Sains Malaysiana 54(2)(2025): 601-609
http://doi.org/10.17576/jsm-2025-5402-24

l_0 Norm Sparse Portfolio Optimisation using Proximal Spectral Gradient Method on Malaysian Stock Market

(Pengoptimuman Portfolio Jarang l_0 Norma menggunakan Kaedah Kecerunan Spektrum Proksimal dalam Pasaran Saham Malaysia)

KEVIN CHOON LIANG YEW[1], WAI KUAN WONG[1,*], HONG SENG SIM[1], YONG KHENG GOH[1], WEI YEING PAN[1] & SHIN ZHU SIM[2]

[1]*Centre for Mathematical Sciences, Universiti Tunku Abdul Rahman, Bandar Sungai Long, 43000 Kajang, Selangor, Malaysia*
[2]*School of Mathematical Sciences, University of Nottingham Malaysia, Jalan Broga, 43500 Semenyih, Selangor, Malaysia*

Received: 2 January 2024/Accepted: 22 November 2024

ABSTRACT

In this paper, we introduce a modified norm-constraint mean-variance portfolio selection method. First, we use the Augmented Lagrangian method (ALM) to convert the objective function to an unconstrained objective function. Then, we apply the proximal spectral gradient method (PSG) onto the unconstrained objective function to find an optimal sparse portfolio. This novel sparse portfolio optimization procedure encourages sparsity in the entire portfolio using l_0 – norm. The PSG utilizes a multiple damping gradient (MDG) method to solve the smooth terms of the function. The step size is computed using the Lipschitz constant. Also, PSG uses the iterative thresholding method (ITH) to solve l_0 – norm and induce the sparsity of the portfolio. The performance of the PSG is illustrated by its application on the Malaysian stock market. It is found that PSG's sparse portfolio outperforms the equal weightage portfolio when the initial portfolio size is around 100 stocks and is prefiltered using the Sharpe ratio or the coefficient of variation.

Keywords: l_0 – norm; sparse portfolio optimization; proximal spectral gradient; Malaysia stock market

ABSTRAK

Dalam kertas ini, kami memperkenalkan kaedah pemilihan portfolio min-varians kekangan norma yang diubah suai. Pada awalnya, kami menggunakan Kaedah Augmented Lagrangian (ALM) untuk mengubah fungsi objektif kepada fungsi objektif tanpa sekatan. Seterusnya, kami memakai kaedah kecerunan proksimal spektrum (PSG) ke atas fungsi objektif tanpa sekatan tersebut untuk mendapat portfolio jarang optimum. Prosedur pengoptimuman portfolio jarang yang novel ini menggalakkan jarang bagi portfolio keseluruhan dengan menggunakan l_0 – norma. PSG menggunakan kaedah kecerunan redaman berbilang (MDG) untuk menyelesaikan sebutan licin pada fungsi tersebut. Saiz langkah dihitung dengan menggunakan pemalar Lipschitz. Tambahan pula, PSG menggunakan kaedah ambang berbilang (ITH) untuk menyelesaikan l_0 – norma dan menggalakkan jarang bagi portfolio. Prestasi PSG ini digambarkan dengan aplikasinya ke atas pasaran saham Malaysia. Didapati bahawa portfolio jarang yang dicadangkan mendahului prestasi portfolio berwajaran sama apabila saiz portfolio permulaan adalah kira-kira 100 saham dan telah ditapis dengan nisbah Sharpe atau pekali variasi.

Kata kunci: l_0 – norma; pengoptimuman portfolio jarang; Kecerunan proksimal spektrum; pasaran saham Malaysia;

INTRODUCTION

Since the introduction of Markowitz' Mean-variance (MV) model (Markowitz 1952), many models have been proposed, all of which are built upon the foundation laid by Markowitz's MV model. Nevertheless, the sample mean and the sample covariance matrix are susceptible to error, especially when dealing with large portfolio size (Chopra & Ziemba 1993; DeMiguel, Garlappi & Uppal 2011; Jagannathan & Ma 2003), it cannot consistently dominate the naïve $1/N$ diversification strategy (Hwang, Xu & In 2018).

Markowitz's MV model was modified in many ways to improve it (Chen, Dai & Zhang 2020; Gotoh & Takeda 2011; Kremer et al. 2020; Olivares-Nadal & DeMiguel 2018). Among the improvements on the MV model, sparse portfolio optimization has been a popular research topic since it reduces the difficulties in managing a handful of stocks and transaction costs. The sparse portfolio optimization is divided into two categories: long-term sparse portfolio optimization which restructures the portfolio with a time interval of a month or a year or even longer, and short-term sparse portfolio optimization which rebalances the portfolio daily (or even shorter) (Wang et al. 2023). In our study, we specifically concentrate on long-term sparse portfolio optimization, aiming to optimize portfolio composition over extended time intervals.

