Degeneracy Analysis of the Portfolio Optimization

by

Zhengzheng Zhou

An essay
presented to the University of Waterloo
in fulfilment of the
essay requirement for the degree of
Master of Mathematics
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2006

I hereby declare that I am the sole author of this essay.

I authorize the University of Waterloo to reproduce this essay by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Abstract

Best and Hlouskova [1] have developed an active set QP algorithm for solving the problem of maximizing an expected utility function of n assets with transaction cost. The algorithm requires the assumption that the optimal solution founded at the end of each iteration is non-degenerate. In the real world, this assumption is impractical. The starting point for the algorithm often is a degenerate point. In the essay, I will present a method for solving this problem. This method is based on breaking the tie using the primal-dual property of the optimization problem.

Acknowledgements

I would like to express my deep thanks to my supervisor, Dr. Michael Best. Without his continues guidance and support, I could not finish this essay. I would also like to thank the reader, Dr. Chek Beng Chua, for his comments and careful reading of the draft.

Thanks my parents for their love and continues encouragement. Although they were not in Canada when I was writing the thesis, I can always feel their support.

Last, I owe great thanks to my life partner, Yu Bai, for his love, encouragement, and being my company for countless good or bad days.

Contents

Chapter 1

Introduction

1.1 Portfolio Optimization Problem with Transaction Cost Model

We consider the problem[3][4][5]

$$\left.\begin{aligned} \text{minimize} \quad &: \quad -t(\mu' x + p' x^+ + q' x^-) + \tfrac{1}{2} x' \Sigma x \\ \text{subject to} \quad &: \quad l' x = 1 \\ & \quad\;\; x - x^+ + x^- = \hat{x} \\ & \quad\;\; 0 \le x^+ \le d \\ & \quad\;\; 0 \le x^- \le e \end{aligned}\right\} \tag{1.1}$$

where $x = (x_1, \ldots, x_n)', \mu = (\mu_1, \ldots, \mu_n)'$ and $\Sigma = [\sigma_{ij}]$, t is a scalar parameter.

Throughout this paper prime ($'$) denotes transposition. All vectors are column vectors unless primed. The notation z_i or $(z)_i$ will be used to denote the ith component of the vector z.

Let x_i denote the proportion of wealth to be invested in asset i, and let μ_i denote the expected return on asset i, $i = 1, \ldots, n$. Let σ_{ij} denote the covariance between assets i and j, $1 \le i, j \le n$. $\Sigma = [\sigma_{ij}]$ is called the covariance matrix for the assets and is symmetric and positive semidefinite. Throughout the essay we will make the stronger assumption that Σ is positive definite. In terms of x, the expected return of the portfolio, μ_p, and the variance of the portfolio, σ_p^2, are given by

$$\mu_p = \mu' x \quad \text{and} \quad \sigma_p^2 = x' \Sigma x$$

Let $l = (1, 1, \ldots, 1)'$; i.e., l is an n vector of ones. Since the components of x are proportions, they must sum to one; i.e., $l'x = 1$. The constraints $l'x = 1$ is usually called the *budget constraint.*

In the portfolio optimization setting, the goal is to choose a value for x which gives a large value for μ_p and a small value for σ_p^2. These two goals tend to be in conflict. By introducing the parameter t, the problem has been solved nicely. For $t \geq 0$, the parameter t balances how much weight is placed on the maximization of $\mu' x$ (expected return) and minimization of $x'\Sigma x$ (risk).

To make the portfolio model more practical, most modern portfolio analysts insist on provision for transaction costs. These are usually incurred relative to some target portfolio, say $\hat{x}$. This may be the presently held portfolio or some industry standard such as S&P500 in the US. The idea is that the portfolio model should include costs for buying or selling assets relative to the target portfolio. Let x^+ denote the vector of purchases and let x^- denote the vector of sales. Then the holdings vector x may be represented as

$$x = \hat{x} + x^+ - x^-, x^+ \geq 0, x^- \geq 0.$$

Let p and q denote the vectors of purchases and sales transactions costs. Having separated out the sales and purchases, the total transaction cost is

$$p'x^+ + q'x^-.$$

In practice, most money managers will not accept very large changes in their holdings. Changes can be controlled by introducing bounds on amount of assets purchased and sold. This constraints can be formulated by introducing the bound constraints. Let d and e denote upper bounds on the amount sold and purchased, respectively. Hence we have the model formulation as presented in (1.1).

1.2 Degeneracy and Its Appearance in Portfolio Optimization

Definition 1.1 *For a general convex quadratic programming problem (QP);*

$$min\{\, c'x + \frac{1}{2}x'Cx \,|\, Ax \,\leq\, b\} \tag{1.2}$$

where c, x are n-vectors, C is an (n, n) symmetric positive semi-definite matrix, A is an (m, n) matrix and b is an m-vector.

*The **feasible region** for (1.2) is the set of points $S = \{\, x \,|\, Ax \leq b \,\}$. A point x_0 is **feasible** for (1.2), if $x_0 \in S$ and **infeasible**, otherwise. Constraints i is **inactive** at x_0 if $a_i'x_0 < b_i$, **active** at x_0 if $a_i'x_0 = b_i$ and **violated** at x_0 if $a_i'x_0 > b_i$.*

Definition 1.2 *A feasible point x_0 is **degenerate** for problem (1.2), if it is feasible and if the gradients of constraints active at x_0 are linearly dependent.*

In practical applications, many problems of the form (1.1) will be solved each time with μ and Σ being updated in response to new data. Consider two consecutive time periods t_1 and t_2 with $t_2 > t_1$. Let μ_1, Σ_1 and $\hat{x}_1$ be the values for μ, Σ and $\hat{x}$ respectively for $t = t_1$. Similarly, let μ_2, Σ_2 and $\hat{x}_2$ be the data for $t = t_2$. If the time periods t_1 and t_2 are close, it is resonable to expect that Σ_1 and Σ_2 will be close as also will be μ_1 and μ_2. Let x_1^* denote the optimal solution for the time period 1 problem. At the end of time period 1, the investor's holdings will be allocated according to x_1^*. It is normal to set $\hat{x}_2 = x_1^*$ so that the transaction costs will be measured against the investor's present holdings ($t = t_1$). The transaction costs may result in many of the asset holdings for the optimal solution of time period 2 being identical to those of the optimal solution of time period 1, namely x_1^*. An active set QP algorithm for the solution of the time period 2 problem can be initiated using $\hat{x}_2 = x_1^*$ as the starting point for the time period 2 version of (1.1). However, a difficulty is that this point is degenerate as shown in the following lemma.

Lemma 1.3 *If the target point $\hat{x}$ satisfies $l'\hat{x} = 1$, then the point $(\hat{x}', (\hat{x}^+)', (\hat{x}^-)')'$ with $\hat{x}^+ = \hat{x}^- = 0$ is a degenerate point for (1.1).*

Proof

$\hat{x}$ is a valid solution to the equation $l'\hat{x} = 1$ by the assumption in the lemma. The constraint $x - x^+ + x^- = \hat{x}$ is also satisfied. So far we have $n + 1$ active constraints. Furthermore, the lower bounds of zero on $\hat{x}^+$ and $\hat{x}^-$ are all active. In total, we have $3n + 1$ active constraints. Problem (1.1) only has $3n$ variables and there are $3n + 1$ constraints active at the point $(\hat{x}', (\hat{x}^+)', (\hat{x}^-)')'$. This implies that the active constraints are linear dependant. Thus $(\hat{x}', (\hat{x}^+)', (\hat{x}^-)')'$ is a degenerate point.

□

Chapter 2

A Solution Algorithm for the Portfolio Model

2.1 Algorithm

A basic difficulty of incorporating transaction costs is that it triples the number of problem variables and requires the addition of $3n$ linear constraints. This gives an optimization problem which is considerably more time consuming to solve.

