Bayesian Optimization and Monte Carlo Simulation in Technology Sector Portfolio Allocation: A Comparative Analysis Using the Sharpe Ratio

Xinyu Hu *

Whiting School of Engineering, Johns Hopkins University, Maryland, The United States

* Corresponding Author Email: xhu18@u.rochester.edu

Abstract. In the dynamic landscape of financial markets, the pivotal role of adept portfolio construction is more prominent than ever. By enabling investors to form combinations of assets, it not only aids in accomplishing specified investment objectives but also in diversifying potential risks. This process employs a systematic investment approach, fostering informed and judicious decision-making strategies in the financial domain. This research seeks to explore the application of Bayesian statistics in parameter estimation and portfolio optimization, particularly focusing on the volatility of the technology sector. Utilizing a conjugate prior to obtaining the posterior and employing Maximum Likelihood Estimation (MLE) to calculate the expected return and covariance matrix, the study constructed portfolio weights based on Mean-Variance Optimization. This portfolio was then compared with weights derived from historical data analysis and a Monte Carlo simulation optimized by the Sharpe ratio. Remarkably, upon empirical testing on an out-of-sample dataset, the Bayesian approach exhibited a superior performance, showcasing robust resilience to potential risks. The study also acknowledges several limitations, suggesting directions for future research to refine the precision and reliability of portfolio optimization in the face of market volatility.

Keywords: Bayesian statistics; portfolio optimization; mean-variance optimization; conjugate prior; empirical data analysis.

1. Introduction

In the constantly evolving financial markets, the essential role of strategic portfolio construction cannot be understated. Portfolio construction serves as a critical tool in wealth management and risk mitigation. It allows investors to strategically combine assets to achieve specific investment goals while diversifying risks. It facilitates a structured investment approach, encouraging informed and rational decision-making. This field of the financial industry has consistently remained a subject of keen interest. The earliest study traces back to the mean-variance optimization model by Markowitz [1]. It effectively optimizes portfolio selection by balancing expected returns against potential volatility, offering a systematic approach to achieving investment objectives. However, the model utilizes statistical measures, mainly expected return and volatility, as known parameters to identify the optimal asset allocation in a portfolio. The accuracy in estimating the mean and variance is then the key to achieving the optimality offered by the mean-variance model. Precise estimates serve as the backbone for constructing the most optimal portfolios that genuinely reflect the risk-return tradeoff investors are willing to embrace, and inaccurate estimation will lead to suboptimality. In essence, the reliability and effectiveness of the mean-variance model lie significantly on the accuracy of mean and variance estimations, underlining the necessity for rigorous analysis in the initial stages of portfolio construction.

Traditionally, these parameters have been estimated using historical data, a method that, although prevalent, contains the potential for estimation errors and subsequent suboptimal portfolio choices. Becker and Xiao endeavored to accurately estimate the required parameters [2]. Their study utilized machine learning techniques to predict equity risk premiums, substituting these forecasts for the first sample moment in portfolio construction. Keller et al. examined the effectiveness of integrating mean-variance optimization with momentum estimations and long-only constraints, demonstrating its superior performance over the 1/N portfolio [3]. Lai et al. attempted to employ Bayes and shrinkage estimators to ascertain the mean and variance of stock returns [4]. They introduced supplementary hyperparameters to address the considerably more complex stochastic optimization challenge brought about by this parameter uncertainty. Among all the approaches, the Bayesian approach is especially attractive. It allows the integration of prior information, which reflects investors' perspectives, within the estimation procedure. This estimation transcends mere historical data analysis by granting analysts to infuse personal insights. Furthermore, it recognizes the inherent risks and uncertainties of parameters within the modeling process by offering a posterior distribution instead of a simple point estimate based on historical data. Additionally, the utilization of fast and reliable numerical algorithms facilitates a simplified implementation process, adept at accurately mirroring complex economic variables. As such, the employment of Bayesian statistics presents a potentially robust method for parameter estimation.

The inception of utilizing Bayesian methods in portfolio decision-making can be traced back to Winkler and Barry's implementation of uninformative prior [5]. Several years later, Klein and Bawa examined the impact of estimation risk on optimal portfolio selections in uncertain situations, particularly when employing non-informative priors [6]. Zelnner and Chetty pioneered in their use of Bayesian predictive distribution within the portfolio decision-making framework while still based on uninformative prior [7]. Jorion then introduced the integration of hyperparameter priors grounded in Bayes-Stein shrinkage [8]. Pástor advocated for the inclusion of asset pricing theories as a source of prior information [9]. Bodnar et al. focused on shrinkage estimation to incorporate the beliefs of investors [10]. Kacperczyk et al. discussed the effectiveness of Bayesian semi-parametric models [11]. Avramov and Zhou provided a throughout literature review of different priors with predictive distribution, including diffused or uninformative prior, conjugate prior, hyperparameter prior, objective prior, etc. [12]. Bauder et al. illustrated the applications of diffuse and conjugate priors in determining the allocation of Markowitz's tangent portfolio [13].

