Dynamic Bayesian learning and optimization in portfolio choice models

by

Shea Daniel Chen

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

in

Engineering – Industrial Engineering and Operations Research

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Andrew E.B. Lim, Chair
Professor David Brillinger
Professor Zuo-Jun “Max” Shen

Spring 2014
Dynamic Bayesian learning and optimization in portfolio choice models

Copyright 2014
by
Shea Daniel Chen
Abstract

Dynamic Bayesian learning and optimization in portfolio choice models

by

Shea Daniel Chen

Doctor of Philosophy in Engineering – Industrial Engineering and Operations Research

University of California, Berkeley

Professor Andrew E.B. Lim, Chair

We develop two dynamic Bayesian portfolio allocation models that address questions of learning and model uncertainty by taking model-specific shortcomings into account.

In our first model, we formulate a multi-period portfolio choice problem in which the investor is uncertain about parameters of the model, can learn these parameters over time from observing asset returns, but is also concerned about robustness. To address these concerns, we introduce an objective function which can be regarded as a Bayesian version of relative regret. The optimal portfolio is characterized and shown to involve a “tilted” posterior, where the tilting is defined in terms of a family of stochastic benchmarks. We have found this model to perform at least as well as a benchmark given the true market parameters, while outperforming it when the market assets have the same trend.

Our next model extends the Black-Litterman portfolio choice model by taking several potential errors into account. We extend Black-Litterman to multiple periods, which allows for us to take into account the pairs of expert forecasts and the realized return. By doing so, we can then perform inference on these experts and discover whether they may have any bias for or against any specific assets. We can also perform similar inference on the market equilibrium distribution, which is typically represented by the capital asset pricing model (CAPM). The result is a model that is analytically intractable but may be solved numerically via Gibbs sampling. Controlled tests show our model performs favorably when Black-Litterman’s model assumptions about the market equilibrium and expert views are violated. Backtests shed light on the model’s ability to account for CAPM’s shortcomings.
To my family, whose unwavering support over the years have guided me toward the completion of this work.
# Contents

**Contents**

List of Figures ............................... iv

List of Tables ............................... v

1 Introduction .................................. 1

2 Dynamic portfolio choice with Bayesian regret .................. 2
   2.1 Introduction .................................. 2
   2.2 Bayesian Regret .............................. 3
      2.2.1 Market Model .............................. 3
      2.2.2 Investor .................................. 4
      2.2.3 Benchmark ................................. 5
   2.3 Formulation of the Bayesian Benchmark Problem ............... 6
      2.3.1 Value Function and Optimal Portfolio ......... 8
   2.4 Tilted Posterior ............................. 9
   2.5 Solution for Known $Q$ .......................... 11
      2.5.1 Tilted Posterior ......................... 11
      2.5.2 Solution .................................. 14
      2.5.3 Numerical Experiments ................. 15
      2.5.4 Backtests ................................ 17

3 A Bayesian Graphical Generalization for the Black-Litterman Model 20
   3.1 Introduction .................................. 20
   3.2 Preliminaries ............................... 22
      3.2.1 The Black-Litterman Model .............. 22
      3.2.2 Gibbs Sampling ........................... 25
   3.3 Graphical Generalized Black-Litterman Model ............ 25
      3.3.1 Model Assumptions ....................... 28
      3.3.2 Conjugate Priors and Hyperparameters .... 31
      3.3.3 Gibbs Sampling Algorithm ............... 31
### Contents

3.4 Numerical Tests .................................................. 33
   3.4.1 Controlled Tests ........................................... 34
   3.4.2 Backtests .................................................. 36

4 Conclusion ......................................................... 42

Bibliography ......................................................... 43
List of Figures

2.1 Backtest of Bayesian regret and classical Bayesian models over the period of 1997-1999 in a market containing only GE, ExxonMobil, and Apple stocks and a risk free asset with \( r = 0.01 \). ............................................................... 19

3.1 Graphical representation of the classical Black-Litterman model. .............. 24
3.2 Graphical representation of Black-Litterman after multiple periods. ............. 27
3.3 Graphical representation of generalized Black-Litterman accounting for expert view bias. ................................................................. 28
3.4 Graphical representation of generalized Black-Litterman model accounting for equilibrium errors. .................................................... 29
3.5 Graphical representation of generalized Black-Litterman, which accounts for violations in Black-Litterman’s assumptions regarding the market equilibrium and expert views. .................................................. 30
3.6 Gibbs samples of \( b^\mu \) for AAPL, 2002-2012 ........................................ 38
3.7 Gibbs samples of \( b^\mu \) for GE, 2002-2012 .......................................... 38
3.8 Portfolio wealths of our extension and Black-Litterman in the 2002-2012 backtest with no expert views ............................................. 39
3.9 Interquartile ranges of the Gibbs sampled expert view biases for AAPL in the 2002-2012 backtest with biased expert views. ......................... 40
3.10 Portfolio wealth processes for Black-Litterman and our extension model in the 2002-2012 backtest with biased expert views. ......................... 41
List of Tables

2.1 Simulation results of model performance under various market conditions . . . . 17

3.1 Summary of the traits of the classical Black-Litterman model and the generalized
Black-Litterman model. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26
3.2 Sharpe ratios of simulations with equilibrium error in the absence of expert views 35
3.3 Sharpe ratios of simulations with view error and no equilibrium errors . . . . . . 36
3.4 Table of the ratio of Sharpe ratios of full model simulations. Values greater than
1 indicate generalized Black-Litterman outperformed Black-Litterman. . . . . . . 36
Acknowledgments

Over the last few years I have been blessed to receive advice and support from many gifted individuals. Professor Andrew Lim has been a mentor and a friend in guiding me through this endeavor. Not enough can be said for his patience and faith in my abilities. I would also like to thank my committee of David Brillinger and Max Shen for their thoughts throughout this process. I would also like to thank George Shanthikumar and Jon Burgstone; I may have been your Graduate Student Instructor, but I learned far more than I taught.

I would also like to thank the IEOR Department Staff, particularly Anayancy Paz and Mike Campbell, for their advice and support in navigating administrative details. Finally, I would like to thank all of my peers for the many research and non-research related conversations over the years.
Chapter 1

Introduction

The birth of modern portfolio theory is often credited to Harry Markowitz in his seminal paper on the topic [24]. In it, Markowitz developed the classical mean-variance framework which featured an objective function that maximized the difference between expected return and the variance of that return. The field has since evolved to utilize a multitude of techniques to expand and improve upon Markowitz’s model.

Our work falls under the category of dynamic portfolio allocation models. Given periodic signals regarding the market, these models seek to properly re-allocate the portfolio weights in every period. These signals can come in a variety of forms, from actual pricing data to expert predictions. A natural framework to consolidate past information with updated information is a Bayesian framework. In working under a Bayesian framework, the tractability of these problems improves dramatically and techniques such as dynamic programming and Gibbs sampling may be utilized as solution methods.

In the portfolio allocation models we will discuss, we focus on opportunities to improve the learning in each model. Bayesian models are often entirely data-driven, basing their decisions on the expectation of future outcomes. There are typically no concerns of relative performance or robustness, which could cause these models to perform poorly in cases of model misspecification. The success of Bayesian models relies heavily on the quality of the observed data. If there are biases or errors in the observations themselves, estimation errors will occur and result in suboptimal solutions.

In this dissertation we improve upon two different approaches to portfolio allocation using dynamic Bayesian methods. Our topic on Bayesian regret in Chapter 2 takes the classical Markowitz problem and changes the objective function to maximize relative regret against a benchmark, rather than simply maximizing total wealth. By doing so, we develop a performance-driven model that is capable of accounting for errors in Bayesian inference while staying within a Bayesian framework. On the other hand, our extension of the Black-Litterman model in Chapter 3 extends the original model in order to account for specific erroneous model assumptions.
Chapter 2

Dynamic portfolio choice with Bayesian regret

2.1 Introduction

Consider an investor living in a world where there are risky assets and a risk-free money market account. Log-returns for the risky assets are i.i.d. normal but the mean and variance are not known to the investor (though they are constant over time). We consider a multiple period model in discrete time where the agent can re-balance his/her portfolio at the start of each investment period.

Portfolio selection falls under the umbrella of optimization under model uncertainty. When parameters of the model are unknown, it is common to study this problem in a Bayesian framework. Usage of a Bayesian framework to address linear-quadratic-Gaussian problems is classical (e.g. [1, 33]) and has also been utilized in portfolio choice problems where the volatility is known [10, 2]. The strength of a Bayesian framework lies in the ability for the user to impose personal views in the form of a prior distribution and to update these views conditional on data. One major limitation of this approach that we address in this paper is that Bayesian frameworks require strong model assumptions and consequently do not consider robustness.

Classical approaches to robust portfolio selection typically utilize a worst-case or max-min objective. While robust in taking all scenarios into consideration, these methods are criticized for being too pessimistic. An alternate approach for incorporating robustness is the notion of relative regret. Relative regret is an approach to robust decision making that defines a “benchmark” for each potential value of the unknown model parameters (commonly, the optimal objective value for a decision maker who optimizes with knowledge of the true parameter values) and evaluates a decision by comparing its performance to that of the benchmark over all possible parameter values. The optimal regret decision is the one that minimizes the worst case relative performance. Early results can be found in [9, 16, 29], whereas more recent analysis of problems utilizing relative regret include [5, 20, 23, 26].
CHAPTER 2. DYNAMIC PORTFOLIO CHOICE WITH BAYESIAN REGRET

Also the related is the notion of asymptotic regret in [12]. One advantage of regret, over the standard worst case approach to robust control (e.g., [3, 13]), is that it is concerned about doing well (relative to the benchmark) over all possible parameter values, and hence is less pessimistic than the usual worst case approach.

In this paper we propose the notion of Bayesian relative regret, which naturally combines ideas from both the Bayesian and regret frameworks, and study solutions of a dynamic optimal portfolio choice problem with this objective. In doing this, we develop a model that is robust to errors in parameter estimation utilizing the feedback provided by relative performance.

The outline of the paper is as follows. We formulate the multi-period market model in Section II, and introduce the Bayesian benchmark problem in Section III. In Section IV, we discuss the interpretation of the “tilted” posterior, an interesting result that arises naturally from our problem formulation. We solve this problem for the special case where the variance is known and provide numerical results in Section V.

2.2 Bayesian Regret

We now lay out the components of our Bayesian regret problem. First we set up the market model.

2.2.1 Market Model

We assume that there is a money market account with a known, constant, continuously compounded interest rate \( r \), and \( n \) risky assets with prices

\[
S_i(t) = S_i(0) \exp \{ \mu_i t + \sigma_i Z(t) \}, \quad i = 1, 2, \ldots, n,
\]

where \( Z(t) \) is a \( n \)-dimensional standard Brownian motion. Prices evolve in continuous time but we assume that the investor only samples at discrete time intervals of size \( \delta \) and sees the prices \( S_1(t\delta), S_2(t\delta), \ldots, S_n(t\delta) \) at each \( t = 0, 1, 2, \ldots, T \). Let \( S(t\delta) = [S_1(t\delta), S_2(t\delta), \ldots, S_n(t\delta)]' \) be the vector of stock prices at time \( t\delta \) and \( \mathcal{S}(t\delta) = \{S(0), S(\delta), \ldots, S(t\delta)\} \) be the history of stock prices observations at \( t\delta \). It follows from this model that log-returns of the stock price \( j \) are i.i.d. normal:

\[
\mathcal{R}_j(t + 1) \overset{\text{def}}{=} \log \frac{S_j((t + 1)\delta)}{S_j(t\delta)} = \mu_j \delta + \sqrt{\delta} \sigma_i Z(t + 1)
\]

where \( \mu_j \in \mathbb{R} \) and \( \sigma_j = [\sigma_{j1}, \ldots, \sigma_{jn}] \in \mathbb{R}^{1 \times n} \) are constants, \( 0 < \delta \ll 1 \) corresponds to the time increment between observations, and \( Z(1), Z(2), \ldots, Z(T) \) are i.i.d. \( n \)-dimensional standard normal random variables. It follows that the vector of log returns

\[
\mathcal{R}(t + 1) = [\mathcal{R}_1(t + 1), \ldots, \mathcal{R}_n(t + 1)]'
= \delta \mu + \sqrt{\delta} \sigma Z(t + 1) \sim N(\delta \mu, \delta Q),
\]

(2.2)
are also normal and i.i.d. with parameters

\[
\mu = \begin{bmatrix}
\mu_1 \\
\mu_2 \\
\vdots \\
\mu_n
\end{bmatrix} \in \mathbb{R}^{n \times 1}, \quad \sigma = \begin{bmatrix}
\sigma_1 \\
\sigma_2 \\
\vdots \\
\sigma_n
\end{bmatrix} \in \mathbb{R}^{n \times n}, \quad Q = \sigma \sigma' \in \mathbb{R}^{n \times n}.
\]

Define \( b \overset{\text{def}}{=} \mu - r \) to be the vector of mean returns of the \( n \) risky assets less the risk-free interest rate. The vector of annualized log-returns, will thus be

\[
B(t + 1) = \frac{R(t + 1)}{\delta} - r \sim N(b, Q/\delta)
\]

We denote by

\[
G_t \overset{\text{def}}{=} \sigma\{B(1), B(2), \ldots, B(T)\} = \sigma\{R(1), R(2), \ldots, R(T)\} = \sigma\{S(0), S(\delta), \ldots, S(t\delta)\} = \sigma\{S(t\delta)\}
\]

the \( \sigma \)-algebra generated by the annualized returns \( B(1), B(2), \ldots, B(t) \) (and equivalently, log returns \( R(1), R(2), \ldots, R(T) \) and stock prices \( S(0), S(\delta), \ldots, S(T\delta) \)) and by \( G \overset{\text{def}}{=} \{G_t\} \) the associated filtration. We assume that the parameters \( H = (Q, b) \) describing the asset price returns are constant and belong to a compact set \( H \subset \mathbb{R}^n \times S^{n \times n}_+ \) where \( S^{n \times n}_+ \) denotes the set of symmetric and strictly positive definite \( n \times n \) matrices with real entries and that there is a constant \( \delta > 0 \) such that \( Q \geq \delta I \) for every \((Q, b) \in H\). We assume that the investor knows the uncertainty set \( H \) but not the specific parameter values \((Q, b)\) of the model generating the returns.