Best and Hlouskova have developed a method for solving the $3n$-dimensional problem by solving a sequence of n-dimensional optimization problems, which account for the transaction costs implicitly rather than explicitly. The method is based on deriving the optimality conditions for the higher dimensional problem solely in terms of lower dimensional quantities. Their method requires the solution of a number of n-dimensional problems without the additional linear constraints and thus with corresponding savings in computer time and storage. Their key idea is to treat the transaction costs implicitly rather than explicitly.

Problem (1.1) is a special case of the following $3n$-dimensional model problem.

$$\left.\begin{array}{rcl} \text{minimize} & : & f(x) + p(x^+) + q(x^-) \\ \text{subject to} & : & x - x^+ + x^- = \hat{x},\ Ax \leq b, \\ & & d \leq x \leq e, \\ & & x^+ \geq 0, x^- \geq 0, \end{array}\right\} \quad (2.1)$$

Where A is an (m, n) matrix, b is an m-vectors, x is an n-vector of asset holdings, d and e are n-vectors of lower and upper bounds on x, respectively, and $-f(x)$ is an expected utility function. The constraints $Ax \leq b$ represent general linear constraints on the assets holdings and the constraints $d \leq x \leq e$ impose explicit bounds on the asset holdings. $\hat{x}, x^+$ and x^- are as in (1.1). For $i = 1, \ldots, n$, the purchase cost for x_i^+ is given by $p_i(x_i^+)$ and the sales cost for x_i^- is given by $q_i(x_i^-)$. The total cost of the transaction is thus

$$p(x^+) + q(x^-),$$

where $p(x^+) = \sum_{i=1}^n p_i(x_i^+)$ and $q(x^-) = \sum_{i=1}^n q_i(x_i^-)$.

Define

$$S \equiv \{x \in \mathbb{R}^n \mid Ax \leq b,\ d \leq x \leq e\} \tag{2.2}$$

The algorithm will require the following two assumptions.

Assumption 2.1 $d < \hat{x} < e$

Assumption 2.2 *let* x, x^+*,and* $x^- \in \mathbb{R}^n$*.*

1. $f(x)$ *is a twice differentiable convex function;*
2. $p(x^+)$ *and* $q(x^-)$ *are separable functions; i.e.,* $p(x^+) = \Sigma_{i=1}^n p_i(x_i^+)$ *and* $q(x^-) = \Sigma_{i=1}^n q_i(x_i^-)$*;*
3. $p_i(x_i^+)$ *and* $q_i(x_i^-)$ *are convex functions of a single variable for* $i = 1, \ldots, n$*;*
4. p*,* q *are twice differentiable functions;*
5. $\nabla p(x^+) \geq 0$ *and* $\nabla q(x^-) \geq 0$*; i.e.,*$p_i(x_i^+)$ *and* $q_i(x_i^-)$ *are increasing functions for* $i = 1, \ldots, n$*.*

Definition 2.3 *Let* $x \in \mathbb{R}^n$*. The index sets of* x *are defined as follows*

$$\begin{aligned} I^+(x) &= \{\, i \mid x_i > \hat{x}_i\} \\ I(x) &= \{\, i \mid x_i = \hat{x}_i\} \\ I^-(x) &= \{\, i \mid x_i < \hat{x}_i\} \end{aligned}$$

Definition 2.4 *Let* $x \in \mathbb{R}^n$

(i) $x^+ \in \mathbb{R}^n$ *is the* ***positive portion of*** x *if*

$$x_i^+ = \begin{cases} x_i - \hat{x}_i & i \in I^+(x) \\ 0 & i \in I(x) \bigcup I^-(x) \end{cases}$$

(ii) $x^- \in \mathbb{R}^n$ *is the* ***negative portion of*** x *if*

$$x_i^- = \begin{cases} 0 & i \in I(x) \bigcup I^+(x) \\ \hat{x}_i - x_i & i \in I^-(x) \end{cases}$$

Note that if x^+ and x^- are the positive and negative portions of x, then $(x', (x^+)', (x^-)')'$ satisfies the constraints $x^+ \geq 0,\ x^- \geq 0$ and $x - x^+ + x^- = \hat{x}$

Definition 2.5 *If* $(x', (x^+)'(x^-)')'$ *is a feasible solution for (2.1) and satisfies the property that not both of* x_i^+, x_i^- *are strictly positive for* $i = 1, \ldots, n$ *then we call* $(x', (x^+)', (x^-)')'$ *is a* **proper feasible solution**. *Additionally, if* $(x', (x^+)'(x^-)')'$ *is a proper feasible solution, which is optimal for (2.1), then we call* $(x', (x^+)'(x^-)')'$ *a* **proper optimal solution**.

In the algorithm we are about to discuss, it searches for an optimal solution for (2.1) to a proper optimal solution.

The algorithm is formulated in terms of solving a sequence of sub-problems. The sub-problems to be solved depend on two n-vectors $\tilde{d}$ and $\tilde{e}$ as follows:

$$\textsc{sub}(\tilde{d}, \tilde{e}): \ \min \ \{f(x) + \tilde{c}(\tilde{d}, \tilde{e}, x) | Ax \leq b, \tilde{d} \leq x \leq \tilde{e}\}$$

The vectors $\tilde{d}$ and $\tilde{e}$ are to be specified. They will always satisfy:

$$\tilde{d}_i = d_i, \tilde{e}_i = \hat{x}_i, \quad \text{or} \quad \tilde{d}_i = \hat{x}_i, \tilde{e}_i = e_i, \quad \text{or} \quad \tilde{d}_i = \tilde{e}_i = \hat{x}_i \tag{2.3}$$

In addition, $\tilde{c}(\tilde{d}, \tilde{e}, x) = \sum_{i=1}^n \tilde{c}_i(\tilde{d}_i, \tilde{e}_i, x_i)$, where for $i = 1, \ldots, n$

$$\tilde{c}_i(\tilde{d}_i, \tilde{e}_i, x_i) = \begin{cases} p_i(x_i - \hat{x}_i), & \text{if } \tilde{d}_i = \hat{x}_i \text{ and } \tilde{e}_i = e_i, \\ 0 & \text{if } \tilde{d}_i = \tilde{e}_i = \hat{x}_i, \\ q_i(\hat{x}_i - x_i) & \text{if } \tilde{d}_i = d_i \text{ and } \tilde{e}_i = \hat{x}_i \end{cases} \tag{2.4}$$

The sub-problem SUB($\tilde{d}, \tilde{e}$) is an n-dimensional problem with linear constraints and a convex, twice differentiable,non-linear objective function. There are many algorithms with demonstrably rapid convergence rates to solve it. See, for example, [2].

Remark 2.1

1. The feasible region for SUB($\tilde{d}, \tilde{e}$) is a compact set. From Assumption (2.2)(1),(4), the objective function for SUB($\tilde{e}, \tilde{e}$) is continuous. These two facts imply the existence of a solution for SUB($\tilde{d}, \tilde{e}$).
2. Note that for (2.1) the upper and lower bounds on x and the continuity of $f(x)$(from Assumption 2.2(1)) imply $f(x)$ is bounded from below over the feasible region of (2.1). Furthermore, from Assumption 2.2(2),(5), $p(x^+) + q(x^-)$ is also bounded from below over the feasible region of (2.1). Therefore, the objective function for (2.1) is bounded from below over the feasible region of (2.1).

The solution method for (2.1) is solely in terms of n-dimensional quantities. This method treats the variable x^+ and x^-, the constraint $x - x^+ + x^- = \hat{x}$, as well as the constraints $x^+ \geq 0, x^- \geq 0$, *implicitly* rather than explicitly. At each iteration j, each x_i is restricted according to one of the possibilities: (i) $d_i \leq x_i \leq \hat{x}_i$, (ii) $x_i = \hat{x}_i$ or (iii) $\hat{x}_i \leq x_i \leq e_i$. The objective function for SUB($\tilde{d}, \tilde{e}$) is created by adding transaction cost terms according to (i)-(iii) as follows: In the case of (i) this term is $q_i(\hat{x}_i - x_i)$, which is the transaction cost for selling the amount $\hat{x}_i - x_i$ of asset i. In the case of (ii) this term is zero. In the case of (iii) this term is $p_i(x_i - \hat{x}_i)$, which is the transaction cost for buying the amount $x_i - \hat{x}_i$ of asset i.

