)

Ω = Diagonal covariance matrix of view error terms (K,N)

Π = Vector of equilibrium excess returns (N,1)

Q = Vector of investor views (K,1)

In case the investor is unconstrained we use 19.1 to determine optimal portfolio weights by,

$$w^* = (\lambda\Sigma)^{-1}\mu$$

……..………21.2

Where μ is the vector derived from 21.1.

21.2 is the solution to the unconstrained optimization problem

$$\max_{w} \; w'\mu - \lambda w'\Sigma w/2$$

Overall, the Black-Litterman model overcomes the most-often cited weaknesses of mean-variance optimization such as highly concentrated portfolios, input-sensitivity, and estimation error-maximization, helping users to realize the benefits of the Markowitz paradigm.

Regularization

2.1 Introduction

In the original formulation of the Markowitz portfolio theory, the underlying process was assumed to be multivariate normal. Accordingly, reward was measured in terms of the expected return, risk in terms of the variance of the portfolio. Unfortunately, the nature of portfolio selection is not compatible with this limit.

Institutional portfolios are large, with N's in the range of hundreds or thousands, while considerations of transaction costs and non-stationarity limit the number of available data points to a couple of hundreds at most. Therefore, portfolio selection works in a region, where size of the portfolio and sample size (i.e. size of the time series available for each asset) are, at best, of the same order of magnitude. This, however, is not the realm of classical statistical methods. It is evident that portfolio theory struggles with the same fundamental difficulty that is underlying basically every complex modeling and optimization task: the high number of dimensions and the insufficient amount of information available about the system.

In order to solve the mean variance problem, the expected return and the covariance matrix of the vector of security return, which are unknown, need to be estimated from available data set. In particular, an estimate of the inverse of the covariance matrix is needed. The sample covariance often used in practice may be the worst choice because it is typically nearly singular, and sometimes not even invertible.

Markowitz mean-variance rule can be viewed as a trade-off between the expected return and the variance of the returns. The mean variance problem consists of choosing the vector x , to minimize the variance of the resulting portfolio $r_{p,t+1}=x' r_{t+1}$, where r_{t+1} is the excess return $R_{t+1} - R_t^f$, for a pre- determined target expected return of the portfolio μ_p :

$$\text{Min}_x \quad \text{Var}[r_p] = w' C w$$

$$\text{S.T.} \quad E[r_p] = w' \mu = \mu_p$$

The optimal portfolio is given by,

$$w^* = \frac{\mu p}{\mu' C \mu} C^{-1} \mu$$

Where μ = Conditional means,

C = Covariance matrix.

The issue of ill-conditioned covariance matrices is important because inverting such a matrix increases dramatically the estimation error and then makes the mean variance solution unreliable.

Many regularization techniques can stabilize the inverse. They can be divided into two classes: Regularization directly applied to the covariance matrix and regularization expressed as a penalized least–squares.

2.2 Regularization applied to the covariance matrix

We will consider here the three most popular regularization techniques: ridge, spectral cut-off, and Landweber Fridman. Each method will give a different estimate of β.

2.2.1 Ridge Regularization

It consists in adding a diagonal matrix to Ω.

$$\beta_\tau = (R'R + \tau I)^{-1} R' 1_T ,$$

$$\beta_\tau = \sum \frac{\lambda j}{\lambda j 2 + \tau} (1_T' v_j) \Phi_j .$$

2.2.2 Spectral cut-off regularization

This method discards the eigen vectors associated with the smallest eigen values.

$$\beta_\tau = \sum_{\lambda j > \tau} \frac{1}{\lambda j} (1_T' v_j) \Phi_j .$$

Interestingly, ν_j are the principal components of **Ω**, so that if r_t follows a factor model, ν_1, ν_2,....... estimate the factors.

