Sensitivity Analysis and Robust Optimization

A Geometric Approach for the Special Case of Linear Optimization

by

Jiyoung Im

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Combinatorics & Optimization

Waterloo, Ontario, Canada, 2018

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, we study the special case of linear optimization to show what may affect the sensitivity of the optimal value function under data uncertainty. In this special case, we show that the robust optimization problem with a locally *smaller* feasible region yields a more conservative robust optimal value than the one with a locally *bigger* feasible region. To achieve that goal, we use a geometric approach to analyze the sensitivity of the optimal value function for linear programming (**LP**) under data uncertainty. We construct a family of proper cones where the strict containment holds for any pair of cones in the family. We then form a family of **LP** problems using this family of cones constructed above; the feasible regions of each pair of **LPs** in the family holds strict containment, every **LP** in the family has the unique optimal solution at the vertex of the cone and has the same objective function, i.e., every **LP** in the family shares the same optimal solution and the same optimal value. We rewrite the **LPs** so that they reflect the given data uncertainty and perform local analysis near the optimal solutions where the local strict containment holds. Finally, we illustrate that an **LP** with a locally *smaller* feasible region is more sensitive than an **LP** with a locally *bigger* feasible region.

Acknowledgements

I would first love to express my deepest gratitude to my advisor, Henry Wolkowicz. Having him as my advisor was the luckiest thing that ever happened to me throughout my academic career. He is the most patient person I have ever met. He never laughed at my silly questions or answers. He consistently allowed this thesis to be my own work, but guided me in the right direction and threw the right questions whenever he thought I needed.

I would also like to express my gratitude to my family for providing me with continuous encouragement throughout my study. I especially want to express my love to my significant other, William Yeung. He is always there for me through ups and downs. No words can describe his unlimited emotional support and countless meals he has cooked for me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Suppose that we want to solve an uncertain linear programming problem (**LP**), i.e., a linear program with uncertain data; and, we are given bounds for each element of the uncertain data. We want to obtain the best possible optimal value in the presence of data uncertainty. *Robust optimization* handles such linear programming problems. Robust optimization tries to find the best uncertainty(perturbation)-immunized solution with given uncertainty set. Given a linear programming problem, we call an optimal value with no uncertainty involved a *nominal optimal value* and call an optimal value with uncertainty involved a *robust optimal value*.

We sometimes obtain a robust optimal value that is very close to its nominal optimal value. In this case, data uncertainty does not play a big role in terms of determining its optimal value. In this thesis, we want to study the question of *why* a robust optimal value may be so well behaved. Similarly, we may obtain a robust optimal value that is far from its nominal optimal value and want to answer the question of *why* a robust optimal value may be poor, more precisely why it is too *conservative*. We show that the behaviour of robust optimal value is related to the *local geometric structure* of the problem. For this, we construct a family of **LPs** that share the same objective function and the feasible regions are of the same structure. At the same time, we have control over the sizes of the feasible regions.

We achieve the goal of this thesis via sensitivity analysis. *Sensitivity analysis* tries to answer how sensitive the optimal value/solution is to small changes in one or more of the parameters/data of the original problem. We first consider an uncertain linear programming and formulate its robust counterpart. We then show that the resulting robust counterpart can be written as a parametric **LP**. We analyze the sensitivity of the optimal value function of the parametric form of the robust counterpart. Intuitively speaking, we want to show that if a nominal optimal solution is determined at a sharp corner of the feasible region, then its robust optimal value is more sensitive than the case where the nominal optimal solution is at a fatter corner.

The topic of this thesis lies in the intersection of the polyhedral theory, robust optimization, parametric linear programming and sensitivity analysis.

This thesis is organized as follows: For the rest of this chapter, we introduce an elementary example and some notations. In Chapter 2, we introduce some definitions and lemmas that we need in the later chapters. In particular, the polyhedral theory and the robust optimization will be extensively used in the later chapters. In Chapter 3, we present how we build polyhedral cones so that we can control their sizes. We then study properties of the cones constructed above. We also study relations between two sets based on the cones generated above. In Chapter 4, we present the main results of this thesis. We first define two classes of **LPs** we wish to study their sensitivities. We consider two classes of **LPs** such that the part of their feasible regions are the cones constructed in Chapter 3. In Chapter 5, we experiment the result presented in Chapter 4. In Chapter 6, we conclude the results presented in this thesis, limitations and further work.

1.1 An Elementary Example

We observe the following example.

Example 1.1.1. *Define the following two families of* ***LPs*** *where* E *is the matrix of ones and* $\epsilon \in [0, 0.1]$:

$$(\mathcal{R}(\epsilon, P)) \qquad \begin{array}{lll} \psi_P(\epsilon) = & \min & \langle \bar{w}, x \rangle \\ & \textit{subject to} & (P - \epsilon E)x \geq p \\ & & x \geq 0, \end{array}$$

$$(\mathcal{R}(\epsilon, Q)) \qquad \begin{array}{lll} \psi_Q(\epsilon) = & \min & \langle \bar{w}, x \rangle \\ & \textit{subject to} & (Q - \epsilon E)x \geq q \\ & & x \geq 0, \end{array}$$

where

$$\bar{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, P = \begin{bmatrix} -\frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}, Q = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}, p = \begin{pmatrix} \frac{1-\sqrt{3}}{2} \\ \frac{1+\sqrt{3}}{2} \end{pmatrix}, q = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}.$$

Figure 1.1.1a shows the feasible regions of $(\mathcal{R}(0, P))$ *and* $(\mathcal{R}(0, Q))$. *We note that the feasible region of* $(\mathcal{R}(0, Q))$ *contains the one of* $(\mathcal{R}(0, P))$.

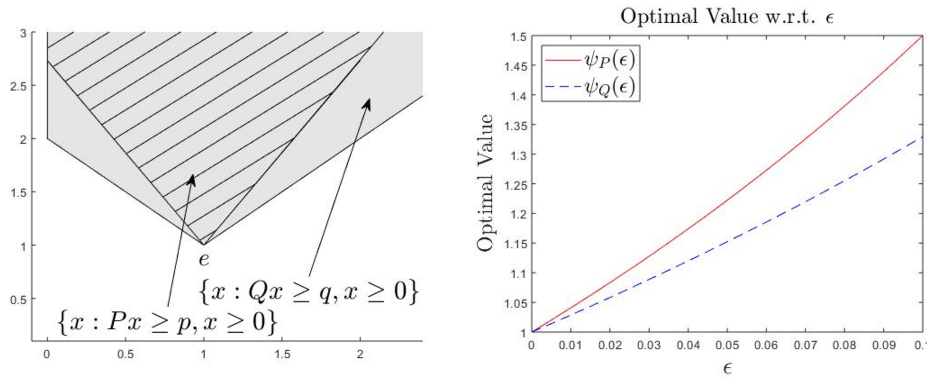
$(\mathcal{R}(0, P))$ *and* $(\mathcal{R}(0, Q))$ *both have the same optimal value and the same optimal solution, which are* 1 *and* e, *respectively. However, for* $\epsilon \in (0, 0.1]$, *we have the following relation:*

$$\psi_P(\epsilon) > \psi_Q(\epsilon). \tag{1.1.1}$$

For example, if $\epsilon_1 = 10^{-10}$ *and* $\epsilon_2 = 10^{-2}$, *we have*

$$\begin{array}{rcl} \psi_P(\epsilon_1) \approx 1 + 4 \cdot 10^{-10} & > & 1 + 2.828 \cdot 10^{-10} \approx \psi_Q(\epsilon_1), \ \textit{and} \\ \psi_P(\epsilon_2) \approx 1 + 5 \cdot 10^{-1} & > & 1 + 3.29 \cdot 10^{-1} \approx \psi_Q(\epsilon_2). \end{array}$$

Figure 1.1.1b shows the computational result of (1.1.1). In this example, we observe that



(a) The feasible region of $(\mathcal{R}(0, P))$ contains the one of $(\mathcal{R}(0, Q))$. (b) $\psi_P(\epsilon) > \psi_Q(\epsilon)$, $\forall \epsilon \in (0, 0.1]$

Figure 1.1.1: The optimal value function $\psi_P(\epsilon)$ is associated with the region filled with diagonal lines on the LHS. We note that $\psi_P(\epsilon)$ yields bigger function values than $\psi_Q(\epsilon)$.

the optimal value of $(\mathcal{R}(\epsilon, P))$ *is more sensitive to its change of data than the optimal value of* $(\mathcal{R}(\epsilon, Q))$. $\square$

We relate the sensitivity of the optimal value functions to the dual optimal solutions of the given **LPs**. We will show in Chapter 4 that $\psi_P(\epsilon)$ is more sensitive than $\psi_Q(\epsilon)$, since the nonzero coordinates of the dual optimal solutions of $(\mathcal{R}(\epsilon, P))$ is bigger than the ones of $(\mathcal{R}(\epsilon, Q))$.

1.2 Notations

We first explain some notations used in this thesis. We use these notations without further explanations in later chapters. We extensively use the definitions from [1, 10, 25].

To represent a part of a matrix, we follow Matlab notations. For example, given a matrix A, $A(:, j)$ denotes the j-th column of a matrix A. Similarly, $A(i, :)$ denotes the i-th row of a matrix A. Given a subset of column indices of A, $A(:, \mathcal{I})$ denotes a submatrix A' of A such that columns of A' are columns of A associated with $\mathcal{I}$. The rank of A is denoted by $\operatorname{rank}(A)$.

Given a set X, we denote $X^\perp$ the orthogonal complement of X; $\operatorname{int}(X)$ denotes the interior of X and $\operatorname{relint}(X)$ denotes the relative interior of X. The null space of X is denoted by $\operatorname{null}(X)$.

We use superscript to denote various vectors. For example, we write w^i to denote different vectors in $\mathbb{R}^n$. However, we use subscript w_i, when the meaning is clear. We use e^n to denote the vector of all one's in $\mathbb{R}^n$. However, we often omit the superscript when the meaning is clear; e_i denotes the i-th column of the identity matrix; E is the matrix of all ones with an appropriate size; and $\operatorname{Ball}(x, \epsilon)$ is the ball centered at x with radius ϵ.

Chapter 2

Preliminaries

2.1 Convex Analysis Background

We present some of the background in convex analysis and polyhedral theory needed in subsequent chapters. In Section 2.1.1, we introduce basic notions of convex analysis and some related results. In Section 2.1.2, we present basic definitions of general cones and some examples. In Section 2.1.3, we present some important lemmas in polyhedral theory used in later chapters of this thesis.