At the jth iteration, SUB($\tilde{d}, \tilde{e}$) is solved to produce optimal solution x^{j+1} and multiplier u^{j+1} for the constraints $Ax \leq b$, then for each $i \in I(x^{j+1})$ the multiplier v_i^{j+1} and w_i^{j+1} can be calculated. If these are all non-negative, then $((x^{j+1})', (x^+)', (x^-)')$ is optimal for (2.1), where x^+ and x^- are the positive and negative portions of x^{j+1}, respectively.

Otherwise, suppose $v_{k_1}^{j+1}$ is the smallest of these multipliers. For the next iteration, the upper bound on x_{k_1} is changed to e_{k_1} and the lower bound on it is changed to $\hat{x}_{k_1}$. If $w_{k_2}^{j+1}$ is the smallest such multiplier, then for the next iteration, the lower bound on x_{k_2} is changed to d_{k_2} and the upper bound is changed to $\hat{x}_{k_2}$. The objective function value for (2.1) for the next iteration will be strictly less than the present one.

Next we give a detailed formulation of the algorithm.

ALGORITHM 2.1

Model Problem: Problem (2.1) under Assumptions 2.1, 2.2 and $S \neq \emptyset$, where S is given by (2.2).

Begin

Initialization:

Start with any $x^0 \in S$. Construct the initial bounds and $\tilde{d}^0$, $\tilde{e}^0$ as follows:

do for $i = 1, \ldots, n$

if $x_i^0 > \hat{x}_i$ set $\tilde{d}_i^0 = \hat{x}_i, \quad \tilde{e}_i^0 = e_i$

elseif $x_i^0 < \hat{x}_i$ set $\tilde{d}_i^0 = d_i, \quad \tilde{e}_i^0 = \hat{x}_i$

else set $\tilde{d}_i^0 = \tilde{e}_i^0 = \hat{x}_i$

endif

enddo

Set $j = 0$ and go to Step 1.

Step 1: Solution of Sub-problem

Solve SUB$(\tilde{d}^j, \tilde{e}^j)$ to obtain optimal solution x^{j+1} and the multiplier vector u^{j+1} for the constraints $Ax \leq b$. Go to Step 2.

Step 2: Update and Optimality Test

For $i \in I(x^{j+1})$ compute

$$\begin{aligned} v_i^{j+1} &= (\nabla f(x^{j+1}) + A'u^{j+1})_i + \frac{dp_i(0)}{dx_i^+}, \\ w_i^{j+1} &= -(\nabla f(x^{j+1}) + A'u^{j+1})_i + \frac{dq_i(0)}{dx_i^-}. \end{aligned}$$

Further, compute k_1 and k_2 such that

$$v_{k_1}^{j+1} = \min \{ v_i^{j+1} \mid i \in I(x^{j+1}) \},$$

$$w_{k_2}^{j+1} = \min \{ w_i^{j+1} \mid i \in I(x^{j+1}) \}.$$

if $v_{k_1}^{j+1} \geq 0$ and $w_{k_2}^{j+1} \geq 0$ then STOP with a proper optimal solution $((x^{j+1})', (x^{j+1,+})', (x^{j+1,-})')'$ for (2.1), where $x^{j+1,+}$ and $x^{j+1,-}$ are positive and negative portions of x^{j+1} with respect to $\hat{x}$, respectively.

elseif $v_{k_1}^{j+1} \leq w_{k_2}^{j+1}$ then

$$\tilde{d}_i^{j+1} = \begin{cases} \tilde{d}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}), \end{cases}$$

$$\tilde{e}_i^{j+1} = \begin{cases} \tilde{e}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}) - \{k_1\}, \\ e_{k_1}, & i = k_1, \end{cases}$$

replace j with $j+1$ and go to Step 1.

else

$$\tilde{d}_i^{j+1} = \begin{cases} \tilde{d}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}) - \{k_2\}, \\ d_{k_2}, & i = k_2, \end{cases}$$

$$\tilde{e}_i^{j+1} = \begin{cases} \tilde{e}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}), \end{cases}$$

replace j with $j+1$ and go to Step 1.

endif

End

Remark 2.2:

Consecutive sub-problems differ in that one or more pairs of bounds have been replaced by others and corresponding changes have been made in the convex separable part of the objective function. Furthermore, the optimal solution for SUB($\tilde{d}^j, \tilde{e}^j$) is feasible for SUB($\tilde{d}^{j+1}, \tilde{e}^{j+1}$) and may be used as a starting point for it. Thus, if $S \neq \emptyset$ then the feasible region of any sub-problem solved by the algorithm is non-empty.

2.2 Finite Termination of the Algorithm

First let's consider the problem

$$\min \; \{f(x) | a_i' x \leq b_i, i = 1, \ldots, m-1, a_m' x = b_m\} \tag{2.5}$$

where $f(x)$ is any twice differentiable convex function, $a_1, \ldots, a_m$ are n-vectors and $b_1, \ldots, b_m$ are scalars. Suppose x^* is an optimal solution for (2.5).Then KKT's for (2.5) assert that there exists an m-vector $u = (u_1, \ldots, u_m)'$ such that

$$\left. \begin{array}{rcll} a_i' x^* & \leq & b, a_m' x^* = b_m, , & i = 1, \ldots, m-1 \\ -\nabla f(x^*) & = & u_1 a_1 + \cdots + u_m a_m, u_i \geq 0 & i = 1, \ldots, m-1 \\ u_i(a_i' x^* - b_i) & = & 0, & i = 1, \ldots, m-1 \end{array} \right\} \tag{2.6}$$

Without loss of generality assume that the first $k-1$ inequality constraints are active at x^* and denote

$$R \equiv \{x | a_i' x = b_i, i = 1, \ldots, k-1, a_i' x \le b_i, i = k, \ldots, m\} \tag{2.7}$$

where $k - 1 < m$.

The following is taken from [1], Lemma (5.1) in Appendix.

Lemma 2.6 *Let R be defined by (2.7), x^* be a non-degenerate optimal solution for (2.7), u_m be the multiplier for constraint $m, u_m < 0$ and $a_i' x^* = b_i$ for $i = 1, \ldots, k-1$. Then there exist points $\tilde{x} \in R$ for which $f(\tilde{x}) < f(x^*)$.*

Proof

Because x^* is non-degenerate, $a_1, \ldots, a_{k-1}$ and a_m are linearly independent. Let $d_{k+1}, \ldots, d_n$ be any vectors such that

$$D' = [\, a_1, \ldots, a_{k-1}, a_m, d_{k+1}, \ldots, d_n]$$

is nonsingular. Let

$$D^{-1} = [\, c_1, \ldots, c_k, c_{k+1}, \ldots, c_n]$$

where c_i, denotes the ith column of D^{-1} for $i = 1, \cdots, n$. Let $s = c_k$. By definition of the inverse matrix,

$$a_i' s = 0, i = 1, \ldots, k-1 \tag{2.8}$$

and

$$a_m' s = 1 \tag{2.9}$$

Consider points of the form $x^* - \sigma s$, where σ is a non-negative scalar. From (2.8)

$$a_i'(x^* - \sigma s) = a_i' x^* = b_i, i = 1, \cdots, k-1$$

so the first $k-1$ constraints remain active at $x^* - \sigma s$ for all $\sigma \ge 0$. Furthermore, from (2.9),

$$a_m'(x^* - \sigma s) = b_m - \sigma < b_m, \quad \text{for} \quad \sigma > 0$$

Thus, constraint m becomes inactive at $x^* - \sigma s$ for all strictly positive σ. Since constraints $k, \ldots, m-1$ are inactive at x^*, this implies that

$$x^* - \sigma s \in R \quad \text{for all positive sufficiently small } \sigma \tag{2.10}$$

From (2.6) and the fact that constraints $k, \ldots, m-1$ are inactive at x^* we have

$$-\nabla f(x^*) = u_1 a_1 + \cdots + u_{k-1} a_{k-1} + u_m a_m$$

From this, (2.8), (2.9) and the hypothesis that $u_m < 0$, it follows that

$$\nabla f(x^*)' s = -u_m > 0 \tag{2.11}$$

From Taylor's series

$$f(x^* - \sigma s) = f(x^*) - \sigma \nabla f(x^*)' s + \frac{1}{2} \sigma^2 s' H(\xi) s, \tag{2.12}$$

where H denotes the Hessian of $f(x)$ and ξ is on the line segment joining x^* and $x^* - \sigma s$. It now follows from (2.10)- (2.12) that $f(x^* - \sigma s) < f(x^*)$ for all positive sufficiently small σ. This completes the proof of the lemma. □

The finite termination property of the algorithm is established in the following theorem.

