ON ANALYZING AND DETECTING MULTIPLE OPTIMA OF PORTFOLIO OPTIMIZATION

YUE QI
China Academy of Corporate Governance & Department of Financial Management
Business School & Collaborative Innovation Center for China Economy
Nankai University, 94 Weijin Road, Tianjin, 300071, China

ZHIHAO WANG AND SU ZHANG*
Department of Financial Management, Business School, Nankai University
94 Weijin Road, Tianjin, 300071, China

(Communicated by Jie Sun)

Abstract. Portfolio selection is widely recognized as the birth-place of modern finance; portfolio optimization has become a developed tool for portfolio selection by the endeavor of generations of scholars. Multiple optima are an important aspect of optimization. Unfortunately, there is little research for multiple optima of portfolio optimization. We present examples for the multiple optima, emphasize the risk of overlooking the multiple optima by (ordinary) quadratic programming, and report the software failure by parametric quadratic programming. Moreover, we study multiple optima of multiple-objective portfolio selection and prove the nonexistence of the multiple optima of an extension of the model of Merton. This paper can be a step-stone of studying the multiple optima.

1. Introduction. Portfolio selection of [16] is widely recognized as the birth-place of modern finance as commented by [22]; portfolio optimization has become a developed tool for portfolio selection by the endeavor of generations of scholars in the following aspects:

- (ordinary) quadratic programming by, e.g., [28], [11], and [12];
- parametric quadratic programming by, e.g., [18], [23], and [13]; and
- model simplification or transformation by, e.g., [17] and [15].

Multiple optima are an important aspect of optimization and can bring instability and incomplete result. Unfortunately, there is little research for multiple optima of portfolio optimization to our knowledge; only the following researchers indirectly study the topic: [7] studies kinks on efficient frontiers with short sales restrictions. [26] derives efficient frontiers in the case of degeneracy and general singularity. [27] comment the work of [7] and study the sufficient conditions for kinks.

2010 Mathematics Subject Classification. Primary: 90B50; Secondary: 90C29.
Key words and phrases. Multiple-objective optimization, portfolio optimization, multiple optimal solutions, multiple efficient solutions, parametric quadratic programming.

The first author is supported by Asian Research Center, Nankai University grant AS1525 and Key Faculty of Major Discipline of Social Science, Nankai University, 2017. The third author is supported by National Natural Science Foundation of China grant 11401322 and Fundamental Research Funds for the Central Universities grant NKZXB1447.

* Corresponding author: Su Zhang.
We present examples for the multiple optima, emphasize the risk of overlooking the multiple optima, and report the software failure. Moreover, because multiple-objective portfolio selection is becoming an interesting topic as surveyed by [25] and [29], we introduce multiple-objective portfolio selection and prove the nonexistence of the multiple optima of an extension of the model of [19].

The rest of this paper is organized as follows: We review multiple-objective optimization, portfolio selection, and portfolio optimization in Section 2. In Section 3, we present examples for multiple optima of portfolio optimization, emphasize the risk of overlooking the multiple optima, and report the failure of portfolio optimization software. We extend the model of [19] by multiple-objective portfolio selection and prove the nonexistence of the multiple optima in Section 4. We conclude the paper in Section 5.

2. Multiple-objective optimization and portfolio selection.

2.1. Multiple-objective optimization. Multiple-objective optimization can be formulated as

\[
\begin{align*}
\max \{ z_1 &= f_1(x) \} \\
\vdots \\
\max \{ z_k &= f_k(x) \} \\
\text{s.t. } x &\in S
\end{align*}
\]

where \( x \in \mathbb{R}^n \) is a decision vector in decision space; \( k \) is the number of objectives; \( f_1(x) \ldots f_k(x) \) are objective functions; \( z = \begin{bmatrix}
    z_1 \\
    \vdots \\
    z_k
\end{bmatrix} \) is a criterion vector in criterion space; \( S \) is the feasible region in decision space; and \( Z = \{ z \mid x \in S \} \) is the feasible region in criterion space.

That \( \bar{z} \) dominates \( z \) is defined as \( \bar{z}_1 \geq z_1 \ldots \bar{z}_k \geq z_k \) with at least one strict inequality. That \( \bar{z} \) is nondominated is defined as that there does not exist \( z \in Z \) such that \( z \) dominates \( \bar{z} \); otherwise, \( \bar{z} \in Z \) is dominated. That \( \bar{x} \) is efficient is defined as that its criterion vector \( \bar{z} \) is nondominated; otherwise, \( \bar{x} \) is inefficient.

The set of all nondominated criterion vectors is called nondominated set and denoted as \( N \). The set of all efficient decision vectors is called efficient set and denoted as \( E \). The purpose of multiple-objective optimization is to compute both \( E \) and \( N \) for the discretion of decision makers. One mechanism to solve (1) is an \( \epsilon \)-constraint approach. By the approach, only one objective function is retained, while the others are transformed into constraints. [24], [20], and [8] describe multiple-objective optimization in details.

For single-objective optimization \( \max f(x), \text{s.t. } x \in S \), suppose that both \( x^1 \) and \( x^2 \) with \( x^1 \neq x^2 \) are optimal solutions; then, we call \( x^1 \) and \( x^2 \) as multiple optimal solutions (multiple optima). We define optimal solution in traditional way (e.g., the definition of [12]).

