Inverse Problems in Portfolio Selection: Scenario Optimization Framework

by

Kaushiki Bhowmick

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Quantitative Finance

Waterloo, Ontario, Canada, 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A number of researchers have proposed several Bayesian methods for portfolio selection, which combine statistical information from financial time series with the prior beliefs of the portfolio manager, in an attempt to reduce the impact of estimation errors in distribution parameters on the portfolio selection process and the effect of these errors on the performance of 'optimal' portfolios in out-of-sample-data.

This thesis seeks to reverse the direction of this process, inferring portfolio managers probabilistic beliefs about future distributions based on the portfolios that they hold. We refer to the process of portfolio selection as the forward problem and the process of retrieving the implied probabilities, given an optimal portfolio, as the inverse problem. We attempt to solve the inverse problem in a general setting by using a finite set of scenarios. Using a discrete time framework, we can retrieve probabilities associated with each of the scenarios, which tells us the views of the portfolio manager implicit in the choice of a portfolio considered optimal.

We conduct the implied views analysis for portfolios selected using expected utility maximization, where the investor's utility function is a globally non-optimal concave function, and in the mean-variance setting with the covariance matrix assumed to be given.

We then use the models developed for inverse problem on empirical data to retrieve the implied views implicit in a given portfolio, and attempt to determine whether incorporating these views in portfolio selection improves portfolio performance out of sample.

Acknowledgements

I would like to thank everyone who made this possible. I am very grateful to my supervisor, Prof. David Saunders, for giving me excellent guidance on my research and for encouraging me to always go that extra mile. His encouragement and support has been crucial during my research and the completion of this thesis.

I would also like to thank the faculty and staff at the University of Waterloo, in particular my readers, Prof. Michael J. Best and Prof. Don L. McLeish, for their continued support and constructive comments regarding my research, as well as Mary Flatt, the MQF Program Coordinator for her help, kindness and encouragement during my year at Waterloo.

I am very grateful for the financial support from Faculty of Mathematics at University of Waterloo in the form of Teaching Assistantships, as well as from Prof. David Saunders and WATRISQ in the form of Research Assistantships.

Finally, I would like to thank my parents and my close friends for the love, patience and encouragement they give me for every decision I make.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Academics as well as finance practitioners are well aware of the limitations of the application of classical portfolio theory to practical problems in portfolio selection. Academic literature has several examples of the impact of estimation errors in distribution parameters on the portfolio selection process[1] and the effect of these errors on the performance of 'optimal' portfolios in out-of-sample-data.[2].

This has led a number of researchers to propose various Bayesian methods for portfolio selection, which combine statistical information from financial time series with the prior beliefs of the portfolio manager[3]. This thesis seeks to reverse the direction of this process, inferring portfolio managers' probabilistic beliefs about future distributions based on the portfolios that they hold. We refer to the process of portfolio selection as the forward problem and the process of retrieving the implied probabilities, given an optimal portfolio, as the inverse problem.

In this thesis, we attempt to solve the inverse problem in a general setting by using a finite set of scenarios. Using a single period framework, we can retrieve probabilities associated with each of the scenarios, which tells us the views of the portfolio manager implicit in the choice of a portfolio considered optimal. Implied views analysis can allow us to determine whether or not a current strategy is consistent with existing views on the future values of the assets. It can also allow us to test whether a given portfolio is optimal or even feasible for a certain set of scenarios.

This thesis will look at the inverse problem where the forward problem is formulated using a expected utility maximization approach to portfolio selection in a discrete-time, one-period scenario framework for a risk-averse investor, subject to a budget constraint.

[1]See (Best & Grauer, 1991), (Huang, 2001).
[2](Broadie, 1993), (DeMiguel, Garlappi, & Uppal, 2009).
[3]See(Black & Litterman, 1992), (Meucci, 2006).

The forward problem in this case is a nonlinear mathematical program with respect to the portfolio weights. This leads to an inverse problem with nonlinear constraints, that depend on the first derivative of the investor's utility function. We will also look at the inverse problem where the portfolios are selected using a mean-variance approach in a scenario framework.

The thesis is structured as follows: The remainder of Chapter 1 summarizes results from relevant literature as it pertains to our topic. Chapter 2 gives a general background on the two types of portfolio selection methods we consider in this thesis and highlights the pertinent results and assumptions that will be used later on. Chapter 3 provides the optimization framework for the forward and the inverse problems, and goes through the steps of formulating the inverse problem from the optimality conditions needed for the forward problem, first in the case of complete markets, and then extending the analysis to incomplete markets. We will then provide examples of the inverse problem with empirical datasets. Chapter 4 is where we will provide an example of using the results from the inverse problems to generate optimal portfolios using a weighted combination of the implied and empirical probabilities, and compare the out-of-sample performance of these portfolios to the $\frac{1}{N}$ portfolio, equal risk-contribution (ERC) portfolio, the original optimal portfolio, and minimum-variance portfolio.

1.1 Literature Review

There is extensive research on inverse problems in finance related to the calibration problem[4], the problem of retrieving implied volatilities[5] from option prices, and the problem of retrieving risk-neutral distributions[6] from the existing assets in the asset universe. The research on inverse problems in portfolio selection is, however, limited to retrieving implied views on asset returns as suggested by a particular choice of portfolio. (Black & Litterman, 1992), (Scherer, 2004), (Disyatat & Gelos, 2001) all look at using inverse portfolio optimization as a risk management tool to compare the mean returns implied in a selected portfolio to the forecasted values of those asset returns. The methods used in Black and Litterman (1992), Scherer (2004), Disyatat and Gelos (2001) rely on portfolios retrieved using the mean-variance (MV) model proposed in Markowitz (1991), or some variant thereof. The approach outlined in the above sources biases the implied returns to a particular model of asset returns, namely that they have a normal distribution, as this is a primary assumption of the MV model. This thesis seeks to find the implied probabilities

[4]Zubelli (2005) gives a brief summary of inverse problems as it applies to calibration.

[5]Mayhew (1995) provides a fairly comprehensive review of methods employed to retrieve implied volatilities.

[6]See Jackwerth (1999) for a review on methods to retrieve risk-neutral distributions and implied binomial trees from option prices.

associated with a finite set of scenarios, instead of just the asset returns. In this thesis, we primarily attempt to find the implied probabilities of scenarios when investors have a well-behaved, concave utility function (not necessarily quadratic). The framework proposed in this thesis also allows us to model asset returns that do not follow a normal distribution.

Hartley and Bakshi (1998) look at the inverse problem for Markowitz models of portfolio selection[7]. Their paper develops algorithms to determine the investor's return-generating process, and the utility function based on the portfolio choices made by the investor. Their approach uses a time-series of actual portfolios for a sample of investors, the time-series of preceding observed prices and dividends of the relevant risky-assets, and a set of socio-economic characteristics for each investor to calibrate the parameter values in each investor utility function and the associated parameters in the returns generating process.

Even though they use a stochastic returns generating process, they have been able to formulate their inverse as a deterministic problem because the inputs to the portfolio selection problem require the *expected returns* and associated *covariance matrix* which are both deterministic even if the actual returns are stochastic. They use a simple auto-regressive model as a basis for the returns generating process, and by formulating the inverse problem as a deterministic problem in terms of the means and the covariances, they ignore the uncertainty aspect that would be indicated by the stochastic returns process - the stochasticity is limited to the error parameter in the auto-regressive model. The authors use a Quasi-Newton Algorithm to estimate the auto-regressive parameters for the rates of returns using the time-series of prior portfolios and asset prices as well as the parameters of the investors' utility function . To implement the algorithm, the authors use the deterministic Neoclassical Econometric Methods (Hartley, 1986, n.d., 1994).

This thesis models the uncertainty in asset prices using a set of discrete scenarios rather than a stochastic returns process. Since we are concerned with retrieving the probabilities of scenarios, and not just the expected returns, we cannot ignore the uncertainty aspect. The inverse problem that we look at also assumes an investor's utility function to be given and seeks to find the probabilities of the scenarios implied by a particular 'optimal' portfolio.

Merkulovich and Rosen (1998), and Dembo and Rosen (2000), on the other hand, use a scenario framework similar to the one proposed in this thesis. [8] The authors assume a set of finite scenarios and a finite number of assets, with the assumption that investors in the market have a probability associated with the realization of each of the scenarios. To formulate the forward problem, the authors use a portfolio replication argument to find a replicating portfolio that behaves identically to the benchmark portfolio under all

[7] See Markowitz (1952, p. 77-91) and Markowitz (1991) for details.

[8] The scenario framework proposed by the authors was the motivation for the framework proposed in this thesis.

scenarios. The benchmark portfolio could be the risk-free rate, or the market portfolio, or any other portfolio chosen by the investor. A measure of risk used by the authors is the expected under-performance of the replicating portfolio with respect to the benchmark, called downside regret. The investor's utility function is modelled as a linear function of the difference between the expected excess profit and the downside regret, subject to the trading constraints. They then go on to find the probabilities associated with the scenarios, which make a given portfolio optimal under the constraints imposed on the forward problem using principles of linear programming and duality theory.

It is important to note that Merkulovich and Rosen (1998), and Dembo and Rosen (2000) use a piecewise linear utility function for their forward problem whereas we use a general (nonlinear) utility function that leads to nonlinear constraints.

Merkulovich and Rosen (1998) also present a framework for retrieving implied views in a mean-variance framework where there are a finite number of assets, and the asset returns follow a multi-normal distribution. In this case, they assume that the forward portfolio selection problem is only subject to a budget constraint. To solve the inverse problem, the authors assume the covariance matrix to be given and attempt to retrieve the vector of asset returns that would make a given portfolio optimal. The authors also present a way to express implied views in the form of expected factor returns when the asset returns are modelled using a factor model. We will show, in this thesis, a way to formulate the inverse problem in a scenario setting, when the asset returns follow a general (non-normal) distribution. We will briefly explore the behaviour of the problem when the covariance matrix is not known and also when it is assumed to be given.

Zagst and Pöschik (2008) approach the inverse problem in a mean-variance framework for portfolios that have been optimized under constraints. First, they present the implied views on the asset returns in an unconstrained case, and then proceed to formulate the inverse problem in the presence of constraints for the forward problem. They conclude that with the addition of constraints, when the covariance matrix is assumed to be given, it is only possible to retrieve views on asset returns relative to those of another asset in the portfolio, as opposed to absolute views on asset returns possible in the unconstrained case. The method proposed by the authors does not give us implied distributions of asset values, but instead attempts to solve for the vector of relative asset returns that would make a particular portfolio optimal.

Best and Grauer (1985) also solve the inverse problem for the MV case in much the same way as we do in this thesis. The authors show that the set of expected return vectors, for which an observed portfolio is mean-variance efficient, is a two parameter family. The derivation of the implied expected return vectors is dependent on a known portfolio, and a known covariance matrix. The authors test the validity of assumption that the mean vector is constant through time in multi-period models using a known covariance matrix and the observed time series of market value weights. The derivation for the implied views for the MV case in this thesis follows the derivation for implied returns by Best and Grauer (1985). However, in this thesis, we use a markedly different framework to retrieve implied views on possible future scenarios themselves, rather than on a set of implied mean returns through time.

In this thesis, we will attempt to solve for the probability distribution implied by an optimal portfolio obtained through MV optimization under a single budget constraint. Since various probabilities can give the same asset returns, we will attempt to minimize the distance between a chosen prior probability vector and the implied probability vector to determine whether the chosen portfolio is optimal under a new set of probabilities.

The next chapter gives an overview of portfolio selection methods used in this thesis: Expected Utility Maximization and Mean-variance optimization.

Chapter 2

Portfolio Selection Methods: A Primer

To properly understand the inverse problem in a portfolio optimization context, it is first necessary to understand the forward portfolio optimization problem.

Portfolio optimization as we know it now attempts to optimize the investor's objective function subject to certain constraints. In Modern Portfolio Theory, the problem of optimal portfolio selection attempts to maximize expected return on the portfolio subject to an acceptable level of risk and other constraints, or alternatively, to minimize the portfolio risk subject to a minimum level of required return on the portfolio, and other constraints.

In a more general setting, we can construct a utility function that meaningfully balances the risk associated with holding the portfolio with the return associated with the portfolio. The set of inputs can be given by the distribution of market factors (such as asset prices, bond yields, risk-free rates) for a finite set of assets sometime in the future, the set of market prices for the same assets at the current time, and a set of constraints such as liquidity constraints, short-selling constraints, minimum required return or maximum allowable risk constraints, etc. Given the above information, we can formulate an optimization problem which can be solved using methods of linear or non-linear programming. Whether or not an optimization problem is a linear program or a non-linear program depends on the objective function (linear or non-linear) and the set of constraints. We traditionally work with linear constraints since it makes it easier to compute the optimal solution.

When we say a solution is optimal, we mean that no other feasible solution provides us with a better value for the objective function. It is necessary to note that depending on the objective function and the feasible region, there could be more than one optimal solution.

The inverse portfolio optimization problem works in the opposite direction. In this case, we have the investor's optimal portfolio and the investor's utility function and constraints; we attempt to solve for the probability distribution that makes the given portfolio optimal. Basically, what we are solving for are the views that the investor has on the market given that he or she considers a given portfolio to be optimal.

Inverse problems are generally ill-posed in that they do not necessarily have a unique solution (See Engl & Kügler, 2005; Aster, Borchers, & Thurber, 2005). An inverse problem will yield different results depending on the conditions imposed on the problem. This is also true in the case of inverse problems in portfolio optimization. Therefore, it is important to note that though the inverse problem might yield an implied distribution, this distribution might not be unique. To obtain useful information about the implied views given an investor's utility function and optimal portfolio, it might be necessary to assume certain parameters as given, even though we might not necessarily know their real values. In this sense, we might want to consider the implied distribution retrieved from the inverse problem as a function of the variables assumed to be given.

We will concern ourselves only with the following portfolio optimization problems: modern portfolio theory models; specifically, the mean-variance portfolio selection, and expected utility maximization; specifically, maximizing the expected utility when the investors' utility preferences are given by a constant relative risk aversion (CRRA) utility function. Characteristics of both problems are described in greater detail below.

2.1 Expected Utility Maximization

In economics, utility is a measure of relative satisfaction. Utility is seen as a way to describe preferences. A utility function is a way of quantifying consumption, such that more preferred patterns of consumption get assigned a higher value than less preferred patterns. [1]

In order to simplify calculations, various assumptions have been made of utility functions. Most utility functions used in modelling or theory are well-behaved. They are usually monotonic, quasi-concave, continuous and there is no unconstrained global optimum. Examples of well-behaved utility functions include, but are not limited to:

- CRRA (constant relative risk aversion, or isoelastic) utility
  The function is given by:
  $U(w) = \begin{cases} \frac{w^{1-\eta}-1}{1-\eta} & for\ \eta \neq 1, w > 0 \\ \log(w) & for\ \eta = 1, w > 0 \end{cases}$

[1] Please see Varian (2005) for more information on consumer preferences and utility.

where η is the risk aversion parameter and w is the level of wealth or consumption. [2]

- Exponential utility
  This function is given by
  $U(w) = 1 - e^{-\eta w}$
  where η is the risk aversion parameter and w is the level of wealth or consumption.

- Quasilinear utility
  This function is given by:
  $U(\mathbf{w}) = w_1 + \nu(w_2, ..., w_n)$
  where w_i is amount of good i consumed. ν is usually a concave function as we want preferences to be quasi-concave. [3]

However, it is possible for preferences not to be representable by a utility function; case in point: lexicographic preferences. In case of lexicographic preferences, a consumer infinitely prefers one good to another. For example, if there are two goods X and Y, and a consumer shows lexicographic preference for good X, then he or she will always pick the bundle of goods that offers the maximum amount of X, regardless of the amount of Y. The amount of good Y will only factor into the decision if all bundles offer the same amount of good X. There is literature in consumer theory about mathematical properties of lexicographic preferences and where they arise in real-world interpretations of consumer utility. (See Ehlers, 2002; Fishburn, 1975)

The expected utility theory deals with the analysis of choices among risky projects with (possibly multidimensional) outcomes. The expected utility model was first proposed by Daniel Bernoulli in 1738 (See (Bernoulli, 1954)). The first important use of the expected utility theory was that of John von Neumann and Oskar Morgenstern who used the assumption of expected utility maximization in their formulation of game theory.(Details can be found in von Neumann & Morgenstern, 1944, second ed. 1947, third ed. 1953) The expected utility theory allows the representation of the level of satisfaction by the sum of utilities from outcomes weighted by the probabilities of these outcomes.

Expected utility is a valid choice of objective in the face of uncertainty for a very good reason - we assume that possible outcomes are independent. When we model uncertainty using randomness, we essentially say that *only one* of the many possible outcomes is going to occur, but at the current time, we can only assign likelihoods of occurrence to the outcomes in question. Utility of each outcome is also independent of the utility of any

[2]This function is also referred to as the *power* utility function for $\eta \neq 1$ and as the *logarithmic* utility function for $\eta = 1$, which is the limiting case of the power utility as $\eta \to 1$.

[3]Arrow and Enthoven (1961) provide conditions for quasi-concave functions under which a point satisfying the Kuhn-Tucker conditions is a point of constrained global maximum. This is a useful result when one is trying to maximize a utility function.

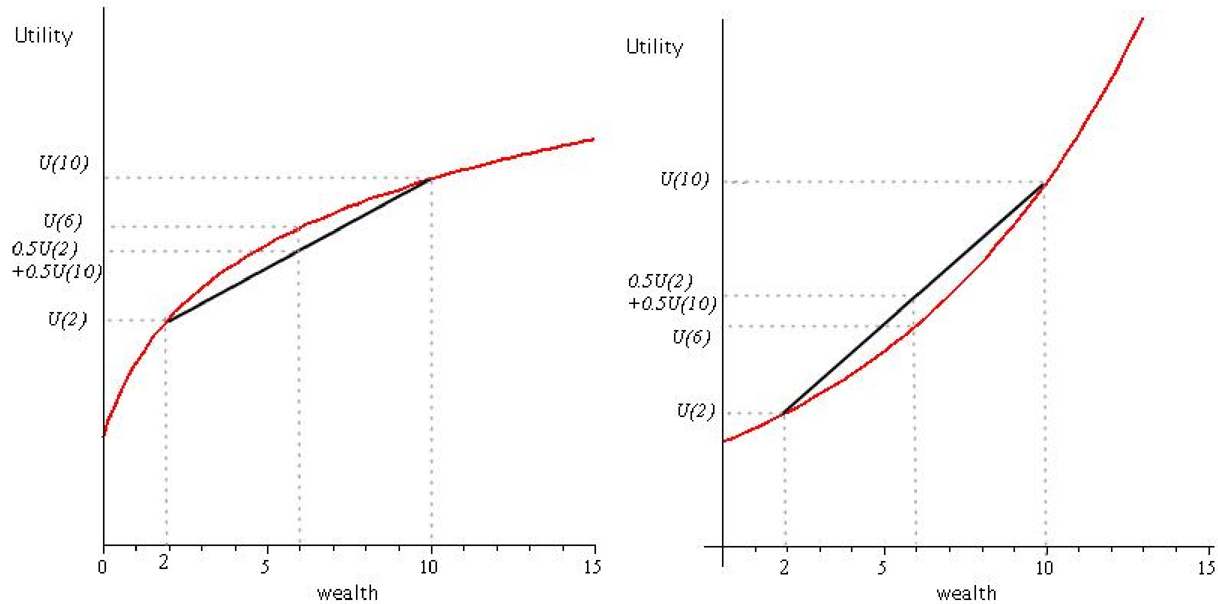
other outcome occurring. To conveniently model the utility of all the possible outcomes sometime in the future at the current instant, it is necessary to weigh the utility of each of the aforementioned outcomes with the probability of each of those outcomes occurring, which gives us the expected utility function.

For example, if we know that there are a finite set of outcomes that can occur at the end of the time period given, then we can denote the value of the commodity in question in outcome i by c_i. Suppose we know that only S possible outcomes can occur and that each outcome has a probability p_i associated with it. Then we can write the expected utility function as, $\sum_{i=1}^{S} p_i U(c_i)$, where $U(\centerdot)$ denotes the consumer's utility function.

The expected utility function also allows us to qualify an investor's preference for risk. We define risk-aversion as follows:

- We say an investor is **risk-averse** if the investor's expected utility function is concave (Figure 2.1a), i.e. the utility of the expected value of wealth is greater than the expected value of the utility of wealth.
- We say an investor is **risk-neutral** if the investor's expected utility function is linear, i.e. the utility of the expected value of wealth is equal to the expected value of the utility of wealth.
- We say an investor is **risk-loving** or **risk-seeking** if the investor's expected utility function is convex (Figure 2.1b), i.e. the utility of the expected value of wealth is less than to the expected value of the utility of wealth.

We refer the reader to Prigent (2007) for more details on expected utility theory as it applies to risk and reward and its role in portfolio optimization.



(a) Utility function of a risk-averse investor (b) Utility function of a risk-loving investor

Figure 2.1: Risk Preference Profiles.

For a risk-averse investor 2.1a, the utility of the expected value of wealth *U(6)* is greater than the expected value of the utility of wealth *0.5U(2)+0.5U(10)*, and for a risk-loving investor 2.1b, the utility of the expected value of wealth *U(6)* is less than the expected value of the utility of wealth *0.5U(2)+0.5U(10)*.

2.2 Mean-Variance Portfolio Selection

Modern Portfolio Theory (MPT), introduced by H.M. Markowitz(1991) is a theory of portfolio optimization that attempts to maximize the portfolio return relative to the level of portfolio risk. In MPT, portfolio risk is denoted by the standard deviation of the portfolio. In the general mean-variance optimization approach an efficient frontier is constructed which consists of all possible optimal portfolios that maximize expected return for a given level of risk which is given by the standard deviation of their portfolios. The general approach assumes that the distribution of all assets is joint normal and given. Best (2010) provides a concise way of generating the efficient frontier using linear algebra and optimization. We replicate some of the results below.

In general, a portfolio consists of various assets held in different amounts. Consider a universe of n assets. Let μ_i denote the return on asset i, $i = 1, 2, ..., n$ and σ_{ij} denote the covariance between assets i and $j, 1 \leq i, j \leq n$. The vector of asset returns is given by $\mu = (\mu_1, \mu_2, ..., \mu_n)^T$ and $\Sigma = [\sigma_{ij}]$ is the variance-covariance matrix for the assets. Σ is symmetric and positive semidefinite. Let x_i denote the fraction invested in asset i $i = 1, 2, ..., n$. So, the vector of portfolio weights is given by $\mathbf{x} = (x_1, x_2, ...x_n)^T$.

We can now denote the expected portfolio return μ_p and portfolio variance σ_p^2 as follows:

$$\mu_p = \mu^T \mathbf{x} \text{ and } \sigma_p^2 = \mathbf{x}^T \Sigma \mathbf{x}$$

We want to find an $\mathbf{x}$ such that the μ_p is large and σ_p^2 is small.

In deriving the results below, we make the following assumptions throughout this section:

Assumption 2.2.1. Σ *is symmetric and positive definite*

Assumption 2.2.2. $\mathbf{l}_n$ *is an* $n \times 1$ *vector of ones, and* μ *is not a multiple of* $\mathbf{l}_n$

Definition 2.2.1. *A portfolio is said to be a* ***variance-efficient portfolio*** *if for a fixed level of portfolio return, there is no other portfolio with a lower variance.*

This definition implies that the variance-efficient portfolios are solutions to the following problem, when μ_p is greater than the return on the minimum variance portfolio:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \Sigma \mathbf{x} \\ \text{s.t: } & \mu^T \mathbf{x} = \mu_p \\ & \mathbf{l}_n^T \mathbf{x} = 1 \end{aligned} \tag{2.2.1}$$

Definition 2.2.2. *A portfolio is said to be a* ***efficient portfolio*** *if for a fixed level of portfolio variance, the portfolio return is greater than the return on the minimum-variance portfolio, or is said to be* ***inefficient*** *otherwise.*

Note that (2.2.1) will also give us the inefficient portfolios when μ_p is less than the return on the minimum variance portfolio.

Definition 2.2.3. *A portfolio is said to be a* ***expected return-efficient portfolio*** *if for a fixed level of portfolio variance, there is no other portfolio with a higher return.*

This definition implies that the efficient portfolios are solutions to the following problem:

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mu^T \mathbf{x} \\ \text{s.t: } & \mathbf{x}^T \Sigma \mathbf{x} = \sigma_p^2 \\ & \mathbf{l}_n^T \mathbf{x} = 1 \end{aligned} \tag{2.2.2}$$

We can combine the return of the portfolio with the risk of the portfolio in the following manner:[4]

$$\begin{aligned} \max_{\mathbf{x}} \quad & \eta\mu^T\mathbf{x} - \tfrac{1}{2}\mathbf{x}^T\Sigma\mathbf{x} \\ & \text{s.t: } \mathbf{l}_n^T\mathbf{x} = 1 \end{aligned} \tag{2.2.3}$$

In the above formulation η is the risk aversion parameter. Larger values of η put more weight on the higher returns at the expense of portfolio risk, whereas lower values of η put more importance on minimizing risk rather than maximizing returns. It is important to note that $\eta = 0$ gives us the minimum-variance portfolio.

Definition 2.2.4. *A portfolio is said to be* ***mean-variance efficient*** *if it is the optimal solution to equation* (2.2.3) *for a given risk-aversion parameter* $\eta \geq 0$.

Each of (2.2.1), (2.2.2), (2.2.3) generates a family of solutions that is identical to the other provided μ is not a multiple of $\mathbf{l}_n$ and $\mu_p \geq \mu_{MVP}$, where μ_{MVP} is the return on the minimum-variance portfolio.

The relationship between the portfolio variance (σ_p^2) and the portfolio return (μ_p) for efficient portfolios is called the efficient frontier. It can be shown using the optimality conditions for (2.2.3) and linear algebra, that the efficient frontier is a parabola in μ_p-σ_p^2 space. The algebraic relationship between the portfolio variance (σ_p^2) and the portfolio return (μ_p) is given by the following equation:

$$\sigma_p^2 - \beta_0 = \frac{(\mu_p - \alpha_0)^2}{\alpha_1} \tag{2.2.4}$$

where,

$$\begin{aligned} \alpha_0 &= \mu^T \left[\frac{\Sigma^{-1}\mathbf{l}_n}{\mathbf{l}_n^T\Sigma^{-1}\mathbf{l}_n} \right] \\ \alpha_1 &= \mu^T \left[\Sigma^{-1}\mu - \frac{\mathbf{l}_n^T\Sigma^{-1}\mu}{\mathbf{l}_n^T\Sigma^{-1}\mathbf{l}_n}\Sigma^{-1}\mathbf{l}_n \right] \\ \beta_0 &= \left[\frac{\Sigma^{-1}\mathbf{l}_n}{\mathbf{l}_n^T\Sigma^{-1}\mathbf{l}_n} \right]^T \Sigma \left[\frac{\Sigma^{-1}\mathbf{l}_n}{\mathbf{l}_n^T\Sigma^{-1}\mathbf{l}_n} \right] \end{aligned}$$

[4] See Best (2010, p. 24)

When there is a risk-free asset in the market, we can modify (2.2.3) as follows:

$$\max_{\mathbf{x}} \quad \eta\, [\mu^T, R] \begin{bmatrix} x \\ x_{rf} \end{bmatrix} - \frac{1}{2} \begin{bmatrix} x \\ x_{rf} \end{bmatrix}^T \begin{bmatrix} \Sigma & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ x_{rf} \end{bmatrix}$$
$$\text{s.t: } \mathbf{l}_n^T x + x_{rf} = 1 \tag{2.2.5}$$

Since Σ is positive definite, we can find the weights in risky assets as:

$$x = \eta \Sigma^{-1}(\mu - R\mathbf{l}_n) \tag{2.2.6}$$

And the risk-free asset as:

$$x_{rf} = 1 - \eta \Sigma^{-1} \mathbf{l}_n^T (\mu - R\mathbf{l}_n) \tag{2.2.7}$$

Please refer to Best (2010, p. 21-51), for complete derivation of the above results.