Theorem 2.7 *Let assumption (2.1),(2.2) be satisfied and let $S \neq \emptyset$, where S is given by (2.2). Begin with any $x^0 \in S$, let the algorithm be applied to (2.1) and let $x^1, x^2, \ldots, x^j, \ldots$ be the points so obtained. Let $x^{j,+}$ and $x^{j,-}$ be the positive and negative portion of x^j, respectively. Assume each $((x^j)', (x^{j,+})', (x^{j,-})')'$ is non-degenerate. Then $G(x^{j+1}, x^{j+1,+}, x^{j+1,-}) < G(x^j, x^{j,+}, x^{j,-})$ for $j = 1, 2, \ldots,$ where $G(x, x^+, x^-)$ is the objective function of (2.1) and for some k, $((x^k)', (x^{k,+})', (x^{k,-})')'$ is a proper optimal solution for (2.1).*

The finite termination of the algorithm relies on the assumption that the optimal solution for the SUB problem is non-degenerate. Practically, this assumption is not always true. The most trivial degenerate point $\hat{x}$ does appear as the optimal solution for the SUB problem. In the next chapter, we will discuss in great detail how to overcome this difficulty.

Chapter 3

Solution for the Degeneracy

3.1 Optimality Condition

In the end of the last chapter, we mentioned that it's not unreasonable to have the target point as the optimal solution for the subproblem in the BH method. Whenever this happens, there is no assurance that the BH method will terminate in finite steps since the assumptions in theorem (2.7) won't hold anymore.

Definition 3.1 *The point x_0 is a **quasi-stationary point** for (1.1) if*

(1) x_0 is in the feasible region of (1.1).

(2) x_0 is optimal for min $\{ -t(\mu'x) + p'x^+ + q'x^- + \frac{1}{2}x'\Sigma x \mid a_i'x = b_i, \quad all \;\; i \in I(x_0)\}$

Lemma 3.2 *The feasible point formed by target point $(\hat{x}', 0', 0')'$ is a quasi-stationary point for (1.1).*

Proof
The feasible region for problem

$$\min \{ -t(\mu'x) + p'x^+ + q'x^- + \frac{1}{2}x'\Sigma x \mid a_i'x = b_i, \quad \text{all} \;\; i \in I((\hat{x}', 0, 0)')\} \tag{3.1}$$

only has one feasible point, $(\hat{x}', 0', 0')'$. Therefore $(\hat{x}', 0', 0')'$ is optimal for (3.1). Hence, $(\hat{x}', 0', 0')'$ is a quasi-stationary point for (1.1). □

We also have shown that $(\hat{x}', 0', 0')'$ is a degenerate point in previous section. Hence, $(\hat{x}', 0', 0')'$ is a degenerate quasi-stationary point. If it were not degenerate, we could just compute the smallest multiplier for active constraints. If that multiplier were non-negative, the present solution is optimal and we are done. If it were negative, we could simply delete that constraint from the active set and carry on.

But the quasi-stationary point is degenerate and this causes problems. First observe that dropping a single bound constraint will not allow a positive move because the budget constraint would be violated. The equality constraints cannot be dropped because they are equality constraints. Therefore, we can't just delete the single constraint with the smallest negative multiplier. There are at least two bound constraints must be dropped; i.e., some component of x^+ and some component of x^- must be increased from 0. Although $\hat{x}$ is not necessary degenerate for SUB problem, $(\hat{x}', 0', 0')$ is degenerate for (1.1).

Let's see two examples first to demonstrate the degeneracy problems in the portfolio optimization.

In the both examples,

$$\mu = \begin{bmatrix} 1 \\ 1.1 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \hat{x} = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$$

$$t = 1 \quad d = e = \infty \quad \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

In the first example, the transaction cost vectors are

$$p = q = \begin{bmatrix} 10^5 \\ 10^5 \end{bmatrix}.$$

Hence the problem (1.1) becomes

Example 3.1

$$\left.\begin{array}{rl} \text{minimize} & : \; -(x_1 + 1.1x_2) + 10^5(x_1^+ + x_2^+ + x_1^- + x_2^+) + \frac{1}{2}(x_1^2 + x_2^2) \\ \text{subject to} & : \; x_1 + x_2 = 1 \\ & \;\; x_1 - x_1^+ + x_1^- = 0.1 \\ & \;\; x_2 - x_2^+ + x_2^- = 0.9 \\ & \;\; x_1^+ \geq 0, x_1^- \geq 0 \\ & \;\; x_2^+ \geq 0, x_2^- \geq 0 \end{array}\right\} \quad (3.2)$$

The transaction costs for selling and buying are very big in this example. Selling or buying α percent of the current holdings will cost the investor \$$(\alpha \times 10^5)$. If the gain from the changing of the holdings can not compensate the cost for the transaction fee, it is unwise for the investor to make any transactions. By observation, the transaction costs will be too big to be compensated by any gain from the change of holdings. Hence, the best action is not to sell or buy any assets. Thus, we guess the optimal solution will be $(\hat{x}', 0', 0')'$.

Lemma 3.3 *The degenerate point $(\hat{x}', 0', 0')'$ is the optimal solution for problem (3.2).*

Proof
Let $(x_0', (x_0^+)', (x_0^-)')' = (\hat{x}', 0', 0')' + d$ be a feasible solution to the problem. In order to satisfy the budget constraint, we must have either asset 1 increase and asset 2 decrease or asset 1 decrease or asset 2 increase. First, let's can assume that

$$x_1^- = x_2^+ = 0. \quad (3.3)$$

Then we must have

$$x_1^+ = x_2^- = \triangle d \neq 0. \quad (3.4)$$

Thus

$$d = [\, \triangle d \quad -\triangle d \quad \triangle d \quad 0 \quad 0 \quad \triangle d \,]'$$

and

$$\begin{array}{rcl} x_0' & = & [\, \hat{x}_1 + \triangle d, \quad \hat{x}_2 - \triangle d \,]' \\ (x_0^+)' & = & [\, \triangle d, \quad 0 \,]' \\ (x_0^-)' & = & [\, 0, \quad \triangle d \,]' \end{array}$$

Let $f(x)$ denotes the objective function. Then

$$\begin{aligned}
f(x_0, x_0^+, x_0^-) &= -((\hat{x}_1 + \triangle d) + 1.1(\hat{x}_2 - \triangle d)) + 10^5(\triangle d + 0 + 0 + \triangle d) + \frac{1}{2}((\hat{x}_1 + \triangle d)^2 + (\hat{x}_2 - \triangle d)^2) \\
&= -(\hat{x}_1 + 1.1\hat{x}_2) + \frac{1}{2}(\hat{x}_1^2 + \hat{x}_2^2) - \triangle d + 1.1\triangle d + 10^5(2\triangle d) + \frac{1}{2}(2\triangle d(\hat{x}_1 - \hat{x}_2) + 2\triangle d^2) \\
&= f(\hat{x}, 0, 0) + 0.1\triangle d + 2 \times 10^5 \triangle d + (\triangle d(\hat{x}_1 - \hat{x}_2) + \triangle d^2) \\
&\geq f(\hat{x}, 0, 0) + (2 \times 10^5 + 0.1)\triangle d + \triangle d(\hat{x}_1 - \hat{x}_2) \\
&= f(x, \hat{0}, 0) + (2 \times 10^5 + 0.1)\triangle d + \triangle d(1 - \hat{x}_2 - \hat{x}_2) \ \text{ since } \ \hat{x}_1 + \hat{x}_2 = 1 \\
&= f(\hat{x}, 0, 0) + (2 \times 10^5 + 0.1 + 1 - 2\hat{x}_2)\triangle d \\
&\geq f(\hat{x}, 0, 0) + (2 \times 10^5 + 0.1 + 1 - 2) \\
&\geq f(\hat{x}, 0, 0)
\end{aligned}$$

Therefore $(\hat{x}', 0', 0')'$ is the optimal solution for (3.2) under assumption (3.3).