2.1.1 Convex Sets

A reader familiar with the basics of convex sets and extreme points can skip to the next section.

Definition 2.1.1 (*convex combination*). *A point x in $\mathbb{R}^n$ is a* convex combination *of the points $x_1, \dots, x_k \in \mathbb{R}^n$, if there exist nonnegative scalars $\lambda_1, \dots, \lambda_k$ such that*

$$x = \sum_{i=1}^{k} \lambda_i x_i \text{ and } \sum_{i=1}^{k} \lambda_i = 1.$$

If all scalars $\lambda_1, \dots, \lambda_k$ are positive, then we call x is a strict convex combination *of the points $x_1, \dots, x_k$.*

Definition 2.1.2 (*convex set*). *A set $X \subseteq \mathbb{R}^n$ is* convex *if X contains all convex combinations of points in X. Equivalently, $X \subseteq \mathbb{R}^n$ is* convex *if for any two points $x_1, x_2 \in X$, the line segment $\{\lambda x_1 + (1-\lambda)x_2 : \lambda \in [0,1]\}$ with endpoints x_1, x_2 is contained in X.*

Definition 2.1.3 (*convex hull*). *Given a set $X \subseteq \mathbb{R}^n$, the* convex hull *of X, denoted by* $\operatorname{conv}(X)$*, is the smallest convex set containing X.*

Definition 2.1.4 (*extreme point*)**.** *We say that $x \in X$ is an* extreme point *of X if*

$$x = \alpha x_1 + (1-\alpha)x_2 \text{ with } \alpha \in (0,1) \text{ and } x_1, x_2 \in X \implies x = x_1 = x_2.$$

We let $\operatorname{ext}(X)$ *denote the set of all extreme points of X.*

Lemma 2.1.5 ([28, Theorem 8.11])**.** [1] *Let X be a compact convex set in $\mathbb{R}^n$. Then for any $p \in \operatorname{int}(X)$, one can fix $a_0 \in \operatorname{ext}(X)$ and find $a_1, \dots, a_n \in \operatorname{ext}(X)$, and write p as a convex combination of $\{a_j\}_{j=0}^n$, with positive coefficient λ_0, i.e.,*

$$p = \sum_{j=0}^{n} \lambda_j a_j, \text{ with } \lambda \geq 0, \lambda_0 > 0 \text{ and } \sum_{j=0}^{n} \lambda_j = 1. \qquad \square$$

Lemma 2.1.6. *Let X be a compact convex set in $\mathbb{R}^n$ and let $\operatorname{ext}(X) = \{x_1, \dots, x_k\}$. Then every point $p \in \operatorname{int}(X)$ can be written as a strict convex combination of all the extreme points, i.e.,*

$$p = \sum_{i=1}^{k} \lambda_i x_i, \text{ for some } \lambda > 0 \text{ and } \sum_{i=1}^{k} \lambda_i = 1.$$

Proof. Let p be an interior point of X. Fix $x_1 \in \operatorname{ext}(X)$ and let $\mathcal{J}_1$ be a subset of indices $\{1, \dots, k\}$ satisfying $|\mathcal{J}_1| = n+1$ and $1 \in \mathcal{J}_1$. Then, by using Lemma 2.1.5, we can write

$$p = \sum_{j \in \mathcal{J}_1} \lambda_j^1 x_j, \quad \text{for some } \lambda^1 \in \mathbb{R}_+^{n+1} \text{ with } \sum_{j \in \mathcal{J}_1} \lambda_j^1 = 1 \text{ and } \lambda_1^1 > 0.$$

In fact, we can write, by letting $\lambda_j^1 = 0, j \in [k] \setminus \mathcal{J}_1$,

$$p = \sum_{j=1}^{k} \lambda_j^1 x_j, \quad \text{for some } \lambda^1 \in \mathbb{R}_+^{k} \text{ with } \sum_{j=1}^{k} \lambda_j^1 = 1 \text{ and } \lambda_1^1 > 0.$$

Similarly, for each $i \in \{1, \dots, k\}$, we can fix $x_i \in \operatorname{ext}(X)$ and write

$$p = \sum_{j=1}^{k} \lambda_j^i x_j, \quad \text{for some } \lambda^i \in \mathbb{R}_+^{k} \text{ with } \sum_{j=1}^{k} \lambda_j^i = 1 \text{ and } \lambda_i^i > 0. \tag{2.1.1}$$

Then adding k equations for each $i \in \{1, \dots, k\}$ of (2.1.1) leads to

$$kp = \sum_{i=1}^{k} \sum_{j=1}^{k} \lambda_j^i x_j \implies p = \frac{1}{k} \sum_{i=1}^{k} \sum_{j=1}^{k} \lambda_j^i x_j.$$

[1] The statement in [28, Theorem 8.11] is modified. The original statement is precisely: Let A be a compact convex subset of $\mathbb{R}^\nu$ of dimension n. Then any point in A is a convex combination of at most $n+1$ extreme points. In fact, for any x, one can fix $e_0 \in \operatorname{ext}(A)$ and find $e_1, \dots, e_n \in \operatorname{ext}(A)$ so x is a convex combination of $\{e_j\}_{j=0}^n$. If $x \in \operatorname{relint}(A)$, then $x = \sum_{j=0}^n \theta_j e_j$ with $\theta_0 > 0$. In particular, $A = \operatorname{conv}(\operatorname{ext}(A))$.

Since $\frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{k}\lambda_j^i = 1$ and each coefficient of x_j is positive, p is a strict convex combination of extreme points of X. $\square$

Remark 2.1.7.

1. *An interior point of a compact convex set may not have a unique strict convex combination.*

2. *An interior point of a compact convex set that is written as a strict convex combination of all extreme points may not use all the extreme points.*

Examples are illustrated in Example 2.1.8 below.

Example 2.1.8. *Given the set*

$$X = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\} = \{e_1, e_2, -e_1, -e_2\},$$

consider the convex set $S = \operatorname{conv}(X)$ (See Figure 2.1.1.). It is clear that all elements in

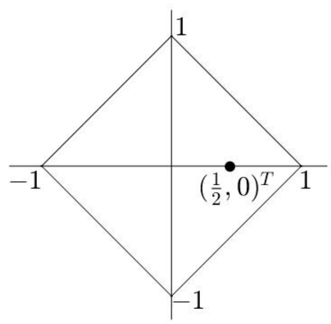


Figure 2.1.1: A convex set with four extreme points in $\mathbb{R}^2$. $\left(\frac{1}{2}, 0\right)^T$ is an interior point of S.

X are extreme points of S.

1. *$\left(\frac{1}{2}, 0\right)^T$ does not have a unique strict convex combination of all extreme points since*

$$\left(\tfrac{1}{2}, 0\right)^T = \tfrac{5}{8}e_1 + \tfrac{1}{8}e_2 + \tfrac{1}{8}(-e_1) + \tfrac{1}{8}(-e_2),$$

and

$$\left(\tfrac{1}{2}, 0\right)^T = \tfrac{6}{10}e_1 + \tfrac{3}{20}e_2 + \tfrac{1}{10}(-e_1) + \tfrac{3}{20}(-e_2).$$

2. *We note that*

$$\left(\tfrac{1}{2}, 0\right)^T = \tfrac{1}{2}e_1 + \tfrac{1}{4}e_2 + \tfrac{1}{4}(-e_2),$$

and the extreme point $-e_1$ is not used to represent $\left(\frac{1}{2}, 0\right)^T$ above. $\square$

Remark 2.1.9. *The statement of Caratheodory's theorem is as follows: Any* $x \in \operatorname{conv}(X) \subset \mathbb{R}^n$ *can be represented as a convex combination of* $n+1$ *elements of* S*. In Example 2.1.8, we note that* S *is in* $\mathbb{R}^2$*. Hence, Caratheodory's theorem states that we need at most three points in* S *to represent* $\left(\frac{1}{2}, 0\right)^T$ *as a convex combination of points in* S *(See Item 2 of Example 2.1.8.).*

2.1.2 General Cones

A reader familiar with basic definitions of cones, polar cones, cone bases and extreme rays may skip to the next section.

Definition 2.1.10 (*cone*)**.** *A nonempty set* $S \subseteq \mathbb{R}^n$ *is a* cone *if* $0 \in S$ *and for every* $x \in S$ *and* $\lambda \geq 0$*,* λx *belongs to* S*. In other words, a nonempty set* S *is a cone if, and only if,* $0 \in S$ *and for every* $x \in S \setminus \{0\}$*,* S *contains the half line starting from the origin in the direction* x*.*

Definition 2.1.11 (*pointed cone*)**.** *A cone* K *is said to be* pointed *if it does not contain a line, i.e.,*

$$x, -x \in K \implies x = 0.$$

Definition 2.1.12 (*pointed cone with vertex*)**.** *Given a pointed cone* K*, a pointed cone with vertex* x *is any set of the form* $x + K$*.*

Definition 2.1.13 (*proper cone*)**.** *A cone* $K \subset \mathbb{R}^n$ *is called a* proper cone *if* K *is convex, closed, pointed and has nonempty interior.*

Definition 2.1.14 (*dual cone*)**.** *The* dual cone K^* *of a set* K *is*

$$K^* := \{x \in \mathbb{R}^n : \langle x, y \rangle \geq 0, \ \forall y \in K\}.$$

Definition 2.1.15 (*conic combination, conical hull*)**.** *A* conic combination *of elements* $x_1, \ldots, x_k$ *is an element of the form* $\sum_{i=1}^k \alpha_i x_i$*, where the coefficients* α_i *are nonnegative. If the coefficients* α_i *are positive,* $\forall i = 1, \ldots, k$*, then we call* $\sum_{i=1}^k \alpha_i x_i$ *a* strict conic combinations *of* $x_1, \ldots, x_k$*.*

The set of all conical combinations from a given nonempty set $S \subset \mathbb{R}^n$ *is the* conical hull *of* S*, denoted by* $\operatorname{cone}(S)$*.*

Definition 2.1.16 (*base for a cone*, e.g., [1])**.** *Let* P *be a cone in a vector space. A nonempty convex subset* B *of* $P \setminus \{0\}$ *is said to be a* base for the cone P *(or, cone base), if for each* $x \in P \setminus \{0\}$*, there exists* $\lambda > 0$ *and* $b \in B$ *both uniquely determined such that* $x = \lambda b$*.*

Lemma 2.1.17 ([1, Corollary 3.8]). *Every closed pointed cone in a finite dimensional vector space has a compact base.* $\square$

By the homogeneity of cones, we may assume that a base B for a pointed cone P in a finite dimensional vector space X is of the form

$$B = P \cap \{x \in X : \langle x, p\rangle = \alpha\},$$

for some $p \in X$ and $\alpha \in \mathbb{R}$ (See Figure 2.1.2.). We note that B is a convex compact set.