In sparse portfolio optimization, many researchers do not use l_0 −norm to sparsify the portfolio but tend to use l_1 −norm as a loose relaxation of l_0 −norm (Chen, Dai & Zhang 2020; Kremer et al. 2020) because the l_0 −norm is non-convex, non-differentiable and non-continuous, posing challenges for solving optimization problems. However, we have chosen to embrace l_0 −norm to construct our sparse portfolio as we believe it is a more natural and suitable method for portfolio selection. Sim et al. (2023) proposed a proximal linearized approach for optimizing sparse equity portfolios. The study effectively employed a gradient method combined with the l_0 −norm in the portfolio selection.

To tackle the optimization problem involving the l_0 −norm, we apply a proximal gradient method called the Proximal Spectral Gradient (PSG) method. It is a gradient method that integrates two essential techniques: the Multiple Damping Gradient (MDG) method (Sim, Leong & Chen 2019) and the iterative thresholding method (ITH). The MDG solves the smooth terms of the objective function while the ITH handles the l_0 −term and controls the number of non-zero stocks in selected portfolios.

In theory, the PSG method has the capability to address large-scale unconstrained optimization problems with minimal storage and less computational time (Woo et al. 2023). While PSG has been used in the field of machine learning, demonstrating robustness and stability in finding sparse solutions for linear systems, it has not yet been explored in the context of portfolio optimization. Therefore, in this study, we aim to extend its application to portfolio optimization and evaluate its efficiency in discovering optimal sparse portfolios. The contributions of this paper are summarized as follows:

1) The construction of a modified norm-constraint MV portfolio selection objective function and a portfolio optimization model that explicitly controls the sparsity of the portfolio. Additionally, the effectiveness of this approach in constructing sparse portfolios is demonstrated by using the Bursa Malaysia Stock Index (FBMKLCI and FBMT100)

2) The adoption of a two-step process, starting with stock filtering based on either the Sharpe ratio or the coefficient of variation, then, applying the PSG to obtain optimal sparse portfolios with varying numbers of active stocks. The performance of the proposed procedure is assessed in terms of average annual return, average Sortino ratio and average M^2 ratio.

The paper is organized in the following order: Methodology section that provides a brief outlines of Markowitz's MV model and our modifications. It discusses also the procedures applying the proximal gradient method which combines MDG and ITH to solve both smooth and non-smooth parts of the objective function. Additionally, the Augmented Lagrangian method (ALM) is introduced to address the constraints in this study. In the Results and Discussion section, the paper presents and discusses the outcomes obtained by applying the method to the Malaysian stock market. Finally, the Conclusion section offers a summary of the findings and key insights derived from the study, emphasizing the effectiveness and potential implications of the proposed approach in portfolio optimization.

METHODOLOGY

MEAN-VARIANCE MODEL

Markowitz's MV model considers both the return and risk of the portfolio. One can think of maximizing the return and minimizing the risk to find the optimal portfolio.

Let $x \in \mathbb{R}^N$ be the weight vector of the N assets, $V \in \mathbb{R}^{N\times N}$ be a positive semidefinite covariance matrix, and $r \in \mathbb{R}^N$ be the mean return vector. A portfolio is defined as a vector of asset weights x_i for assets $i = 1,\dots,N$ that represents the proportion of capital to be invested in each asset. Consider the following constrained optimization problem:

$$\min_{x\in\mathbb{R}^N} \frac{1}{2} x^T V x - r^T x, \tag{1}$$

$$s.t. \sum_{i=1}^{N} x_i = 1,$$

$$x_i \geq 0.$$

Here, $x^T V x$ represents the risk of the portfolio and $r^T x$ represents the portfolio mean return. In this study, we include a budget constraint and no short sell constraints. Now, by letting β_1, θ, β_2, and γ be the positive coefficient parameters, we present our modified norm-constrained portfolio selection model as follows:

$$\min_{x\in\mathbb{R}^N} \frac{\beta_1}{2} x^T V x - \theta r^T x + \frac{\beta_2}{2} \parallel x \parallel_2^2 + \gamma \parallel x \parallel_0 \tag{2}$$

$$s.t. \sum_{i=1}^{N} x_i = 1,$$

$$x_i \geq 0.$$

Among different regularization constraints in the function, l_2 −norm encourages diversification and stability to the portfolio, while l_0 −norm encourages sparse solutions. As mentioned in (1), the investor's capital must be fully invested and short selling is not permitted. However, solving constrained optimization problems can be challenging. To address this, we introduce the ALM into the MV function to convert the constrained objective function to an unconstrained function, thereby facilitating the optimization process.

The augmented Lagrangian function can be expressed as:

$$\mathcal{L}_A(x, \lambda; \delta) \overset{def}{=} f(x) + \sum_{i \in \epsilon} \lambda_i c_i(x) + \frac{\delta}{2} \sum_{i \in \epsilon} c_i^2(x)$$

where δ is the penalty parameter and λ_i is the estimate of the Lagrange multiplier for the equality constraint i, $c_i(x)$.