### 2.2.2 Investor

Suppose we have an investor with initial wealth \( x(0) \) who is able to invest in the financial market. The investor is able to change his/her portfolio at each decision epoch in response to realized returns. More specifically, the investors portfolio \( \pi(t) = [\pi_1(t), \ldots, \pi_n(t)]' \) at decision epoch \( t \) determines the proportion of his/her wealth invested in each of the \( n \) risky assets between decision epochs \( t \) and \( t + 1 \) (equivalently, on the time interval \([t\delta, (t + 1)\delta)\)), and can depend on the history of returns \( R(1), R(2), \ldots, R(t) \) up to and including decision epoch \( t \) (and equivalently, the stock prices \( S(0), S(\delta), \ldots, S(t\delta) \) up to time \( t\delta \)). With this in mind, we define the class of admissible investment portfolios by

\[
A = \{\pi = (\pi(0), \pi(1), \pi(2), \ldots, \pi(T - 1))|\pi(t) \in G_t\}. \tag{2.4}
\]
The investor does not know the parameter values \((Q, b)\) beyond the fact that they belong to the set \(\mathcal{H}\).

We assume that the investor’s wealth process satisfies

\[ x(t+1) = x(t) e^{\delta \left[ r + b' \pi(t) - \frac{1}{2} \pi(t)' Q \pi(t) \right] + \sqrt{\pi(t)' \sigma Z(t+1)}}, \quad (2.5) \]

which is an approximation for \(x(t+1)\), i.e.,

\[
\begin{align*}
\sum_{i=1}^{n} \frac{x(t) \pi_i(t)}{S_i(t)} S_i(t+1) + x(t) \left( 1 - \sum_{i=1}^{n} \pi_i(t) \right) e^{r \delta} \\
= x(t) \sum_{i=1}^{n} \pi_i(t) e^{\pi_i(t+1)} + x(t) \left( 1 - \sum_{i=1}^{n} \pi_i(t) \right) e^{r \delta} \quad (2.6)
\end{align*}
\]

which becomes exact when \(\delta \downarrow 0\) (see also ).

Using (2.3) and (2.2), (2.5) can be rewritten as

\[ x(t+1) = x(t) e^{\delta \left[ r + b' \psi(t) - \frac{1}{2} \psi(t)' Q \psi(t) \right] + \sqrt{\psi(t)' \sigma Z(t+1)}}. \quad (2.7) \]

### 2.2.3 Benchmark

In this section we define and construct a benchmark process which will be used in our formulation of the Bayesian regret problem. Suppose that the model generating the stock prices has parameters \(H \in \mathcal{H}\). Adopting (2.5) and (2.7), the wealth of an investor who adopts the portfolio \(\psi(t)\) will evolve as

\[
\begin{align*}
y(t+1) &= y(t) e^{\delta \left[ r + b' \psi(t) - \frac{1}{2} \psi(t)' Q \psi(t) \right] + \sqrt{\psi(t)' \sigma Z(t+1)}} \\
&= y(t) e^{\delta \left[ r + \psi(t)' B(t+1) - \frac{1}{2} \psi(t)' Q \psi(t) \right]} \quad (2.8)
\end{align*}
\]

We define the benchmark associated with model parameters \(H = (Q, b)\) as the wealth corresponding to optimal investment (in the sense we now define) for this model. More specifically, we assume that the investment portfolio \(\psi_H(t)\) adopted by the benchmark investor is the optimal solution of the multi-period problem

\[
\begin{align*}
\max_{\psi} E_H U(y_H(t)) &\equiv \max_{\psi} \frac{1}{\eta} E_H \left[ y_H(T)^{\eta} \right] \\
\text{subject to} &
\begin{align*}
y_H(t+1) &= y_H(t) e^{\delta \left[ r + b' \psi(t) - \frac{1}{2} \psi(t)' Q \psi(t) \right] + \sqrt{\psi(t)' \sigma Z(t+1)}} \\
y_H(0) &\text{known}.
\end{align*}
\end{align*}
\quad (2.10)
\]
Note that the benchmark investor’s utility is \( U^B(y) = \frac{1}{\eta}y^\eta \). It is well known that

\[
\psi^H(t) = \frac{1}{1 - \eta}Q^{-1}b
\]

is the solution to this problem. Note that \( \psi^H(t) \) depends explicitly on the parameter \( H \).

The wealth of the benchmark investor is given by

\[
y^H(t + 1) = y^H(t)e^{\delta \left[ r + \frac{1 - 2\eta}{\pi(1 - \eta)}b'Q^{-1}b + \sqrt{\frac{\pi}{1 - \eta}}b'Q^{-1}b \right]}
\]

which is obtained by substituting \( \psi^H \) into (2.8) and (2.9).

### 2.3 Formulation of the Bayesian Benchmark Problem

We are almost ready to state our Bayesian benchmark problem. The objective will be similar to a relative regret objective through its use of the comparison function \( U(z) = \frac{1}{\gamma}z^\gamma \), where \( \gamma < 1 \). How this models differs from a relative regret formulation is that the investor now also has a prior \( \mu(dH) \) on the set of unknown parameters \( \mathcal{H} \).

We are interested in solving the Bayesian relative regret problem

\[
\max_{\pi} \frac{1}{\gamma} \mathbb{E} \left( \frac{y(T)}{y(T)} \right)^\gamma
\]

subject to

\[
x(t + 1) = x(t)e^{\delta \left[ r + \frac{1 - 2\eta}{\pi(1 - \eta)}b'Q^{-1}b + \sqrt{\frac{\pi}{1 - \eta}}b'Q^{-1}b \right] + \sqrt{\pi}Q \pi(t) \sigma Z(t+1)}
y(t + 1) = y(t)e^{\delta \left[ r + \frac{1 - 2\eta}{\pi(1 - \eta)}b'Q^{-1}b + \sqrt{\frac{\pi}{1 - \eta}}b'Q^{-1}b \right] + \sqrt{\pi}Q \pi(t) \sigma Z(t+1)}
\]

\[ x(0), y(0) \text{ given.} \]

\[ \mu(dH) \sim \text{distribution on } \mathcal{H} \]

\[ \pi \in \mathcal{A} \]

In particular, the prior distribution \( \mu(dH) \) on the family of models \( \mathcal{H} \) needs to be specified. The parameter \( H = (Q, b) \) generating the data is not known by the investor, and it follows that the benchmark process \( y(t) \) is also not observable since it depends on the model \( H \) through \( \psi^H \). In other words (2.14) is a partially observed stochastic control problem in which \((H, y(t))\) are not observable. We adopt the observable pair \((x(t), S_t)\) as the state for (2.14).

An advantage of this model is that (2.14) allows the decision maker to express subjective views about the unknown parameter \( H \) through the prior distribution. It is worth comparing
(2.14) to the classical Bayesian formulation of the portfolio selection problem

$$\begin{align*}
\max_\pi \frac{1}{\gamma} \mathbb{E} (x(T))^{\gamma} \\
\text{subject to} \\
x(t+1) = x(t) e^{\delta [r+b'\pi(t)-\frac{1}{2}\pi(t)'Q\pi(t)] + \sqrt{3}\pi(t)'\sigma Z(t+1)} \\
x(0) \text{ given.} \\
\mu(dH) \sim \text{distribution on } \mathcal{H} \\
\pi \in \mathcal{A}
\end{align*}$$

(2.15)

which does not "benchmark" the objective.

For another perspective on the role of the benchmark process, define the function-valued stochastic process

$$y(t) \overset{\text{def}}{=} \{y^H(t)|H \in \mathcal{H}\}, \ t \in [0,T].$$

For each $t$, $y(t)$ is a function of $H$, where $y^H(t)$ (the function $y(t)$ evaluated at $H$) is the wealth of the benchmark investor for model $H$ conditional on the realization $S_t$ of the stock prices. Observe that $y^H(t)$ is a $\mathcal{G}_t$-adapted stochastic process (for $H \in \mathcal{H}$ fixed) and $y(t)$ is a $\mathcal{G}_t$-adapted function valued stochastic process. The objective in (2.14) can be written in the form

$$\frac{1}{\gamma} \mathbb{E} \left[ \left( \frac{x(T)}{y(T)} \right)^{\gamma} \right] = \frac{1}{\gamma} \mathbb{E} \left\{ \left( \frac{x(T)}{y(T)} \right)^{\gamma} | S_T \right\}.$$ 

Conditioning on the parameter $H$, the inner expectation becomes

$$\mathbb{E} \left\{ \left( \frac{x(T)}{y(T)} \right)^{\gamma} | S_T \right\} = \mathbb{E} \left\{ \mathbb{E} \left( \left( \frac{x(T)}{y^H(T)} \right)^{\gamma} | H, S_T \right) | S_T \right\}.$$ 

Since

$$\mathbb{E} \left( \left( \frac{x(T)}{y(T)} \right)^{\gamma} | H, S_T \right) = \left( \frac{x(T)}{y^H(T)} \right)^{\gamma}$$

due to the $\mathcal{G}_t$-measurability of $x(T)$ and the definition of $y^H(T)$. It follows that

$$\mathbb{E} \left\{ \left( \frac{x(T)}{y(T)} \right)^{\gamma} | S_T \right\} = \int_{H \in \mathcal{H}} \left( \frac{x(T)}{y^H(T)} \right)^{\gamma} \mu_T(dH)$$

and hence

$$\frac{1}{\gamma} \mathbb{E} \left[ \left( \frac{x(T)}{y(T)} \right)^{\gamma} \right] = \frac{1}{\gamma} \mathbb{E} \left\{ \int_{H \in \mathcal{H}} \left( \frac{x(T)}{y^H(T)} \right)^{\gamma} \mu_T(dH) \right\}.$$ 

(2.16)

The equality (2.16) shows that the objective in our problem (2.14) is equivalent to another objective (the right hand side of (2.16)) that compares the terminal wealth $x(T)$ with the terminal benchmark value $y^H(t)$ by computing $\frac{1}{\gamma} \left( \frac{x(T)}{y^H(t)} \right)^{\gamma}$ for each model $H \in \mathcal{H}$, combines
all these values by integrating over all models with respect to the posterior distribution $\mu_T(dH)$ (the integral in (2.16), and then computes the expectation with respect to the prior $\mu$ (the outer expectation). Roughly speaking, the investor is rewarded for performing well relative to the benchmark for models that appear more likely according to the data (i.e., where $\mu_T(dH)$ takes a large value). That is, it appears that the optimal portfolio for such an objective will seek to perform well relative to the entire family of benchmarks when there is insufficient data to eliminate most of the models (i.e., when the posterior is still diffuse), and concentrates on performing well relative to the most likely models when the posterior variance decreases.