It is important to note that mean-variance analysis relies heavily on proper estimation of the mean vector and variance-covariance matrix of the assets in the market. It has been shown by Best and Grauer (1991) that optimal portfolios are significantly sensitive to parameter values, especially the asset means. Mean-variance portfolios are also sensitive to the estimation errors in the variance-covariance matrix but to a much lesser degree.

In the next chapter, we introduce our optimization framework and derive the inverse problem for portfolios selected using expected utility maximization and for portfolios selected using mean-variance optimization. For the expected utility maximization case, we derive the results for complete markets first, and then for incomplete markets. We also provide numerical examples, using empirical data, to illustrate how the proposed model performs in practice. We also briefly discuss the stability of the problem, the condition number, and Tikhonov regularization as it applies to the stability of the solution of our inverse problem.

Chapter 3

Optimization Framework

In a portfolio optimization problem, we generally make certain assumptions about the distribution of asset returns. Most often, asset returns are assumed to be normally distributed. However, is known that in reality, asset returns often display heavier tails, i.e. they do not follow a normal distribution[1]. We also tend to estimate the distribution parameters based on past data, but there might not always be enough data to accurately estimate the true parameters of the distribution (Broadie, 1993). The discrete-time, single-period framework allows us to work with non-normal distributions because it models discrete market scenarios observable at some time in the future (Merkulovich & Rosen, 1998).

We will consider a single period model where only a finite set of scenarios $s, s \in \mathbb{Z}$ can occur. At the end of the time period, only one out of the s scenarios can occur, but at time 0, there is uncertainty associated with the outcomes. Let us define this set of independent scenarios as $\Omega := \{\omega_1, \omega_2, ..., \omega_s\}$, where ω_i denotes scenario i. We also assume that investors have a certain probability, p_i, for each of the ω_i scenarios occurring, $i = 1, ..., s$.

The scenarios in question can be any number of risk factors affecting the price of the assets in the asset universe. Provided we know the relationship between these factors and the asset prices, we can convert the risk factor realizations into asset prices at the end of the period. For the sake of simplicity, we will assume that scenarios at the end of the period are asset prices in those respective scenarios.

Here, we introduce the basic definitions of terms, the structure of our optimization framework, and assumptions that will be used throughout the thesis.

We consider the following inputs for the optimization problem. We consider an asset universe of n distinct assets and s possible scenarios. We denote the scenarios by the $s \times n$

[1]McDonald (2006) discusses the validity of the normality assumption for continuously compounded returns.

scenario matrix $\mathcal{A}$, where each element a_{ij} is the value of the asset i in ω_j. We also have a $n \times 1$ price vector, $\mathbf{z}$ denoting the known prices of the n assets at the beginning of the period. Moreover, we also have a $s \times 1$ vector of probabilities corresponding to each of the s scenarios. In the forward optimization problem, we are trying to find the $n \times 1$ vector of optimal portfolio allocations,$\mathbf{y}^*$, where y_i is the number of *units* or *volume* of asset i in our portfolio.

The structure of the inputs is given below:

$$\mathcal{A}_{s\times n} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & a_{s3} & \dots & a_{sn} \end{bmatrix}, \ \mathbf{z}_{n\times 1} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}, \ \mathbf{y}_{n\times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \ \mathbf{p}_{s\times 1} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_s \end{bmatrix} \tag{3.0.1}$$

Our portfolio optimization problem is subject to the available initial wealth (budget) given by $B > 0$.

This means that the initial value of our portfolio is given by:

$$\mathbf{z}^T\mathbf{y} \leq B. \tag{3.0.2}$$

If we denote the portfolio weights by $\mathbf{x}$ such that $\mathbf{x}^T\mathbf{l}_n = 1$, we can define this $n \times 1$ vector as $\mathbf{x} = [x_1, x_2, ..., x_n]^T$, where $x_i = \dfrac{z_i y_i}{B}$, $i = 1...n$. This is an important distinction to make as $\mathbf{x}$ denotes the proportion of total wealth invested in each asset, whereas $\mathbf{y}$ denotes the actual units of each asset held in the portfolio. To simplify our analysis of the problems posed later in this chapter, we will allow an investor to hold fractional units of any asset.[2]. To make this distinction clear, we will refer to $\mathbf{y}$ as *portfolio allocation (by volume)* and $\mathbf{x}$ as *portfolio weights*.

If there is a risk-free asset in the market, (3.0.2) is always satisfied with equality.

Assumption 3.0.3. *There exists a risk-free asset in the market with a positive rate of return r, such that the total return in all scenarios is given by $R = (1 + r)$.*

The results in this thesis assume 3.0.3 to be true always unless stated otherwise.

[2]In reality, asset units are integer values. However, imposing the restriction that asset volumes be integer quantities turns our portfolio optimization problem into an integer programming problem. Integer linear programming problems belong to the family of NP-hard problems (Papadimitriou, 1981). Nonlinear integer programs also belong to the family of NP-hard problems and are more complex than the integer linear programming problems (Jünger et al., 2009, p. 561-612).

Since $\mathbf{y}$ is the vector of asset volumes in our portfolio, and $\mathcal{A}$ is the possible set of asset values at a future date, we can say that the wealth in each scenario is given by $\mathcal{A}\mathbf{y}$.

Let $\mathcal{A}\mathbf{y} = \mathbf{w}$.

Note that $\mathbf{w}$ is a vector with s elements, and w_i is the wealth in ω_i. Let $\mathbf{A_i}$ be the i^{th} row of the matrix $\mathcal{A}$. If the portfolio weights are given by $\mathbf{y}$, then

$$w_i = \mathbf{A_i}\,\mathbf{y} \tag{3.0.3}$$

Figure 3.1 explains the relationship between the above inputs, as it applies to our optimization problem, in a slightly more graphical manner.

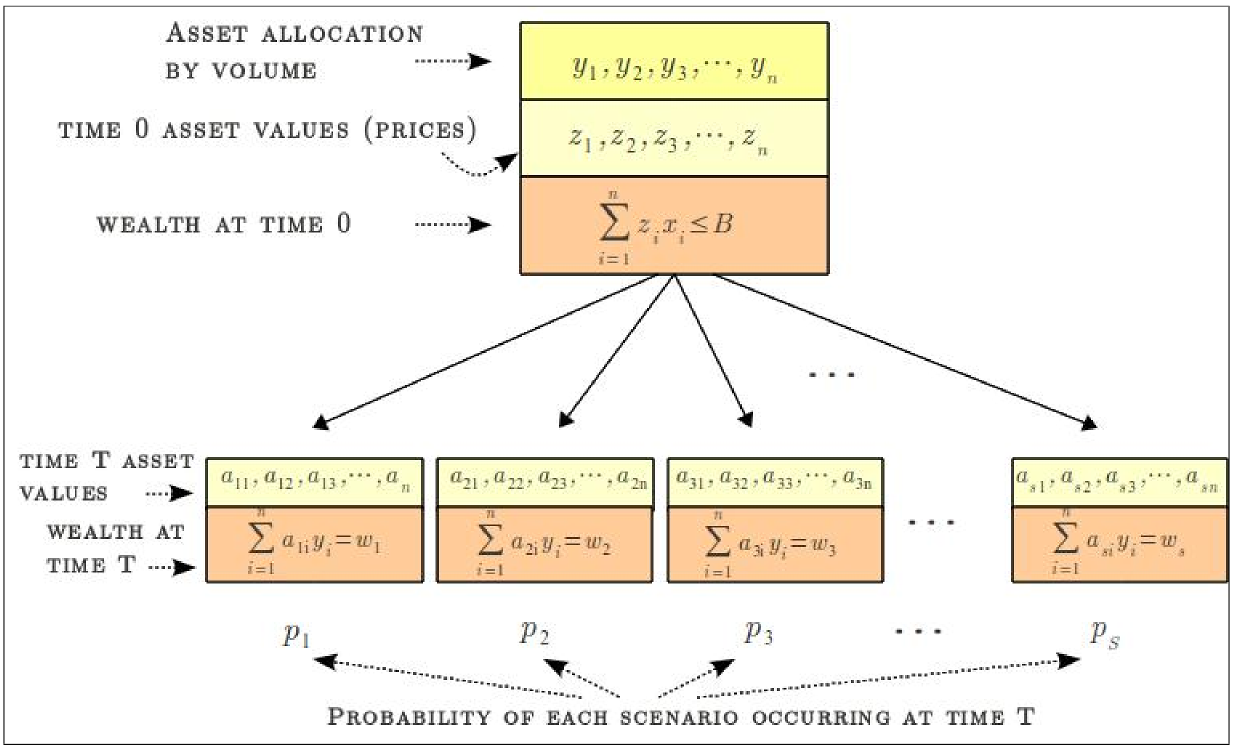


Figure 3.1: Structure of the Optimization Framework

There are n assets and s scenarios. At time 0, the values of the assets are given by the asset prices z_i. The inequality denoting wealth at time 0 is always satisfied with equality if there is a risk-free asset in the market. At time T, the values of the assets are given by a_{ji} for each of the scenarios j, $j = 1...s$ and assets i, $i = 1...n$. The wealth in each scenario is denoted by w_i. The probabilities p_j, $j = 1...s$ all sum to 1 and $p_j \geq 0$

Sometimes, as in the case of mean-variance optimization, we want a set of future asset returns rather than future asset values as inputs to the mathematical program. We can get an associated matrix of asset returns from the above inputs very easily.

Let us define the new matrix of asset returns as $\bar{\mathcal{A}}$.

Each entry $\bar{a}_{ij}$ of $\bar{\mathcal{A}}$ is defined as $\dfrac{a_{ij}}{z_i}, a_{ij} > 0, z_i > 0.$ [3]

$$\bar{\mathcal{A}}_{s\times n} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} & \dots & \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} & \dots & \bar{a}_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{s1} & \bar{a}_{s2} & \bar{a}_{s3} & \dots & \bar{a}_{sn} \end{bmatrix}, \tag{3.0.4}$$

If we assume that the scenario matrix encompasses all possible states of the market for the chosen securities, we can classify whether a market is complete or incomplete.

Definition 3.0.5. *We define a* **contingent claim** *as a random variable V on Ω.*

Definition 3.0.6. *In general, a market is said to be* **complete** *if all contingent claims can be hedged, i.e. $\exists$ a replicating portfolio θ such that $\mathcal{A}\theta = V \ \ \forall V \in \mathbb{R}^s$.*

In terms of the inputs described above, an equivalent condition for completeness is if $Rank(\mathcal{A}) = s$.

We will define a market to be **complete without redundant assets** *if it is* complete *and $n = s$.*

Definition 3.0.7. *We say the market is* **incomplete** *if $Rank(\mathcal{A}) < s$, i.e. there are more scenarios than the (non-redundant) assets in the market. If we have more scenarios than assets, it means that all assets are non-redundant.*

If $n > s$, then we have two cases that we need to check.

Case 1 $n > s$ and $Rank(\mathcal{A}) = s$

[3]In most cases, we can safely assume that the asset values are strictly non-negative and that prices at time 0 are positive. However, in some cases, such as with futures and forwards, it is possible for the asset values to take on negative values at the end of the time period, and for the price at time 0 to be 0. In such cases, we can define the cost of the futures or forwards as the funding requirements or the capital requirements for owning a position in the derivative.

- By definition of $Rank$, we know that $\mathcal{A}$ has exactly s linearly independent(L.I.) columns and that $Col(\mathcal{A})$ is the span of these L.I. columns.
- By the Spanning Set Theorem in Appendix A.1, we can successively remove those columns (redundant assets) that are a linear combination of the L.I columns, till we are left with only s linearly independent columns.
- This leaves us with s independent assets and s scenarios, which gives us a complete market as per the definition above.

Case 2 $n > s$ and $Rank(\mathcal{A}) < s$

- Let $Rank(\mathcal{A}) = k$. Again, by the definition of $Rank$, we know that $\mathcal{A}$ has exactly k L.I. columns. And by the Rank Theorem in Appendix A.2, we know that $\mathcal{A}$ has exactly k leading ones.
- Note that the rows of $\mathcal{A}$ denote the asset values in a particular scenario, ω_i. If $k < s$, then there is at least one redundant row, i.e. at least one row is linearly dependant on the s independent rows denoted by the leading ones. It should be clear that we cannot have $Rank(\mathcal{A}) < s$, since by the definition of our inputs, scenarios are independent, i.e. they cannot be linear or convex combinations, or scalar multiples of other scenarios.

Since we can always convert a complete market where $n > s$ into a complete market without redundant assets, we will only consider markets without redundant assets for our analysis.

3.1 Generating the Scenario Matrix

The optimization framework proposed above assumes the future asset values for Ω to be given. In practice, however, we seldom know what the future will look like. Even as we try to model the uncertainty inherent in estimating future asset values by using scenarios, we can seldom assign objective probabilities to those scenarios. How then, do we get the scenario matrix for our portfolio optimization problem?

Since this thesis focuses on inverse problems in portfolio optimization, we will only concern ourselves with the problem of generating a viable scenario matrix. The inverse problems focus on retrieving the probabilities associated with each of the scenarios given an "optimal" portfolio.

The easiest method to get a scenario matrix would be to generate the possible set of scenarios using Monte Carlo(MC) simulations. To generate these joint scenarios for asset values, we first need to calibrate a model using historical time-series. Once we have the calibrated parameters, we can generate several joint scenarios quite easily from various non-normal distributions using MC techniques.[4]

We can also use Quasi-Monte Carlo(QMC) techniques to get a better coverage of the scenario set. The advantage with using QMC methods to simulate the joint scenarios is the fact that QMC sequences give us points that are very evenly spaced in the probability space.(See Joy, Boyle, & Tan, 1996).

3.2 Expected Utility Maximization in Complete Markets

We formulate our problem based on the expected utility hypothesis detailed in Section 2.1. It is important to note that different people could have a different utility for the same payoff due to different personal preferences.

We will hold the following assumption to be satisfied throughout the thesis:

Assumption 3.2.1. *for expected utility maximization problems, the investor's utility function is given by* $\mathcal{U}$*, such that* $\mathcal{U} \in \mathcal{C}^{(2)}(\mathcal{D})$ [5]*,* $\mathcal{U}' > 0$ *and* $\mathcal{U}$ *is concave, i.e.* $\mathcal{U}'' < 0$.

This family of utility functions reflects the preferences of a rational investor who will invest more money only if the utility increases of investing increases with the amount invested. However, we will only concern ourselves with risk-averse investors, whose marginal utility of wealth decreases as wealth increases.

To solve our optimization problem, we will need to use the method of Lagrange multipliers. Inputs in (3.0.1) and Assumption 3.2.1 guarantee that our problem formulation below satisfies the necessary and sufficient conditions for optimality as stated in the *Lagrange Multiplier Theorem* in Bertsekas (1995), Proposition 3.1.1 and Proposition 3.2.1 (restated in Appendix B for convenience).

Note that maximizing $[\mathcal{U}(x_1), \ldots \mathcal{U}(x_m)]\mathbf{c}$ is equivalent to minimizing $-1 \times [\mathcal{U}(x_1), \ldots \mathcal{U}(x_m)]\mathbf{c}$, where $\mathbf{c}$ is a $m \times 1$ vector of positive constants.

It is easy to verify Proposition 3.1.1 since we only have one linear constraint as stated below. Additionally, both the objective function and the constraint are twice continuously differentiable, by definition.

[4]See McLeish (2005) for several examples of generating asset returns from non-normal distributions.

[5]$\mathcal{C}^{(2)}(\mathcal{D})$ is the family of twice-differentiable functions on the domain $\mathcal{D}$, where $\mathcal{D}$ denotes the domain of $\mathcal{U}$.

To verify Proposition 3.2.1, note that $\nabla^2_{xx} L(x^*, \lambda^*)$ in our case is given by,$-1[\mathcal{U}(x_1), \ldots \mathcal{U}(x_m)]\, \mathbf{c}$, which is strictly greater than 0 by definition of $\mathcal{U}$. So, $y^T\, \nabla^2_{xx} L(x^*, \lambda^*)\, y > 0, \quad \forall y \neq 0$ with $\nabla h(x^*) = 0$.

Let us assume that the form of the inputs for our optimization is given by (3.0.1). However, since we are dealing with complete markets, the scenario matrix $\mathcal{A}$ is square and invertible. The fact that our market is complete means that we can use properties of invertible matrices from linear algebra to get "nice" results. Appendix A.3 states the main results for invertible matrices that will be used below.

Suppose the total amount of money that one can invest is given by B, our optimization problem can be written as a convex programming problem as follows:

$$\begin{aligned} &\max_{y} \;\; \mathbf{p^T}\mathcal{U}(\mathcal{A}\mathbf{y}) \\ &\text{s.t}: \;\; \mathbf{z^T y} = B \end{aligned} \tag{3.2.1}$$

We can solve the forward problem using the method of Lagrange multipliers as follows:

$$\mathcal{L} \;=\; \mathbf{p^T}\mathcal{U}(\mathcal{A}\mathbf{y}) - \lambda(\mathbf{z^T y} - B) \tag{3.2.2}$$

But before we proceed further, let us define $\widetilde{\mathcal{U}'}$ as an $s \times s$ diagonal matrix.

$$\widetilde{\mathcal{U}'}_{s\times s} \;\equiv\; \begin{bmatrix} \mathcal{U}'(\mathbf{A_1 y}) & 0 & \ldots & 0 \\ 0 & \mathcal{U}'(\mathbf{A_2 y}) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathcal{U}'(\mathbf{A_s y}) \end{bmatrix}, \tag{3.2.3}$$

Now, taking the partial derivatives with respect to $\mathbf{y}$ and with respect to λ gives us the following equations:

$$\nabla_{\mathbf{y}}\mathcal{L} \;=\; \mathbf{p^T}\widetilde{\mathcal{U}'}\mathcal{A} - \lambda \mathbf{z^T} \tag{3.2.4}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} \;= \mathbf{z^T y} - B \tag{3.2.5}$$

We can find the optimal portfolio, $\mathbf{y}$ by setting (3.2.4) and (3.2.5) equal to zero and solving the resulting equations.

There is an equivalent formulation for finding the optimal portfolio as well.

Supposing we say that the real world probabilities are given by $\mathbf{p}$ and the corresponding risk neutral probabilities are given by $\mathbf{q}$. We can find the risk-neutral probabilities $\mathbf{q}$ by solving $\mathcal{A}^T\mathbf{q} = \mathbf{z}(1+r)$, where r is the risk-free rate and all $q_i > 0$ and $\sum q_i = 1$. Since $\mathcal{A}$ denotes a complete market, $\mathcal{A}^T\mathbf{q} = \mathbf{z}(1+r)$, has a unique solution, i.e. there is a unique risk-neutral probability vector $\mathbf{q}$. This is also true when there are redundant assets in the market.[6]

Let $\mathcal{A}\mathbf{y} = \mathbf{w}$. We can now restate the optimization problem as follows:

$$\begin{aligned} &\max_{w} \ \mathbf{p^T}\mathcal{U}(\mathbf{w}) \\ &\text{s.t}: \ \mathbf{q^T w} = B(1+r) \end{aligned} \tag{3.2.6}$$

The Lagrangian in this case is given by:

$$\mathcal{L} \ = \ \mathbf{p^T}\mathcal{U}(\mathbf{w}) - \gamma(\mathbf{q^T w} - B(1+r)) \tag{3.2.7}$$

Taking the partial derivatives with respect to $\mathbf{w}$ and with respect to γ gives us the following equations:[7]

$$\nabla_{\mathbf{w}}\mathcal{L} \ = \ \mathbf{p^T}\widetilde{\mathcal{U}'} - \gamma\mathbf{q^T} \tag{3.2.8}$$

$$\frac{\partial\mathcal{L}}{\partial\gamma} \ = \mathbf{q^T w} - B(1+r) \tag{3.2.9}$$

We can find the optimal terminal wealth, $\mathbf{w}$ by setting (3.2.8) and (3.2.9) equal to zero and solving the resulting equations.

If we use this alternate formulation, it would be nice to establish the connection between (3.2.4) and (3.2.8). More importantly, we want to know the relation between λ and γ.

If we multiply the (3.2.8) by $\mathcal{A}$ on the right, we get the following:

$$\mathbf{p^T}\widetilde{\mathcal{U}'}\mathcal{A} - \gamma\mathbf{q^T}\mathcal{A}$$

[6]Proof of this fact can be found in Dana and Jeanblanc-Picqué (2003, p. 18).
[7]Note that (3.0.3) means

$$\widetilde{\mathcal{U}'}_{s\times s} \equiv \begin{bmatrix} \mathcal{U}'(\mathbf{A_1 y}) & 0 & \dots & 0 \\ 0 & \mathcal{U}'(\mathbf{A_2 y}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathcal{U}'(\mathbf{A_s y}) \end{bmatrix} = \begin{bmatrix} \mathcal{U}'(w_1) & 0 & \dots & 0 \\ 0 & \mathcal{U}'(w_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathcal{U}'(w_s) \end{bmatrix}$$

Since $\mathcal{A}^{\mathbf{T}}\mathbf{q} = \mathbf{z}(1+r)$ we can make the following substitution:

$$\mathbf{p}^{\mathbf{T}}\widetilde{\mathcal{U}'}\mathcal{A} - \gamma\mathbf{z}^{T}(1+r)$$

$$\text{Which is equivalent to (3.2.4), if we set } \lambda = \gamma(1+r) \tag{3.2.10}$$

We now set (3.2.8) and (3.2.9) to 0, and solve for the stationary points. Once we find the optimal $\mathbf{w}^*$, we can retrieve the optimal portfolio $\mathbf{y}^*$ as a solution of $\mathcal{A}\mathbf{y} = \mathbf{w}$. Since $\mathcal{A}$ is invertible, we have a unique solution given by $\mathbf{y}^* = \mathcal{A}^{-1}\mathbf{w}^*$.

3.2.1 Formulation of the Inverse Problem

To formulate our inverse problem in the case of complete markets, let us revisit the Lagrangian formulation in the previous section.

Setting (3.2.5) equal to 0 returns the budget constraint, i.e. checks for primal feasibility.

Setting (3.2.4) equal to zero gives us $\mathbf{p}^{\mathbf{T}}\widetilde{\mathcal{U}'}\mathcal{A} = \lambda\mathbf{z}^{\mathbf{T}}$.

Or equivalently, taking the transpose of both sides gives us: $\mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}\mathbf{p} = \lambda\mathbf{z}$. (3.2.11)

For the inverse problem, we know the optimal portfolio $\mathbf{y}^*$, and it is the probability vector $\mathbf{p}$ that is unknown. Since the optimal portfolio is given, we can calculate the value of $\widetilde{\mathcal{U}'}$ explicitly. We then have a problem with respect to $\mathbf{p}$ with the constraints given by (3.2.11). Since $\mathbf{p}$ is a probability vector, the feasible region is given by the constraint set:

$$\begin{aligned} &\mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}\mathbf{p} = \lambda\mathbf{z} \\ &\sum_{i=1}^{s} p_i = 1 \\ &p_i \geq 0, \quad \forall i = 1...s \end{aligned} \tag{3.2.12}$$

This formulation assumes that we know the value of the Lagrange multiplier associated with the budget constraint. In reality, however, we might not know the real value of λ associated with the optimal portfolio.[8] We can modify the feasible region as follows:

[8]Note that in case of complete markets, we always know the value of λ as it is unique. However, in case of incomplete markets, the value of λ is not unique. This is explained further in the Section 3.3.

$$\mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}\mathbf{h} = \mathbf{z} \qquad (3.2.13)$$
$$h_i \geq 0, \quad \forall i = 1...s$$

Where $h_i = \dfrac{p_i}{\lambda}$.

This is a valid substitution because λ is always greater than zero.[9] The reason for the above modification is explored further in Section 3.3.2.

Existence of the Probability Vector

In formulating the inverse problem, one question we want to answer is, "Is the given portfolio optimal/feasible for our scenario set and asset prices?" To answer this question, we want to know whether there exists a vector of probabilities that would make the given portfolio optimal/feasible. In the case of complete markets, our scenario matrix is square and invertible, and by extension, the transpose of the scenario matrix is also invertible. Also, since $\widetilde{\mathcal{U}'}$ is a diagonal matrix with strictly positive entries, it is also invertible. It follows from linear algebra, that the product of the two matrices will also be invertible, i.e. nonsingular.

It follows quite nicely from Appendix A.3, that if $\mathcal{A}^T\widetilde{\mathcal{U}'}$ is nonsingular, then (3.2.13) has a unique solution for every $\mathbf{z}$.

As a matter of fact, in case of complete markets, there is only one feasible point in the feasible region, and the solution is given by:

$$\mathbf{h} = (\widetilde{\mathcal{U}'})^{-1}(\mathcal{A}^T)^{-1}\mathbf{z} \qquad (3.2.14)$$

It is clear from (3.2.13) that all elements of $\mathbf{h}$ are greater than or equal to 0. However, $\mathbf{h}$ is not the zero vector since $\mathbf{z} > 0$.

3.2.2 Example 1: 3 assets and 3 scenarios

This example looks at the expected utility maximization problem with 3 assets and 3 scenarios. Asset 1 is the risk free asset with a return of 5% at the end of the period. The scenario matrix $\mathcal{A}$ and price vector $\mathbf{z}$ are defined as follows:

[9] Please refer to Appendix C for complete statement and proof.

$$\mathcal{A}_{3\times3} = \begin{bmatrix} 1.05 & 4 & 3 \\ 1.05 & 6 & 10 \\ 1.05 & 7 & 12 \end{bmatrix}, \ \mathbf{z}_{3\times1} = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix} \tag{3.2.15}$$

Let us assume that the investor's utility function is given by the following:

$$U(\mathbf{w}) = U(w_1, \ldots, w_m) = \begin{cases} \left[\frac{w_1^{1-\eta}-1}{1-\eta}, \frac{w_2^{1-\eta}-1}{1-\eta}, \ldots, \frac{w_m^{1-\eta}-1}{1-\eta}\right] & for\ \eta \neq 1, w_i > 0 \\ [\log(w_1), \log(w_2), \ldots, \log(w_m)] & for\ \eta = 1, w_i > 0 \\ -\infty & for\ \ w_i = 0 \end{cases} \tag{3.2.16}$$

Note 1. *We are only using the positive arm of the power utility. Negative values of wealth are not in our domain.*

Forward Problem

Supposing we say that the real world probabilities are given by $\mathbf{p}$ and the corresponding risk neutral probabilities are given by $\mathbf{q}$. We can find the risk-neutral probabilities $\mathbf{q}$ by solving $\mathcal{A}' * \mathbf{q} = \mathbf{z}(1+r)$, all $q_i \geq 0$ and $\sum q_i = 1$.

Let $\mathcal{A} * \mathbf{y} = \mathbf{w}$.

Given that the budget is B and the risk-free rate is r, we can formulate our utility maximization problem as follows:

$$\begin{aligned} \max_{\mathbf{w}} \quad & p_1U(w_1) + p_2U(w_2) + p_3U(w_3) \\ s.t: \quad & q_1w_1 + q_2w_2 + q_3w_3 = B(1+r) \end{aligned} \tag{3.2.17}$$

Note that in this case, we also have the non-negativity constraint associated with $\mathbf{w}$.

