Industrial and Systems Engineering

Multiobjective optimization via parametric optimization: models, algorithms and applications

Tamás Terlaky
Lehigh University

Oleksandr Romanko
McMaster University

Alireza Ghaffari-Hadigheh
Azarbaijan University of Tarbiat Moallem

Report: 10T-007

Annals of Operations Research manuscript No.
(will be inserted by the editor)

Multiobjective optimization via parametric optimization: models, algorithms and applications

Oleksandr Romanko · Alireza Ghaffari-Hadigheh · Tamás Terlaky

Received: 15 July 2010 / Accepted: date

Abstract In this paper we highlight the relationships between multiobjective optimization and parametric optimization that is used to solve such problems. Solution of a multiobjective problem is the set of Pareto efficient points, known in the literature as Pareto efficient frontier or Pareto front. Pareto points can be obtained by using either weighting the objectives or by ε-constrained (hierarchical) method for solving multiobjective optimization models. Using those methods we can formulate them as parametric optimization problems and compute their efficient solution set numerically. We present a methodology for conic quadratic optimization that allows tracing the Pareto efficient frontier without discretization of the objective space and without solving the corresponding optimization problem at each discretization point.

Keywords Multiobjective optimization · Parametric optimization · Conic quadratic optimization · Pareto front · Efficient frontier · Financial optimization

1 Introduction

Multicriteria decision making or *multicriteria analysis* is a complex process of finding the best compromise among alternative decisions. A decision maker first describes the problem based on relevant assumptions about the real world problem. After that, alternative decisions are generated and evaluated. Optimization serves as a tool for solving multicriteria analysis problems when those problems are formulated as *multiobjective optimization* problems.

Oleksandr Romanko
Department of Computing and Software, McMaster University,
1280 Main Street West, Hamilton, Ontario, L8S4K1, Canada
E-mail: romanko@mcmaster.ca

Alireza Ghaffari-Hadigheh
Department of Mathematics, Azarbaijan University of Tarbiat Moallem,
35 Kms Tabriz Maragheh Road, Tabriz, East Azarbayjan, Iran
E-mail: hadigheha@azaruniv.edu

Tamás Terlaky
Department of Industrial and Systems Engineering, Lehigh University,
Harold S. Mohler Laboratory, 200 West Packer Avenue, Bethlehem, PA, USA
E-mail: terlaky@lehigh.edu

Classical meaning of the word "optimization" refers to single-objective optimization, which is a technique used for searching extremum of a function. This term generally refers to mathematical problems where the goal is to minimize (maximize) an objective function subject to some constraints. Depending on the nature and the form of the objective function and the constraints, continuous optimization problems are classified to linear, quadratic, conic and general nonlinear optimization problems.

Linear optimization (LO) is a highly successful operations research model. Therefore, it was natural to generalize the LO model to handle more general nonlinear relationships. However, this gives rise to many difficulties such as lack of strong duality, possible non-convexity and consequently problems with global versus local optimums, lack of efficient algorithms and software, etc.

In the recent decade, a new class of convex optimization models that deals with the problem of minimizing a linear function subject to an affine set intersected with a convex cone has appeared. It is known as conic optimization. Although the conic optimization model seems to be restrictive, any convex optimization problem can be cast as a conic optimization model and there are efficient solution algorithms for many classes of conic models such as conic quadratic optimization (CQO) and conic linear optimization (CLO). While it sounds counterintuitive, CQO is a sub-class of CLO. Conic optimization has many interesting applications in engineering, image processing, finance, economics, combinatorial optimization, etc.

Conic linear optimization (CLO) is the extension of LO to the classes of problems involving more general cones than the positive orthant. As our results heavily rely on duality theory (for review of the topic consult [?]), we present both primal and dual formulations of problems belonging to the CLO class. General form of CLO problem is:

$$\begin{array}{cc} \textit{Primal problem} & \textit{Dual problem} \\ \min_{x}\{c^T x \,:\, Ax = b,\, x \in \mathscr{K}\} & \max_{y,s}\{b^T y \,:\, A^T y + s = c,\, s \in \mathscr{K}^*\}, \end{array} \qquad (1)$$

where $\mathscr{K} \in \mathbb{R}^n$ is a closed, convex, pointed and solid cone, $\mathscr{K}^* = \{s \in \mathbb{R}^n \,:\, s^T x \geq 0 \ \forall x \in \mathscr{K}\}$ is the dual cone of $\mathscr{K}$; $A \in \mathbb{R}^{m\times n}$, $\operatorname{rank}(A) = m$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ are fixed data; $x, s \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ are unknown vectors. Often $x \in \mathscr{K}$ is also denoted as $x \geq_{\mathscr{K}} 0$. Moreover, $x \geq_{\mathscr{K}} y$ means $x - y \geq_{\mathscr{K}} 0$.

Examples of pointed convex closed cones include:

- the nonnegative orthant:

$$\mathbb{R}^n_+ = \mathscr{K}_\ell = \{x \in \mathbb{R}^n \,:\, x \geq 0\},$$

- the quadratic cone (also know as Lorentz cone, second order cone or ice-cream cone):

$$L^n = \mathscr{K}_q = \{x \in \mathbb{R}^n \,:\, x_1 \geq \|x_{2:n}\|\},$$

- the semidefinite cone:

$$S^n_+ = \mathscr{K}_s = \{X \in \mathbb{R}^{n\times n} \,:\, X = X^T,\, X \succeq 0\},$$

- any linear transformation and finite direct product of such cones.

Each of the three "standard" cones $\mathscr{K}_\ell$, $\mathscr{K}_q$ and $\mathscr{K}_s$ are closed, convex and pointed cones with nonempty interior. Moreover, each of these cones are self-dual, which means that the dual cone $\mathscr{K}^*$ is equal to the original cone $\mathscr{K}$. The same holds for any (finite) direct product of such cones.

If x and (y,s) are feasible for (1), then the *weak duality* property holds

$$c^T x - b^T y = x^T s \geq 0. \tag{2}$$

The *strong duality* property $c^T x = b^T y$ does not always hold for CLO problems. A sufficient condition for strong duality is the primal-dual Slater condition, which requires the existence of a feasible solution pair x and (y,s) for (1) such that $x \in \text{int}\mathscr{K}$ and $s \in \text{int}\mathscr{K}^*$. In this case, the primal-dual optimal set of solutions (x,y,s) is

$$\begin{aligned} Ax &= b,\ x \in \mathscr{K}, \\ A^T y + s &= c,\ s \in \mathscr{K}^*, \\ x^T s &= 0. \end{aligned} \tag{3}$$

System (3) is known as the *optimality conditions*.

Conic quadratic optimization (CQO) is the problem of minimizing a linear objective function subject to the intersection of an affine set and the direct product of quadratic cones. CQO is the sub-class of CLO and, consequently, CQO problems are expressed in the form of (1). More information on CLO and CQO problems, their properties and duality results can be found in [?]. The CQO problem subclasses described in this section include *linear optimization* (LO), *convex quadratic optimization* (QO), quadratically constrained quadratic optimization (QCQO) and *second order conic optimization* (SOCO). CLO, among others, includes CQO and semidefinite optimization (SDO). In all these cases CLO problems can be solved efficiently by Interior Point Methods (IPMs).

LO, QO and SOCO formulations are presented below and their parametric (multiobjective) counterparts are discussed in Section 3. We leave parametric semidefinite optimization outside of this paper, even though there are some results available for this class of problems [?]. However, according to our best knowledge, there are no algorithms for parametric semidefinite optimization that are implementation-ready and can be used in practice.

Linear optimization problems are formulated as:

$$\begin{array}{llll} & \textit{Primal problem} & & \textit{Dual problem} \\ (LP) & \min\limits_x \{c^T x \,:\, Ax = b,\ x \geq 0\} & (LD) & \max\limits_{y,s} \{b^T y \,:\, A^T y + s = c,\ s \geq 0\}, \end{array} \tag{4}$$

where $A \in \mathbb{R}^{m\times n}$, $\text{rank}(A) = m$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $x,s \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.

Convex quadratic optimization problems contain a convex quadratic term in the objective function:

$$\begin{array}{llll} & \textit{Primal problem} & & \textit{Dual problem} \\ & \min\limits_x\ c^T x + \frac{1}{2} x^T Q x & & \max\limits_{x,y,s}\ b^T y - \frac{1}{2} x^T Q x \\ (QP) & \text{s.t. } Ax = b & (QD) & \text{s.t. } A^T y + s - Qx = c \\ & \quad x \geq 0 & & \quad s \geq 0, \end{array} \tag{5}$$

where $Q \in \mathbb{R}^{n\times n}$ is a symmetric positive semidefinite matrix.

In *second order conic optimization* problems the variables are restricted to lie in the Lorentz cone leading to the following formulation:

$$\begin{array}{llll} & \textit{Primal problem} & & \textit{Dual problem} \\ & \min\limits_x\ c^T x & & \max\limits_{y,s}\ b^T y \\ (SOCP) & \text{s.t. } Ax = b & (SOCD) & \text{s.t. } A^T y + s = c \\ & \quad x_1^i \geq \|x_{2:n_i}^i\|,\ i = 1,\dots,I & & \quad s_1^i \geq \|s_{2:n_i}^i\|,\ i = 1,\dots,I, \end{array} \tag{6}$$

where $x = (x_1^1, x_2^1, \ldots, x_{n_1}^1, x_1^2, x_2^2, \ldots, x_{n_2}^2, \ldots, x_1^I, x_2^I, \ldots, x_{n_I}^I)^T \in \mathbb{R}^n$ and $s = (s_1^1, s_2^1, \ldots, s_{n_1}^1, s_1^2, s_2^2, \ldots, s_{n_2}^2, \ldots, s_1^I, s_2^I, \ldots, s_{n_I}^I)^T \in \mathbb{R}^n$ with $n = \sum_{i=1}^I n_i$. Second order cone constraints of the type $x_1^i \geq \|x_{2:n_i}^i\|$ are often written as $(x_1^i, \overline{x}^i) \in \mathscr{K}_q^i$, where $\overline{x}^i = x_{2:n_i}^i$, or just $(x_1^i, \overline{x}^i) \geq_{\mathscr{K}_q^i} 0$.

As $(x_1^1, \ldots, x_{n_1}^1)^T \in \mathscr{K}_q^1$, $(x_1^1, \ldots, x_{n_2}^2)^T \in \mathscr{K}_q^2, \ldots, (x_1^I, \ldots, x_{n_I}^I)^T \in \mathscr{K}_q^I$ and $\mathscr{K} = \mathscr{K}_q^1 \times \mathscr{K}_q^2 \times \ldots \times \mathscr{K}_q^I$ we can also rewrite problem (6) in its shorter form (1). In the remainder of the paper, cone $\mathscr{K}$ denotes the quadratic cone (direct product of linear cones $\mathscr{K}_\ell$ and quadratic cones $\mathscr{K}_q$), unless otherwise specified.

In addition to LO and QO problems, SOCO also includes quadratically constrained quadratic optimization (QCQO). Details about the QCQO problem formulation and its transformation to SOCO formulation can be found in [?].

LO, QO and SOCO problems presented in this section are single-objective convex optimization problems. Most of real-life optimization problems are multiobjective in their nature and in many cases those can be formulated as multiobjective LO, QO or SOCO problems. Theoretical background and solution techniques for multiobjective optimization are discussed in Section 2. In that section we also highlight the relationships between multiobjective optimization and parametric optimization that is used to solve such problems. Parametric optimization algorithms for LO, QO and SOCO optimization problems is the subject of Section 3. Extensions to other classes of optimization problems, e.g., convex non-linear optimization, are briefly mentioned. Finally, we present financial applications of multiobjective optimization and numerically solve three examples in Section 4.

2 Multiobjective and Parametric Optimization

Let x be an n-dimensional vector of *decision variables*. The multiobjective optimization problem, where the goal is to optimize a number of possibly conflicting objectives simultaneously, is formulated as:

$$\begin{array}{l} \min \ \{f_1(x), f_2(x), \ldots, f_k(x)\} \\ \text{s.t. } x \in \Omega, \end{array} \tag{7}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, k$ are (possibly) conflicting objectives and $\Omega \subseteq \mathbb{R}^n$ is a feasible region. Each of the functions f_i represent an attribute or a *decision criterion* that serves the base for the decision making process.

Multiobjective optimization is a subclass of *vector optimization*, where the vector-valued objective function $f_0 = \{f_1(x), f_2(x), \ldots, f_k(x)\}$ is optimized with respect to a proper convex cone $\mathscr{C}$ which defines preferences. When a vector optimization problem involves the cone $\mathscr{C} = \mathbb{R}_+$, it is known as a *multicriteria* or *multiobjective optimization* problem.

In this paper we consider *convex multiobjective conic optimization* problems and most of the results hereafter are restricted to that problem class. Moreover, we also mention some of the results available for general multiobjective problems. Problem (7) is a *convex* multiobjective optimization problem if all the objective functions $f_1, \ldots, f_k$ are convex, and the feasible set Ω is convex as well. For example, it can be defined as $\Omega = \{x : g_j(x) \leq 0,\ h_j(x) = 0\}$, where the inequality constraint functions $g_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, l$ are convex and the equality constraint functions $h_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, m$ are affine. For LO, QO and SOCO problems the set of constraints can be written as $\Omega = \{x : Ax = b,\ x \geq_{\mathscr{K}} 0\}$, where $\mathscr{K}$ is an appropriate convex cone and $Ax = b$ are the equality constraints with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The set Ω, is called the feasible region in the decision space or just the *decision space*.

Definition 1 A vector $x^* \in \Omega$ is Pareto optimal (or efficient solution) if there does not exist another $x \in \Omega$ such that $f_i(x) \leq f_i(x^*)$ for all $i = 1, \ldots, k$ and $f_j(x) < f_j(x^*)$ for at least one index j.

