**Bayesian Estimates of the Parameters for Portfolio Optimization**

**Michael Quinn**

**University of Illinois at Urbana-Champaign**

**msquinn2@illinois.edu[[1]](#footnote-1)**

***Abstract:*** *Bayesian estimation techniques are proposed to find the parameters for a minimum variance portfolio within the Markowitz framework. Motivation for this method comes from a series of scenarios relating to an analyst’s confidence in the generalizability of very recent stock data. The paper posits that an optimal stock allocation relies on a balance between recent and long-term stock behavior. The usage of prior distributions for the parameters allows for this balance. Monte Carlo sampling techniques are used to validate results.*

***Keywords:*** *portfolio optimization, Bayesian statistics, Gibbs Sampling, Markov Chain Monte Carlo*

1. Introduction

Most of contemporary financial theory traces its lineage back to Harry Markowitz (1952) and a solution to a very specific optimization problem: given a selection of multiple assets, what’s the best way to make a portfolio? There are nearly countless answers to this question, each depending on their own assumptions about ways in which investors try to maximize personal utility. Investors could, for example, seek to maximize long-term wealth, or, just as easily, they could seek to maximize immediate gain. Either perspective depends on the investor’s personal attitude towards risk.

Markowitz’s model does not try to answer any questions about personal preferences about risk. All investors are different, and all will have different reasons for preferring different risk levels. Instead, Markowitz assumes general *risk aversion*. That is, given a certain attitude towards risk, any investor will prefer to maximize their return without increasing the risk of this return. This is the same as saying that an investor will seek to minimize the risk of attaining any level of return on their investment.

How can they do this? The answer is surprisingly simple: diversify. A broad basket of returns will minimize the risk of any one asset under-performing, allowing for a reduction in risk. Similarly, given the choice of a wide variety of assets and the option to borrow or lend at a risk free rate, any investor is capable of forming a portfolio that maximizes their possible return at a given risk level. Markowitz named this the *efficient portfolio frontier*.

Markowitz’s model is not without its criticisms. For the purposes of this paper, we will focus on three:

* The outputs of the Markowitz model are very sensitize to input parameters (Best and Grauer, 1991). This is especially true for expected returns.
* Stock data is often very noisy, which leads to a high chance for estimation error (Chopra and Ziemba, 1993). Considering that asset fundamentals are dynamic, an analyst is challenged to balance relevant and representative data.
* The Markowitz model often encounters a “corner problem,” where an extreme allocation of a limited number of assets is favored over a broad diversification of risks (Black and Litterman, 1992) . This goes against the intuition of the theory and should be avoided.

Rachev *et. al.* (2008) offer a Bayesian framework for avoiding these three problems. First, a Bayesian methodology helps make the estimate more robust by properly capturing the risk of estimation error. This results in a full distribution for each of the model’s parameters. Furthermore, using Bayesian techniques also allows for a greater amount of information incorporated into the model. The averages of analysts’ forecasts have shown to have some predictive ability, especially when accounting for certain systematic biases (Clement, 1999). The traditional Markowitz model cannot take this into account.

Bayesian methods differ from frequentist statistical methods through the usage of priors (Hoff, 2009). Frequentist methods consider data to be generated by random processes, while the parameters governing these processes are fixed. The best example would be a Physics experiment, where laws of motion govern observed process and data are generated from measurements. Bayesians reverse this relationship: the data are fixed, while parameters are random. In practice this means that a Bayesian approaches a problem with a set of beliefs about a problem and then uses data to update these beliefs. This can be a source of high risk and reward. On the one hand, incorrect priors will result in incorrect models, and many can be troubled by this subjective element added to statistical modeling. On the other hand, the existence of a prior allows for the incorporation of diverse sources of information, including previous publications, experts’ beliefs and personal experience (Kruschke, 2010). These are invaluable tools in many contexts, providing insight that is not available in data.

Bayesian methods were previously limited in application because of the computational challenges they often pose. Thanks to powerful sampling techniques like Markov Chain Monte Carlo, this is no longer the case. Most statistical software can easily handle these sorts of problems. In this case, I will use R.

The examples in this paper highlight how Bayesian methods help analysts manage uncertainty and non-representative data. This is a problem faced by many investors in Central Asia, and the techniques described here could help overcome these challenges. A Bayesian framework can provide rigorous support to the heuristic judgment currently employed for making financial decisions, optimizing portfolios and reducing risks.