For multiple-objective optimization, an efficient set typically has many efficient solutions and the solutions typically have different criterion vectors. Then, we extend multiple optima as follows: For (1), suppose that both \( x^1 \) and \( x^2 \) with \( x^1 \neq x^2 \) are efficient and have identical criterion vectors (i.e., \( z^1 = z^2 \)); then, we call \( x^1 \) and \( x^2 \) as multiple efficient solutions.
For notation, we use bold-face symbols (e.g., \(x\)) to denote vectors or matrices and use normal symbols (e.g., \(a\)) to denote scalars. For stock \(A\), we use \(\mu_A\), \(\sigma_A\), \(x_A\), and \(x_A\) to respectively denote the expected return, standard deviation, portfolio weight vector, and element for stock \(A\) in a portfolio weight vector. We sometimes use \(A\) instead of stock \(A\), \(P_1\) instead of portfolio \(P^1\), and \(A\) to \(C\) instead of stocks \(A\) to \(C\). We respectively denote the covariance and correlation between \(A\) and \(B\) by \(\sigma_{AB}\) and \(\rho_{AB}\). We deploy similar notation for other stocks. We also interchangeably use \(x_A\) and \(x_1\), although \(x_A\) is more indicative than \(x_1\).

2.2. Portfolio selection. [16] formulates portfolio selection as

\[
\begin{align*}
\min \{ z_1 &= x^T \Sigma x \} \\
\max \{ z_2 &= \mu^T x \}
\end{align*}
\tag{2}
\]

subject to \(x \in S\) where for \(n\) stocks, \(x\) is an \(n\)-vector for portfolio weights; \(\Sigma\) is an \(n \times n\) covariance matrix of stock returns; \(\mu\) is an \(n\)-vector for stock expected returns; \(z_1\) measures portfolio variance; and \(z_2\) measures portfolio expected return. [16] calls the nondominated set of (2) as an efficient frontier. By differentiating decision space with criterion space, we reserve efficient for decision space and reserve nondominated for criterion space. [19] analyzes

\[
\begin{align*}
\min \{ z_1 &= x^T \Sigma x \} \\
\max \{ z_2 &= \mu^T x \}
\end{align*}
\tag{3}
\]

subject to \(1^T x = 1\) where \(1\) is an \(n\)-vector of ones. On one hand, unrealistically unlimited weights are allowed in (3). On the other hand, most result of (3) can be analytically derived. This analyticity brings substantial advantage in research and teaching (e.g., the text of [14]); (3) also serves as the foundation of asset pricing models. [19] applies an \(e\)-constraint approach to (3) and gets

\[
\begin{align*}
\min \{ z_1 &= x^T \Sigma x \} \\
\text{s.t. } &\mu^T x = e_2 \\
&1^T x = 1
\end{align*}
\tag{4}
\]

where \(e_2\) is a parameter. The criterion vector set of all the optimal solutions of (4) with varying \(e_2\) is called the minimum-variance frontier of (3).

2.3. Portfolio optimization. We demonstrate a comprehensive picture of portfolio optimization by depicting the (whole) feasible region \(Z\) as shaded in Figure 1. \(C\) is called the maximum expected-return portfolio and can be obtained by

\[
\begin{align*}
\max \{ z_2 &= \mu^T x \}
\end{align*}
\tag{5}
\]

subject to \(x \in S\)

\(B\) is called the minimum variance portfolio and can be obtained by

\[
\begin{align*}
\min \{ z_1 &= x^T \Sigma x \}
\end{align*}
\tag{6}
\]

subject to \(x \in S\)
The boundary from $B$ to $C$ as a solid thick curve is the nondominated set and can be obtained by (2). $A$ is called the minimum expected-return portfolio, doesn’t belong to the nondominated set, and can be obtained by

\[
\min \{ z_2 = \mu^T x \} \\
\text{s.t. } x \in S
\]

The boundary from $A$ to $B$ as a solid thin curve can be obtained by

\[
\min \{ z_1 = x^T \Sigma x \} \\
\min \{ z_2 = \mu^T x \} \\
\text{s.t. } x \in S
\]

The boundary from $A$ to $C$ as a broken curve can be obtained by

\[
\max \{ z_1 = x^T \Sigma x \} \\
\text{s.t. } x \in S \\
\mu^T x = e_2
\]

where $e_2 \in [\mu_A, \mu_C]$; $\mu_A$ and $\mu_C$ are respectively the expected returns of $A$ and $C$. Model (7) is difficult to solve, because we are maximizing a convex function. Therefore, we typically rarely see $Z$ of portfolio selection.

To compute the nondominated set, the major style (e.g., Chapter 7 of [1]) is as follows:

1. computing $C$ by (5);
2. computing $B$ by (6);
3. computing the mid portfolios by

\[
\min \{ z_1 = x^T \Sigma x \} \\
\text{s.t. } x \in S \\
\mu^T x = e_2
\]

where $\mu_B$ and $\mu_C$ are respectively the expected returns of $B$ and $C$, a set of $e_2 \in [\mu_B, \mu_C]$ is taken, and (8) is repetitively solved for each $e_2$; and
4. connecting all the portfolios in (standard deviation, expected return) space.

However, the major style has the following weakness: First, it obtains only an approximation of the nondominated set. Second, it can’t reveal some structure (e.g.,
ON ANALYZING PORTFOLIO MULTIPLE OPTIMA

Theorem 2.1) of the nondominated set. Last, it doesn’t consider multiple optima and can obtain incorrect result; we’ll demonstrate this weakness in Section 3.