The set of all Pareto optimal (or efficient) solutions $x^* \in \Omega$ is called the Pareto optimal (efficient solution) set Ω_E.

As values of the objective functions are used for making decisions by the decision maker, it is conventional for multiobjective optimization to work in the space of the objective functions, which is called the *objective space*. By mapping the feasible region into the objective space, we get:

$$Z = \{z \in \mathbb{R}^k : z = ((f_1(x), f_2(x), \ldots, f_k(x))^T \, \forall x \in \Omega)\}.$$

The set Z is the set of objective values of feasible points, it is referred to as the set of *achievable objective values*. Points in the achievable set Z can be ranked into efficient and non-efficient points (see Figure 1) that leads to the definition of Pareto optimality.

Analogous definition of Pareto optimality can be stated for an objective vector $z^* \in Z$. Equivalently, z^* is Pareto optimal if the decision vector x^* corresponding to it is Pareto optimal [?].

Definition 2 For a given multiobjective problem (7) and Pareto optimal set Ω_E, the Pareto front is defined as:

$$Z_N = \{z^* = (f_1(x^*), \ldots, f_k(x^*))^T \mid x^* \in \Omega_E\}.$$

A set Z_N of Pareto optimal (also called nondominated or efficient) solutions z^* forms the *Pareto efficient frontier* or *Pareto front*. The Pareto front, if $k = 2$, is also known as the *optimal trade-off curve* and for $k > 2$ it is called the *optimal trade-off surface* or the *Pareto efficient surface*.

Solution methods are designed to help the decision maker to identify and choose a point on the Pareto front. Identifying the whole frontier is computationally challenging, and often it cannot be performed in reasonable time. Solution methods for multiobjective optimization are divided into the following categories [?]:

- *a priori methods* are applied when the decision maker's preferences are known a priory; those include the value function method, lexicographic ordering and goal programming.
- *iterative methods* guide the decision maker to identify a new Pareto point from an existing one (or existing multiple points), the process is stopped when the decision maker is satisfied with the actual efficient point.
- *a posteriori methods* are used to compute the Pareto front or some of its parts; those methods are based on the idea of scalarization, namely transforming the multiobjective optimization problem into a series of single-objective problems; a posteriori methods include *weighting methods*, the *ε-constrained method* and related scalarization techniques.

Computing the Pareto front can be challenging as it does not posses known structure in most of the cases, and, consequently, discretization in the objective space is frequently used to compute it. The problem is that discretization is computationally costly in higher dimensions, and discretization is not guaranteed to produce all the (or desired) points on the Pareto front.

It turns out that for some classes of multiobjective optimization problems the structure of the efficient frontier can be identified. Those include multiobjective LO, QO and SOCO optimization problems. For those classes of problems, the Pareto efficient frontier can be sub-divided into pieces (subsets) that have specific properties. These properties allow the identification of each subsets of the frontier. The piece-wise structure of the Pareto front also provides additional information for the decision maker.

Before looking at the scalarization solution techniques for multiobjective optimization, that allow us to identify all nondominated (Pareto efficient) solutions, we need to introduce a number of concepts and some theoretical results.

Definition 3 An objective vector $z^* \in Z$ is weakly Pareto optimal if there does not exist another decision vector $z \in Z$ such that $z_i < z_i^*$ for all $i = 1, \dots, k$.

The set of weakly Pareto efficient (nondominated) vectors is denoted by Z_{wN}. It follows that $Z_N \subseteq Z_{wN}$. When unbounded trade-offs between objectives are not allowed, Pareto optimal solutions are called *proper* [?]. The set of properly efficient vectors is denoted as Z_{pN}.

Both sets Z_{wN} (weak Pareto front) and Z_N (Pareto front) are connected if the functions f_i are convex and the set Ω satisfies one of the following properties [?]:

- Ω is a compact, convex set;
- Ω is a closed, convex set and $\forall z \in Z$, $\Omega(z) = \{x \in \Omega : f(x) \leq z\}$ is compact.

Let us denote by $\mathbb{R}^k_+ = \{z \in \mathbb{R}^k : z \geq 0\}$ the nonnegative orthant of $\mathbb{R}^k$. Consider the set:

$$\mathscr{A} = Z + \mathbb{R}^k_+ = \{z \in \mathbb{R}^k : f_i(x) \leq z_i,\ i = 1, \dots, k,\ x \in \Omega\},$$

that consists of all values that are worse than or equal to some achievable objective value. While the set Z of achievable objective values need not be convex, the set $\mathscr{A}$ is convex, when the multiobjective problem is convex [?].

Definition 4 A set $Z \in \mathbb{R}^k$ is called $\mathbb{R}^k_+$-convex if $Z + \mathbb{R}^k_+$ is convex.

A point $x \in \mathscr{C}$ is a minimal element with respect to componentwise inequality induced by $\mathbb{R}^k_+$ if and only if $(x - \mathbb{R}^k_+) \cap \mathscr{C} = x$. The minimal elements of $\mathscr{A}$ are exactly the same as the minimal elements of the set Z. This also means that any hyperplane tangent to the Pareto efficient surface is a supporting hyperplane – the Pareto front is on one side of the hyperplane [?]. It follows that the Pareto front must belong to the boundary of Z [?].

Proposition 1 $Z_N = \left(Z + \mathbb{R}^k_+\right)_N \subset \mathrm{bd}(Z)$.

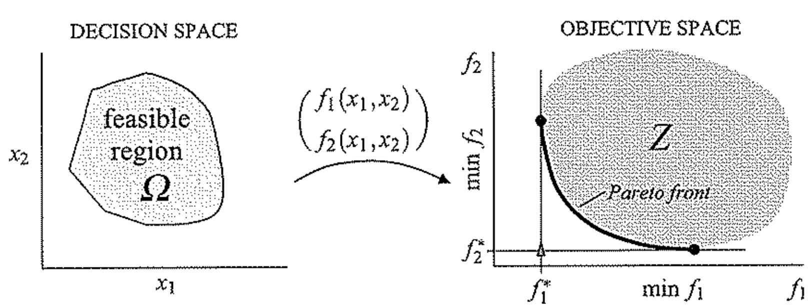


Fig. 1: Mapping the Decision Space into the Objective Space.

When talking about convex multiobjective optimization problems, it is useful to think of the Pareto front as a function, and not as a set. Under assumptions about convexity of the functions f_i and the set Ω for bi-objective optimization problems ($k = 2$), the (weakly) Pareto front is a convex function [?]. Unfortunately, when $k > 2$ it is not the case even for linear multiobjective optimization problems.

Most a posteriori methods for solving multiobjective optimization problems are based on scalarization techniques. Let us consider the two most popular scalarization methods:

- weighting method;
- ε-constraint method.

2.1 Weighting Method

The idea of the weighting method is to assign weights to each objective function and optimize the weighted sum of the objectives. A multiobjective optimization problem can be solved with the use of the *weighting method* by optimizing single-objective problems of the type

$$\begin{aligned} \min \ & \sum_{i=1}^{k} w_i f_i(x) \\ \text{s.t. } & x \in \Omega, \end{aligned} \tag{8}$$

where f_i is linear, convex quadratic or second order conic function in our case, $\Omega \subseteq \mathbb{R}^n$ (convex), $w_i \in \mathbb{R}$ is the weight of the i-th objective, $w_i \geq 0, \forall i = 1, \ldots, k$ and $\sum_{i=1}^{k} w_i = 1$. Weights w_i define the importance of each objectives. Due to the fact that each objectives can be measured in different units, the objectives may have different magnitudes. Consequently, for the weight to define the relative importance of objectives, all objectives should be normalized first. Some of the normalization methods are discussed in [?]. As we intend to compute the whole Pareto front, normalization is not required.

It is known that the weighting method produces weakly efficient solutions when $w_i \geq 0$ and efficient solutions if $w_i > 0$ for all $i = 1, \ldots, k$ [?]. For convex multiobjective optimization problems any Pareto optimal solution x^* can be found by the weighting method.

Let us denote by $\mathscr{S}(w, Z) = \{\hat{z} \in Z : \hat{z} = \operatorname{argmin}_{z \in Z} w^T z\}$ the set of optimal points of Z with respect to w. In addition, we define

$$\mathscr{S}(Z) = \bigcup_{w > 0, \ \sum_{i=1}^{k} w_i = 1} \mathscr{S}(w, Z), \qquad \mathscr{S}_0(Z) = \bigcup_{w \geq 0, \ \sum_{i=1}^{k} w_i = 1} \mathscr{S}(w, Z).$$

As Z is $\mathbb{R}^k_+$-convex set in our case, we get [?]:

$$\mathscr{S}(Z) = Z_{pN} \subset Z_N \subset \mathscr{S}_0(Z) = Z_{wN}. \tag{9}$$

In addition, if $\hat{z}$ is the unique element of $\mathscr{S}(w, Z)$ for some $w \geq 0$, then $\hat{z} \in Z_N$ [?]. The last observation combined with (9), allows us identifying the whole (weak) Pareto front with the use of the weighting method.

2.2 ε-Constrained Method

For illustration purposes, we first consider a problem with two objective functions. Multi-objective optimization can be based on ranking the objective functions in descending order of importance. Each objective function is then minimized individually subject to a set of additional constraints that do not allow the values of each of the higher ranked functions to exceed a prescribed fraction of their optimal values obtained in the previous step. Suppose that f_2 has higher rank than f_1. We then solve

$$\min\{f_2(x) : x \in \Omega\},$$

to find the optimal objective value f_2^*. Next, we solve the problem

$$\begin{aligned} &\min f_1(x) \\ &\text{s.t. } f_2(x) \le (1+\varepsilon) f_2^*, \\ &\quad x \in \Omega. \end{aligned}$$

Intuitively, the hierarchical ranking method can be thought as saying "f_2 is more important than f_1 and we do not want to sacrifice more than ε percentage of the optimal value of f_2 to improve f_1."

Considering the general case of k objective functions and denoting the right-hand-side term of the constraints on the objective functions' values by $\varepsilon_j = (1+\varepsilon_j) f_j^*$, we get the following single-objective optimization problem, which is known as the *ε-constrained method*:

$$(MOC_\varepsilon) \qquad \begin{aligned} &\min f_\ell(x) \\ &\text{s.t. } f_j(x) \le \varepsilon_j,\ j = 1,\dots,k,\ j \ne \ell \\ &\quad x \in \Omega. \end{aligned} \qquad (10)$$

Every solution x^* of the ε-constrained problem (10) is weakly Pareto optimal [?], so formulation (10) can be used to compute weak Pareto front Z_{wN}.

Let x^* solve (10) with $\varepsilon_j^* = f_j(x^*)$, $j \ne \ell$. Then x^* is *Pareto optimal* [?,?] if:

1) x^* solves (10) for every $\ell = 1,\dots,k$;
2) x^* is the unique solution of (10);
3) Lin's conditions [?,?].

The third set of necessary and sufficient conditions for (strong) Pareto optimality of optimal solutions is described in [?] based on the results of Lin [?,?]. Let us define

$$\phi_\ell(\varepsilon) = \min\{f_\ell(x) : x \in \Omega,\ f_j(x) \le \varepsilon_j \text{ for each } j \ne \ell\}.$$

The following theorem [?] establishes that x^* is Pareto optimal if the optimal value of (MOC_{ε^0}) is strictly greater than $f_\ell(x^*)$ for any $\varepsilon^0 \le \varepsilon^*$.

Theorem 1 *Let x^* solve (10) with $\varepsilon_j^* = f_j(x^*)$, $j \ne \ell$. Then x^* is Pareto optimal solution if and only if $\phi_\ell(\varepsilon) > \phi_\ell(\varepsilon^*)$ for all ε such that $\varepsilon \le \varepsilon^*$ and for each ε (10) has an optimal solution with finite optimal value.*

In many cases, conditions 2) and 3) can be verified to identify the Pareto front Z_N. For instance, the second condition holds when all the objective functions $f_j(x)$ are strictly convex. Condition 3) can be verified if function $\phi_\ell(\varepsilon)$ is computed by parametric optimization techniques, see Section 2.4.

2.3 Parametric Optimization

Optimization models typically contain two types of variables: those that can be changed, controlled or influenced by the decision maker are called *parameters*, the remaining ones are the decision variables. Parameters arise because the input data of the problem is not accurate or is changing over time. The main interest of sensitivity analysis is to determine how known characteristics of the problem are changing by small perturbations of the data. However, if we go farther from the current parameter value, not only the current properties of the problem might not be valid, but also the problem may change significantly. Study of this situation is referred to as *parametric optimization*.

Let us consider a general convex parametric optimization problem

$$\phi(\lambda) = \min\{f(x,\lambda) \ : \ x \in \mathscr{M}(\lambda),\ \lambda \in \Lambda\}, \tag{11}$$

with a parameter vector λ and function $f(x,\lambda)$ that is convex in terms of x. Let $\phi(\lambda)$ be the optimal value function and $\psi(\lambda)$ is the optimal set map of (11), and for our purpose

$$\mathscr{M}(\lambda) = \{x \in X \ : \ g_i(x) \leq \lambda_i, i = 1,2,\ldots,m\}, \tag{12}$$

where g_i are real-valued functions defined on X. Observe that both $\mathscr{M}(\lambda)$ and $\psi(\lambda)$ are two point-to-set maps.

In parametric optimization, in addition to the optimal value function $\phi(\lambda)$, the optimal solution set $\psi(\lambda)$ is considered as function of the parameter vector. Investigating their behavior is the aim of parametric optimization.

Among many properties defined for optimization models, the following two are important ones: *uniqueness of optimal solution* and *stability*. When there is a unique optimal solution for each parameter, the solution set map is a real-valued map (as opposed to set-valued maps). Uniqueness of optimal solution is often violated in practice, especially for large-scale problems. For LO and QO problems it leads to degeneracy of optimal solutions, causing difficulties and ambiguities in post-optimality analysis [?]. Having multiple optimal solutions converts the optimal solution map to a point-to-set map. Continuous feasible perturbation of the parameter imply continuous changes of the feasible set and the optimal value function. The optimal value function describes the behavior of the objective function regardless of the uniqueness of the optimal solution.