On top of that, as the computational power surges, an increasing number of algorithms have been incorporated into the portfolio construction process. Among the most significant is the Monte Carlo Simulation. This method bypasses the constraints of the mean-variance model and parameter uncertainty. It directly generates weights on a substantial scale, subsequently selecting the highest-performing weights from all the samples. Its simplicity of implementation and absence of any prerequisite conditions or assumptions renders it an excellent alternative for comparison with Bayesian mean-variance estimation.

This research aims to conduct a comparative analysis of the application of Bayesian statistics in parameter estimation and portfolio optimization. It seeks to draw comparisons with the traditional estimated mean-variance model and Monte Carlo simulation methods utilizing empirical data. Given the pronounced volatility in the technology sector, it is a great source of data to examine the robustness of these three methodologies. The outcomes will subsequently be analyzed using the Sharpe ratio as the benchmark.

2. Methodology

2.1. Data Source

2.1.1 Data and visualization

This study engages with 5 prominent stocks from the tech sector to compose the portfolio, namely NVDA (NVIDIA Corporation), MSFT (Microsoft Corporation), AMD (Advanced Micro Devices, Inc.), AAPL (Apple Inc.), and AMZN (Amazon.com, Inc.). The adjusted closing prices of these stocks have been sourced from Yahoo Finance, segregated distinctly for the purposes of parameter estimation and portfolio formulation. The data used for estimation includes three years of data from 1/1/2019 to 12/31/2021. Following this, the portfolio construction or the test set utilized a full year's data, ranging from 1/1/2022 to 12/31/2022.

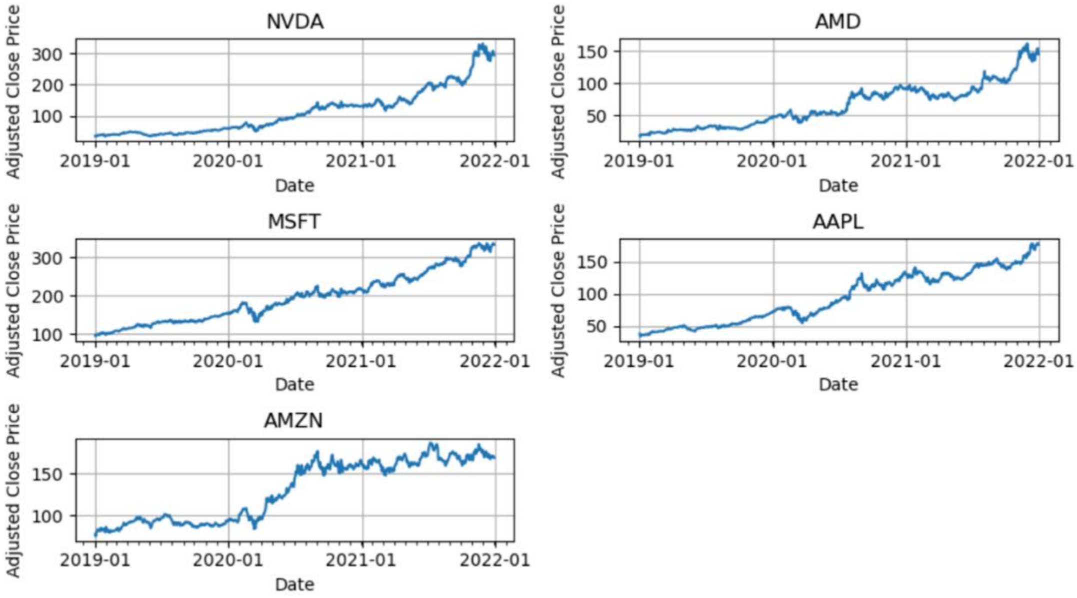

Fig. 1 Visualiazation of Data

Figure 1 is a visualization of the three-year data for five stocks selected for the portfolio. These stocks exhibit considerable volatility and seemingly align in a similar trend pattern, implying a positive correlation among them. Consequently, the constructed portfolio might have a considerable degree of risk. This scenario serves as an excellent ground for evaluating the robustness of the various models in focus.

2.1.2 Data preprocessing

Before the Bayesian statistics approach was implemented, the data was preprocessed to ensure more dependable results. Illustrating NVDA as an example, the initial step involved converting the data to percentage changes to obtain the stock's return.