Instead of the major style, [18] and some scholars (reviewed in Section 1) completely compute the nondominated set by parametric quadratic programming. Moreover, [18] (p.176) demonstrate the structure of the nondominated set by proposing the following theorem:

**Theorem 2.1.** For (2), if the $S$ is formed by linear constraints, the nondominated set is piece-wisely made up by connected hyperbolas in (standard deviation, expected return) space.

However, parametric quadratic programming is difficult. To our knowledge, the active, independent, and free software is just Optimizer based on the work of [18] and CIOS based on the work of [13]. [18] and [13] may not consider multiple optima either; Optimizer and CIOS fail in some examples in Section 3.

3. Examples of portfolio optimization with multiple optima. In this section, we typically utilize standard portfolio selection model of [18] (p.3) as follows:

$$
\begin{align*}
\text{min} \{ z_1 = x^T \Sigma x \} \\
\text{max} \{ z_2 = \mu^T x \} \\
\text{s.t. } 1^T x = 1 \\
x \geq 0
\end{align*}
$$

(9)

where $0$ is a vector of zeros. Also in this section, we typically use symbols $A$, $B$, and $C$ for three stocks; the $A$, $B$, and $C$ in all figures are just symbols and don’t necessarily denote the same stock (e.g., $A$ in Figure 2 and $A$ in Figure 4 are for different stocks). We start with portfolios of two stocks. [1] in Chapter 7 illustrate that the portfolios lie on a curve passing through the two stocks in (standard deviation, expected return) space. Specifically, [9] (p.67) describe the following lemma for correlation 1:

**Lemma 3.1.** If the correlation of stocks $A$ and $B$ is 1, the portfolios of $A$ and $B$ lie on a linear segment from $A$ to $B$ in (standard deviation, expected return) space.

3.1. For maximum expected-return portfolio. We analyze an example of 3-stock portfolios with multiple optima for the maximum expected-return portfolio.

For stocks $A$, $B$, and $C$, the expected return vector, covariance matrix, and correlation matrix are

$$
\mu = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 12 \\ 0 & 12 & 16 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.
$$

For simplicity, we typically use integers for $\mu$ and $\Sigma$. [3] (pp.33&35) define covariance matrices as symmetric and positive semidefinite. The eigenvalues of $\Sigma$ are 0, 4, and 25; therefore, $\Sigma$ is a (legitimate) covariance matrix. We plot $A$ to $C$ in the right part of Figure 2. Then, we

- build portfolios of $A$ and $B$ and plot the portfolios as a curve from $A$ to $B$;
- build portfolios of $A$ and $C$ and plot the portfolios as a curve from $A$ to $C$; and
- build portfolios of $B$ and $C$ and plot the portfolios as a linear segment from $B$ to $C$ by Lemma 3.1. The segment is horizontal, because $\mu_B = \mu_C$.

We obtain the $Z$ of this example by proposing the following lemma:
Lemma 3.2. The curve from A to B, curve from A to C, and linear segment from B to C are the boundary of the Z.

Proof. \( \forall \mathbf{x} = \begin{bmatrix} x_A \\ x_B \\ x_C \end{bmatrix} \in S \) (i.e., \( \begin{bmatrix} x_A \\ x_B \\ x_C \end{bmatrix} \geq 0 \) and \( x_A + x_B + x_C = 1 \)), \( \mathbf{\mu}^T \mathbf{x} = x_A + 2x_B + 2x_C \leq 2x_A + 2x_B + 2x_C = 2 = \mu_B \). Thus, the linear segment from B to C is the upper boundary of Z.

The variance of \( \mathbf{x} \) is \( \mathbf{x}^T \Sigma \mathbf{x} = 4x_A^2 + 9x_B^2 + 16x_C^2 + 24x_Bx_C \). We take \( \begin{bmatrix} x_A \\ x_B + x_C \end{bmatrix} \in S \) as a portfolio of A and B. With

\[
\begin{bmatrix} x_A \\ x_B + x_C \end{bmatrix} \sum \begin{bmatrix} x_A \\ x_B + x_C \end{bmatrix} \leq \mathbf{x}^T \Sigma \mathbf{x} \quad \mathbf{\mu}^T \begin{bmatrix} x_A \\ x_B + x_C \end{bmatrix} = \mathbf{\mu}^T \mathbf{x}
\]

the curve from A to B is the left boundary of Z.

Similarly, we take \( \begin{bmatrix} x_A \\ 0 \\ x_B + x_C \end{bmatrix} \in S \) as a portfolio of A and C. With

\[
\mathbf{x}^T \Sigma \mathbf{x} = 4x_A^2 + 9x_B^2 + 16x_C^2 + 24x_Bx_C \leq \begin{bmatrix} x_A \\ 0 \\ x_B + x_C \end{bmatrix} \sum \begin{bmatrix} x_A \\ 0 \\ x_B + x_C \end{bmatrix} \quad \mathbf{\mu}^T \begin{bmatrix} x_A \\ 0 \\ x_B + x_C \end{bmatrix} = \mathbf{\mu}^T \mathbf{x}
\]

the curve from A to C is the right boundary of Z.