By stability we mean having some important invariant properties of the problem such as continuity of the optimal value function or its differentiability. Even though the notion of "*stability*" stands for many different properties, there are two main ones. In [?], it is used for describing the semicontinuity of the optimal value function as well as the study of upper Hausdorff semicontinuity of the optimal solution set, that is a set-valued function in general. This approach to stability for QO has been also studied in [?]. Most probably optimal value function has critical points. If a parameter value is not a critical point of the optimal value function (or far enough from a critical point), it assures the decision-maker that this solution is stable.

There is another notion, "*critical region*", used in parametric optimization with numerous meanings. In the LO literature it may refer to the region for parameter values where the given optimal basis (might be degenerate and not unique) remains optimal [?]. In QO and LO [?] it might also refer to the region where the active constraint set remains the same [?]. It is worth mentioning that the existence of a strictly complementary optimal solution is guaranteed in LO. Strictly complementary optimal solutions define the optimal partition of the index set (Section 3.1). The optimal partition is unique for any LO problem, and it has direct relation to the behavior of the optimal value function. In parametric optimization, one may be interested in studying the behavior of the optimal solution set, while one also might want to investigate the behavior of the optimal value function. These differences lead to diverse results in the context of parametric optimization. Here, we consider stability and critical regions in the context of optimal partition (see Section 3.1).

Recall that every convex optimization problem is associated with a Lagrangian dual optimization problem (see, e.g. [?]). The optimal objective function value of the dual problem is a lower bound for the optimal value of the primal objective function value (weak duality property). When these two optimal values are equal (strong duality), i.e., when the duality gap is zero, then optimality is achieved. However, the strong duality property does not hold in general, but when the primal optimization problem is convex and the *Slater constraint qualification* holds, i.e., there exists a strictly feasible solution, strong duality is guaranteed. In some special cases like LO, QO and SOCO, parameters in the right-hand side of the constraints are translated as parameters in the coefficients of the objective function of its dual problem. This helps us to unify the analysis.

In this paper, we consider optimal partition invariancy for LO problems [?] and for QO problems [?] as a criterion for analyzing the behavior of the optimal value function (see Section 3.1). We also extend these definitions to SOCO problems.

2.4 Multiobjective Optimization via Parametric Optimization

By now, the reader may have understood that multiobjective optimization problems are closely related to, and can be represented as parametric optimization problems. Consequently, we may use algorithms of parametric optimization to solve multiobjective optimization problems and to compute the Pareto fronts. Before defining the relations between multiobjective optimization and parametric optimization more formally, we mention that multiobjective LO, QO and, to some extent, SOCO problems can be efficiently solved by parametric optimization algorithms. Parametric optimization techniques exist for wider classes of problems, but computational complexity may prevent using those directly to identify efficient frontiers.

The main idea of this paper is that we can solve multiobjective optimization problems using parametric optimization techniques. *A posteriori multiobjective optimization techniques* are based on parameterizing (scalarizing) the objective space and solving the resulting parametric problem. Consequently, parametric optimization algorithms can be utilized to solve multiobjective optimization problems.

Based on the weighting method (8) and choosing the vector of weights as $w = (\lambda_1, \ldots, \lambda_{k-1}, 1)^T \geq 0$, as w can be scaled by a positive constant, for the weighted objective function $\sum_i w_i f_i(x)$, we can formulate the parametric optimization problem with the λ_i parameters in the objective function as

$$\begin{array}{rll} \phi(\lambda_1, \ldots, \lambda_{k-1}) = & \min & \lambda_1 f_1(x) + \ldots + \lambda_{k-1} f_{k-1}(x) + f_k(x) \\ & \text{s.t.} & x \in \Omega, \end{array} \tag{13}$$

for computing weakly Pareto optimal solutions, or $(\lambda_1, \ldots, \lambda_{k-1})^T > 0$ for computing Pareto optimal solutions. Formulation (13) is known as the *Lagrangian problem* [?] and possesses almost identical properties as the weighting problem (8).

Based on the ε-constrained method (10) we can present the following parametric problem:

$$\begin{aligned} \phi(\varepsilon_1,\ldots,\varepsilon_{k-1}) = \min\ & f_k(x) \\ \text{s.t}\ & f_i(x) \le \varepsilon_i,\ i=1,\ldots,k-1 \\ & x \in \Omega, \end{aligned} \tag{14}$$

where $\varepsilon_1,\ldots,\varepsilon_{k-1}$ are parameters in the right-hand-side of the constraints. In this case, the optimal value function $\phi(\varepsilon_1,...,\varepsilon_{k-1})$ includes the Pareto front as a subset.

It is not hard to observe that the parametric problems (13) and (14) are equivalent to (8) and (10), respectively, but they are just written in the forms used in the parametric optimization literature. The relationships between those formulations and their properties are extensively studied in [?].

Algorithms and techniques developed for solving parametric optimization problems are described in Section 3. Note that the optimal value function $\phi(\varepsilon)$ of problem (14) is the boundary of the set $\mathscr{A}$ and the Pareto front is a subset of that boundary. These results are illustrated by examples in Section 4.

2.5 Multiobjective and Parametric Quadratic Optimization

Results described by now in Section 2 apply to general convex multiobjective optimization problems. In contrast, parametric optimization techniques discussed in this paper apply to LO, QO and SOCO problems only. In this section we specialize the formulations presented in Section 2.4 to the parametric optimization problem classes described in Section 3.

We define the *multiobjective quadratic optimization* problem as a convex multiobjective problem with one convex quadratic objective function f_k and $k-1$ linear objectives $f_1,\ldots,f_{k-1}$ subject to linear constraints. For the multiobjective QO problem the weighted sum formulation (13) specializes to

$$\begin{aligned} \phi(\lambda_1,\ldots,\lambda_{k-1}) = \min\ & \lambda_1 c_1^T x + \ldots + \lambda_{k-1} c_{k-1}^T x + \tfrac{1}{2} x^T Q x \\ \text{s.t.}\ & Ax = b \\ & x \ge 0, \end{aligned} \tag{15}$$

and the ε-constrained formulation (14) becomes

$$\begin{aligned} \phi(\varepsilon_1,\ldots,\varepsilon_{k-1}) = \min\ & \tfrac{1}{2} x^T Q x \\ \text{s.t.}\ & c_i^T x \le \varepsilon_i,\ i=1,\ldots,k-1 \\ & Ax = b \\ & x \ge 0. \end{aligned} \tag{16}$$

Parametric QO formulations (15) and (16) can be solved with algorithms developed in Section 3. The uni-parametric case corresponds to an optimization problem with two objectives. A bi-parametric QO algorithm allows solving multiobjective QO problems with three objectives. Multiobjective problems with more than three objectives require multi-parametric optimization techniques (see Section 3.4.1). Note that in formulations (15) and (16), parameters appear in the objective function and in the right-hand side of the constraints, respectively.

Multiobjective QO problems are historically solved by techniques that approximate the Pareto front [?,?]. An alternative approach is the parametric optimization discussed in this paper. Examples of multiobjective QO problems appearing in finance are solved with parametric QO techniques in Section 4.

If we allow for more than one convex quadratic objective in the multiobjective optimization problem, formulations (15) and (16) become parametric QOCO. It happens due to the fact that now quadratic functions appear in the constraints as well. Parametric SOCO, that includes parametric QCQO problems, is a more general class of problems. Preliminary results for solving parametric SOCO problems are described in Section 3.3. Properties of multiobjective optimization problems with more than one convex quadratic objectives and linear constraints are discussed in [?].

As we learned in this section, multiobjective optimization problems can be formulated as parametric optimization problems. Some classes of multiobjective optimization problems that include linear and convex quadratic optimization problems can be efficiently solved using parametric optimization algorithms. Parametric optimization allows not only computing Pareto efficient frontiers (surfaces), but also identifying piece-wise structures of those frontiers. Structural description of Pareto fronts gives functional form of each of its pieces and thus helps decision makers to make better decisions.

3 Solving Parametric Optimization Problems

Utilizing different approaches for solving multiobjective optimization problems via parametric optimization (see Sections 2.4 and 2.5), we review methods and results of uni- and bi-parametric LO, QO and SOCO problems in this section. Our methodology is based on the notion of optimal partition and we study the behavior of the optimal value function that contains the Pareto front. To save space, all proofs are omitted, and we refer the interested reader to [?] and the related publications listed there for more details.

3.1 Uni-parametric Linear and Convex Quadratic Optimization

The primal and dual solutions sets of QO are denoted by $\mathscr{QP}$ and $\mathscr{QD}$, respectively. Observe that the QO problem reduces to a LO problem where $Q=0$. For a primal-dual optimal solution (x^*,y^*,s^*), the complementarity property $x^{*T}s^*=0$ holds, that is equivalent to $x_j^*s_j^*=0$ for $j\in\{1,2,\ldots,n\}$. A strictly complementary optimal solution further satisfies $x^*+s^*>0$. The existence of this kind of optimal solutions is true only for LO, while there is no guarantee to have strictly complementary optimal solution for any other class of optimization problems. For QO and SOCO problems the existence of maximally complementary optimal solution is proved. A primal-dual maximally complementary optimal solution has the maximum number of positive components for both x and s. In this case the *optimal partition* can be uniquely identified by:

$$\mathscr{B}=\{j\,:\,x_j>0,\,x\in\mathscr{QP}\},$$
$$\mathscr{N}=\{j\,:\,s_j>0,\,(y,s)\in\mathscr{QD}\},$$
$$\mathscr{T}=\{1,2,\ldots,n\}\setminus(\mathscr{B}\cup\mathscr{N}).$$

As mentioned previously, for LO the set $\mathscr{T}$ is always empty.

The *uni-parametric QO* problem, with the parameter in the right-hand-side of the constraints, is defined as

$$(QP_\varepsilon)\quad \phi(\varepsilon)=\min\{c^Tx+\frac{1}{2}x^TQx\,:\,Ax=b+\varepsilon\triangle b,\,x\geq 0\}, \tag{17}$$

with its dual as

$$(QD_\varepsilon) \quad \max\{(b+\varepsilon\triangle b)^T y - \frac{1}{2}x^T Qx \; : \; A^T y + s - Qx = c, \; s \geq 0, \; x \geq 0\}, \tag{18}$$

where $\triangle b \in \mathbb{R}^m$ is the fixed perturbing vector. The corresponding sets of feasible solutions are denoted by $\mathscr{QP}_\varepsilon$ and $\mathscr{QD}_\varepsilon$, and the optimal solution sets as $\mathscr{QP}^*_\varepsilon$ and $\mathscr{QD}^*_\varepsilon$, respectively.

The optimal value function $\phi(\varepsilon)$ is a piecewise convex (linear when $Q=0$) quadratic function over its domain. Points where the optimal partition changes are referred to as *transition points*. These are precisely the points where the representation of the optimal value function changes too. At these points, the optimal value functions fails to have first or second order derivatives. As the number of tri-partitions of the index set is finite, there are a finite number of quadratic pieces of the optimal value function.

To find the representation of the optimal value function on the (invariancy) intervals between two consequent transition points, we only need to have primal-dual optimal solutions for two parameter values from that interval.

Theorem 2 *For two values $\varepsilon_1 < \varepsilon_2$ with identical optimal partition, the optimal partition is the same for all $\varepsilon \in [\varepsilon_1, \varepsilon_2]$. Moreover, if (x^1, y^1, s^1) and (x^2, y^2, s^2) are maximally complementary primal-dual optimal solutions at ε_1 and ε_2, then the optimal value function can be represented as*

$$\phi(\varepsilon) = \phi(0) + \varepsilon\triangle b^T y^1 + \frac{1}{2}\varepsilon^2 \triangle b^T (y^2 - y^1), \tag{19}$$

where $\phi(0)$ corresponds to the optimal value of the unperturbed problem at $\varepsilon = 0$ and $0 \in [\varepsilon_1, \varepsilon_2]$.

Proof See [?], p. 6-21. □

Observe that for the LO case, when the optimal partition is fixed then the dual optimal solution set is invariant and consequently, the coefficient of the square term is zero. Thus, the objective value function is linear. To find an invariancy interval one needs to solve two auxiliary LO problems when a primal-dual optimal solution is available at an arbitrary parameter value in this interval. Let $\sigma(v)$ denotes the support set of the vector v, i.e., the index set of nonzero components of the vector v.

Theorem 3 *Let $x^* \in \mathscr{QP}^*_\varepsilon$ and $(x^*, y^*, s^*) \in \mathscr{QD}^*_\varepsilon$ be given for arbitrary ε. Let $(\varepsilon_\ell, \varepsilon_u)$ denotes the invariancy interval that includes ε. Moreover, let $T = \{1, 2, \ldots, n\} \setminus (\sigma(x^*) \cup \sigma(s^*))$. Then*

$$\begin{aligned}
\varepsilon_\ell = \min\{\varepsilon : \; & Ax - \varepsilon\triangle b = b, x \geq 0, x^T s^* = 0, x_T = 0,\\
& A^T y + s - Qx = c, s \geq 0, s^T x^* = 0, s_T = 0\},\\
\varepsilon_u = \max\{\varepsilon : \; & Ax - \varepsilon\triangle b = b, x \geq 0, x^T s^* = 0, x_T = 0,\\
& A^T y + s - Qx = c, s \geq 0, s^T x^* = 0, s_T = 0\}.
\end{aligned}$$

Proof See [?], p. 6-26. □

Remark 1 In case of LO, Theorem 3 reduces to solving the following simpler problems:

$$\begin{aligned}
\varepsilon_\ell &= \min\{\varepsilon : Ax - \varepsilon\triangle b = b, x \geq 0, x^T s^* = 0\},\\
\varepsilon_u &= \max\{\varepsilon : Ax - \varepsilon\triangle b = b, x \geq 0, x^T s^* = 0\}.
\end{aligned}$$

Observe that this interval might be unbounded from one side if the corresponding auxiliary LO problem is unbounded, and the interval is singleton (transition point), when the given optimal solution is maximally (strictly in LO case) complementary and then $\varepsilon_\ell = \varepsilon_u$.