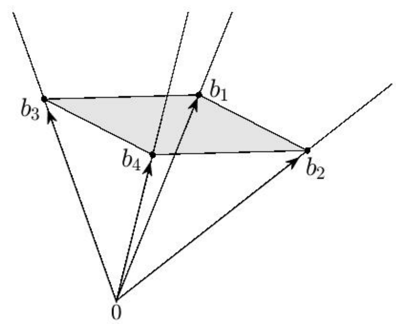


Figure 2.1.2: An illustration of a cone base in $\mathbb{R}^3$: The shaded region is a cone base of the given cone.

Example 2.1.18. *Suppose that $P = \mathbb{R}^n_+$. Then we can choose a base*

$$B = \mathbb{R}^n_+ \cap \{x \in \mathbb{R}^n : \langle e, x\rangle = 1\},$$

i.e., if $x \in \mathbb{R}^n_+$, then we let $\lambda = \langle e, x\rangle$ and $b = \frac{1}{\lambda}x$ to get the unique representation. Similarly, if $P = \mathbb{S}^n_+$ (i.e., the cone of positive semi-definite matrices), then we can choose

$$B = \mathbb{S}^n_+ \cap \{X \in \mathbb{S}^n_+ : \langle I, X\rangle = 1\} = \mathbb{S}^n_+ \cap \{X \in \mathbb{S}^n_+ : \mathrm{trace}(X) = 1\},$$

i.e., if $X \in \mathbb{S}^n_+$, then we let $\lambda = \langle I, X\rangle$ and $B = \frac{1}{\lambda}X$ to get the unique representation. $\square$

Definition 2.1.19 (*ray*). *Given a vector $a \in \mathbb{R}^n$, we denote the half-line generated by a*

$$\mathrm{ray}(a) = \mathrm{cone}(\{a\}) = \{ka : k \geq 0\}.$$

Definition 2.1.20 (*extremal vector, extreme ray*). *Let $P \subset \mathbb{R}^n$ be a cone. A nonzero vector $v \in P$ is said to be an* extremal vector *of P if*

$$x \in P \text{ and } v - x \in P \implies x = \lambda v,$$

for some $\lambda \geq 0$. In this case, the half-ray ray(r) *is called an* extreme ray.

Example 2.1.21. *Given*

$$a_1 = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}, a_2 = \begin{pmatrix} 1/2 \\ 1/3 \end{pmatrix}, \text{ and } a_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

define $P := \operatorname{cone}(\{a_1, a_2, a_3\})$. *Then* $\operatorname{ray}(a_1)$ *and* $\operatorname{ray}(a_2)$ *are extreme rays of* P *and* a_1 *and* a_2 *are extremal vectors of* P *(See Figure 2.1.3.). In fact, for any* $\lambda \in \mathbb{R}_{++}$, $\lambda a_1, \lambda a_2$ *are extremal vectors of* P. □

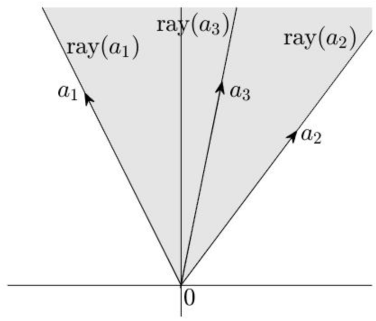


Figure 2.1.3: The shaded region is $\operatorname{cone}(\{a_1, a_2, a_3\})$.

Definition 2.1.22 (*recession cone*)**.** *The* recession cone *of the closed convex set* C *is the closed convex cone* C_∞ *defined by*

$$C_\infty := \{d \in \mathbb{R}^n : x + td \in C, \ \forall t > 0 \text{ and } \forall x \in C\}.$$

We call each element in C_∞ *a* recession direction*.*

An *extreme direction* d, with $\|d\| = 1$, of a convex set is a recession direction of the set that cannot be represented as a strict conic combination of two distinct recession directions d^1 and d^2 with $\|d^1\| = \|d^2\| = 1$.

Lemma 2.1.23 ([25, Theorem 18.5])**.** *Let* C *be a close convex set containing no lines, and let* S *be the set of all extreme points and extreme directions of* C*. Then* $C = \operatorname{conv}(S)$. □

Definition 2.1.24 (*polar cone*)**.** *Let* K *be a convex cone. The* negative polar cone *of* K *is*

$$K^\circ := \{s \in \mathbb{R}^n : \langle s, x \rangle \leq 0, \ \forall x \in K\}.$$

The positive polar cone *of* K *is*

$$K^* := \{s \in \mathbb{R}^n : \langle s, x \rangle \geq 0, \ \forall x \in K\},$$

which is precisely the dual cone of K *(See Definition 2.1.14.).*

Lemma 2.1.25. *Given two cones K_1, K_2 satisfying $K_1 \subseteq K_2$, polarization is order-reversing, i.e.,*

$$K_1 \subseteq K_2 \implies K_1^\circ \supseteq K_2^\circ,$$
$$K_1 \subseteq K_2 \implies K_1^* \supseteq K_2^*.$$

Proof. Suppose that $x \in K_2^\circ$. Then $\langle s, x\rangle \leq 0,\ \forall s \in K_2$, by Definition 2.1.24. Since K_2 contains K_1, we have $\langle u, x\rangle \leq 0,\ \forall u \in K_1$. Hence $x \in K_1^\circ$, by Definition 2.1.24 again. Replacing $\leq$ with $\geq$ in the preceding proof gives the proof for the dual cone. □

2.1.3 Polyhedral Theory

In the previous section, we studied some definitions of general cones. In this section, we further study a special class of cones, namely, the *polyhedral cones*.

Definition 2.1.26 (*polyhedral cone*)**.** *A set $P \subseteq \mathbb{R}^n$ is a* polyhedral cone *if P is the intersection of a finite number of halfspaces containing the origin on their boundaries. That is, $P := \{x \in \mathbb{R}^n : Ax \geq 0\}$ for some $A \in \mathbb{R}^{m\times n}$.*

Definition 2.1.27 (*finitely generated cone*)**.** *A set $P \subseteq \mathbb{R}^n$ is a* finitely generated cone *if P is the convex cone generated by a finite set of vectors $r^1, \dots, r^k$, for $k \geq 1$. We write $P = \operatorname{cone}\left(\{r^1, \dots, r^k\}\right)$, and we say that $r^1, \dots, r^k$ are the* generators *of P. If R is the $n \times k$ matrix with columns $r^1, \dots, r^k$,*

$$\operatorname{cone}\left(\{r^1, \dots, r^k\}\right) = \{x \in \mathbb{R}^n : \exists \lambda \geq 0 \text{ such that } x = R\lambda\}.$$

Lemma 2.1.28 ([10, Theorem 3.11], Minkowski, Weyl)**.** *A subset of $\mathbb{R}^n$ is a finitely generated cone if, and only if, it is a polyhedral cone.* □

We can also make a statement that resembles Lemma 2.1.6 in terms of proper polyhedral cones. We first present a lemma on the cone base of a cone.

Lemma 2.1.29 ([1, Theorem 1.48])**.** *Let B be a base of a cone K. Then a vector $b \in B$ is an extremal vector if and only if b is an extreme point of the convex set B.* □

With Lemma 2.1.17 and Lemma 2.1.29, we can state the cone version of Lemma 2.1.6. In other words, an interior point in a proper cone can be written as a strict convex combination of its extremal vectors.

Lemma 2.1.30. *Let K be a pointed cone generated by $\{a^i\}_{i=1,\dots,k}$, i.e., $K = \operatorname{cone}(\{a^i\}_{i=1,\dots,k})$. Assume that each a^i, $i = 1, \dots, k$ is an extremal vector of K. Then a point $p \in \operatorname{int}(K)$ can be written as a strict conic combination of extremal vectors of K, i.e.,*

$$p = \sum_{i=1}^{k} \lambda_i a^i, \text{ with } \lambda_i > 0, \forall i = 1, \dots, k.$$

Proof. By Lemma 2.1.17, we can choose a compact convex cone base B for K. Since K has a finite number of extremal vectors, B has a finite number of extreme points in B, say $B = \{b^1, \dots, b^k\}$, by Lemma 2.1.29. Since $p \in \text{int}(K)$, $\gamma p = \bar{p} \in \text{int}(B)$, for some $\gamma > 0$. Then, by Lemma 2.1.6,

$$\bar{p} = \sum_{i=1}^{k} \mu_i b^i, \text{ for some } \mu > 0 \text{ and } \sum_{i=1}^{k} \mu_i = 1.$$

Let $\lambda_i = (1/\gamma)\mu_i$, for all $i = 1, \dots, k$. Then, we have

$$p = \sum_{i=1}^{k} \lambda_i b^i, \text{ for some } \lambda > 0. \qquad \square$$

In Section 3.1, we will construct a polyhedral cone P with specified properties. The cone is constructed by generating its extreme rays and is defined by the convex hull of these extreme rays, i.e., $P = \text{cone}(\{p^i : i \in I\})$, where I is the set of indices of extreme rays. By Lemma 2.1.28, we know that we need a finite number of extreme rays to construct a polyhedral cone. Depending on the number of its extreme rays, the cone can be a *nondegenerate* cone or a *degenerate* cone. In the rest of this section, we explore how the number of extreme rays of a cone determines its nondegeneracy/degeneracy.

Definition 2.1.31 (*nondegenerate/degenerate polyhedral cone*)**.** *A proper cone P formed from an intersection of halfspaces is said to be* nondegenerate*, if exactly n distinct halfspaces are active at its vertex. If there are more than n active halfspaces at its vertex, then we call P* degenerate*.*

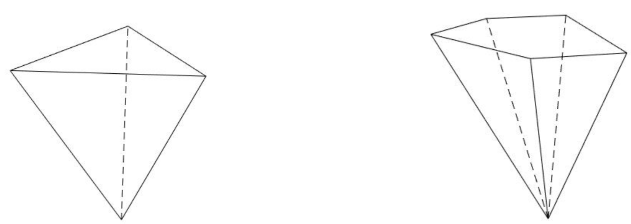


Figure 2.1.4: A non-degenerate cone(LHS) and a degenerate cone(RHS) in $\mathbb{R}^3$.