We depict the Z as shaded in the right part of Figure 2. With the Z, we readily locate the N as the upper part of the curve from A to B and mark the N by a thick curve. In the left part of Figure 2, we depict the S in \( \mathbb{R}^3 \) as shaded; S is formed by a triangle passing through \( \mathbf{x}^A = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \), \( \mathbf{x}^B = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \), and \( \mathbf{x}^C = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \). Corresponding to the N, we mark the E by a thick linear segment.

The meaningfulness of this example is as follows:

1. Neither Optimizer nor CIOS can correctly solve this example.
2. By the major style of portfolio optimization (reviewed in Section 2), we try to compute the maximum expected-return portfolio by (5). Because the linear
Figure 3. Incorrectly approximating the $N$ of the example in subsection 3.1 by the major style of portfolio optimization

segment from $B$ to $C$ is correspondingly horizontal and there is no further modeling for the maximum expected-return portfolio, we typically obtain $P_1$ instead of $B$. We depict $P_1$ to $P_4$ in Figure 3.

3. By the style, we obtain $P_2$ and $P_3$ as the mid portfolios and $P_4$ as the minimum variance portfolio. We connect portfolios $P_1$ to $P_4$ to approximate the $N$ and mark the connection by a broken thick curve.

4. However, the style typically misses $B$ and consequently over-estimates the curve from $B$ to $P_2$.

We generalize this example in the following different styles (to save space, we omit the description which will be available upon requests):

• setting $\mu = \begin{bmatrix} \mu_A \\ \mu_B \\ \mu_B \end{bmatrix}$ with $0 \leq \mu_A < \mu_B$ and $\Sigma = \begin{bmatrix} \sigma_A^2 & 0 & 0 \\ 0 & \sigma_B^2 & \sigma_B \sigma_C \\ 0 & \sigma_B \sigma_C & \sigma_C^2 \end{bmatrix}$ with $\sigma_B < \sigma_C$; and

• making both $\Sigma$ and $\Omega$ invertible (thus positive definite) by $\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 9 \\ 0 & 9 & 16 \end{bmatrix}$ with eigenvalues 2.84, 4, and 22.16, and $\Omega = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.75 \\ 0 & 0.75 & 1 \end{bmatrix}$.

• for general 3-stock (or more than 3-stock) portfolio optimization with $C$ as the only maximum expected-return portfolio, i.e., $\mu = \begin{bmatrix} \mu_A \\ \mu_B \\ \mu_C \end{bmatrix}$ with $\mu_A < \mu_C$ and $\mu_B < \mu_C$ and $\Sigma = \begin{bmatrix} \sigma_A^2 & \sigma_{AB} & \sigma_{AC} \\ \sigma_{AB} & \sigma_B^2 & \sigma_{BC} \\ \sigma_{AC} & \sigma_{BC} & \sigma_C^2 \end{bmatrix}$. We add stock $D$ and set, for $A$ to $D$, $\mu = \begin{bmatrix} \mu_A \\ \mu_B \\ \mu_C \\ \mu_D \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sigma_A^2 & \sigma_{AB} & \sigma_{AC} & \sigma_{AD} \\ \sigma_{AB} & \sigma_B^2 & \sigma_{BC} & \sigma_{BD} \\ \sigma_{AC} & \sigma_{BC} & \sigma_C^2 & \sigma_{CD} \\ \sigma_{AD} & \sigma_{BD} & \sigma_{CD} & \sigma_D^2 \end{bmatrix}$ with $\sigma_C < \sigma_D$.

Then, any convex combination of $C$ and $D$ is the multiple optima of maximum expected-return portfolio.
3.2. **For minimum variance portfolio.** We analyze an example of 3-stock portfolios with multiple optima for the minimum variance portfolio. For \( A \) to \( C \), \( \mu = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \); \( \Sigma = \begin{bmatrix} 4 & 0 & 4 \\ 0 & 9 & 0 \\ 4 & 0 & 4 \end{bmatrix} \) with eigenvalues 0, 8, and 9; and \( \Omega = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \). We plot \( A \) to \( C \) in the left part of Figure 4. Then, we

- build portfolios of \( A \) and \( B \) and plot the portfolios as a curve from \( A \) to \( B \);
- build portfolios of \( C \) and \( B \) and plot the portfolios as a curve from \( C \) to \( B \).

By studying the two curves, we find that the two curves are symmetric with respect to the horizontal line (Expected return = 3); i.e., we can take the line as a symmetric axis, vertically flip the curve from \( A \) to \( B \), and get the curve from \( C \) to \( B \).

We depict the line in gray color in the left part of Figure 4. To verify the symmetric relationship, we propose the following lemma:

**Lemma 3.3.** The curve from \( A \) to \( B \) and curve from \( C \) to \( B \) are symmetric with respect to the horizontal line. For any portfolio \( P_1 \) with portfolio weight \( x^{P_1} = \begin{bmatrix} x_A \\ x_B \\ 0 \end{bmatrix} \in S \) on the curve from \( A \) to \( B \), there exists a portfolio \( P_2 \) with \( x^{P_2} = \begin{bmatrix} 0 \\ x_B \\ x_A \end{bmatrix} \in S \) on the curve from \( C \) to \( B \), so the correlation between \( P_1 \) and \( P_2 \) is 1, and \( P_1 \) and \( P_2 \) are symmetric with respect to the line.