We can try solving this using the method of Lagrange multipliers:

$$\mathcal{L} = \quad p_1U(w_1) + p_2U(w_2) + p_3U(w_3) - \lambda(q_1w_1 + q_2w_2 + q_3w_3) \tag{3.2.18}$$

Taking the partials and setting them equal to zero gives us the following equations to solve:

$$\frac{\partial \mathcal{L}}{\partial w_i} = \quad \frac{p_i}{w_i^\eta} - q_i\lambda = 0 \quad \forall i,\ i = 1, 2, 3 \tag{3.2.19}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = q_1w_1 + q_2w_2 + q_3w_3 - B(1+r) = 0 \tag{3.2.20}$$

For $B \neq 0$, we can find an explicit solution to this problem.

In this case, let us assume $\eta = 2$, and $\mathbf{p} = [0.2859, 0.4295, 0.2846]'$. The risk-neutral probabilities associated with the scenarios and prices in (3.2.15) are given by $\mathbf{q} = [0.3833, 0.6000, 0.0167]'$. Let us assume that $B = 100$ and $r = 0.05$.

Substituting in the value of η into each of the equations (3.2.19) $\forall i$,(3.2.20) we get the following:

$$p_i - q_i \lambda w_i^2 = 0 \quad \forall i,\ i = 1, 2, 3 \tag{3.2.21}$$

$$q_1 w_1 + q_2 w_2 + q_3 w_3 - B(1+r) = 0 \tag{3.2.22}$$

Summing equations (3.2.21) over all i and simplifying gives us:

$$\lambda = \frac{1}{[q_1 w_1^2 + q_2 w_2^2 + q_3 w_3^2]} \tag{3.2.23}$$

Setting the equations in (3.2.21) equal to each other gives us the following:

$$w_1 = \pm\sqrt{\frac{p_1 q_2}{q_1 p_2}} w_2 \tag{3.2.24}$$

$$w_3 = \pm\sqrt{\frac{p_3 q_2}{q_3 p_2}} w_2 \tag{3.2.25}$$

Substituting (3.2.24) and (3.2.25) into (3.2.22) gives us:

$$w_2 = \frac{B(1+r)}{\left[\pm\sqrt{\frac{p_1 q_2}{q_1 p_2}} q_1 + q_2 \pm \sqrt{\frac{p_3 q_2}{q_3 p_2}} q_3\right]}. \tag{3.2.26}$$

Solving this system based on the given data yields 4 points, only one of which is in our domain.

$$\mathbf{w}_1 = \begin{bmatrix} 99.9149 \\ 97.8853 \\ 478.0845 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} -369.4227 \\ 361.9185 \\ 1767.6566 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 117.7927 \\ 115.3999 \\ -563.6280 \end{bmatrix}, \mathbf{w}_4 = \begin{bmatrix} -841.8181 \\ 824.7179 \\ -4028.0284 \end{bmatrix}$$

The values of λ corresponding to each of these 4 points is given by $[0.00007471, 0.00000547, 0.00005375, 0.00000105]$.

Evaluating the function at the only feasible point, $\mathbf{w_1}$, gives us the optimal value of 0.0078. We can therefore say that the optimal point is at $\mathbf{w}^* = \mathbf{w_1}$ with a corresponding value of λ given by $\lambda^* = 0.00007471$

Since the scenario matrix is square and invertible, we can solve for the optimal portfolio: $\mathbf{y}^* = \mathcal{A}^{-1}\mathbf{w}^*$.

In this case,

$$\mathbf{y}^* = \begin{bmatrix} -2563.4240 \\ 888.4845 \\ -254.1427 \end{bmatrix} \tag{3.2.27}$$

Inverse Problem

Now we want to see whether our formulation actually returns the initial probability vector. We use optimization toolbox in `MATLAB` and `CVX`, a package for specifying and solving convex programs (Grant & Boyd, 2011). The code is provided in Appendix D.1

Let us assume now that we know the optimal portfolio $\mathbf{y}^*$ as well as the information given in (3.2.15), but we do not know the set of probabilities $\mathbf{p}$ associated with the scenarios.

We can find the set of all probabilities by solving the system of linear equations given by (3.2.12). For the sake of simplicity we will minimize the least squares distance between the prior probability distribution and the probability associated with the optimal portfolio. However, as mentioned before, the feasible region only has one point.

For this example, let us use the optimal portfolio given by equation (3.2.27), and the prior probability distribution $\widetilde{\mathbf{p}} = [0.2859, 0.4295, 0.2846]'$. We should get the optimal probability vector $\mathbf{p}^* = \widetilde{\mathbf{p}}$.

The first derivative of the investor's utility function is given by the first derivative of equation (3.2.16):

$$\mathcal{U}'(\mathbf{w}) = \mathcal{U}'(w_1, \dots, w_m) = \begin{cases} \left[w_1^{-\eta}, w_2^{-\eta}, \dots, w_m^{-\eta}\right] & for\ \eta \neq 1 \\ \left[\frac{1}{w_1}, \frac{1}{w_2}, \dots, \frac{1}{w_m}\right] & for\ \eta = 1 \end{cases} \tag{3.2.28}$$

Let $\mathcal{A}^* = \mathbf{w}^*$, $\mathbf{w}^*$ as given above.

In this example, the $\widetilde{\mathcal{U}'}$ is given by:

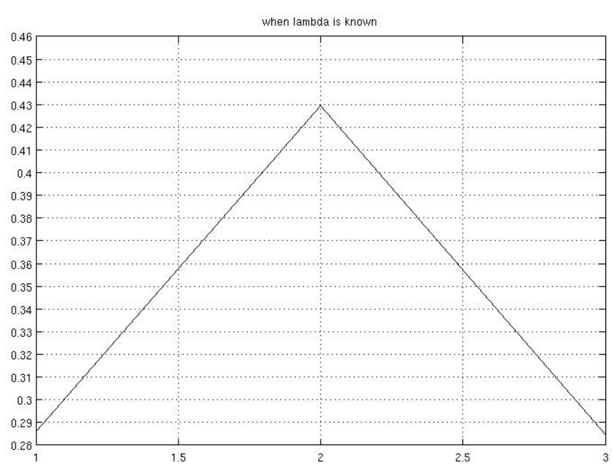
$$\widetilde{\mathcal{U}'} = \begin{bmatrix} w_1^{-2} & 0 & 0 \\ 0 & w_2^{-2} & 0 \\ 0 & 0 & w_3^{-2} \end{bmatrix}, \tag{3.2.29}$$

Using the value of λ^* given in the solution to the forward problem, we can formulate the feasible region for our inverse problem as stated in (3.2.12). Solving the least squares minimization problem in MATLAB gives us back the prior probability vector as expected.

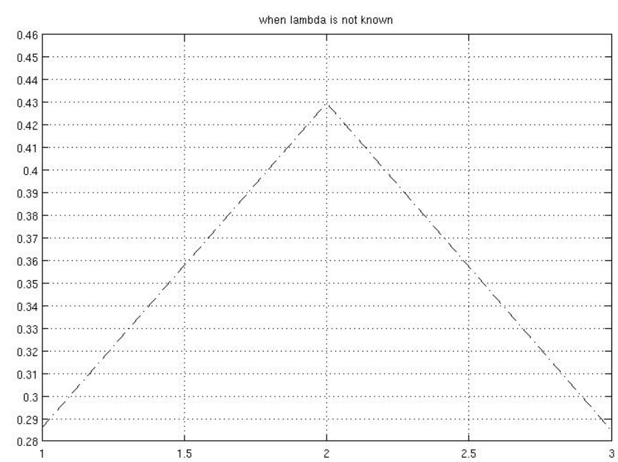
We can also retrieve the probabilities using the formulation given in (3.2.13). In this case, we get a slightly different probability vector. $\mathbf{p_{inv}} = [0.2859, 0.4297, 0.2844]'$. And solving the problem using the closed form solution in (3.2.14) also gives us $\mathbf{p_{inv}}$.

The slight difference in values can be attributed to the instability that is to be expected with matrices that have a large condition number. In our case, the matrix given by $\mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}$ has a condition number of 2220.19047, which indicates that the matrix $\mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}$ is nearly singular.

The graphs of the implied probabilities are given in figure 3.2



(a) formulation (3.2.12)



(b) formulation (3.2.13) and closed form solution (3.2.14)

Figure 3.2: implied probabilities for 3 asset case

As we can see, the graphs are practically identical, though numerically, there are slight differences.

3.3 Expected Utility Maximization in Incomplete Markets

A more interesting case for portfolio selection is when you have incomplete markets, i.e. when the rank of the scenario matrix $\mathcal{A} < s$, where s is the number of scenarios. When the market is not complete, our scenario matrix $\mathcal{A}$ is not invertible. However, there are still quite a few similarities between this case and the case above.

Firstly, the forward problem formulation remains unchanged. We can still solve for the optimal portfolio using the formulation in (3.2.1). There does exist an alternate formulation, similar to (3.2.6). The attainable payoff method in this case modifies the optimization problem as follows:

$$\begin{aligned} \max_{w} \ & \mathbf{p^T}\mathcal{U}(\mathbf{w}) \\ \text{s.t}: \ & \mathbf{q^T w} = B(1+r) \\ & \mathbf{V^T w} = \vec{0} \end{aligned} \tag{3.3.1}$$

Here $\mathbf{V}$ denotes an orthonormal basis for $Null(\mathcal{A}^T)$. It is important to note that we can find a state-price vector ψ_0 in incomplete markets by solving $\mathcal{A}^T\psi = \mathbf{z}$, all $\psi_i > 0$. However, in case of incomplete markets, if you have one solution, you have infinitely many solutions given by $(\psi_0 + t\mathbf{v})$, where $\mathbf{v} \in Null(\mathcal{A}^T)$. We can then find a risk-neutral vector by normalizing the state-price vector.

It is advisable to use the alternate formulation if the scenario matrix $\mathcal{A}$ is nearly singular (Further discussion on this in Section 3.3.3.)

The Lagrangian in this case is given by:

$$\mathcal{L} \ = \ \mathbf{p^T}\mathcal{U}(\mathbf{w}) - \alpha(\mathbf{q^T w} - B(1+r)) - \beta^T(\mathbf{V^T w}). \tag{3.3.2}$$

where α = Lagrange multiplier associated with the budget constraint,

β = Vector of Lagrange multipliers for constraints given by $\mathbf{V^T w} = 0$.

Taking the partial derivatives with respect to $\mathbf{w}$ and with respect to α and β gives us the following equations:

$$\nabla_{\mathbf{w}}\mathcal{L} \ = \ \mathbf{p^T}\widetilde{\mathcal{U}'} - \alpha\mathbf{q^T} - \beta^T(\mathbf{V^T}), \tag{3.3.3}$$

$$\frac{\partial \mathcal{L}}{\partial \alpha} = -\mathbf{q^T w} + B(1+r) \tag{3.3.4}$$

$$\nabla_\beta \mathcal{L} \;\; = -(\mathbf{V^T w}) \tag{3.3.5}$$

If we use this alternate formulation, we can establish the connection between (3.2.4) and (3.3.3) and between λ and α similar to the analysis for λ and γ.

If we multiply the (3.3.3) by $\mathcal{A}$ on the right, we get the following:

$$\mathbf{p^T}\widetilde{\mathcal{U}'}\mathcal{A} - \alpha\mathbf{q^T}\mathcal{A} - \beta^T\mathbf{V^T}\mathcal{A}$$

Since $\mathcal{A}^\mathbf{T}\mathbf{q} = \mathbf{z}(1+r)$ we can make the following substitution:

$$\mathbf{p^T}\widetilde{\mathcal{U}'}\mathcal{A} - \alpha\mathbf{z}^T(1+r) - \beta^T\underbrace{\mathbf{V^T}\mathcal{A}}_{=\,0},$$

which is equivalent to (3.2.4), if we set

$$\lambda = \alpha(1+r). \tag{3.3.6}$$

We now set (3.2.8) and (3.2.9) to 0, and solve for the stationary points. Once we find the optimal $\mathbf{w}^*$, we can retrieve the optimal portfolio $\mathbf{y}^*$ as a solution of $\mathcal{A}\mathbf{y} = \mathbf{w}$. The constraint $\mathbf{V^T w} = \vec{0}$ guarantees the existance of a solution, in that it restricts the possible vectors $\mathbf{w}$ to $Col(\mathcal{A})$.

3.3.1 Formulation of the Inverse Problem

The formulation of the inverse problem in incomplete markets is identical to the formulation of the problem in case of complete markets. However, there are some key differences that emerge as a result of the scenario matrix being neither square nor invertible.

The formulation of the inverse problem follows the exact same steps as in the case of complete markets, the only difference being that the scenario matrix $\mathbf{A}$ is not a full-row-rank matrix.

The matrix of first derivatives $\widetilde{\mathcal{U}'}$ is still an $s \times s$ matrix defined by (3.2.3).

The feasible region is, once again, given by (3.2.12), when we know the value of λ and by (3.2.13) when the value of λ is unknown. The feasible region in this case has more than one point, so to retrieve a meaningful solution, we need to impose certain conditions on the objective function.

Merkulovich and Rosen (1998) suggest finding solutions that optimize various solution functionals, subject to data constraints. In terms of financial applications, as is the case here, they also suggest criteria that can be imposed on the feasible set that could provide useful information for risk managers. For the framework proposed in their paper, and by extension, the framework proposed in this thesis, they suggest finding solutions that are (a) Closest to a prior distribution, (b) Maximum deviations from a prior, and (c)Minimum and maximum parameter values (i.e. bounds on the probabilities of a given scenario occurring).

In this thesis, we will attempt to find closest to prior solutions by minimizing the least squares distance between the prior and the implied probabilities, and the bounds on the probabilities of scenarios, for example, the probability of extreme events in our scenario set occurring as implied by the choice of "optimal" portfolio.

Existence of the Probability Vector

In case of incomplete markets, our scenario matrix is not invertible and neither is the transpose of the scenario matrix, which means that verifying the existence of the probability vector in this case is not as easy.

It is clear from principles of linear algebra, that if there exists one solution to 3.2.13, then there exist infinitely many solutions given by: $\mathcal{A}^T\widetilde{\mathcal{U}'}(\mathbf{h}+t\mathbf{v}) = \mathbf{z}$, where $t \in \mathbb{R}$ and all $\mathbf{v} \in Null(\mathcal{A}^T\widetilde{\mathcal{U}'})$

If there exists a vector $\mathbf{y}$ such that, $0 \leq \mathbf{y}^T\mathcal{A}^T\widetilde{\mathcal{U}'}$ and $\mathbf{y}^T\mathbf{z} > 0$, then there is no solution. (By Farkas' Lemma in Appendix D.1).

For instance, if we take a price vector $\mathbf{z}$ such that $\mathbf{z} \in Null(\widetilde{\mathcal{U}'}\mathcal{A})$, then $\mathcal{A}^T\widetilde{\mathcal{U}'}\mathbf{h} = \mathbf{z} \Rightarrow 0 = \mathbf{z}^T\mathcal{A}^T\widetilde{\mathcal{U}'}\mathbf{h} = \mathbf{z}^T\mathbf{z} > 0$, which is inconsistent, i.e., there is no solution.

3.3.2 Example 2: 9 assets and 21 scenarios

In this section, we look at the behaviour of the inverse optimization problem in case of incomplete markets. We use real-world data using puts on the S&P 500 Index expiring on June 11, 2011. The data was accessed on April 11th, 2011 from the CBOE website. We also use the risk free rate of 0.1% on the 1 month T-bill. Let the initial budget B be \$1000.[10].

Since we only have one underlying asset, we use a grid to generate the scenario set. We consider a range of strike prices for the puts, and possible scenarios for the index prices given in Table 3.1.

[10]The yield on the 1 month US T-bill on April 11th was 0.1% and for the 3 month US T-bill, the yield was 0.3%. Data available at: `http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2011`

Strike Prices K							
950	1100	1225	1340	1345	1450	1550	1650

Index Prices S										
800	850	900	950	1000	1050	1100	1150	1200	1250	1300
	1350	1400	1450	1500	1550	1600	1650	1700	1750	1800

Table 3.1: Table of Strike Prices (assets) and Stock Prices (scenarios)

The complete set of scenarios used is given in Table 3.2.

$\mathbf{p_0}$	Put Values $(K-S)^+$								
0	1.0002	150	300	425	540	545	650	750	850
0	1.0002	100	250	375	490	495	600	700	800
0	1.0002	50	200	325	440	445	550	650	750
0	1.0002	0	150	275	390	395	500	600	700
0	1.0002	0	100	225	340	345	450	550	650
0.0002	1.0002	0	50	175	290	295	400	500	600
0.0016	1.0002	0	0	125	240	245	350	450	550
0.0153	1.0002	0	0	75	190	195	300	400	500
0.0638	1.0002	0	0	25	140	145	250	350	450
0.1462	1.0002	0	0	0	90	95	200	300	400
0.2313	1.0002	0	0	0	40	45	150	250	350
0.2414	1.0002	0	0	0	0	0	100	200	300
0.1698	1.0002	0	0	0	0	0	50	150	250
0.0831	1.0002	0	0	0	0	0	0	100	200
0.0329	1.0002	0	0	0	0	0	0	50	150
0.0110	1.0002	0	0	0	0	0	0	0	100
0.0026	1.0002	0	0	0	0	0	0	0	50
0.0008	1.0002	0	0	0	0	0	0	0	0
0	1.0002	0	0	0	0	0	0	0	0
0	1.0002	0	0	0	0	0	0	0	0
0	1.0002	0	0	0	0	0	0	0	0

Asset Prices at Time 0								
1.0000	0.0060	0.4535	3.8388	25.1473	26.6550	110.5743	210.5450	310.5197

Table 3.2: Probabilities, Scenario Matrix and Asset Prices

We want to first calculate the optimal portfolio for this scenario set. To do that, we will need the probability of each scenario occurring. We use the historical mean and variance of returns on the S&P500 Index to simulate 100000 MC realizations of index values, and then use the relative frequency to calculate the probability of each scenario occurring. For sake of simplicity, we assume that the returns are normally distributed. However, we could simulate probabilities under non-normal distributions as well and our proposed framework will work. The probability of each scenario is also given in Table 3.2. We can now calculate the optimal portfolio using formulation (3.3.1). We use `MATLAB` and `CVX`(Grant & Boyd, 2011) to solve the forward problem. Complete code is given in Appendix E.1.

The optimal portfolio and optimal Lagrange multiplier for the inputs in Table3.2 is given by:

$$\mathbf{y}^* = \begin{bmatrix} 2211.2875 \\ 19.1487 \\ 2.4105 \\ -7.9466 \\ -86.9864 \\ 92.0823 \\ 123.5647 \\ -242.6240 \\ 115.8418 \end{bmatrix}, \ \lambda^* = 9.0852 \times 10^{-7} \tag{3.3.7}$$

Now that we have an optimal portfolio, we can check to see if the inverse problem retrieves the probabilities we used to calculate the optimal portfolio if we use p_0 as the prior vector.

We will minimize the least squares distance between the prior vector and the implied probability vector, first using the value of λ^* from (3.3.7) and the formulation of the feasible region given in (3.2.12), and then using the formulation of the feasible region given in 3.2.13. The full formulation of the optimization problem and the results are given below.

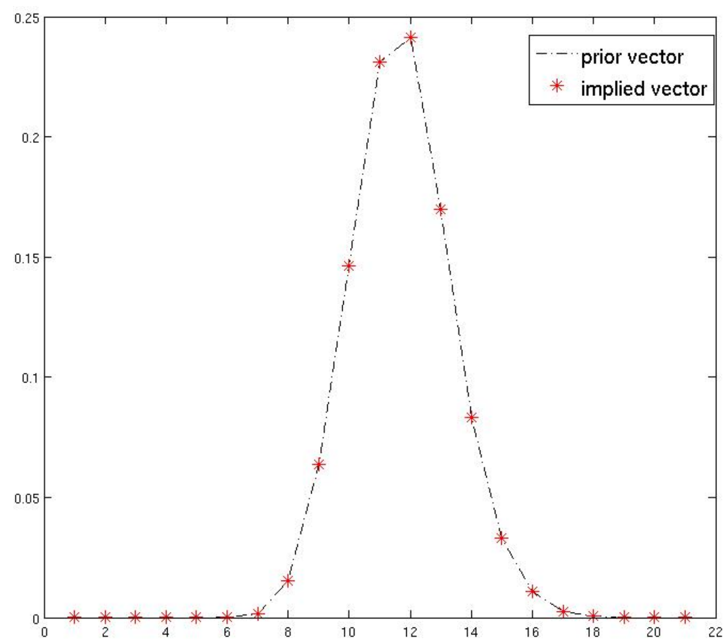
$$\begin{aligned} &\min_{p} \ ||\mathbf{p_o} - \mathbf{p}|| \\ &\text{s.t} : \mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}}'\mathbf{p} = \lambda\mathbf{z} \\ &\qquad \sum_{i=1}^{s} p_i = 1 \\ &\qquad p_i \geq 0, \ \ \forall i = 1...s \end{aligned} \tag{3.3.8}$$

When we try to solve this problem using the exact formulation above using `CVX` and `CPLEX`, `MATLAB` says the problem is infeasible, and there could be two reasons why that is the case. First, our constraints are inconsistent, or second, it is a numerical issue. On further investigation, we can see that the matrix formed by the left-hand side of set of constraints is full row-rank. We know from linear algebra that inconsistencies occur if the left-hand sides of the equations in a system are linearly dependent, and the constant terms do not satisfy the dependence relation. A system of equations whose left-hand sides are linearly independent is always consistent. We can, therefore, conclude that the problem is numerical. If we modify the problem as follows,

$$\begin{aligned} &\min_{p} \;\; ||\mathbf{p_o} - \mathbf{p}|| \\ &\text{s.t} : \mathcal{A}^{\mathbf{T}} \widetilde{\mathcal{U}'} \mathbf{p} \leq \lambda \mathbf{z} \\ &\qquad \sum_{i=1}^{s} p_i = 1 \\ &\qquad p_i \geq 0, \;\; \forall i = 1 ... s, \end{aligned} \tag{3.3.9}$$

we get the results in Figure 3.3.

$\mathbf{p_0}$	$\mathbf{p_{implied_{cvx}}}$	$\mathbf{p_{implied_{cplex}}}$
0	0.0000	0.0000
0	0.0000	0.0000
0	0.0000	0.0000
0	0.0000	0.0000
0	0.0000	0.0000
0.0002	0.0002	0.0002
0.0016	0.0016	0.0016
0.0153	0.0153	0.0153
0.0638	0.0638	0.0638
0.1462	0.1462	0.1462
0.2313	0.2326	0.2326
0.2414	0.2414	0.2414
0.1698	0.1698	0.1698
0.0831	0.0831	0.0831
0.0329	0.0329	0.0329
0.0110	0.0110	0.0110
0.0026	0.0026	0.0029
0.0008	0.0008	0.0000
0	0.0000	0.0000
0	0.0000	0.0000
0	0.0000	0.0000



Graphs for the prior and the implied probability

Figure 3.3: Prior probabilities and implied probabilities retrieved using (3.3.9) and `CVX` and `CPLEX`

As we can see, the implied probabilities are identical to the prior vector to 4 decimal places. In (3.3.9), we changed the hard equality constraints to soft inequality constraints.

In doing so, we increased the size of the feasible region. We want to find out next, within what tolerance of the right hand side does the actual solution lie. To do this, we modify the problem as following:

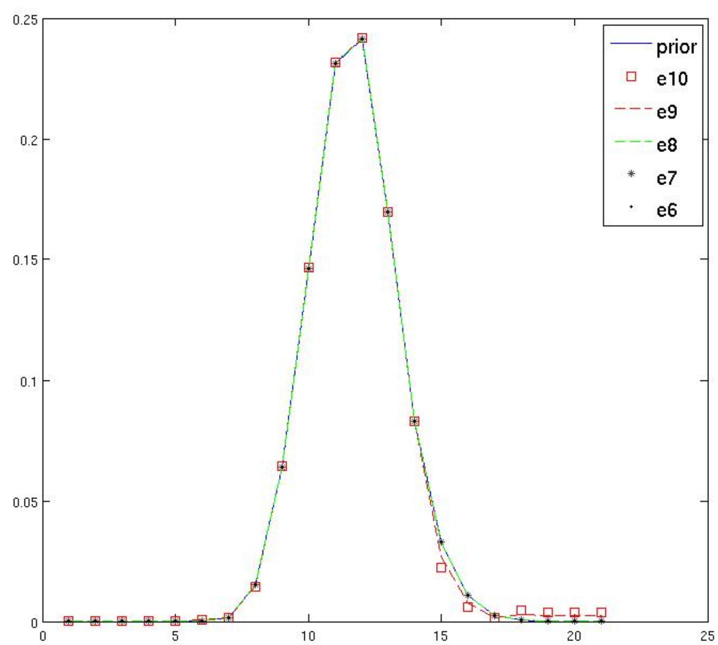
$$\begin{aligned}
&\min_{p} \;\; ||\mathbf{p_o} - \mathbf{p}|| \\
&\text{s.t} : \lambda\mathbf{z} - \epsilon \leq \mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}}'\mathbf{p} \leq \lambda\mathbf{z} + \epsilon \\
&\qquad \sum_{i=1}^{s} p_i = 1 \\
&\qquad p_i \geq 0, \;\; \forall i = 1...s
\end{aligned} \tag{3.3.10}$$

We vary the value of ϵ from 10^{-10} to 10^{-2}, increasing it by a factor of 10 for each iteration. We stop as soon as the implied probabilities equal the prior probabilities. We will use `CVX`(Grant & Boyd, 2011) to solve the above problem. Figure 3.4 shows the implied probabilities and the corresponding values of ϵ.

We can see from the table and the graph in Figure 3.4 that ϵ_6 is the one that returns the prior vector. What we might want to ask now is, ”What does this value of ϵ mean?”. What this means is that the true solution lies between a very narrow band of error tolerance given by 2×10^{-6}. As a matter of fact, we can go further, and say that the true solution lies between $\lambda\mathbf{z} \leq \mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}}'\mathbf{p} \leq \lambda\mathbf{z} + 10^{-6}$, because the initial feasible region, which had the entire area less than $\lambda\mathbf{z}$ already contained the prior vector in it.

However, this is just an arbitrary tolerance that we picked. What if we were to use a vector of tolerances set as a percentage of the right hand side? We start off with setting the tolerance as 0.1% of the right hand side and we increase the tolerance by 0.1% at each iteration till we increase the tolerance to 0.9% of the right hand side. The results are given in Table 3.3.

p_0	$\epsilon_{10} = 10^{-10}$	$\epsilon_9 = 10^{-9}$	$\epsilon_8 = 10^{-8}$	$\epsilon_7 = 10^{-7}$	$\epsilon_6 = 10^{-6}$
0	0.0001	0.0001	0.0000	0.0000	0.0000
0	0.0001	0.0001	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0001	0.0001	0.0002	0.0001	0.0000
0.0002	0.0006	0.0005	0.0003	0.0002	0.0002
0.0016	0.0016	0.0015	0.0016	0.0016	0.0016
0.0153	0.0142	0.0147	0.0153	0.0153	0.0153
0.0638	0.0642	0.0640	0.0638	0.0638	0.0638
0.1462	0.1469	0.1466	0.1462	0.1462	0.1462
0.2313	0.2315	0.2316	0.2313	0.2313	0.2313
0.2414	0.2418	0.2419	0.2418	0.2415	0.2414
0.1698	0.1699	0.1698	0.1699	0.1698	0.1698
0.0831	0.0830	0.0829	0.0831	0.0831	0.0831
0.0329	0.0222	0.0266	0.0327	0.0329	0.0329
0.0110	0.0059	0.0077	0.0108	0.0110	0.0110
0.0026	0.0016	0.0014	0.0024	0.0026	0.0026
0.0008	0.0046	0.0031	0.0006	0.0008	0.0008
0	0.0038	0.0024	0.0000	0.0000	0.0000
0	0.0038	0.0024	0.0000	0.0000	0.0000
0	0.0038	0.0024	0.0000	0.0000	0.0000



Graphs for the prior and the implied probabilities

Figure 3.4: Prior probabilities and implied probabilities retrieved using (3.3.10) and CVX

$\mathbf{p_0}$	Tolerance Level as a Percentage of RHS								
	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.7%	0.8%	0.9%
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002	0.0002
0.0016	0.0017	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016
0.0153	0.0153	0.0153	0.0153	0.0153	0.0153	0.0153	0.0153	0.0153	0.0153
0.0638	0.0638	0.0638	0.0638	0.0638	0.0638	0.0638	0.0638	0.0638	0.0638
0.1462	0.1462	0.1462	0.1462	0.1462	0.1462	0.1462	0.1462	0.1462	0.1462
0.2313	0.2313	0.2313	0.2313	0.2313	0.2313	0.2313	0.2313	0.2313	0.2313
0.2414	0.2414	0.2414	0.2414	0.2414	0.2414	0.2414	0.2414	0.2414	0.2414
0.1698	0.1698	0.1698	0.1698	0.1698	0.1698	0.1698	0.1698	0.1698	0.1698
0.0831	0.0831	0.0831	0.0831	0.0831	0.0831	0.0831	0.0831	0.0831	0.0831
0.0329	0.0329	0.0329	0.0329	0.0329	0.0329	0.0329	0.0329	0.0329	0.0329
0.0110	0.0110	0.0110	0.0110	0.0110	0.0110	0.0110	0.0110	0.0110	0.0110
0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026	0.0026
0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 3.3: Implied probabilities using (3.3.10) with ϵ as a function of RHS

Implied probabilities in the last column of Table 3.3 appears to be the closest to the prior vector p_0, to 4 decimal places. We will use the 0.1% and the 0.8% columns to graph the implied probabilities given in Figure 3.5. The graphs show that the probabilities are virtually identical.