Given a cone P, knowing all extreme rays of P makes it very easy to obtain its polar cone.

Lemma 2.1.32 ([1, Theorem 3.36])**.** *Let a polyhedral cone P be generated by its extremal vectors $\{a^i\}_{i \in I}$. Then aside from scalar multiples, the dual cone P^* is precisely the inter-*

section of halfspaces determined by the same a^i's, i.e.,

$$P^* = \{y \in \mathbb{R}^n : \langle a^i, y\rangle \geq 0,\ \forall i \in I\}.^2$$

Proof. Suppose that we are given any points

$$x \in \text{cone}(\{a^i\}_{i\in I}) \ \text{ and } \ y \in \{y \in \mathbb{R}^n : \langle a^i, y\rangle \geq 0,\ \forall i \in I\}.$$

Then, for some $\lambda \geq 0$,

$$\langle x, y\rangle = \left\langle \sum_{i=1}^{k} \lambda_i a^i, y \right\rangle = \sum_{i=1}^{k} \lambda_i \left\langle a^i, y\right\rangle \geq 0.$$

Hence, $\{y \in \mathbb{R}^n : \langle a^i, y\rangle \geq 0,\ \forall i \in I\} \subset P^*$.

To show the equality, suppose to the contrary that there exists

$$y' \in P^* \setminus \{y \in \mathbb{R}^n : \langle a^i, y\rangle \geq 0,\ \forall i \in I\}.$$

Then there exists $j \in I$ such that $\langle a^j, y'\rangle < 0$. Note that $\langle a^j, y'\rangle \geq 0$ since $a^j \in P$. Hence, we have $0 \leq \langle a^j, y\rangle < 0$ and this yields a contradiction. □

A set $\{x \in \mathbb{R}^n : \langle a^i, x\rangle \leq b_i, \langle a^j, x\rangle = b_j, i \in \mathcal{I}, j \in \mathcal{J},\ \text{for some}\ \mathcal{I}, \mathcal{J}\}$ is said to be in *minimal representation* for P if all its constraints are irredundant. The following lemma is written with polyhedra. Since cones are polyhedra, Lemma 2.1.33 still works with cones. Lemma 2.1.33 shows that there is a unique representation of the dual cone P^*, up to scalar multiples.

Lemma 2.1.33 ([10, Corollary 3.31])**.** *Let P be a full-dimensional polyhedron and let $Ax \leq b$ be a minimal representation of P. Then $Ax \leq b$ is uniquely defined up to scaling, i.e., up to multiplying inequalities by a positive scalar.* □

In this thesis, we assume that the full-dimensional polyhedral cones are given in their minimal representations. By Lemma 2.1.33, we may assume that each normal vector a^i to the halfspace $\{x \in \mathbb{R}^n : \langle a^i, x\rangle \geq 0\}$ is of length 1.

Let $\{p^j\}_{j\in J}$ be the set of extremal vectors of the dual cone, $(\text{cone}(\{a^i\}_{i\in I}))^*$. Then, Table 2.1.1 is the consequence of Lemma 2.1.32. We observe from Table 2.1.1 that each extremal vector of the cone P gives a normal vector of a half-space determining its dual cone P^*. In other words, it is very easy to find halfspaces determining a cone as long as we know the extremal vectors of its dual cone. Similarly, if we know the extremal vectors of the dual cone, we can easily find P as an intersection of halfspaces. The following example

[2] The statement in Lemma 2.1.32 is modified. The original statement in [1, Theorem 3.36] is precisely: Let a polyhedral cone P is generated by the minimal set of inequalities $\{(a^i)^T x \geq 0 : i \in 1, \ldots, l\}$. Aside from scalar multiples, its inequalities are precisely the ones given by any collection of the l extremal vectors that generate the dual cone P^*.

Cone	P	P^*
Extremal vectors	$a^1, \dots, a^k$	$p^1, \dots, p^j$
Minimal set of inequalities	$p^1, \dots, p^j$	$a^1, \dots, a^k$

Table 2.1.1: Conic duality

illustrates how easily we can obtain the halfspace description of a cone once we know all extremal vectors of the primal and dual cones.

Example 2.1.34. *Given the vectors in* $\mathbb{R}^3$,

$$a^1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}, \, a^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \, a^3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{and } a^4 = \begin{pmatrix} -2/3 \\ 2/3 \\ 1/3 \end{pmatrix},$$

define $P := \operatorname{cone}(\{a^1, \dots, a^4\})$. *Then by Lemma 2.1.32, we have*

$$P^* = \{x \in \mathbb{R}^3 : \langle a^i, x\rangle \geq 0, \ i = 1, \dots 4\}.$$

The extremal vectors of P^* *are given by*

$$p^1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \, p^2 = \begin{pmatrix} 1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \end{pmatrix}, \, p^3 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \quad \text{and } p^4 = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{pmatrix}$$

(We are going to see how to obtain extremal vectors of the dual cone in Algorithm 2.1.1 later in this section.). Then by Lemma 2.1.32 again, the dual cone of P^* *is*

$$P^{**} = \{x \in \mathbb{R}^3 : \langle p^i, x\rangle \geq 0, \ i = 1, \dots 4\}.$$

Therefore, we have

$$\operatorname{cone}(\{a^1, \dots, a^4\}) = P = P^{**} = \{x \in \mathbb{R}^3 : \langle p^i, x\rangle \geq 0, \ i = 1, \dots 4\}.$$

See Figure 2.1.5 for an illustration. □

The following lemma shows a characterization for an extremal vector of a pointed polyhedral cone.

Lemma 2.1.35 ([10, Theorem 3.35])**.** *Let* P *be a pointed polyhedral cone and let* $\operatorname{ray}(r)$ *be a ray in* P. *Then* r *is an extremal vector of* P *if, and only if,* r *satisfies* $n-1$ *linear independent constraints of* $Ax \geq 0$ *with equality.* □

An extremal vector of a pointed polyhedral cone $P := \{x \in \mathbb{R}^n : Ax \geq 0\}$, with $A \in \mathbb{R}^{m \times n}$, can be obtained by using Lemma 2.1.35. We assume that P is in its minimal representation, i.e., all halfspaces of P are irredundant.

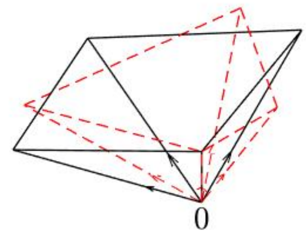


Figure 2.1.5: The solid black lines represent $P = \mathrm{cone}(\{a^1, \ldots, a^4\})$ and the red dashed lines represent $P^* = \mathrm{cone}(\{p^1, \ldots, p^4\})$.

1. We choose $n-1$ linearly independent rows of A and form a $(n-1) \times n$ submatrix A' of A as follows:
$$A' = \begin{bmatrix} (a^1)^T \\ \vdots \\ (a^{n-1})^T \end{bmatrix}.$$

2. We compute the null space of A'. Then we have
$$\mathrm{null}(A') = \{kp : k \in \mathbb{R}\}, \text{ for some } p \in \mathbb{R}^n.$$
(Note that $\dim(\mathrm{null}(A')) = 1$, since the size of A' is $(n-1) \times n$ and the rows of A' are linearly independent.)

3. We need to check the feasibility of p. This process is necessary as p might not be a part of the cone P:
   (a) If $Ap \geq 0$, then p is an extremal vector.
   (b) If $Ap \leq 0$, then $-p$ is an extremal vector of P.

   If p satisfies neither $Ap \leq 0$ nor $Ap \geq 0$, p cannot be made to be feasible to $Ax \geq 0$.

The above process Item 1-3 was to obtain just one extremal vector. Hence we need to find all possible $n-1$ linearly independent rows of A and repeat the above process Item 1-3 for each submatrix of A that has $n-1$ linearly independent rows of A. We present this in Algorithm 2.1.1 below.

As we are able to obtain all extremal vectors of a pointed polyhedral cone of the form $\{x \in \mathbb{R}^n : Ax \geq 0\}$ using Algorithm 2.1.1, we can write the dual cone $K^* = \{x \in \mathbb{R}^n : \langle a^i, x\rangle \geq 0, i = 1, \ldots, k\}$ as the convex hull of its extreme rays, say $K^* = \mathrm{cone}\left(\{p^j\}_{j\in J}\right)$. Then by Lemma 2.1.32 again, we obtain
$$K = K^{**} = \{x \in \mathbb{R}^n : \langle p^j, x\rangle \geq 0,\ j \in J\}.$$

Algorithm 2.1.1: Compute Extremal Vectors of a Cone

Input: A pointed polyhedral cone $C = \{x \in \mathbb{R}^n : Ax \geq 0\}$, where $A \in \mathbb{R}^{m \times n}$ with $m \geq n$.

Make a list $L \in \mathbb{R}^{{}^mC_{n-1} \times n-1}$ of all possible combinations of $\{1, \ldots, m\}$ taken $n-1$ at a time, where each row of L contains a combination.

for $i = 1 : {}^mC_{n-1}$ **do**
  $I := L(i, :)$
  $\bar{A} := A(I, :)$
  if rank $\left(\bar{A}\right) = n - 1$ **then**
    compute $p = \text{null}\left(\bar{A}\right)$
    normalize p
    if $Ap \geq 0$ **then**
      save p ▷ p is an extremal vector of C.
    end if
    if $Ap \leq 0$ **then**
      save $-p$ ▷ $-p$ is an extremal vector of C.
    end if
  end if
end for

Return: A matrix P with each row corresponds to an extremal vector $(p^i)^T$ of C.