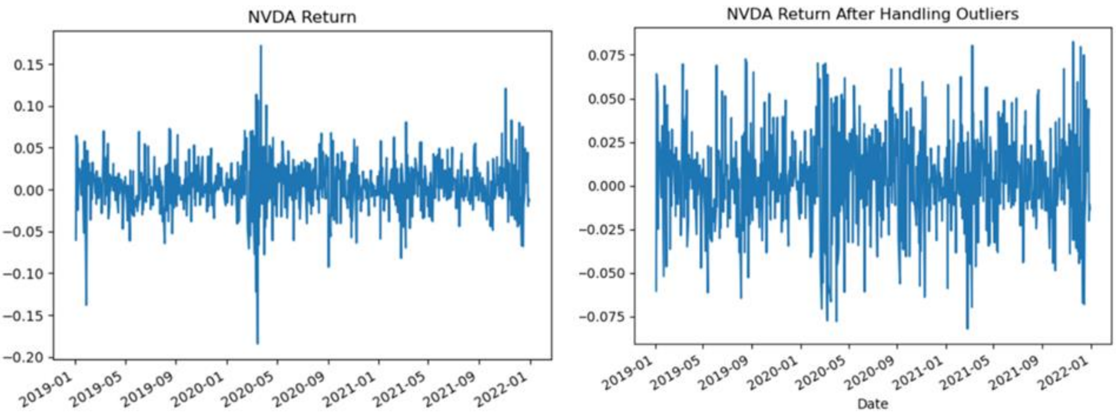

Fig. 2 Return of NVDA before and after processing

The left part of figure 2 displays a graphical representation of NVDA's returns. While the data fluctuated around zero, distinct outliers were observed, especially around May 2020. To enhance the robustness of the likelihood function, these outliers were removed, and the data was tested for stationarity. In accordance with the prevalent rule of thumb, any values boasting a z-score greater than 3 or less than -3 were eliminated. The refined data is depicted in the right part of figure 2. The Augmented Dickey-Fuller test was then employed to verify the stationarity of the data. The ADF yielded a test statistic of -29.556978. This significantly negative figure indicated strong evidence against the null hypothesis of the ADF test, thereby suggesting that the data is stationary. The remaining four stocks were processed in a similar manner, and the ADF test confirmed that all of them exhibited stationarity. To determine the potential priors of the processed data, histograms were generated.

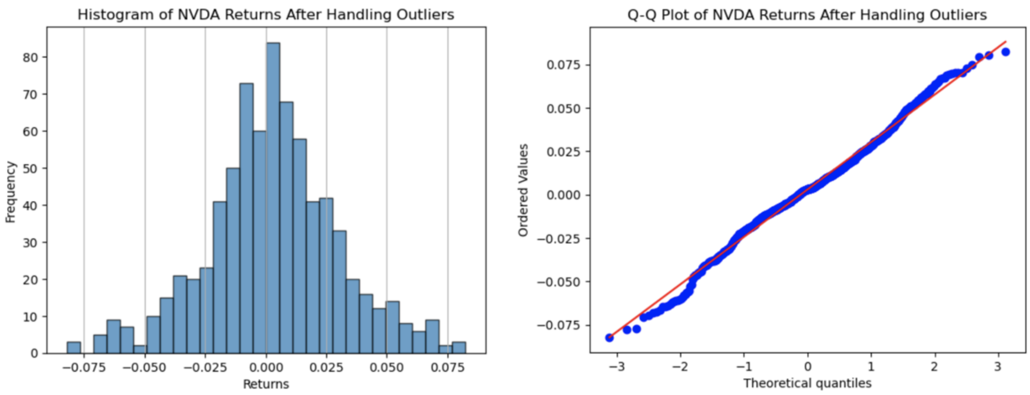

Fig. 3 Histogram and Q-Q plot of NVDA Returns

The histogram in figure 3 exhibits characteristics of a normal distribution, and the Q-Q plot further consolidated it. The points on the Q-Q plot largely align with the straight line, with points slightly below in the beginning and slightly above at the end. This indicated that the distribution has thinner tails at the lower quantiles and fatter tails at the upper quantiles, implying a mild positive skewness of the data.

2.2. Modeling

2.2.1 Mean-variance optimization

Serving as the foundation of this study, it is advantageous to review of the mean-variance optimization theory developed by Markowitz. It aims to choose weights of the portfolio that maximize the quadratic utility function.

$$U(w) = E[R_p] - \frac{\gamma}{2} Var[R_p] \tag{1}$$

where w denoted the weights of the portfolio, and R_p denoted the portfolio return. γ is the risk aversion coefficient. Considering k stocks in the portfolio, by the Lagrangian multiplier method, the optimal weight can be easily deducted.

$$w^* = \frac{1}{\gamma}\Sigma^{-1}\mu \tag{2}$$

where μ is a k-demensional vector of stock returns, and Σ is the covariance matrix.

In the context of traditional mean-variance optimization, μ and Σ are replaced by point estimations derived from historical data.