Let $\delta > 0$ be a given parameter, the augmented Lagrangian function of (2) is defined as follows:

$$\begin{aligned}\mathcal{L}_A(x, \lambda; \delta) &= \tfrac{\beta_1}{2} x^T V x - \theta r^T x + \tfrac{\beta_2}{2} \| x \|_2^2 \\ &+ \lambda(e^T x - 1) + \tfrac{\delta}{2}(e^T x - 1)^2\end{aligned} \qquad (3)$$

GRADIENT METHOD OPTIMIZATION

In the field of optimization, gradient methods are widely used for solving optimization problems, primarily because they only require the gradient of the objective function. This property makes gradient methods computationally efficient and applicable to a wide range of optimization tasks. Given the unconstrained optimization problem as follows:

$$\min_{x \in \mathbb{R}^n} f(x),$$

where f is a twice continuously differentiable function. The method generates a sequence x_k by the following rule:

$$x_{k+1} = x_k + \alpha_k d_k,$$

where d_k is the search direction and $\alpha_k > 0$ is the steplength.

Cauchy (1847) introduced the steepest descent method and it is the simplest gradient method for solving large-scale unconstrained optimization problems. It has a low computational cost per iteration and only low storage, $O(n)$ is required. Despite these advantages, it has a slow convergence rate and 'zig-zagging' behaviour when almost reaching the optimum point.

To solve this issue, a modified gradient method called the Multiple Damping Gradient method (MDG) is proposed (Sim, Leong & Chen 2019). It aims to align the step length for each gradient component with a parameter corresponding to some values within the spectrum of the local Hessian matrix and eventually force the algorithm in the one-dimensional subspace to be spanned by eigen-direction.

Later, the proximal spectral gradient method (PSG) is proposed for solving sparse optimization (Woo et al. 2023). In this paper, Woo et al. (2023) incorporated the MDG with a proximal method. The numerical results were promising and outperformed the Proximal Broyden–Fletcher–Goldfarb–Shanno and Proximal Steepest Descent. Therefore, we will apply the PSG for sparse portfolio optimization in this study.

PROXIMAL GRADIENT METHOD

The modified MV objective function in (2) can be separated into two parts: the smooth function terms and the non-smooth l_0 −norm. The smooth function terms will be solved using MDG whereas the l_0 −norm will be solved using ITH.

MDG belongs to the spectral gradient method family and is a novel approach (Sim, Leong & Chen 2019). It requires only $O(n)$ storage by replacing the full rank Hessian matrix with a $N \times N$ diagonal matrix B_{k+1}^{-1}. The inverse of B_{k+1} which is denoted by

$$H_{k+1} = B_{k+1}^{-1},$$

gives us the approximation of the inverse Hessian matrix.

Let $B_{k+1} = \mathrm{diag}\left(B_{k+1}^{(1)}, \ldots, B_{k+1}^{(n)}\right)$ and $s_k = \left(s_k^{(1)}, \ldots, s_k^{(n)}\right)$. Given the Lagrange multiplier ω, the i^{th} diagonal matrix $B_{k+1}^{(i)}$ is given as follows:

$$B_{k+1}^{(i)} = \frac{1}{1+\omega\left(s_k^{(i)}\right)^2}, i = 1,2, \ldots, n, \qquad (4)$$

and the cumulative function is written as:

$$F(\omega) = \left(\sum_{i=1}^{n} \frac{\left(s_k^{(i)}\right)^2}{1+\omega\left(s_k^{(i)}\right)^2}\right) - s_k^T y_k, \qquad (5)$$

where $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$, with g denotes the gradient of the function. We can obtain the Lagrange multiplier ω through approximation by one iteration of Newton-Raphson, starting from $\underline{\omega} = 0$. When $s_k^T s_k > s_k^T y_k$, $F(\omega)$ yields a unique positive solution, thus the approximation of ω_k is as follows:

$$\omega_k \approx \underline{\omega} - \frac{F(\underline{\omega})}{F'(\underline{\omega})}$$

$$= \frac{s_k^T s_k - s_k^T y_k}{\sum_{i=1}^{n}\left(s_k^{(i)}\right)^4}. \qquad (6)$$

Therefore, by combining (4) and (6), B_{k+1} is updated as follow:

$$B_{k+1} = \begin{cases} \operatorname{diag}\left(B_{k+1}^{(1)}, \dots, B_{k+1}^{(n)}\right), & \text{if } s_k^T s_k > s_k^T y_k, \\ \frac{s_k^T y_k}{s_k^T s_k} I, & \text{otherwise,} \end{cases} \tag{7}$$

where $\frac{s_k^T y_k}{s_k^T s_k}$ is the Oren-Luenberger scaling (Luenberger & Ye 2021).