Within this paper, three Gibbs samplers will be employed to estimate portfolio parameters. The first will use an unknown mean and known variance. It corresponds to an analyst not having reliable information about asset return, while keeping a sense of asset riskiness. The second will use an unknown mean and variance along with an uninformative prior. This corresponds to the analyst having limited to no information about the future of the market of the assets in her portfolio. Last, an informative prior will be used to estimate both mean and variance, incorporating an analyst’s area knowledge about future possible outcomes.

**2. Methodology**

Several methodological issues need to be addressed. I’ll begin with a discussion of portfolio allocation, following the optimization algorithm described by Constantinides and Malliaris (1995) and Zivot (2013). Next I’ll address issues concerning the probability distributions of capital asset returns. I will follow that with a discussion of the Gibbs sampler, which I rely on heavily for making Bayesian parameter estimates. Last, I’ll highlight Monte Carlo techniques for validating results.

**2.1 Optimal Portfolio Allocation**

Let’s define a portfolio as the weighted average over a vector consisting of random assets.[[2]](#footnote-2) I’ll call this vector

|  |  |  |
| --- | --- | --- |
|  |  | ( 1 ) |

Assume this vector follows multivariate normal distribution that depends on expected individual asset returns and variance .[[3]](#footnote-3)

|  |  |  |
| --- | --- | --- |
|  |  | ( 2 ) |

The linear combination of these random variables across a vector of weights is our portfolio.

|  |  |  |
| --- | --- | --- |
|  |  | ( 3 ) |

Portfolio weights are determined by the amount of total wealth allocated towards each asset. I will limit the weight of each asset in the portfolio to fall between -0.5 and 0.5, as I hope to reduce the effects of extreme distributions. A negative allocation is equivalent to short-selling that asset.

Thus, the expected return from a portfolio is just the weighted average of the returns of each of the assets,

|  |  |  |
| --- | --- | --- |
|  |  | ( 4 ) |

We will define risk as the variance of these returns. Since the movements of individual prices, dividends and returns within the stock market are correlated, a portfolio can either magnify or mitigate the risks of holding two assets. This can be captured in the covariance matrix.

|  |  |  |
| --- | --- | --- |
|  |  | ( 5 ) |

The portfolio variance is a scalar that depends on the asset allocation weights and the covariance matrix.

|  |  |  |
| --- | --- | --- |
|  |  | ( 6 ) |

The investor holds the portfolio over the limited time period. From this, the optimization problem can be expressed in two parts. First, the investor wishes to minimize risk while exceeding a minimum level of returns . That is to say,

|  |  |  |
| --- | --- | --- |
|  | subject to | ( 7 ) |

Above, is a vector of 1’s with the same dimension as . In other words, the weights must always sum to 1. At the same time, the investor wishes to maximize returns while not exceeding a maximum acceptable level of risk .

|  |  |  |
| --- | --- | --- |
|  | subject to | ( 8 ) |

Without constraints on the maximum and minimum allocations of assets in the portfolio, this problem can be solved using only matrix algebra. The investor’s minimum variance portfolio is

|  |  |  |
| --- | --- | --- |
|  |  | ( 9 ) |

The investor can add both risk and return up to his or her predefined *risk limit*. For each desired level of return , the optimal weights are found in the solution to the following linear system.

|  |  |  |
| --- | --- | --- |
|  |  | ( 10 ) |

In the preceding formula,

|  |  |  |
| --- | --- | --- |
|  |  | ( 11 ) |

The parameter vector is chosen to satisfy the following Lagrangian equation. As can be seen in the equation, the vector has two elements.

|  |  |  |
| --- | --- | --- |
|  |  | ( 12 ) |

The matrix and vector allow for the creation of the three constraints in this problem:

* The return of the portfolio must exceed the desired return at that risk level. This allows the solver to iterate over a range of possible returns and variances.
* The weights must sum to 1.
* The weights must fall within -0.5 and 0.5.

I implemented this problem in R using the quadprog package and the example of Matuszak (2013). With a loop, you can generate an efficient portfolio across a range of risk levels. The result is the following efficient frontier (see Figure 1), and I’ve included a basic one below. This plot illustrates the trade-off between risk (the portfolio’s standard deviation) and return in a given portfolio. It was chosen using the basic method of calculating means and variances from the last three months of weekly returns (see section 3 for the particular assets). While every asset allocation along the frontier can be considered optimal, most analysts select the portfolio with the highest Sharpe ratio as a rule of thumb.