2.2.3 Landweber – Fridman regularization

The solution to (4) can be computed iteratively as

$$\Psi_k = (I - cR'R)\ \psi_{k-1} + cR'1_\tau$$

With $0 < c < 1/ R^2$. Alternatively, we can write

$$\beta_\tau = \sum \frac{1}{\lambda j} \{1 - (1 - c\lambda_j^{\ 2})^{1/\tau}\} (1_\tau' \ v_j)\ \Phi_j$$

Here, the regularization parameter τ is such that $1/\ \tau$ represents the number of iterations. The three methods involve a regularization parameter τ which needs to converge to zero with T at a certain rate for the solution to converge.

2.3 Regularization scheme as penalized least-square

The traditional optimal Markowitz portfolio x* is the solution to (1) that can be reformulated by exploiting the relation $C = E(r_t r_t') - \mu\mu'$ as

$$x^* = \text{argmin}_x\ E\,[(\mu_p - x'r_t)^2]$$

$$\text{S.T. } x'\mu = \mu_p$$

If one replaces the expectation by sample average μ, the problem becomes :

$$x^* = \text{argmin}_x \frac{1}{T} \left\| \mu p - x' rt \right\|_2^2$$

$$\text{S.T. } x'\mu = \mu_p$$

As mentioned before, the solution of this problem may be very unreliable if RR' is nearly singular. To avoid having explosive solutions, we can penalize the large values by introducing a penalty term applied to a norm of x. Depending on the norm we choose, we end up with different regularization techniques. X_i

2.3.1 Bridge method

For Y > 0 the Bridge estimate is given by

$$x_\tau^* = \text{argmin}_x \frac{1}{T} \left\| \mu p - x' rt \right\|_2^2 + \tau \sum_{i=1}^{p} \left| x_i \right|^Y$$

Where τ is the penalty term.

The Bridge method includes two special cases. For Y = 1 we have Lasso regularization, while Y = 2 leads to the Ridge method. The term $\sum_{i=1}^{p}$ xi Y can be interpreted as a transaction cost. It is linear for Lasso, but quadratic for the ridge.

2.3.2 Least Absolute Shrinkage and Selection Operator (LASSO)

The Lasso regularized solution is obtained by solving :

$$x_\tau^* = \mathrm{argmin}_x \ \|\mu p 1t - x' rt\|_2^2 + \tau \ \|x\|_1$$

$$\text{S.T.} \quad x' \mu = \mu_p$$

For two differently penalty constants τ_1 and τ_2 the optimal regularized portfolio satisfies : $(\tau_1 - \tau_2) \left(\|x^{[\tau_2]}\|_1 - \|x^{[\tau_1]}\|_1 \right) >= 0$ then the higher the l_1 – penalty constant (τ), the sparser the optimal weights. So that a portfolio with non-negative entries corresponds to the largest values of τ and thus to the sparsest solution. In particular the same solution can be obtained for all τ greater than some value τ_0.

Brodie et al. consider models without a risk free asset. Using the fact that all the wealth as invested ($x' 1_N = 1$), they use the equivalent formulation for the objective function as :

$$\|\mu_p 1_T - Rx\|_2^2 + 2\tau \sum_{i \ with \ x(i)<0} |x_i| + \tau$$

which is equivalently to a penalty on the short positions. The Lasso regression then regulates the amount of shortening in the portfolio designed by the optimization process, so that the problem stabilizes.

The general from of the l_1- penalized regression with linear constraints is:

$$x_\tau^* = \mathrm{argmin}_{x \epsilon H} \ \|b - Ax\|_2^2 + \tau \ \|x\|_1$$

H is an affine subspace defined by linear constraints. The regularized optimal can be found using an adaptaion of the homotopy / LARS algorithm as described in Brodie et al. .

Mathematical Formulation

The basic formulation which can solve the optimal portfolio selection problem is

$$\text{Min} \quad \{ c^T x + x^T C x \} \qquad \text{-------} \ (1)$$

$$\text{S.T.} \quad \left| \begin{array}{l} Ax = b \\ x >= 0 \end{array} \right.$$

Consider the Karush – Kuhn – Tucker conditions.