To solve the non-smooth, nonconvex, and nondifferentiable l_0 − norm, the MDG is incorporated with the ITH. ITH is simple to apply, and it can induce sparsity of the portfolio. For every $\mu_{sparse} > 0$ and $x \in \mathbb{R}^N$, we can express the proximal operator of l_0 − norm in terms of element-wise thresholding:

$$prox_{\mu, |\cdot|_0}(x)_i = f(x) = \begin{cases} 0, & |x_i| < \sqrt{2\mu_{sparse}}, \\ x_i, & |x_i| = \sqrt{2\mu_{sparse}}, \\ x_i, & |x_i| > \sqrt{2\mu_{sparse}}. \end{cases}$$

Now we are ready to present the algorithm for the PSG.
PSG Algorithm:

1. Set $k = 0$, initialize $x_0 \in \mathbb{R}^N$ and Lagrange multiplier λ_0. Set a sparsity control parameter $\mu_{sparse} > 0$, convergence tolerance ϵ and B_0 as $N \times N$ identity matrix.

2. Compute g_0, the derivative of (3) and $d_0 = -B_0^{-1} g_0$.

3. For every $k \geq 0$, compute $x_{k+1} = x_k + \alpha_k d_k$ iteratively.

4. Update x_{k+1} by the proximal method.

$$x_{k+1}^{(i)} = \begin{cases} 0, & \text{if } x_{k+1}^{(i)} < \sqrt{2\mu_{sparse}}, \\ x_{k+1}^{(i)}, & \text{if } x_{k+1}^{(i)} = \sqrt{2\mu_{sparse}}, \\ x_{k+1}^{(i)}, & \text{if } x_{k+1}^{(i)} > \sqrt{2\mu_{sparse}}. \end{cases}$$

5. Update $\lambda_{k+1} = \lambda_k + \delta(e^T x - 1)$.

6. If $\left|\sum x_{k+1}^{(i)} - 1\right| \leq \epsilon$, stop, else compute g_{k+1}, B_{k+1}, and $d_{k+1} = -B_{k+1}^{-1} g_{k+1}$, where B_{k+1} is defined by (7).

7. Set $k = k + 1$, and go to step 3.

Note that we compute the step size α using the Lipschitz constant, L. Given two points x and y, a function is Lipschitz continuous if there exists a real positive constant L such that

$$\| f'(x) - f'(y) \| \leq L \| x - y \|.$$

Based on (3), the first-order partial derivative is as follows:

$$\mathcal{L}'_A(x, \lambda; \delta) = \beta_1 V x - \theta r + \beta_2 x + \lambda e + \delta(e^T x - 1)e.$$

Therefore, by applying the triangular inequality:

$$\begin{aligned} \| \mathcal{L}'(x) - \mathcal{L}'(y) \| &= \| \beta_1(Vx - Vy) + \beta_2(x - y) + \delta e e^T (x - y) \| \\ &= \| \beta_1 V + \beta_2 I + \delta e e^T \| \| x - y \| \\ &\leq (\beta_1 \| V \|_2 + \beta_2 \| I \|_2 + \delta \| e e^T \|_2) \| x - y \| \\ &= \left(\beta_1 \sqrt{\operatorname{tr}(VV^T)} + \beta_2 \sqrt{n} + \delta n\right) \| x - y \|. \end{aligned}$$

Based on the inequality, we can approximate the Lipschitz constant L as

$$\beta_1 \sqrt{\operatorname{tr}(VV^T)} + \beta_2 \sqrt{n} + \delta n.$$

Here, as $L > 1$, we obtain our α with $\min\left\{1, \frac{1}{L}\right\}$.

RESULTS AND DISCUSSION

In our study, we conducted all our analysis using Python 3.11 software on Dell Vostro 14 5468 (2.7 GHz Intel Core i5 8GB). We first gather the daily adjusted closing prices, Adj_t of the stocks using the *Yfinance* package in Python and compute the daily excess returns, r_t using the formula:

$$r_t = \frac{Adj_t - Adj_{t-1}}{Adj_{t-1}} - r_{f,daily},$$

where $r_{f,daily}$ is the daily risk-free rate that can be obtained by dividing the annual risk-free rate by 252 trading days. The adjusted closing price is taken into consideration as it reflects the stock's value after accounting for any corporate actions, such as stock splits, dividends, and right offerings. In this study, we set the risk-free rate at 1.75% per annum following Malaysia's central bank rate in March 2022.

With these daily returns, we normalize the daily returns using a modified Manly Box-Cox transformation (Hawkins & Weisberg 2017; Manly 1976). The modified Manly Box-Cox transformation is based on a two-step procedure: Step 1: Transform x to $y = 0.5(x + \sqrt{x^2 + c^2})$, and Step 2: Apply the Manly Box-Cox transformations:

$$w = \begin{cases} \frac{\exp(\phi y) - 1}{\phi}, & \phi \neq 0 \\ y & \phi = 0 \end{cases}$$

where c is a constant and the parameter ϕ is estimated using the *Yeo-Johnson* method that provided by `Scipy.stats` package in Python. In this study, we set $c = 0.01$. After that, we continue to find the annualized mean excess return and the standard deviation of the stocks, thus, computing the necessary mean return vector and covariance matrix to apply in sparse portfolio optimization. However, when analysing the performance of the resulting sparse portfolio, we use the original daily returns to compute the annual returns and other indicators.