**Proof.** The expected returns of \( P_1 \) and \( P_2 \) respectively are \( \mu^T \begin{bmatrix} x_A \\ x_B \\ 0 \end{bmatrix} = 2x_A + 3x_B \) and \( \mu^T \begin{bmatrix} 0 \\ x_B \\ x_A \end{bmatrix} = 4x_A + 3x_B \). With \( x_A + x_B = 1 \), the average of the two expected returns is

\[
\frac{(2x_A + 3x_B) + (4x_A + 3x_B)}{2} = \frac{(6x_A + 6x_B)}{2} = 3
\]

(10)

The variances of \( P_1 \) and \( P_2 \) respectively are

\[
[x_A \ x_B \ 0] \Sigma \begin{bmatrix} x_A \\ x_B \\ 0 \end{bmatrix} = 4x_A^2 + 9x_B^2, \quad [0 \ x_B \ x_A] \Sigma \begin{bmatrix} 0 \\ x_B \\ x_A \end{bmatrix} = 4x_A^2 + 9x_B^2
\]

(11)

The covariance between \( P_1 \) and \( P_2 \) is \([x_A \ x_B \ 0] \Sigma \begin{bmatrix} 0 \\ x_B \\ x_A \end{bmatrix} = 4x_A^2 + 9x_B^2 \). Thus, \( \rho_{P_1, P_2} = 1 \). By (10) and (11), \( P_1 \) and \( P_2 \) are symmetric.

In the left part of Figure 4, we depict \( P_1 \) and \( P_2 \) and construct portfolios of \( P_1 \) and \( P_2 \). Because \( \rho_{P_1, P_2} = 1 \) by Lemma 3.3, the portfolios of \( P_1 \) and \( P_2 \) lie on a linear segment from \( P_1 \) to \( P_2 \) by Lemma 3.1. Moreover, the linear segment is vertical, because \( P_1 \) and \( P_2 \) have identical variances by (11). We depict the portfolios of \( P_1 \) and \( P_2 \) by a broken line. As \( P_1 \) moves along the curve from \( A \)
to $B$ and $P^2$ correspondingly moves along the curve from $C$ to $B$, the broken line also moves. We depict five broken lines in the left part of Figure 4. As the broken line moves, we obtain a region and depict the region as shaded in the right part of Figure 4.

Moreover, we argue that the region is exactly the $Z$ of this example. To verify the argument, we study the decision space. For $P_1$ and $P_2$, because $x_A + x_B = 1$, we rewrite $x^{P_1}$ and $x^{P_2}$ as

$$x^{P_1} = \begin{bmatrix} x_A \\ 1 - x_A \\ 0 \end{bmatrix} \text{ with } x_A \in [0,1], \quad x^{P_2} = \begin{bmatrix} 0 \\ 1 - x_A \\ x_A \end{bmatrix} \text{ with } x_A \in [0,1] \quad (12)$$

The portfolios of $P_1$ and $P_2$ are $\lambda x^{P_1} + (1 - \lambda)x^{P_2}$ with $\lambda \in [0,1]$. The portfolio weight set of the shaded region (depicted in Figure 4) is

$$R = \{\lambda x^{P_1} + (1 - \lambda)x^{P_2} | \lambda \in [0,1], x_A \in [0,1]\}$$

We propose the following lemma to verify the argument:

**Lemma 3.4.** Set $R$ equals the $S$ of this example (i.e., $S = \{x \in \mathbb{R}^3 | 1^T x = 1, x \geq 0\}$). Thus, the shaded region is the $Z$ of this example.

**Proof.** First, we want to prove $R \subseteq S$. For all $x \in \mathbb{R}^3$, $1^T x = 1, x \geq 0$, we rewrite $x$ as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$ 

We want to obtain

$$\begin{bmatrix} \lambda x_A \\ 1 - x_A \\ (1 - \lambda)x_A \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 1 - x_1 - x_2 \end{bmatrix};$$

i.e.,

$$\begin{align*}
\lambda x_A &= x_1 & (13) \\
1 - x_A &= x_2 & (14) \\
(1 - \lambda)x_A &= 1 - x_1 - x_2 & (15)
\end{align*}$$

Figure 4. The $Z$ and $N$ of the example in subsection 3.2
We add (13) and (14) and get (15), so we drop (15) as redundant. By solving (13) and (14), we obtain \( x_A = 1 - x_2 \) and \( \lambda = \frac{x_1}{1-x_2} \) with \( x_2 < 1 \). Then, \( x_A \in [0, 1] \), because \( x_2 \in [0, 1] \); \( \lambda \geq 0 \), because \( x_1 \geq 0 \) and \( 1 - x_2 \geq 0 \); and \( \lambda \leq 1 \), because \( x_1 + x_2 \leq 1 \) and \( x_1 \leq 1 - x_2 \) and thus \( \frac{x_1}{1-x_2} \leq 1 \). With the obtained \( x_A \) and \( \lambda \),
\[
\mathbf{x} = \lambda \mathbf{p}_3 + (1 - \lambda) \mathbf{p}_4 \quad \text{with} \quad x_2 < 1.
\]
If \( x_2 = 1 \), \( \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \) with \( \lambda = \frac{1}{2} \) and \( x_A = 0 \). Thus, \( \mathbf{x} \in \mathbb{R}; \mathcal{S} \subset \mathbb{R} \).

Last, the lemma holds by \( \mathcal{R} \subset \mathcal{S} \) and \( \mathcal{S} \subset \mathcal{R} \).