KKT conditions provide the necessary conditions for optimality.

For,

$$\text{Min} \quad f(x)$$

$$\text{S.T.} \quad \left| \begin{array}{l} g_i(x) >= 0 \ , \ i=1 \ldots . \ m \\ x >= 0 \end{array} \right.$$

The necessary conditions for optimality of x^* :

$$-\nabla f(x^*) + \sum_{i=1}^{m} (\lambda^*_i \nabla g_i(x^*)) <= 0$$

$$[-\nabla f(x^*) + \sum_{i=1}^{m} (\lambda^*_i \nabla g_i(x^*))] \; x^* = 0$$

$$\sum_{i=1}^{m} \lambda^*_i \nabla g_i(x^*) = 0$$

$$\lambda^* >= 0$$

The above stated conditions are sufficient too, if f(x) is convex and $g_1,\ldots,g_m$ are concave.

Using the KKT conditions, we get the following conditions for x to be optimal in the set of equations in (1).

$$c - 2Cx + A^T\lambda + v = 0$$

$$v^Tx = 0$$

$$v >= 0$$

For the sake of conversion into a standard form λ can be expressed as a sum of two non –negative integers:

$\lambda = \lambda^+ - \lambda^-$

$\lambda^+, \lambda^- >= 0$

Substituting the above equations in the KKT necessary conditions,

$c - 2Cx + A^T\lambda^+ - A^T\lambda^- + v = 0$

Now, the problem is to find λ, x, V which satisfy the following formulation,

$$\text{Min} \quad \sum_{i=1}^{n} u_i$$

$$\text{S.T.} \quad \left| \begin{array}{l} Ax = b \\ -2Cx + A^T\lambda^+ - A^T\lambda^- + v + Fu = -c \\ x, \lambda^+, \lambda^-, v, u >= 0 \end{array} \right.$$

Where F is a diagonal (n x n) matrix $F(i,i) = \left| \begin{array}{l} 1 \text{ , if } c(i) >= 0 \\ -1 \text{ , if } c(i) < 0 \end{array} \right.$

Example

The above stated formulation has been tested using the following data:

A= [-1 1 1 -1 1]

b= [1]

C=

$$\begin{bmatrix} 0.00421276 & 0.00004712 & -0.00080459 & -0.00022827 & -0.00104229 \\ 0.00004712 & 0.00350925 & 0.0011045 & 0.00172536 & -0.00085248 \\ -0.00080459 & 0.0011045 & 0.00254941 & -0.00001498 & 0.00105894 \\ -0.00022827 & 0.00172536 & -0.00001498 & 0.00173681 & -0.00080253 \\ -0.00104229 & -0.00085248 & 0.00105894 & -0.00080253 & 0.00110889 \end{bmatrix}$$

$$c = \begin{bmatrix} 0.77 \\ 0.82 \\ 0.8 \\ 0.62 \\ 0.6 \end{bmatrix}$$

Running the code in appendix 1 the following solution is arrived upon as optimal:

$$X = \begin{bmatrix} 77.072 \\ 115.47 \\ 0 \\ 37.399 \\ 0 \end{bmatrix}$$

$$\lambda^{+} = [0]$$

$$\lambda^{-} = [0]$$

$$V = \begin{bmatrix} 0 \\ 0 \\ -1.306 \\ 0 \\ -1.1444 \end{bmatrix}$$

$$u=\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}$$

Appendix

Matlab Code

```
[n1, n1] = size(C1);

I=eye(n1);

Z1=zeros(1,2*n1+2);

for i=1:n1

for j=1:n1

if ((i==j) & ( c1(i) >= 0))

F(i,j)=1;

elseif ((i==j) & ( c1(i)< 0))

F(i,j)=-1;

else F(i,j)=0;

end

end
```

```
end

A = [-2*C1, a' , -a' , I , F; a , Z1];

b = [-c1; b1];

c=-[zeros(2*n1+2,1);ones(n1,1)];

B=[1 2 3 15 16 17];

eps= 1e-3;

% Solves: Maximize c^Tx subject to Ax = b, x >= 0

% We will assume that the LP is nondegenerate

% We are given an initial feasible basis B

% [obj,x,y] = revised_simplex(c,A,b,eps,B)

% eps is a suitable optimality tolerance, say 1e-3

% Output parameters:- obj is the optimal objective value

% x is the primal optimal solution

% y is the dual optimal solution

[m,n] = size(A);

%%%%%%%%%%%%%%%%%%%%

% Step 1:- We are given an initial basis B
```

```
N = setdiff([1:n],B);
% B = find(x0);
% N = find(ones(n,1) - abs(sign(x0)));
xB = A(:,B)\b;
% xB = x0(B);
iter = 0;
while 1==1,
iter = iter + 1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Step 2 :- Solve A_B^Ty = c_B and compute s_N = c_N - A_N^Ty
% Declare optimality if s_N <= 0
% Else find the entering non-basic variable x_{N(k)}
y = A(:,B)'\c(B);
sN = c(N) - A(:,N)'*y;
[sNmax,k] = max(sN);
if sNmax <= eps,
fprintf('We are done\n');
fprintf('Number of iterations is %d\n',iter);
x = zeros(n,1);
x(B) = xB;
fprintf('Optimal objective value is %f\n',c'*x);
obj = c'*x;
return;
end;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Step 3 :- Solve A_Bd = a_{N(k)}
% Find theta = Min_{i=1,...,m|d_i > 0} xB(i)/d(i)
% Let theta = xB(l)/d(l)
% x_{B(l)} is the leaving basic variable
% Also check for unboundedness if d <= 0
d = A(:,B)\A(:,N(k));
zz = find(d > eps)';
if (isempty(zz))
error('System is unbounded\n');
```

```
end

[qw,er]= ismember(N(k),1:n1);

[qw1,as]= ismember(N(k) + n1+2,B);

[qw2,ty]= ismember(N(k),n1+3:2*n1+2);

[qw1,df]= ismember(N(k) - n1-2,B);

if er~=0 & as~=0

l= find(B== (N(k) +n1+2));

theta = xB(l)/d(l);

elseif ty~= 0 & df~=0;

l= find(B== (N(k) -n1-2));

theta = xB(l)/d(l);

else [theta,ii] = min(xB(zz)./d(zz));

l= zz(ii(1));

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Step 4:- Update B and N

% Also x(B(i)) = x(B(i)) - theta*d(i), i=1,...,m and i not equal to l
% x(B(l)) = theta
temp = B(l);
B(l) = N(k);
N(k) = temp;
xB = A(:,B)\b;
end; % while
```

References

[1] H. M. Markowitz, “Portfolio selection”, The Journal of Finance, Vol. 7, No. 1 , pp.77-91, 1952

[2] H. M. Markowitz, “Portfolio Selection: Efficient Diversification of Investments”, 1959

[3] N. Chapados, and Y. Bengio, “Cost functions and model combination for VaR-based asset allocation using neural networks.” IEEE Transactions on Neural Networks, 12, 890-906, 2001

[4] M. Carrasco, “A regularization approach to the many instrument problems.” , mimeo, 2009.

[5] R. Kan and G. Zhou, "Optimal portfolio choice with parameter uncertainty", Journal of Financial and Quantitative Analysis, 42, 621-656, 2007

[6] O. Ledoit, M. Wolf, “A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices.” , Journal of Multivariate Analysis, 88, 365-411, 2004

[7] Hal Varian, “A Portfolio of Nobel Laureates: Markowitz, Miller and Sharpe.”, The Journal of economic Perspectives, Vol 7, No. 1, pp 159-169, 1996

[8] Marine Carrasco and Nèrèe Noumon, “Optimal Portfolio Selection using Regularization.” Universitè de Montréal, working paper , Feb 2010