In this paper, we set the convergence tolerance as $\epsilon = 10^{-6}$. We stop the algorithm when $\left|\sum x_{k+1}^{(i)} - 1\right| \leq \epsilon$. Since our focus in this study was not to find the best parameter, we fix our coefficient of the terms as $\beta_1 = 1$, $\theta = 1$, $\beta_2 = 1, \lambda = 1, \delta = 80$. The δ is set large enough to ensure the resulting weights are non-negative.

FBMKLCI AND FBMT100

In this section, we evaluate the application of PSG on two Bursa Malaysia indexes: FTSE Bursa Malaysia KLCI (FBMKLCI) and FTSE Bursa Malaysia Top 100 Cap Index (FBMT100). The FBMKLCI comprises the 30 largest companies listed on the Bursa Malaysia Main Board by full market capitalization from the 13 sectors and most of them are blue chips companies, while FBMT100 comprises the top 100 companies for Corporate Governance disclosure by rank and performance with market capitals ranging from RM30 million to RM10+ billion (0.0017% to 0.58% of total Malaysia's market capital).

In this section, we select a study period from 2018 to 2021, to assess the performance of the PSG method both before and during the COVID-19 pandemic period. We apply the PSG to the two indexes in 2018, 2019, and 2020 and fix the number of active stocks in the sparse portfolio each year to be 10. After obtaining the weights of the 10 active stocks for each year, we retain the sparse portfolios for one year and analyse their performance in terms of annual return, the Sharpe ratio and the Sortino ratio.

Sharpe ratio measures the stock's performance by comparing it to a risk-free investment while considering the risk.

$$Sharpe\ Ratio = \frac{r - r_f}{\sigma}, \sigma = \sqrt{\frac{\sum(r_{daily} - \bar{r})^2}{number\ of\ trading\ days}},$$

where σ is the standard deviation of the stock.

On the other hand, the Sortino ratio measures the stock's performance relative to its downside risk or negative volatility.

$$Sortino\ Ratio = \frac{r - r_f}{\sigma_d}, \sigma_d = \sqrt{\frac{\sum min(r_{daily} - r_{f,daily}, 0)^2}{number\ of\ trading\ days}},$$

where σ_d is the downside deviation.

Tables 1 and 2 show the performance of the optimal sparse portfolio generated by PSG, with an equal-weightage (EW) portfolio as a benchmark. We also include the performance of the index itself during the year as a comparison.

From Table 1, PSG's sparse portfolio outperforms the FBMKLCI but could not outperform the EW portfolio in 2019. Its annual return, Sharpe ratio and Sortino ratio are 2.3550%, 0.1522 and 0.2610, which are half the result of the EW portfolio. In 2020, when the unprecedented COVID-19 pandemic occurred, Malaysia's economy faced a structural break which impacted all the stocks. The blue chip companies were no exception. Nevertheless, PSG's sparse portfolio performs better than the EW portfolio and FBMKLCI as it manages to capture the companies in 2019 that were resistant to the COVID-19 pandemic's impact. However, we can see it fails to predict the price trend of the stocks in 2021. In 2021, some of the companies in PSG's sparse portfolio obtained in 2020 were impacted negatively, causing it to have the worst annual return (-5.2327%), but it still managed to perform better than FBMKLCI in Sharpe ratio (-0.0556) and Sortino ratio (-0.0439). Practically, since the desired number of stocks to be held in a portfolio is within 10 to 30 stocks and FBMKLCI comprises good-performing blue-chips stocks, it is not desirable to apply the PSG for further portfolio optimization. In this case, the EW method is enough to give the investors a good return and risk diversification.

In Table 2, we can see that PSG's sparse portfolio consistently performs better than EW portfolio and

TABLE 1. Performance of the PSG portfolio using FBMKLCI

Method	Year	Annual Return (%)	Sharpe Ratio	Sortino Ratio
FBMKLCI		-3.5899	-0.7107	-0.6503
EW	2019	**13.6893**	**0.2454**	**0.4323**
PSG		2.3550	0.1522	0.2610
FBMKLCI		2.0498	0.1142	0.2859
EW	2020	15.0784	0.1953	0.3221
PSG		**16.5156**	**0.2298**	**0.3515**
FBMKLCI		-2.5380	-0.3582	-0.2783
EW	2021	**-2.3252**	**0.0074**	**0.0348**
PSG		-5.2327	-0.0556	-0.0439

TABLE 2. Performance of the PSG portfolio using FBMT100

Method	Year	Annual Return (%)	Sharpe Ratio	Sortino Ratio
FBMT100		-1.7039	-0.4344	-0.2785
EW	2019	14.2729	**0.3450**	**0.6857**
PSG		**14.9992**	0.1552	0.5582
FBMT100		2.6953	0.1476	0.3228
EW	2020	10.5592	0.2049	0.3714
PSG		**20.3025**	**0.4728**	**0.7322**
FBMT100		-2.9004	-0.4420	-0.3662
EW	2021	0.0734	0.0301	0.1085
PSG		**6.3221**	**0.3527**	**0.5693**

the FBMT100 in all three years despite the COVID-19 pandemic, except the Sharpe ratio and the Sortino ratio in 2019 (0.1552 and 0.5582), which are lower than EW portfolio's (0.3450 and 0.6857). In short, PSG works in portfolio optimization when the portfolio size is large as it can scan through more stocks that have better performance and retain them in the portfolio.