2.3.1 Value Function and Optimal Portfolio

The following result characterizes the value function $V^\mu(t,x,S_t)$ and optimal policy $\pi^*$ for the Bayesian benchmarking problem (2.14) for a given prior distribution $\mu$. Observe as in (2.16) that the conditional expectations $E[y(T)^{-\gamma}|S_T]$ which appears in the definition of the value function can equivalently be written as

$$E[y(T)^{-\gamma}|S_T] \equiv \int_{H \in \mathcal{H}} y^H(T)^{-\gamma} \mu_T(dH)$$

and similarly for other such terms.

**Proposition 2.3.1.** The value function for (2.14) is

$$V^\mu(T,x,S_T) = \frac{x^\gamma}{\gamma} E[y(T)^{-\gamma}|S_T]$$

$$V^\mu(t,x,S_t) = \frac{x^\gamma}{\gamma} E[y(t)^{-\gamma}|S_t]$$

$$= \frac{1}{\gamma} E \left[ \left( \frac{x}{y(t)} \right)^\gamma e^{\gamma h(t,S_t)} \right| S_t]$$

(2.17)

where $h(t,S_t)$ is a $\mathcal{G}_t$ adapted process defined by $h(T,S_T) = 0$ and

$$e^{\gamma h(t,S_t)} E[y(t)^{-\gamma}|S_t] = \max_{\pi} E \left[ y(t)^{-\gamma} \exp \left\{ \gamma \delta \left( b'\pi - \frac{1}{2}\pi'Q\pi - \frac{1-2\eta}{2(1-\eta)\gamma} b'Q^{-1}b \right) \\ + \gamma \sqrt{\delta} \left( \pi - \frac{1}{1-\eta} Q^{-1}b \right)' \sigma Z(t+1) + \gamma h(t+1,S_{t+1}) \right] \right| S_t].$$

(2.18)

The optimal policy is

$$\pi^\mu(t,S_t) = \arg \max_{\pi} E \left[ y(t)^{-\gamma} \exp \left\{ \gamma \delta \left( b'\pi - \frac{1}{2}\pi'Q\pi - \frac{1-2\eta}{2(1-\eta)\gamma} b'Q^{-1}b \right) \\ + \gamma \sqrt{\delta} \left( \pi - \frac{1}{1-\eta} Q^{-1}b \right)' \sigma Z(t+1) + \gamma h(t+1,S_{t+1}) \right] \right| S_t].$$

(2.19)
In principle, we need to solve (2.18) recursively to determine \( h(T, S_T) \). Dynamic programming implies that the optimal policy does not depend on the benchmarks because \( \pi \) family of benchmarks establishes (2.17) as well as the recursive relationship (2.18) for the process \( h(t, S_t) \).

**Proof.** Clearly \( V^\mu(T, x; S_T) \) is given by (2.17) and it follows that \( h(T, S_T) = 0 \). Suppose there is some \( t+1 \) such that

\[
V^\mu(t+1, x, S_{t+1}) = \frac{1}{\gamma} E \left[ \left( \frac{x}{y(t+1)} \right)^\gamma e^{\gamma h(t+1, S_{t+1})} | S_{t+1} \right].
\]

Dynamic programming implies that

\[
V^\mu(t, x, S_t) = \max_{\pi} \mathbb{E} \left\{ V^\mu(t+1, x(t+1), S_{t+1} \big| x(t) = x, S_t) \right\}
\]

\[
= \max_{\pi} \mathbb{E} \left\{ \frac{1}{\gamma} \left( \frac{x(t+1)}{y(t+1)} \right)^\gamma e^{\gamma h(t+1, S_{t+1})} \bigg| x(t+1), S_{t+1} \right\} \bigg| x(t) = x, S_t) \right\}
\]

\[
= \max_{\pi} \mathbb{E} \left\{ \frac{1}{\gamma} \left( \frac{x(t+1)}{y(t+1)} \right)^\gamma e^{\gamma h(t+1, S_{t+1})} \bigg| x(t+1), S_{t+1} \right\} \bigg| x(t) = x, S_t) \right\}
\]

\[
= \frac{x^\gamma}{\gamma} \max_{\pi} \mathbb{E} \left\{ y(t)^{-\gamma} e^{\gamma \delta (\nu - \frac{\nu}{\gamma} Q_{\pi} - \frac{1-\nu}{\gamma^2} Q^{-1} b)} + \gamma \sqrt{\mathbb{V}(\pi - \frac{1}{\gamma} Q^{-1} b)'} \sigma Z(t+1) + \gamma h(t+1, S_{t+1}) \bigg| S_t \right\}
\]

This establishes (2.17) as well as the recursive relationship (2.18) for the process \( h(t, S_t) \). The characterization (2.19) of the optimal policy follows from the observation that this is the maximizer in the dynamic programming recursion. \( \square \)

Proposition 2.3.1 shows that the optimal policy depends on the data \( S_t \) through the family of benchmarks \( y(t) \equiv \{y^H(t), H \in \mathcal{H} \} \) as well as the posterior distribution \( \mu_t(dH) \). In principle, we need to solve (2.18) recursively to determine \( h(t, S_t) \) in order to find the optimal portfolio.

There are several special cases where the recursion is relatively easy to solve. The first of these is when the comparison function is logarithmic, \( U(z) = \log(z) \), i.e., \( \gamma = 0 \). In this case, the optimal policy does not depend on the benchmarks because \( \mathbb{E}[\log(z)] = \mathbb{E} \left[ \log \left( \frac{z}{y} \right) \right] = \mathbb{E}[\log(x)] - \mathbb{E}[\log(y)] \), so optimizing the objective (2.14) is equivalent to maximizing expected log utility of the (un-benchmarked) terminal wealth \( \mathbb{E}[\log(x)] \).

The other case is when we let \( \gamma \to \infty \) while keeping \( \delta > 0 \) constant. In this case, it is optimal to solve a sequence of single-period regret problems without learning. The most interesting case corresponds to intermediate values of \( \gamma \) when the comparison function is not logarithmic, where dependence on both the posterior and benchmarks cannot be avoided.

### 2.4 Tilted Posterior

In this section we give another perspective on the interplay between benchmarks and the posterior. Specifically, we show that the benchmarks define a likelihood ratio (for each time period) which “tilts” the posterior towards models that have performed well for the
data. This interpretation allows us to reformulate the recursion (2.18), the characterization (2.19) of the optimal policy, and the problem of parameter estimation in terms of the tilted posterior.

To begin, dividing (2.18) through by \(\mathbb{E}[y(t)^{-\gamma}\mid S_t]\) suggests that we define

\[
M_t^\gamma \overset{\text{def}}{=} \frac{y(t)^{-\gamma}}{\mathbb{E}[y(t)^{-\gamma}\mid S_t]} \equiv \frac{y(t)^{-\gamma}}{\int_\mathcal{H} y^H(t)^{-\gamma} \mu_t(dH)} \tag{2.20}
\]

or equivalently,

\[
M_t^\gamma(H) \overset{\text{def}}{=} \frac{y^H(t)^{-\gamma}}{\int_\mathcal{H} y^H(t)^{-\gamma} \mu_t(dH)} \equiv \frac{y^H(t)^{-\gamma}}{\mathbb{E}[y(t)^{-\gamma}\mid S_t]} \equiv \frac{y^H(t)^{-\gamma}}{\mathbb{E}_{\mu_t}[y(t)^{-\gamma}]}. \tag{2.21}
\]

Conditional on \(S_t\), \(M_t^\gamma(H)\) is a function defined on the set of models \(\mathcal{H}\). Observing that \(M_t^\gamma > 0\) \(\mathbb{P}\)-a.s., and

\[
\mathbb{E}[M_t^\gamma \mid S_t] \equiv \mathbb{E}_{\mu_t}[M_t^\gamma] \overset{\text{def}}{=} \int_\mathcal{H} M_t^\gamma(H) \mu_t(dH) = 1
\]

it follows that \(M_t^\gamma\) is the Radon-Nikodym derivative/likelihood ratio defining a change of measure from the posterior \(\mu_t\) to another measure \(\rho_t\) on \(\mathcal{H}\). This point is worth dwelling on (and perhaps rephrasing) as it is an intuitive idea. At time \(t\) we have observations \(S_t\) and the posterior distribution \(\mu_t(dH)\) which is a probability measure on \(\mathcal{H}\). The random variable \(M_t^\gamma : \mathcal{H} \to [0, \infty)\) obtained from the benchmark function \(y(t)\) defines a new probability measure \(\rho_t(dH)\) on \(\mathcal{H}\) via

\[
\rho_t(dH) = M_t^\gamma(H) \mu_t(dH). \tag{2.21}
\]

Observe that \(M_t^\gamma(H)\) is “large” for values of \(H\), i.e., models, where \(y^H(t)\) is large and small where \(y^H(t)\) is small. That is, \(M_t^\gamma(dH)\) tilts the posterior in favor of models that performed well historically which, as we have already mentioned, seems like a natural thing to do.

Recognizing that \(\rho_t(dH)\) is a probability measure, we can now define, for any matrix \(A\),

\[
\mathbb{E}_{\rho_t}[A] \equiv \mathbb{E}[M_t^\gamma A \mid S_t], \text{ etc.}
\]

In particular, we can now write the recursion (2.18) for \(h(t, S_t)\)

\[
e^{-h(t, S_t)} = \max_{\pi} \mathbb{E}_{\rho_t}\left[ \exp\left\{ \gamma \delta \left( b' \pi - \frac{1}{2} \pi'Q\pi - \frac{1-2\eta}{2(1-\eta)^2} b'Q^{-1}b \right) \right. \right.
+ \gamma \sqrt{\delta} \left( \pi - \frac{1}{1-\eta} Q^{-1}b \right)' \sigma Z(t+1) + \gamma h(t+1, S_{t+1}) \bigg]\mid S_t. \tag{2.22}
\]
and the optimal policy (2.19)

$$
\pi^\mu(t, S_t) = \arg \max_\pi \mathbb{E}_\rho \left[ \exp \left\{ \gamma \delta \left( b' \pi - \frac{1}{2} \pi' Q \pi - \frac{1-2\eta}{2(1-\eta)} b' Q^{-1} b \right) \right\} + \gamma \sqrt{\delta} \left( \pi - \frac{1-\eta}{2} Q^{-1} b \right)' \sigma Z(t+1) + \gamma h(t+1, S_{t+1}) \right] | S_t. \right. \tag{2.23}
$$

in terms of the tilted posterior $\rho_t$.

2.5 Solution for Known $Q$

We assume now that the investor knows the covariance matrix $Q$ but not the true mean $b$ of the annualized log-normal returns. To model this, our prior $\mu_0$ will follow a multivariate normal distribution over the mean returns $b$.

2.5.1 Tilted Posterior

Let the prior at time $t = 0$ be denoted as $\mu_0(dH) \sim N(b_0, \Sigma_0)$ where the hyperparameter $b_0$ is the prior estimate of $b$ and $\Sigma_0$ is the covariance matrix which indicates the initial confidence of the estimate $b_0$. Recall $B(t) \overset{d}{=} \frac{R(t)}{\delta} - r \sim N(b, Q/\delta)$, the annualized log-returns less $r$. Note that this form gives us observations with mean $b$, our variable of interest. The normality of the log-returns allows us to utilize conjugate priors when updating $\mu_t$. If at time $t$, the prior is $\mu_t(dH) \sim N(b_t, \Sigma_t)$, it is well known that after observing $B(t+1)$, the posterior $\mu_{t+1}(dH)$ is also normal with hyperparameters

$$
\begin{align*}
& b_{t+1} = [\Sigma_t^{-1} + \delta Q^{-1}]^{-1} \left[ \Sigma_t^{-1} b_t + \delta Q^{-1} B(t+1) \right] \\
& \Sigma_{t+1} = [\Sigma_t^{-1} + \delta Q^{-1}]^{-1}
\end{align*}
$$

Using (2.21) along with this fact, we may state the following proposition.