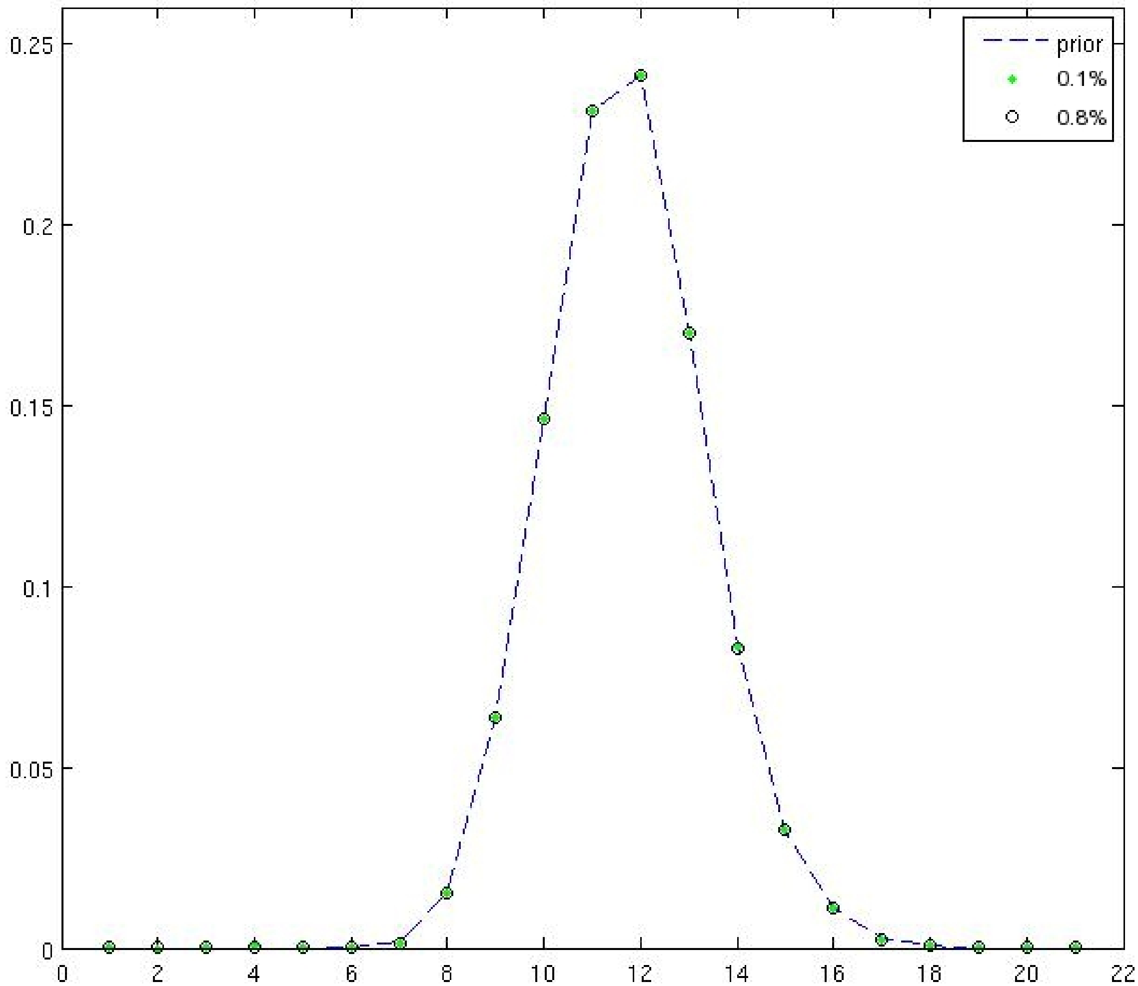


Figure 3.5: Prior vector and implied probability vectors.

Another instance that we would like to consider is when the Lagrange multiplier is unknown. In this case, we use the following formulation.

$$\begin{aligned} &\min_{h} \ ||\mathbf{p_0} - \mathbf{h}|| \\ &\text{s.t} : \mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}}'\mathbf{h} = \mathbf{z} \\ &\qquad h_i \geq 0, \ \ \forall i = 1...s \end{aligned} \tag{3.3.11}$$

Where $h_i = \dfrac{p_i}{\lambda}$.

This formulation is slightly different because we are minimizing the distance between $\mathbf{h} = \frac{1}{\lambda}\mathbf{p}$ and $\mathbf{p_0}$ rather than between $\mathbf{p}$ and $\mathbf{p_0}$. Once we find a suitable vector $\mathbf{h}$, we normalize it to get the vector of implied probabilities. The results are given in Figure 3.6

$\mathbf{p_0}$	$\mathbf{p_{implied_{cvx}}}$	$\mathbf{p_{implied_{cplex}}}$
0	0.0000	0.0000
0	0.0000	0.0000
0	0.0001	0.0001
0	0.0000	0.0000
0	0.0000	0.0000
0.0002	0.0000	0.0000
0.0016	0.0000	0.0000
0.0153	0.0222	0.0222
0.0638	0.0758	0.0758
0.1462	0.1326	0.1326
0.2313	0.2468	0.2468
0.2414	0.2592	0.2592
0.1698	0.1580	0.1580
0.0831	0.1001	0.1001
0.0329	0.0026	0.0026
0.0110	0.0009	0.0009
0.0026	0.0013	0.0013
0.0008	0.0001	0.0001
0	0.0001	0.0001
0	0.0001	0.0001
0	0.0001	0.0001

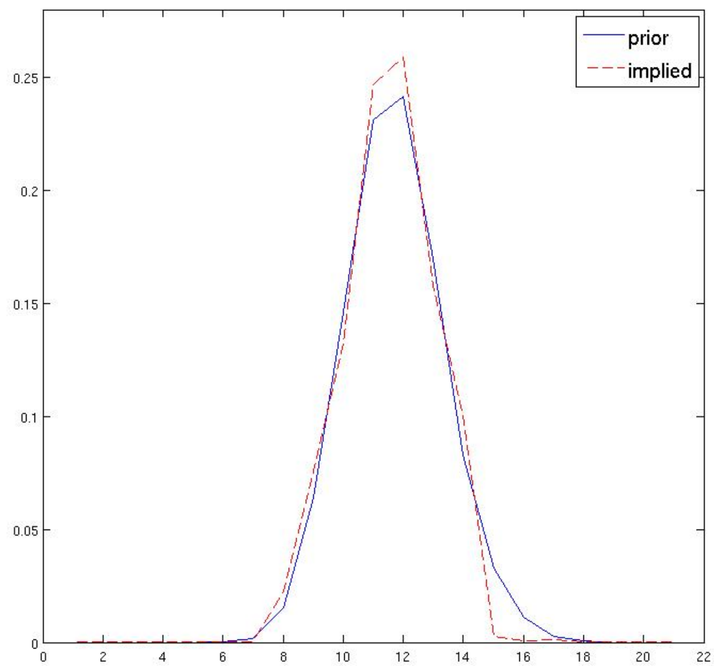


Figure 3.6: Prior probabilities and implied probabilities retrieved using (3.3.11) and `CVX` and `CPLEX`

Once again, `CVX` and `CPLEX` give identical solutions, though the implied probabilities are not identical to the prior probabilities.

Note that $h_i = \frac{1}{\lambda}p_i$. If we sum both sides over all i we get $\sum_{i=1}^{s} h_i = \frac{1}{\lambda}\sum_{i=1}^{s} p_i$. $\sum_{i=1}^{s} p_i = 1$ always, by definition of $\mathbf{p}$. So, we have,$\sum_{i=1}^{s} h_i = \frac{1}{\lambda}$. Which means we can approximate λ as $\lambda = \frac{1}{\sum_{i=1}^{s} h_i}$.

In theory, all values of $\mathbf{h}$ should give us the same value of λ. However, if the system is ill-conditioned, as it is in our case, different values of $\mathbf{h}$ give us different values of λ. The difference, however, is on the order of magnitude of 10^-7, i.e. value of λ is correct to 6 decimal places.

3.3.3 The Condition Number

The above example displays a certain instability in the system, which is to be expected in inverse problems (Engl & Kügler, 2005; Aster et al., 2005). One of the ways to quantify the stability of a system or a problem is to look at the condition number of the matrix of constraints, which in the case of the forward problem formulation (3.2.1) is given by $\mathcal{A}$ and in case of all the inverse problem formulations is given by $\mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}$.

The condition number of a matrix A is denoted by $\kappa(A)$ and is the ratio of the largest singular value to the smallest singular value. We can find the singular values of a matrix by the singular value decomposition (Appendix A.4). For an $m \times n$ matrix A, $\kappa(A) = ||A||||A^{+}||$, where $||A^{+}||$ is the Moore-Penrose (MP) inverse of A as given in A.4.3. For our purposes, $||\centerdot||$ denotes the spectral norm of a matrix. In this thesis, we will be referring to the spectral norm of a matrix when we use the word 'norm'. Please refer to Appendix A.5 for a formal definition of an induced matrix norm. For a square matrix, the condition number is the product of the matrix norm and the norm of the inverse.

A matrix with a condition number close to 1 is said to be well-conditioned, and matrices with a large condition number are said to be ill-conditioned. Though is hard to say what constitutes a *large* condition number, the general rule of thumb is that if $\kappa(A) \sim 10^k$, then we can expect to lose at most k digits of precision in solving the system $A\mathbf{x} = \mathbf{b}$ (Cheney & Kincaid, 2007). It is important to note that unitary matrices have condition number of 1, which means that if we can use a unitary matrix instead of an ill-conditioned one, our solution is more stable.

If we go back to formulations (3.2.1) and (3.3.1), it is evident that if $\mathcal{A}$ has a large condition number, (3.3.1) is a more stable formulation. With (3.3.1) to retrieve the optimal portfolio, we only need to concern ourselves with $\mathcal{A}$ twice: once to get the risk-neutral probability and again to get the optimal portfolio allocation (by volume) vector $\mathbf{y}^*$. With (3.2.1), we have to solve for the vector $\mathbf{y}$ at each iteration to determine if it satisfies the budget constraint, which is problematic if the system is ill-conditioned because we cannot guarantee optimality.

For our inverse problem formulations, we regularize the problem by first using Tikhonov regularization.

In Tikhonov regularization, we consider all solutions with $||A\mathbf{x} - \mathbf{b}||^2 \leq \delta$, and select the one that minimizes the norm of $\mathbf{x}$. Note that as δ increases, the set of feasible models expands, and the minimum value of of the objective function decreases. It is also possible to get the same set of solutions by considering problems of the form:

$$\{\min ||A\mathbf{x} - \mathbf{b}||^2 \ : \ ||\mathbf{x}||^2 \leq \epsilon\}.$$

A third option is to consider the damped least squares problem $\min ||A\mathbf{x} - \mathbf{b}||_2^2 + \alpha ||\mathbf{x}||_2^2$. We refer the reader to Aster et al. (2005), chapter 5 for a deeper understanding of the Tikhonov regularization technique as it applies to discrete inverse problems.

In our case, we will use the first formulation, where we minimize the least-squares distance between the implied probability and the prior probabilities, where the error tolerance δ is set as a fixed constant which we called ϵ, and then again with the error tolerance set as a percentage of the RHS vector.

3.3.4 Example 2 Revisited: A Different Prior Distribution

Let us consider the optimal portfolio allocation (by volume), $\mathbf{y}^*$, and the scenario matrix $\mathcal{A}$ from Section 3.3.2. However, now let the prior vector be given by p_{ht}. The difference between the original probability vector $\mathbf{p}_0$ used to calculate $\mathbf{y}^*$ and p_{ht} can be seen clearly in Figure 3.7: The new prior distribution has significantly heavier tails than the original distribution. We want to see what the implied views from our optimal portfolio are when we minimize the distance to this new prior with heavier tails.

$\mathbf{p_0}$	$\mathbf{p_{ht}}$
0	0.0087
0	0.0142
0	0.0192
0	0.0311
0	0.0386
0.0002	0.0540
0.0016	0.0691
0.0153	0.0783
0.0638	0.0826
0.1462	0.0863
0.2313	0.0766
0.2414	0.0821
0.1698	0.0726
0.0831	0.0690
0.0329	0.0555
0.0110	0.0507
0.0026	0.0402
0.0008	0.0314
0	0.0231
0	0.0166
0	0

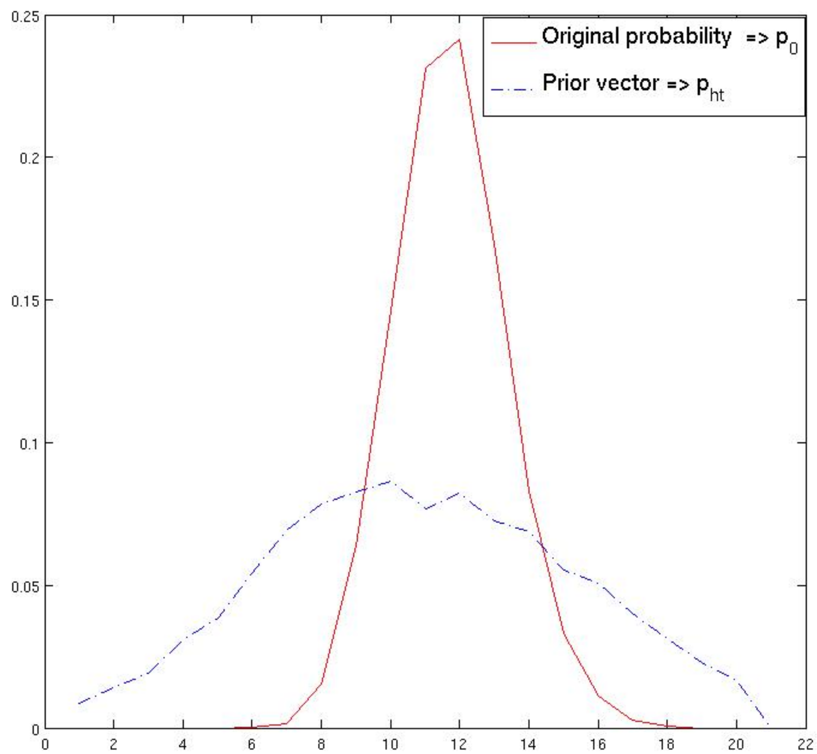


Figure 3.7: Original probability vector $\mathbf{p}_0$ and new prior vector $\mathbf{p}_{ht}$

We want to see what the probability distribution implied by the portfolio $\mathbf{y}^*$ tells us about the views on certain scenarios, if we minimize the distance to this distribution with heavier tails.

To do that, we will first find a feasible $\mathbf{h}$ using the feasible region given by (3.2.13) and a constant objective function. Then, we can find the approximate value of λ, let us call it,

$\tilde{\lambda}$ as $\tilde{\lambda} = \dfrac{1}{\sum_{i=1}^{s} h_i}$. Once we have the value of $\tilde{\lambda}$, we use the following formulation:

$$\begin{aligned}
&\min_{p} \ ||\mathbf{p_{ht}} - \mathbf{p}|| \\
&\text{s.t} : \tilde{\lambda}\mathbf{z} - \epsilon \leq \mathcal{A}^{\mathbf{T}}\tilde{\mathcal{U}}'\mathbf{p} \leq \tilde{\lambda}\mathbf{z} \\
&\quad \sum_{i=1}^{s} p_i = 1 \\
&\quad p_i \geq 0, \ \ \forall i = 1...s
\end{aligned} \tag{3.3.12}$$

Here, ϵ is set as a percentage of $\tilde{\lambda}\mathbf{z}$. We start off with a value of $\epsilon = 0.001\tilde{\lambda}\mathbf{z}$ and we solve (3.3.12) successively, increasing the value of ϵ at each iteration till the feasible region is equivalent to the one given in (3.3.9). The increments in ϵ and the corresponding implied probability vector are given in Figure 3.8.

As we can see, smaller values of ϵ give implied probabilities closer to the original probability, and larger values give implied probabilities closer to the prior. However, even when the feasible region is given by (3.3.9), we do not see heavier left tails. It appears to be that the current portfolio implies that the index prices will drop below 1100 in the next two months.

Code for this implementation can be found in Appendix E.2

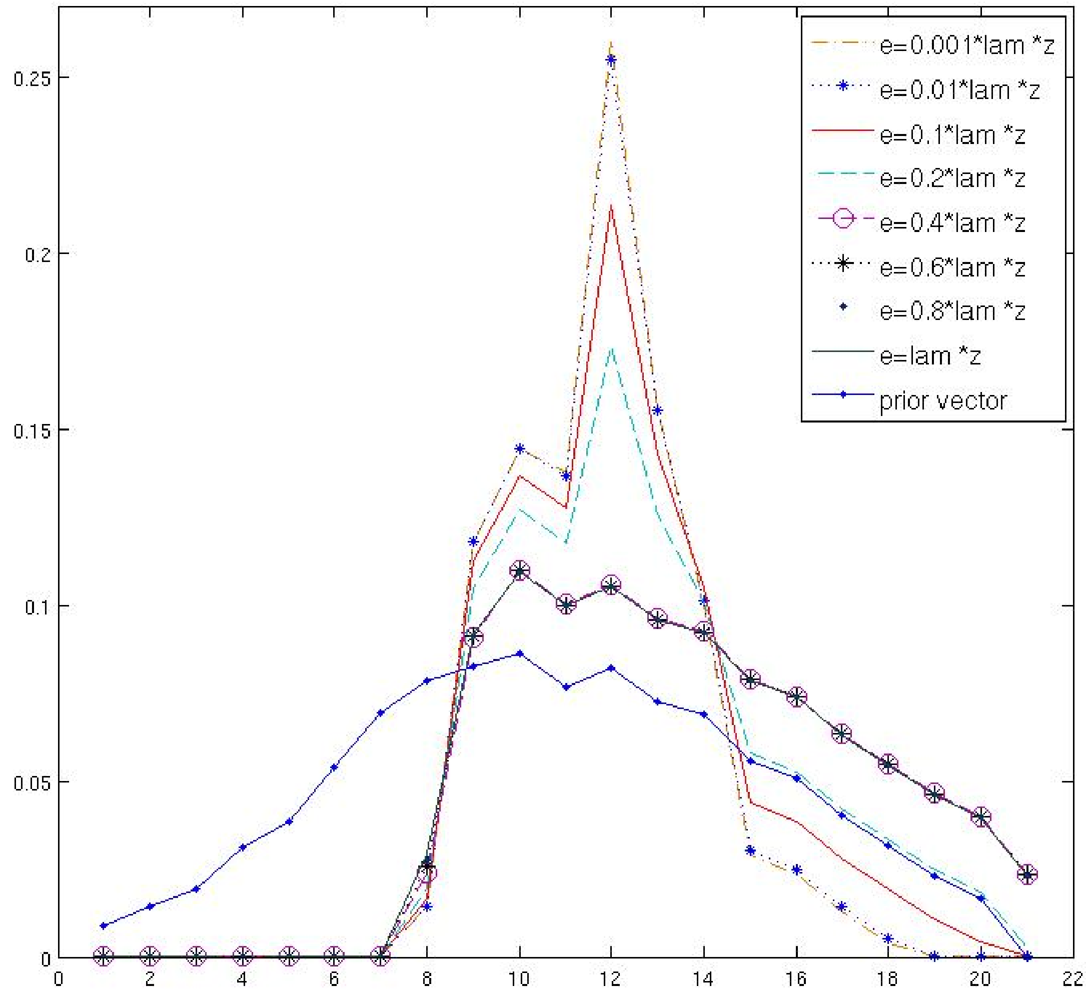


Figure 3.8: Prior probabilities and implied probabilities retrieved using (3.3.12) and `CVX`

3.3.5 Example 2 Revisited: Minimum or Maximum Implied Probability of an Event

We can also use the inverse problem formulation to get a maximum or a minimum probability of an event occurring.

Consider the inputs in Table 3.2. We can use the optimal portfolio allocation vector $\mathbf{y}^*$ to get the maximum and minimum probabilities of each of the scenarios, ω_i, occurring. We can calculate the value of $\widetilde{\lambda}$ as in the previous section.

We calculate the maximum and minimum probability of ω_i occurring as follows:

We define a $s \times 1$ vector of coefficients $\mathbf{c}_i$ such that all elements are zero, except for the

i^{th} element. Let $\epsilon = 0.01\,\widetilde{\lambda}\,\mathbf{z}$.

The maximum probability of ω_i as implied by our portfolio is then calculated using the following linear program:

$$\begin{aligned} \max_{p}\ & \mathbf{c}_i^T\,\mathbf{p} \\ \text{s.t}: & \widetilde{\lambda}\mathbf{z} - \epsilon \leq \mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}}'\mathbf{p} \leq \widetilde{\lambda}\mathbf{z} \\ & \sum_{i=1}^{s} p_i = 1 \\ & p_i \geq 0, \ \ \forall i = 1...s \end{aligned} \tag{3.3.13}$$

The minimum probability of ω_i as implied by our portfolio can be calculated using the exact same formulation as 3.3.13, except we minimize the objective function instead of maximizing.

We calculated the maximum and minimum probabilities of each of the scenarios using code in Appendix E.2. The results are given in Figure 3.9

$\mathbf{p_{min}}$	$\mathbf{p_{max}}$
0	0.0002
0	0.0002
0	0.0001
0	0.0001
0	0.0004
0	0.0021
0	0.0106
0	0.0369
0	0.1896
0	0.3423
0.0590	0.3681
0.1847	0.3466
0	0.3237
0	0.2201
0	0.0717
0	0.0706
0	0.0715
0	0.0763
0	0.0763
0	0.0763
0	0.0763

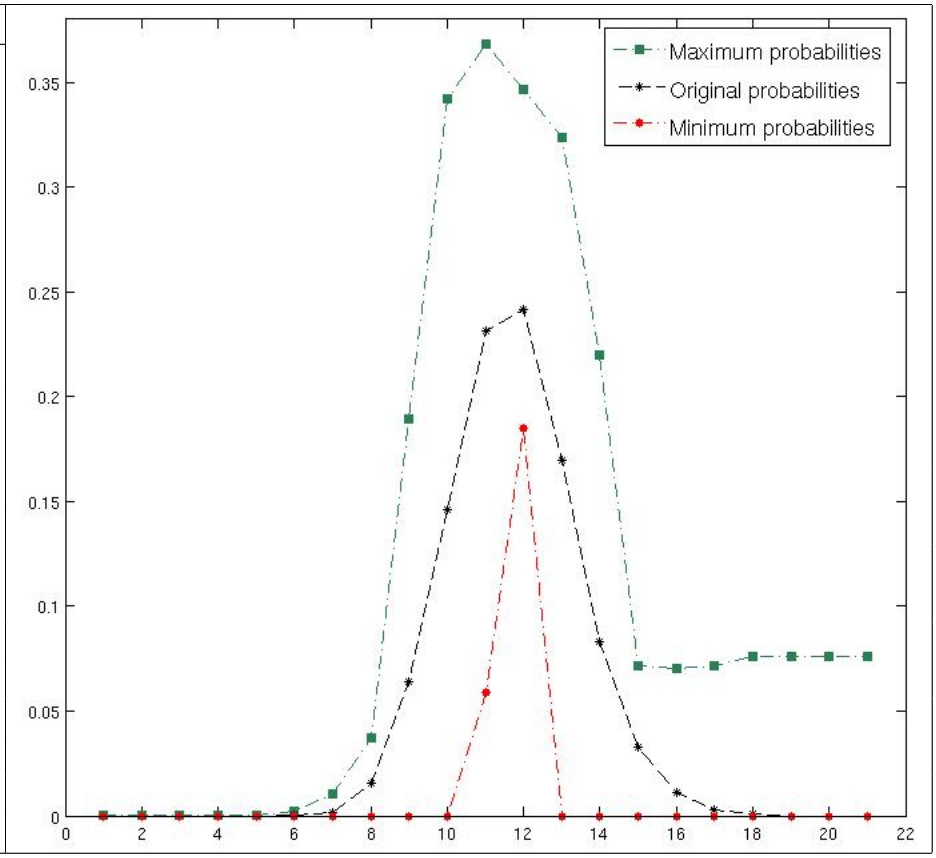


Figure 3.9: Minimum and maximum probabilities of each scenario occurring

3.4 Inverse Problems in the Mean-Variance Setting

Supposing we know that the optimal portfolio was chosen using mean-variance portfolio analysis. The investor's objective and constraint set is given by (2.2.3) and the set of inputs is given by (3.0.1) and (3.0.4).