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t \tag{3}$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})^T \tag{4}$$

2.2.2 Bayesian statistics

In their literature review, Avramov and Zhou determined that utilizing a Bayesian approach centered on the diffuse or the data-based uninformative priors does not lead to significantly different portfolio decisions when compared to the traditional framework. Therefore, this study focused on evaluating the efficacy of the conjugate prior within this setting. In this context, mu was assigned a k-dimensional multivariate normal prior conditioned on Σ, while Σ was assigned with an inverted Wishart prior. The conjugate priors are listed as follows.

$$\mu|\Sigma \sim N_k(\mu_0, \frac{1}{\kappa_0}\Sigma) \tag{5}$$

$$\Sigma \sim IW(v_0, \Phi_0) \tag{6}$$

Their joint distribution is thus given by Normal-Inversed Wishart$(\mu_0, \frac{1}{\kappa_0}\Phi_0, v_0, \Phi_0)$.

$$p(\mu,\Sigma) \propto |\Sigma|^{-\left(\frac{vo+k}{2}+1\right)} \exp\{-\frac{1}{2}tr(\Phi_0\Sigma^{-1}) - \frac{\kappa_0}{2}(\mu-\mu_0)^T\Sigma^{-1}(\mu-\mu_0)\} \tag{7}$$

Based on the histogram and Q-Q plot analysis, it is reasonable to assume that the data follows a normal distribution.

$$y_1, \dots, y_n \sim N_k(\mu,\Sigma) \tag{8}$$

The likelihood function is given by:

$$L(Y|\mu,\Sigma) = L(y_1,\dots,y_n|\mu,\Sigma) = \prod_{i=1}^{n}(2\pi)^{-\frac{k}{2}}|\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y_i-\mu)^T\Sigma^{-1}(y_i-\mu)\right\} \tag{9}$$

$$\propto |\Sigma|^{-\frac{n}{2}} \exp\{-\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^T\Sigma^{-1}(y_i-\mu)\}, \tag{10}$$

$$\propto |\Sigma|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}[n\mu^T\Sigma^{-1}\mu - 2\mu^T\Sigma^{-1}\sum_{i=1}^{n}y_i]\right\}, \tag{11}$$

$$\propto |\Sigma|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\mu^T A_1\mu + \mu^T b_1\right\}, \tag{12}$$

Where $A_1 = n\Sigma^{-1}$ and $b_1 = n\Sigma^{-1}\bar{y}$.

By the Bayes theorem, the posterior distribution of mu can be expressed as

$$p(\mu|\Sigma,Y) \propto L(Y|\mu,\Sigma)p(\mu|\Sigma) \tag{13}$$

Where prior of mu is given by

$$p(\mu) = (2\pi)^{-\frac{k}{2}}\left|\frac{1}{\kappa_0}\Sigma\right|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\mu-\mu_0)^T(\frac{1}{\kappa_0}\Sigma)^{-1}(\mu-\mu_0)\} \tag{14}$$

Let $S_0 = \frac{1}{\kappa_0}\Sigma$, then, $p(\mu) \propto |S_0|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\left(\mu^T S_0{}^{-1}\mu + \mu_0^T S_0{}^{-1}\mu_0 - 2\mu^T S_0{}^{-1}\mu_0\right)\}$. Since $\mu_0^T S_0{}^{-1}\mu_0$ is a constant, $p(\mu)$ can be rewrite as:

$$\propto |S_0|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\left(\mu^T S_0{}^{-1}\mu - 2\mu^T S_0{}^{-1}\mu_0\right)\} \tag{15}$$

$$\propto |S_0|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\mu^T A_0\mu + \mu^T b_0\} \tag{16}$$

where $A_0 = S_0{}^{-1}$ and $b_0 = S_0{}^{-1}\mu_0$.

Equation (13) can be further developed as

$$p(\mu|\Sigma,Y) \propto |\Sigma|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\mu^T A_1\mu + \mu^T b_1\right\} * |S_0|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\mu^T A_0\mu + \mu^T b_0\} \tag{17}$$

$$\propto |\Sigma|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\mu^T A_1\mu + \mu^T b_1\right\} * |S_0|^{-\frac{1}{2}} \exp\{-\frac{1}{2}\mu^T A_0\mu + \mu^T b_0\} \tag{18}$$

Where

$$A_n = A_0 + A_1 = n\Sigma^{-1} + S_0{}^{-1} \tag{19}$$

$$b_n = b_0 + b_1 = n\Sigma^{-1}\bar{y} + S_0{}^{-1}\mu_0 \tag{20}$$

The distribution of μ conditioned on the data and Σ is therefore.

$$\mu|\Sigma, \mathrm{Y} \sim MVN(\mu_n, \Sigma_n) \tag{21}$$

Where

$$\mu_n = \Sigma_n(n\Sigma^{-1}\bar{y} + S_0{}^{-1}\mu_0) \tag{22}$$

$$\Sigma_n = (n\Sigma^{-1} + S_0{}^{-1})^{-1} \tag{23}$$

It should be noted that the posterior mean can be articulated as a sum of the weighted prior mean and the sample mean.