Hence, by using Algorithm 2.1.1 and Lemma 2.1.32, we can convert the cone defined by the convex hull of its extreme rays into the cone defined by the intersection of halfspaces. We present a diagram of this procedure in Table 2.1.2:

$$\begin{array}{ccc}
K = \text{cone}(\{a^i\}_{i \in I}) & \xrightarrow{\text{Lemma 2.1.32}} & K^* = \{x \in \mathbb{R}^n : \langle a^i, x\rangle \geq 0,\ i \in I\} \\
\| & & \downarrow \text{Algorithm 2.1.1} \\
K^{**} = \{x \in \mathbb{R}^n : \langle p^j, x\rangle \geq 0,\ j \in J\} & \xleftarrow{\text{Lemma 2.1.32}} & K^* = \text{cone}(\{p^j\}_{j \in J})
\end{array}$$

Table 2.1.2: Starting from $K = \text{cone}(\{a^i\}_{i \in I})$, we can convert K as an intersection of halfspaces.

We now present a lemma on the number of extreme rays.

Lemma 2.1.36. *If $P = \{x \in \mathbb{R}^n : Ax \geq 0\}$ is a proper nondegenerate cone, then there are exactly n extreme rays in P.*

Proof. Suppose that P is nondegenerate. Then the intersection of exactly n distinct half-spaces $\{x \in \mathbb{R}^n : (a^i)^T x \leq 0,\ i = 1, \ldots, n\}$ determine the cone P. Let $(a^i)^T$ be the i-th row of A. Then there are exactly $\binom{n}{n-1}$ one-dimensional null spaces determined by $n - 1$

halfspaces. Suppose without loss of generality that A' is obtained by removing the last row of A. Let d be the basis of null(A'). Then we get

$$Ad = \begin{bmatrix} A' \\ (a^n)^T \end{bmatrix} d = \begin{bmatrix} A'd \\ \langle a^n, d\rangle \end{bmatrix} = \begin{bmatrix} 0 \\ \langle a^n, d\rangle \end{bmatrix}.$$

Therefore, d must satisfy either $\langle a^n, d\rangle > 0$ or $\langle a^n, d\rangle < 0$. If $\langle a^n, d\rangle > 0$, d is a feasible extremal vector. If $\langle a^n, d\rangle < 0$, we obtain $\langle a^n, -d\rangle > 0$ and hence $-d$ is an extremal vector. Therefore, every $n-1$ of n halfspaces determine an extreme ray of P. Thus, P has exactly n extreme rays. □

Corollary 2.1.37. *If a cone P is the convex hull of exactly n extreme rays, then P is nondegenerate.*

Proof. Suppose that $P = \text{cone}(\{a^i\}_{i=1,\dots,n})$, where each $a^i, i = 1, \dots, n$, is an extremal vector of P. Then by Lemma 2.1.32, the dual cone P^* is

$$\{x \in \mathbb{R}^n : \langle a^i, x\rangle \geq 0, i = 1, \dots, n\}.$$

By Lemma 2.1.36, the dual cone is the convex hull of exactly n extreme rays. Denote each extremal vector by p^i. Then, by the Lemma 2.1.32 again,

$$P^{**} = \{x : \langle p^i, x\rangle \geq 0, i = 1, \dots, n\}.$$

Thus, P is a non-degenerate cone. □

Corollary 2.1.38. *If P is proper and nondegenerate, then so is P^*.*

Proof. It is clear by the proof of Corollary 2.1.37. □

2.2 The Generalized Gradient

This section contains contents from Section 2.1 in [9]. In this section, we work with an arbitrary Banach space X, i.e., a complete normed space.

Definition 2.2.1 (*locally Lipschitz near x*)**.** *Let Y be a subset of X. Let $f : Y \to R$ be a given function, and let $x \in Y \subseteq X$. The function f is said to be* locally Lipschitz near x *if there exists a scalar K and a positive number ϵ such that the following holds:*

$$|f(y'') - f(y')| \leq K\|y'' - y'\|, \ \forall y'', y' \in \text{Ball}(x, \epsilon) \cap Y.$$

Intuitively speaking, when a function f is locally Lipschitz near x, the function values near x cannot be fluctuating too wildly.

Remark 2.2.2. *We want to point out that 'locally Lipschitz* near *x' and 'locally Lipschitz* at *x' are not the same notion. A function $f : Y \to \mathbb{R}$ is said to be locally Lipschitz at $x \in Y$ if*

$$|f(x) - f(y)| \leq K\|x - y\|,\ \forall y \in \text{Ball}(x, \epsilon) \cap Y.$$

While 'locally Lipschitz near *x' implies 'locally Lipschitz* at *x', however the other direction does not hold (See Example 2.2.3.). Hence, 'locally Lipschitz* near *x' is a stronger property than 'locally Lipschitz* at *x'.*

Example 2.2.3. *Consider the function $f : \mathbb{R} \to \mathbb{R}$ (See Figure 2.2.1.)*

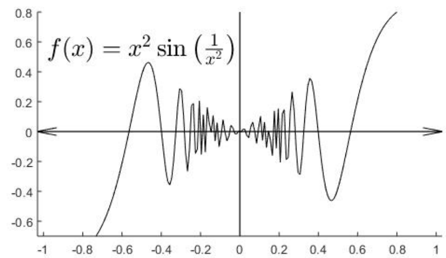


Figure 2.2.1: A function that is locally Lipschitz at 0, but not locally Lipschitz near 0

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & ,\ \textit{if } x \neq 0, \\ 0 & ,\ \textit{if } x = 0. \end{cases}$$

We note that $f(x)$ is differentiable at all $x \in \mathbb{R}$. In particular, $f'(0) = 0$, since $\lim_{h\to 0} \left(h^2 \sin(1/h^2)\right)/h = 0$. *Hence, by the definition of differentiability, we have*

$$\forall \epsilon > 0, \exists \delta > 0 \textit{ such that } |h| < \delta \implies \left|\frac{f(h) - f(0)}{h}\right| < \epsilon.$$

In other words, there exists $\delta > 0$ such that $|f(h) - f(0)| < 1 \cdot |h|$. Thus, $f(x)$ is Lipschitz continuous at 0.

Suppose that $f(x)$ is Lipschitz continuous near 0. Then there exists $\epsilon > 0$ such that $\forall x'', x' \in (-\epsilon, \epsilon)$, we have $|f(x'') - f(x')| \leq L|x'' - x'|$, for some L. Let $x = 1/\sqrt{2n\pi + \delta}$ and $y = 1/\sqrt{2n\pi}$ such that $0 < x < y < \epsilon$. We note that

$$\sin\left(\frac{1}{y^2}\right) = 0 \textit{ and } x - y = \frac{\sqrt{2n\pi} - \sqrt{2n\pi + \delta}}{\sqrt{2n\pi}\sqrt{2n\pi + \delta}}.$$

Then, we have

$$
\begin{aligned}
|f(x)-f(y)| &= \left|x^2\sin\left(\tfrac{1}{x^2}\right)-y^2\sin\left(\tfrac{1}{y^2}\right)\right| \\
&= \left|x^2\sin\left(\tfrac{1}{x^2}\right)\right| \\
&= \left|\tfrac{1}{2n\pi+\delta}\sin(2n\pi+\delta)\right| \\
&= \left|\sin(2n\pi+\delta)\frac{\sqrt{2n\pi}}{\sqrt{2n\pi+\delta}(\sqrt{2n\pi}-\sqrt{2n\pi+\delta})}\frac{\sqrt{2n\pi}-\sqrt{2n\pi+\delta}}{\sqrt{2n\pi}\sqrt{2n\pi+\delta}}\right| \\
&\le \left|\sin(2n\pi+\delta)\frac{\sqrt{2n\pi}}{\sqrt{2n\pi+\delta}(\sqrt{2n\pi}-\sqrt{2n\pi+\delta})}\right||x-y| \\
&= \left|\sin(2n\pi+\delta)\frac{\sqrt{2n\pi}\left(\sqrt{4n^2\pi^2+2\delta n\pi}+2n\pi+\delta\right)}{-2\delta n\pi-\delta^2}\right||x-y| \\
&\le L\,|x-y|\,.
\end{aligned}
$$

However, letting $n \to \infty$ yields a contradiction. Thus, $f(x)$ is not Lipschitz continuous near 0. □

The conventional directional derivative of f at x in the direction d is defined as

$$f'(x;d) := \lim_{t\downarrow 0}\frac{f(x+td)-f(x)}{t}.$$

A reader may refer to [18, Section D.1.1] for more arguments on the directional derivative.

We define the generalized directional derivative of a nonsmooth function as well.

Definition 2.2.4. *(generalized directional derivative) Let f be locally Lipschitz near x, and let v be any other vector in X. The* generalized directional derivative *of f at x in the direction v is given by*

$$f^\circ(x;v) := \limsup_{y\to x,\lambda\downarrow 0}\frac{f(y+\lambda v)-f(y)}{\lambda}\left(=\lim_{\lambda\downarrow 0}\sup_{y\in \mathrm{Ball}(0,\delta),\delta\in(0,\lambda)}\frac{f(x+y+\delta v)-f(x+y)}{\delta}\right).$$

Definition 2.2.5 (*generalized gradient*)**.** *The* generalized gradient *of f at x is*

$$\partial f(x) := \{\xi\in X^* : f^\circ(x;v)\ge\langle v,\xi\rangle,\ \forall v\in X\},$$

where X^* is the dual space of X.

Example 2.2.6. *Let f be a real valued function on* $\mathbb{R}$ *(See Figure 2.2.2.)*:

$$f(x) = \begin{cases} 2x & ,\ \textit{if } x\ge 0, \\ -\frac{1}{2}x & ,\ \textit{if } x<0. \end{cases}$$

If $x>0$, then we have

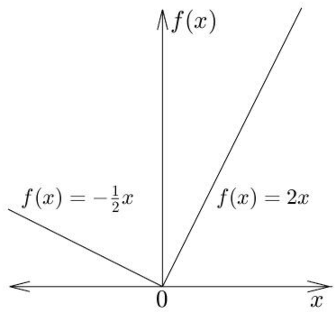


Figure 2.2.2: A nonsmooth function on $\mathbb{R}$.

$$f^{\circ}(x;v) = \limsup_{y\to x, t\downarrow 0} \frac{2(y+tv)-2y}{t} = 2v, \ \textit{and } \partial f(x) = \{2\}.$$

If $x < 0$, then we have

$$f^{\circ}(x;v) = \limsup_{y\to x, t\downarrow 0} \frac{-1/2(y+tv)+1/2y}{t} = -\frac{1}{2}v, \ \textit{and } \partial f(x) = \{-1/2\}.$$

Now suppose that $x = 0$. If $v \geq 0$, then we have

$$f^{\circ}(x;v) = \lim_{t\downarrow 0} \sup_{|y|<\delta, \delta\in(0,t)} \frac{f(y+\delta v)-f(y)}{\delta} = \lim_{t\downarrow 0} \frac{2\delta v}{\delta} = 2v,$$

and $2v \geq \xi v$ implies that $2 \geq \xi$. If $v < 0$, we have

$$f^{\circ}(x;v) = \lim_{t\downarrow 0} \sup_{|y|<\delta, \delta\in(0,t)} \frac{f(y+\delta v)-f(y)}{\delta} = \lim_{t\downarrow 0} \frac{-1/2\delta v}{\delta} = -\frac{1}{2}v,$$

and $-1/2v \geq \xi v$ implies that $-1/2 \leq \xi$. Thus, by Definition 2.2.5, we have $\partial f(x) = [-1/2, 2]$. □

We introduced Example 2.2.6 for illustrative purposes. In practice, it is not easy to compute the generalized gradients. Hence, we introduce the following characterization between the generalized directional derivative and the generalized gradient for later use.