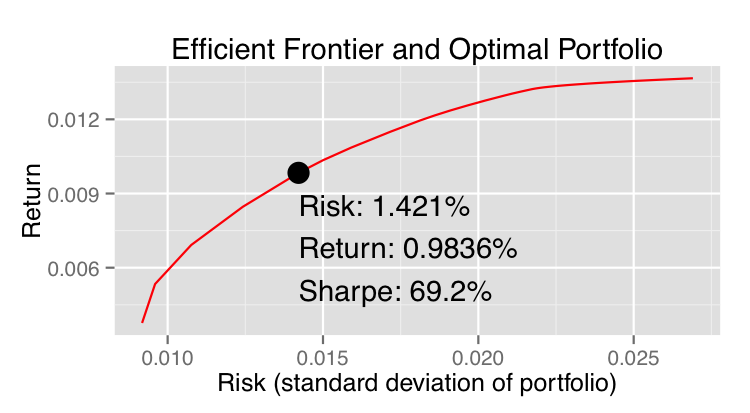


Figure 1: Efficient portfolio frontier and optimal allocation point.

**2.2 Bayesian Inference with the Multinormal Distribution**

Unless otherwise noted, all of the Bayesian forms of the multivariate normal distribution come from Gelman *et. al.* (2013). I will implement three versions of the Bayesian multivariate normal model. First, let’s define the multivariate normal distribution as follows. Given a vector random variables:

|  |  |  |
| --- | --- | --- |
|  |  | ( 13 ) |

As mentioned above, this is the simulation distribution of our asset returns. This model has the parameter vector of means of length and the parameter matrix of variances and covariances is by . The probability density function of this model is

|  |  |  |
| --- | --- | --- |
|  |  | ( 14 ) |

For independently and identically distributed (i.i.d.) observations, the likelihood is

|  |  |  |
| --- | --- | --- |
|  |  | ( 15 ) |

is the sum of squares matrix relative to and is defined thusly,

|  |  |  |
| --- | --- | --- |
|  |  | ( 16 ) |

When variance is known, the *posterior distribution for with known*  is as follows,

|  |  |  |
| --- | --- | --- |
|  |  | ( 17 ) |

The precision matrix is the inverse of the covariance, and is easier to work with in certain distributions. I will define it and the parameterized version of as follows,

|  |  |  |
| --- | --- | --- |
|  |  | ( 18 ) |

where and are the prior mean vector and variance matrix for the conjugate prior distribution of . is the sample variance and is the sample mean.

For sampling purposes it is usually easier to work with the posterior conditional marginal distributions of subvectors of with a known variance. Let the index or indicate the absence of an element with the index or from the vector of means or matrix of variances. Then appropriate conditional and marginal distributions are,

|  |  |  |
| --- | --- | --- |
|  |  | ( 19 ) |

Thus, the preceding marginal distribution allows us to sample for a single mean in the means vectors. The coefficients to find the conditional mean and the conditional percision matrix are defined as follows,

In general, write out the integers from zero through nine and express other integers as numbers: “Nine economists” but “10 sociologists.” Express all negative integers and all non-integers as numbers.

*Personal pronouns and possessives*. Do not use “their” or “them” when referring to one person. Use “her” or “his,” not “his/her” or “his or her”; “she” or “he” but not “he/she” or “he or she.” When in doubt, use the female form. Thus write: “Each economist should stick to her specialty” – not “each economist should stick to their specialty”, since it is unclear here whether the specialty to be adopted is that of our particular economist or of economists in general.

*Active and passive voices*. Generally, the active voice is clearer and more informative than the passive voice. Write: “I will argue that…” or “this paper will argue that…” -- not “it will be argued that…” since it is unclear who will be arguing, the writer or the reader. However, the passive voice has its uses.

*Punctuation*. When using quotation marks, place periods and commas within the marks. For example, write

the industrial base remains “structurally weakened.”

not

the industrial base remains “structurally weakened”.

Omit periods from abbreviations: The US, not the U.S.

Italicize unfamiliar foreign words.

2. Section headings

2.1 Section subheadings

3. Illustrations and tables

3.1 General

Illustrations and tables should be originals or sharp prints. All illustrations, both line art and photographs, should be positioned on or near the page where they are discussed, preferably at the top or bottom of the page. Cite the data source at the bottom of the table.