**Proposition 2.5.1.** Suppose that prior to time $t$, the investor knows the covariance $Q$ and has a prior distribution on the mean log-returns of each stock (less the risk free rate), $b$. At time $t$, let the posterior be distributed as $\mu_t(dH) \sim N(b_t, \Sigma_t)$. Given a hidden benchmark process with utility function $U^B(y) = \frac{1}{\eta} y^\eta$, $\eta < 1$, the tilted posterior is normally distributed with mean

$$
\left[ \Sigma_t^{-1} + \frac{-\gamma t \delta}{(1-\eta)^2} Q^{-1} \right]^{-1} \left[ \Sigma_t^{-1} b_t + \frac{-\gamma t \delta}{1-\eta} Q^{-1} b_t \right]
$$

and covariance matrix

$$
\left[ \Sigma_t^{-1} + \frac{-\gamma t \delta}{(1-\eta)^2} Q^{-1} \right]^{-1}
$$

where

$$
\bar{B}_t \overset{d}{=} \frac{1}{t} \sum_{s=1}^t B(s) = \frac{1}{t} \sum_{s=1}^t \frac{R(s)}{\delta} - r.
$$
Proof. Recall (2.21) is what allows us to change measure from $\mu_t$ to $\rho_t$. We will first need the definition of $y^H(t)$ for this case, with respect to $\mathcal{R}$. Rearranging (2.2) with respect to $Z(t+1)$, substituting into (2.12) and working backwards recursively gives us

$$y^H(t) = y(0) \exp \left\{ t\delta \left[ r - \frac{1}{2(1-\eta)^2}b'Q^{-1}b + \frac{1}{1-\eta}b'Q^{-1}\overline{b}_t \right] \right\}$$

where $\overline{b}_t$ is as defined in the proposition. By (2.21),

$$\rho_t(dH) = M^\gamma_t(H) \mu_t(dH)$$

$$= \frac{y^H(t)^{-\gamma}}{\mathbb{E}[y(t)^{-\gamma}|\mathcal{S}_t]} \mu_t(dH)$$

$$\propto y^H(t)^{-\gamma} \mu_t(dH)$$

$$\propto y(0)^{-\gamma} \exp \left\{ -\gamma t\delta \left[ r - \frac{1}{2(1-\eta)^2}b'Q^{-1}b + \frac{1}{1-\eta}b'Q^{-1}\overline{b}_t \right] \right\}$$

$$\times \exp \left\{ -\frac{1}{2} \left[ (b-b_t)'\Sigma_t^{-1}(b-b_t) \right] \right\}$$

$$\propto \exp \left\{ -\frac{1}{2} \left[ -\frac{\gamma t\delta}{1-\eta}b'Q^{-1}b + b'\Sigma_t^{-1}b + \frac{2\gamma t\delta}{1-\eta}b'Q^{-1}\overline{b}_t - 2b'\Sigma_t^{-1}b_t \right] \right\}$$

$$= \exp \left\{ -\frac{1}{2} \left[ b'Ab - 2b'B \right] \right\}$$

where $A \overset{\text{def}}{=} \Sigma_t^{-1} + \frac{-\gamma t\delta}{(1-\eta)^2}Q^{-1}$ and $B \overset{\text{def}}{=} \Sigma_t^{-1}b_t + \frac{-\gamma t\delta}{1-\eta}Q^{-1}\overline{b}_t$. This further simplifies to

$$\rho_t(dH) \propto \exp \left\{ -\frac{1}{2} [(b-v)'C^{-1}(b-v)] \right\}$$

(2.25)

where

$$C \overset{\text{def}}{=} A^{-1} = \left[ \Sigma_t^{-1} + \frac{-\gamma t\delta}{(1-\eta)^2}Q^{-1} \right]^{-1}$$

and

$$v \overset{\text{def}}{=} CB = \left[ \Sigma_t^{-1} + \frac{-\gamma t\delta}{(1-\eta)^2}Q^{-1} \right]^{-1} \left[ \Sigma_t^{-1}b_t + \frac{-\gamma t\delta}{1-\eta}Q^{-1}\overline{b}_t \right]$$

Due to the form in (2.25), which contains all of the relevant $b$ terms, we conclude that given the multivariate normal posterior $\mu_t(dH)$ and the benchmark process $y^H$, the tilted posterior $\rho_t(dH)$ is also multivariate normal with mean $\left[ \Sigma_t^{-1} + \frac{-\gamma t\delta}{(1-\eta)^2}Q^{-1} \right]^{-1} \left[ \Sigma_t^{-1}b_t + \frac{-\gamma t\delta}{1-\eta}Q^{-1}\overline{b}_t \right]$ and covariance matrix $\left[ \Sigma_t^{-1} + \frac{-\gamma t\delta}{(1-\eta)^2}Q^{-1} \right]^{-1}$.

Note the similarities between the hyperparameters in the tilted posterior and posterior (2.24). In a sense, a tilted posterior is indeed a posterior updated after observations. The
difference is that the “data” observed is mean annualized return, adjusted by the risk aversion of both the investor and benchmark. Furthermore, as $\eta \downarrow -\infty$, the tilted posterior is simply the posterior itself.

Based on Proposition 2.5.1, we can develop hyperparameter update rules for the tilted posterior $\rho_t$.

**Proposition 2.5.2.** Suppose the tilted posterior at time $t$ is distributed as

$$\rho_t(dH) \sim N(m_t, A_t).$$

After observing $B(t+1) = \frac{R(t+1)}{\delta} - r$, the tilted posterior at time $t+1$ is then

$$\rho_{t+1}(dH) \sim N(m_{t+1}, A_{t+1})$$

where

$$A_{t+1} = \left[ A_t^{-1} + \left( 1 + \frac{-\gamma}{(1-\eta)^2} \right) \delta Q^{-1} \right]^{-1}$$

and

$$m_{t+1} = \left[ A_t^{-1} + \left( 1 + \frac{-\gamma}{(1-\eta)^2} \right) \delta Q^{-1} \right]^{-1} \left[ A_t^{-1} m_t + \left( 1 + \frac{-\gamma}{1-\eta} \right) \delta Q^{-1} B(t) \right].$$

**Proof.** To show this, we will derive $\rho_t$ and $\rho_{t+1}$ from $\mu_t$ and $\mu_{t+1}$, respectively, and show that their parameters have been updated as stated in the proposition. Let $\mu_t(dH) \sim N(b_t, \Sigma_t)$. Then $\mu_{t+1}(dH)$ follows the distribution in (2.24), specifically,

$$\begin{align*}
\left\{ 
\begin{array}{l}
b_{t+1} = \left[ \Sigma_t^{-1} + \delta Q^{-1} \right]^{-1} \left[ \Sigma_t^{-1} b_t + \delta Q^{-1} B(t+1) \right] \\
\Sigma_{t+1} = \left[ \Sigma_t^{-1} + \delta Q^{-1} \right]^{-1}
\end{array}
\right.
\end{align*}$$

From Proposition 2.5.1 we know that

$$\rho_t \sim N \left( \left[ \Sigma_t^{-1} + \frac{-\gamma \delta t}{(1-\eta)^2} Q^{-1} \right]^{-1} \left[ \Sigma_t^{-1} b_t + \frac{-\gamma \delta t}{1-\eta} Q^{-1} B_t \right], \left[ \Sigma_t^{-1} + \frac{-\gamma \delta t}{(1-\eta)^2} Q^{-1} \right]^{-1} \right),$$

i.e.,

$$\begin{align*}
m_t &= \left[ \Sigma_t^{-1} + \frac{-\gamma \delta t}{(1-\eta)^2} Q^{-1} \right]^{-1} \left[ \Sigma_t^{-1} b_t + \frac{-\gamma \delta t}{1-\eta} Q^{-1} B_t \right] \\
A_t &= \left[ \Sigma_t^{-1} + \frac{-\gamma \delta t}{(1-\eta)^2} Q^{-1} \right]^{-1}.
\end{align*}$$
Now we will derive \( \rho_{t+1}(dH) \) from \( \mu_{t+1}(dH) \). Applying Proposition 2.5.1 to (2.26) gives us

\[
A_{t+1} = \left[ \Sigma_{t+1}^{-1} + \frac{-\gamma \delta (t + 1)}{(1 - \eta)^2} Q^{-1} \right]^{-1} \\
= \left[ \Sigma_{t}^{-1} + \delta Q^{-1} + \frac{-\gamma \delta (t + 1)}{(1 - \eta)^2} Q^{-1} \right]^{-1} \\
= \left[ \Sigma_{t}^{-1} + \frac{-\gamma \delta t}{(1 - \eta)^2} Q^{-1} + \delta Q^{-1} \frac{-\gamma \delta}{(1 - \eta)^2} Q^{-1} \right]^{-1} \\
= \left[ A_{t}^{-1} + \left( 1 + \frac{-\gamma}{(1 - \eta)^2} \right) \delta Q^{-1} \right]^{-1}
\]

and

\[
m_{t+1} = \left[ \Sigma_{t+1}^{-1} + \frac{-\gamma \delta (t + 1)}{(1 - \eta)^2} Q^{-1} \right]^{-1} \left[ \Sigma_{t+1}^{-1} b_{t+1} + \frac{-\gamma \delta (t + 1)}{1 - \eta} Q^{-1} b_{t+1} \right] \\
= A_{t+1} \left[ \Sigma_{t}^{-1} b_{t} + \delta Q^{-1} B(t + 1) + \frac{-\gamma \delta (t + 1)}{1 - \eta} Q^{-1} \sum_{s=1}^{t+1} \frac{B(s)}{t+1} \right] \\
= A_{t+1} \left[ \Sigma_{t}^{-1} b_{t} + \frac{-\gamma \delta t}{1 - \eta} Q^{-1} \sum_{s=1}^{t} \frac{B(s)}{t} + \delta Q^{-1} B(t + 1) + \frac{-\gamma \delta}{1 - \eta} Q^{-1} B(t + 1) \right] \\
= \left[ A_{t}^{-1} + \left( 1 + \frac{-\gamma}{(1 - \eta)^2} \right) \delta Q^{-1} \right]^{-1} \left[ A_{t}^{-1} m_{t} + \left( 1 + \frac{-\gamma}{1 - \eta} \right) \delta Q^{-1} B(t + 1) \right]
\]

thus completing our proof.

\( \Box \)

2.5.2 Solution

With an understanding of how the tilted posterior is updated, we are now poised to solve the case where the investor knows the covariance matrix \( Q \) and not the mean annualized log-returns \( b \), but can apply a prior on the mean. Recall (2.22),

\[
e^{\gamma h(t, S_t)} = \max_{\pi} \mathbb{E}_{\rho_t} \left\{ \exp \left\{ \gamma \delta q(\pi, H) s + \gamma \sqrt{\delta} u(\pi, H) \sigma Z(t + 1) + \gamma h(t + 1, S_{t+1}) \right\} \right| S_t \}.
\]

The difficulty in deriving the solution lies in this recursive definition of \( h(t, S_t) \). Given all of our assumptions up to this point, we make one more, which is that \( h(t, S_t) \) is quadratic. Specifically, we assume \( h(t, S_t) \) takes the form \( m_{t+1}' M_{t+1} m_{t+1} \), where \( M_{t+1} \) is an \( n \times n \) matrix and \( m_{t+1} \) is the mean of the tilted posterior at time \( t + 1 \). With this assumption, we can use induction to work backwards from \( t = T \), where \( h(T, S_T) = 0 \iff M_T = 0, c_T = 0 \), to find \( h(t, S_t) \). With \( h(t, S_t) \) defined, deriving the optimal policy \( \pi^* \) from (2.22) becomes straightforward.
The optimal policy for this case is

$$\pi^* = \left(1 + \frac{-\gamma}{1-\eta}\right) \left[\begin{array}{c} P_{t+1}^{-1} - \gamma Q - \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 Q P_{t+1} L_{t+1} \\ L_{t+1} + \gamma M_{t+1} + \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 L_{t+1} P_{t+1} M_{t+1} \end{array}\right] A_t^{-1} m_t$$

(2.27)

where

$$P_{t+1} = \left[Q - \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right) M_{t+1}\right]^{-1}$$

$$L_{t+1} = \left[A_t^{-1} + \delta \left(1 + \frac{-\gamma}{1-\eta}\right) Q^{-1} - \delta \left(1 + \frac{-\gamma}{1-\eta}\right)^2 P_{t+1}\right]^{-1}.$$

and the backwards recursive definition of $M_t$ is

$$M_t = M_{t+1} - \frac{1}{\gamma} A_t + \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 M_{t+1} P_{t+1} M_{t+1}$$

$$+ \frac{1}{\gamma} \left[I + \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 P_{t+1} M_{t+1}\right]' L_{t+1} \left[I + \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 P_{t+1} M_{t+1}\right]$$

$$+ \delta \left(1 + \frac{-\gamma}{1-\eta}\right)^2 \left[L_{t+1} + \gamma M_{t+1} + \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 L_{t+1} P_{t+1} M_{t+1}\right]'$$

$$\times \left[(P_{t+1} Q P_{t+1})^{-1} - \gamma P_{t+1}^{-1} - \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 L_{t+1}\right]^{-1}$$

$$\times \left[L_{t+1} + \gamma M_{t+1} + \delta \gamma \left(1 + \frac{-\gamma}{1-\eta}\right)^2 L_{t+1} P_{t+1} M_{t+1}\right].$$

### 2.5.3 Numerical Experiments

We performed many simulations to test different facets of our model, from prior sensitivity to the choice of $\gamma$, but in the interest of brevity, we present our main results: the performance of Bayesian regret against the ideal benchmark under different market conditions.