Proposition 2.2.7 ([9, Proposition 2.1.2])**.** *Let f be Lipschitz continuous near $x \in X$. Then, for every v in X,*

$$f^{\circ}(x;v) = \max\{\langle \xi, v\rangle : \xi \in \partial f(x)\}.$$ □

Lemma 2.2.8. *If a function f is continuously differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$, i.e., $\partial f(x)$ is a singleton.* □

2.3 Linear Programming

In this section, we introduce some basic definitions and lemmas related to linear programming theory. We follow the approach from [17]. A reader familiar with the *basics* of linear programming can skip to the next section.

Definition 2.3.1 (*linear programming (**LP**)*)**.** *A* linear programming ***(LP)*** *is the problem of maximizing or minimizing an affine function subject to a finite number of linear equality and/or inequality constraints.*

Given a minimization **LP** denoted (P), we call a vector x a *feasible solution* if x satisfies all the constraints of (P). If the objective value of a feasible solution x^* is at least as small as any other feasible solution, then we call x^* an *optimal solution* to (P). An **LP** is said to be in *standard equality form* (**SEF**), if it is of the form

$$(\textbf{SEF}) \qquad \begin{array}{rl} \max_x & \langle c, x\rangle + \bar{z} \\ \text{subject to} & Ax = b \in \mathbb{R}^m \\ & x \in \mathbb{R}^n_+, \end{array}$$

where $\bar{z}$ is a scalar in $\mathbb{R}$, and Below we use the given linear system in (**SEF**) and assume that the matrix $A \in \mathbb{R}^{m\times n}$ has linearly independent rows. In addition, we refer to a polyhedron that is formed from the constraints in (**SEF**).

Given a subset B of column indices of $A \in \mathbb{R}^{m\times n}$, let A_B be the matrix formed by columns $A(:, i)$ for all $i \in B$.

Definition 2.3.2. *(basis, basic/nonbasic variable) Given the linear system $Ax = b$ as in (**SEF**), we say that a set of column indices B forms a* basis*, if the matrix A_B is square and non-singular. Then, the variables $x_j, j \in B$ are said to be* basic*; and the variables $x_j, j \in N := \{1, \dots, n\} \setminus B$, are said to be* nonbasic*.*

Definition 2.3.3 (*basic solution!basic feasible solution*)**.** *A vector $\bar{x}$ is said to be a* basic solution *of $Ax = b$ for a basis B, if*

$$A\bar{x} = b \text{ and } \bar{x}_N = 0.$$

$\bar{x}$ is called a basic feasible solution*, if $\bar{x} \geq 0$.*

Definition 2.3.4 (*degenerate/nondegenerate basic solution*)**.** *Given the polyhedron $\{x \in \mathbb{R}^n : Ax = b,\ x \geq 0\}$ with $A \in \mathbb{R}^{m\times n}$, let $\bar{x}$ be a basic solution. The vector $\bar{x}$ is said to be a* degenerate *basic solution, if more than $n - m$ of the coordinates of $\bar{x}$ are zero. If the basic solution $\bar{x}$ has exactly $n - m$ of nonzero coordinates, $\bar{x}$ is said to be a* nondegenerate *basic solution.*

Readers not familiar with duality may refer to [7, Chapter 5] or [17, Chapter 4]. In this thesis we concentrate on two primal-dual pairs. We list two pairs of primal-dual **LPs**, which we will use frequently, the standard and symmetric forms, respectively, e.g., [22]:

min	$\langle c, x\rangle$	max	$\langle b, y\rangle$
subject to	$Ax = b$	subject to	$A^T y \leq c$
	$x \geq 0$		

min	$\langle c, x\rangle$	max	$\langle b, y\rangle$
subject to	$Ax \geq b$	subject to	$A^T y \leq c$
	$x \geq 0$		$y \geq 0$

The following strong duality theorem follows from standard **LP** duality theorems in texts on **LP**, see e.g., [22].

Theorem 2.3.5 (primal-dual strong duality for **LP**)**.** *Let (P) and (D) be a feasible primal-dual pair. Then there exist optimal solutions $\bar{x}$ of (P) and $\bar{y}$ of (D). Moreover, the objective value with $\bar{x}$ in (P) equals the objective value with $\bar{y}$ in (D).*

Given a minimization **LP**(P) $\min\{\langle c, x\rangle : Ax \geq b,\ x \geq 0\}$, we denote

$$\operatorname{Argmin}\{\langle c, x\rangle : Ax \geq b,\ x \geq 0\}$$

as the set of optimal solutions to (P). Similarly, given a maximization **LP**(D) $\max\{\langle b, y\rangle : A^T y \leq c,\ y \geq 0\}$, we denote

$$\operatorname{Argmax}\{\langle b, y\rangle : A^T y \leq c,\ y \geq 0\}$$

as the set of optimal solutions to (D).

There are many interesting statements involving primal-dual pairs of **LPs**. For example, primal-dual strong duality holds for any primal-dual LP pair as long as one of them has a finite optimal value. In fact, the optimal values of the pairs are always the same unless both are infeasible. The primal **LP** has Slater points if, and only if, the dual optimal set is bounded. One can never find a primal-dual pair in standard form where both have bounded feasible regions.

2.4 Linear Robust Optimization Theory

Robust optimization is a field of optimization that tries to find the best uncertainty (perturbation)-immunized solution with given uncertainty set. A candidate solution x needs to satisfy all possible realizations of uncertain data. We call the set of all uncertain data an *uncertainty set*. Usually, the uncertainty set contains infinitely many elements and hence we can have infinitely many constraints. Such systems are called *semi-infinite* systems. (If the uncertainty set has finitely many elements, then we generally reduce to a finite constrained problem.) It is important to convert the infinite system of constraints to a finite system of inequalities and a finite number of variables. That is, we wish to convert the semi-infinite system into a *tractable* form.

In Section 2.4.1, we introduce some basic definitions of robust optimization theory. In Section 2.4.2, we show how to convert a semi-infinite system into a tractable system with box uncertainty.

2.4.1 Definitions

We follow the definitions and the preliminaries from [6].

Definition 2.4.1 (*uncertain linear programming*)**.** *An* uncertain linear programming *problem is a collection*

$$(\boldsymbol{LP}_{\mathcal{U}}) \qquad \left\{\min_x \left\{\langle c, x\rangle + d : Ax \le b\right\}\right\}_{(c,d,A,b)\in\mathcal{U}}$$

of $\boldsymbol{LP}$ *problems* $\min_x \{\langle c, x\rangle + d : Ax \le b\}$ *with a common structure (i.e., with the same number of constraints and the same number of variables) with the data varying in a given uncertainty set* $\mathcal{U}$*.*

We assume that the uncertainty set $\mathcal{U}$ is affinely parametrized by a perturbation vector ζ varying in a given perturbation set $\mathcal{Z}$, i.e.,

$$\mathcal{U} = \left\{(c, d, A, b) : \left[\begin{array}{c|c} c^T & d \\ \hline A & b \end{array}\right] = \left[\begin{array}{c|c} c_0^T & d_0 \\ \hline A_0 & b_0 \end{array}\right] + \sum_{\ell=1}^{L} \zeta_\ell \left[\begin{array}{c|c} c_\ell^T & d_\ell \\ \hline A_\ell & b_\ell \end{array}\right] \text{ for } \zeta \in \mathcal{Z}\right\}.$$

Here, (c_0, d_0, A_0, b_0) is the *nominal data* given in $(\mathbf{LP}_{\mathcal{U}})$. We may partition the uncertainty set $\mathcal{U}$ as follows:

$$\mathcal{U} = \mathcal{U}_{A,b} \times \mathcal{U}_{c,d},$$

where $\mathcal{U}_{A,b} := \{(A, b) : (A, b, c, d) \in \mathcal{U}\}$ and $\mathcal{U}_{c,d} := \{(c, d) : (A, b, c, d) \in \mathcal{U}\}$.

Definition 2.4.2 (*robust feasible*)**.** *A vector* $\bar{x} \in \mathbb{R}^n$ *is a* robust feasible *solution to* $(\boldsymbol{LP}_{\mathcal{U}})$*, if it satisfies all realizations of the constraints from the uncertainty set* $\mathcal{U}_{A,b}$*. We denote* $\mathcal{F}_{\mathcal{U}}(A, b)$ *as the set of robust feasible solutions (robust feasible set), i.e.,*

$$\mathcal{F}_{\mathcal{U}}(A, b) := \{x \in \mathbb{R}^n : Ax \le b,\ \forall (A, b) \in \mathcal{U}_{A,b}\}.$$

Definition 2.4.3 (*robust counterpart*)**.** *The* robust counterpart *of an uncertain* $\boldsymbol{LP}$ *problem* $(\boldsymbol{LP}_{\mathcal{U}})$ *is the optimization problem*

$$\begin{aligned} p^*_{RC} &= \min_x \sup_{(c,d,A,b)\in\mathcal{U}} \{\langle c, x\rangle + d : Ax \le b\} \\ &= \min_{x\in\mathcal{F}_{\mathcal{U}}(A,b)} \sup \{\langle c, x\rangle + d : (c, d) \in \mathcal{U}_{c,d}\} \end{aligned} \tag{2.4.1}$$

of minimizing the robust value of the objective over all robust feasible solutions to the uncertain problem.