With the \( \mathbf{Z} \), we readily locate the \( \mathcal{N} \) as a thick curve in the right part of Figure 4. The \( \mathcal{S} \) is identical to the \( \mathcal{S} \) (depicted in Figure 2) of the example in subsection 3.1; the \( \mathcal{E} \) of this example is a part of the linear segment from \( \mathbf{x}^B \) to \( \mathbf{x}^C \).

The meaningfulness of this example is as follows:

1. For the software of parametric quadratic programming, Optimizer can correctly solve this example, but CIOS can’t.
2. By the major style of portfolio optimization, we compute \( \mathbf{C} \) as the maximum expected-return portfolio and \( \mathbf{P}_3 \) and \( \mathbf{P}_4 \) as the mid portfolios. We depict \( \mathbf{D} \) and \( \mathbf{F} \) and \( \mathbf{P}_3 \) to \( \mathbf{P}_5 \) in Figure 5.
3. By the style, we try to compute the minimum variance portfolio by (6). Because the linear segment from \( \mathbf{D} \) to \( \mathbf{F} \) is correspondingly vertical and there is no further modeling for the minimum variance portfolio, we typically obtain \( \mathbf{P}_5 \) instead of \( \mathbf{D} \). We connect \( \mathbf{C}, \mathbf{P}_3, \mathbf{P}_4, \) and \( \mathbf{P}_5 \) to approximate the \( \mathcal{N} \) and mark the connection by a broken thick curve in Figure 5.
4. However, the style typically misses \( \mathbf{D} \) and consequently over-estimates the curve from \( \mathbf{D} \) to \( \mathbf{P}_3 \).

We generalize \( \mathbf{\mu} \) and \( \mathbf{\Sigma} \) by \( \mathbf{\mu} = \begin{bmatrix} \mu_A \\ \mu_B \\ \mu_C \end{bmatrix} \) with \( \mu_A < \mu_B < \mu_C \) and by \( \mathbf{\Sigma} = \begin{bmatrix} \sigma_A^2 & 0 & \sigma_A^2 \\ 0 & \sigma_B^2 & 0 \\ \sigma_A^2 & 0 & \sigma_C^2 \end{bmatrix} \) with \( \sigma_A < \sigma_B \); except the symmetric relationship, this example still holds. To save space, we omit the description; we depict the generalization in Figure 6.
4. A multiple-objective portfolio selection model without multiple optima. By multiple-objective optimization, we extend (3) by proposing

\[
\begin{align*}
\min \{ z_1 = & x^T \Sigma x \} \\
\max \{ z_2 = & \mu^T x \} \\
\max \{ z_3 = & \ell^T x \} \\
s.t. \quad 1^T x = 1
\end{align*}
\]

where \( \ell \) is an \( n \)-vector for stock expected liquidity for illustrative purposes; [25] and [29] list candidates for this third objective function of (16); the other symbols have been described in (2) and (3). We make the following assumptions:

**Assumption 1.** Matrix \( \Sigma \) is invertible and thus positive definite.

**Assumption 2.** Vectors \( \mu \), \( \ell \), and \( 1 \) are linearly independent.

Assumptions 1 and 2 are not strict prerequisite, because

- by traditional index models in finance (e.g., Chapter 8 of [1]), \( \Sigma = \beta \Omega \beta^T + D \) where for \( n \) stocks and \( i \) indexes, \( \beta \) is an \( n \times i \) coefficient matrix; \( \Omega \) is an \( i \times i \) covariance matrix of the \( i \) indexes; and \( D \) is an \( n \times n \) diagonal matrix with positive diagonal elements. With \( \beta \Omega \beta^T \) as positive semidefinite and \( D \) as positive definite, \( \Sigma \) is positive definite.
- Supposing that \( \mu \) and \( \ell \) are linearly dependent, we can drop objective function \( z_3 = \ell^T x \) since \( z_3 = \ell^T x \) and \( z_2 = \mu^T x \) are basically identical.

We handle (16) by the following \( e \)-constraint approach:

\[
\begin{align*}
\min \{ z_1 = & x^T \Sigma x \} \\
s.t. \quad \mu^T x = & e_2 \\
\ell^T x = & e_3 \\
1^T x = & 1
\end{align*}
\]

where \( e_2 \) and \( e_3 \) are parameters. By Assumption 2, (17) is feasible; i.e., \( \forall e_2 \in \mathbb{R}, \forall e_3 \in \mathbb{R}, \{ x \in \mathbb{R}^n \mid \mu^T x = e_2, \ell^T x = e_3, 1^T x = 1 \} \neq \emptyset \). The criterion vector
We premultiply (18) by \( \Sigma^{-1} \), obtain \( x = \frac{1}{2}(g_2 \Sigma^{-1} \mu + g_3 \Sigma^{-1} \ell + g_4 \Sigma^{-1} 1) \), and substitute this expression of \( x \) into (19)-(21) as follows:

\[
g_2 \mu^T \Sigma^{-1} \mu + g_3 \mu^T \Sigma^{-1} \ell + g_4 \mu^T \Sigma^{-1} 1 = 2e_2
\]

\[
g_2 \mu^T \Sigma^{-1} \ell + g_3 \ell^T \Sigma^{-1} \ell + g_4 \ell^T \Sigma^{-1} 1 = 2e_3
\]

\[
g_2 \mu^T \Sigma^{-1} 1 + g_3 \ell^T \Sigma^{-1} 1 + g_4 1^T \Sigma^{-1} 1 = 2
\]