$$\mu_n = \frac{\kappa_0}{n+\kappa_0}\mu_0 + \frac{n}{n+\kappa_0}\bar{y} \tag{24}$$

For the deduction of posterior of Σ, the likelihood function is firstly reorganized.

$$L(\Sigma|Y) = \prod_{i=1}^{n}(2\pi)^{-\frac{k}{2}}|\Sigma|^{-\frac{1}{2}}\exp\{-\frac{1}{2}(y_i-\mu)^T\Sigma^{-1}(y_i-\mu)\} \tag{25}$$

$$\propto |\Sigma|^{-\frac{1}{2}}\exp\{-\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^T\Sigma^{-1}(y_i-\mu)\} \tag{26}$$

It can be rewrite as

$$\log\big(L(\Sigma|Y)\big) \propto -\frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^T\Sigma^{-1}(y_i-\mu)\} \tag{27}$$

Where

$$-\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu)^T\Sigma^{-1}(y_i-\mu)\} = -\frac{1}{2}tr\ (S\Sigma^{-1}). \tag{28}$$

Let $S = \sum_i (y_i-\bar{y})(y_i-\bar{y})^T$, then.

$$L(\Sigma|Y) \propto |\Sigma|^{-\frac{n}{2}}\exp\{-\frac{1}{2}tr\ (S\Sigma^{-1})\} \tag{29}$$

By the Bayes theorem, the posterior distribution of Σ can be expressed as

$$p(\Sigma|Y) \propto L(\Sigma|Y) * p(\Sigma) \tag{30}$$

$$\propto |\Sigma|^{-\frac{n}{2}}\exp\left\{-\frac{1}{2}tr(S\Sigma^{-1})\right\} * |\Sigma|^{-\frac{v_0+k+1}{2}}\exp\{-\frac{1}{2}tr(\Phi_0\Sigma^{-1})\} \tag{31}$$

$$\propto |\Sigma|^{-\frac{v_0+n+k+1}{2}}\exp\{-\frac{1}{2}tr((S+\Phi_0)\Sigma^{-1})\} \tag{32}$$

To incorporate the prior of mu, there is an additional term to consider.

$$\Phi_n = \Phi_0 + S + \frac{\kappa_0 n}{\kappa_0+n}(\bar{y}-\mu_0)(\bar{y}-\mu_0)^T \tag{33}$$

Let $\kappa_n = \kappa_0 + n$, the posterior of Σ is therefore.

$$\Sigma|Y \sim IW(v_n, \Phi_n) \tag{34}$$

The joint distribution of the parameters is consequently represented by Normal-Inversed Wishart $(\mu_n, \frac{1}{\kappa_n}\Phi_n, v_n, \Phi_n)$.

$$p(\mu,\Sigma) \propto |\Sigma|^{-\left(\frac{v_n+k}{2}+1\right)}\exp\{-\frac{1}{2}tr(\Phi_n\Sigma^{-1}) - \frac{\kappa_n}{2}(\mu-\mu_n)^T\Sigma^{-1}(\mu-\mu_0)\} \tag{35}$$

The initial estimates for the mean μ_0 and variance σ_0 were derived from historical data, aiming to incorporate all the information from the past. The likelihood, or the $\bar{y}$, was based on the refined data to ensure a robust resistance to outliers. The subsequent hyperparameters were then initialized with arbitrary values. n = number of data points. k = 5 which corresponds to the five stocks included in the portfolio. κ_0 = 0.01 where a small value was selected for κ_0 to indicate that the prior is relatively weak, implying that the analysis should be substantially influenced by the data rather than historical outliers. v_0 = k+2 since the degrees of freedom v_0 must be at least equivalent to k to maintain statistical validity. In this case, v_0 was chosen to be slightly larger than k to instill a moderate level of informativeness into the prior.

2.2.3 Monte carlo simulation

Monte Carlo simulation managed to circumvent the challenge of prior estimation by directly generating weights. To align with the Bayesian, mean-variance approach, short selling was allowed in this framework. 50,000 sets of random values ranging from -1 to 1 were generated and subsequently normalized to ensure that they have a sum of 1. Based on the three-year historical data, the portfolio mean and volatility were computed for each distinct weight set. To select the optimal portfolio, the modified Sharpe Ratio was employed as the benchmark criterion, guiding the selection process toward a portfolio that promises superior risk-adjusted returns.

$$Modified\ Sharp\ Ratio = \frac{E[R_p]}{\sigma_p} \tag{36}$$

3. Results and Discussion

3.1. The Result Analysis

3.1.1 Bayesian statistics

A total of 10,000 samples from the posterior distribution were generated utilizing the Gibbs sampler, incorporating a burn-in period of 2,000 samples to stabilize the sampling process. Following this, a convergence diagnosis was undertaken on the samples after the burn-in period to confirm the adequacy and stability of the sampled data set.