Our market consists of 3 stocks and a money market account with continuously compounded interest rate of $r = 0.01$. We considered four alternative models,

$$\mathcal{H} = \{(b_1, Q), (b_2, Q), (b_3, Q), (b_4, Q)\},$$

for our risky assets. Note that each model has the same covariance matrix

$$Q = \begin{bmatrix} 0.036769309 & -0.02754215 & 0.001618045 \\ -0.027542150 & 0.07462111 & -0.029954251 \\ 0.001618045 & -0.02995425 & 0.017116896 \end{bmatrix}.$$
Each simulation used one of four configurations of excess returns, given by

\[
b_1 = \begin{bmatrix} 0.08 \\ 0.05 \\ 0.07 \end{bmatrix}, \quad b_2 = \begin{bmatrix} -0.08 \\ -0.05 \\ -0.07 \end{bmatrix}, \quad b_3 = \begin{bmatrix} 0.08 \\ -0.05 \\ 0.07 \end{bmatrix}, \quad b_4 = \begin{bmatrix} -0.08 \\ 0.05 \\ -0.07 \end{bmatrix}.
\]

For each \( b_i \), we performed 1000 simulations where each simulation consisted of one trading year (252 days) of trading data generated by \((b_i, Q)\). We adopted a daily rebalancing period.

For every simulation run, we observed two processes: the Bayesian regret (BR) investor’s wealth, and the “perfect information” (PI) investor’s wealth. The BR model invests according to our solution (2.27) while the PI investor has foreknowledge of the true value of \( b_i \) and simply applies the static policy of (2.11). Note that the PI investor’s process is the hidden process \( y(t) \) in our model. Because we know the true parameters, we can observe this process and find the true value of the objective.

We chose the parameters \( \eta = -350 \) and \( \gamma = -80 \) for the PI investor’s utility function and the BR investor’s comparison function, respectively. We picked \( \eta \) so that the benchmark investor would yield a 1-3% annual return whereas \( \gamma \) was picked to be less risk-averse in order to allow the investor to have the flexibility to use leverage to outperform the benchmark when necessary.

The BR model was given the zero vector for the prior mean and a diagonal covariance matrix with a very large variance. In other words, we gave our investor the difficult task of managing his/her respective portfolios with, effectively, no prior information.

Table 2.1 summarizes the results of these simulations. Using the terminal wealths of each process, we calculated the proportion of simulations where BR outperformed PI. We also calculated the mean of the log ratios of the two processes. Specifically, we took the mean of \( \log \left( \frac{x(T)}{y(T)} \right) \), which quantifies BR’s performance against PI over all of the simulations. We note that performing on par with the benchmark (taking values close to zero) would not be a bad result because the benchmark has the advantage of knowing the correct parameters of the problem.

In the market models where all of the assets are experiencing overall gains \((b_1)\) or losses \((b_2)\), BR picks up the pattern and outperforms PI in approximately 90% of the simulations. Not only did it outperform the benchmark, it consistently did so as the means of the log ratios were positive and performed well, with 0.05749 and 0.05542 for \( b_1 \) and \( b_2 \), respectively. Simulations where the means have mixed signs \((b_3 \text{ and } b_4)\) showed BR outperforming PI in over 65% of the simulations, but not as definitively as in the previous two cases, with means of the log ratios at 0.01983 and 0.01902 for \( b_3 \) and \( b_4 \), respectively. The worst relative performance by BR against PI, over all simulations, is obtaining 97% of PI’s final wealth.
CHAPTER 2. DYNAMIC PORTFOLIO CHOICE WITH BAYESIAN REGRET

Table 2.1: Simulation results of model performance under various market conditions

<table>
<thead>
<tr>
<th>Simulation Model</th>
<th>BR vs. PI</th>
<th>$\mathbb{E} \left[ \log \left( \frac{y(T)}{y(T)} \right) \right]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1 = \begin{bmatrix} 0.08 \ 0.05 \ 0.07 \end{bmatrix}$</td>
<td>0.902</td>
<td>0.05749</td>
</tr>
<tr>
<td>$b_2 = \begin{bmatrix} -0.08 \ -0.05 \ -0.07 \end{bmatrix}$</td>
<td>0.889</td>
<td>0.05542</td>
</tr>
<tr>
<td>$b_3 = \begin{bmatrix} 0.08 \ -0.05 \ 0.07 \end{bmatrix}$</td>
<td>0.655</td>
<td>0.01983</td>
</tr>
<tr>
<td>$b_4 = \begin{bmatrix} 0.05 \ -0.07 \end{bmatrix}$</td>
<td>0.665</td>
<td>0.01902</td>
</tr>
</tbody>
</table>

In summary, Bayesian regret performed well against the perfect information benchmark under various market conditions when given no prior information. This is very encouraging not only because it outperformed the benchmark, but because the benchmark was given the best case scenario of knowing the correct mean of returns. Furthermore, this benchmark is the hidden process $y(t)$ as defined in (2.14), which is precisely the process our model was designed to outperform. This demonstrates the ability of Bayesian regret to utilize learning to leverage appropriately in order to perform well relative to our benchmark of interest.

2.5.4 Backtests

Several backtests were performed in a setup similar to our numerical experiments in the previous section. Our market consisted of three stocks, GE, ExxonMobil, and Apple, and a money market account with continuously compounded interest rate of $r = 0.01$. We obtained daily closing price data for the three assets from 1985 through 2009 and implemented both our Bayesian regret model as well as the classical Bayesian model for various subsets of the data. Each model was given $100 with which to invest. We used various lengths of time to “learn” the covariance matrix and establish a prior mean for each backtest simulation. A period of time from one to three years were then treated as “future” observations and simulated the sample paths the two models accordingly.

In cases where the returns of the assets were relatively stationary, Bayesian regret outperformed its classical Bayesian counterpart in terminal wealths. An example of this would
be when the tests spanned 1997-1999 with a learning period of 1985-1996 to establish the covariance matrix and prior mean, as seen in Figure 2.1. This three year span turns out to be a profitable and interesting period for all three assets. GE’s revenues surpassed $100 Billion in 1998. Exxon’s merger with Mobil took place during this time, which may account for some of the high variance movement, but certainly had a positive impact on the asset overall. Following the return of Steve Jobs as interim CEO in September of 1997, Apple began experiencing a comeback with the introduction of the iMac and iBook.

The 1997-1999 case demonstrates the strength of the model when the model assumptions are mostly satisfied. Bayesian regret manages to have a slightly higher portfolio value but never drifts far away from the classical Bayesian model over the first half of the simulation due to abrupt jumps in asset prices. This non-stationary behavior causes Bayesian regret’s portfolio value to even drop below that of classical Bayesian approximately 450 days into the backtest simulation. From this point on, however, the returns of all three assets begin to stabilize, a pattern that Bayesian regret was better equipped to use to its advantage, especially in the final 50 days of trading. This example indicates that a natural extension to our model would be to consider non-stationarity in the mean of the log-returns. Overall we are satisfied with the performance of this model under conditions that closely match our modeling assumptions.
Figure 2.1: Backtest of Bayesian regret and classical Bayesian models over the period of 1997-1999 in a market containing only GE, ExxonMobil, and Apple stocks and a risk free asset with $r = 0.01$. 
Chapter 3

A Bayesian Graphical Generalization for the Black-Litterman Model

3.1 Introduction

In applications from business to medicine, many decisions are made by taking both data and professional opinions into account. The complexity of this type of decision making increases when multiple professional views are expressed. Assuming each expert’s performance is consistent and recorded, one can learn the reliability of each expert’s view and consolidate this evaluation with the data available. In finance, a seminal model that sought to combine expert views with market data is the Black-Litterman model.

The Black-Litterman model is a portfolio allocation model developed by Fischer Black and Robert Litterman in Goldman Sachs, first published in [8] and given more elaboration in [7]. The Black-Litterman model establishes a posterior distribution for future returns by combining the market equilibrium distribution and tilting the distribution with respect to expert views. Given the posterior distribution, the allocation rule is decided by the user. This paper will focus on the estimation aspect and simply employ a mean-variance portfolio to judge performance.

A comprehensive literature review and detailed implementation of the model can be found in [31]. What follows summarizes much of the overview provided by Walters. The original papers by Black and Litterman were important but also lacking in the details of its implementation. The derivation utilized Theil’s mixed estimation model described in [30] and it was not until [27] and [11] that the Black-Litterman model started to be viewed in a Bayesian framework.

One instrumental aspect of the Black-Litterman model is its choice of the equilibrium market portfolio as the prior. Utilizing historical returns can yield unreasonably extreme portfolios while uninformative priors are lacking in an “intuitive connection back to the market,” according to Walters. The equilibrium market portfolio is the portfolio that “clears the market” and is thus a natural, market-neutral starting point for portfolio construction.
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

Black and Litterman use the capital asset pricing model and reverse engineering to obtain these distributions. In-depth justification for the use of the equilibrium market portfolio as the prior can be found in [7].

The second major contribution of the Black-Litterman model is its incorporation of expert views to establish posterior distributions of excess returns. By utilizing Bayes’ rule, Black and Litterman developed a natural framework for combining the market equilibrium with expert views. In a Bayesian framework, the Black-Litterman model is a straightforward implementation of conjugate normal distributions, the details of which we will cover in Section 3.2.1.

Much of the literature has been primarily concerned with the proper tuning of the Black-Litterman model’s parameters. Due to CAPM’s limitations, alternatives for modeling the equilibrium market portfolio have been explored, such as in [19], where the authors developed an extension of Black-Litterman that incorporated additional uncorrelated market factors. Specification of the link matrix can either depend on a market capitalization weighted method as in [17] and [18], or with equal weights as in [28]. How the user specifies the covariance matrix of the views can be very diverse, ranging from setting it to be proportional to the prior as is used in [17] and [25], to the use of confidence intervals [18] and factor models ([4]). Furthermore, the parameters $\tau$ in Black and Litterman’s original formulation has been a point of confusion. A detailed discussion regarding this parameter can be found in [32].

Estimation of the market’s covariance matrix is also a concern that [21] and [22] address by detailing the estimation process used at Goldman Sachs.

To our knowledge, there has been little work done in extending or generalizing the Black-Litterman model to account for shortcomings in the model beyond the misspecification of inputs. One work of note is that of [6], which views the Black-Litterman model in an inverse optimization framework. This interpretation allowed them to generalize the model in several respects, ranging from a complete characterization of the inputs to allowing for more general views and relaxing distributional assumptions on market returns.

Our contribution in our extension is twofold. First, we are interested in learning about the experts, specifically whether these experts may have some underlying bias in their views. Secondly, we account for potential errors in the market equilibrium. Instead of improving the input of the equilibrium distribution, such as in [19], we allow for the use of a suboptimal estimate of the market equilibrium and utilize a correction term to account for errors. Generally speaking, our model, which we will refer to as the generalized Black-Litterman model, allows for Black-Litterman to be implemented, but corrects for potential errors via Bayesian inference.

The remainder of this paper is as follows: Section 3.2 covers the background of the Black-Litterman model, its graphical representation and Gibbs sampling. Section 3.3 introduces the generalized Black-Litterman model. We end by discussing numerical results in Section 3.4.
3.2 Preliminaries

In the interest of keeping this paper self-contained, we will briefly cover the Black-Litterman model, its graphical representation and Gibbs Sampling.

3.2.1 The Black-Litterman Model

The Black-Litterman model estimates future returns of assets given two components: the market equilibrium and expert views. By combining these two pieces, the Black-Litterman model is able to limit the sensitivity of the optimal allocation function to the input parameters. To do this, Black-Litterman cleverly utilizes Bayes’ rule to provide a posterior distribution of the future returns.

**General Definition**

Consider a market whose returns are represented by the multivariate random variable \( X \). Assume the distribution of \( X \) is known via its probability density function \( f_X \). One then obtains an expert opinion from an investor, whom will provide an assessment of the outcome of the market, \( V \). Specifically, when \( X \) takes a realization \( x \), the investor’s view \( V \) is modeled to be a perturbation of the realization \( x \), and thus the view is the conditional distribution \( V|x \). Furthermore, the investor may not express an opinion regarding the entire vector \( X \) but rather a subset defined by the function \( g(X) \). Thus, conditional on a realization \( x \), the view is of the form \( V|g(x) \) with conditional p.d.f. \( f_{V|g(x)} \).