G-COV AND G-SR

With the previous findings, it is crucial to note that PSG works efficiently only when the portfolio size is large. Also, when we analyse the stocks individually in the whole Malaysia market, there are more stocks that are not in the FBMT100 but are performing better than some stocks in FBMT100. Therefore, we take a step further by creating a new portfolio using 100 stocks with the smallest positive coefficient of variation; with the best risk-return trade-off in each year. We name this portfolio as G-COV. To determine the stocks' coefficient of variation, we use the one-year preceding daily adjusted closing price.

We apply the PSG to generate five sparse portfolios, each with 10, 15, 20, 25 and 30 active stocks, respectively (PSG-10, PSG-15, PSG-20, PSG-25, PSG-30). We carry out our analysis each year from 2012 to 2021, eventually summarizing the results into 3 years (2019-2021), 5 years (2017-2021), and 10 years (2012-2021) and averaging the results based on the duration of the timeframes. We compare these sparse portfolios in terms of average annual returns, average Sortino ratio and average M^2 ratio. We also add the EW portfolio as a benchmark. M^2 ratio evaluates the performance of the portfolio relative to another portfolio as a benchmark; in this study, the benchmark is the FBMT100 index. It is calculated with the formula herewith:

$$M^2 ratio = (S_p - S_m)\sigma_m$$

where S_p is Sharpe ratio of the portfolio; S_m is the Sharpe ratio of the market index; and σ_m is the standard deviation of the market index. In Table 3, all the sparse portfolios perform better than the EW portfolio in 3 years, 5 years, and 10 years timeframe in all the indicators. In the average annual return section, the PSG-20 portfolio obtains the highest 3 years average annual return (19.3713%), 5 years average annual return (12.0532%) and 10 years average annual return (15.1996%). On the other hand, the PSG-10 portfolio has the highest average Sortino ratio except in 3 years timeframe (0.7596) which is slightly lower than PSG-15 (0.7622). It also has the highest average M^2 ratio among the sparse portfolios. Therefore, this shows that it can perform well even during a market downfall. Although PSG-10 may not have the highest average annual return, it is the best-performing sparse portfolio among other sparse portfolios when we investigate other indicators.

Other than G-COV, we have constructed another portfolio with 100 stocks with the best Sharpe ratio using one-year preceding data and named this portfolio G-SR. The difference between the Sharpe ratio and the coefficient of variation is that the Sharpe ratio uses the mean excess return after deducting the risk-free rate. In contrast, the coefficient of variation uses the realized mean return. We conduct the same procedure by applying the PSG onto G-SR. The results of the performance of the sparse portfolios are shown in Table 4.

In Table 4, the PSG's sparse portfolios' performance surpasses the EW portfolio when we use G-SR. In the average annual return section, the PSG-15 has the lowest average annual return among all the sparse portfolios with 11.4280% in the 3 years timeframe, 7.5255% in the 5 years timeframe, and 10.6492% in the 10 years timeframe. It is also the only sparse portfolio that has a lower average annual return than the EW portfolio in the 3 years and 5 years timeframe. Nevertheless, when we look at its average Sortino ratio and average M^2 ratio, PSG-15 has a better average Sortino ratio and average M^2 ratio than the EW portfolio, showing that it is still performing better than the EW portfolio. What we need to consider, too, is that PSG-15 only has 15 active stocks in the portfolio whereas the EW portfolio has 100 active stocks. This difference in the number of stocks impacts transaction costs. With a smaller number of active stocks, PSG-15 potentially incurs lower transaction costs compared to the EW portfolio. Therefore, when taking transaction costs into account, PSG-15 appears to be a more promising option due to its potential for reduced transaction costs.

On the other hand, among the sparse portfolios, PSG-20 has the highest 3 years (19.7821%), 5 years (12.4460%) and 10 years' (14.6134%) average annual return, and the highest 5 years and 10 years' average Sortino ratio (0.4982 and 0.7274). However, its M^2 ratio is not the highest of all the sparse portfolios, showing that PSG-20 is riskier than some sparse portfolios, but if an investor is willing to accept higher risk in pursuit of higher returns, they have the option to consider PSG-20. We have shown that PSG works efficiently on G-COV and G-SR with one-year preceding daily adjusted closing price. We are curious too if we can obtain better results with a longer timeframe of preceding data. Thus, we conduct the same analysis again using G-COV and G-SR, but we use the two-year preceding daily adjusted closing price to compute the mean return, standard deviation, Sharpe ratio and the coefficient of variation of the stocks. The results are shown in Tables 5 and 6.