Let,$\bar{A}_i$ denote the i^{th} row of $\bar{\mathcal{A}}$, μ_i denote the expected return of asset i, and σ_{ij} denote the covariance between asset i and asset j. We can translate the above inputs into expected returns and the variance-covariance matrix of the assets as follows:

$$\mu_{n\times 1} = \bar{\mathcal{A}}^T \mathbf{p} \quad \text{or} \quad \mu_i = \bar{A}_i^T \mathbf{p} \tag{3.4.1}$$

$$\sigma_{ij} = \sum_{k=1}^{s} (p_k \cdot \bar{a}_{ki} \cdot \bar{a}_{kj}) - \mu_i \cdot \mu_j \tag{3.4.2}$$

The optimality conditions (Best, 2010) for (2.2.3), in partitioned form are:

$$\eta \begin{bmatrix} \mu \\ R \end{bmatrix} - \begin{bmatrix} \Sigma & 0 \\ 0^T & 0 \end{bmatrix} \begin{bmatrix} x \\ x_{n+1} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{l}_n \\ 1 \end{bmatrix} \tag{3.4.3}$$

(3.4.3) gives us:

$$\lambda = \eta r \tag{3.4.4}$$

$$\Sigma x = \eta(\mu - R\mathbf{l}_n) \tag{3.4.5}$$

If we are given the optimal portfolio $\begin{bmatrix} x^* \\ 1 - x^* \end{bmatrix}$, and it is the probability vector $\mathbf{p}$ that is unknown, (3.4.4) gives us :

$$\eta\mu - \Sigma x^* = \eta \, R \, \mathbf{l}_n \tag{3.4.6}$$

Using (3.4.2) we can write Σ as follows:

$$\Sigma = \begin{bmatrix} \sum_{k=1}^{s}(p_k \cdot \bar{a}_{k1}^2) - \mu_1^2 & \sum_{k=1}^{s}(p_k \cdot \bar{a}_{k1} \cdot \bar{a}_{k2}) - \mu_1 \cdot \mu_2 & \cdots & \sum_{k=1}^{s}(p_k \cdot \bar{a}_{k1} \cdot \bar{a}_{kn}) - \mu_1 \cdot \mu_n \\ \sum_{k=1}^{s}(p_k \cdot \bar{a}_{k2}^2 \cdot \bar{a}_{k1}) - \mu_2 \cdot \mu_1 & \sum_{k=1}^{s}(p_k \cdot \bar{a}_{k2}^2) - \mu_2^2 & \cdots & \sum_{k=1}^{s}(p_k \cdot \bar{a}_{k2} \cdot \bar{a}_{kn}) - \mu_2 \cdot \mu_n \\ \cdots & \cdots & \ddots & \cdots \\ \sum_{k=1}^{s}(p_k \cdot \bar{a}_{kn}^2 \cdot \bar{a}_{k1}) - \mu_n \cdot \mu_1 & \sum_{k=1}^{s}(p_k \cdot \bar{a}_{kn} \cdot \bar{a}_{k2}) - \mu_n \cdot \mu_2 & \cdots & \sum_{k=1}^{s}(p_k \cdot \bar{a}_{kn}^2) - \mu_n^2 \end{bmatrix} \tag{3.4.7}$$

It is important to note that Σ has elements that are quadratic in $\mathbf{p}$ since μ_i is also a function of $\mathbf{p}$. See (3.4.1). This means that our feasible region for the inverse problem given by (3.4.6) is an optimization problem with quadratic constraints. This is a significantly complicated problem that we do not explore further in this thesis.

However, if we assume that Σ is given to us, this problem is greatly simplified; we can rewrite (3.4.6) as follows:

$$\mu = \bar{\mathcal{A}}^T \mathbf{p} = \frac{1}{\eta} \left(\eta \, R \, \mathbf{l}_n + \Sigma x^* \right) \tag{3.4.8}$$

The feasible region for our inverse optimzation problem is given by:

$$\begin{aligned} \text{s.t} : \bar{\mathcal{A}}^T \mathbf{p} &= \frac{1}{\eta} \left(\eta \, R \, \mathbf{l}_n + \Sigma x^* \right) \qquad (3.4.9) \\ \sum_{i=1}^{s} p_i &= 1 \\ p_i &\geq 0, \quad \forall i = 1 ... s \end{aligned}$$

If the matrix $\bar{\mathcal{A}}^T$ has full rank, there is a unique $\mathbf{p}$. If it isn't, then either there are no solutions or infinitely many solutions to the system given by (3.4.8)

3.4.1 Example 2 Revisted in a Mean-Variance Setting

Consider the inputs for the optimization given in Section 3.3.2 in Table 3.2. For purposes of mean-variance optimization, we only consider the scenarios with the risky assets. So our input matrix is slightly modified, in that it does not have the first column of asset values in Table 3.2. Let us assume that Σ is given. We calculate Σ using the `cov` command in `MATLAB`. μ is calculated as per (3.4.1). Let $\eta = 0.5$.[11] The optimal portfolio in this case is given by:

$$\mathbf{x}^* = \begin{bmatrix} -3.0749 \\ 0.0000 \\ 0.0033 \\ 0.0475 \\ -3.4036 \\ 2.6614 \\ 1.5138 \\ -5.1215 \\ 8.3738 \end{bmatrix}$$

Using this portfolio to solve a least squares minimization with the feasible region given by 3.4.9 gives us $\mathbf{p_{implied1}}$ in Figure 3.10. Code for the implementation is given in Appendix E.4.2.

If we use a different prior vector, say p_{ht} as in the previous section, we get $\mathbf{p_{implied2}}$ in Figure 3.10. The results are given in Figure 3.10.

As we can see, the implied probability vector is very different from both the original vector and the prior vector. It appears that the choice of prior does not seem to affect the resulting implied probability vector in a significant manner if we use the feasible region given by (3.4.9).

[11] It is important to note that in this case, higher values of eta imply lower risk aversion. Refer to 2.2.

$\mathbf{p}_{\mathbf{ht}}$	$\mathbf{p}_{\mathbf{implied1}}$	$\mathbf{p}_{\mathbf{implied2}}$	$\mathbf{p}_{\mathbf{original}}$
0.0087	0.0000	0.0000	0
0.0142	0.0000	0.0000	0
0.0311	0.0000	0.0022	0
0.0386	0.0015	0.0000	0
0.0540	0.0037	0.0000	0.0002
0.0691	0.0071	0.0001	0.0016
0.0783	0.0084	0.0249	0.0153
0.0826	0.0444	0.0447	0.0638
0.0863	0.1079	0.0977	0.1462
0.0766	0.1677	0.1711	0.2313
0.0821	0.3637	0.3672	0.2414
0.0726	0.2198	0.2128	0.1698
0.0690	0.0607	0.0642	0.0831
0.0555	0.0144	0.0144	0.0329
0.0507	0.0000	0.0000	0.0110
0.0402	0.0000	0.0000	0.0026
0.0314	0.0006	0.0006	0.0008
0.0231	0.0000	0.0000	0
0.0166	0.0000	0.0000	0
0	0.0000	0.0000	0

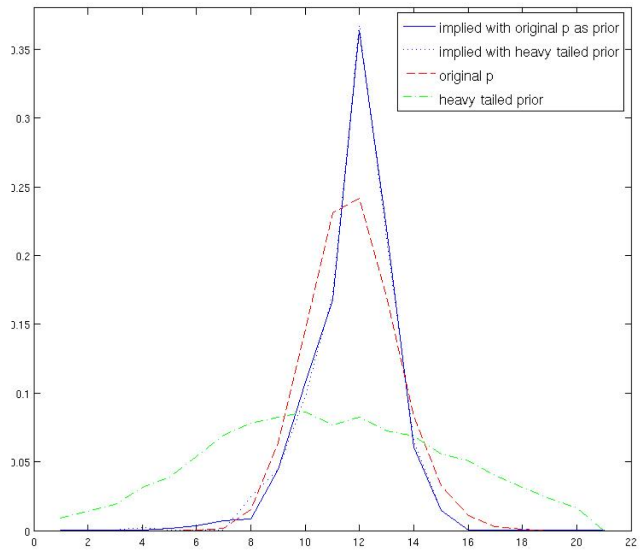


Figure 3.10: Prior probabilities and implied probabilities retrieved using (3.4.9) and original probability used to calculate means

Note that the condition number of $\bar{\mathcal{A}}$ in this case is 4.5132×10^5, which is quite large. It stands to reason that the system is ill-conditioned.

If we modify the feasible region as follows, we get more reasonable values for our implied distribution:

$$(1-\epsilon)\,\frac{1}{\eta}\,(\eta\,R\,\mathbf{l}_n+\Sigma x^*) \leq \bar{\mathcal{A}}^T\mathbf{p} \leq (1+\epsilon)\,\frac{1}{\eta}\,(\eta\,R\,\mathbf{l}_n+\Sigma x^*) \qquad (3.4.10)$$

$$\sum_{i=1}^{s} p_i = 1$$

$$p_i \geq 0, \quad \forall i = 1...s$$

Once again, ϵ is the error tolerance as a percentage of the RHS in (3.4.9).

If we minimize the distance to the original probability vector, with $\epsilon = [0.01, 0.05, 0.1, 0.5, 1]$, we get implied probabilities shown in Figure 3.11. As we can see, the implied probabilities given by $\epsilon = 0.5$ and $\epsilon = 1$ are the closest to the original probability. Numerically, $\epsilon = 1$ returns the prior exactly, as expected.

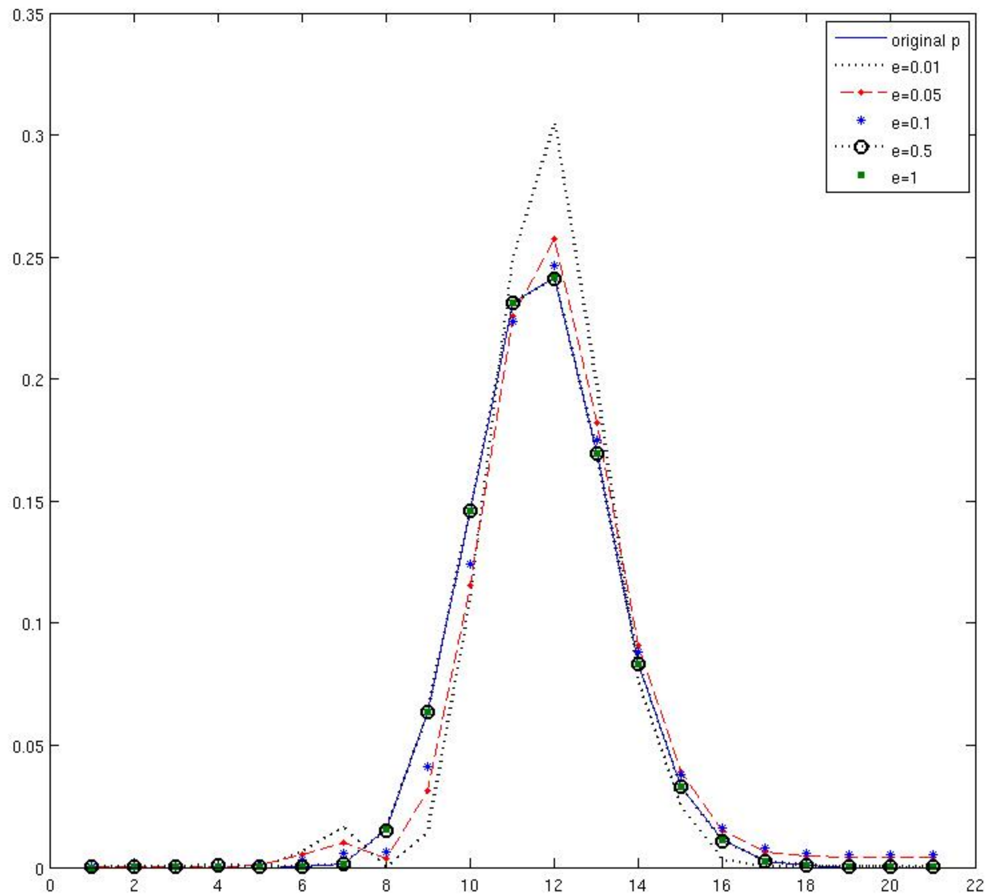


Figure 3.11: Implied probabilities retrieved using (3.4.10) and values of ϵ given above for prior given by $\mathbf{p_{original}}$.

If we repeat the exercise above, but with the heavy-tailed probability vector as a prior instead of the original probabilities, we get the implied probabilities shown in Figure 3.12. As we can see, the results are similar to the Expected Utility Maximization case, where the implied probabilities of the left tail events is close to zero.

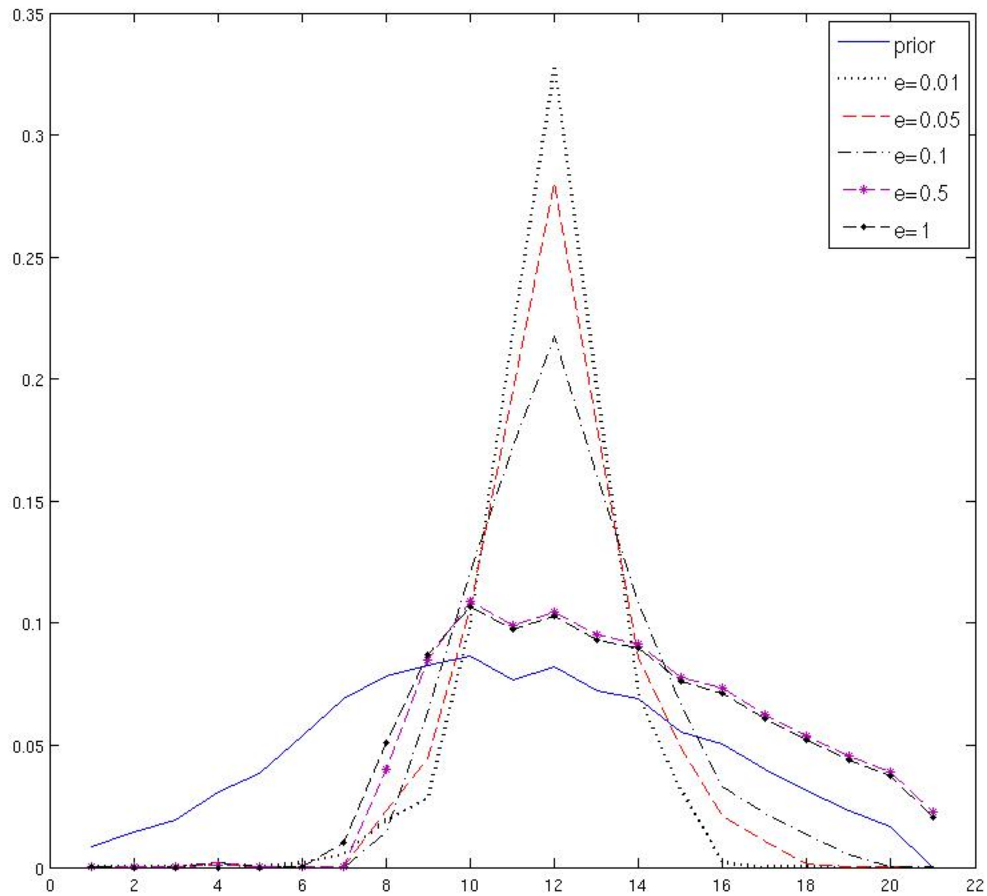


Figure 3.12: Implied probabilities retrieved using (3.4.10) and values of ϵ given above for prior given by $\mathbf{p_{ht}}$.

Chapter 4

Incorporating Implied Views in Portfolio Selection

Once we find the implied views on a portfolio, we can check to see if incorporating those views in portfolio selection affects portfolio performance in any way.

In this chapter, we show how one might go about retrieving the optimal portfolio of asset allocations by volume for an investor with a logarithmic utility function (or isoelastic utility function as in Chapter 3 with $\eta = 1$) given the portfolio weights. We look at two examples. In the first example, look at implied views from an actual portfolio of assets held by a Canadian Mutual Fund, and in the second example, we conduct a "rolling window" test to generate a series of out-of-sample portfolio values using the implied returns on a constant mix portfolio with 60% invested in equities and 40% in the risk-free asset. For the rolling window tests, we use the S&P 500 index as a proxy for equities and the US 1 month T-bill rates for the risk-free interest rate. [1]. We use the monthly index values from January 1962 to December 2010 to conduct our tests.

For the rolling-window tests, the results are as expected. If we assume that the investor is highly risk-averse, incorporating implied views in portfolio selection during "bubble" years gives lower returns compared to the portfolio selected using the empirical probability distribution for future scenarios. On the other hand, incorporating the views implied by a portfolio held by a risk-averse investor during the "bust" years improves portfolio performance compared to the portfolio selected using the empirical distribution. On the whole, for the period from 2002 to 2010 for which we have out-of-sample portfolios, the average portfolio value was highest for the 60-40 portfolio, followed by the $\frac{1}{N}$ portfolio, the mixture portfolio, and lastly, the portfolio selected using the empirical distribution.

[1] Data obtained from `https://wrds-web.wharton.upenn.edu/wrds/`.

4.1 Implied Views from a Canadian Mutual Fund: One Period Out-of-Sample Test

To generate the scenarios for our first example, we calibrate our asset returns to a mixture-of-normals model using the historical monthly prices on the S&P/TSX Sectors. Once we have the required parameters, we generate the scenarios using QMC methods.

Once we have the scenario set, we use the portfolio allocation retrieved from portfolio weights to solve the inverse optimization problem to get the implied probabilities associated with the scenarios.

Once we have this implied probability vector, we re-solve the portfolio optimization problem with a new probability vector which is an equal mix of the implied probability vector and the empirical probability vector.

We then compare the one-period out-of-sample performance to other portfolio selection methods such as equal risk contribution (ERC) portfolio, minimum-variance portfolio, the original fund portfolio and the equally-weighted portfolio. We find that the new portfolio generated by mixing the implied and the empirical probabilities performs better than all other portfolios except for the minimum-variance portfolio.

However, this analysis is only done over one period. The results might be very different under other periods with different market conditions. Also, there is ill-conditioning in the system, which can lead to errors in implied probabilities that we retrieve and the optimal portfolio that we get using the mixed probability vector.

One way to remedy this problem would be to set higher tolerances. However, the tolerances that we pick are rather arbitrary, so there is a great deal of subjectivity associated with the implied probability vector. Finding criteria for reasonable tolerance numbers could be a possible direction for future research.

In this chapter, we use the Acuity All Cap 30 Canadian Equity Fund as of July 31st, 2011. Fund data can be accessed at `http://www.theglobeandmail.com/globe-investor/funds-and-etfs/funds/summary/?id=53551&cid=Acuity%20Funds%20Ltd.`. A screen-capture of the fund profile on August 28th, 2011 can be seen in Figure 4.1.

Fund Facts

Fund Sponsor:	Acuity Funds Ltd.	**Globe 5-Star Rating:**	--
Managed by:	Acuity Investment Management	**RRSP Eligible:**	Canadian
Fund Type:	MF Trust	**Min. Invest (initial):**	$500
Inception Date:	September 2000	**Subsequent:**	$100
Asset Class:	Canadian Focused Small/Mid Cap Equity	**Min. Invest (initial RRSP):**	$500
Quartile Rank:	3 (3YR ending July 31, 2011)	**Subsequent RRSP:**	$500
Total Assets:	$227.0 million	**Closed:**	No
Mgmt Exp. Ratio (MER):	2.95%	**Restricted:**	Yes
Management Fee:	2.50%	**Restriction:**	Not Available in NUV
Load Type:	Optional		

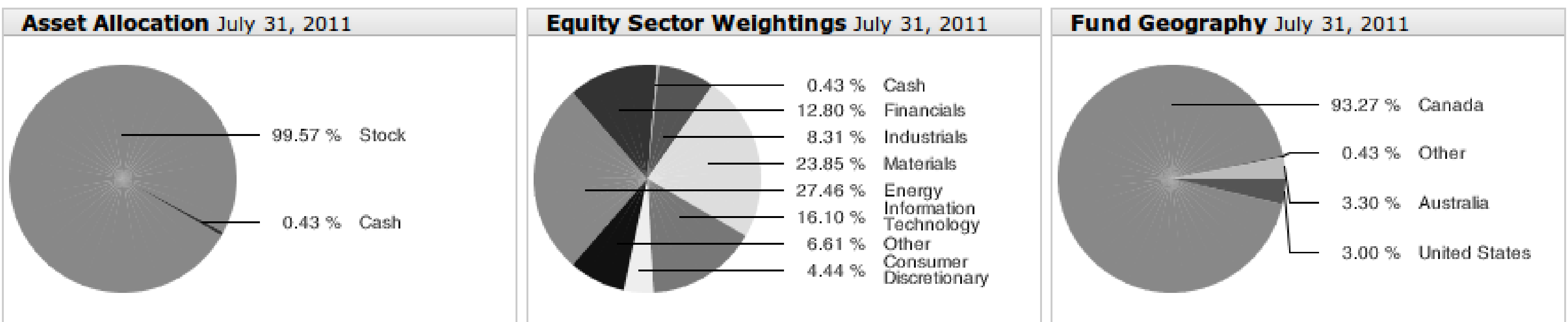


Figure 4.1: Fund Profile for Acuity All Cap 30 Canadian Equity Fund

We use the equity sector weightings as the optimal portfolio weights $\mathbf{x}^*$, and we use the corresponding sector index values as a proxy for the aggregate value of stocks in that respective sector. We use 112 months of data to estimate parameter values for the distribution of assets. We assume that returns follow a joint normal distribution and that index values follow a log-normal distribution. We use a mixture of normals model to generate our scenario matrix. We use the `MATLAB` function `gmdistribution.fit` to calibrate our mixture of normals model with 2 components. We then use this information to generate 150 scenarios using the Sobol set in dimension 7.

The sector classes and their corresponding indices and portfolio weights are given in Table 4.1.

x	**Sector**	**Corresponding S&P/TSX Index**
0.0043	Cash	Government of Canada 1-month T-bill
0.1280	Financials	S&P/TSX Capped Financials Index
0.0831	Industrials	S&P/TSX Capped Industrials Index
0.2385	Materials	S&P/TSX Capped Materials Index
0.2746	Energy	S&P/TSX Capped Energy Index
0.1610	Information Technology	S&P/TSX Capped Information Technology Index
0.0444	Consumer Discretionary	S&P/TSX Capped Consumer Discretionary Index
0.0661	Other	S&P/TSX Capped Diversified Metals and Minerals

Table 4.1: Portfolio weights, sectors and corresponding assets

We retrieve the implied probabilities for the Expected Utility Maximization under the following assumptions:

1. the investors have a logarithmic utility function,
2. the risk-free rate of interest is given by the yield on the 1-month Govt. of Canada T-bill and is calculated as (*quoted yield*) $\times \dfrac{31}{365}$. We use the quoted yield of 0.91%[2] and,
3. the prior distribution $\mathbf{p_0}$ is given by the empirical distribution, i.e. each scenario has a probability $p_i = \dfrac{1}{s}$, $s = 150$.

[2] 1 month T-bill yields for the past 3 months can be found here: `http://www.bankofcanada.ca/rates/interest-rates/t-bill-yields/selected-treasury-bill-yields-10-year-lookup/?rangeType=dates&rangeValue=1&rangeWeeklyValue=1&rangeMonthlyValue=5&lP=lookup_tbill_yields.php&sR=2001-09-10&se=L_V122529&dF=2011-06-11&dT=2011-08-11` .

Since the initial investment is \$227.0 Million, it stands to reason that our sensitivity parameter λ will be very small for the logarithmic utility function. Our scenario matrix has a fairly high condition number (9.756010^5), which means that we need to introduce tolerances to retrieve meaningful solutions.

We solve the inverse problem using the following formulation:

$$\begin{aligned} \min_{h} \ & ||\mathbf{p_0} - \mathbf{h}|| \\ \text{s.t} : & \ \mathbf{z} - 0.05\,\mathbf{z} \leq \mathcal{A}^{\mathbf{T}}\widetilde{\mathcal{U}'}\mathbf{h} \leq \mathbf{z} + 0.05\,\mathbf{z} \\ & h_i \geq 0, \ \ \forall i = 1...s, \end{aligned} \tag{4.1.1}$$

where $0.05\,\mathbf{z}$ are the error tolerances.

We can retrieve the implied probability vector by normalizing $\mathbf{h}$. Let the normalized vector be denoted by $\mathbf{p}_{imp}$.

Suppose we want to test the performance of the given portfolio with the performance of the portfolio calculated using our prior distribution and the portfolio calculated using the mixture of our implied probabilities and prior probabilities.[3] We can denote a new probability vector $\mathbf{p}_{new}$ as the weighted combination of $\mathbf{p}_{imp}$ and $\mathbf{p}_0$. Let the weighing parameter α be 0.5, so $\mathbf{p}_{new} = 0.5\mathbf{p}_{imp} + 0.5\mathbf{p}_0$. If we want to calculate the new optimal portfolio, using formulation 3.3.1, we need a risk-neutral probability vector. In this case, we find the state-price vector by solving the following linear program:

$$\begin{aligned} \min_{h} \ & ||\psi|| \\ \text{s.t} : & \ \mathbf{z} - 0.05\,\mathbf{z} \leq \mathcal{A}^{\mathbf{T}}\psi \leq \mathbf{z} + 0.05\,\mathbf{z} \\ & h_i \geq 0, \ \ \forall i = 1...s \end{aligned} \tag{4.1.2}$$

Where $0.05\,\mathbf{z}$ are the error tolerances.

The risk-neutral vector $\mathbf{q}$ is the normalized ψ vector.

However, using formulation 3.3.1 leads to errors in the calculation of the optimal portfolio allocation vector, and portfolio weights do not sum to 1. This error can be attributed to the errors involving computations that use the Moore-Penrose inverse of our scenario matrix: namely the calculation of the risk-neutral probability and the optimal portfolio allocation from the optimal wealth. This entire system is very unstable - solving it multiple times yields different results. However, in every iteration, the optimal allocation has the majority of the wealth allocated to the risk-free asset. For the sake of demonstration, we use the following portfolio.

[3] In our case, for the sake of simplicity, we chose the empirical probabilities. In reality, we could have a very different prior vector more representative of the portfolio manager's views on the market.

The graphical relationship between the different probabilities for this instance is shown in Figure 4.2. The horizontal axis has the scenarios sorted in ascending order by terminal wealth, and the vertical axis shows the probabilities associated with each of the scenarios.

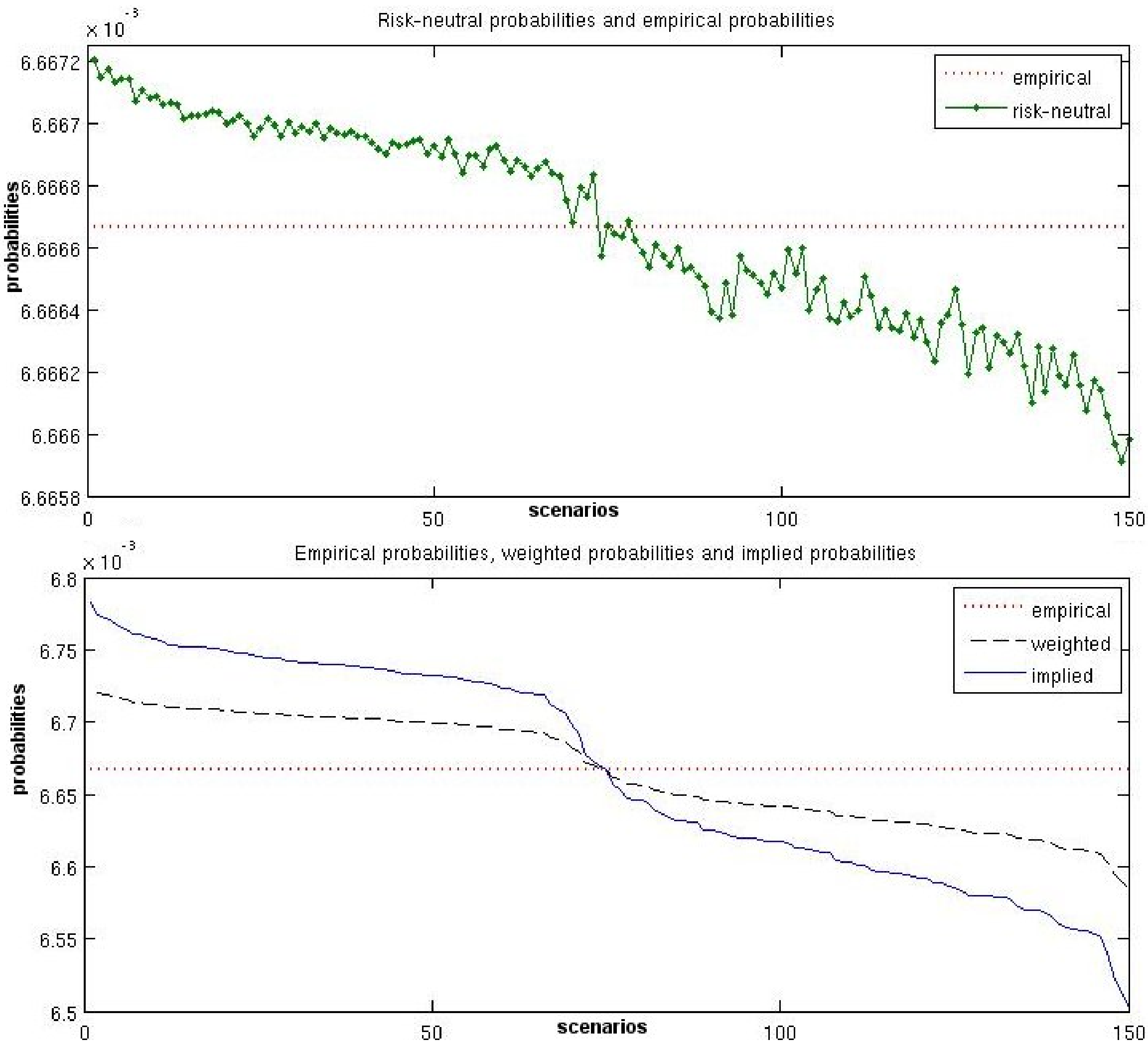


Figure 4.2: Different probability vectors

The optimal portfolio allocation by volume and the associated optimal portfolio weights are given by:

$$\mathbf{y}^*_{new} = 10^6 \times \begin{bmatrix} 237.600319 \\ -0.010144 \\ -0.003369 \\ -0.015328 \\ -0.040753 \\ -0.015754 \\ -0.0013920 \\ -0.000461 \end{bmatrix}, \mathbf{x}^*_{new} = \begin{bmatrix} 1.0467 \\ -0.0090 \\ -0.0047 \\ -0.0129 \\ -0.0060 \\ -0.0083 \\ -0.0025 \\ -0.0029 \end{bmatrix} \quad (4.1.3)$$

The optimal portfolio allocation by volume and the associated optimal portfolio weights when we just use the empirical probabilities are given by:

$$\mathbf{y}^*_{empirical} = 10^6 \times \begin{bmatrix} 229.196161 \\ -0.005228 \\ -0.001479 \\ -0.005067 \\ -0.030647 \\ -0.004490 \\ -0.000495 \\ 0.000011 \end{bmatrix}, \mathbf{x}^*_{empirical} = \begin{bmatrix} 1.0096 \\ -0.0046 \\ -0.0021 \\ -0.0043 \\ -0.0046 \\ -0.0024 \\ -0.0009 \\ 0.0001 \end{bmatrix}. \quad (4.1.4)$$

We can now compare the out-of-sample performance of the above portfolio, with the original portfolio, the $\frac{1}{N}$ portfolio and the minimum-variance portfolio.