Finding the left and right derivatives of the optimal value function at a transition point requires the solution of two LO problems, provided an arbitrary primal-dual optimal solution is given at this point.

Theorem 4 *With the notation of Theorem 3, let (x^*, y^*, s^*) be a given optimal solution pair at the specific transition point ε. The left and right derivatives of the optimal value function $\phi(\varepsilon)$ are given by the optimal values of the following optimization problems:*

$$\begin{aligned}
\phi'_- = \min\nolimits_{x,y,s}\{\triangle b^T y : & \; Ax - \varepsilon\triangle b = b, x \geq 0, x^T s^* = 0, x_{\mathcal{T}} = 0, \\
& A^T y + s - Qx = c, s \geq 0, s^T x^* = 0, s_{\mathcal{T}} = 0\}, \\
\phi'_+ = \max\nolimits_{x,y,s}\{\triangle b^T y : & \; Ax - \varepsilon\triangle b = b, x \geq 0, x^T s^* = 0, x_{\mathcal{T}} = 0, \\
& A^T y + s - Qx = c, s \geq 0, s^T x^* = 0, s_{\mathcal{T}} = 0\}.
\end{aligned}$$

Proof See [?], p. 6-27. □

Remark 2 In case of LO, Theorem 4 reduces to solving the following simpler problems:

$$\begin{aligned}
\phi'_- &= \min\{\triangle b^T y : A^T y + s = c, \; s \geq 0, \; s^T x^* = 0\}, \\
\phi'_+ &= \max\{\triangle b^T y : A^T y + s = c, \; s \geq 0, \; s^T x^* = 0\}.
\end{aligned}$$

Observe that at a transition point ϕ'_- or ϕ'_+ might be infinite as one of the associated auxiliary LO problems might be unbounded. It is not possible to have both LO problems in Theorem 4 feasible and unbounded, and they are both bounded and feasible when the transition point is not an end point of the domain of the optimal value function. Moreover, if ε is not a transition point, then $\phi'_- = \phi'_+$ which in this case is the derivative of the optimal value function.

We refer the interested reader to [?] for the results when, in case of parametric QO, perturbation exists in the linear term of the objective function. Results for the case, when both the right-hand-side and the linear term of the objective function is perturbed simultaneously with the same parameter can be found in [?].

3.2 From Uni- to Bi-Parametric Optimization

Going from uni-parametric to bi-parametric optimization has different problem formulations. One formulation of bi-parametric optimization is to have one of the parameters in the right-hand-side of the constraints and the second one in the objective function data. This point of view to bi-parametric optimization has been considered extensively in LO and QO [?,?,?].

Another formulation is considering these two parameters either both in the right-hand-side of the constraints, or both in the objective function data. From now on, by bi-parametric optimization, we mean having both parameters in the objective or both in the right-hand-side data. Analogous to the previous discussion, we omit the LO problem as a special case, and review the results for bi-parametric QO problem with parameters in the right-hand-side of the constraints. The bi-parametric QO problem is defined as follows:

$$(QP_{\varepsilon,\lambda}) \quad \phi(\varepsilon,\lambda) = \min_x\{c^T x + \frac{1}{2}x^T Qx \; : \; Ax = b + \varepsilon\triangle b^1 + \lambda\triangle b^2, \; x \geq 0\},$$

and its dual as

$$(QD_{\varepsilon,\lambda})\quad \max_{x,y,s}\{(b+\varepsilon\triangle b^1+\lambda\triangle b^2)^T y-\frac{1}{2}x^T Qx\,:\,A^T y+s-Qx=c,\,s\geq 0\},$$

where $\triangle b^1, \triangle b^2 \in \mathbb{R}^m$ are the given perturbing vectors. The case, when so-called critical regions are defined as regions where a given optimal basis remains optimal has been discussed thoroughly in [?]. As mentioned in the uni-parametric LO case, the invariancy region where the optimal partition remains invariant, includes possibly exponentially many critical regions. Moreover, on an invariancy region, the optimal value function has a specific representation, and two disjoint regions correspond to two different representations of the optimal value function. In multiobjective optimization, we are interested in the behavior of the optimal value function (as Pareto front) instead of the optimal solutions set. Thus, we investigate a technique for identifying all invariancy regions and describing the behavior of the optimal value function for bi-parametric optimization problems.

Let the optimal partition be known for $(\varepsilon,\lambda)=(0,0)$. The invariancy region that includes the origin is a (possibly unbounded) polyhedral convex set. This region is denoted here as $\mathscr{IR}(\triangle b^1, \triangle b^2)$ and referred to as the actual invariancy region. To identify this region, we refine the algorithmic approach used in [?]. The results and methodology is analogous to the case of bi-parametric QO [?].

3.2.1 Detecting the Boundary of an Invariancy Region

Observe that an invariancy region might be a singleton or a line segment. We refer to these type of regions as *trivial* regions. In this section, we describe the tools to identify a non-trivial invariancy region. Recall that for $\varepsilon=\lambda$, the bi-parametric QO problem reduces to uni-parametric QO problem. This trivial observation suggests developing a method to convert the bi-parametric QO problem into a series of uni-parametric QO problems. We start with identifying points on the boundary of the invariancy region. To accomplish this, we select the lines passing through the origin as

$$\lambda = t\varepsilon. \tag{20}$$

For now, we assume that the slope t is positive. Substituting (20) into the problem $(QP_{\varepsilon,\lambda})$ converts it into the following uni-parametric QO problem:

$$\min\{c^T x+\frac{1}{2}x^T Qx\,:\,Ax=b+\varepsilon\overline{\triangle b},\,x\geq 0\}, \tag{21}$$

where $\overline{\triangle b}=\triangle b^1+t\triangle b^2$. Now, we can solve two associated auxiliary LO problems from Theorem 3 to identify the range of variation for parameter ε when equation (20) holds. These two auxiliary LO problems are:

$$\begin{aligned}\varepsilon_\ell = \min_{\varepsilon,x,y,s}\{\,\varepsilon : & \,Ax-\varepsilon\overline{\triangle b}=b,\,x_{\mathscr{B}}\geq 0,\,x_{\mathscr{N}\cup\mathscr{T}}=0,\\ & A^T y+s-Qx-\lambda\overline{\triangle c}=c,\,s_{\mathscr{N}}\geq 0,\,s_{\mathscr{B}\cup\mathscr{T}}=0\,\},\end{aligned} \tag{22}$$

and

$$\begin{aligned}\varepsilon_u = \max_{\varepsilon,x,y,s}\{\,\varepsilon : & \,Ax-\varepsilon\overline{\triangle b}=b,\,x_{\mathscr{B}}\geq 0,\,x_{\mathscr{N}\cup\mathscr{T}}=0,\\ & A^T y+s-Qx-\lambda\overline{\triangle c}=c,\,s_{\mathscr{N}}\geq 0,\,s_{\mathscr{B}\cup\mathscr{T}}=0\,\},\end{aligned} \tag{23}$$

where $\pi=(\mathscr{B},\mathscr{N},\mathscr{T})$ is the optimal partition for $\varepsilon=\lambda=0$.

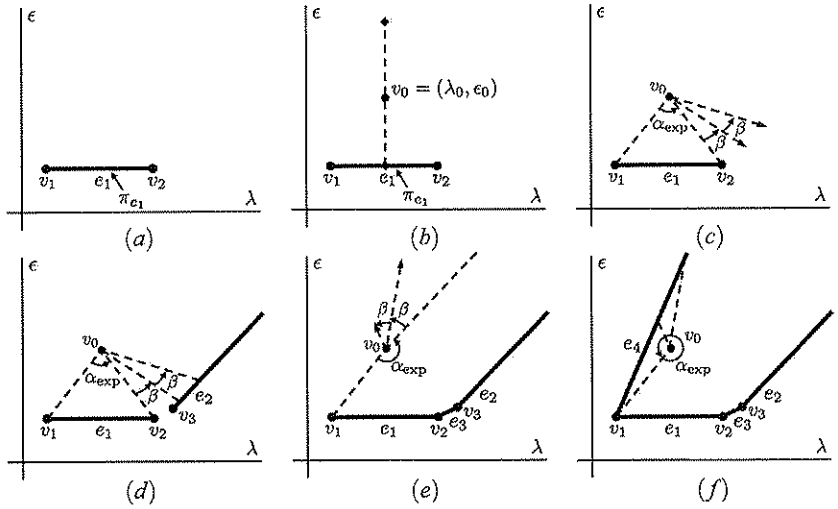


Fig. 2: Invariancy Region Exploration Algorithm for Bi-Parametric QO.

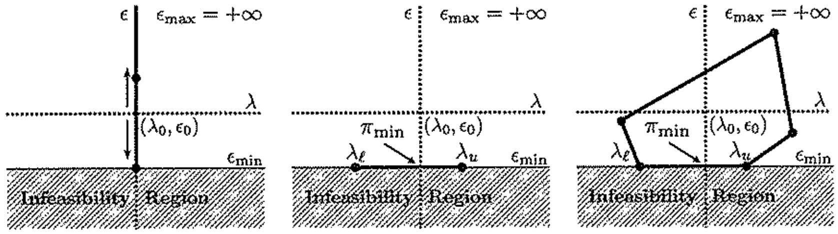
Now, we can summarize the procedure for identifying all transition points (vertices) and transition lines (edges) in an invariancy region. Let's assume that we know an initial inner point of the invariancy region and one of the edges (Figure 2(a) and (b) shows how to find an inner point of the region). We are going to "shoot" by solving sub-problem (22) or (23) counter-clockwise from the initial point to identify each edge (see Figure 2(c-f)). As we already know one of the edges, we exclude all the angles α_{exp} between the initial point and the two vertices v_1 and v_2 of the known edge from the candidate angles to shoot. So, we shoot in the angle $v_0 - v_2$ plus in the small angles β and 2β and identify the optimal partition in the two points we get. Here we find the invariancy region boundary between the vertex v_2 and the point we get when shooting in the angle 2β. If the optimal partition is the same for the points in the directions β and 2β, we compute the vertices of this new edge e_2 and verify if one of those correspond to a vertex of the previously known edge e_1. If it is not the case, then bisection is used to identify the missing edges between e_1 and e_2. We continue in this manner until all edges of the invariancy region are identified.

3.2.2 Transition from an Invariancy Region to the Adjacent Invariancy Regions

The first step of the algorithm is to determine the bounding box for the values of ε. Due to the fact that ε is the parameter appearing in the constraints, the problem $(QP_{\varepsilon,\lambda})$ may become infeasible for large or small ε values. Determining the bounding box is done as in many computational geometry algorithms [?,?]. To find the range of ε where the parametric problem $(QP_{\varepsilon,\lambda})$ is feasible, we solve the following problem starting from the initial point $(\lambda_0, \varepsilon_0)$:

$$\min\{c^T x + \frac{1}{2} x^T Qx \;:\; Ax = b + \varepsilon \triangle b^1 + \lambda_0 \triangle b^2,\; x \geq 0\}. \tag{24}$$

Solving problem (24) gives the values of $\varepsilon_{\min}$ and $\varepsilon_{\max}$ that (see Figure 3(a)) are the lower and the upper feasibility bounds for the bi-parametric problem $(QP_{\varepsilon,\lambda})$. Observe that we may have either $\varepsilon_{\min} = -\infty$ or $\varepsilon_{\max} = +\infty$.



(a) (b) (c)

Fig. 3: The Initialization of the Bi-Parametric QO Algorithm.

After identifying the feasibility bounds in the "$\varepsilon - \lambda$" plane, we choose $\varepsilon_{\min} \neq \infty$ or $\varepsilon_{\max} \neq \infty$. Let $\varepsilon = \varepsilon_{\min}$ and the optimal partition at the point $(\lambda_0, \varepsilon_{\min})$ is $\pi_{\min} = (\mathscr{B}_{\min}, \mathscr{N}_{\min}, \mathscr{T}_{\min})$. Then we can solve problems in Theorem 3 with the optimal partition $\pi = \pi_{\min}$ and $\lambda \triangle b^2$ replaced by $\varepsilon_{\min} \triangle b^2$ to identify the edge on the line $\varepsilon = \varepsilon_{\min}$ (see Figure 3(b)). If the point $(\lambda_0, \varepsilon_{\min})$ is a singleton, we find the invariancy interval to the right from it. Now, we have an edge of one of the invariancy regions and we can get an initial inner point of that invariancy region selecting a point on the edge and utilizing Algorithm 6.3 from [?]. Using that initial inner point, we can identify the first non-trivial invariancy region including all of its edges and vertices as described in subsection 3.2.1 (see Figure 3(c)).

To enumerate all invariancy regions in the bounding box, we use concepts and tools [?,?] from computational geometry. The algorithm that we are going to present possess some similarities with polygon subdivision of the space and planar graphs. Our algorithm is essentially the subdivision of the bounding box into convex polyhedrons that can be unbounded.