Table 5 shows an intriguing pattern observed in the sparse portfolios. It appears that as the number of active stocks increases, the overall performance of the portfolios diminishes. In other words, PSG-10 is the best-performing sparse portfolio in all three indicators, followed by PSG-15, PSG-20, PSG-25 and PSG-30. It is noteworthy that all the sparse portfolios exhibit better performance compared to the EW portfolio, except for PSG-25 and PSG-30, which show a slightly lower average annual return over the three-year period than the EW portfolio. Upon comparing the findings presented in Table 5 with the results obtained from G-COV with the one-year preceding daily adjusted closing price (Table 3), the two-year preceding daily adjusted closing price only brings slight improvement to some sparse portfolios in different indicators, but overall, it exhibits poorer performance. This indicates that when employing G-COV in portfolio optimization, it is more

TABLE 3. Performance of the sparse portfolios on G-COV with one year preceding daily adjusted closing price

	Methods						
	years	EW	PSG-10	PSG-15	PSG-20	PSG-25	PSG-30
Average	3	12.7100	16.2034	13.6743	**19.3713**	15.1606	13.3077
Annual	5	7.8495	12.0123	9.5980	**12.0532**	9.9170	9.0678
Return (%)	10	11.3101	15.1373	14.5936	**15.1996**	14.8931	13.9596
Average	3	0.4836	0.7596	**0.7622**	0.7332	0.6049	0.5856
Sortino	5	0.3208	**0.6928**	0.6270	0.5589	0.4749	0.4673
Ratio	10	0.5972	**0.9727**	0.8913	0.8201	0.8064	0.7907
Average	3	4.1667	**7.4355**	7.0221	6.0739	5.3248	5.1546
M^2	5	4.5412	**10.7199**	8.3798	6.8748	6.3837	6.4576
Ratio	10	3.6595	**7.5641**	6.1798	5.3290	5.2389	5.2163

TABLE 4. Performance of the sparse portfolios on G-SR with one year preceding daily adjusted closing price

	Methods						
	years	EW	PSG-10	PSG-15	PSG-20	PSG-25	PSG-30
Average	3	12.7100	15.9535	11.4280	**19.7821**	15.9586	13.9763
Annual	5	7.7816	9.6218	7.5255	**12.4460**	10.2190	9.0486
Return (%)	10	10.5000	12.1045	10.6492	**14.6134**	12.9524	11.6113
Average	3	0.4836	**0.7607**	0.6680	0.7204	0.5993	0.5589
Sortino	5	0.3149	0.4614	0.4510	**0.4982**	0.4482	0.4149
Ratio	10	0.6031	0.6917	0.6498	**0.7274**	0.6893	0.6441
Average	3	4.1667	**6.8833**	6.3380	6.4077	5.4865	5.2215
M^2	5	4.4983	5.4479	5.7221	5.6546	5.9037	**5.9215**
Ratio	10	3.6828	4.1608	4.2066	4.5110	**4.6082**	4.4722

TABLE 5. Performance of the sparse portfolios on G-COV with two years preceding daily adjusted closing price

	Methods						
	years	EW	PSG-10	PSG-15	PSG-20	PSG-25	PSG-30
Average	3	11.4107	**17.4323**	15.0601	12.5226	11.3404	11.3353
Annual	5	7.4568	**12.3364**	12.0332	9.0441	8.9115	8.2327
Return (%)	10	9.8526	**14.9784**	14.8616	12.1272	11.5505	10.8087
Average	3	0.4246	**0.7578**	0.6621	0.5525	0.5253	0.4454
Sortino	5	0.3418	**0.6907**	0.5876	0.4518	0.4665	0.3862
Ratio	10	0.6217	**0.9514**	0.8517	0.7423	0.7227	0.6628
Average	3	3.5648	**5.7441**	5.1603	4.7250	4.4102	3.7301
M^2	5	5.2078	**10.2512**	8.4039	6.4754	7.3641	6.1794
Ratio	10	4.0398	**7.4063**	6.2020	5.0102	5.3438	4.6043

TABLE 6. Performance of the sparse portfolios on G-SR with two years preceding daily adjusted closing price

	Methods						
	years	EW	PSG-10	PSG-15	PSG-20	PSG-25	PSG-30
Average	3	11.4480	**39.7300**	31.6336	24.8149	19.8004	16.9809
Annual	5	6.9503	**20.5944**	13.8313	14.1015	12.0650	11.0655
Return (%)	10	9.6809	**18.3246**	17.8339	14.3584	13.2885	12.0484
Average	3	0.4286	**1.0017**	0.9532	0.7711	0.6671	0.5021
Sortino	5	0.3135	0.5735	**0.6091**	0.4554	0.4759	0.3806
Ratio	10	0.6054	**0.8423**	0.8022	0.6857	0.6915	0.6070
Average	3	3.5897	**7.6529**	7.0470	6.1856	5.5271	4.2747
M^2	5	4.9861	**6.9286**	6.7955	5.7040	6.5879	6.1710
Ratio	10	3.8974	**5.5532**	5.2102	4.4089	4.8194	4.4130

advantageous to utilize the one-year preceding daily adjusted closing price.