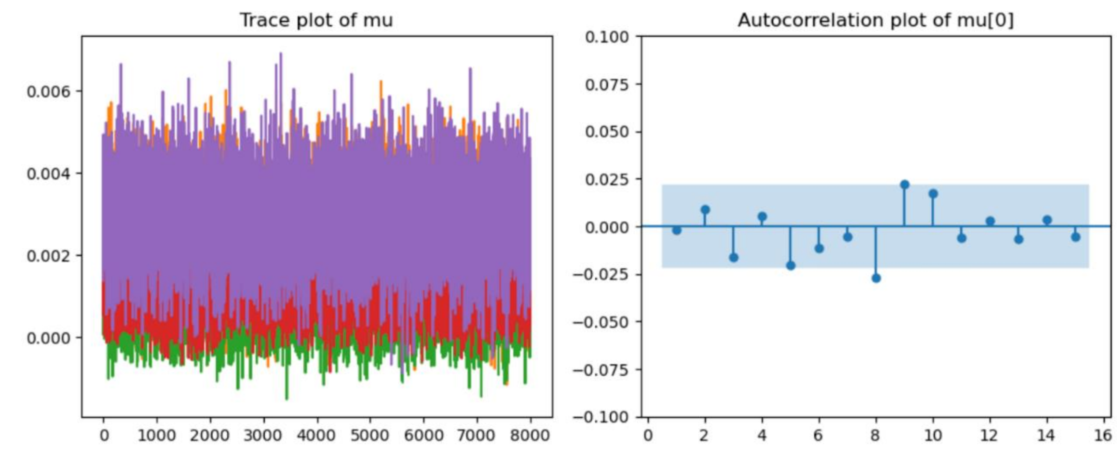

Fig. 4 Traceplot and ACF of μ

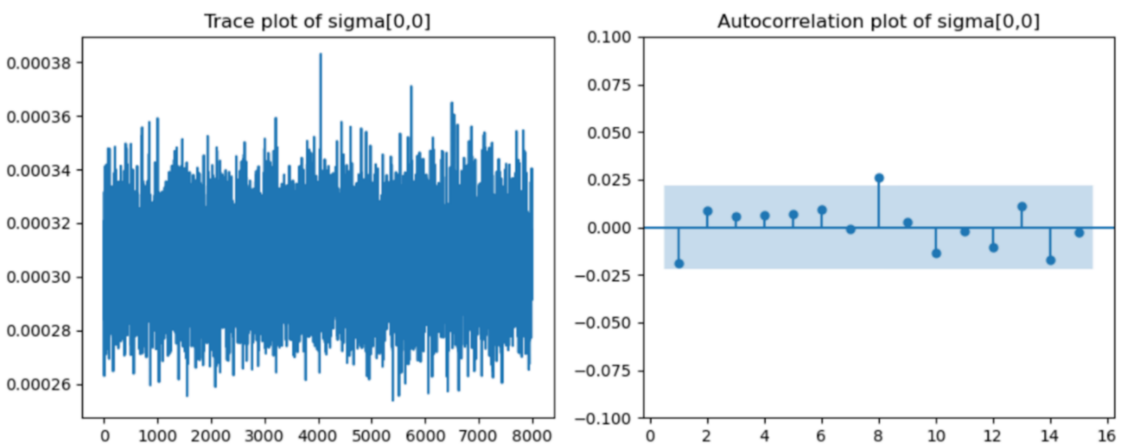

Fig. 5 Traceplot and ACF of Σ

The above figures 4 and 5 are the traceplot and ACF graph for one of the values in mu and Σ. The trace plot doesn't show any discernible trends, suggesting a random walk pattern which is a positive sign of convergence. Additionally, the ACF graph illustrates a rapid decline in autocorrelations, indicating a high degree of independence between successive samples. Collectively, these graphs have shown that the sampling process has successfully achieved convergence. Subsequently, the values of mu and Σ were estimated by the Maximum Likelihood Estimation (MLE) approach, based on these generated samples from the posterior distribution.

3.1.2 Monte carlo simulation results

Utilizing the approach detailed in section 2.2.3, one of the generation processes was visualized, and the respective optimal portfolio weights were documented.