The geometric objects involved in the given problem are vertices, edges and cells (faces), see Figure 4. Cells correspond to the non-trivial invariancy regions. Edges and vertices are trivial invariancy regions, each edge connects two vertices. It is important to notice that cells can be unbounded if the corresponding invariancy region is unbounded. That is why we need to extend the representation of the vertex to allow incorporating the information that the vertex can represent the virtual endpoint of the unbounded edge if the corresponding cell is unbounded. For instance, edge e_1 in Figure 4 is unbounded, so in addition to its first endpoint v_1, we add another virtual endpoint being any point on the edge except v_1. Consequently, each vertex need to be represented not only by its coordinates (x,y), but also by the third coordinate z that indicates if it is a virtual vertex and the corresponding edge is unbounded. Another note to make is that the optimal partition may not be unique for each vertex or edge. First, at every virtual vertex, the optimal partition is the same as on the corresponding edge. Second, we may have situations when the optimal partition is the same on the incident edges and vertices if those are on the same line (edges e_2 and e_7 and vertex v_3 have the same optimal partition in Figure 4).

To enumerate all invariancy regions we use two queues that store indices of the cells that are already investigated and to be processed. At the start of the algorithm, the first cell enters the to-be-processed queue and the queue of completed cells is empty (c_1 is entering the to-be-processed queue in Figure 4). After that, we identify the cell c_1 including all faces and vertices starting from the known edge e_1 and moving counter-clockwise (note that the virtual vertices corresponding to the unbounded edges are not shown in Figure 4). We continue in

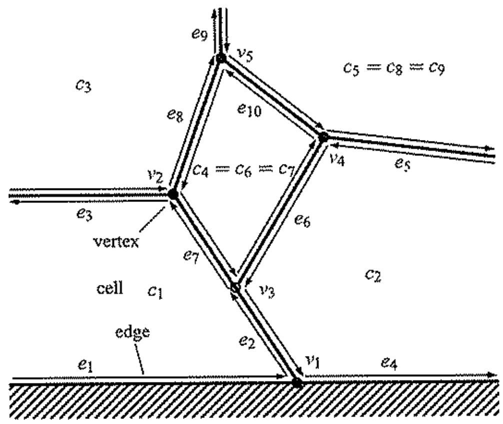


Fig. 4: Bi-Parametric QO – Computational Geometry Problem Representation.

that manner until to-be-processed queue is empty and we have identified all the invariancy regions.

Data: The CQO optimization problem and $\triangle b^1$, $\triangle b^2$
Result: Optimal partitions on all invariancy intervals, optimal value function
Initialization: compute bounding box in the "$\varepsilon - \lambda$" plane and compute inner point in one of the invariancy regions;
while *not all invariancy regions are enumerated* **do**
  run sub-algorithm to compute all edges and vertices of the current invariancy region;
  add all unexplored regions corresponding to each edge to the to-be-processed queue and move the current region to the queue of completed region indices;
  if *to-be-processed queue of the unexplored regions is not empty* **then**
    pull out the first region from the to-be-processed queue;
    compute an inner point of the new region;
  else
    return the data structure with all the invariancy regions, corresponding optimal partitions and optimal value function;
  end
end

Algorithm 1: Algorithm for Enumerating All Invariancy Regions

Algorithm 1 runs in linear time in the output size (the constant $C \cdot n$ is 3). But, by the nature of the parametric problem, the number of vertices, edges and faces can be exponential in the input size. In our experiences worst case does not happen in practise very often though.

Remark 3 Similar to uni-parametric case, it is easy to verify that the optimal value function on an invariancy interval is a quadratic function in terms of the two parameters ε and λ, and it fails to have first or second directional derivatives passing the boundary of an invariancy region. Moreover, the optimal value function is a convex piecewise quadratic function on the "$\varepsilon - \lambda$" plane.

Remark 4 As we already mentioned, considering $Q = 0$ reduces a QO problem to a LO problem. Consequently, the outlined Algorithm 1 works for LO problems without any modifications. In bi-parametric LO case, the optimal value function is a piecewise linear function in two parameters ε and λ, and it fails to have directional derivative passing a transition line separating two adjacent invariancy regions.

3.3 Discussions on Parametric Second-Order Conic Optimization

Parametric SOCO is a natural extension of parametric analysis for LO and QO. As we point out in Section 2.5, parametric SOCO allows solving multiobjective quadratic optimization problems with more than one quadratic objective. The *optimal basis approach* to parametric optimization in LO cannot be directly generalized to parametric optimization in SOCO [?]. In contrast, it is promising to generalize the *optimal partition approach* of parametric LO and QO to SOCO. We describe ideas and preliminary results related to parametric SOCO in this section.

The standard form SOCO problem is defined in Section 1. Primal problem (*SOCP*) and dual problem (*SOCD*) are specified by equation (6). Before defining parametric SOCO formally, we describe the geometry of second-order (quadratic) cones. An extensive review of SOCO problems can be found in [?].

Unlike LO and QO, in SOCO we work with blocks of primal and dual variables, see the definition of (*SOCP*) and (*SOCD*) problems in Section 1. Those primal-dual blocks (x^i, s^i), $i = 1, \ldots, I$ of variables compose the decision vectors of SOCO problems $x = (x^1, \ldots, x^I)^T$ and $s = (s^1, \ldots, s^I)^T$, where $x^i, s^i \in \mathbb{R}^{n_i}$. We also refer to cone $\mathscr{K}$ as a second-order cone when it is a product cone $\mathscr{K} = \mathscr{K}_q^1 \times \ldots \times \mathscr{K}_q^I$, where $x^i \in \mathscr{K}_q^i$, $i = 1, \ldots, I$. As a linear cone $\mathscr{K}_\ell^i$ is a one-dimensional quadratic cone $\mathscr{K}_q^i$ $(x_1^i \geq 0)$, we treat linear variables as one-dimensional blocks. As before, $\mathscr{K}^*$ is the dual cone of $\mathscr{K}$.

The bi-parametric SOCO problem in primal and dual form is expressed as:

$$
(SOCP_{\varepsilon,\lambda}) \qquad \begin{aligned} \phi(\varepsilon,\lambda) = \ & \min c^T x \\ & \text{s.t. } Ax = b + \varepsilon \triangle b^1 + \lambda \triangle b^2 \\ & \qquad x \in \mathscr{K}, \end{aligned} \tag{25}
$$

and

$$
(SOCD_{\varepsilon,\lambda}) \qquad \begin{aligned} & \max\ (b + \varepsilon \triangle b^1 + \lambda \triangle b^2)^T y \\ & \text{s.t. } A^T y + s = c \\ & \qquad s \in \mathscr{K}^*, \end{aligned} \tag{26}
$$

where $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = m$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ are fixed data; $x, s \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ are unknown vectors; $\lambda, \varepsilon \in \mathbb{R}$ are the perturbation parameters. Note that constraints $x \in \mathscr{K}$ and $s \in \mathscr{K}^*$ are replaced by $x_1^i \geq \|x_{2:n_i}^i\|$ and $s_1^i \geq \|s_{2:n_i}^i\|$, $i = 1, \ldots, I$ for computational purposes.

Algebraic representation of the optimal partition for SOCO problems is required for computational purposes. It will allow identification of invariancy intervals for parametric SOCO problems.

Yildirim [?] has introduced an optimal partition concept for conic optimization. He took a geometric approach in defining the optimal partition while using an algebraic approach is necessary for algorithm design. Although, the geometric approach has the advantage of being independent from the representation of the underlying optimization problem, it has some deficiencies. The major difficulty is extracting the optimal partition from a high-dimensional geometric object and, consequently, it is inconvenient for numerical calculations. In contrast, the algebraic approach is directly applicable for numerical implementation.

More recent study [?] provided the definition of the optimal partition that can be adapted to algebraic approach. We describe the details and compare the definitions of the optimal partition for SOCO in [?] (its algebraic form) and [?] in this section. Before defining the optimal partition for SOCO formally, we introduce the necessary concepts and notation. The interior and boundary of second-order cones are defined as follows.

Definition 5 The interior of second-order cone $\mathcal{K}_q \in \mathbb{R}^n$ is

$$\mathrm{int}\mathcal{K}_q = \{x \in \mathcal{K}_q : x_1 > \|x_{2:n}\|\}.$$

Definition 6 The boundary of second-order cone $\mathcal{K}_q \in \mathbb{R}^n$ without the origin 0 is

$$\mathrm{bd}\mathcal{K}_q = \{x \in \mathcal{K}_q : x_1 = \|x_{2:n}\|,\ x \neq 0\}.$$

Assuming strong duality, the *optimality conditions* for SOCO problems are:

$$\begin{aligned} Ax - b &= 0,\ x \in \mathcal{K}, \\ A^T y + s - c &= 0,\ s \in \mathcal{K}, \\ x \circ s &= 0, \end{aligned}$$

where the multiplication operation "$\circ$" is defined as $x \circ s = (x^1 \circ s^1, \ldots, x^I \circ s^I)^T$ and $x^i \circ s^i = ((x^i)^T s^i,\ x_1^i s_2^i + s_1^i x_2^i,\ \ldots,\ x_1^i s_{n_i}^i + s_1^i x_{n_i}^i)^T$.

Strict complementarity for SOCO problems [?] is defined as $x^i \circ s^i = 0$ and $x^i + s^i \in \mathrm{int}\mathcal{K}_q^i$, $i = 1, \ldots, I$. Interior point methods for SOCO produce maximally complementary solutions that maximize the number of strictly complementary blocks i.

With respect to its cone $\mathcal{K}_q^i$ each block x^i can be in one of three states:

- block x^i is in the interior of $\mathcal{K}_q^i$:

$$\mathrm{int}\,\mathcal{K}_q^i = \{x^i \in \mathcal{K}_q^i : x_1^i > \|x_{2:n_i}^i\|\},$$

- block x^i is on the boundary of $\mathcal{K}_q^i$:

$$\mathrm{bd}\,\mathcal{K}_q^i = \{x^i \in \mathcal{K}_q^i : x_1^i = \|x_{2:n_i}^i\| \,\mathrm{and}\, x^i \neq 0\},$$

- block x^i equals 0:

$$x^i = 0.$$

The same results are valid for the dual blocks of variables $s^i \in (\mathcal{K}_q^i)^*$. As second-order cones are self-dual $\mathcal{K} = \mathcal{K}^*$, we are going to denote both primal and dual cones by $\mathcal{K}$ in the remainder of this chapter.

The optimal partition for SOCO has four sets, so it is a 4-partition $\pi = (\mathcal{B}, \mathcal{N}, \mathcal{R}, \mathcal{T})$ of the index set $\{1, 2, \ldots, I\}$. The four subsets are defined in [?] as:

$$\begin{aligned} \mathcal{B} &= \{\, i : x_1^i > \|x_{2:n_i}^i\|\ (x^i \in \mathrm{int}\mathcal{K}_q^i) \text{ for a primal optimal solution } x \,\}, \\ \mathcal{N} &= \{\, i : s_1^i > \|s_{2:n_i}^i\|\ (s^i \in \mathrm{int}\mathcal{K}_q^i) \text{ for a dual optimal solution } (y, s) \,\}, \\ \mathcal{R} &= \{\, i : x^i \neq 0 \neq s^i\ (x^i \in \mathrm{bd}\mathcal{K}_q^i \text{ and } s^i \in \mathrm{bd}\mathcal{K}_q^i) \\ &\qquad \text{for a primal-dual optimal solution } (x, y, s) \,\}, \\ \mathcal{T} &= \{\, i : x^i = s^i = 0, \text{ or } s^i = 0 \text{ and } s^i \in \mathrm{bd}\mathcal{K}_q^i, \text{ or } s^i = 0 \text{ and } x^i \in \mathrm{bd}\mathcal{K}_i^q \\ &\qquad \text{for a primal-dual optimal solution } (x, y, s) \,\}. \end{aligned}$$

Now we can state all possible configurations for primal-dual blocks of variables at optimality, those are summarized in Table 1 and serve as a basis for defining the optimal partition. Cases that are not geometrically possible, as those do not satisfy the optimality conditions, are shown as "×" in Table 1.

Table 1: Optimal Partition for SOCO.

s^i \ x^i	0	$\text{bd}\,\mathscr{K}_q^i$	$\text{int}\,\mathscr{K}_q^i$
0	$i \in \mathscr{T}$	$i \in \mathscr{T}$	$i \in \mathscr{B}$
$\text{bd}\,\mathscr{K}_q^i$	$i \in \mathscr{T}$	$i \in \mathscr{R}$	×
$\text{int}\,\mathscr{K}_q^i$	$i \in \mathscr{N}$	×	×

For the set $\mathscr{R}$ of the optimal partition it holds that $x^i \neq 0 \neq s^i$, and those blocks x^i and s^i lie on the boundary of $\mathscr{K}$ (i.e., $x_1^i = \|x_{2:n_i}^i\| \neq 0$ and analogous relation holds for the dual). Let $(\bar{x}, \bar{y}, \bar{s})$ be a maximally complementary solution of problems $(SOCP)$ and $(SOCD)$ defined by (6), then as $x^i \circ s^i = 0$ we have

$$x^i \in \{\, \alpha \bar{x}^i : \alpha \geq 0 \,\},$$

$$s^i \in \{\, \beta(\bar{x}_1^i, -\bar{x}_{2:n_i}^i) : \beta \geq 0 \,\},$$

is equivalent to the primal and dual blocks belonging to orthogonal boundary rays of the cone $\mathscr{K}_q^i$.

We can replace the definition of the set $\mathscr{R}$ of the optimal partition by:

$$\mathscr{R}(\bar{x}) = \{\, (i, \bar{x}^i) : x^i \neq 0 \neq s^i,\ x^i \in \{\alpha \bar{x}^i : \alpha \geq 0\},\ s^i \in \{\beta(\bar{x}_1^i, -\bar{x}_{2:n_i}^i) : \beta \geq 0\} \text{ for a primal-dual optimal solution } (x, y, s) \,\}.$$

Based on the results of Yildirim [?], we can alternatively define the optimal partition in algebraic form as $\pi_r = (\mathscr{B}, \mathscr{N}, \mathscr{R}(\bar{x}), \mathscr{T})$. The difference from the definition of π is that for primal-dual boundary blocks, it holds that $x^i \in$ a specific boundary ray of $\mathscr{K}_q^i$ and $s^i \in$ the orthogonal boundary ray of $\mathscr{K}_q^i$, instead of $x^i \in \text{bd}\,\mathscr{K}_q^i$ and $s^i \in \text{bd}\,\mathscr{K}_q^i$.