We introduce notation \( C = \begin{bmatrix} \mu^T \Sigma^{-1} \mu & \mu^T \Sigma^{-1} \ell & \mu^T \Sigma^{-1} 1 \\ \mu^T \Sigma^{-1} \ell & \ell^T \Sigma^{-1} \ell & \ell^T \Sigma^{-1} 1 \\ \mu^T \Sigma^{-1} 1 & \ell^T \Sigma^{-1} 1 & 1^T \Sigma^{-1} 1 \end{bmatrix} \) and reexpress the three equations above as

\[
C \begin{bmatrix} g_2 \\ g_3 \\ g_4 \end{bmatrix} = \begin{bmatrix} 2e_2 \\ 2e_3 \\ 2 \end{bmatrix}
\]

[21] prove \( C \) as invertible. We respectively compute the determinant and inverse of \( C \) as \( |C| = adf + 2be - ace - bcf - ccd > 0 \) and \( C^{-1} = \frac{1}{|C|} \begin{bmatrix} df - ce & cc - bf & be - cd \\ ce - bf & af - cc & bc - ae \\ be - cd & bc - ae & ad - bb \end{bmatrix} \).

We premultiply (22) by \( C^{-1} \) and obtain

\[
\begin{bmatrix} g_2 \\ g_3 \\ g_4 \end{bmatrix} = C^{-1} \begin{bmatrix} 2e_2 \\ 2e_3 \\ 2 \end{bmatrix} = \frac{2}{|C|} \begin{bmatrix} \frac{1}{2}(e_2(df - ce) + e_3(ce - bf) + (be - cd)) \\ \frac{1}{2}(e_2(ce - bf) + e_3(af - cc) + (bc - ae)) \\ \frac{1}{2}(e_2(be - cd) + e_3(bc - ae) + (ad - bb)) \end{bmatrix}
\]
We substitute this \[
\begin{bmatrix}
g_2 \\
g_3 \\
g_4
\end{bmatrix}
\] into the previously derived \( x = \frac{1}{2} (g_2 \Sigma^{-1} \mu + g_3 \Sigma^{-1} \ell + g_4 \Sigma^{-1} 1) \), and obtain
\[
x = \frac{1}{|C|} [(e_2(df - ee) + e_3(ce - bf) + (be - cd)) \Sigma^{-1} \mu \\
+ (e_2(ce - bf) + e_3(af - cc) + (bc - ae)) \Sigma^{-1} \ell \\
+ (e_2(be - cd) + e_3(bc - ae) + (ad - bb)) \Sigma^{-1} 1]
\] (23)

We demonstrate the single optimality of \( x \) in the following theorem:

**Theorem 4.1.** The \( x \) of (23) is the only optimal solution (i.e., no multiple optima) of (17); therefore, (16) doesn’t have multiple efficient solutions.

**Proof.** (17) is a convex optimization problem as defined by [2] (pp.136&137); the optima exist because (17) has a lower bound 0. For (17), the method of Lagrange multipliers locates all the extrema including all the locally optima and all the (globally) optima. [2] (p.138) document that any locally optima of convex optimization problems are also (globally) optimal. Consequently, the method locates all the optima of (17). Following the method from (18) to (23), we find only one solution of the system of linear equations (18)-(21); the solution (23) is the only optimal solution of (17) and there is no multiple optima.

Suppose that (16) has multiple efficient solutions (defined in Section 2); i.e., there exist \( x^1 \in E \) and \( x^2 \in E \) with \( x^1 \neq x^2 \) and with \( z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \) as the common criterion vector of \( x^1 \) and \( x^2 \). Then, we take \( e_2 = z_2 \) and \( e_3 = z_3 \) for (17); \( x^1 \) and \( x^2 \) are both the optimal solutions of (17) but \( x^1 \neq x^2 \). Therefore, the status of the two optimal solutions contradicts the conclusion of the only optimal solution of (17) in the previous paragraph. The supposition is incorrect; (16) doesn’t have multiple efficient solutions.

**5. Conclusions and future directions.** The model of [16] has been criticized in the following aspects: about its applicability by, e.g., [5]; about its formulation especially constraints by, e.g., [4]; and about its input parameters by, e.g., [6]. However, the model of [16] is still widely described in classic textbooks (e.g., those of [1] and [14]); moreover, [15], [10], and the scholars for Issue 2 Volume 234 of European Journal of Operational Research refine the model.

Since subprime mortgage crisis, in practice, finance industry, especially mutual fund industry as “the most important of these financial intermediaries” (as described by [1, p.92]), has been reexamining the theoretical underpinnings of financial application (including the model of [16]); in academia, scholars (e.g., the researchers for Issue 2 Volume 234 of European Journal of Operational Research, for Harry Markowitz’s contribution to portfolio theory and operations research) have been showing a renewed interest in portfolio theory. Under such circumstances, this paper can be useful for enhancing the robustness of portfolio optimization.

We are in the early stage of analyzing the multiple optima of portfolio optimization. Therefore, when research is conducted, more questions may be raised than answered. In this regard, we envision the following research projects: 1. studying
the existence of the necessary condition of the multiple efficient solutions by analyzing the Karush-Kuhn-Tucker condition for parametric quadratic programming, and 2. computationally studying the structure of the nondominated sets of different portfolio selection models.