Table 6 shows the same pattern as Table 5 where PSG-10 has the best performance while PSG-30 has the worst performance. There is only one time when PSG-15 performs better than PSG-10 in the 5 years average Sortino ratio. Also, all the sparse portfolios are performing better than the EW portfolio. When we compare this result with the result using G-SR with one-year preceding daily adjusted closing price (Table 4), we can see a remarkable improvement in most of the indicators except for PSG-20 and PSG-30 in average Sortino ratio and average M^2 ratio. Thus, PSG works more favourably on G-SR when we use the two-year preceding daily adjusted closing price.

CONCLUSIONS

In conclusion, our study demonstrates the effectiveness of the PSG method in constructing sparse portfolios. We begin by formulating a modified norm-constraint mean-variance portfolio selection objective function using l_0 norm, l_2 norm and ALM. Next, we apply PSG comprising of MDG and ITH onto two of Malaysia's market indexes: FBMKLCI and FBMT100. The result shows that PSG can create a sparse portfolio that outperforms the EW portfolio when the portfolio size is large and the market conditions are favourable. We investigate two sets of portfolios with 100 stocks: G-COV and G-SR. We apply the PSG to construct the sparse portfolios with varying numbers of active stocks. We find that the sparse portfolios outperform the EW portfolio in G-COV and G-SR using either one-year or two-year preceding daily adjusted closing price. Also, the sparse portfolio with 10 to 20 active stocks yields the best performance compared to other sparse portfolios. These findings highlight the potential of PSG for enhanced portfolio optimization and diversification strategies.

ACKNOWLEDGMENTS

This research was supported by the Ministry of Higher Education (MoHE), Malaysia through Fundamental Research Grant Scheme (FRGS/1/2022/STG06/UTAR/02/2). The authors gratefully acknowledge Universiti Tunku Abdul Rahman which supported this work under UTARRF (IPSR/RMC/UTARRF/2021-C1/W02).

REFERENCES

Cauchy, A. 1847. Méthode générale pour la résolution des systemes d'équations simultanées. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* 25: 536-538.

Chen, J., Dai, G. & Zhang, N. 2020. An application of sparse-group LASSO regularization to equity portfolio optimization and sector selection. *Annals of Operations Research* 284(1): 243-262.

Chopra, V.K. & Ziemba, W.T. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management* 19(2): 6-11.

DeMiguel, V., Garlappi, L. & Uppal, R. 2011. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? In *Heuristics: The Foundations of Adaptive Behavior*, edited by Gigerenzer, G., Hertwig, R. & Pachur, T. Oxford Academic. pp. 644-667.

Gotoh, J. & Takeda, A. 2011. On the role of norm constraints in portfolio selection. *Computational Management Science* 8(4): 323-353.

Hawkins, D.M. & Weisberg, S. 2017. Combining the Box-Cox power and generalised log transformations to accommodate nonpositive responses in linear and mixed-effects linear models. *South African Statistical Journal* 51(2): 317-328.

Hwang, I., Xu, S. & In, F. 2018. Naive versus optimal diversification: Tail risk and performance. *European Journal of Operational Research* 265(1): 372-388.

Jagannathan, R. & Ma, T. 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance* 58(4): 1651-1683.

Kremer, P.J., Lee, S., Bogdan, M. & Paterlini, S. 2020. Sparse portfolio selection via the sorted l_1-norm. *Journal of Banking & Finance* 110: 105687.

Luenberger, D.G. & Ye, Y. 2021. *Linear and Nonlinear Programming.* 5th ed. Cham Switzerland: Springer.

Manly, B.F.J. 1976. Exponential data transformations. *Journal of the Royal Statistical Society Series D: The Statistician* 25(1): 37-42.

Markowitz, H. 1952. Portfolio selection. *The Journal of Finance* 7(1): 77-91.

Olivares-Nadal, A.V. & DeMiguel, V. 2018. Technical note - A robust perspective on transaction costs in portfolio optimization. *Operations Research* 66(3): 733-739.

Sim, H.S., Leong, W.J. & Chen, C.Y. 2019. Gradient method with multiple damping for large-scale unconstrained optimization. *Optimization Letters* 13(3): 617-632.

Sim, H.S., Ling, W.S.Y., Leong, W.J. & Chen, C.Y. 2023. Proximal linearized method for sparse equity portfolio optimization with minimum transaction cost. *Journal of Inequalities and Applications* 152.

Wang, H., Zhang, W., He, Y. & Cao, W. 2023. l_0-norm based short-term sparse portfolio optimization algorithm based on alternating direction method of multipliers. *Signal Processing* 208: 108957.

Woo, G.Y.H., Sim, H.S., Goh, Y.K. & Leong, W.J. 2023. Proximal variable metric method with spectral diagonal update for large scale sparse optimization. *Journal of the Franklin Institute* 360(7): 4640-4660.

*Corresponding author; email: wongwk@utar.edu.my