Comparing the two definitions of the optimal partition, π and π_r, it is worth to mention a couple of differences. When the optimal partition is defined as π, it partitions the index set $\{1, 2, \ldots, I\}$ of the blocks of variables. Consequently, it directly extends the definition of the optimal partition for QO (see Section 3.1) by adding the additional set $\mathscr{R}$ that corresponds to primal-dual optimal solutions being on the boundary of the cone, i.e., the case that does not exist for QO. In contrast, when the optimal partition is defined as π_r, it partitions not only the index set $\{1, 2, \ldots, I\}$, but also the space, as the set $\mathscr{R}(\bar{x})$ includes both indices of the blocks i and vectors $\bar{x}^i$ that define specific boundary rays. Definition of the optimal partition π_r is similar to the definition of the optimal partition for semidefinite optimization (SDO) in [?], which partitions the space and not the index set. Note that the real meaning of the partition set $\mathscr{R}(\bar{x})$ is that primal and dual vectors should be on the boundary of the cone and belong to a specific ray on that boundary. If the optimal solution stays on the boundary, but moves to another boundary ray when the problem ($SOCP$) is perturbed, the optimal partition π_r changes, while π remains invariant.

Lets us consider the bi-parametric SOCO problem (25)-(26). We assume that the unperturbed problem ($SOCP_{0,0}$), where $\lambda = \varepsilon = 0$, has non-empty primal and dual optimal solution sets and strong duality holds for it, i.e., the duality gap is zero. For now, we use the definition π_r of the optimal partition.

Similar to parametric QO in Section 3.2, we can transform the bi-parametric SOCO problem into a series of uni-parametric problems. For simplicity, let us assign $\lambda = \varepsilon$. Moreover, let (x^*, y^*, s^*) be a maximally complementary optimal solution for $\varepsilon = 0$ with the optimal partition $\pi_r = (\mathscr{B}, \mathscr{N}, \mathscr{R}(x^*), \mathscr{T})$, the endpoints of the invariancy interval containing ε can be computed as:

$$\begin{aligned}\varepsilon_\ell = \min_{\varepsilon,x,y,s,\alpha,\beta} \{\, \varepsilon : & Ax - \varepsilon(\Delta b^1 + \Delta b^2) = b,\ x_{\mathscr{B}\cup\mathscr{T}} \in \mathscr{K}_{\mathscr{B}\cup\mathscr{T}},\ x_{\mathscr{N}} = 0,\ x_{\mathscr{R}} = \alpha x^*_{\mathscr{R}},\ \alpha \geq 0, \\ & A^T y + s = c,\ s_{\mathscr{N}\cup\mathscr{T}} \in \mathscr{K}_{\mathscr{N}\cup\mathscr{T}},\ s_{\mathscr{B}} = 0,\ s_{\mathscr{R}} = \beta s^*_{\mathscr{R}},\ \beta \geq 0 \,\},\end{aligned}$$

$$\begin{aligned}\varepsilon_u = \max_{\varepsilon,x,y,s,\alpha,\beta} \{\, \varepsilon : & Ax - \varepsilon(\Delta b^1 + \Delta b^2) = b,\ x_{\mathscr{B}\cup\mathscr{T}} \in \mathscr{K}_{\mathscr{B}\cup\mathscr{T}},\ x_{\mathscr{N}} = 0,\ x_{\mathscr{R}} = \alpha x^*_{\mathscr{R}},\ \alpha \geq 0, \\ & A^T y + s = c,\ s_{\mathscr{N}\cup\mathscr{T}} \in \mathscr{K}_{\mathscr{N}\cup\mathscr{T}},\ s_{\mathscr{B}} = 0,\ s_{\mathscr{R}} = \beta s^*_{\mathscr{R}},\ \beta \geq 0 \,\},\end{aligned}$$

where $\mathscr{K}_{\mathscr{B}\cup\mathscr{T}}$ is the Cartesian product of the cones $\mathscr{K}^i_q$ such that $i \in \mathscr{B}\cup\mathscr{T}$, $\mathscr{K}_{\mathscr{N}\cup\mathscr{T}}$ is defined analogously. Proof of this result for computing ε_ℓ and ε_u can be found in Theorem 4.1 in [?]. Alternatively, the constraints of the problems above can be completely rewritten in terms of the solution set instead of the index set, i.e., constraints $\{x_{\mathscr{B}\cup\mathscr{T}} \in \mathscr{K}, x_{\mathscr{N}} = 0, x_{\mathscr{R}} = \alpha x^*_{\mathscr{R}}, \alpha \geq 0\}$ can be written as $\{x \in \mathscr{K},\ x \circ s^* = 0\}$.

The optimization problems for computing the endpoints ε_ℓ and ε_u of the current invariancy interval are SOCO optimization problems due to the fact that constraints of the form $x_{\mathscr{R}} = \alpha x^*_{\mathscr{R}}$, $\alpha \geq 0$ are linear (the invariancy interval can be a singleton, unlike in the QO case). In contrast, if we use the definition of the optimal partition π, constraints $x_{\mathscr{R}} \in \mathrm{bd}\,\mathscr{K}$ are non-linear and are not second-order cone representable.

The results obtained by Yildirim [?] for the simultaneous perturbation case in conic optimization and by using the geometric definition of the optimal partition are directly linked to our findings. In his paper, Yildirim proved that the optimal value function is quadratic on the current invariancy interval. Although Yildirim's and our results are very interesting in the light of extending the parametric optimization techniques to SOCO problems, the obstacles, discussed in the remaining of this section, prevent direct mapping of them to algorithm design and implementation.

Unlike for parametric LO and QO problems, the optimal partition π_r for SOCO may change continuously, that poses difficulties for identifying all invariancy intervals for parametric SOCO. For the intervals of the parameter ε, where the optimal partition π_r is not changing continuously, the optimal value function is quadratic (see Proposition 5.1 in [?]). Another way to say it, for parametric SOCO we can have a continuum of changing transition points until we find an invariancy interval. In general, the optimal value function is piecewise-quadratic and it is quadratic on every invariancy interval. For the intervals, where the optimal partition changes continuously, we obtain the regions of non-linearity of $\phi(\varepsilon)$ and there is no known way of describing $\phi(\varepsilon)$ completely on those intervals.

The intervals where the optimal partition π_r changes continuously, represent a curve on the boundary of the quadratic cone. Similarly, if the optimal partition is defined as π, the intervals with $\mathscr{R} \neq \emptyset$ represent a curve on the quadratic cone surface. Characterization of those curves and finding a computable description of them will allow identifying all invariancy intervals and computing the optimal value function. While those curves are conjectured to have hyperbolic shape, there no results characterizing those curves that we are aware of. To get a computational algorithm for parametric SOCO, this characterization is a missing ingredient. Another remaining open problem is to find a rounding procedure for SOCO problems to identify exact optimal solutions.

Algorithms for computing the optimal value function $\phi(\varepsilon,\lambda)$ for parametric SOCO problems are subject of future research as there are no algorithms for parametric SOCO optimization. Invariancy regions corresponding to the definition of the optimal partition π are illustrated by an example in Section 4.3. That example also highlights the difficulties that arise during bi-parametric SOCO.

3.4 Parametric Optimization: Extensions

3.4.1 Multi-Parametric Optimization

In this section we discuss how ideas from uni- and bi-parametric optimization in Sections 3.1 and 3.2 extend to multi-parametric case. Some multi-parametric results exist for LO and QO from the optimal basis invariancy [?] and optimal partition invariancy [?]. Bi-parametric optimization algorithm from Section 3.2 can be extended to multi-parametric case as well.

We would like to mention some differences of our algorithmic approach to parametric QO optimization in Section 3.2 and the algorithm described in [?] which is implemented in [?]. First, in our study we allow simultaneous perturbation in the right-hand-side of the constraints and the linear term of the objective function with different parameters, while in [?] and related publications only perturbation in either the right-hand-side or the linear term of the objective is considered. Second, in [?] the authors define a critical region as the region of parameters where active constraints remain active. As the result, an important precondition for analysis in [?] is the requirement for either making non-degeneracy assumption or exploiting special tools for handling degeneracy, while, our algorithm does not require any non-degeneracy assumptions. Finally, the algorithm for parametric quadratic optimization described in [?] uses a different parameter space exploration strategy than ours. Their recursive algorithm identifies a first critical (invariancy) region, and after that reverses the defining hyperplanes one by one in a systematic process to get a subdivision of the complement set. The regions in the subdivision are explored recursively. As the result, each critical (invariancy) region can be split among many regions and, consequently, all the parts has to be detected. Thus, each of the potentially exponential number of invariancy regions may be split among exponential number of regions, which makes their algorithm computationally expensive.

3.4.2 Nonlinear Parametric Optimization

Let us consider convex non-linear parametric problem (11). When continuity of the functions $g_i(x,\lambda)$ for all (x,λ) and the convexity of these functions on $\mathbb{R}^n$ for all $\lambda \in \Lambda$ are added to the solution set (12), one can derive stronger results (see Section 3.2 in [?]).

This is the case we encounter in multiparametric LO and QO problems in some sense. With these assumptions, $\mathscr{M}$ is Hausdorff upper semi-continuous at λ_0 if $M(\lambda_0)$ is bounded and an $x_0 \in X$ exists such that $g(x_0) < \lambda_0$ ($g(x) = (g_1(x),\ldots,g_m(x))^T$ and λ_0 is an m-vector) (see Theorem 3.3.1 in [?]). It means that x_0 must be an interior point of the parametric solution set (12).

Moreover, for $X = \mathbb{R}^n$, if $\mathscr{M}(\lambda)$ is nonempty for all $\lambda \in \Lambda$ and $\mathscr{M}(\lambda_0)$ be affine subspace, then $\mathscr{M}$ is Hausdorff-continuous at λ_0 (see Theorem 3.3.3.2 in [?]). This is the case we have in multiparametric LO and QO problems when perturbation occurs in the right-hand-side of constraints.

4 Multiobjective Optimization Applications and Examples

In this section we present examples of multiobjective optimization problems that can be formulated and solved via parametric optimization. Multiobjective optimization problems arise in many areas including engineering (maximize vehicle speed and maximize its safety), finance (maximize profit and minimize risk), environmental economics (maximize profit and minimize environmental impact) and health care (kill tumor and spare healthy tissues). Examples described in this chapter are financial optimization problems from the area of risk management and portfolio selection. For examples of multiobjective optimization problems appearing in engineering we refer the reader to consult a vast literature on multi-disciplinary design [?]. Health care applications include Intensity Modulated Radiation Therapy (IMRT) planning for cancer treatment among others. For instance, a multiobjective linear IMRT problem is studied in [?], where the authors formulate an optimization problem with three objectives and compute an approximation of Pareto efficient surface.

In portfolio optimization, the goal of investors is to obtain optimal returns in all market environments when risk is involved in every investment, borrowing, lending and project planning activity. From the multicriteria analysis point of view, investors need to determine what fraction of their wealth to invest in which asset in order to maximize the total return and minimize the total risk of their portfolio. There are many risk measures used for quantitative evaluation of portfolio risk including variance, portfolio beta, Value-at-Risk (VaR) and expected shortfall (CVaR) among others. In addition to risk measures, there are portfolio performance indicators: expected market return, expected credit loss, price earnings ratio, etc. The most famous portfolio management model that involves a risk-return tradeoff is the mean-variance portfolio optimization problem introduced by Markowitz [?]. The conflicting objectives in the Markowitz model are minimizing portfolio variance (risk) and maximizing expected return.

Multiobjective optimization is a natural tool for portfolio selection models as those involve minimizing one or several risk measures, and maximizing a number of portfolio performance indicators. We describe three variants of multiobjective portfolio optimization problems and their corresponding parametric formulations:

1. Parametric LO (three linear objectives) in Section 4.1.
2. Parametric QO (two linear objectives and one quadratic objective) in Section 4.2.
3. Parametric SOCO (one linear objective, one quadratic objective and one second-order conic objective) in Section 4.3.

4.1 Portfolio Selection with Multiple Linear Objectives

Here, we discuss the multiobjective portfolio selection problem, where the objective functions are linear. Those models are rooted in the Capital Asset Pricing Model (CAPM), where the risk measure of an asset or portfolio is given by its beta coefficient. CAPM is the equilibrium version of mean-variance theory. Due to measuring risk in terms of the beta coefficients, the objective function in the risk minimization problem is linear in portfolio weights. In [?] a decision tool for the selection of stock portfolios based on multiobjective LO is described. Linear objective functions of the problem are the return, price earnings ratio, volume of transactions, dividend yield, increase in profits and risk, which is expressed as the linear function of betas. The authors apply portfolio selection to a set of fifty-two stocks from the Athens Stock Exchange. We are going to briefly describe their model including objective functions and constraints and compute the Pareto front for three out of six objectives considered in [?]. Readers interested in full details of the formulation and data for the model may consult [?].

The decision variables in portfolio selection problems are the portfolio weights x_i, $i = 1,\ldots,N$, where N is the total number of assets available for investment. Portfolio weights define a proportion of total wealth (or total budget) invested in the corresponding stock. As a matter of convenience, sum of portfolio weights is normalized to one $\sum_{i=1}^{N} x_i = 1$. Denoting by r_i the expected market return of an asset i, allows us to compute the portfolio market return as $r_P = \sum_{i=1}^{N} r_i x_i = r^T x$.