An optimal solution to the robust counterpart is called a *robust optimal solution* to $\mathbf{LP}_{\mathcal{U}}$ and the optimal value of the robust counterpart is called the *robust optimal value* of $\mathbf{LP}_{\mathcal{U}}$. The meaning of the robust counterpart (2.4.1) is that we are looking for a solution that gives the best possible guaranteed value. The word *best* corresponds to *min*, and the word *guaranteed* corresponds to *sup*.

2.4.2 Reformulation of LP with Box Uncertainty

In this section, we will explore how a linear semi-infinite system

$$Ax \leq b, \ \forall (A, b) \in \mathcal{U}_{A,b}$$

can be reformulated as a system of finite number of inequalities and variables under box uncertainty.

We first focus on a single uncertainty-affected linear inequality $\langle a, x\rangle \leq \beta, x \in \mathbb{R}^n$. We let $\langle a^0, x\rangle \leq \beta^0$ be a single nominal constraint of $Ax \leq b$. Given an uncertainty set

$$\mathcal{U} := \left\{ [a;\beta] = \left[a^0;\beta^0\right] + \sum_{\ell=1}^{L} \zeta_\ell \left[a^\ell;\beta^\ell\right], \text{ for some } \zeta = (\zeta_\ell) \in \mathcal{Z} \right\}, \tag{2.4.2}$$

we want to represent the family of linear inequalities

$$\left\{\langle a, x\rangle \leq \beta\right\}_{(a,\beta)\in\mathcal{U}}$$

using a finite number of inequalities and a finite number of variables. Note that in (2.4.2), elements in $\mathcal{U}$ are parametrized by the set $\mathcal{Z}$, which we call the *perturbation set*.

Let the perturbation set be defined as

$$\mathcal{Z} := \mathrm{Box}_1 := \left\{\zeta \in \mathbb{R}^L : \|\zeta\|_\infty \leq 1\right\}.$$

Then, we have

$$
\begin{aligned}
x \text{ is robust feasible } &\iff a^T x \le \beta,\ \forall [a;\beta] \in \left\{ [a^0;\beta^0] + \sum_{\ell=1}^{L} \zeta_\ell \left[a^\ell;\beta^\ell\right],\ \zeta \in \mathcal{Z} \right\} \\
&\iff \left[a^0\right]^T x + \sum_{\ell=1}^{L} \zeta_\ell \left[a^\ell\right]^T x \le \beta^0 + \sum_{\ell=1}^{L} \zeta_\ell \beta^\ell,\ \forall \zeta \in \{\zeta : \|\zeta\|_\infty \le 1\} \\
&\iff \sum_{\ell=1}^{L} \zeta_\ell \left[\left[a^\ell\right]^T x - \beta^\ell\right] \le \beta^0 - \left[a^0\right]^T x,\ \forall (\zeta : |\zeta_\ell| \le 1, \ell = 1, \ldots, L) \\
&\iff \max_{-1 \le \zeta_\ell \le 1} \left[\sum_{\ell=1}^{L} \zeta_\ell \left[\left[a^\ell\right]^T x - \beta^\ell\right] \right] \le \beta^0 - \left[a^0\right]^T x \\
&\iff \sum_{\ell=1}^{L} \left| \left[a^\ell\right]^T x - \beta^\ell \right| \le \beta^0 - \left[a^0\right]^T x \\
&\iff \left[a^0\right]^T x + \sum_{\ell=1}^{L} \left| \left[a^\ell\right]^T x - \beta^\ell \right| \le \beta^0 \\
&\iff \left[a^0\right]^T x + \sum_{\ell=1}^{L} \underbrace{\left| \left[a^\ell\right]^T x - \beta^\ell \right|}_{u_\ell} \le \beta^0 \\
&\iff \left[a^0\right]^T x + \sum_{\ell=1}^{L} u_\ell \le \beta^0 \text{ and } -u_\ell \le \left[a^\ell\right]^T x - \beta^\ell \le u_\ell,\ \ell = 1, \ldots, L.
\end{aligned}
\tag{2.4.3}
$$

Note that we started with a semi-infinite system of inequalities and ended up with a system of linear inequalities with a finite number of constraints and a finite number of variables (though there was some dramatic increase in the number of constraints and variables.).

Remark 2.4.4. *A robust counterpart can change its character. The class of problem that a robust counterpart lies in depends on its perturbation set. For example, if the given perturbation set is an ellipsoid, the resulting constraint of the robust counterpart is a conic quadratic constraint. See [16, page 5] for a summary.*

We considered a single linear inequality so far. Now suppose that more than one linear inequality of $Ax \le b$ are uncertain and each uncertain inequality is associated with its own uncertainty set

$$
\mathcal{U}_i = \left\{ [a_i; b_i] = \left[a_i^0; b_i^0\right] + \sum_{\ell=1}^{L} \zeta_\ell \left[a_i^\ell; b_i^\ell\right] : \zeta \in \mathcal{Z}_i \right\},
$$

where $[a_i^0; b_i^0]$ is the data of the i-th linear inequality of $Ax \le b$. We repeat (2.4.3) for each uncertain inequality. We note that each reformulation will require additional variables, u's.

Robust Counterpart of a Special Case

Given a linear inequality $\langle a, x\rangle \le \beta$, suppose that each coefficient of a is uncertain and is known to vary in the interval $[-\epsilon, \epsilon]$, i.e.,

$$\langle \tilde{a}, x\rangle \le \beta,\ \forall \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon],\ \forall i \in [n]. \tag{2.4.4}$$

Then the robust counterpart of (2.4.4) can be rewritten as

$$\max \{\langle \tilde{a}, x\rangle : \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon],\ \forall i \in [n]\} \le \beta. \tag{2.4.5}$$

We may introduce additional variables and turn the semi-infinite constraint (2.4.5) into a finite system of linear deterministic inequalities as shown in (2.4.3). However in a special case, we have a simpler way to reformulate (2.4.5). Rewriting (2.4.5) gives

$$\langle a, x\rangle + \max \big\{\langle u, x\rangle : u \in \{y : \|y\|_\infty \le \epsilon\}\big\} \le \beta. \tag{2.4.6}$$

We use the notion of dual norm to convert (2.4.6).

Definition 2.4.5 (*dual norm*)**.** *Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^n$. The* dual norm $\|\cdot\|^*$ *of $\|\cdot\|$ is defined by*

$$\|s\|^* := \max\{\langle s, x\rangle : \|x\| \le 1\}.$$

It is well-known that the dual norm of the ℓ_∞ norm is the ℓ_1 norm:

$$\|x\|_\infty^* = \|x\|_1.$$

Therefore, we have

$$\begin{aligned} \max \{\langle u, x\rangle : u \in \{y : \|y\|_\infty \le \epsilon\}\} &= \max \{\langle u, x\rangle : u \in \epsilon\{y : \|y\|_\infty \le 1\}\} \\ &= \epsilon \cdot \max \{\langle u, x\rangle : u \in \{y : \|y\|_\infty \le 1\}\} \\ &= \epsilon\|x\|_\infty^* \\ &= \epsilon\|x\|_1. \end{aligned}$$

Hence, (2.4.6) becomes

$$\langle a, x\rangle + \epsilon\|x\|_1 \le \beta.$$

If we impose nonnegativity on the variable x, we have

$$\langle a, x\rangle + \epsilon\langle e, x\rangle \le \beta. \tag{2.4.7}$$

It is easy to see that if the given inequality is of the form $a^T x \ge \beta$ (i.e., with the opposite inequality), its robust counterpart is

$$\min \big\{\langle \tilde{a}, x\rangle : \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon],\ \forall i \in [n]\big\} \ge \beta, \tag{2.4.8}$$

which is equivalent to

$$\langle a, x\rangle - \epsilon\langle e, x\rangle \ge \beta. \tag{2.4.9}$$

We note that if $a = \epsilon e$ and $\beta > 0$, (2.4.9) is violated. In this case, x is not robust feasible since x does not satisfy all possible realizations of uncertain data. We will frequently make use of (2.4.7) and (2.4.9) later. We also note that (2.4.7) and (2.4.9) are in a nice parametric form.

Optimistic Counterpart

The contents in this section are not needed until Section 4.5.4. Readers may want to come back to this section later.

Given an uncertainty set $\mathcal{U}$, define the following **LP**(P):

$$(P) \qquad \begin{array}{ll} \min\limits_{x} & g(x) \\ \text{subject to} & f_i(x; u_i) \geq 0,\ i = 1, \ldots, m,\ u_i \in \mathcal{U},\ x \in \mathbb{R}^n. \end{array} \tag{2.4.10}$$

Definition 2.4.6 (*optimistic feasible solution*, [4]). *A vector x is an* optimistic feasible *solution of (P) if it satisfies the constraints for at least one realization of the uncertainty set. That is, x is an optimistic feasible solution if, and only if, for every $i = 1, \ldots, m$, $f_i(x; u_i) \geq 0$ for some $u_i \in \mathcal{U}$.*

The optimistic counterpart of problem (P) consists of minimizing the best possible objective function (i.e., minimum with respect to the parameters) over the set of optimistic feasible solutions:

$$\begin{array}{ll} \min\limits_{x} & \min\limits_{u \in \mathcal{U}} g(x) \\ \text{subject to} & f_i(x; u_i) \geq 0 \text{ for some } u_i \in \mathcal{U}_i, i = 1, \ldots, m. \end{array} \tag{2.4.11}$$

Let the constraint system in (P) is given by $Ax - b \geq 0$. Let $\langle a, x \rangle \geq \beta$ be one of the inequalities of $Ax - b \geq 0$. Suppose that each coefficient of a is uncertain and is known to vary in the interval $[-\epsilon, \epsilon]$, i.e.,

$$\langle \tilde{a}, x \rangle \geq \beta,\ \forall \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon],\ \forall i \in [n]. \tag{2.4.12}$$

Then, the optimistic counterpart of (2.4.12) can be written as

$$\max \{\langle \tilde{a}, x \rangle : \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon],\ \forall i \in [n]\} \geq \beta. \tag{2.4.13}$$

Note that 'max' was used in (2.4.13) (in optimistic counterpart), while 'min' was used in (2.4.8) (in robust counterpart). If we impose *nonnegativity* on the variables, we have the following:

$$\begin{aligned}\max\{\langle u,x\rangle : u\in\{y:\|y\|_\infty\le\epsilon\}\} &= \max\{\langle u,x\rangle : u\in\epsilon\{y:\|y\|_\infty\le 1\}\}\\ &= \epsilon\cdot\max\{\langle u,x\rangle : u\in\{y:\|y\|_\infty\le 1\}\}\\ &= \epsilon\|x\|_\infty^*\\ &= \epsilon\|x\|_1\\ &= \epsilon\langle e,x\rangle.\end{aligned}$$

Hence, the uncertain inequality becomes

$$\langle a,x\rangle+\epsilon\langle e,x\rangle\ge\beta. \tag{2.4.14}$$

We will make use of (2.4.14) in Section 4.5.4.