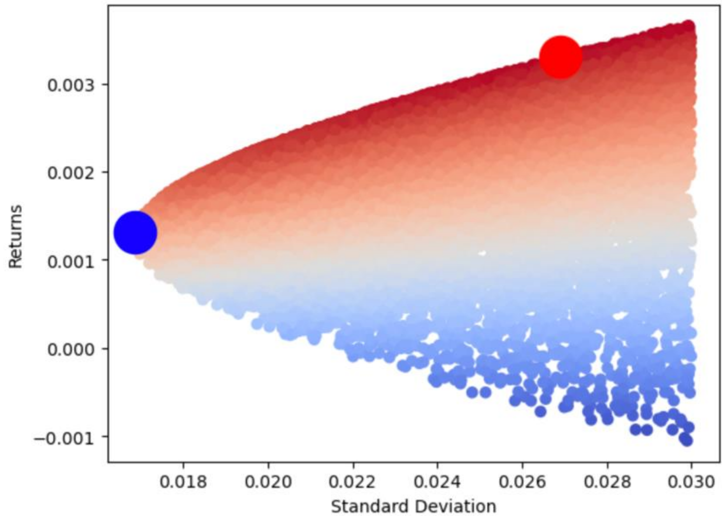

Fig. 6 Efficient Frontier

Figure 6 offers a graphical representation of the 50,000 randomized portfolios. The upper segment of the depicted parabola aligns with the efficient frontier as delineated in Markowitz's modern portfolio theory. The point illustrated in blue represents the minimum variance portfolio. The red point is a marker of the portfolio that has the highest modified sharp ratio among all the samples, representing an optimum balance of risk and return. The red point sample is the targeted weight under this method.

3.1.3 The model results

Based on the details provided above, both the prior and the posterior means and variance were collected and structured into tables a comprehensive comparison (Table 1 and 2).

Table 1. Prior and Posterior mean

	AAPL	AMD	AMZN	MSFT	NVDA
prior	0.002269	0.003259	0.001211	0.001832	0.003341
posterior	0.002243	0.002562	0.000817	0.001176	0.002940

Table 2. Prior and Posterior Covariance matrix

	AAPL	AMD	AMZN	MSFT	NVDA
prior	0.000463	0.000406 0.001117	0.000261 0.000339 0.000343	0.000318 0.000374 0.000247	0.000437 0.000730 0.000345
				0.000366	0.000417
					0.000928
posterior	0.000305	0.000239	0.000161	0.000172	0.000273
		0.000811	0.000216	0.000187	0.000496
			0.000278	0.000161	0.000271
				0.000247	0.000271
					0.000744

Despite being derived from the same dataset, the prior and posterior yield very different values. Assuming an ambiguous risk aversion coefficient of 1, the posterior estimations were then plugged into equation (2) to obtain the optimal weight within the framework of mean-variance optimization.

Table 3. Weights generated by Three Method

	AAPL	AMD	AMZN	MSFT	NVDA
Historical	0.66753	0.23216	-0.47905	0.15226	0.42710
Bayesian	1.26015	0.12441	-0.62618	-0.23650	0.47812
Monte Carlo	0.21170	0.21211	0.41760	-0.43766	0.59626

Table 4. Sharp Ratios generated by Three Method

	SR of training set	Return of training set	Volatility of training set	SR of testing set	Return of testing set	Volatility of testing set
Historical	0.12257	0.00340	0.02772	-0.04054	-0.00126	0.03118
Bayesian	0.11869	0.00369	0.03106	-0.02963	-0.00097	0.03278
Monte Carlo	0.12249	0.00330	0.02692	-0.04175	-0.00127	0.03051

The weights deduced from the three models vary significantly. The Bayesian approach appears to adopt a more aggressive strategy compared to the others, evidenced by more short-selling positions, which theoretically indicates a portfolio with higher risk. When evaluating the training set, the dataset used for estimation, the Bayesian method displayed the lowest modified sharp ratio, while the other two methods presented nearly identical values. As projected, the portfolio constructed via the Bayesian method demonstrated higher risk in comparison to the other strategies (Table 3 and 4).

Remarkably, the outcomes from the test set are out of expectations. The portfolio constructed through the Bayesian method exhibited the highest modified sharp ratio and the least amount of loss within the test set. It is worth noting that all three portfolios, constructed using distinct methods, endured losses in the test phase. This could potentially be attributed to the pronounced volatility and unpredictability inherent in the technology sector, coupled with the non-diversified risks embedded within the portfolio structure. Under these circumstances, the Bayesian portfolio demonstrated superior resilience to such risks, recording the minimum loss. A comparison between the historical and the Monte Carlo methods revealed fairly similar outcomes, albeit the Monte Carlo approach seemed to offer the lowest risk resistance.