The beta coefficient β is a relative measure of systematic (non-diversifiable) risk, it reflects the tendency of an asset to move with the market. As beta measures correlation with the market portfolio, it is calculated as $\beta_i = \frac{\mathrm{Cov}(r_i, r_M)}{\mathrm{Var}(r_M)}$, where r_i is the asset i return and r_M is the return of the market portfolio. If $\beta_i < 1$ then asset i has less systematic risk than the overall market and the opposite holds for $\beta_i > 1$. As a result, portfolio risk minimization can be expressed as the linear function of asset weights, namely $\{\min_x \beta^T x\}$.

Among the other six objectives that are considered in [?] is maximizing return $\{\max_x r^T x\}$ and minimizing Price Earnings Ratio (P/E) $\{\min_x d^T x\}$. The Price Earnings Ratio d_i for each stock is computed as share price in the stock market at time period t divided by earnings per share at period $t-1$. We could have computed the Pareto efficient surface for more than three objectives here, but we restrict our attention to only those three due to well known difficulties with visualizing surfaces in more than 3 dimensions. Denoting the three objectives as $f_1 = -r^T x$, $f_2 = \beta^T x$ and $f_3 = d^T x$, we obtain the following parametric optimization problem:

$$\begin{array}{l} \min\ -r^T x + \lambda_1 \beta^T x + \lambda_2 d^T x \\ \text{s.t.}\ \ x \in \Omega, \end{array} \tag{27}$$

where Ω in [?] is the set of linear constraints that includes no-short-sales restriction $x \geq 0$; upper limits for the capital allocations $x_i \leq u_i$, $i = 1,\ldots,52$; specific preferences for some stocks of the form $x_j \geq l_i$; and the constraints on betas of the form that portion y of the capital will be allocated to stocks with $\beta \in \{\beta_1, \beta_2\}$ that are expressed as $\sum_{i \subseteq I} x_i = y$. Note that maximizing $r^T x$ is equivalent to minimizing $-r^T x$.

The parametric optimization problem that follows from the ε-constrained multiobjective formulation is the following:

$$\begin{array}{rl} \min\limits_{x,t} & -r^T x \\ \text{s.t.} & \beta^T x + t_1 = \varepsilon_1 \\ & d^T x + t_2 = \varepsilon_2 \\ & \sum_i x_i = 1 \\ & \sum_{i \in I} x_i = 0.2 \\ & x \geq 0,\ t \geq 0, \end{array} \tag{28}$$

where t_1, t_2 are the slack variables used to convert the linear inequality constrains into equality constraints and $\varepsilon = (\varepsilon_1, \varepsilon_2)^T$ is the vector of parameters. We have used a subset of the

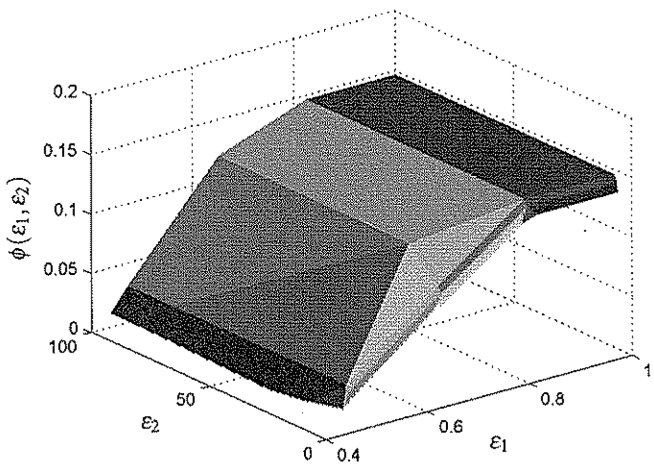


Fig. 5: The Optimal Value Function for the Parametric Linear Portfolio Optimization Problem.

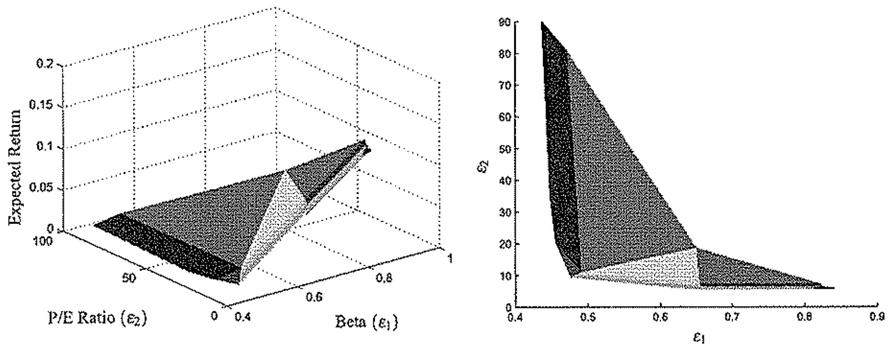
constraints $x \in \Omega$ from [?] for the ease of exposition and included the no short-sales constraint $x \geq 0$ and the constraint $\sum_{i \in J} x_i = 0.2$ stating that 20% of capital is allocated to stocks with a beta coefficient less than 0.5. Formulation (28) is parametric LO problem with two parameters in the right-hand-side of the constraints.

The optimal value function for problem (28) is shown in Figure 5. We can use the optimal partition for the variables t_1 and t_2 to determine the Pareto-efficient surface. For the invariancy regions corresponding to Pareto-efficient solutions, $t_1 \in \mathscr{N}$ and $t_2 \in \mathscr{N}$, meaning that those variables belong to the subset $\mathscr{N}$ of the optimal partition. The invariancy regions corresponding to the Pareto efficient solutions are shown in Figure 6(b) and the Pareto front is depicted in Figure 6(a). The Pareto front is a piecewise linear function. The knowledge of invariancy intervals and optimal value function on those intervals gives us the structure of the Pareto front.

4.2 Mean-Variance Optimization with Market Risk and Transaction Cost

The Markowitz mean-variance model is commonly used in practice in the presence of market risk. From an optimization perspective, minimizing variance requires solving a QO problem. Denoting a vector of expected market returns by r as before and a variance-covariance matrix of returns by Q, the mean-variance portfolio optimization problem is formulated as a QO problem where the objectives are to maximize the expected portfolio return $\{\max_x r^T x\}$ and to minimize variance $\{\min_x x^T Q x\}$. The multiobjective optimization problem can be formulated as the weighted sum problem

$$\begin{array}{ll} \min_x & -\lambda r^T x + \frac{1}{2} x^T Q x \\ \text{s.t.} & x \in \Omega, \end{array} \tag{29}$$



(a) (b)

Fig. 6: The Pareto Front for the Multiobjective Linear Portfolio Optimization Problem (a) and the Invariancy Regions Corresponding to It (b).

or as the ε-constrained problem

$$\begin{array}{rl} \min_x & \frac{1}{2}x^T Q x \\ \text{s.t.} & -r^T x \leq \varepsilon, \\ & x \in \Omega, \end{array} \tag{30}$$

where Ω is the set of linear constraints on portfolio weights.

A portfolio may incur transaction cost associated with each trading. Denoting the linear transaction cost by ℓ_i, we add the third objective of minimizing the trading cost $\ell^T x$ of a portfolio to the mean-variance portfolio optimization problems (29)-(30).

We use a small portfolio optimization problem to illustrate the multiobjective optimization methodology. The problem data is presented in Tables 2 and 3. Table 2 shows expected market returns per unit transaction cost for 8 securities, as well as their weights in the initial portfolio x_0.

We put non-negativity bounds $x \geq 0$ on the weights disallowing short-sales and optimize three objectives:

1) minimize the variance of returns;
2) maximize expected market return;
3) minimize transaction cost.

Moreover, we also need to add a constraint that makes the sum of the weights equal to one.

Thus, the multiobjective portfolio optimization problem looks like:

$$\begin{array}{rl} \min & f_1(x) = -r^T x,\ f_2(x) = \ell^T x,\ f_3(x) = \frac{1}{2}x^T Q x \\ \text{s.t} & \sum_i x_i = 1, \\ & x_i \geq 0\ \forall i. \end{array} \tag{31}$$

Table 2: Portfolio Data for Mean-Variance Optimization with Market Risk and Transaction Cost.

Security	x_0	$r = \mathbb{E}$(Market Return)	$\ell =$ (Transaction Cost)
1	0	0.095069	0.009830
2	0.44	0.091222	0.005527
3	0.18	0.140161	0.004001
4	0	0.050558	0.001988
5	0	0.079741	0.006252
6	0.18	0.054916	0.000099
7	0.13	0.119318	0.003759
8	0.07	0.115011	0.007334

Table 3: The Return Covariance Matrix Q for Mean-Variance Optimization with Market Risk and Transaction Cost.

Security	1	2	3	4	5	6	7	8
1	0.002812	0.002705	-0.001199	0.000745	-0.000064	0.001035	-0.000336	0.000178
2	0.002705	0.015664	-0.003000	0.001761	-0.002282	0.007129	0.000596	-0.003398
3	-0.001199	-0.003000	0.008842	-0.000155	0.003912	0.001424	0.001183	-0.001710
4	0.000745	0.001761	-0.000155	0.002824	0.001043	0.003874	0.000225	-0.001521
5	-0.000064	-0.002282	0.003912	0.001043	0.007213	-0.001946	0.001722	0.001199
6	0.001035	0.007129	0.001424	0.003874	-0.001946	0.013193	0.001925	-0.004506
7	-0.000336	0.000596	0.001183	0.000225	0.001722	0.001925	0.002303	-0.000213
8	0.000178	-0.003398	-0.001710	-0.001521	0.001199	-0.004506	-0.000213	0.006288

We solve problem (31) as a parametric problem corresponding to the ε-constraint multiobjective formulation:

$$\begin{array}{l} \min\limits_{x,t} \frac{1}{2}x^T Q x \\ \text{s.t} \ -r^T x + t_1 = \varepsilon_1 \\ \quad \ell^T x + t_2 = \varepsilon_2 \\ \quad \sum_i x_i = 1, \\ \quad x \geq 0, \ t \geq 0, \end{array} \tag{32}$$

where t_1, t_2 are the slack variables used to convert the linear inequality constrains into equality constraints and $\varepsilon = (\varepsilon_1, \varepsilon_2)^T$ is the vector of parameters.

The optimal value function for problem (32) is shown in Figure 7 and the corresponding invariancy regions – in Figure 8(a). We can utilize the optimal partition for the variables t_1 and t_2 to determine the Pareto efficient surface. For the invariancy regions corresponding to Pareto efficient solutions, $t_1 \neq \mathscr{B}$ and $t_2 \neq \mathscr{B}$, meaning that those variables do not belong to the subset $\mathscr{B}$ of the optimal partition. The invariancy regions corresponding to Pareto efficient solutions are shown in Figure 8(b) and the Pareto front is depicted in Figure 9.

Invariancy regions have a very intuitive interpretation for portfolio managers and financial analysts as inside each invariancy region the portfolio composition is fixed. By fixed composition we mean that the pool of assets included in the portfolio remains unchanged while asset weights can vary. For instance, on the invariancy region π_1 in Figure 8(b) the optimal partition is $\mathscr{N}\mathscr{N}\mathscr{B}\mathscr{B}\mathscr{N}\mathscr{B}\mathscr{B}\mathscr{N}$ which means that the portfolio is composed of

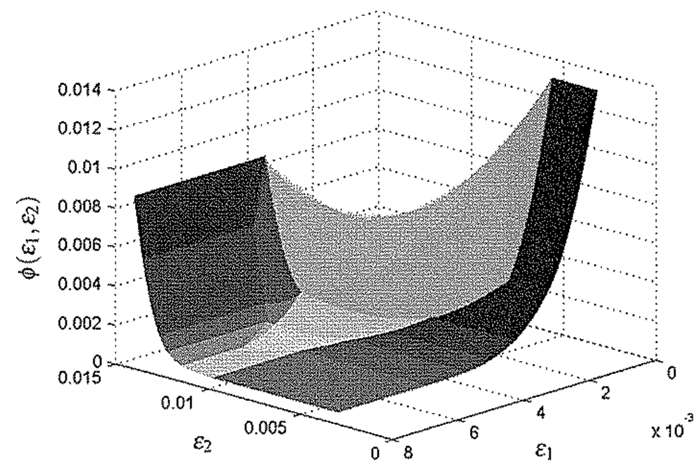


Fig. 7: The Optimal Value Function for the Mean-Variance Portfolio Problem in the Presence of Transaction Cost.

securities 3, 4, 6 and 7. The functional form of the Pareto front on the invariancy region π_1 is $f_3 = 0.1 - 0.4f_1 - 23.7f_2 + 13.4f_1^2 + 11999.4f_2^2 - 621.9f_1f_2$.

4.3 Robust Mean-Variance Optimization

One of the common criticisms of mean-variance optimization is its sensitivity to return estimates. As the consequence of that fact, small changes in the return estimates can result in big shifts of the portfolio weights x. One of the solutions to this problem is *robust optimization*, which incorporates uncertainties into the optimization problem. For a review of the robust optimization applied to portfolio management we refer the reader to [?].

We consider a variant of robust portfolio selection problems proposed by Ceria and Stubbs [?]. In their model, instead of the uncertainty set being given in terms of bounds, they use ellipsoidal uncertainty sets. In [?] the authors assume that only r, the vector of estimated expected returns, is uncertain in the Markowitz model (29). In order to consider the worst case of problem (29), it was assumed that the vector of true expected returns r is normally distributed and lies in the ellipsoidal set:

$$(r - \hat{r})^T \Theta^{-1} (r - \hat{r}) \leq \kappa^2,$$

where $\hat{r}$ is an estimate of the expected return, Θ is covariance matrix of the estimates of expected returns with probability η, and $\kappa^2 = \chi_N^2(1-\eta)$ with χ_N^2 being the inverse cumulative distribution function of the chi-squared distribution with N degrees of freedom.