[9] O. Ledoit and M. Wolf, “Improved estimation of the covariance matrix of stock returns with an application to portfolio selection.”, 2003

[10] E. J. Elton and M. J. Gruber, “Modern Portfolio Theory and Investment Analysis.” Wiley, New York, 1995

[11] R. Jagannathan and T. Ma, “Risk reduction in large portfolios: Why imposing the wrong constraints helps," Journal of Finance, vol. 58, pp. 1651-1684, 2003

[12] O. Ledoit and M. Wolf, “Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection.," Journal of Empirical Finance, vol. 10, no. 5, pp. 603-621, 2003

[13] O. Ledoit and M. Wolf, “Honey, I Shrunk the Sample Covariance Matrix," J.Portfolio Management, vol. 31, p. 110, 2004

[14] V. DeMiguel, L. Garlappi, F. J. Nogales, and R. Uppal, “A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms.," Management Science, vol. 55, pp. 798-812, 2009

[15] X. Huang, Mean-semivariance models for fuzzy protfolio selection, Journal of Computational and Applied Mathematics 217 (2008) 1-8.

[16] H. Markowitz, Computation of mean-semivariance effcient sets by the critical line algorithm, Annals of Operational Research 45 (1993) 307-317.

[17] H. Konno, K. Suzuki, A mean-variance-skewness optimization model, Journal of Operations Research Society of Japan 38 (1995) 137-18

[18] S.C. Liu, S.Y. Wang, W.H. Qiu, A mean-variance-skewness model for portfolio selection with transaction costs, International Journal of Systems Science 34 (2003) 255-262.

[19] C. Pornchai, D. Krushnan, H. Shatid, J.P. Arun, Portfolio selection and skewness: Evidence from International Stock Markets, Journal of Banking and Finance, 21 (1997) 143-16

[20] J.D. Jobson, B. Korkie, V. Ratti, Improved Estimation for Markowitz Portfolios using James-Stein Type Estimators, Proceedings of the American Statistical Association, Business and Economics Statistics Section 41 (1979), 279-284

[21] J.D. Jobson, B. Korkie, Estimation for Markowitz efficient portfolios, Journal of the American Statistical Association, 75:544-544

[22] C. Carlsson, R. Fuller, P. Majlender, A possibilistic approach to selecting portfolios with highest utility score, Fuzzy Sets and Systems 131 (2002) 13-21.

[23] T. Leon, V. Liern, E. Vercher, Viability of infeasible portfolio selection problems: a fuzzy approach, European Jounal of Operational Research 139 (2002) 178-189

[24] E. Vercher, J.D. Bermudez, J.V. Segura, Fuzzy portfolio optimization under downside risk measures, Fuzzy Sets and Systems 158 (2007) 769-782

[25] H. Tanaka, P. Guo, Portfolio selection based on upper and lower exponential possibility distributions, European Journal of Operational Research 114 (1999) 115-126.

[26] H. Tanaka, P. Guo, I.B. TÄurksen, Portfolio selection based on fuzzy probabilities and possibility distributions, Fuzzy Sets and Systems 111 (2000) 387-397

[27] M. Arenas-Parra, A. Bilbao-Terol, M.V. Rodrguez-Ura, A fuzzy goal programming approach to portfolio selection, European Journal of Operational Research 133 (2001) 287-29

[28] B.R. Feiring, W.L. Wong, M. Poon and Y.C. Chan, Portfolio selection in downsiderisk optimization approach - application to the Hong Kong stock market. International Journal of Systems Science 25 (1994) 1921-1929

[29] H. Konno, H. Shirakawa, H. Yamazaki, A mean-absolute deviation-skewness portfolio optimization model, Annals of Oper. Res (1993), 45 205-220

[30] Frost, P. A., and Savarino, J. E. (1986), “An Empirical Bayes Approach to Efficient Portfolio Selection,”Journal of Financial and Quantitative Analysis, 21, No. 3, 293–305. (1988), “For better performance; Constrain portfolio weights,” Journal of Portfolio Management,14, 29–34.