Instead of having assumption (3.3), we assume that

$$x_1^+ = x_2^- = 0.$$

Under this assumption, the statement $f(x_0, x_0^+, x_0^-) \geq f(\hat{x}, 0, 0)$ is still true using the similar proof as before. Hence $(\hat{x}', 0', 0')'$ is optimal under this assumption.

In all, we have showed that $(\hat{x}', 0', 0')'$ is the optimal solution in both cases. We can conclude that the optimal solution for (3.2). □

The next example will show that the target point is not always the optimal solution. In this example

$$p = q = \begin{bmatrix} 10^{-5} \\ 10^{-5} \end{bmatrix}.$$

In this case, the transaction costs are very small. In most of the cases, the gain from the change of the current holdings will be much greater than the lose from the transaction costs. Hence, the investor will change the holdings if the resulting portfolio has the greater return than the present one. This time, the target portfolio may not be the optimal holdings anymore. The following demonstrates this idea.

The example problem will be (1.1) becomes

Example 3.2

$$\left.\begin{array}{lll} \text{minimize} & : & -(x_1 + 1.1x_2) + 10^{-5}(x_1^+ + x_2^+ + x_1^- + x_2^+) + \frac{1}{2}(x_1^2 + x_2^2) \\ \text{subject to} & : & x_1 + x_2 = 1 \\ & & x_1 - x_1^+ + x_1^- = 0.1 \\ & & x_2 - x_2^+ + x_2^- = 0.9 \\ & & x_1^+ \geq 0, x_1^- \geq 0 \\ & & x_2^+ \geq 0, x_2^- \geq 0 \end{array}\right\} \quad (3.5)$$

To show that the target point $\hat{x}$ is not the optimal solution, we only need to find a point with a smaller objective value. Consider point $x_0 = [0.5 \quad 0.5 \quad 0.4 \quad 0 \quad 0 \quad 0.4]'$. By direction calculation, we have

$$\begin{aligned} f(x_0) &= 0.5 + 1.1 \times 0.5 + 10^{-5} \times 0.4 + 10^{-5} \times 0.4 + \frac{1}{2}(0.5^2 + 0.5^2) \\ &= 1.05 + 0.8 \times 10^{-5} + 0.25 \\ &= 1.3 + 0.8 \times 10^{-5} \\ &< 1.51 = f(\hat{x}, 0, 0) \end{aligned}$$

This verifies that the $(\hat{x}', 0', 0')'$ is not the optimal solution for the problem (3.5).

Using these two examples, we have showed the degenerate point could either be the optimal solution to the problem or a non-optimal solution. Since the size of practical problems usually is very big, we can't use the analysis as we did in the previous example. We have developed a systematic approach to determine if the current degenerate point is the optimal solution.

To test for the optimality for the degenerate point x_0, the following lemmas will be used.

Lemma 3.4 *Let x_0 be a quasi-stationary point for*

$$min\ \{\ f(x) \mid Ax \leq b\ \} \quad (3.6)$$

Let A_0 be a subset of the rows of A corresponding to those constraints active at x_0. If the LP

$$min\ \{\ \nabla f'(x_0)s \mid A_0 s \leq 0\ \} \quad (3.7)$$

has an optimal solution $s = 0$, then x_0 is optimal for (3.6). If $s = 0$ is not optimal, then (3.7) is unbounded from below and there exists an s such that $A(x_0 + s) \leq b$ and $\nabla f'(x_0)s < 0$.

Proof

Suppose that $s = 0$ is an optimal solution for (3.7) and x_0 is not the optimal solution for (3.6). Then there exists a solution $x' = x_0 + s_0$ such that

$$\begin{aligned} f(x_0) &> f(x'). && (3.8)\\ Ax' &\leq b && (3.9) \end{aligned}$$

Since $Ax_0 \leq b$ and $Ax' = A(x_0 + s_0) \leq b$,

$$As_0 \leq 0 \tag{3.10}$$

This implies that s_0 is feasible to problem (3.7).

By the Taylor Theorem, we know

$$f(x') = f(x_0) + \nabla f(x_0)' s_0 + \frac{1}{2} s_0' Hf(x_0) s_0. \tag{3.11}$$

$f(x)$ is a convex function. This implies that

$$s_0' Hf(x_0) s_0 \geq 0. \tag{3.12}$$

Combine (3.8),(3.11) and (3.12), we have

$$\nabla f(x_0)' s_0 < 0.$$

This contradicts to the fact that $\nabla f(x_0)' s = 0$ is the optimal value for (3.7). Hence, our assumption that x_0 is not the optimal solution is not true. i.e., x_0 is the optimal solution for problem (3.6).

If $s = 0$ is not the optimal solution for (3.7), then there must exist another feasible solution $\tilde{s}$ such that

$$\nabla f'(x_0)\tilde{s} < \nabla f'(x_0) s = 0, \quad A\tilde{s} \leq 0. \tag{3.13}$$

Let $\hat{s} = \sigma \tilde{s}$, where $\sigma > 0$. Then

$$\begin{aligned} A_0 \hat{s} &= A_0(\sigma \tilde{s}) \\ &= \sigma (A_0 \tilde{s}) \\ &< 0 \text{ by (3.13)} \end{aligned}$$

Therefore, $\hat{s}$ is a feasible solution as well. Also,

$$\begin{aligned} \nabla f'(x_0)\hat{s} &= \nabla f'(x_0)(\sigma \tilde{s}) \\ &= \sigma(\nabla f'(x_0)\tilde{s}) \\ &\rightarrow -\infty \text{ as } \sigma \rightarrow \infty \end{aligned}$$

Hence, the problem (3.7) is unbounded from below.

Let $A = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}$ and $b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$, where A_0 are a subset of A corresponding to those constraints active at x_0. Then

$$A_0 x_0 = b_0 \ , A_1 x_0 < b_1.$$

let $s = -\sigma \tilde{s}$, where σ is positively sufficiently small. Then

$$\begin{aligned} A(x_0 + s) &= Ax_0 + As \\ &= \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} x_0 + \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} s \\ &= \begin{bmatrix} A_0 x_0 - \sigma A_0 \tilde{s} \\ A_1 x_0 - \sigma A_1 \tilde{s} \end{bmatrix} \\ &\leq \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}. \end{aligned}$$

Also, $\nabla f'(x_0)s = \sigma \nabla f'(x_0)\tilde{s} < 0$. In all, we have showed that there exists an s such that $A(x_0 + s) \leq b$ and $\nabla f'(x_0)s < 0$. □

The version of (3.7) for problem (1.1)is

$$\left.\begin{aligned} \text{minimize} &: \nabla f'(\hat{x})s + p's^+ + q's^- \\ \text{subject to} &: s - s^+ + s^- = 0 \\ & \quad l's = 0 \\ & \quad s^+ \geq 0, s^- \geq 0 \end{aligned}\right\} \tag{3.14}$$

Using the above lemma, we can check if the target solution is the optimal solution. One way of checking the optimality of $s = 0$ for (3.14) is to use the KKT conditions. This approach could be tedious if the size of the problem is very large, which is always the case in the real applications.