Let $\hat{x}$ be the optimal portfolio on the estimated frontier for a given target risk level. Then, the worst case (maximal difference between the estimated expected return and the actual expected return) of the estimated expected returns with the given portfolio $\hat{x}$ can be

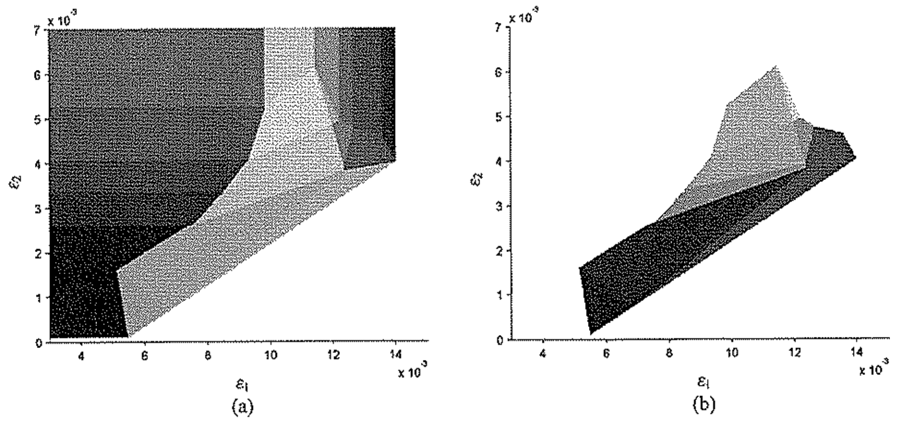


Fig. 8: Invariancy Regions (a) and Invariancy Regions Corresponding to the Pareto Efficient Solutions (b) for the Mean-Variance Portfolio Optimization Problem with Transaction Cost.

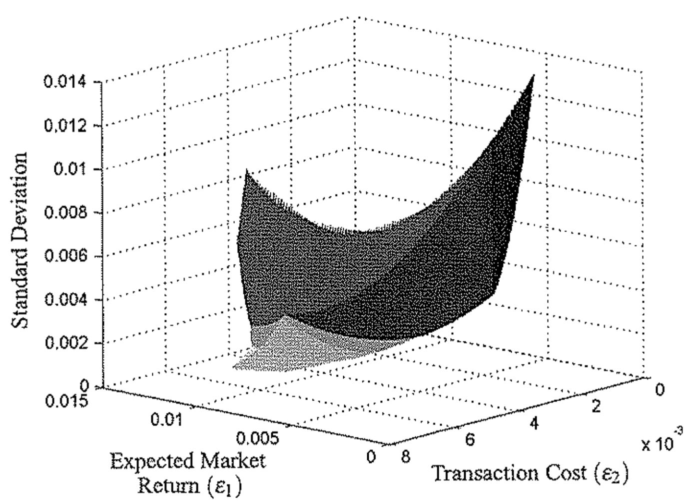


Fig. 9: The Pareto Efficient Surface for the Mean-Variance Portfolio Optimization Problem with Transaction Cost.

formulated as:

$$\begin{aligned} \max_{\hat{r}-r} \ & (\hat{r}-r)^T \hat{x} \\ \text{s.t.} \ & (r-\hat{r})^T \Theta^{-1} (r-\hat{r}) \leq \kappa^2. \end{aligned} \tag{33}$$

As derived in [?], by solving problem (33) we get that the optimal objective value $(\hat{r}-r)^T\hat{x}$ is $\kappa\|\Theta^{1/2}\hat{x}\|$. So, the true expected return of the portfolio can be expressed as $r^T\hat{x} = \hat{r}^T\hat{x} - \kappa\|\Theta^{1/2}\hat{x}\|$.

Table 4: Expected Returns and Standard Deviations with Correlations = 20% for Robust Mean-Variance Optimization, Optimal Weights for Two Portfolios.

Security	r^1	r^2	σ
Asset 1	7.15%	7.16%	20%
Asset 2	7.16%	7.15%	24%
Asset 3	7.00%	7.00%	28%

Security	Portfolio A	Portfolio B
Asset 1	38.1%	84.3%
Asset 2	69.1%	15.7%
Asset 3	0.0%	0.0%

Now, problem (29) becomes a robust portfolio selection problem

$$\begin{array}{l} \min_x \; -\lambda \hat{r}^T x + \frac{1}{2} x^T Q x + \kappa \|\Theta^{1/2} x\|, \\ \text{s.t. } x \in \Omega. \end{array} \tag{34}$$

Problem (34) is SOCO problem, moreover, it is a parametric optimization problem. We solve an instance of problem (34) rewriting its formulation as:

$$\begin{array}{ll} \min & -\hat{r}^T x + \kappa \|\Theta^{1/2} x\| + \lambda x^T Q x \\ \text{s.t.} & \sum_{i=1}^n x_i = 1 \\ & x \geq 0, \end{array} \tag{35}$$

where $\hat{r}$ is the vector of expected returns, Θ is the covariance matrix of estimated expected returns, Q is the covariance matrix of returns, κ is the estimation error aversion, and λ is the risk aversion.

Formulation (35) is a parametric SOCO problem with two parameters κ and λ. Preliminary results on parametric SOCO are discussed in Section 3.3. If we look at it in the multi-objective sense, it is the problem of maximizing expected return, minimizing risk (variance of returns) and minimizing estimation error for the expected return. The problem formulation emphasizes the differences between the true, the estimated, and the actual Markowitz efficient frontiers [?].

To demonstrate the influence that sensitivities in the return estimates can potentially have on the portfolio selection, Ceria [?] presented a simple portfolio consisting of three assets. Table 4 shows expected returns for the two estimates and standard deviations for the assets. As Table 4 also shows, completely different portfolio weights can be obtained while optimizing the portfolio with expected return estimates r^1 and r^2. Taking r^1 as the estimate of the expected returns, we solve the multiobjective problem (35) to find all possible trade-offs between the three objectives – maximizing expected return, minimizing variance and minimizing estimation error.

As $x^T Q x \leq \sigma_1^2$ ($Q = RR^T$) and $\|\Theta^{1/2} x\| = \sqrt{x^T \Theta x} \leq \sigma_2$, we can rewrite problem (35) in the form:

$$\begin{array}{ll} \min & -\hat{r}^T x + \lambda_1 u_0 + \lambda_2 v_0 \\ \text{s.t.} & \sum_{i=1}^n x_i = 1 \\ & x \geq 0 \\ & \Theta^{1/2} x - u = 0 \\ & R^T x - v = 0 \\ & (u_0, u) \in \mathcal{K}_q, \; (v_0, v) \in \mathcal{K}_q, \end{array} \tag{36}$$

where parameters $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ and $\mathcal{K}_q$ is the second-order cone. Parametric problem (36) represents the weighting method for multiobjective optimization.

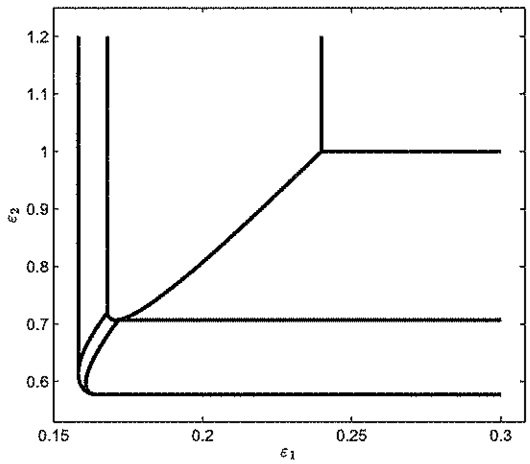


Fig. 10: The Invariancy Regions for the Robust Mean-Variance Portfolio Optimization Problem.

Formulating the parametric problem corresponding to the ε-constrained method for multiobjective optimization we get:

$$
\begin{array}{ll}
\min & -\hat{r}^T x \\
\text{s.t.} & \sum_{i=1}^{n} x_i = 1 \\
& x \geq 0 \\
& \Theta^{1/2} x - u = 0 \\
& R^T x - v = 0 \\
& u_0 = \varepsilon_1 \\
& v_0 = \varepsilon_2 \\
& (u_0, u) \in \mathscr{K}_q,\ (v_0, v) \in \mathscr{K}_q,
\end{array}
\tag{37}
$$

where parameters $\varepsilon_1 \geq 0$ and $\varepsilon_2 \geq 0$, and Θ is the identity matrix.

The optimal value function of the parametric SOCO formulation (37) with parameters $(\varepsilon_1, \varepsilon_2)$ in the constraints is shown in Figure 11. The corresponding invariancy regions are presented by Figure 10. To identify the invariancy regions that correspond to Pareto efficient solutions we need to restrict our attention to the regions where the second order conic blocks u and v belong to the subsets $\mathscr{R}$ or $\mathscr{T}$ of the optimal partition. Those invariancy regions and the corresponding Pareto efficient surface is shown in Figure 12.

5 Conclusions and Future Directions

In this paper we considered techniques for solving multiobjective optimization problems and their parametric counterparts. By formulating and solving *multiobjective optimization* problems as *parametric optimization* problems we bridged the gap between the two fields and unified the theory and practice of multiobjective and parametric optimization. Some classes of multiobjective optimization problems that include linear, convex quadratic and

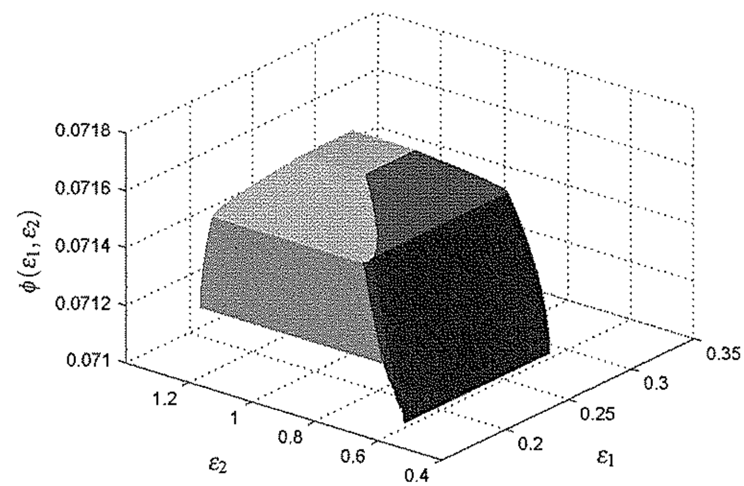


Fig. 11: The Optimal Value Function for the Robust Mean-Variance Portfolio Optimization Problem.

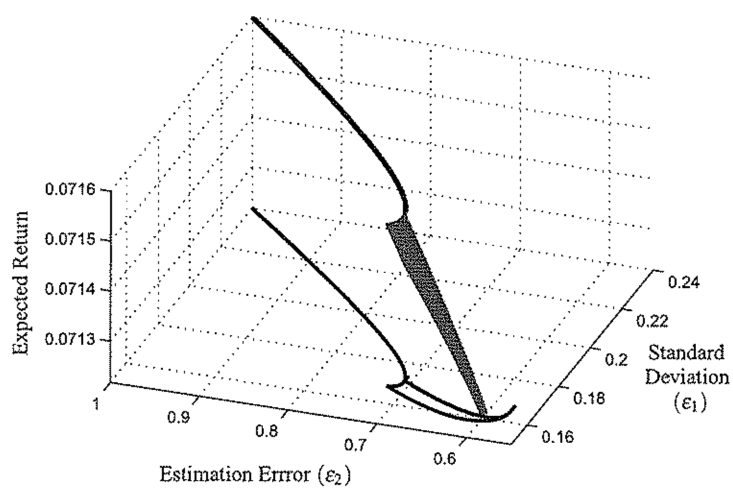


Fig. 12: The Pareto Efficient Surface for the Robust Mean-Variance Portfolio Optimization Problem.

potentially second-order conic optimization problems can be efficiently solved using parametric optimization algorithms. In particular, parametric optimization techniques described in this paper give us a practical tool for solving multiobjective quadratic optimization problems. Parametric optimization allows not only computing Pareto fronts (efficient surfaces), but also identifying piece-wise structure of those surfaces. Structural description of Pareto fronts gives functional form of each of its pieces and thus helps decision makers to make better decisions.

Even though some techniques exist for solving convex non-linear parametric problems, those are not widely used. So, solving multiobjective convex non-linear problems in practice is one of the hot research areas. If a multiobjective problem is non-convex (i.e, mixed integer), different approximations can be used allowing tracing Pareto efficient frontier with parametric optimization [?].

Integration of parametric optimization techniques that use optimal bases, optimal set invariancy and optimal partition invariancy into a unified framework remains to be done. There are many publications that address different aspects of parametric optimization, but there is no study that puts those techniques together and describe how well those perform for different classes of optimization problems. Additional work has to be done on classifying multiobjective optimization problems for which the Pareto efficient frontier has identifiable structure.

Implementing parametric optimization into optimization software packages remains one of the challenges. Unfortunately, available software for parametric optimization is very limited. Commercial optimization packages such as CPLEX [?] and MOSEK [?] include basic sensitivity analysis for LO that is based on an optimal basis. MOSEK is the only package that provides optimal partition based sensitivity analysis for LO as an experimental feature. As parametric optimization is the generalization of sensitivity analysis, techniques for identifying invariancy and stability regions have to be implemented on the top of sensitivity analysis available in those packages. Experimentation with active-set based multi-parametric optimization for LO and QO can be performed with MPT (Multi-Parametric Toolbox for MATLAB) [?] and this toolbox can be called from the YALMIP modeling environment [?].

Acknowledgements The authors' research was partially supported by the NSERC Discovery Grant #48923, the Canada Research Chair Program, grant from Lehigh University and MITACS. We are grateful to Helmut Mausser, Alexander Kreinin and Ian Iscoe from Algorithmics Inc. for valuable discussions on the practical examples of multiobjective optimization problems in finance. We would like to thank Antoine Deza from McMaster University, Imre Pólik from Lehigh University and Yuri Zinchenko from University of Calgary for their helpful suggestions.