The minimum-variance portfolio, ERC portfolio and the equally weighted portfolios are given by:

$$\mathbf{x}^*_{mv} = \begin{bmatrix} 0 \\ 0.9171 \\ 0.0235 \\ 0.0356 \\ 0.0056 \\ 0.0003 \\ 0.0047 \\ 0.0130 \end{bmatrix}, \mathbf{x}^*_{erc} = \begin{bmatrix} 0.9990 \\ 0.0006 \\ 0.0001 \\ 0.0001 \\ 0.0001 \\ 0.0001 \\ 0.0000 \\ 0.0000 \end{bmatrix}, \mathbf{x}^*_{\frac{1}{N}} = \begin{bmatrix} 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.1250 \end{bmatrix}. \quad (4.1.5)$$

The associated portfolio allocations by volume are given by:

$$\mathbf{y}^*_{mv} = 10^6 \times \begin{bmatrix} 0 \\ 1.032998 \\ 0.016918 \\ 0.042272 \\ 0.037305 \\ 0.000630 \\ 0.002684 \\ 0.002074 \end{bmatrix}, \mathbf{y}^*_{erc} = 10^6 \times \begin{bmatrix} 226.767716 \\ 0.000689 \\ 0.000068 \\ 0.000140 \\ 0.000377 \\ 0.000102 \\ 0.000025 \\ 0.000007 \end{bmatrix}, \mathbf{y}^*_{\frac{1}{N}} = 10^6 \times \begin{bmatrix} 28.375000 \\ 0.140791 \\ 0.008980 \\ 0.014829 \\ 0.083975 \\ 0.023745 \\ 0.007062 \\ 0.001988 \end{bmatrix}. \quad (4.1.6)$$

The ending wealth for each of these portfolios as calculated using asset values on August 2nd, 2011, is given below in descending order:

Portfolio Selection Method	Terminal Wealth (in Millions)
Minimum-variance ($\mathbf{y}^*_{mv}$)	227.7194
Weighted probabilities ($\mathbf{y}^*_{new}$)	227.3982
Empirical probabilities ($\mathbf{y}^*_{emp}$)	227.2095
ERC ($\mathbf{y}^*_{erc}$)	227.1745
Equally Weighted$\frac{1}{N}$ ($\mathbf{y}^*_{\frac{1}{N}}$)	225.0082
Original portfolio ($\mathbf{y}^*$)	223.7614

Table 4.2: Portfolio Value on August 21, 2011.

As we can see, the original optimal portfolio is actually the worst performing portfolio. However, in this case, incorporating the implied probabilities into the empirical probabilities gives us a better performing portfolio compared to the one calculated using empirical probabilities and the original portfolio. [4]

These results look promising, but they should be taken with a grain of salt. The scenario matrix has a high condition number, so there is a high likelihood of the resulting portfolio allocations being inaccurate to some degree.

As mentioned before, to determine whether incorporating implied views into portfolio selection actually improves out-of-sample performance, we should consider repeating the above analysis for the same fund with portfolios at different points in time with a longer time-series. This is exactly what we do in the next section with a constant mix portfolio.

[4]Code for the above exercise is provided in Appendix E.4.1.

4.2 Implied Views from Constant Mix Portfolio: Rolling Window Out-of-Sample Test

In this section, we consider an investor who invests precisely 60% in equities and 40% in the risk-free asset regardless of market conditions. If we assume that this investor's utility function is given by 3.2.16, with $\eta = 2$, we can repeat the above exercise of retrieving implied views and incorporating them in portfolio selection for the constant mix portfolio over a longer time frame.

We use the S&P 500 index as a proxy for the equities in the market and we use the yield on the 1-month US T-bills to calculate the risk-free rate to be used for our problem. We assume that at each iteration, the investor has \$1000.00 to invest at the beginning of the period.

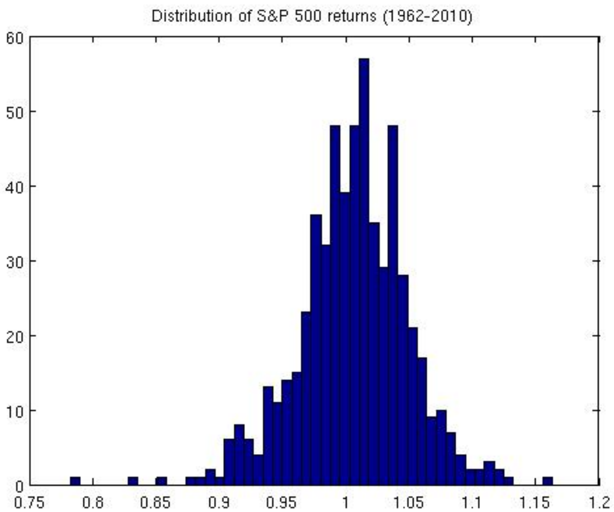
We use data from January 1962 to December 2010 as our available data set (588 points). We use a rolling window of 480 months (or 40 years) to estimate the parameters of our our mixture-of-normals model for the monthly returns on the S&P 500.

At each iteration, we discard the data for first month used in the estimation window, and add the out-of-sample data for the previous iteration. This gives us 108 out-of sample portfolio values to compare for the original portfolio ('60-40'), the new portfolio('Mixed'), the one calculated using the empirical distribution ('Empirical'), and the equally-weighted portfolio ('50-50'). We can compare the portfolio performances over the period from January 2002 to December 2010, which include the bubble years as well as the 2007-2008 crash and subsequent 2 years.

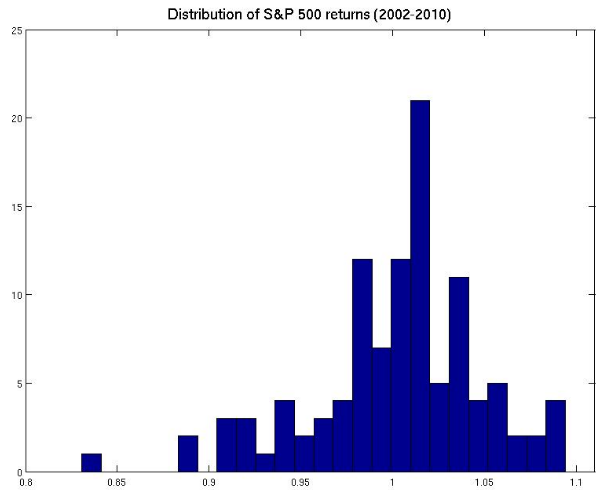
Summary statistics for the returns on S&P 500 data series are given in Table 4.3 and the histograms for returns are given in Figure 4.3.

	Returns calculated over:			
	1962-2010	2002-2010	2002-2006	2007-2010
Mean	1.0059	1.0019	1.0043	0.9991
Variance	0.0019	0.0021	0.0013	0.0032
Skewness	-0.4257	-0.7385	-0.6138	-0.6274
Kurtosis	4.7311	4.1730	4.4002	3.2033

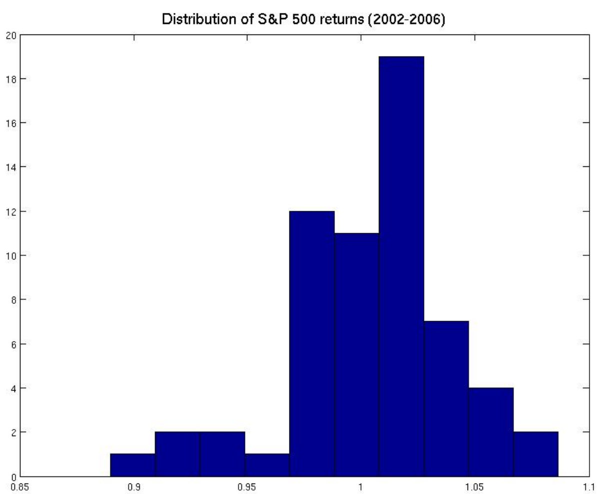
Table 4.3: Summary statistics for returns on S&P 500 data series.



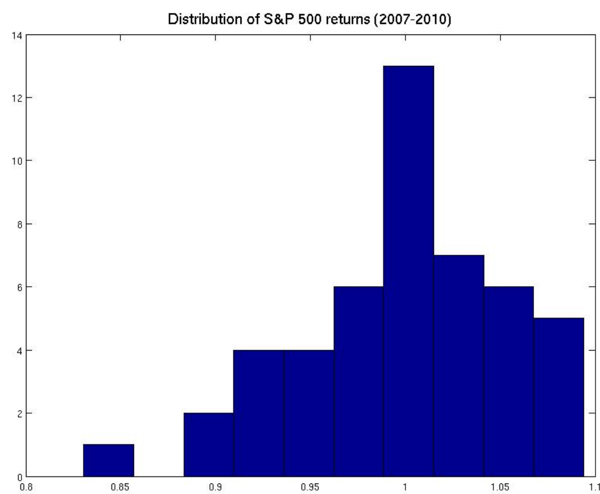
(a) Returns on S&P 500 (1962-2010)



(b) Returns on S&P 500 (2002-2010)



(c) Returns on S&P 500 (2002-2006)



(d) Returns on S&P 500 (2007-2010)

Figure 4.3: Histograms of returns on the S&P 500 over different periods.

The time-series for the out-of-sample portfolio values is shown in Figure 4.4. The average returns for each of the portfolios over the different time periods- 2002-2010, 2002-2006, and 2007-2010 - are shown in Table 4.4.

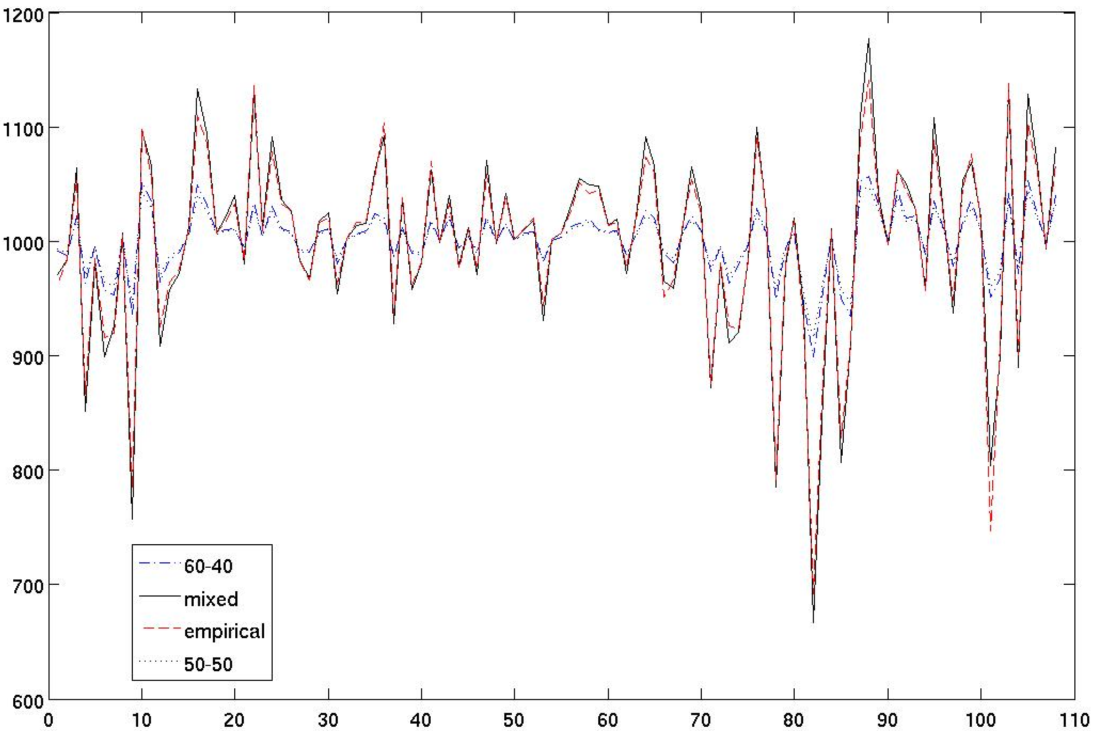


Figure 4.4: Realized portfolio values from 2002 to 2010.

	60-40	Mixed	Empirical	50-50
Average over 2002 to 2010	1001.7560	1000.1758	999.0271	1001.7133
Bubble (2002 - 2006)	1003.0977	1006.7718	1006.8921	1002.8315
Bust (2007 - 2010)	1000.0788	991.9307	989.1960	1000.3157

Table 4.4: Average out-of-sample portfolio performance

The results show that the average return for the 'Mixed' portfolio lies between the average return for the 'Empirical' portfolio and the original '60-40' portfolio. In this case, it appears as though incorporating implied views in portfolio selection during "bubble" years gives lower returns compared to the 'Empirical' portfolio, while incorporating the views implied by a portfolio held by a risk-averse investor during the "bust" years improves portfolio performance compared to the 'Empirical' portfolio. On the whole, for the period from 2002 to 2010 for which we have out-of-sample portfolios, the average portfolio value was highest for the '60-40' portfolio, followed by the '50-50' portfolio, the 'Mixed' portfolio, and lastly, the 'Empirical' portfolio.

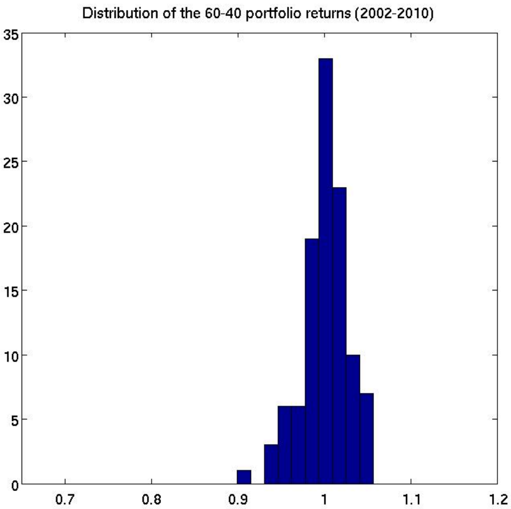
We can also look at the distribution of the portfolio returns that we get from each of these 4 strategies. The summary statistics for the returns on these 4 portfolios are given in Table 4.5. We can see that the returns are skewed to the left for all the strategies, but the skew is more extreme for the 'Mixed' and the 'Empirical' portfolio returns. What this means is that we have heavier left tails for the returns from these two strategies compared to the returns from the '60-40' or '50-50' strategies. The values for kurtosis are also higher than 3, which means that the distributions for these returns have a higher peak than the normal distribution. In conclusion, the returns clearly do not follow a normal distribution, a fact corroborated by the histograms of returns shown in Figure 4.5.

	60-40	Mixed	Empirical	50-50
Mean	1.0018	1.0002	0.9990	1.0017
Variance	0.0008	0.0066	0.0058	0.0005
Skewness	-0.7385	-1.1351	-1.2716	-0.7385
Kurtosis	4.1730	5.5211	5.8403	4.1730

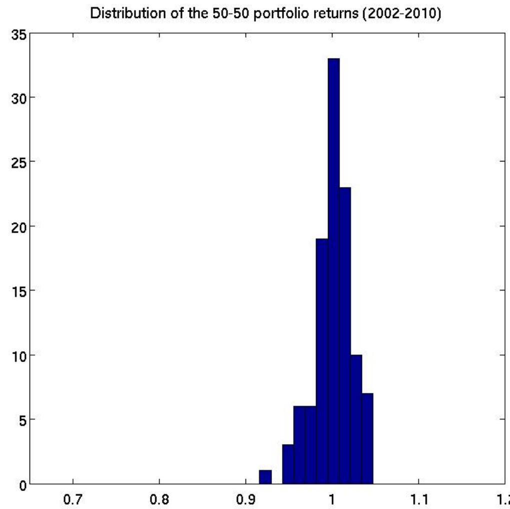
Table 4.5: Summary statistics for returns on out-of-sample returns for the 4 different portfolio selection strategies.

We can also compare the realized Sharpe ratio for the 4 strategies, and calculate the confidence interval (C.I.) associated with each Sharpe Ratio. As seen above, the returns are not IID normal, so we cannot use the asymptotic distribution of the Sharpe ratio calculated under the assumption of IID normal returns. To calculate our C.I.'s, we use the formula for the asymptotic variance of the estimator of the Sharpe ratio $Var(\hat{S})$ without the assumption of normality given by Christie (2005, eq. (21)). The Sharpe ratio is calculated as $\hat{S} = \dfrac{\mathbb{E}(R_t - R_{ft})}{\sqrt{Var(\hat{S})}}$. The half-width of the 95% C.I. is calculated as $1.92 \times \sqrt{\dfrac{Var(\hat{S})}{T - M}}$, where T is the total length of our time series (588), and M is the length of our rolling window (480).

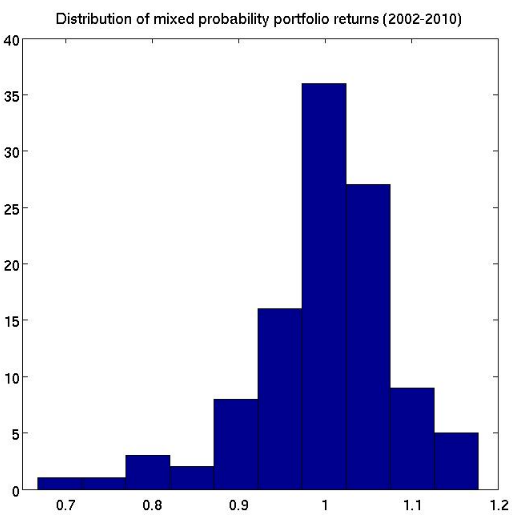
The results for the 95% confidence interval for the Sharpe ratios for the 4 strategies are given in Table 4.6



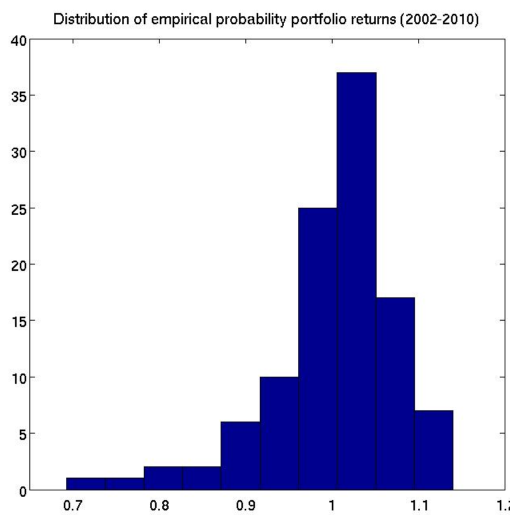
(a) Realized returns for the '60-40' portfolio



(b) Realized returns for the '50-50' portfolio



(c) Realized returns for the 'Mixed' portfolio



(d) Realized returns for the 'Empirical' portfolio

Figure 4.5: Histograms of realized returns for different portfolio selection strategies.

	60-40	Mixed	Empirical	50-50
Lower Bound	-0.9508	-0.6367	-1.3579	-0.9793
Sharpe Ratio	0.0038	-0.0181	-0.0343	0.0027
Upper Bound	0.9584	0.6005	1.2893	0.9848

Table 4.6: Confidence intervals for the Sharpe ratios from the 4 different portfolio selection strategies.

As we can see, the intervals are quite large, and there is a fair amount of overlap. As a matter of fact, the intervals for 'Mixed' portfolio is contained within the interval for the '60-40' portfolio, which is contained within the interval for the '50-50' portfolio, which in turn is contained within the interval for the 'Empirical' portfolio. This means that the difference between our Sharpe ratios is not really statistically significant.

Another thing we might be interested in is the cost associated with implementing each of these strategies. Since we do not consider transaction costs, we use turnover as a proxy. We calculate turnover as per the formula given by DeMiguel et al. (2009, p. 12 eq. 15). Results for the 4 strategies are given in Table 4.7. DeMiguel et al. (2009) interpret the turnover as the average percentage of wealth traded in each period. So, the smaller our turnover, the lower the cost of implementing the strategy. We can see, that the turnover is the lowest for the '50-50' portfolio and highest for the 'Empirical' portfolio. Incorporating the implied views from the '60-40' portfolio appears to reduce our turnover slightly.

	60-40	Mixed	Empirical	50-50
Turnover	0.0213	1.1363	1.2540	0.0180

Table 4.7: Turnover for the 4 different portfolio selection strategies.

The code for this implementation can be found in Appendix E.4.2.

Based on the analysis above, it appears that in this case, incorporating implied views into portfolio selection does improve portfolio performance slightly. We suggest repeating the analysis with a different proportions for mixing the empirical and implied probability distributions. As an extension, it would also be an interesting exercise to use a different distribution, say normal or log-normal, for scenario probabilities instead of the empirical distribution to generate different optimal portfolios. It would, no doubt, be informational to explore the effect of incorporating implied probabilities on portfolio performance when the prior distribution is not the empirical distribution.

Chapter 5

Conclusions

In summary, this thesis introduces a portfolio selection framework where we consider a finite, discrete set of scenarios to occur sometime in the future. Assuming we know that the investor's utility function is given, and that there is a portfolio that is optimal for our scenario set, we then develop a framework to find the likelihood of occurrence of each of those scenarios as implied by the choice of "optimal" portfolio. This thesis generalizes the inverse problem for expected utility maximization to non-linear (concave) utility functions for risk-averse investors. We show conditions for the utility functions under which our proposed framework solves the inverse problem.

We use principles of mathematical programming and linear algebra to derive results for the inverse problem, first, in case of complete markets, and then, we extend this analysis to incomplete markets.

Exploring the problem in complete markets provides us with a key understanding as to how the inverse problem is related to the forward problem when all assets in the asset universe can be replicated. Basically, in a complete market, the probability vector associated with the optimal portfolio is unique, and since the scenario matrix is invertible, there exists a closed form solution to our inverse problem.

Also, the inverse problem depends on the first derivative of the investor's utility function evaluated at the optimal portfolio for each of the scenarios. This can create additional instability in our system if the investor's utility function is extremely flat or extremely steep at the point where it is being evaluated. Extremely flat functions give us extremely small positive first derivative values, and extremely steep functions give us extremely large positive first derivative values. In both cases, extreme values can cause computing problems. This can create computational problems while we solve for the optimal inverse solution.

When we extend the inverse problem framework to an incomplete market, where we have more scenarios than assets. In this case, we do not have a nice closed-form solution as our matrix is not invertible. In theory, the feasible region has equality constraints, however, in practice, the system is highly ill-conditioned, and we need to use regularization techniques, or additional error tolerances to retrieve reasonable solutions. However, the choice of these tolerances is entirely arbitrary, and we have to use trial-and-error to retrieve solutions that do not have a lot of noise.

When we are minimizing the distance to a prior distribution, it is important to note that the larger our tolerance, the closer our solution is to the prior. However, it can be very different from the original probabilities used to calculate the optimal portfolio we are given. It would be nice to establish a range for the tolerances based on the information present in our system. Determining a 'reasonable' tolerance that will return the most accurate implied probability vector could be a direction for further research.

We also show how one might go about getting the maximum or minimum probability of an event occurring as implied by an optimal portfolio.

Once we have implied probabilities, as given by a particular choice of portfolio, we get a new vector of probabilities associated with the scenarios by mixing the implied probabilities with the empirical probabilities. We then solve the forward optimization problem again, this time with the new probabilities, to get a different optimal portfolio. We then compare the out-of-sample performance of this portfolio with the out-of-sample performance of the original optimal portfolio and the optimal portfolio that we get using the empirical probabilities over several time periods. As expected, the average portfolio value given by the new portfolio lies between value of the portfolio calculated using the empirical probabilities and the original portfolio.

In conclusion, we show how to construct a feasible region for the inverse problem given the investor's utility function, an optimal portfolio, current asset prices, and the scenario matrix. We then go on to provide numerical examples demonstrating how we use the formulation of the feasible region proposed above. In the presence of numerical instabilities, when the constraint matrix for our inverse problem is ill-conditioned, we also suggest a couple of ways to regularize the problem.

We also provide two examples of how one might incorporate views implied by a given portfolio by mixing them with a different distribution and compare the values of the resulting portfolios over time. We discover that in periods of high positive returns, risk-averse portfolios have a poorer performance than the mixed portfolios, and in periods of low returns, the risk-averse portfolios show better performance than the mixed portfolios. One direction for future research could be to determine at what point would incorporating implied views from risk-averse portfolios in portfolio selection yield higher returns than simply using the empirical distribution.

APPENDICES

Appendix A

Important Results from Linear Algebra

A.1 The Spanning Set Theorem

Definition A.1.1. *Let H be a subspace of a vector space V. An indexed set of vectors $\mathcal{B} = \{\mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_p}\}$ in V is a* **basis** *for H if*

1. *$\mathcal{B}$ is a linearly independent set, and*

2. *$H = Span\{\mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_p}\}$*

Theorem A.1.1. *Let $S = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_p}\}$ be a set in V, and let $H = Span\{\mathbf{b_1}, \mathbf{b_2}, ..., \mathbf{b_p}\}$.*

1. *If one of the vectors in S, for instance, $\mathbf{v_k}$, is a linear combination of the remaining vectors in S, then the set formed from S by removing $\mathbf{v_k}$ still spans H.*

2. *If $H \neq \{0\}$, some subset of S is a basis for H*

Please refer to Lay (2003), section 4.3, pg 240 for a proof of the above theorem.