3.2 Tables

Center tables. Set table number and title (normal, 12 point) justified below table adjusting text to table width. Horizontal lines should be placed above and below table headings, above the subheadings and at the bottom of the table above any notes. Avoid vertical lines.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Pilot plant | |  | Full scale plant | |
|  | Influent | Effluent |  | Influent | Effluent |
| Method-C cyanide | 4.1 | 0.05 |  |  | 0.02 |
| Thiocyanide | 60.0 | 1.0 |  | 50.0 | <0.10 |
| Ammonia | 6.0 | 0.50 |  | 0.10 |  |
| Copper | 1.0 | 0.04 |  | 1.0 | 0.05 |
| Suspended solids |  |  |  |  | <10.0 |

Table 1: Typical plant emissions. Use style and adjust text to table width. Source: Smith (1976).

3.3 Figures

Center figures and images. Place the caption below the illustration, leaving about six points between the caption and the text and six points between the text and the top of the figure.

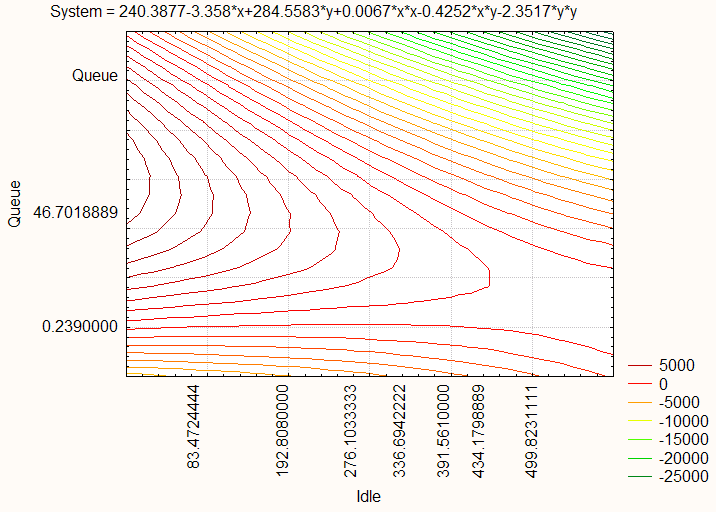


Figure 1. Caption and source here.

4. References

*A book reference:*

1. Fishwick, P. A. (1995), *“Simulation model design and execution: Building digital worlds,”* Prentice-Hall

*A journal reference:*

1. Harper, P., Phillips, S., and Gallagher, J. (2005), “Geographical simulation modeling for the regional planning of oral and maxillofacial surgery,” *Journal of the Operational Research Society*, Vol. 56, pp. 134-143

*A URL reference:*

1. Therapeutic Goods Administration (2004), “Department of Health and Ageing,” Canberra, viewed 13 October, 2004, <http://www.tga.gov.au/recalls/index.htm>

*A proceedings reference:*

1. Ahmed, S., and Amagoh, F., (2007), “AED patient arrival policy in a hospital,” *Proceedings of the International Academy of Business and Public Administration Disciplines*, Orlando, Florida, pp. 21-33

*A book chapter reference:*

1. Gould, S.J., (2000), “More things in Heaven and Earth,” in H. Rose & S. Rose (eds.), *Alas, poor Darwin: Arguments against evolutionary psychology*, Harmony Books, New York

*A reference for an unpublished manuscript:*

1. Jones, J., (2014), “Underground classic,” Obscure University (manuscript)

For style questions not covered in this document, please see *The Chicago manual of style*.

1. \* A project repository is available on Github. This includes the original .Rnw file for generating this paper and supplementary R code used in this project. See here: <https://github.com/michaelquinn32/Bayesian_Portfolio_Paper>. [↑](#footnote-ref-1)
2. All discussions of assets will center on returns, and data is formatted accordingly. A typical stock offers capital returns in the form of price increases and dividends. Bonds offer returns in the form of coupon payments and the difference between market and face value. [↑](#footnote-ref-2)
3. This is a simplifying assumption that does not hold in all market conditions. See Mandelbrot (2004). Plenty of research, including that of Markowitz, has shown that an optimal portfolio can still be found after relaxing this assumption. See (Rachev, Ortobelli, and Schwartz, 2004). [↑](#footnote-ref-3)