3.2. Discussion and Further Improvement

This study comes with several limitations. Firstly, the portfolios were not restricted by a no-short-selling constraint, potentially leading to compliance issues and additional fees. Given the inherent high risks associated with the technology sector, it would be both reasonable and practical to explore portfolios that prohibit short selling. Secondly, the prior and the likelihood were calculated using data spanning three years. As suggested by Keller, Butler, and Kipnis in their research, utilizing shorter lookback periods, up to a maximum of 12 months, could be more effective than the extended horizons employed for estimation in this study. Future improvements could involve a more thoughtful selection of the estimation horizon to enhance the efficacy of the estimation process. Thirdly, the empirical study highlighted the pronounced volatility inherent in the technology sector. Therefore, in addition to pursuing maximum utility, further studies within the technology sector should accord a higher priority to risk management and mitigation. Moreover, this research focused exclusively on the conjugate prior, while various other priors have demonstrated effectiveness within this framework, as highlighted by Avramov and Zhou. Incorporating more contemporary priors could be a significant enhancement to this study. Last, but certainly not least, the study limited its estimation to the Maximum Likelihood Estimation (MLE) on the posterior distribution of parameters. A notable advantage of the Bayesian method is its ability to yield a parameter distribution, rather than a singular point estimation. Subsequent research should aim to extract more information from the posterior distribution, perhaps by employing confidence intervals to determine a range of potential optimal weights. This approach might pave the way for integrating fundamental analysis in the selection of the final optimal weights, offering a richer, more nuanced perspective in portfolio optimization.

4. Conclusion

In summary, this research attempted on a journey to evaluate the efficacy of employing Bayesian statistics in the estimation of expected return and covariance matrix, and its subsequent application in mean-variance optimization for portfolio analysis. Utilizing a conjugate prior to deriving the posterior and Maximum Likelihood Estimation for obtaining expected return and covariance, the study formulated portfolio weights in line with the Bayesian approach. In addition, it drew comparisons by establishing optimal weights based exclusively on historical data and through the use of Monte Carlo simulation, which generated sets of weights and selected according to the Sharpe ratio.

A critical aspect of this research was the empirical testing of these distinct sets of weights on an out-of-sample dataset, a procedure that yielded noteworthy results. Although the methods grounded in historical data and Monte Carlo simulation delivered comparable performances, the former exhibited a slight edge. The Bayesian methodology appeared superior, demonstrating remarkable resilience to risks and outperforming the other techniques in a volatile technical sector setting. It is easy to implement and serves as a viable alternative to relying exclusively on historical data. The utilization of the conjugate prior proves to be beneficial in this context.

Despite its promising findings, the study acknowledges its boundaries, including the absence of a no-short-selling constraint and the potential enhancement of estimation efficiency through shorter lookback periods. Moreover, it recognizes the opportunity to broaden the scope by considering a wider variety of priors, and the potential to delve deeper into the rich information encapsulated in the posterior distribution, beyond a singular MLE estimation.

Moving forward, it is prudent to address these limitations, perhaps by incorporating portfolios without short selling, contemplating shorter estimation horizons, and expanding the utilization of contemporary priors. Moreover, future research should attempt to utilize the full information captured in the posterior distribution, fostering a more comprehensive analysis that might integrate fundamental analytical components in the determination of final optimal weights.

References

[1] Markowitz H. Mean-variance analysis in portfolio choice and capital markets. Journal of Finance, 1952, 7: 77–91.

[2] Becker H, et al. Improving Mean Variance Portfolio Construction With Machine Learning Stock Price Forecasts: A different approach to forecasting the sample mean for mean variance optimization. Erasmus School of Economics Bachelor Thesis Econometrics and Operations Research, 2020.

[3] Keller W, Butler A and Kipnis I. Momentum and Markowitz: A Golden Combination. Social Science Research Network, 2015.

[4] Lai T, Xing H, Chen Z. Mean–variance portfolio optimization when means and covariances are unknown. The Annals of Applied Statistics, Ann. Appl. Stat., 2021, 5(2): 798-823.

[5] Winkler R, Barry C. A Bayesian Model for Portfolio Selection and Revision. The Journal of Finance, 1975, 30(1): 179–92.

[6] Klein R, Bawa V. The effect of estimation risk on optimal portfolio choice. Journal of Financial Economics, 1976, 3(3): 215-231.

[7] Zellner A, Chetty K. Prediction and Decision Problems in Regression Models from the Bayesian Point of View, Journal of the American Statistical Association, 1965, 60: 608-616.

[8] Jorion P. Bayes-Stein Estimation for Portfolio Analysis. Journal of Financial and Quantitative Analysis, 1986, 21(3): 279-292.

[9] Pástor L. Portfolio selection and asset pricing models. Journal of Finane, 2000, 55(1): 179-223.

[10] Bodnar, el al. Estimation of the global minimum variance portfolio in high dimensions. European Journal of Operational Research, 2018, 226(1): 371-390.

[11] Kacperczyk M, Damien P, Walker S. A new class of Bayesian semi-parametric models with applications to option pricing, Quantitative Finance, 2013, 13(6): 967-980.

[12] Avramov D, Zhou, G. Bayesian Portfolio Analysis. Annual Review of Financial Economics, 2010, 2: 25-47.

[13] Bauder D, et al. Bayesian inference for the tangent portfolio. International Journal of Theoretical and Applied Finance, 2018, 21(08).