Researching multiple optima of portfolio optimization is challenging; we present the examples and demonstrate the meaningfulness. However, the researching is worthwhile, because portfolio optimization will then no longer be a one-size-fits-all affair after enough result is accumulated. Portfolio optimization will then have the ability to adapt itself to consider the unexpected and thus become a robust tool for investments.

Acknowledgments. The authors thank the editor and the anonymous referees for their careful reading of the paper and valuable comments.

REFERENCES

[1] Z. Bodie, A. Kane and A. J. Marcus, Investments, 10th edition, McGraw-Hill/Irwin, Boston, 2013.
[2] S. P. Boyd and L. Vandenberghe, Convex Optimization, 1st edition, Cambridge University Press, Cambridge, Britain, 2004.
[3] P. J. Brockwell and R. A. Davis, Time Series: Theory and Methods, 1st edition, Springer Verlag, New York, 1987.
[4] V. DeMiguel, L. Garlappi, F. J. Nogales and R. Uppal, A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms, Management Science, 55 (2009), 798–812.
[5] V. DeMiguel, L. Garlappi and R. Uppal, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, Review of Financial Studies, 22 (2009), 1915–1953.
[6] D. Disatnik and S. Katz, Portfolio optimization using a block structure for the covariance matrix, Journal of Business Finance & Accounting, 39 (2012), 806–843.
[7] P. H. Dybvig, Short sales restrictions and kinks on the mean variance frontier, The Journal of Finance, 39 (1984), 239–244.
[8] M. Ehrgott, Multicriteria Optimization, Volume 491 of Lecture Notes in Economics and Mathematical Systems, 2nd edition, Springer Verlag, Berlin, 2005.
[9] E. J. Elton, M. J. Gruber, S. J. Brown and W. N. Goetzmann, Modern Portfolio Theory and Investment Analysis, 9th edition, John Wiley & Sons, New York, 2014.
[10] J. B. Guerard, Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques, Springer, New York, 2010.
[11] G. Hadley, Nonlinear and Dynamic Programming, Addison Wesley, Reading, Massachusetts, 1964.
[12] F. S. Hillier and G. J. Lieberman, Introduction to Operations Research, Third edition. Holden-Day, Inc., Oakland, Calif., 1980.
[13] M. Hirschberger, Y. Qi and R. E. Steuer, Large-scale MV efficient frontier computation via a procedure of parametric quadratic programming, European Journal of Operational Research, 204 (2010), 581–588.
[14] C. Huang and R. H. Litzenberger, Foundations for Financial Economics, Prentice Hall, Englewood Cliffs, New Jersey, 1988.
[15] B. I. Jacobs, K. N. Levy and H. M. Markowitz, Portfolio optimization with factors, scenarios, and realistic short positions, Operations Research, 53 (2005), 586–599.
[16] H. M. Markowitz, Portfolio selection, The Journal of Finance, 7 (1952), 77–91.
[17] H. M. Markowitz and A. F. Perold, Portfolio analysis with factors and scenarios, The Journal of Finance, 36 (1981), 871–877.
[18] H. M. Markowitz and G. P. Todd, Mean-Variance Analysis in Portfolio Choice and Capital Markets, Frank J. Fabozzi Associates, New Hope, Pennsylvania, 2000.
[19] R. C. Merton, An analytical derivation of the efficient portfolio frontier, Journal of Financial and Quantitative Analysis, 7 (1972), 1851–1872.
[20] K. Miettinen, Nonlinear Multiojective Optimization, Kluwer Academic Publishers, Boston, 1999.
[21] Y. Qi, R. E. Steuer and M. Wimmer, An analytical derivation of the efficient surface in portfolio selection with three criteria, *Annals of Operations Research*, 251 (2017), 161–177.

[22] M. Rubinstein, Markowitz’s “portfolio selection”: A fifty-year retrospective, *The Journal of Finance*, 57 (2002), 1041–1045.

[23] M. Stein, J. Branke and H. Schmeck, Efficient implementation of an active set algorithm for large-scale portfolio selection, *Computers & Operations Research*, 35 (2008), 3945–3961.

[24] R. E. Steuer, *Multiple Criteria Optimization: Theory, Computation, and Application*, John Wiley & Sons, New York, 1986.

[25] R. E. Steuer and P. Na, Multiple criteria decision making combined with finance: A categorized bibliography, *European Journal of Operational Research*, 150 (2003), 496–515.

[26] J. Vörös, The explicit derivation of the efficient portfolio frontier in the case of degeneracy and general singularity, *European Journal of Operational Research*, 32 (1987), 302–310.

[27] J. Vörös, J. Kriens and L. W. G. Strijbosch, A note on the kinks at the mean variance frontier, *European Journal of Operational Research*, 112 (1999), 236–239.

[28] P. Wolfe, The simplex method for quadratic programming, *Econometrica*, 27 (1959), 382–398.

[29] C. Zopounidis, E. Galariotis, M. Doumpos, S. Sarri and K. Andriosopoulos, Multiple criteria decision aiding for finance: An updated bibliographic survey, *European Journal of Operational Research*, 247 (2015), 339–348.

Received July 2016; revised November 2016.

E-mail address: yorkche@nankai.edu.cn
E-mail address: wang-zhihao@hotmail.com
E-mail address: zhangsu@nankai.edu.cn