Using Bayes’ rule, one can calculate the distribution of the market conditioned on the investor’s opinion \( X|v \):

\[
 f_{X|v}(x|v) = \frac{f_{V|g(x)}(v|x)f_X(x)}{\int f_{V|g(x)}(v|x)f_X(x)dx}, \tag{3.1}
\]

Thus in a Bayesian context, one can interpret the distribution \( f_X \) to be the prior, \( f_{V|g(x)} \) to be the likelihood and \( f_{X|v} \) to be the posterior distribution.

**Linear Expertise on Normal Markets**

Utilizing conjugate distributions, this model becomes tractable and straightforward to solve. [7] used these modeling assumptions in their original model and generalized Black-Litterman will be an extension of this version. Suppose the \( N \)-dimensional market vector \( X \) is normally distributed, i.e.,

\[ X \sim N(\mu, \Sigma). \]

The mean and covariance are assumed to be known or obtainable via estimation techniques. A common choice of \( \mu \) is the expected returns vector obtained through the capital asset pricing model (CAPM).
Suppose the investor’s area of expertise is a linear function of the market, specifically,
\[ g(x) \equiv Px \]
where \( P \) is the “link” or “pick” matrix. Given there are \( K \) views, \( P \) is a \( K \times N \) matrix where each row corresponds to one view and selects the linear combination of the assets involved in the view. How \( P \) is defined can depend on a market capitalization weighted method, as proposed in [17] and [18], or with equal weights as in [28]. Because the details regarding the construction of the link matrix do not directly impact our discussion of Black-Litterman and generalized Black-Litterman, we refer the reader to the above references for further information.

Investor views can either be absolute or relative. Absolute views typically express an opinion of the return of a single asset, e.g., “Apple’s return will be 10%.” On the other hand, relative views tell about the difference between two or more assets, e.g., “Amazon will outperform IBM by 2%.” Relative views are more common among practitioners but knowledge of the current expected returns of the assets in question are required to understand their implications. In our example of Amazon outperforming IBM by 2%, if Amazon had an expected return that was 8% and IBM’s was 5% prior to this view, then our view implies that Amazon has lost ground to IBM, despite still outperforming IBM. On the other hand, if Amazon’s expected return was 6% instead of 8%, then the view implies that Amazon has gained ground on IBM.

Suppose our market contains only Apple, Amazon and IBM as assets and is indexed in this order. The two views we stated above would then be represented in the link matrix \( P \) as
\[
P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}.
\]
Rows expressing absolute views sum to 1 whereas rows expressing relative views will sum to 0.

The conditional distribution of the investor’s views given the outcome of the market is assumed to be normal,
\[ V|Px \sim N(Px, \Omega) \]
where the symmetric and positive matrix \( \Omega \) corresponds to the confidence in the investor’s view. There are several ways to approach defining \( \Omega \), but generally this is determined at the discretion of the practitioner. Lastly, the investor’s view will be a realization \( v \) derived from the distribution \( V|Px \). In other words, the investor’s view is a noisy observation with a mean centered around the relevant realized returns.

Using (3.1), it can be shown that the Black-Litterman distribution is normal:
\[ X|v \sim N(\mu_{BL}, \Sigma_{BL}), \]
where
\[
\mu_{BL}(v, \Omega) = (\Sigma^{-1} + P'\Omega^{-1}P)^{-1}[\Sigma^{-1}\mu + P'\Omega^{-1}v],
\]
\[
\Sigma_{BL}(\Omega) = (\Sigma^{-1} + P'\Omega^{-1}P)^{-1}.
\]
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

$$r \sim N(\mu, \Sigma)$$

$$V \sim N(Pr, \Omega)$$

Figure 3.1: Graphical representation of the classical Black-Litterman model.

Note that this form is consistent with the Bayesian interpretation of Black-Litterman that sets the market distribution as a normally distributed prior and the view as normal likelihood, yielding the above hyperparameters for the posterior distribution. Another way to think of this is that the posterior is an average of the prior mean and the view weighted by the confidence of each, as determined by the inverse of the covariance matrices. A note regarding this formulation compared to [7] is that we have dropped the scalar $\tau$, which represents uncertainty in the market equilibrium vector, in favor of incorporating it into $\Sigma$. We refer the reader to [32] for further discussion regarding the role and choice of $\tau$.

It is also worth noting that multiple expert views could be incorporated in the classical Black-Litterman model. The link matrices of multiple experts can be consolidated into one united link matrix. The covariance of the experts will need to be given but as long as it is well defined, generalizing to multiple experts is of little concern.

**Graphical Representation of Black-Litterman**

Bayesian graphical models are acyclic graphs where nodes represent random variables and directed arcs represent conditional independence. These are useful representations of what may otherwise be complex models and allow for straightforward applications of probabilistic inference. Because generalized Black-Litterman is an extension of Black-Litterman, and thus is more complex, this representation will be a useful tool to better understand it. An additional note with respect to the graphical representations in this paper is that we will depict the nodes of unobserved random variables with circles and the nodes of observed random variables with squares.

The linear expertise example of the Black-Litterman model from Section 3.2.1 is shown as a Bayesian graphical model in Figure 3.1. The Black-Litterman model has a straightforward graphical representation, one that we will expand upon greatly in generalized Black-Litterman.
3.2.2 Gibbs Sampling

Black-Litterman is a useful model due to its tractability and can exploit the use of conjugate distributions in the oft-used linear expertise model outlined in Section 3.2.1. Due to the complexity of generalized Black-Litterman, the analytical form of our posterior is intractable. We thus turn to Gibbs Sampling, a numerical algorithm that generates samples of high dimensional joint distributions, to sample from the posterior distribution.

Gibbs sampling is a Markov chain Monte Carlo algorithm that allows the user to obtain samples from a joint distribution whose direct samples may be difficult to obtain. This technique takes advantage of the fact that sampling from a conditional distribution is simpler than sampling from a joint distribution.

Suppose we would like to obtain samples of the multivariate random variable \( X = (x_1, x_2, \ldots, x_n) \) with joint distribution \( P(x_1, x_2, \ldots, x_n) \). Let \( X^{(i)} = (x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}) \) be the \( i \)th sample. The Gibbs sampling algorithm is as follows.

**Initialize:** Choose initial values for \( X^{(0)} \)

For iteration \( i + 1 \) in the algorithm, we are given \( X^{(i)} \) and construct \( X^{(i+1)} \)

**Sample:** Sample each element \( x_j^{(i+1)} \) from the conditional distribution

\[
P \left( x_j \left| x_1^{(i+1)}, x_2^{(i+1)}, \ldots, x_{j-1}^{(i+1)}, x_{j+1}^{(i)}, \ldots, x_n^{(i)} \right. \right)
\]

**Repeat:** Repeat the sampling step to obtain as many observations of \( X^{(i)} \) as necessary.

Given these samples, one can then sample from the marginal distributions by only considering the samples of the corresponding subset. Furthermore, expected values and other statistics may be obtained by taking the sample statistics of the Gibbs samples.

Due to the Markov chain structure of this process, there will be correlation between each sample and thus, it is common to use every \( n \)th sample instead of every sample in order to obtain samples with less dependence. Additionally, it is also common to omit the first samples of the algorithm and have a “burn-in period” to mitigate the effect the initial values may have on the rest of the sample. We refer the reader to [15] for more details on the determining a proper burn-in period.

Given advances in computational power, Gibbs sampling has become a viable and useful tool for generating samples of joint distributions.

3.3 Graphical Generalized Black-Litterman Model

Before we discuss the details of the generalized Black-Litterman model, we summarize the features of both Black-Litterman and generalized Black-Litterman in Table 3.1. Both models have the same goal of estimating the return, but differ in assumptions and available
<table>
<thead>
<tr>
<th>Model Traits</th>
<th>Black-Litterman Model</th>
<th>Generalized Black-Litterman Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected return</td>
<td>Expected returns equal the CAPM estimate:</td>
<td>Expected returns do not equal the CAPM estimate:</td>
</tr>
<tr>
<td></td>
<td>• Expected returns (μ) are constant</td>
<td>• Expected returns μ_t vary every period</td>
</tr>
<tr>
<td></td>
<td>• Expected returns are assumed to equal CAPM estimate (μ = α)</td>
<td>• CAPM estimate α contains systemic bias b_μ</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Expected return will have mean centered around CAPM with the bias offset (μ_t ∼ N(α + b_μ, β))</td>
</tr>
<tr>
<td>Returns</td>
<td>• Returns are normally distributed with mean μ and known covariance Σ</td>
<td>• Returns are normally distributed with mean μ_t, which varies every period, and known covariance Σ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expert Views</td>
<td>• Views (V) are unbiased</td>
<td>• Views (V_t) may be biased (b_V_t)</td>
</tr>
<tr>
<td></td>
<td>• Views are “centered” around next period’s realized returns (V</td>
<td>r ∼ N(Pr, Ω))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Bias is normally distributed with constant mean and covariance (b_V_t ∼ N(δ_V, Θ_V))</td>
</tr>
<tr>
<td>Unknown Variables</td>
<td>• r: return</td>
<td>• r_t: return</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• μ_1, . . . , μ_t: expected returns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• b_μ: error in CAPM estimate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• b_V_1, . . . , b_V_t: bias in expert views</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• δ_V, Θ_V: mean and covariance of expert bias</td>
</tr>
<tr>
<td>Known Variables</td>
<td>• μ: expected return, also the CAPM estimate</td>
<td>• r_1, . . . , r_t−1: history of returns</td>
</tr>
<tr>
<td></td>
<td>• V: reported expert view</td>
<td>• V_1, . . . , V_t: historical and current expert view</td>
</tr>
<tr>
<td></td>
<td>• Σ: covariance of return</td>
<td>• Σ: covariance of return</td>
</tr>
<tr>
<td></td>
<td>• Ω: covariance of expert views</td>
<td>• Ω_1, . . . , Ω_t: historical and current expert covariance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• α: CAPM estimate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• β: covariance of expected return</td>
</tr>
</tbody>
</table>

Table 3.1: Summary of the traits of the classical Black-Litterman model and the generalized Black-Litterman model.
information. In formulating generalized Black-Litterman, we looked into how assumptions about the expected returns, returns and expert views may be incorrect and accounted for them accordingly. Under these new assumptions, we extended the model to have access to a history of expert views and returns in order to better estimate the return.

We will now detail the steps in extending Black-Litterman. The first aspect we consider is what occurs in a multi-period setting. As a graphical model, shown in Figure 3.2, there is a noticeable change in comparison to Figure 3.1: the returns prior to period \( t \) are now observed and known. Given this history of view-return pairs, inference may be performed on the expert whose views were previously expressed. A reason for why we may want to learn more about our expert is that the allegedly unbiased expert may, in fact, be a biased investor that favors certain assets over others. By performing inference on the expert, we can learn about the tendencies of the expert and account for them in the final estimation.

We believe that view bias would not be identical every period, but should have an overall tendency. Thus, we model the bias of the views of each period, \( b^V_t \), to be a normal random variable with parameters (\( \delta^V, \Theta^V \)). Experts then report a view based around both the actual return \( r_t \) and the bias \( b^V_t \). The graphical model thus extends to the one found in Figure 3.3.

After modeling the expert bias, we now consider whether there may be other types of errors in the “top half” of our graphical model. Classically, the returns are modeled to be random variables whose parameters are based on the market portfolio. Because the market portfolio is unobservable, the market equilibrium needs to be estimated, which leads to uncertainty in this estimate. A common method for estimating the market equilibrium is using the capital asset pricing model (CAPM). CAPM, however, has been shown to have poor empirical success. A detailed review of CAPM summarizing the model and its criticisms can be found in [14]. To account for uncertainty in the expected returns, [8] uses the scalar \( \tau \) to scale the covariance matrix of the expected returns vector. All these factors point toward a need to account for errors in the estimation of the market equilibrium.

Our approach to modeling uncertainty in the expected returns is two-fold. First, we assume that the expected return is itself a normal random variable with a new realization every period, \( \mu_t \). Secondly, we add a correction term to the mean of this random variable,
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

Figure 3.3: Graphical representation of generalized Black-Litterman accounting for expert view bias.

$b^\mu$, making the expected returns a sum of the reported market equilibrium vector plus the correction term. These modeling assumptions extend Figure 3.2 to the graphical model found in Figure 3.4.

The entire generalized Black-Litterman model, when viewed as a graphical model, combines Figures 3.3 and 3.4 to obtain 3.5. With this blueprint in mind, we can now state the specifics of the generalized Black-Litterman model.