If $s = 0$ is not the optimal solution, then we need to find an s such that $As \leq 0$ and $\nabla f'(\hat{x})s < 0$.

The following lemmas provide an alternating way of checking optimality for (3.14) and find a search direction if $s = 0$ is not the optimal solution.

Lemma 3.5 *If there exist integers i and $j, i \neq j$ such that*

$$\nabla f'(\hat{x})_i + p_i < \nabla f'(\hat{x})_j - q_j. \tag{3.15}$$

Then (3.14) is unbounded from below. A search direction which causes this is

$$\begin{aligned} s &= (0, \ldots, s_i^+, 0, \ldots, -s_j^-, 0, \ldots, 0)', \\ s^+ &= (0, \ldots, 0, s_i^+, 0, \ldots, 0)', \\ s^- &= (0, \ldots, 0, s_j^-, 0, \ldots, 0)' \end{aligned} \tag{3.16}$$

with $s_i^+ = s_j^- = \alpha > 0$. Note s_i^+ is in the ith position of s^+ and s_j^- is in the jth position of s^-.

Proof
Suppose there exists integers i and j satisfy (3.15). Let s be the vector as defined in (3.16). Then

$$\begin{aligned} \nabla f'(\hat{x})s + p's + q's &= \nabla f'(x_0)_i s_i^+ - \nabla f'(x_0)_j s_j^- + p_i s_i^+ + q_j s_j^-, \\ &= \alpha(\nabla f'(x_0)_i + p_i - \nabla f'(x_0)_j + q_j) \\ &< 0 \end{aligned}$$

As $s_i = s_j = \alpha$ approaching infinity, $\nabla f'(\hat{x})s + p's + q's$ approaches negative infinity.

The search direction $(s', (s^+)', (s^-)')'$ is feasible for (3.14) since

$$\begin{aligned} l's &= s_i^+ - s_j^- + \sum_{k \neq i,j} s_k \\ &= \alpha - \alpha + \sum_{k \neq i,j} 0 \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} s - s^+ + s^- &= \sum_{k \neq i,j} (s_k - s_k^+ + s_k^-) + s_i^+ - s_i^+ - s_j^- + s_j^- \\ &= \sum_{k \neq i,j} 0 + 0 \\ &= 0. \end{aligned}$$

Thus, Problem (3.14) is unbounded from below. By lemma (3.4), we can use $(s', (s^+)', (s^-)')'$ as the search direction. Along this search direction, the objective function value of (1.1) is decreasing. Hence we can jump out of the degenerate point for problem (1.1). □

Remark 3.1

To find a pair of i and j satisfy (3.15), we just need to calculate $u = min\{\ \nabla f(x_0)_i + p_i \mid i = 1, \ldots, n\ \}$ and $v = max\{\ \nabla f(x_0)_j - q_j \mid j = 1, \ldots, n\ \}$. If $u < v$, then we have found a pair of i and j which satisfies (3.15). Otherwise, we can conclude there is no such pair can satisfy (3.15).

Lemma 3.6 *If there is no pair of i and j satisfies (3.15), then $\hat{x}$ is the optimal solution for problem (1.1). i.e. $\forall\ i, j \in \{1, \ldots, n\}$*

$$\nabla f(\hat{x})_i + p_i \geq \nabla f(\hat{x})_j - q_j. \tag{3.17}$$

Proof
We will proceed the proof with the dual of the problem. The Dual problem for (3.14) is the following:

$$\begin{aligned}
\text{maximize} \quad &: \quad 0'\mu && (3.18)\\
\text{subject to} \quad &: \quad \mu_1 + \mu_2 l = -\nabla f(\hat{x}) && (3.19)\\
& \quad -\mu_1 - \mu_3 = -p && (3.20)\\
& \quad \mu_1 - \mu_4 = -q && (3.21)\\
& \quad \mu_3 \geq 0 && (3.22)\\
& \quad \mu_4 \geq 0 && (3.23)
\end{aligned}$$

Since the objective function is a constant, any feasible solution will be the optimal solution for this optimization problem.

Suppose that μ_1, μ_2, μ_3 and μ_4 is a set of feasible solution for the above problem. Then using the above constraints, we can derive the following equations:

$$\begin{aligned}
(3.20) + (3.21) \quad &\Longrightarrow \quad (\mu_3)_i + (\mu_4)_i = p_i + q_i, && (3.24)\\
(3.19) + (3.20) \quad &\Longrightarrow \quad \mu_2 l' - \mu_3 = -\nabla f(\hat{x}) - p && (3.25)\\
(3.19) - (3.21) \quad &\Longrightarrow \quad \mu_2 l' - \mu_4 = -\nabla f(\hat{x}) + q && (3.26)
\end{aligned}$$

We know that there exits an extreme feasible point which is the optimal solution for the optimization problem. This point has at least $3n + 1$ active constraints since there are $3n + 1$ variables. There are $3n$ equations and $2n$ inequalities for the dual problem. Hence, we need to have at least one more equality to make a feasible point to be an extreme point. Therefore, it is either $(\mu_3)_i = 0$ or $(\mu_4)_i = 0$ for some $i \in \{1, \ldots, n\}$.

Without lose generosity, let $(\mu_3)_i = 0$. Use this in (3.25), we will have

$$(\mu_2 l')_i - 0 = -(\nabla f(\hat{x}) + p)_i.$$

Therefore

$$\mu_2 = -(\nabla f(\hat{x}) + p)_i. \tag{3.27}$$

Substitute μ_2 into (3.19), we have

$$\mu_1 - (\nabla f(\hat{x}) + p)_i l = -\nabla f(\hat{x})$$

Therefore

$$\mu_1 = -\nabla f(\hat{x}) + (\nabla f(\hat{x}) + p)_i l \tag{3.28}$$

Substitute μ_1, μ_2 into(3.20), we have

$$\nabla f(\hat{x}) - (\nabla f(\hat{x}) + p)_i l - \mu_3 = -p$$

Simplify further, we solved μ_3.

$$\mu_3 = \nabla f(\hat{x}) - (\nabla f(\hat{x}) + p)_i l + p \tag{3.29}$$

Substitute μ_1 into (3.21),

$$-\nabla f(\hat{x}) + (\nabla f(\hat{x}) + p)_i l - \mu_4 = -q$$

Solve above equation, we have

$$\mu_4 = -\nabla f(\hat{x}) + (\nabla f(\hat{x}) + p)_i l + q. \tag{3.30}$$

Now we only need to check if inequality constraints have been violated.

$$\begin{aligned} (\mu_3)_k &= \nabla f(\hat{x})_k - \nabla f(\hat{x})_i - p_i - p_k,\ k \in \{1, \ldots, n\} \\ &= (\nabla f(\hat{x}) + p)_k - (\nabla f(\hat{x}) + p)_i \end{aligned}$$

We will choose i with the following property so that $(\mu_3)_k \geq 0$.

$$i = \min \{\ (j\ |\nabla f(\hat{x}) + p)_j, j = 1, 2, \ldots n\ \}$$

Now we will check if $(\mu_4)_k \geq 0$, $\forall\ k \in \{1, \ldots, n\}$.

$$\begin{aligned} (\mu_4)_k &= -\nabla f(\hat{x})_k + \nabla f(\hat{x})_i + p_i + q_k, \\ &= (\nabla f(\hat{x}) + p)_i - (\nabla f(\hat{x}) - q)_k \\ &\geq 0 \text{ by } (3.17) \end{aligned}$$

Now we can conclude that $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ is a optimal solution of the dual problem. Since the objective function for the dual problem is constant (0), the optimal value for the dual problem is 0. By Strong Duality theorem, the optimal value for the primal problem is 0 as well. Hence $s = 0$ is the optimal solution for problem (3.14). Therefore $\hat{x}$ is the optimal solution for (1.1) by Lemma (3.5).