[31] Jorion, P. (1992), “Portfolio Optimization in Practice,” Financial Analysts Journal, January/February 68-74.

[32] Michaud, R. (1998), “Efficient Asset Management,” Cambridge: Harvard Business School Press.

[33] Polson, N.G., and Tew, B.V. (2000), “Bayesian Portfolio Selection: An Empirical Analysis of the S&P 500 Index 1970-1996,” Journal of Business and Economic Statistics, 18, 2, 164–73.

[34] Britten-Jones, M. (2002),“Portfolio optimization and Bayesian regression,” Unpublished working paper,London Business School.

[35] Chopra,V. K. and W. T. Ziemba (1993). The effect of errors in mean, variance and co-variance estimates on optimal portfolio choice. Journal of Portfolio Management 19,6-11.

[36] Michael J. Best and Robert R. Grauer. "On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results." Review of Financial Studies, 4, 1991, 315-342.

[37] Jorion, P (1986), “Bayes-Stein Estimation of Portfolio Analysis,” Journal of Financial and QuantitativeAnalysis, 21, 279–292.

[38] Treynor, J. L., and Black, F. (1973), “How to use security analysis to improve portfolio selection,”Journal of Business, January, 66-88.

[39] Kandel, S., and Stambaugh, R. F. (1996), “On the predictability of stock returns: An asset allocation perspective,” Journal of Finance, 51, 2, 385–424.

[40] Zellner, A., and Chetty, V. K. (1965), “Prediction and decision problems in regression models from the Bayesian point of view,” Journal of the American Statistical Association, 60, 608-616.

[41] Klein, R. W., and Bawa, V. (1976), “The effect of estimation risk on optimal portfolio choice,” Journal of Financial Economics, 3, 215–231.

[42] Brown, S. J. (1978), “The portfolio choice problem: Comparison of certainty equivalent and optimal Bayes portfolios,” Communications in Statistics: Simulation and Computation B7, 321–334.

[43] P´astor, L., and Stambaugh, R. F. (2000), “Comparing asset pricing models: An investment perspective,”Journal of Financial Economics, 56, 335–381.

[44] P´astor, L. (2000), “Portfolio selection and asset pricing models,” Journal of Finance, 55, 179–223.

[45] Black, F., and Litterman, R. (1992), “Global portfolio optimization,” Financial Analysts Journal,September-October 1992.

[46] Bollerslev, T., Engle, R., Wooldridge, J.M., 1988. A capital asset pricing model with time varying covariances. Journal of Political Economy 96, 116-131.

[47] Engle(2002) “Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models, Journal of Business and Economic Statistics, 20, pp339-350

[48] Goetzmann, W. N., Li, L. and Rouwenhorst, K. G.: 2005, Long-term global market correlations, Journal of Business 78, 1.38.

[49] Longin, F. and B. Solnik (1995), 'Is the International Correlation of Equity Returns Constant: 1960-1990?' Journal of International Money and Finance, February, 3-26

[50] Ang, A. and Chen, J.: 2002, Asymmetric correlations of equity portfolios, Journal of Financial Economics 63, 443.494.

[51] Ledoit, O., Santa-Clara, P. and Wolf, M.: 2003, Flexible multivariate GARCH modeling with an application to international stock markets, Review of Economics and Statistics 85, 735.747.

[52] Erb, C., Harvey, C. and Viskanta, T.: 1994, Forecasting international equity correlations, Financial Analyst Journal November/December, 32.45.

[53] "An Analysis of Covariance Risk and Pricing Anomalies." Tobias J. Moskowitz; Review of Financial Studies, 2003, 16(2), pp. 417-57.

[54] Michael J. Best, "Portfolio Optimization"(book), 2010 , Chapman & Hall/CRC Finance Series