A.2 The Rank Theorem

Theorem A.2.1. *The dimensions of the column space and the row space of an $m \times n$ matrix A are equal. This common dimension, $Rank(A)$ also equals the number of pivot positions in A and satisfies the equation*

$$Rank(A) + dim(Null(A)) = n$$

Please refer to Lay (2003), section 4.6, pg 265-266 for a proof of the above theorem.

A.3 Invertible Matrix Theorem

Let A be a square $n \times n$ matrix, and let I be the $n \times n$ identity matrix. Then the following statements are equivalent.

1. A is an invertible matrix.
2. A is row-equivalent to the $n \times n$ identity matrix.
3. A has n pivot positions.
4. The equation $A\mathbf{x} = \mathbf{0}$ only has the trivial solution.
5. The columns of A form a linearly independent set.
6. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
7. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
8. The equation $A\mathbf{x} = \mathbf{b}$ has exactly one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
9. The columns of A span $\mathbb{R}^n$.
10. There exists an $n \times n$ matrix C such that $CA = I$ and in this case, $A = C^{-1}$
11. There exists an $n \times n$ matrix D such that $AD = I$ and in this case $A = D^{-1}$.
12. A^T is an invertible matrix.
13. The columns of A form a basis for $\mathbb{R}^n$.
14. Col $A = n$.

15. dim Col $A = n$.

16. Rank $A = n$.

17. Null $A = \mathbf{0}$.

18. dim Null $A = 0$.

19. det $A \neq 0$.

20. The number 0 is not an eigenvalue of A. (If 0 is an eigenvalue of A then A is not invertible.)

21. $(\text{Col } A)^{\perp} = \{\mathbf{0}\}$.

22. $(\text{Null } A)^{\perp} = \mathbb{R}^n$.

23. Row $A = n$.

24. A has n non-zero singular values.

Proofs for statements (1)-(12) can be found in section 2.3 of Lay (2003).
Proofs for statements (13)-(18) can be found in section 2.9 of Lay (2003).
Proof for (19) can be found in section 3.2 of Lay (2003).
Proof for (20) can be found in section 5.2 of Lay (2003).
Proofs for statements (21)-(24) can be found in section 7.4 of Lay (2003).

A.4 The Singular Value Decomposition (SVD)

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix S of the form

$$S = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

for which the diagonal entries in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and there exists an $m \times m$ unitary matrix U and an $n \times n$ unitary matrix V such that

$$A = USV^T \tag{A.4.1}$$

We refer the reader to Lay (2003), section 7.4 and Björck (1996), section 1.2 (Theorem 1.2.1) for details on calculating the component matrices of the SVD.

When S contains rows or columns of zeros, a more compact decomposition of A is possible. We can partition the matrices U and V into submatrices, whose first blocks contain the r columns corresponding to the r singular values in D, as follows

$$U = [U_r \quad U_{m-r}] \quad \text{and } V = [V_r \quad V_{n-r}].$$

We can rewrite A as

$$A = [U_r \quad U_{m-r}] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} [V_r \quad V_{n-r}] = U_r D V_r^T \tag{A.4.2}$$

We can write the Moore-Penrose inverse of A as

$$A^+ = V_r D^{-1} U_r^T \tag{A.4.3}$$

It is important to note that columns of U_{m-r} form an orthonormal basis for $Null(A^T)$. Proof can be found in Lay (2003), pg 478, example 6.

A.5 Matrix Norm

Let A be an $m \times n$ matrix. Given any vector norm $||\,\centerdot\,||_v$, the corresponding *subordinate matrix norm* of A, also denoted by $||\,\centerdot\,||_v$, is defined by

$$||A||_v = \max_{||\mathbf{x}||_v \neq 0} \frac{||A\mathbf{x}||_v}{||\mathbf{x}||_v}$$

which satisfies the following properties:

$$1. \quad ||A||_v \geq 0, \qquad A \in \mathbb{R}^{m\times n} \tag{A.5.1}$$
$$2. \quad ||A+B||_v \leq ||A||_v + ||B||_v, \qquad A, B \in \mathbb{R}^{m\times n} \tag{A.5.2}$$
$$3. \quad ||\theta A||_v = |\theta|||A||_v, \qquad A \in \mathbb{R}^{m\times n}, \theta \in \mathbb{R} \tag{A.5.3}$$
$$4. \quad ||AB||_v \leq ||A||_v||B||_v, \qquad A, B \in \mathbb{R}^{m\times n} \tag{A.5.4}$$

For our case, we will be considering the 2-norm ($v = 2$).

Note that for the 2-norm (also called the **spectral norm**), the norm of a unitary matrix is 1. This is easy to verify using the above definition. If Q is an $n \times n$ unitary matrix, we know that $Q^T Q = \mathbb{I}_n$, where $\mathbb{I}_n$ is the identity matrix. $||Q\mathbf{x}||_2^2 = \mathbf{x}^T Q^T Q \mathbf{x} = \mathbf{x}^T \mathbf{x} = ||\mathbf{x}||_2^2$. The result then follows from the definition.

Appendix B

Lagrange Multiplier Theorem

Proposition 3.1.1 [1]: Let x^* be a local minimum of f subject to $h(x) = 0$, and assume that the constraint gradients $\nabla h_1(x^*), \nabla h_2(x^*), \ldots, \nabla h_m(x^*)$ are linear independent. Then there exists a unique vector $\lambda^* = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ called a *Lagrange multiplier vector*, such that,

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(x^*) = 0$$

If in addition f and h are twice continuously differentiable, we have

$$y^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \right) y \geq 0, \qquad \forall y \in V(x^*)$$

where $V(x^*)$ is the subset of the first-order feasible variations

$$V(x^*) = \{y | \nabla h_i(x^*)^T y = 0 \ , i = 1, \ldots, m\}$$

Proposition 3.2.1 [2]: Assume that f and h are twice continuously differentiable, and let $x^* \in \mathbb{R}^n$ and $\lambda^* \in \mathbb{R}^m$ satisfy

$$\nabla_x L(x^*, \lambda^*) = 0, \qquad \nabla_\lambda L(x^*, \lambda^*) = 0$$

If,

$$y^T \nabla_{xx}^2 L(x^*, \lambda^*) y > 0, \qquad \forall y \neq 0 \text{ with } \nabla h(x^*)^T y = 0.$$

Then x^* is a strict local minimum of f subject to $h(x) = 0$.

[1] Bertsekas (1995, p. 278-279)

[2] Bertsekas (1995, p. 296)

Appendix C

Lower Bound on the Lagrange Multiplier

Theorem C.0.1. *For an expected utility maximization problem in a discrete-time, single period framework outlined in section 3.2 and section 3.3, we hold the following assumptions to be true:*

- *There is a risk-free asset in the market with a positive rate of return r, such that total return is given by $R = (1 + r)$.*
- *The optimization problem is only subject to a budget constraint.*
- *The investor's utility function is given by $\mathcal{U}$, such that $\mathcal{U} \in \mathcal{C}^{(2)(D)}$, $\mathcal{U}' > 0$ and $\mathcal{U}$ is concave.*

If the above assumptions are satisfied, then the Lagrange multiplier associated with the budget constraint is strictly greater than zero.

Proof. Let the scenario set be given by the input matrix $\mathcal{A}$ in (3.2.15). For sake of simplicity, let the last asset be the risk-free asset, and let the value of the risk-free asset at time 0 be 1 and the value at the end of the time period be given by $R > 0$. Let the objective value function in (3.2.17) be denoted by: $f(B) = \mathbb{E}_p(\mathcal{U}(A\mathbf{y}))$.

We can see from (3.2.17) that $\mathbf{y}$ is a function of the budget B. Also, the objective function is a non-negative linear combination of concave functions, and is therefore a concave function itself (Boyd & Vandenberghe, 2004, p. 79). Since the Lagrange multiplier λ

measures the sensitivity of the objective value function to the constraint (Boyd & Vandenberghe, 2004, p. 251), we can write:

$$\lambda = \frac{df(B)}{dB} = \lim_{\epsilon \to 0} \frac{f(B+\epsilon) - f(B)}{\epsilon}, \qquad \text{for some } \epsilon > 0 \tag{C.0.1}$$

Let us assume that we have an optimal feasible portfolio $\mathbf{y}^*$ that satisfies the budget constraint. If we increase the right hand side of the constraint by some $\epsilon > 0$, we get,

$$\begin{aligned} \mathbf{z}^{\mathbf{T}}\mathbf{y}^* &\leq B + \epsilon \\ z_1 y_1 + z_2 y_2 + \cdots + 1 y_n &\leq B + \epsilon \end{aligned} \tag{C.0.2}$$

$\mathbf{y}^*$ is no longer feasible for the new budget, because for a portfolio to be feasible, the budget constraint must be satisfied with equality. So, to satisfy (C.0.2) with equality we can assign the additional ϵ funds to the risk-free asset, i.e. $z_1y_1 + z_2y_2 + \cdots + 1(y_n + \epsilon) = B + \epsilon$.

We can denote this new portfolio as $\tilde{\mathbf{y}} = [y_1, y_2, ..., y_n + \epsilon]^T$, which is feasible for $B + \epsilon$.

In this case, the wealth in each scenario can be represented as $A\mathbf{y}^* + R\epsilon\, \mathbf{l}_n$, where $\mathbf{1}_n$ is a $n \times 1$ vector of ones.

We can rewrite (C.0.1) as follows:

$$\begin{aligned} \lambda = \frac{df(B)}{dB} &= \lim_{\epsilon \to 0} \frac{f(B+\epsilon) - f(B)}{\epsilon} \\ &\geq \lim_{\epsilon \to 0} \frac{\mathbb{E}_p\left[\mathcal{U}(A\mathbf{y}^* + R\epsilon\, \mathbf{l}_n)\right] - \mathbb{E}_p[\mathcal{U}(A\mathbf{y}^*)]}{\epsilon} \\ &= \lim_{\epsilon \to 0} \frac{\mathbb{E}_p[\mathcal{U}(A\mathbf{y}^* + R\epsilon\, \mathbf{l}_n) - \mathcal{U}(A\mathbf{y}^*)]}{\epsilon} \\ &= \mathbb{E}_p\left[\lim_{\epsilon \to 0} \frac{\mathcal{U}(A\mathbf{y}^* + R\epsilon\, \mathbf{l}_n) - \mathcal{U}(A\mathbf{y}^*)}{\epsilon}\right] \end{aligned}$$

Multiplying the numerator and denominator by R gives us

$$\lambda = \frac{df(B)}{dB} = R\,\mathbb{E}_p\left[\lim_{\epsilon \to 0} \frac{\mathcal{U}(A\mathbf{y}^* + R\epsilon\, \mathbf{l}_n) - \mathcal{U}(A\mathbf{y}^*)}{R\,\epsilon}\right]$$

$$= R\,\mathbb{E}_p\left[\mathcal{U}'(\mathcal{A}\mathbf{y}^*)\right] \geq R \min\left[\mathcal{U}'(\mathcal{A}\mathbf{y}^*)\right]$$

Since $R > 0$ and $\mathcal{U}'(\bullet) > 0$, $R \min\left[\mathcal{U}'(\mathcal{A}\mathbf{y}^*)\right] > 0 \Longrightarrow \lambda > 0$ □

Appendix D

Theorems of the Alternative

D.1 Farkas' Lemma

For any points $a^1, a^2, ..., a^m$ and c in $\mathbb{R}^n$, exactly one of the following systems has a solution:

$$\sum_{i=1}^{m} u_i \, a^i = c, \quad 0 \leq u_1, u_2, ...u_m \tag{D.1.1}$$

$$\left\langle a^i, x \right\rangle \leq 0 \;\text{ for } i = 1, 2, ..., m, \quad \langle c, x \rangle > 0, \;\; x \in \mathbb{R}^n \tag{D.1.2}$$

Please refer to Borwein and Lewis (2006), pg 24 for proof.

Appendix E

Matlab Code

E.1 Example 1: Simple 3 Asset case

example1.m

```
delete ('example1.txt')
clear all
diary('example1.txt')
% ========================================================================
% Inputs
% ========================================================================

% eta=2;

A =[1.0500,4.0000,3.0000;1.0500,6.0000,10.0000;1.0500,7.0000,12.0000];

z =[1;5;7];

p =[0.2859;0.4295;0.2846];

r=0.05;

q=linsolve(A',z*(1+r))
B=100;

% ========================================================================
% Analytical Solution for the Forward Problem
```

```
% ========================================================================
% reassigning the constant values
% Ax=w. w=[w1 w2 w3]; Refer to equation

a=p(1); b=p(2); c=p(3); u=q(1); s=q(2); t=q(3);

% coefficient vectors
w1coeffs=[sqrt((a*s)/(u*b)); -sqrt((a*s)/(u*b)); sqrt((a*s)/(u*b));...
          -sqrt((a*s)/(u*b))];
w3coeffs=[sqrt((c*s)/(t*b)); sqrt((c*s)/(t*b)); -sqrt((c*s)/(t*b));...
          -sqrt((c*s)/(t*b))];

% possible values of w2
w2=1.05*B*(u*w1coeffs+t*w3coeffs+s).^-1
w1=[sqrt((a*s)/(u*b)); -sqrt((a*s)/(u*b)); sqrt((a*s)/(u*b));...
        -sqrt((a*s)/(u*b))].*w2
w3=[sqrt((c*s)/(t*b)); sqrt((c*s)/(t*b)); -sqrt((c*s)/(t*b));...
        -sqrt((c*s)/(t*b))].*w2;
format long, w3
format short

lambda1=(w1.^2*u+w2.^2*s+w3.^2*t).^-1

%check to see if budget constraint is satisfied
u*w1+s*w2+t*w3

%evaluate objective at stationary points
f=1-(w1.^-1*a+w2.^-1*b+w3.^-1*c)
[val, ind]=max(f);

%find the optimum and the corresponding value of lambda
w=[w1(ind) w2(ind) w3(ind)]'
lam=lambda1(ind)
x0=inv(A)*w;
format long, x0
format short

% ========================================================================
% Checking with Matlab Optimizer
% ========================================================================
```

```
% there is a problem with this. It depends on which starting vector I use.
% If I use the optimum from above, the optimizer tells me it is optimal.
% If I use a different starting vector, I get a different answer.

[w1,FVAL,EXITFLAG,OUTPUT,LAMBDA] = ...
                 fmincon(@(w)rnutil(w,p,2),w,[],[],q', B*1.05, [] ,[], [])

if (w1~=w)
    display('Something is wrong. Check the calculations.')
end

% ========================================================================
% Solving the Inverse Problem
% ========================================================================

U=UPrime('power',A,x0,2);
C=A'*U;
lam1=(1+r)*lam;
lb=[0 0 0]'; ub=[1 1 1]';

% inverse problem with the feasible region set as inequalities.
% Matlab cannot find a feasible point if I use strict equalities.
% However, when I the optimization is done, the constraints are satisfied.
[p_var,FVAL,EXITFLAG,OUTPUT,LAMBDA] = ...
 FMINCON(@(p_var)distToPrior(p_var,p),q,C,lam1*z,ones(1,3),1,lb,ub)

% inverse problem in the absence of lambda. But setting the feasible region
% as strict equalities without the sum to 1 constrant.
% normalized solution is the same as the previous p vector.
% This is not the case when the # of columns is greater than the # of rows.
% example 2 deals with that.

[h,FVAL,EXITFLAG,OUTPUT,LAMBDA] = ...
    FMINCON(@(p_var)distToPrior(p_var,p),q,[],[],C,z,lb,[])

p_inv=h/sum(h);

Comparison=[p_inv p];  diary('off')
```

E.2 Example 2: Puts on the S&P 500 Index

example2A.m

```
% ========================================================================
% Inputs
% ========================================================================
% load the data matrix
load data.mat

% scenario matrix
A;

% price vector
z;

% initial budget constraint.
B=1000;

% risk-free rate
r=A(1,1)-1;

% original vector of probabilities associated with the scenarios
% used to calculate the optimal portfolio in the forward problem
p0=[0 0 0 0 0 0.0002 0.0016 0.0153 0.0638 0.1462 0.2313,...
    0.2414 0.1698 0.0831 0.0329 0.0110 0.0026 0.0008 0 0 0]';

% prior vector with heavy tails
pht=[0.00872 0.01422 0.01922 0.03112 0.03862 0.05400 0.06910 0.07830,...
    0.08260 0.08630 0.07660 0.08210 0.07260 0.06900 0.05550 0.05070,...
    0.04020 0.03140 0.02310 0.01660 0]';

p=p0;

% get dimensions of the scenario matrix
[s,n]=size(A);

% ========================================================================
% Risk-neutral Probabilities
% ========================================================================
```

```
% reset solver and solver precision
cvx_solver sedumi
cvx_precision default

% find the state-price vector using cvx
cvx_begin
variable psi(s)
dual variable bet1
minimize (norm(psi))
    subject to
        A'*psi==z;
        psi>=zeros(s,1);
cvx_end

% check to see if the state-price vector exists
if (~strcmp(cvx_status, 'Solved')||min(psi)<=0)
 Display('Market is not arbitrage free')
 min(psi)
 break;
end

% normalize the state-price vector to get the risk-neutral probabilities
q=psi/sum(psi);

% ========================================================================
% Solve the forward problem
% ========================================================================

% find the basis for the null space of A'
V=null(A');
[m,s]=size(V');

% construct the constraints to solve the optimization using the attainable
% payoff method of calculating the optimal portfolio.

Aeq=[q'; V'];
beq=[B*(1+r); zeros(m,1)];

% set the risk-aversion parameter
```

```
eta=2;

% solve the forward optimization problem.
cvx_begin
variable w1(s)
dual variable gam1
maximize (-1*rnutil(w1,p,2))
    subject to
        gam1: Aeq*w1 == beq;
              w1>=0;
cvx_end

% get the optimal portfolio and the associated lagrange multiplier value
x1=pinv(A)*w1;

% note that we need to use the absolute value due to the difference in
% formulation of the forward problem. The CVX solver sets the Lagrangian
% as, (objective)+lambda*(constraint), whereas we set the Lagrangian as,
% (objective) - lambda*(constraint).
lam1=abs((1+r)*gam1(1));


% ========================================================================
% Solving the inverse using (3.3.9)
% ========================================================================

% get the diagonal matrix of first derivatives associated with the optimal
% portfolio, scenario matrix and the risk-aversion parameter.
U1=UPrime('power',A,x1,eta);

% The matrix of constraints
C=A'*U1;

% set the lower bounds and upper bounds on the probability values
lb=zeros(21,1); ub=ones(21,1);

lam=lam1;

% solve the least squares minimization in (3.3.9) using CVX
cvx_begin
```

```
    variable p1(21)
    minimize (norm(p1-p))
    subject to
        C*p1<=lam*z;
        sum(p1)==1;
        p1>=0;
cvx_end

% solve the least squares minimization in (3.3.9) using CPLEX
[pInv,resnorm,residual,exitflag,output,lambda] = ...
    cplexlsqnonneglin(eye(21),p,C,lam*z, ones(1,21),1);


% plot the two values, pInv and p1 against the original vector
plot(p, 'k--')
hold on
plot(pInv, 'r*')
plot(p1, 'r.')
legend('prior vector', 'implied vector')

% =========================================================================
% Solving the inverse using (3.3.10)
% =========================================================================

% code to solve the inverse optimization using (3.3.10)
% initial epsilon
eps=10^-10;
bol=false;
ctr=1;
P2=[p];
while(bol==false)
    % suppress CVX output
    cvx_quiet(true)
    % solve inverse problem for given epsilon
    cvx_begin
        variable p1(21)
        minimize (norm(p1-p))
        subject to
            C*p1<=lam*z+eps;
            C*p1>=lam*z-eps;
```

```
            sum(p1)==1;
            p1>=0;
    cvx_end

    % increment the value of epsilon
    eps=eps*10;

    % check to see if the optimization problem was solved successfully.
    % if it was solved, save the resulting vector and check to see if the
    % resulting vector is within a certain tolerance of the prior.
    if(strcmp(cvx_status,'Solved'))
        eps;
        cvx_optval
        ctr
        if (floor(log(cvx_optval)/log(10))<=-7 || ctr==7)
            bol=true;
        end
        P2=[P2 p1];
    end

    % increase counter
    ctr=ctr+1;
end

% plot the resulting vectors. The plots in the document were edited
% (colour and legend) using the plot editor in MATLAB.
plot(P2)

% solving (3.3.10) with the tolerance as a percentage of the RHS
P3=[];
tol=0.001*lam*z; % setting the tolerance as 0.1% of RHS
bol=false;
cvx_quiet(true)
cvx_begin
    variable p2(21)
    minimize (norm(p2-p))
    subject to
        C*p2<=lam*z;
        C*p2>=lam*z-tol;
        sum(p2)==1;
```

```
        p2>=0;
cvx_end
P3=[p p2];

% solving again with tolerance as 1% of RHS
tol=0.01*lam*z;
bol=false;
cvx_quiet(true)
cvx_begin
variable p2(21)
minimize (norm(p2-p))
subject to
    C*p2<=lam*z;
    C*p2>=lam*z-tol;
    sum(p2)==1;
    p2>=0;
cvx_end
P3=[P3 p2];

% solving with tolerance as 10%, 20%, 40%, 60%, 80% and 100% of RHS
ctr=1;
tol=0.1*lam*z;
inc=tol;
while(bol==false)
    cvx_quiet(true)
    cvx_begin
    variable p2(21)
    % solve the least squares optimization
    minimize (norm(p2-p))
    subject to
        C*p2<=lam*z;
        C*p2>=lam*z-tol;
        sum(p2)==1;
        p2>=0;
    cvx_end

    %increment tolerance
    if(ctr>=2)
        tol=tol+2*inc;
    else
```

```
        tol=tol+inc;
    end

    % check to see if the optimization problem was solved successfully.
    % if it was solved, save the resulting vector and check to see if the
    % resulting vector is within a certain tolerance of the prior.
    if(strcmp(cvx_status,'Solved'))
       if (floor(log(norm(p2-p))/log(10))<=-7 || ctr==7)
            bol=true;
        end
        P3=[P3 p2];
    end
    % increment counter
    ctr=ctr+1;

end

% ========================================================================
% Solving the inverse using (3.3.11)
% ========================================================================

% solve the optimization using (3.3.11)
cvx_quiet(false)
cvx_begin
    variable h(21)
    minimize (norm(h-p))
    subject to
        C*h==z;
        h>=0
cvx_end
% normalize the vector
p2=h/sum(h);

figure
plot(p)
hold on
plot(p2, 'r--')
legend('prior','implied')

% ========================================================================
```

```
% Example 2 Revisited: A Different Prior Distribution
% Solving the inverse using (3.3.12)
% ===========================================================================
% calculating the value of lambda
cvx_quiet(false)
cvx_begin
    variable h(21)
    minimize 0
    subject to
        C*h==z;
        h>=0
cvx_end
lam3=1/sum(h);

P4=[];

% setting the tolerance as 0.1% of RHS.
tol=0.001*lam3*z;
bol=false;
cvx_quiet(true)
cvx_begin
    variable p2(21)
    minimize (norm(p2-pht))
    subject to
        C*p2<=lam3*z;
        C*p2>=lam3*z-tol;
        sum(p2)==1;
        p2>=0;
cvx_end

P4=[pht p2];

% setting the tolerance as 1% of RHS.
tol=0.01*lam3*z;
bol=false;
cvx_quiet(true)
cvx_begin
variable p2(21)
minimize (norm(p2-pht))
subject to
```

```
    C*p2<=lam3*z;
    C*p2>=lam3*z-tol;
    sum(p2)==1;
    p2>=0;
cvx_end
P4=[P4 p2];

% setting the tolerance as 10%, 20%, 40%, 60%, 80% and 100% of RHS.
ctr=1;
tol=0.1*lam3*z;
inc=tol;
while(bol==false)
    cvx_quiet(true)
    cvx_begin
    variable p2(21)
    minimize (norm(p2-pht))
    subject to
        C*p2<=lam3*z;
        C*p2>=lam3*z-tol;
        sum(p2)==1;
        p2>=0;
    cvx_end

    % increment tolerance
    if(ctr>=2)
        tol=tol+2*inc;
    else
        tol=tol+inc;
    end

    % check to see if the optimization problem was solved successfully.
    % if it was solved, save the resulting vector and check to see if the
    % resulting vector is within a certain tolerance of the prior.
    if(strcmp(cvx_status,'Solved'))
        if (floor(log(norm(p2-pht))/log(10))<=-5 || ctr==6)
            bol=true;
        end
        P4=[P4 p2];
    end
```

```
    ctr=ctr+1;

end

% ========================================================================
% Example 2 Revisited: Minimum or Maximum Implied Probability of an Event
% Solving the inverse using (3.3.13)
% ========================================================================

tol=0.001*lam3*z;
inc=tol;
p_max=zeros(21,1);
p_min=zeros(21,1);


for i=1:s
    c=zeros(21,1);
    c(i)=1;
    %--------------------------------------------------------------------
    % Solving for the maximum probability of each event
    %--------------------------------------------------------------------
    bol=false;
    ctr=1;
    tol=0.01*lam3*z;
    inc=tol;
    while(bol==false)
        cvx_quiet(true)
        cvx_begin
        variable p1(21)
        maximize (c'*p1)
        subject to
            C*p1<=lam3*z;
            C*p1>=lam3*z-tol;
            sum(p1)==1;
            p1>=0;
        cvx_end

        tol=tol+inc;

        if(strcmp(cvx_status,'Solved')||ctr==100)
```

```
            bol=true;
            ctr
            p_max(i)=cvx_optval;
        end
        ctr=ctr+1;
    end
    %---------------------------------------------------------------------
    % Solving for the maximum probability of each event
    %---------------------------------------------------------------------
    bol=false;
    ctr=1;
    tol=0.01*lam3*z;
    inc=tol;
    while(bol==false)
        cvx_quiet(true)
        cvx_begin
        variable p2(21)
        minimize (c'*p2)
        subject to
            C*p2<=lam3*z;
            C*p2>=lam3*z-tol;
            sum(p2)==1;
            p2>=0;
        cvx_end

        tol=tol+inc;

        if(strcmp(cvx_status,'Solved')||ctr==100)
            bol=true;
            ctr
            p_min(i)=cvx_optval;
        end
        ctr=ctr+1;
    end
end

figure
plot(p_max, 'g-.')
hold on
plot(p, 'k-.')
```

```
plot(p_min, 'r-.')
legend('Maximum probabilities','Original probabilities','Minimum probabilities')
```