□

3.2 Degeneracy Algorithm

We will give a detailed formulation of the modified Best/Hlouskova algorithm, which resolves degeneracy as in Section 3.1.

ALGORITHM 3.1

Model Problem: Problem (1.1) and $S \neq \emptyset$, where S is the feasible region for (1.1).

Begin

Initialization:

Start with any $x^0 \in S$. Construct the initial bounds and $\tilde{d}^0$, $\tilde{e}^0$ as follows:

do for $i = 1, \ldots, n$

if $x_i^0 > \hat{x}_i$ set $\tilde{d}_i^0 = \hat{x}_i$, $\tilde{e}_i^0 = e_i$

elseif $x_i^0 < \hat{x}_i$ set $\tilde{d}_i^0 = d_i$, $\tilde{e}_i^0 = \hat{x}_i$

else set $\tilde{d}_i^0 = \tilde{e}_i^0 = \hat{x}_i$

endif

enddo

Set $j = 0$ and go to Step 1.

Step 1: Solution of Sub-problem

Solve SUB$(\tilde{d}^j, \tilde{e}^j)$ to obtain optimal solution x^{j+1} and the multiplier vector u^{j+1} for the constraints $Ax \leq b$.

if x^{j+1} is not a degenerate point for the original problem, Go to Step 2.

else Go to Step 3.

endif

Step 2: Update and Optimality Test

For $i \in I(x^{j+1})$ compute

$$\begin{aligned} v_i^{j+1} &= (\nabla f(x^{j+1}) + A' u^{j+1})_i + \frac{dp_i(0)}{dx_i^+}, \\ w_i^{j+1} &= -(\nabla f(x^{j+1}) + A' u^{j+1})_i + \frac{dq_i(0)}{dx_i^-}. \end{aligned}$$

Further, compute k_1 and k_2 such that

$$v_{k_1}^{j+1} = \min \{\, v_i^{j+1} \mid i \in I(x^{j+1}) \,\},$$

$$w_{k_2}^{j+1} = \min \{\, w_i^{j+1} \mid i \in I(x^{j+1}) \,\}.$$

if $v_{k_1}^{j+1} \geq 0$ and $w_{k_2}^{j+1} \geq 0$ then STOP with a proper optimal solution $((x^{j+1})', (x^{j+1,+})', (x^{j+1,-})')'$

for (1.1), where $x^{j+1,+}$ and $x^{j+1,-}$ are positive and negative portions of x^{j+1} with respect to $\hat{x}$, respectively.

elseif $v_{k_1}^{j+1} \leq w_{k_2}^{j+1}$ then

$$\tilde{d}_i^{j+1} = \begin{cases} \tilde{d}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}), \end{cases}$$

$$\tilde{e}_i^{j+1} = \begin{cases} \tilde{e}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}) - \{k_1\}, \\ e_{k_1}, & i = k_1, \end{cases}$$

replace j with $j+1$ and go to Step 1.

else

$$\tilde{d}_i^{j+1} = \begin{cases} \tilde{d}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}) - \{k_2\}, \\ d_{k_2}, & i = k_2, \end{cases}$$

$$\tilde{e}_i^{j+1} = \begin{cases} \tilde{e}_i^j, & i \in \{1, \ldots, n\} - I(x^{j+1}), \\ \hat{x}_i, & i \in I(x^{j+1}), \end{cases}$$

replace j with $j+1$ and go to Step 1.

endif

Step 3: Degeneracy Handling

Compute

$$k = \min \{ k \mid \nabla f'(x^{j+1})_k + p_k,\ k = 1, \ldots, n \}$$

and

$$l = \max \{ j \mid \nabla f'(x^{j+1})_j - q_j,\ j = 1, \ldots, n \}.$$

if $\nabla f'(x^{j+1})_k + p_k \geq \nabla f'(x^{j+1})_l - q_l$, Go to Step 3.1

else Go to Step 3.2

endif

Step 3.1 STOP output the optimal solution $((x^{j+1})', (x^{j+1,+})', (x^{j+1,-})')'$ for (1.1), where $x^{j+1,+}$ and $x^{j+1,-}$ are positive and negative portions of x^{j+1} with respect to $\hat{x}$, respectively.

Step 3.2

$$\tilde{d}_i^{j+1} = \begin{cases} \tilde{d}_i^j, & i \in \{1, \ldots, n\} - \{l\}, \\ d_i, & i = l, \end{cases}$$

$$\tilde{e}_i^{j+1} = \begin{cases} \tilde{e}_i^j, & i \in \{1, \ldots, n\} - \{k\} \\ e_k, & i = k \end{cases}$$

replace j with $j+1$ and go to Step 1.

End

The algorithm is demonstrated using Example (3.3) and (3.4), which in addition show the values of the objective function at each iteration.

Example 3.3 $f(x) = x_1 + 1.1x_2 + \frac{1}{2}(x_1^2 + x_2^2)$, $\hat{x} = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$, $p(x^+) = 10^5(x_1^+ + x_2^+)$, $q(x^-) = 10^5(x_1^- + x_2^-)$. The constraints for the problem are

$$x_1 + x_2 = 1$$

$$\begin{aligned} x_1 - x_1^+ + x_1^- &= \hat{x}_1 \\ x_2 - x_2^+ + x_2^- &= \hat{x}_2 \end{aligned}$$

Initialization: We choose $x^0 = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$ as the starting point for the sub-problem. Then $\tilde{d}^0 = \tilde{e}^0 = \hat{x}$ and thus $\tilde{c}^0(x) = 0$.

Step 1: SUB$(\tilde{d}^0, \tilde{e}^0)$ is precisely the problem:

$$\min\{\, x_1 + 1.1x_2 + \frac{1}{2}(x_1^2 + x_2^2) \,|\, x_1 + x_2 = 1,\, 0.1 \le x_1 \le 0.1,\, 0.9 \le x_2 \le 0.9 \,\},$$

which has the optimal solution $x^1 = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$. It is a degenerate point for the original problem. Therefore, we proceed to step 3.

Step 3: $\nabla f(x^1) = \begin{bmatrix} 1.1 \\ 2 \end{bmatrix}$, and $p = q = \begin{bmatrix} 10^5 \\ 10^5 \end{bmatrix}$. Therefore, $k = 1$ with $\nabla f(x^1)_1 + p_1 = 1.1 + 10^5$ and $l = 2$ with $\nabla f(x^1)_2 - q_2 = 2 - 10^5$. Since

$$\nabla f(x^1)_1 + p_1 > \nabla f(x^1)_2 - q_2,$$

we stop with the optimal solution $(x, x^+, x^-)' = (0.1, 0.9, 0, 0, 0, 0)'$

Example 3.4 We obtain a second example from Example 3.1 by leaving $f(x), \hat{x}, x^0, A, b, d$ and e unchanged but taking $p = q = \begin{bmatrix} 10^{-5} \\ 10^{-5} \end{bmatrix}$

Initialization: Then $\tilde{d}^0 = \tilde{e}^0 = \hat{x}$ and thus $\tilde{c}^0(x) = 0$.

Step 1: SUB$(\tilde{d}^0, \tilde{e}^0)$ is precisely the problem:

$$\min\{\, x_1 + 1.1x_2 + \frac{1}{2}(x_1^2 + x_2^2) \,|\, x_1 + x_2 = 1,\, 0.1 \le x_1 \le 0.1,\, 0.9 \le x_2 \le 0.9 \,\},$$

which has the optimal solution $x^1 = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$. It is a degenerate point for the original problem. Therefore, we proceed to step 3.