3.3.1 Model Assumptions

Suppose we have a market that consists of $N$ assets and a risk-free asset. Excess returns of the assets, $r_t \in \mathbb{R}^N$, evolve over discrete time periods $t = 1, 2, \ldots$ and are observed at the end of each period. Define $\mathcal{R}_t \equiv \{r_1, r_2, \ldots, r_t\}$ to be the history of observed excess returns up to, and including period $t$. Let the mean of the excess returns be $\mu_t \in \mathbb{R}^N$, which itself is a normal random variable such that

$$\mu_t \sim N(\alpha + b^\mu, \beta),$$
where $b^\mu \in \mathbb{R}^N$ is an unknown correction term. In our setting, $\alpha \in \mathbb{R}^N$ will be the estimated expected market returns, e.g., expected returns calculated from CAPM, and thus is specified by the user. Similarly, we assume $\beta \in \mathbb{R}^{N \times N}$ is given or estimated. Furthermore, the returns $r_t$ are normally distributed as

$$r_t \sim N(\mu_t, \Sigma)$$

and the covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$ is assumed to be known. With these changes, the graphical model would extend to the graph found in Figure 3.4.

To allow for inference on the bias of expert views, we will assume our expert will express his/her $K$ views on the same selection of assets every period. In other words, our expert uses the same link matrix $P \in \mathbb{R}^{K \times N}$ for every period. Suppose our investor’s views $V_t \in \mathbb{R}^K$ for period $t$ have mean $Pr_t + b^V_t$, where $b^V_t$ is a $K$-dimensional, normal random variable whose parameters are constant but unknown. Additionally, we allow for the confidence of the views to vary every period, and denote the covariance matrix of the views for period $t$ as $\Omega_t$. In other words,

$$V_t|r_t \sim N(Pr_t + b^V_t, \Omega_t)$$

where

$$b^V_t \sim N(\delta^V, \Theta^V)$$
Figure 3.5: Graphical representation of generalized Black-Litterman, which accounts for violations in Black-Litterman’s assumptions regarding the market equilibrium and expert views.
and the parameters \((\delta^V, \Theta^V)\) are unknown. The graphical representation of this extension to the Black-Litterman model can be found in Figure 3.3.

The entire graphical model can be found in Figure 3.5. The goal of generalized Black-Litterman, at the start of time \(t\), is to calculate the distribution of \(r_t\) given the expert view \(V_t\) and the history of returns \(R_{t-1}\). What makes solving this model challenging is that \(r_t\) is no longer the only unobserved random variable.

### 3.3.2 Conjugate Priors and Hyperparameters

Despite not directly observing them, we can still perform Bayesian inference on the newly introduced unknown variable \(b^\mu\) and the hidden process \(B^\mu \equiv \{b^\mu_1, b^\mu_2, \ldots, b^\mu_t\}\), provided we specify prior distributions for these quantities. To further simplify the process, we will define prior distributions that are conjugate distributions. Specifically, let the prior for \(b^\mu\) be \(N(\delta^\mu_0, \Theta^\mu_0)\). This definition will allow for the posterior distribution of \(b^\mu\) to also be normally distributed.

The \(b^\mu_j, j = 1, 2, \ldots, t\) are modeled to be normally distributed with unknown mean and covariance matrix \((\delta^V, \Theta^V)\). The conjugate distribution for this case is a normal inverse-Wishart distribution, which is a generalization of the gamma distribution. Let us set the prior on \((\delta^V, \Theta^V)\) to be \(NIW(\eta^V_0, \kappa^V_0, \nu^V_0, \Lambda^V_0)\). What this means is that the covariance matrix \(\Theta^V\) has a prior inverse-Wishart distribution with parameters \(\nu^V_0 \in \mathbb{R}, \Lambda^V_0 \in \mathbb{R}^{K \times K}\) and, given a particular value of \(\Theta^V\),

\[
\delta^V \sim N\left(\eta^V_0, \frac{1}{\kappa^V_0} \Theta^V\right)
\]

where \(\eta^V_0 \in \mathbb{R}^{K}\) and \(\kappa^V_0 \in \mathbb{R}\).

Specification of the parameters is determined by the user. If one’s prior belief is to be consistent with the classical Black-Litterman framework, then the prior means \(\delta^\mu_0\) and \(\eta^V_0\) would simply be zero vectors. The respective prior variability may then be determined based on one’s confidence in this belief.

### 3.3.3 Gibbs Sampling Algorithm

Using Figure 3.5 as a guide, we can now develop our Gibbs sampling algorithm for generalized Black-Litterman.

The quantity we would like to estimate is

\[
P \left( r_t | r_1, r_2, \ldots, r_{t-1}, V_1, V_2, \ldots, V_t \right), \tag{3.2}
\]

the posterior distribution of \(r_t\) given the observations of the returns and views up to the start of period \(t\). Because there are several random variables we cannot observe in addition to \(r_t\), we will use Gibbs sampling to sample from the joint distribution of all unknown random variables in generalized Black-Litterman,

\[
P \left( r_t, b^\mu, \mu_1, \ldots, \mu_t, b^V_1, \ldots, b^V_t, \delta^V, \Theta^V | r_1, r_2, \ldots, r_{t-1}, V_1, V_2, \ldots, V_t \right). \tag{3.3}
\]
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

Given samples of \((r_t, b^\mu, \mu_1, \ldots, \mu_t, b^V_1, \ldots, b^V_t, \delta^V, \Theta^V)\), we can obtain samples of the marginal distribution of \(r_t\) by collecting the \(r_t\) elements within these Gibbs samples. The samples from this marginal distribution are precisely samples from the posterior distribution (3.2).

Define the \(i^{th}\) Gibbs sample to be

\[
X^{(i)} \equiv \left( b^{\mu(i)}, \mu_t^{(i)}, r_t^{(i)}, B_t^{V(i)}, \delta^{V(i)}, \Theta^{V(i)} \right) \forall i,
\]

where

\[
\mu_t^{(i)} \equiv \{ \mu_1^{(i)}, \mu_2^{(i)}, \ldots, \mu_t^{(i)} \},
\]

\[
B_t^{V(i)} \equiv \{ b_1^{V(i)}, b_2^{V(i)}, \ldots, b_t^{V(i)} \}.
\]

Given \(\{\alpha, \beta, \Sigma, P, \Omega_t\}\) and prior hyperparameters \(\{\delta^\mu_0, \Theta^\mu_0, \eta^V_0, \kappa^V_0, \nu^V_0, \Lambda^V_0\}\), the Gibbs sampling algorithm will then be as follows:

**Initialize** Choose initial \(X^{(0)}\)

For iteration \(i + 1\) in the loop, we are given \(X^{(i)}\) and construct \(X^{(i+1)}\) in the following manner:

**Step 1** Sample \(b^{\mu(i+1)} \sim N\left(\hat{\alpha}_{b^\mu}, \hat{\beta}_{b^\mu}\right)\) where

\[
\hat{\alpha}_{b^\mu} = (t\beta^{-1} + \{\Theta^\mu_0\}^{-1})^{-1} \left[ t\beta^{-1} \left( \bar{\mu}_t^{(i)} - \alpha \right) + \{\Theta^\mu_0\}^{-1} \delta^\mu_0 \right]
\]

\[
\hat{\beta}_{b^\mu} = (t\beta^{-1} + \{\Theta^\mu_0\}^{-1})^{-1},
\]

and

\[
\bar{\mu}_t^{(i)} \equiv \frac{1}{t} \sum_{j=1}^{t} \mu_j^{(i)},
\]

the sample mean of \(\mu_t^{(i)}\).

**Step 2** Sample each component of \(\mu_t^{(i+1)}\) as \(\mu_j^{(i+1)} \sim N\left(\hat{\alpha}_{\mu,j}, \hat{\beta}_{\mu,j}\right)\) for \(j = 1, 2, \ldots, t\) where

\[
\hat{\alpha}_{\mu,j} = (\Sigma^{-1} + \beta^{-1})^{-1} \left[ \Sigma^{-1} r^* + \beta^{-1} \left( \alpha + b^{\mu(i+1)} \right) \right]
\]

\[
\hat{\beta}_{\mu,j} = (\Sigma^{-1} + \beta^{-1})^{-1},
\]

and

\[
r^* = \begin{cases} 
  r_j & \text{if } j = 1, 2, \ldots, t - 1 \\
  r_t^{(i)} & \text{if } j = t.
\end{cases}
\]
Step 3 Sample $r_t^{(i+1)} \sim N(\hat{\alpha}_r, \hat{\beta}_r)$ where

$$\hat{\alpha}_r = (P'\Omega_t^{-1}P + \Sigma^{-1})^{-1} \left[ P'\Omega_t^{-1}(V_t - b_t^{V(i)}) + \Sigma^{-1}\mu_t^{(i+1)} \right]$$

$$\hat{\beta}_r = (P'\Omega_t^{-1}P + \Sigma^{-1})^{-1}.$$

Step 4 Sample each component of $B_t^{V(i+1)}$ as $b_j^{V(i+1)} \sim N(\hat{\alpha}_{bV,j}, \hat{\beta}_{bV,j})$ for $j = 1, 2, \ldots, t$

where

$$\hat{\alpha}_{bV,j} = \left( \Omega_j^{-1} + \Theta^{V(i)} \right)^{-1} \left[ \Omega_j^{-1}(V_j - Pr^*) + \Theta^{V(i)}\delta^{V(i)} \right]$$

$$\hat{\beta}_{bV,j} = \left( \Omega_j^{-1} + \Theta^{V(i)} \right)^{-1}.$$ and

$$r^* = \begin{cases} r_j & \text{if } j = 1, 2, \ldots, t - 1 \\ r_t^{(i+1)} & \text{if } j = t. \end{cases}$$

Step 5 Sample $\delta^{V(i+1)}, \Theta^{V(i+1)} \sim NIW(\eta_t^V, \kappa_t^V, \nu_t^V, \Lambda_t^V)$ where

$$\eta_t^V = \frac{\kappa_0^V}{\kappa_0^V + t} \eta_0^V + \frac{t}{\kappa_0^V + t} \overline{b_t^{V(i+1)}}$$

$$\kappa_t^V = \kappa_0^V + t$$

$$\nu_t^V = \nu_0^V + t$$

$$\Lambda_t^V = \Lambda_0^V + \sum_{j=1}^t \left( b_j^{V(i+1)} - \overline{b_t^{V(i+1)}} \right) \left( b_j^{V(i+1)} - \overline{b_t^{V(i+1)}} \right)'$$

$$+ \frac{\kappa_0^V}{\kappa_0^V + t} \left( \overline{b_t^{V(i+1)}} - \eta_0^V \right) \left( \overline{b_t^{V(i+1)}} - \eta_0^V \right)'$$

and $\overline{b_t^{V(i+1)}} = \frac{1}{t} \sum_{j=1}^t b_j^{V(i+1)}$.

Repeat Repeat Steps 1–5 to generate as many samples of $X^{(i)}$ as necessary.

As one can see, the steps of our Gibbs sampling algorithm follow Figure 3.5 “from top to bottom.” Using Figure 3.5, the contribution of each component is clear and can be easily converted into the respective Gibbs sampling step.

### 3.4 Numerical Tests

To test generalized Black-Litterman against the classical Black-Litterman model, we performed controlled tests and backtests. Controlled testing allows for us to test both models
under correct modeling assumptions while backtesting provides a glimpse into how well the models perform with real data. We note that due to the difficulty of obtaining expert data, all views expressed in the simulations were generated by following the assumptions of generalized Black-Litterman.

3.4.1 Controlled Tests

For our controlled simulations, we simulated a market of three assets following the assumptions of the model. Each simulation consisted of 50 periods of “history” followed by one period of portfolio allocation of both the Black-Litterman model and generalized Black-Litterman. This process was repeated 500 times and the statistics of the returns of the last period were calculated.

We used the covariance matrix

$$
\Sigma = \begin{pmatrix}
0.015838413 & -0.01307935 & 0.001936589 \\
-0.013079347 & 0.03004989 & -0.012119245 \\
0.001936589 & -0.01211924 & 0.007054950 
\end{pmatrix}
$$

for the covariance of the returns, which results in standard deviations ranging from 8% to 17%. We chose \( \alpha = (0.09, 0.05, -0.04) \) and set \( \beta \) depending on the requirements of the test. To reduce the complexity of the tests, we set the link matrix \( P \) to be the identity, which means all views are absolute views.