E.3 Example 2: Puts on the S&P 500 Index in the Mean-Variance Setting

example2mv.m

```
% ========================================================================
% Inputs
% ========================================================================
load data.mat;
% total return on the risk-free asset
r=A(1,1);

% original vector of probabilities associated with the scenarios
% used to calculate the optimal portfolio in the forward problem
p=[0 0 0 0 0 0.0002 0.0016 0.0153 0.0638 0.1462 0.2313,...
    0.2414 0.1698 0.0831 0.0329 0.0110 0.0026 0.0008 0 0 0]';

% prior vector with heavy tails
pht=[0.00872 0.01422 0.01922 0.03112 0.03862 0.05400 0.06910 0.07830,...
    0.08260 0.08630 0.07660 0.08210 0.07260 0.06900 0.05550 0.05070,...
    0.04020 0.03140 0.02310 0.01660 0]';


% ========================================================================
% Calculating the returns matrix and the covariance matrix
% ========================================================================
Z=repmat(z',21,1);
eta=0.5;
A_=A./Z;
A_=A_(:,2:end);
Sig=cov(A_);
mu=A_'*p;

% ========================================================================
% Calculating the optimal portfolio and minimizing distance to original
% probability vector
% ========================================================================
x=0.5*inv(Sig)*(mu-r*ones(8,1));
x_opt=[1-sum(x); x];
```

```
g=p;

cvx_begin
    variable p1(21);
    minimize (norm(g-p1))
    subject to
        A_'*p1==(eta*r*ones(8,1) + Sig*x)/eta;
        p1>=0;
        sum(p1)==1;
cvx_end

figure
plot(p1)
hold on

% =========================================================================
% Minimizing the distance to a heavy tailed prior.
% =========================================================================
g=pht;

cvx_begin
    variable p2(21);
    minimize (norm(g-p2))
    subject to
        A_'*p2<=(eta*r*ones(8,1) + Sig*x)/eta;
        p2>=0;
        sum(p2)==1;
cvx_end

plot(p2, 'b:')
hold on
plot(p, 'r--')
plot(pht,'g-.')

% plots edited using the plot editor in MATLAB.

cond(A_)

% =========================================================================
% Minimizing the distance to original probability vector with tolerances.
```

```
% =========================================================================
epsilon = [0.01 0.05 0.1 0.5 1];

P5=[p];
g=p;

for i = 1:length(epsilon)
tol = epsilon(i);

cvx_begin
variable p2(21);
minimize (norm(g-p2))
subject to
A_'*p2<=(1+tol)*(eta*r*ones(8,1) + Sig*x)/eta;
A_'*p2>=(1-tol)*(eta*r*ones(8,1) + Sig*x)/eta;
p2>=0;
sum(p2)==1;
cvx_end
P5=[P5 p2];
end
plot(P5)
% plots edited for clarity using the plot editor in MATLAB.


% =========================================================================
% Minimizing the distance to heavy-tailed prior with tolerances.
% =========================================================================
epsilon = [0.01 0.05 0.1 0.5 1];

P6=[pht];
g=pht;

for i = 1:length(epsilon)
tol = epsilon(i);

cvx_begin
variable p2(21);
minimize (norm(g-p2))
subject to
A_'*p2<=(1+tol)*(eta*r*ones(8,1) + Sig*x)/eta;
```

```
A_'*p2>=(1-tol)*(eta*r*ones(8,1) + Sig*x)/eta;
p2>=0;
sum(p2)==1;
cvx_end
P6=[P6 p2];
end
plot(P6)
% plots edited for clarity using the plot editor in MATLAB.
```

E.4 Performance Tests

E.4.1 Incorporating Implied Views from a Fund Portfolio

pt1.m

```
% ======================================================================
% Inputs
% ======================================================================
load PTdata.mat

% calculate the sector returns
returns=Data(2:end-1,:)./Data(1:end-2,:)-1;

[f,g]=size(returns);

%setting the seed to replicate results.
sd = RandStream.create('mt19937ar','seed',859);

ind=randi(sd,2,f,1);


% fit the returns to a mixture of normals model
obj = gmdistribution.fit(returns,2,'Start',ind);
ComponentMeans = obj.mu;
ComponentCovariances = obj.Sigma;
MixtureProportions = obj.PComponents;

% number of scenarios.
s=150;

% calculate the monthly risk-free rate.
r=0.0091*31/365;

% price at time 0
P0=Data(end-1,:);

% price after 1 month
P1=Data(end,:);
```

```
% separating the two mean vectors
mu1=ComponentMeans(1,:);
mu2=ComponentMeans(2,:);

% separating the two variance vectors
sig1=sqrt(diag(ComponentCovariances(:,:,1)))';
sig2=sqrt(diag(ComponentCovariances(:,:,2)))';

%cholesky decomposition of the two covariance matrices for correlation.
C1=chol(ComponentCovariances(:,:,1)); C2=(ComponentCovariances(:,:,2));

% ========================================================================
% Generating Scenarios
% ========================================================================
% Get the points in the Sobol Set in dimension 7
P=sobolset(7,'Skip',1e3);
% get the first 150 points
V=net(P,s);

% use the inverse normal cdf to get normal values
Z=norminv(V);

% establish the correlation for both normal models.0
Z1=C1*Z'; Z2=C2*Z';


% get a vector of standard uniform random variables to get a 50-50
% mixture
U=rand(sd,s,7);

% get the proper mixture of means and standard deviations
mu=repmat(mu2,s,1).*(U>0.5)+repmat(mu1,s,1).*(U<=0.5);
sig=repmat(sig1,s,1).*(U>0.5)+repmat(sig2,s,1).*(U<=0.5);

% get the correct mixture of normal variables
ZZ=Z2'.*(U>0.5);
Z=Z1'.*(U<=0.5);
Z=Z+ZZ;

% generate asset values.
```

```
Y1=(mu-sig.^2/2)+sig.*Z;
S=repmat(P0,s,1).*exp(Y1);

% generate scenario matrix.
A=[(1+r)*ones(s,1) S];

% ========================================================================
% Fund Data and Sorting the Scenarios by Wealth in Ascending Order
% ========================================================================
% portfolio weights
x=[0.43;12.80;8.31;23.85;27.46;16.10;4.44;6.61]/100;

% initial budget
B=227000000;

% initial price vector.
z=[1; P0'];

% vector of portfolio allocation by volume.
y=x*B./z;

% calculating wealth in each scenario
W=A*y;
scenarios=[W A];

% sorting scenarios by wealth.
scenarios=sortrows(scenarios,1);
A=scenarios(:,2:end);


% ========================================================================
% Retrieving implied views and the risk-neutral probabilities
% ========================================================================

% Calculating our matrix of first derivatives
U=UPrime('power',A,y,1);
% calculating the coefficient matrix for constraints
C=A'*U;

% setting the vector of empirical probabilities
```

```
p0=ones(s,1)/s;

% solving for the implied probabilities
cvx_quiet(false)
cvx_begin
variable h(s)
minimize (norm(p0-h))
subject to
C*h<=z+0.05*z;
C*h>=z-0.05*z;
h>=0
cvx_end

% normalizing to get the vector of implied probabilities
p4=h/sum(h);

% solving for the state-price vector.
cvx_begin
variable psi(s)
dual variable bet1
minimize (norm(psi))
    subject to
        A'*psi<=z+0.05*z;
        A'*psi>=z-0.05*z;
        psi>=zeros(s,1);
cvx_end
q=psi/sum(psi);

% ======================================================================
% Solving the various portfolio selection models
% ======================================================================
% solving the forward problem with the new mixture of probabilities.

V=null(A');
[m,s]=size(V');
Aeq=[q'; V'];
beq=[B*(1+r); zeros(m,1)];
n=9;

% new probability : 50% implied + 50% empirical
```

```
p=0.5*p0+0.5*p4;

cvx_begin
variable w2(s)
dual variable gam1
maximize (-1*rnutil(w2,p,1))
    subject to
        gam1: Aeq*w2 == beq;
              w2>=0;
cvx_end

% new optimal portfolio allocation by volume (mixture)
y2=pinv(A)*w2;

% corresponding portfolio weights
x2=y2.*z/B;

lam1=abs((1+r)*gam1(1));


% solving the above problem, but with the empirical probabilities only
cvx_begin
variable w_emp(s)
dual variable gam1
maximize (-1*rnutil(w_emp,p0,1))
    subject to
        gam1: Aeq*w_emp == beq;
              w_emp>=0;
cvx_end

% new optimal portfolio allocation by volume (empirical)
y_emp=pinv(A)*w_emp;

% corresponding portfolio weights
x_emp=y_emp.*z/B;


Prices=repmat(z',s,1);
eta=0.5;
A_=A./Prices;
```

```
A_=A_(:,2:end);

Sig1=cov(A_);

% Solving for the minimum-variance portfolio.
cvx_begin
variable mv(7)
dual variable bet1
minimize (mv'*Sig1*mv)
    subject to
        bet1: sum(mv)==1;

cvx_end

%portfolio weights
x_mv=[0; mv];

% corresponding optimal portfolio allocation by volume (minimum-variance)
y_mv=x_mv*B./z;


% Solving for the ERC portfolio.
c=5;
cvx_begin
variable erc(8)
dual variable bet1
minimize (erc'*[zeros(1,8);zeros(7,1) Sig1]*erc)
    subject to
        sum(log(erc))>=c;
erc>=0;
cvx_end
% portfolio weights
x_erc=erc/sum(erc);

% corresponding optimal portfolio allocation by volume (ERC)
y_erc=x_erc*B./z;

% equally weighted portfolio.
x_ew=ones(8,1)/8;
y_ew=x_ew*B./z;
```

```
% ==========================================================================
% Comparing out-of-sample performance.
% ==========================================================================

z1=[(1+r); P1'];

val = z1'*y
val_mixed = z1'*y2
val_emp= z1'*y_emp
val_mv= z1'*y_mv
val_erc= z1'*y_erc
val_ew= z1'*y_ew

% NOTE THAT ALL PLOTS IN THE DOCUMENT WERE CREATED USING THE PLOT EDITOR
% IN MATLAB.
```

E.4.2 Incorporating Implied Views from a Constant Mix Portfolio

pt2.m

```
% ===========================================================================
% Inputs and Scenarios
% ===========================================================================
clear all;
load data6040.mat;

Mn=[]; Sig=[]; Prop = []; M=480; SP=SP500_RF(:,1); RF=SP500_RF(:,2);
s=150; OPTIONS = STATSET('MaxIter',300);


Data = SP(2:end)./SP(1:end-1);

figure
hist(Data,50); title('Distribution of S&P 500 returns (1962-2010)')

figure
D1=Data(480:587);
hist(D1,25); title('Distribution of S&P 500 returns (2002-2010)')

figure
D2=Data(480:540);
hist(D2,10); title('Distribution of S&P 500 returns (2002-2006)')

figure
D3=Data(540:587);
hist(D3,10); title('Distribution of S&P 500 returns (2007-2010)')



%Summary Statistics for the Data
M1_D=[mean(Data) mean(D1) mean(D2) mean(D3)];
M2_D=[var(Data) var(D1) var(D2) var(D3)];
M3_D=[skewness(Data) skewness(D1) skewness(D2) skewness(D3)];
M4_D=[kurtosis(Data) kurtosis(D1) kurtosis(D2) kurtosis(D3)];
```

```
[M1_D; M2_D; M3_D; M4_D]

% calculate the returns for the rolling window. At each snap-shot, the
% return means, variances and proportions are saved in a new column in
% their respective matrices. Mn = means, Sig=variances, Prop = proportion

reset(RandStream.getDefaultStream);
ind=randi(2,M-2,1);

ctr1=M;
for i=1:(length(SP)-M)
% get the 480 pieces of data needed for estimation.
    est=SP(i:ctr1);
    returns=(est(2:end-1,1)./est(1:end-2))-1;

    %fit a mixture-of-normals model to the data
    obj = gmdistribution.fit(returns,2,'Start',ind,'Options',OPTIONS);
    ComponentMeans = obj.mu;
    ComponentCovariances = obj.Sigma;
    ComponentCov=[ComponentCovariances(:,:,1);ComponentCovariances(:,:,2)];
    MixtureProportions = obj.PComponents';

    %increment counter
    ctr1=ctr1+1;

%save the fitted parameter values
    Mn=[Mn ComponentMeans];
    Sig=[Sig sqrt(ComponentCov)];
    Prop=[Prop MixtureProportions];
end

% calculate the scenario set using a Sobol Set with dimension (588-480)
[m,n]=size(Mn);
Q=sobolset(n,'Skip',1e3);
V=net(Q,s);
size(V);
Z=norminv(V);

%setting the seed to replicate results.
sd = RandStream.create('mt19937ar','seed',5000);
```

```
% standard uniform random variables
U=rand(sd,s,n);

prop=repmat(Prop(1,:),s,1);

% getting the proper mix of parameters for the scenario forecast
mu=repmat(Mn(2,:),s,1).*(U>prop)+repmat(Mn(1,:),s,1).*(U<=prop);
sig=repmat(Sig(2,:),s,1).*(U>prop)+repmat(Sig(1,:),s,1).*(U<=prop);
Z1=sig.*Z;

Y1=exp((mu-sig.^2/2)+Z1);

% ==========================================================================
% Generating Scenarios and conducting the out-of-sample tests
% ==========================================================================
% The constant mix portfolio
x=[0.4; 0.6];

% starting budget at each iteration.
B=1000;

% risk-aversion parameter for the investor with a constant mix portfolio.
eta=2;

% matrix to store the results from the tests
Results=[]; P=[]; X=[]; X2=[]; Xemp=[]; Xew=[];
X_=[]; X2_=[]; Xemp_=[]; Xew_=[];

%setting initial counter values.
ctr=M;
index=M+1;

%begin calculations
for i=1:n
% initial starting price
    z=[100; SP(ctr)];

    % vector of portfolio allocation by volume
y=x*B./z;
```

```
X=[X x];

% Scenario matrix.
    A=[100*ones(s,1)*(1+RF(M)) repmat(SP(ctr),s,1).*Y1(:,i)];

    % sorting scenarios by wealth.
    W=A*y;
    scenarios=[W A];

    scenarios=sortrows(scenarios,1);
    A=scenarios(:,2:end);

    %calculating the matrix of first derivatives
    U=UPrime('power',A,y,eta);

    % calculating our coefficient matrix for constraints
    C=A'*U;

% empirical probability vector
    p0=ones(s,1)/s;

% solving for the implied probabilities.
    cvx_quiet(true)
    cvx_begin
        variable h(s)
        minimize (norm(p0-h))
        subject to
        C*h<=z+0.1*z;
        C*h>=z-0.1*z;
        h>=0
    cvx_end
    p4=h/sum(h);
    lam2=1/sum(h);

    P=[P p4];

    % solving for the risk-neutral probabilities
    cvx_begin
        variable psi(s)
        dual variable bet1
```

```
        minimize (norm(psi))
            subject to
                A'*psi==z;
                psi>=zeros(s,1);
    cvx_end
    q=psi/sum(psi);

% check to see if market is arbitrage-free.
    if(sum(psi<=10^-16)>0)
        disp(['market is not arbitrage-free ', i])
    end

% solving the forward optimization problem with the mixture of
% the implied and empirical probabilities

    V=null(A');
    [m,s]=size(V');
    Aeq=[q'; V'];
    beq=[B*(1+RF(ctr)); zeros(m,1)];

   % mixing the two probabilities
    p=0.5*p0+0.5*p4;


    cvx_begin
        variables w2(s)
        dual variable gam1
        maximize (-1*rnutil(w2,p,eta))
            subject to
            gam1: Aeq*w2 == beq;
                      w2>=0;
    cvx_end
    % new optimal portfolio allocation by volume (mixture)
y2=pinv(A)*w2;

% corresponding portfolio weights
x2=y2.*z/B;
X2=[X2 x2];

    lam1=abs((1+RF(ctr))*gam1(1));
```

```
% solving the above problem using the empirical probabilities
    cvx_begin
        variable w_emp(s)
        dual variable gam1
        maximize (-1*rnutil(w_emp,p0,eta))
            subject to
                gam1: Aeq*w_emp == beq;
                      w_emp>=0;
    cvx_end
    % new optimal portfolio allocation by volume (empirical)
y_emp=pinv(A)*w_emp;

% corresponding portfolio weights
x_emp=y_emp.*z/B;
Xemp=[Xemp x_emp];

% portfolio allocation by volume (equally weighted)
    y_eq=[0.5;0.5]*B./z;
Xew=[Xew [0.5;0.5]];

% price one month into the future.
    z1=[A(1,1); SP(index)];

% increment counters
    ctr=ctr+1;
    index=index+1;

% save the results
    Results=[Results;[ctr z1'*y z1'*y2  z1'*y_emp z1'*y_eq]];
    X_=[X_ z1.*y/B];
    X2_=[X2_ z1.*y2/B];
    Xemp_=[Xemp_ z1.*y_emp/B];
    Xew_=[Xew_ z1.*y_eq/B];

end

[T,num]=size(Results);

mu1=mean(Results(:,2:end))
```

```
mu2=mean(Results(1:60, 2:end))
mu3=mean(Results(61:108, 2:end))


Turnover=[sum(sum(abs(X_(:,1:end-1)-X(:,2:end))))/(107)
sum(sum(abs(X2_(:,1:end-1)-X2(:,2:end))))/(107)
sum(sum(abs(Xemp_(:,1:end-1)-Xemp(:,2:end))))/(107)
sum(sum(abs(Xew_(:,1:end-1)-Xew(:,2:end))))/(107)]';

% Summary statistics for realized out-of-sample returns
m1=[mean(Results(:,2:end))/B];
m2=[var(Results(:,2:end)/B)];
m3=[skewness(Results(:,2:end)/B)];
m4=[kurtosis(Results(:,2:end)/B)];

[m1;m2;m3;m4]

rt_rf=(Results(:,2:end))/1000-1 - repmat(RF(481:588),1,4);
R=mean(rt_rf);
Si=sqrt(m2);
si=repmat(Si,T,1);
SR=R./Si;
S=repmat(SR,T,1);

VarSharpe=(1/T)*sum((S.^2).*repmat(m4,T,1)./(4*si.^4) ...
-S.*((rt_rf).*((Results(:,2:end)/1000) -...
 repmat(m1,T,1)).^2-(rt_rf).*si.^2)./si.^3+(rt_rf.^2)./si.^2 ...
 - 2*S.*rt_rf./si + 3*S.^2/4);

c=norminv(.9725);
CI=[SR-c*sqrt(VarSharpe/T); SR; SR+c*sqrt(VarSharpe/T)];



figure
plot(Results(:,2), 'b-.')
hold on
plot(Results(:,3), 'k')
plot(Results(:,4), 'r--')
plot(Results(:,5), 'k:')
```

```
legend('60-40','mixed','empirical','50-50')

figure
hist(Results(:,2))
title('Distribution of the 60-40 portfolio returns (2002-2010)')

figure
hist(Results(:,3))
title('Distribution of mixed probability portfolio returns (2002-2010)')

figure
hist(Results(:,4))
title('Distribution of empirical probability portfolio returns (2002-2010)')

figure
hist(Results(:,5))
title('Distribution of the 50-50 portfolio returns (2002-2010)')
```

E.5 Helper Functions

UPrime.m

```
function [U]=UPrime(utilfunc,A,x,ra_param)
%function calculates the first derivative of the utility function
% identified by 'utilfunc' (string).
% exponential utility denoted by 'exp'. Power utility denoted by
% 'power'. mean-variance denoted by 'mv'

%size of the scenario matrix. s scenarios and n assets.
    [s,n]=size(A);
    if (strcmp(utilfunc,'exp'))
        for i=1:s
            u(i)=expUtilPrime(ra_param,A(i,:),x);
        end
        U=diag(u);
    elseif (strcmp(utilfunc,'power'))
        for i=1:s
            u(i)=logUtilPrime(ra_param,A(i,:),x);
        end
        U=diag(u);
    end
end
```

logUtilPrime.m

```
function [u]=logUtilPrime(eta,a,x)
    % note: if a'*x <0 and eta<1, we have a problem. We cannot take even
    % roots of negative numbers. Also, we cannot take logarithms of
    % negative numbers. And when a'x=0, we get u=Inf.
    % eta is the risk aversion parameter.
    if(eta==1)
        u=real((a*x)^-1);
    else
        u=real((a*x)^-eta);
    end
```

expUtilPrime.m

```
function [u]=expUtilPrime(alpha,a,x)
    %derivative of the exponential utility function.
    % alpha is the risk aversion parameter.
    u=alpha*exp(-alpha*a'*x);
```

rnutil.m

```
function val=rnutil(w,p,eta)
% this is the power utility function.
  if eta==1
val = -1*p'*log(w);
  else
  val=-1*p'*pow_p(w,(1-eta))/(1-eta);
  end
end
```

References

Ahuja, R. K., & Orlin, J. B. (2001). Inverse Optimization. *Operations Research*, *49*(5), 771-783. (http://www.jstor.org/stable/3088574)

Arrow, K. J., & Enthoven, A. C. (1961, Oct.). Quasi-concave programming. *Econometrica*, *29*(4), 779-800. (http://www.jstor.org/stable/1911819) 8

Aster, R. C., Borchers, B., & Thurber, C. H. (2005). *Parameter estimation and inverse problems* (Vol. 90). Amsterdam ; Boston: Elsevier Academic Press. 7, 40, 41

Bernoulli, D. (1954, Jan.). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23-36. (Originally published in 1738 and translated by Dr. Lousie Sommer) 8

Bertsekas, D. P. (1995). *Nonlinear programming*. Athena Scientific. (ISBN:9781886529144) 19, 73

Best, M. J. (2010). *Portfolio Optimization*. Boca Raton, FL: Chapman and Hall/CRC. 10, 12, 13, 47

Best, M. J., & Grauer, R. R. (1985, Mar.). Capital Asset Pricing Compatible with Observed Market Value Weights. *The Journal of Finance*, *40*(1), 85-103. 4, 5

Best, M. J., & Grauer, R. R. (1991). On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *The Review of Financial Studies*, *4*(2), 315-342. 1, 13

Björck, Å. (1996). *Numerical methods for least squares problems*. SIAM. (ISBN:9780898713602) 71

Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, *48*(5), 28-43. (http://www.jstor.org/stable/4479577) 1, 2

Borwein, J. M., & Lewis, A. S. (2006). *Convex Analysis and Nonlinear Optimization: Theory and Examples*. USA: Springer. 76

Boyd, S. P., & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press . 74, 75

Broadie, M. (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research*, *45*, 21-58. 1, 14

Cheney, E., & Kincaid, D. (2007). *Numerical mathematics and computing*. Brooks/Cole. (ISBN: 9780495114758) 40

Christie, S. (2005, May). Is this Sharpe Ratio Useful in Asset Allocation? *MAFC Research Papers*(31). (`http://www.mafc.mq.edu.au`) 63

Dana, R.-A., & Jeanblanc-Picqué, M. (2003). *Financial markets in continuous time.* Springer. (ISBN:9783540434030) 21

Dembo, R. (1991). Scenario Optimization. *Annals of Operations Research*, *30*(1), 63-80. Available from `http://dx.doi.org/10.1007/BF02204809` (10.1007/BF02204809)

Dembo, R., & Rosen, D. (2000). The Practice of Portfolio Replication. *ALGO RESEARCH QUARTERLY*, *3*(2), 11-22. 3, 4

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*(5), 1915-1953. (`http://rfs.oxfordjournals.org/content/22/5/1915.full.pdf+html`) 1, 65

Disyatat, P., & Gelos, R. G. (2001). The Asset Allocation Of Emerging Market Mutual Funds [Working Paper Series]. SSRN. 2

Ehlers, L. (2002). Multiple public goods and lexicographic preferences: replacement principle. *Journal of Mathematical Economics*, *37*(1), 1 - 15. (`http://www.sciencedirect.com/science/article/pii/S0304406802000022`) 8

Engl, H. W. (2007, Jul.). Calibration problems An inverse problems view. In *Wilmott* (p. 16-20).

Engl, H. W., & Kügler, P. (2005). Nonlinear inverse problems: Theoretical aspects and some industrial applications. In H.-G. Bock et al. (Eds.), *Multidisciplinary methods for analysis optimization and control of complex systems* (Vol. 6, p. 3-47). Springer Berlin Heidelberg. (`http://dx.doi.org/10.1007/3-540-27167-8_1`) 7, 40

Fishburn, P. C. (1975, July). Axioms for Lexicographic Preferences. *The Review of Economic Studies*, *42*(3), 415-419. (Published by: Oxford University Press. `http://www.jstor.org/stable/2296854`) 8

Grant, M., & Boyd, S. (2011, Apr.). *CVX: Matlab software for disciplined convex programming, version 1.21.* `http://cvxr.com/cvx`. 26, 32, 35

Hartley, M. J. (n.d.). Neoclassical Econometrics: The Kernel. *Journal of Mathematical Analysis and Applications*(2). 3

Hartley, M. J. (1986). Calibration of macroeconomic models with incomplete data. *Computers and Mathematics with Applications: An International Journal Series A, Bellman Memorial Issue 6*, 769-776. 3

Hartley, M. J. (1994). A systems approach to the calibration of deterministic nonlinear simultaneous equation models with incomplete data. *Applied Stochastic Models and Data Analysis*. 3

Hartley, M. J., & Bakshi, G. S. (1998, Oct). Markowitz Models of Portfolio Selection:The Inverse Problem. *Advances in Investment Analysis and Portfolio Management*, *7*, 55-89. 3

Huang, J. (2001). *Sensitivity Analysis for Changes in The Covariance Matrix in Portfolio*

Optimization. Unpublished master's thesis, University of Waterloo. 1

Iyengar, G., & Kang, W. (2005). Inverse conic programming with applications. *Operations Research Letters*, *33*(3), 319 - 330. (http://www.sciencedirect.com/science/article/pii/S016763770400063X)

Jackwerth, J. C. (1999). Option Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review. *Journal of Derivatives*, *2*(7), 66-82. 2

Jagannathan, R., & Ma, T. (2003, Aug.). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *The Journal of Finance*, *58*(4), 1651-1683. (http://www.jstor.org/stable/3648224)

Joy, C., Boyle, P. P., & Tan, K. S. (1996, Jun.). Quasi-monte carlo methods in numerical finance. *Management Science*, *42*(6), 926–938. 19

Jünger, M., Liebling, T. M., Naddef, D., Nemhauser, G., Pulleyblank, W. R., Reinelt, G., et al. (2009). *50 years of integer programming 1958-2008: From the early years to the state-of-the-art*. Springer. 15

Kritzman, M., Page, S., & Turkington, D. (n.d.). In Defense of Optimization: The Fallacy of 1/N [Accepted Paper Series]. *Financial Analysts Journal*.

Lay, D. C. (2003). *Linear algebra and its applications*. Boston: Addison Wesley. 69, 70, 71, 72

Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77-91. 3

Markowitz, H. M. (1991). *Portfolio Selection*. Oxford, UK: Blackwell Publishers Inc. 2, 3, 10

Mayhew, S. (1995, Jul.-Aug.). Implied Volatility. , 8-20. (http://www.jstor.org/stable/4479853) 2

McDonald, R. L. (2006). *Derivatives Markets* (2nd. ed.). Boston: Addison-Wesley. 14

McLeish, D. L. (2005). *Monte Carlo simulation and finance*. Hoboken, NJ: John Wiley & Sons. 19

Merkulovich, L., & Rosen, D. (1998). Images from a portfolio [Working Paper]. Algorithmics Inc. 3, 4, 14, 29

Meucci, A. (2006, Feb.). Beyond Black-Litterman: views on nonnormal markets. *Risk*, 87-92. 1

Miller, N. (2006, Aug.). *Chapter 4: Topics in Consumer Theory.* available at: http://www.hks.harvard.edu/nhm/notes2006/notes4.pdf.

Papadimitriou, C. H. (1981, Oct.). On the complexity of integer programming. *J. ACM*, *28*(4), 765–768. (http://doi.acm.org/10.1145/322276.322287) 15

Prigent, J.-L. (2007). *Portfolio Optimization and Perfomance Analysis*. Boca Raton, FL: Chapman and Hall/CRC. 9

Scherer, B. (2004). *Portfolio Construction and Risk Budgeting* (2nd. ed.). London: Riskbooks. 2

Tarantola, A., & Valette, B. (1982, May). Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion. *Reviews of Geophysics and Space Physics*, *20*(2), 219–232.

Varian, H. R. (2005). *Intermediate Microeconomics: A Modern Approach*. New York and London: W.W. Norton & Company. 7

von Neumann, J., & Morgenstern, O. (1944, second ed. 1947, third ed. 1953). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press. 8

Zagst, R., & Pöschik, M. (2008). Inverse portfolio optimization under constraints. *Journal of Asset Management*, *9*(3), 239-253. (`http://www.jstor.org/stable/3088574`) 4

Ziemba, W. T., & Vickson, R. G. (2006). *Stochastic Optimization Models in Finance.* Singapore: World Scientific.

Zubelli, J. P. (2005). Inverse Problems in Finance: A Short Survey on Calibration Techniques. In *Proceedings of the Second Brazilian Conference on Statistical Modelling in Insurance and Finance.* (p. 64-76). 2