Step 3: $\nabla f(x^1) = \begin{bmatrix} 1.1 \\ 2 \end{bmatrix}$, and $p = q = \begin{bmatrix} 10^{-5} \\ 10^{-5} \end{bmatrix}$. Therefore, $k = 1$ with $\nabla f(x^1)_1 + p_1 = 1.1 + 10^{-5}$ and $l = 2$ with $\nabla f(x^1)_2 - q_2 = 2 - 10^{-5}$. Since

$$\nabla f(x^1)_1 + p_1 < \nabla f(x^1)_2 - q_2,$$

we have

$$\tilde{d}_i^{j+1} = \begin{cases} \tilde{d}_i^j = \hat{x}_i, & i = 1 \\ -\infty, & i = 2, \end{cases}$$

$$\tilde{e}_i^{j+1} = \begin{cases} \tilde{e}_i^j = \hat{x}_i, & i = 2 \\ \infty, & i = 1 \end{cases}$$

Step 1: $\mathrm{SUB}(\tilde{d}^0, \tilde{e}^0)$ is precisely the problem:

$$\min \{ x_1 + 1.1x_2 + \frac{1}{2}(x_1^2 + x_2^2) \,|\, x_1 + x_2 = 1,\, 0.1 \le x_1,\, x_2 \le 0.9 \},$$

which has the optimal solution $x^2 = \begin{bmatrix} 0.55 \\ 0.45 \end{bmatrix}$.

Step 2: $I(x^2) = \emptyset$, $u_1 = 1.55 - 10^{-5}, u_2 = u_3 = 10^{-5}$. Since $u_1 \ge 0, u_2 \ge 0$ and $u_3 \ge 0$, $((x^2)', (x^{2,+})', (x^{2,-})')'$ is optimal for (1.1) with the given data, where $x^{2,+} = \begin{bmatrix} 0.35 \\ 0 \end{bmatrix}$, $x^{2,-} = \begin{bmatrix} 0 \\ 0.35 \end{bmatrix}$. Furthermore, $G(x^2, x^{2,+}, x^{2,-}) = 1.29751$.

Theorem 3.7 *Begin with any feasible point for problem (1.1), let the algorithm 3.1 be applied to the problem and let* $x^1, x^2, \ldots, x^j, \ldots$ *be the points so obtained. Let* $x^{j,+}$ *and* $x^{j,-}$ *be the positive* *and negative portion of x^j, respectively. Then $G(x^{j+1}, x^{j+1,+}, x^{j+1,-}) < G(x^j, x^{j,+}, x^{j,-})$ for $j = 1, 2, \ldots$, where $G(x, x^+, x^-)$ is the objective function of (1.1) and for some k, $((x^k)', (x^{k,+})', (x^{k,-})')'$ is a proper optimal solution for (1.1).*

Proof

At the end of step 1 in iteration $j-1$, if the optimal point $((x^j)', (x^{j,+})', (x^{j,-})')'$ is not degenerate, we proceed to step 2 as in Algorithm 2.1. By Theorem (2.7), we either have that $((x^j)', (x^{j,+})', (x^{j,-})')'$ is the optimal solution for problem (1.1) or $G((x^{j+1})', (x^{j+1,+})', (x^{j+1,-})') < G((x^j)', (x^{j,+})', (x^{j,-})')$.

If $((x^j)', (x^{j,+})', (x^{j,-})')'$ is degenerate, then we proceed to step 3. If

$$\nabla f'(x_0)_k + p_k \geq \nabla f'(x_0)_j - q_j,$$

where k, j are defined as in Step 3 of the algorithm, then according to Lemma (3.6), $((x^j)', (x^{j,+})', (x^{j,-})')'$ is an optimal solution for (1.1). This is in consistent with step 3.1 in the Algorithm 3.1

If

$$\nabla f'(x_0)_k + p_k < \nabla f'(x_0)_j - q_j,$$

then according to Lemma (3.5), we can find a search direction s as defined in the lemma. Along this search direction, the objective function of the problem will decrease. And the search direction in step 3.2 is defined based on the Lemma (3.5). Hence

$$G((x^{j+1})', (x^{j+1,+})', (x^{j+1,-})') < G((x^j)', (x^{j,+})', (x^{j,-})').$$

□

Remark 3.2

In each iteration, the objective function value is strictly decreased and the problem is bounded from below over the feasible region (Remark 2.1). Hence there are only finitely many iterations and none of the sub-problems can be repeated. Thus, the algorithm terminates in a finite number of steps with an optimal solution for (1.1).

3.3 Generalization of Algorithm

The algorithm we have discussed in the previous section is for solving problem (1.1). And this problem is a special case for (2.1). With minor modification, Algorithm 3.1 can be applied to the following problem.

$$\left.\begin{array}{rcl} \text{minimize} & : & f(x) + p(x^+) + q(x^-) \\ \text{subject to} & : & l'x = 1 \\ & & x - x^+ + x^- = \hat{x} \\ & & 0 \le x^+ \le d \\ & & 0 \le x^- \le e \end{array}\right\} \tag{3.31}$$

where x is an n-vector of asset holdings, d and e are n-vectors of lower and upper bounds on x, respectively, and $-f(x)$ is an expected utility function. t is a scalar parameter.$\hat{x}, x^+$ and x^- are as in (1.1). For $i = 1, \ldots, n$, the purchase cost for x_i^+ is given by $p_i(x_i^+)$ and the sales cost for x_i^- is given by $q_i(x_i^-)$. The total cost of the transaction is thus

$$p(x^+) + q(x^-),$$

where $p(x^+) = \sum_{i=1}^n p_i(x_i^+)$ and $q(x^-) = \sum_{i=1}^n q_i(x_i^-)$.

In Algorithm 3.1, we search for a pair of i and $j, i \neq j$ such that

$$\nabla f'(x_0)_i + p_i < \nabla f'(x_0)_j - q_j.$$

Since the transaction cost functions in (3.31) are no longer linear functions, we search for a pair of i and $j, i \neq j$ such that

$$\nabla f'(x_0)_i + \nabla(p_i(x_i^+)) < \nabla f'(x_0)_j - \nabla(q_i(x_i^-)). \tag{3.32}$$

Lemma 3.8 *Modifying the Algorithm 3.1 according to (3.32), the new algorithm is a solution algorithm for problem (3.31).*

Proof

The accuracy of Algorithm 3.1 is based on Lemma (3.4), (3.5) and (3.6). The convexity of the objective function is untouched by the modification to the problem since both $p(x^+)$, $q(x^+)$ are still convex function as before. In the proof of lemma (3.5) and (3.6), p_i and q_j are only be used as scalar for comparison. The linear property of the transaction cost function is not significant to the structure of the proofs. By substituting the scalar p_i with $\nabla(p_i(x_i^+))$ and q_j with $\nabla(q_j(x_j^-))$ will not change the structure of the dual problem which is defined in the proof of the lemmas. This is due to the constraints for the main problem has not been modified. The inequalities derived from the dual problem also won't be affected either. Hence, the conclusion drawn from those inequalities still applies to problem (3.31). Therefore, the modified the Algorithm 3.1 is a solution algorithm for problem (3.31). □

3.4 Conclusion

We have consider the problem of maximizing an expected utility function of n assess with the occurrence of degeneracy. This work is based on the existing solution algorithm developed by Best and Hlouskova. In order to jump out the degenerate point, we have developed the solution algorithm using primal and dual property. The method finds a search direction without involving solving linear system of equations. The easy degeneracy handling steps came from the simplicity of the constraint functions defined in the problem.

Bibliography

[1] M. J. Best and J. Hlouskova "An algorithm for portfolio optimization with transaction costs" *Management Science*, Vol 51, No 11, November 2005 pp 1676-1688.

[2] M. J. Best and K. Ritter " A class of accelerated conjugate direction methods for linearly constrained minimizationproblems" *Mathematics of Computation*,30/135(1976) 478-504

[3] Z. Bodie, A. Kane and A. J. Marcus *Investments*, Irwin McGraw-Hill, 4th edition, 1999

[4] D. G. Luenberger *Investment Science*, Oxford University Press, New York, 1998

[5] H. M. Markowitz *Portfolio Selection: Efficient Diversification of Investment*, Johm Wiley, New York, Yake Univeristy Press, New Haven, 1959