We set priors such that the prior mean would be zero, i.e., we presumed there to be no bias in the views nor errors in the equilibrium estimate. However, we also allowed enough prior variance for the model to learn if this is not the case. The initialization step of the Gibbs sampling algorithm thus used relevant prior hyperparameters as well as known model parameters such as \( \alpha \).

There were three categories of controlled tests: 1) equilibrium assumptions violated with no views, 2) view assumptions violated while equilibrium assumptions hold, and 3) both assumptions violated.

Equilibrium Error

In the absence of expert views, both the Black-Litterman model and generalized Black-Litterman reduce to a portfolio allocation problem given the equilibrium distribution. The difference between the models is that Black-Litterman will simply use CAPM whereas generalized Black-Litterman will attempt to correct for any errors in the CAPM estimate based on historical performance. We simply set \( \Omega \) to be incredibly large in order for both models to ignore all views.

Another subtle difference between the extended Black-Litterman model and the classical model is that generalized Black-Litterman accounts for variance in the expected returns, modeled using the covariance matrix \( \beta \). Therefore, even when there is neither equilibrium
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

<table>
<thead>
<tr>
<th>Model</th>
<th>-0.03</th>
<th>-0.02</th>
<th>-0.01</th>
<th>0</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended BL</td>
<td>0.3565</td>
<td>0.6242</td>
<td>0.8742</td>
<td>1.2257</td>
<td>1.4269</td>
<td>1.6939</td>
<td>1.9603</td>
</tr>
<tr>
<td>Black-Litterman</td>
<td>0.1196</td>
<td>0.4612</td>
<td>0.7862</td>
<td>1.1936</td>
<td>1.4858</td>
<td>1.8273</td>
<td>2.1688</td>
</tr>
</tbody>
</table>

Table 3.2: Sharpe ratios of simulations with equilibrium error in the absence of expert views

error nor any views, the two models will still differ from each other due to this point. For these tests, we set \( \beta = 0.0005I \) to allow for a noticeable amount of variability which result in expected means that deviate up to 0.045 from \( \alpha + b^\mu \).

For our simulations involving errors in the equilibrium estimate, we set \( b^\mu = ce \), where \( e \) is the vector whose elements are all 1 and \( c \) is a constant we varied from test to test. We recorded the Sharpe ratios of each model following each test and summarize the results in Table 3.2. As mentioned previously, even when there is no equilibrium error \( (c = 0) \), the two models do not perform identically and, in fact, generalized Black-Litterman outperforms Black-Litterman.

For cases where \( c \) is negative, generalized Black-Litterman out-performed Black-Litterman, indicating that the extension managed to utilize historical data to adjust the portfolio allocation. As for when \( c \) is positive, Black-Litterman outperforms generalized Black-Litterman by a bit, which we attribute to generalized Black-Litterman’s built-in robustness when considering \( \beta \) in the portfolio allocation. Because expected returns are actually higher than reported, aggressive portfolio allocation reaps higher rewards than a robust portfolio allocation. In summary, generalized Black-Litterman accounts for equilibrium errors in a fair and expected manner.

View Error

We now look at the other aspect of generalized Black-Litterman, errors in the views. We eliminate all equilibrium error, which means \( b^\mu = 0 \) and \( \beta = 0 \), thus the true expected return is indeed \( \alpha \). We set the mean of the bias to be \( \delta^V = ce \) and varied \( c \) throughout the test. The variance of the bias and views were set to \( \Theta^V = 0.015^2I \) and \( \Omega_t = 0.02I \), respectively.

Similar to the previous section, we expect generalized Black-Litterman to differ from Black-Litterman even when \( \delta^V = 0 \) due to taking into account variance in the views. Table 3.3 summarizes the results of these tests for various values of \( c \). Across the board, generalized Black-Litterman performs very consistently whereas Black-Litterman was susceptible to view bias. This clearly demonstrates generalized Black-Litterman’s ability to be robust to view bias.
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

<table>
<thead>
<tr>
<th>Model</th>
<th>-0.2</th>
<th>-0.15</th>
<th>-0.1</th>
<th>-0.05</th>
<th>0</th>
<th>0.05</th>
<th>0.1</th>
<th>0.15</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extension</td>
<td>1.1993</td>
<td>1.2025</td>
<td>1.1785</td>
<td>1.2490</td>
<td>1.1871</td>
<td>1.2552</td>
<td>1.1941</td>
<td>1.2164</td>
<td>1.2153</td>
</tr>
<tr>
<td>BL</td>
<td>0.8047</td>
<td>0.8704</td>
<td>0.8588</td>
<td>1.0304</td>
<td>0.9719</td>
<td>1.1315</td>
<td>1.0666</td>
<td>1.1551</td>
<td>1.1852</td>
</tr>
</tbody>
</table>

Table 3.3: Sharpe ratios of simulations with view error and no equilibrium errors

<table>
<thead>
<tr>
<th>View Bias ($\delta^V$)</th>
<th>-0.10</th>
<th>-0.05</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.02</td>
<td>1.2317</td>
<td>1.2357</td>
<td>1.2418</td>
<td>1.2498</td>
<td>1.2593</td>
</tr>
<tr>
<td>-0.01</td>
<td>1.0943</td>
<td>1.0777</td>
<td>1.0647</td>
<td>1.0546</td>
<td>1.0466</td>
</tr>
<tr>
<td>0.00</td>
<td>1.0792</td>
<td>1.0522</td>
<td>1.0306</td>
<td>1.0133</td>
<td>0.9991</td>
</tr>
<tr>
<td>0.01</td>
<td>1.1870</td>
<td>1.1269</td>
<td>1.0979</td>
<td>1.0746</td>
<td>1.0557</td>
</tr>
<tr>
<td>0.02</td>
<td>1.4095</td>
<td>1.3434</td>
<td>1.2895</td>
<td>1.2454</td>
<td>1.2089</td>
</tr>
</tbody>
</table>

Table 3.4: Table of the ratio of Sharpe ratios of full model simulations. Values greater than 1 indicate generalized Black-Litterman outperformed Black-Litterman.

**Full Model**

Tests of the full model with both view bias and market equilibrium errors included were conducted with similar parameters as the previous tests. Generalized Black-Litterman was set up to have a prior of no view bias or equilibrium error but contained enough prior variance to allow for it to learn both tendencies if present. The ratio of generalized Black-Litterman’s Sharpe ratio to classical Black-Litterman’s Sharpe ratio was recorded for each series and can be found in Table 3.4. All values greater than 1 indicate that the Sharpe ratio for generalized Black-Litterman was higher than that of the classical Black-Litterman model.

The lowest value of these ratios appeared when there was no equilibrium bias. Within those simulations, Black-Litterman performed better when the view bias was more positive. As is to be expected, Black-Litterman performs better when more of its assumptions are satisfied and there are no equilibrium errors or view bias.

**3.4.2 Backtests**

For our backtests, we simulated a market consisting three assets, Apple (AAPL), Exxon-Mobil (XOM), and General Electric (GE). The 13-week treasury bill (‘IRX) was used as a risk-free asset. We obtained weekly closing prices from 1992-2012, and utilized the first ten years as “history” and simulated weekly performance of Black-Litterman as well as
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

generalized Black-Litterman over the last 120 weeks. Every week, CAPM was calculated over the previous 120 weeks of closing prices, thus the equilibrium estimate varied every period.

We took a naive approach of obtaining the sample covariance matrix to use in determining $\beta$ and $\Sigma$. Since both $\beta$ and $\Sigma$ are sources of variance in the sample covariance matrix, we simply estimated the contribution of $\beta$ to be $\frac{1}{5}$ of the total variance, and thus $\frac{1}{3}$ of the standard deviation, and took $\Sigma$ as to be the remainder of the sample covariance matrix.

We performed one backtest with no views and a second with the views incorporated. Due to fact that expert views were simulated, backtests with views alone were not performed.

Backtest with No Views

In performing a backtest with no expert views, we are testing to see whether generalized Black-Litterman picks up on any errors in the CAPM estimate. As it was with the controlled tests, we started generalized Black-Litterman off with prior means of no errors, but allowed enough variance to pick up on bias should it exist. Figures 3.6 and 3.7 are plots of the interquartile range of the Gibbs samples for $b^\mu$, the correction term for CAPM. As time goes on, it is clear that CAPM does not become a very reliable estimator of the expected return.

In the case of Apple, generalized Black-Litterman detected that the expected return appears to increase over time. This is consistent with the company’s history of success that started around 2002. A CAPM estimate that takes the previous 10 years into account will hardly be prepared to give a reliable estimate when the next ten years turn out to be astronomically profitable. On the other hand, GE had been a historically reliable asset which has fallen into tougher times as of late and Figure 3.7 certainly demonstrates this. Once again, a CAPM estimate that is based upon the previous 10 years of history has its shortcomings. However, these are shortcomings generalized Black-Litterman was able to detect and account for.

In Figure 3.8, we pit generalized Black-Litterman against the classical Black-Litterman model and recorded their portfolio wealth processes, allowing each portfolio to re-balance each week.

Full Model Backtest

As was done in the controlled tests, we generated biased views according to the model assumptions. Specifically, we generated pessimistic bias where mean biases were $\delta^V = -0.6e$ and $\Theta^V = 0.08^2 I$. Thus, view biases were clearly negative, ranging from -0.72 to -0.44. Our extension model was able to pick out the errors in the CAPM estimate and produced figures similar to Figures 3.6 and 3.7 in the previous section.

The negative bias of expert views were easily detected by generalized Black-Litterman as demonstrated by the interquartile range plot of the Gibbs sampled expert biases for Apple, Inc. in Figure 3.9. Worth noting is that the model was able to detect this behavior despite noisy views and errors in the market equilibrium estimate.
CHAPTER 3. A BAYESIAN GRAPHICAL GENERALIZATION FOR THE BLACK-LITTERMAN MODEL

Figure 3.6: Gibbs samples of $b^\mu$ for AAPL, 2002-2012

Figure 3.7: Gibbs samples of $b^\mu$ for GE, 2002-2012
Figure 3.8: Portfolio wealths of our extension and Black-Litterman in the 2002-2012 backtest with no expert views

The addition of expert views, despite their bias, aided both models tremendously in their portfolio allocation. In Figure 3.8 where there were no views, both models were susceptible to the flaws of CAPM due to the events of 2008. Figure 3.10, however, tells a different story and shows, generally, a positive earnings trend for both models. Rather than a drop in performance, both models enjoyed an appreciable increase in portfolio value in 2008. The contribution of expert views was quite significant in this case and the ability to determine view bias was what separated our extension from the classical Black-Litterman model.
Figure 3.9: Interquartile ranges of the Gibbs sampled expert view biases for AAPL in the 2002-2012 backtest with biased expert views.
Figure 3.10: Portfolio wealth processes for Black-Litterman and our extension model in the 2002-2012 backtest with biased expert views.
Chapter 4

Conclusion

In this dissertation, we developed two dynamic Bayesian portfolio allocation models that address questions of learning and model uncertainty with respect to the context of each problem.

In Chapter 2, we formulated the problem of portfolio selection with parameter uncertainty in the framework of Bayesian relative regret. For this problem, the solution involved a “tilted” posterior, where the tilting is done by way of a likelihood ratio which depends on the family of benchmarks and favors the model which has performed the best on the history of realized returns. We solved the case where the covariance of the returns is known and showed that the solution can be reduced to the Bayesian solution under special circumstances. Numerical results are promising, as the Bayesian regret model outperformed a benchmark given full information to the parameters of the problem. In all, we have shown that Bayesian regret utilizes learning to perform well relative to relevant benchmarks to the point that it can even outperform the benchmark that invests according to the correct market parameters.

In Chapter 3, we generalized the Black-Litterman model to take into account errors in the market equilibrium estimate and the possibility for expert view bias. This method results in a more robust portfolio allocation due to its ability to correct for errors in the model inputs. The complexity of this model resulted in the utilization of Gibbs sampling as the solution method. Due to the vast improvements in computational power, this method allows for solutions to be calculated in a reasonable amount of time. Our numerical tests indicate that the generalized Black-Litterman model performs well in comparison to its classical counterpart under a variety of conditions. Notably, generalized Black-Litterman was able to detect and account for monthly discrepancies between the CAPM estimate and the actual returns for historical data. Furthermore, our model was able to learn the behavior of biased experts and balance the portfolio accordingly.

With these models, we found model-specific solutions to possible shortcomings in the learning process of these dynamic Bayesian portfolio allocation models. In doing so, we discovered some interesting side effects of these adjustments, such as the tilted posterior in Chapter 2. Furthermore, ideas such as Bayesian regret and graphical model extensions are not specific to finance and can be utilized in other dynamic Bayesian contexts as well.
Bibliography