2.5 Linear Sensitivity Analysis

LP sensitivity analysis is a very well-established field. There is a considerable amount of literature on **LP** sensitivity analysis. There is a literature even dating back to 1954 by Saaty and Gass [27] and to 1956 by Mills [23]. A reader may refer to [2, 13, 20, 21, 31] for general understanding in this field. Arsham and Oblak [2] suggested classifications in **LP** postoptimal analysis (See Table 2.5.1.). A book by Gal [14] contains comprehensive arguments on postoptimal analysis with an ample number of examples.

Classical sensitivity analysis questioned how much given data could be perturbed while keeping the current optimal basis. However this approach has a grave shortcoming: In the presence of degeneracy(i.e., the existence of multiple optimal bases) of the primal optimal solution, we might get incorrect information in terms of sensitivity. A good example is illustrated in the short paper by Strum [30]. An approach using an optimal partition was suggested to remove the concerns that arise with the degeneracy of the optimal solution by Jansen et al. [21].

Most of the classical sensitivity analysis is performed on the RHS vector or the objective vector changes using the simplex tableau. In addition, we often encounter sensitivity analysis on only *one* component of the RHS vector or the objective vector for easier analysis. When the RHS (the cost vector, respectively) is perturbed, only the RHS (the cost vector, respectively) is affected in the final tableau. However, when matrix coefficients are perturbed, not only the matrix coefficients are affected but also both the RHS and the cost vector are affected in the final tableau. Moreover, there could be some additional implicit constraints imposed on each matrix coefficient due to the modelling. Hence, sensitivity analysis on the matrix coefficients using bases is relatively less studied because the analysis is more complex.

There are many interesting statements; one parameter change in the RHS and the cost vector yield the optimal value function to be piecewise linear convex and concave, respectively (A reader may refer to [31] for properties.). However, matrix coefficient changes

Perturbation Analysis

Tolerance Analysis

Symmetric Tolerance Analysis

Parametric Analysis

Sensitivity Analysis

Allowable change in one RHS or cost or/and element of matrix coefficient A

Simultaneous change in a given direction for RHS or cost or coefficients of matrix A

Allowable equal percentage change over all RHS and/or cost and/or coefficients of matrix A

Allowable percentage change in either direction for each RHS, cost and/or coefficients of matrix A

Largest set of any simultaneous changes in the RHS, cost and coefficients of matrix A

Table 2.5.1: Hierarchy in **LP** postoptimal analysis suggested by Arsham and Oblak [2]. Parametric analysis is often referred as sensitivity analysis.

do not guarantee the convexity nor concavity of the optimal value function and the optimal value function is often nonlinear (See the example in [32].). We can also find an example that the optimal value function is not even continuous with respect to the changes in matrix coefficients (See [14, Example 8-2].).

The sensitivity of an **LP** heavily relies on the dual optimal solutions. Each component of dual optimal solutions is often referred as the shadow price (or marginal value, Lagrange multiplier) associated with a particular constraint. The Shadow price plays an important role in analysis of economic models. It gives the change in the optimal value function per unit increase in the RHS value for a specific constraint, while all other problem data remains unchanged. A recent paper by Gisbert et al [15] contains the calmness of objective value function in the variation of the objective vector and the RHS vector based on sets of dual optimal solutions (A reader who wishes to explore the definition of calmness may refer to [26, Section 8.F].).

Regardless of perturbations on the RHS, the objective vector or the matrix coefficients, we see that the dual optimal solutions play a very important role in sensitivity analysis. Hence, when we have a full knowledge on the set of primal-dual optimal solutions, the sensitivity analysis gets easier. In this thesis, we study the sensitivity of the optimal value function over *all the matrix coefficients*. Freund [12] performed the postoptimal analysis under simultaneous changes in matrix coefficients involving the optimal bases. In the case of nondegeneracy, Freund suggests a very easy analysis. However, as we mentioned above, postoptimal analysis using the bases is difficult when the degeneracy is present. De Wolf [32] suggested a formula using the *generalized gradient* for sensitivity analysis without involving arguments on bases in the statement. However it requires full knowledge on the set of primal-dual optimal solutions. In this thesis, we extensively use the results given in [12, 32]. The results from [12, 32] will be stated in Chapter 4.

Chapter 3

Construction of Cones and the Properties of Constructed Cones

3.1 Generating the Polyhedral Cones

We now generate proper polyhedral cones with special properties that allow us to control their sizes. In other words, we generate proper cones satisfying the strict containment using a parameter (or an angle) in $(0, \pi/2)$. In Sections 3.1.1 and 3.1.2, we show how we build the desired cones that hold the strict containment; while in Section 3.1.3 we study some of the properties of these cones.

3.1.1 Construction of Generator c in a Two-Dimensional Subspace

We first focus on a two-dimensional subspace of $\mathbb{R}^n$ and construct a unit vector c to be used to generate the polyhedral cone. Suppose that we are given two orthonormal vectors $\bar{w}$ and w', $(\langle \bar{w}, w' \rangle = 0,\ \|\bar{w}\| = \|w'\| = 1)$, and the scalar $\theta \in (0, \pi/2)$. We want to find a vector $c \in \mathbb{R}^n$ such that c lies in $\operatorname{span}(\{\bar{w}, w'\})$ and the angle between c and $\bar{w}$ is θ, i.e.,

$$c = \alpha \bar{w} + \beta w', \quad \text{for some } \alpha, \beta \in \mathbb{R}, \text{ and } \arccos\left(\frac{\langle c, \bar{w} \rangle}{\|c\| \|\bar{w}\|}\right) = \theta.$$

Then letting $\cos(\theta) = \alpha/1$ and $\sin(\theta) = \beta/1$ gives the desired vector c (See Figure 3.1.1.);

$$c = \cos(\theta)\bar{w} + \sin(\theta) w'. \tag{3.1.1}$$

3.1.2 Building a Polyhedral Cone

We assume that the vectors satisfying the following hypothesis are given.

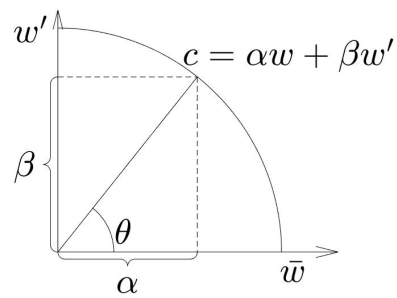


Figure 3.1.1: Given $\bar{w}$ and w' satisfying $\langle \bar{w}, w' \rangle = 0$: construction of c such that $\langle c, \bar{w} \rangle = \cos\theta$.

Hypothesis 3.1.1. *Let $\mathcal{K} = \{1, \ldots, k\}$, and suppose we are given k distinct unit vectors $w^1, \ldots, w^k \in \mathbb{R}^n$ such that the following hold:*

(1) $k \geq n$,

(2) $\|w^i\| = 1, \ \forall i \in \mathcal{K}$,

(3) $\dim(\operatorname{span}(\{w^i\}_{i\in\mathcal{K}})) = n - 1$,

(4) $0 \in \operatorname{relint} \operatorname{conv}\left(\{w^i\}_{i\in\mathcal{K}}\right)$. □

Remark 3.1.2. *For the case $n = 2$, we note that k cannot be greater than 2. If $k > n = 2$, then (2) and (3) of Hypothesis 3.1.1 and the assumption 'distinct vectors $w^1, \ldots, w^k \in \mathbb{R}^n$' cannot be satisfied simultaneously.*

By Item (3) of Hypothesis 3.1.1, the set $\{w^i\}_{i\in\mathcal{I}}$ spans a $n-1$ dimensional subspace of $\mathbb{R}^n$, that is, we can fix $\bar{w} \in \mathbb{R}^n$, $\|\bar{w}\| = 1$ such that the orthogonal complement

$$\bar{w}^\perp = \operatorname{span}(\{w^i\}_{i\in\mathcal{K}}).$$

Given $\theta \in (0, \pi/2)$, for each $i = 1, \ldots, k$, we define

$$\boxed{c^i := \cos\theta \bar{w} + \sin\theta w^i.} \tag{3.1.2}$$

We note that $\langle c^i, \bar{w} \rangle = \cos\theta$, for each $i = 1, \ldots, n$. We construct the desired polyhedral cone using the c^i's, i.e., as a convex hull of the rays $\operatorname{ray}(c^i)$, $\forall i = 1, \ldots, k$:

$$\boxed{K = \operatorname{cone}\left(\{c^i\}_{i=1,\ldots,k}\right).} \tag{3.1.3}$$

We study the properties of these generated cones K in Section 3.1.3, below. In the rest of this section, we present a simple example and discuss how we may generate data that satisfies our Hypothesis 3.1.1.

Example 3.1.3. *Suppose that the following vectors and angle are given:*

$$w^1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, w^2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, w^3 = \begin{pmatrix}-1\\0\\0\end{pmatrix}, w^4 = \begin{pmatrix}0\\-1\\0\end{pmatrix}, \quad \theta = \pi/6. \tag{3.1.4}$$

Let $\bar{w} = (0,0,1)^T$ *and define vectors*

$$c^i := \cos\theta\bar{w} + \sin\theta w^i = \frac{\sqrt{3}}{2}\bar{w} + \frac{1}{2}w^i,\ \forall i = 1,\dots,4.$$

Then we have

$$c^1 = \begin{pmatrix}1/2\\0\\\sqrt{3}/2\end{pmatrix}, c^2 = \begin{pmatrix}0\\1/2\\\sqrt{3}/2\end{pmatrix}, c^3 = \begin{pmatrix}-1/2\\0\\\sqrt{3}/2\end{pmatrix}, c^4 = \begin{pmatrix}0\\-1/2\\\sqrt{3}/2\end{pmatrix}.$$

We define $K := \operatorname{cone}\left(\{c^i\}_{i=1,\dots,4}\right)$ *(See Figure 3.1.2.).* □

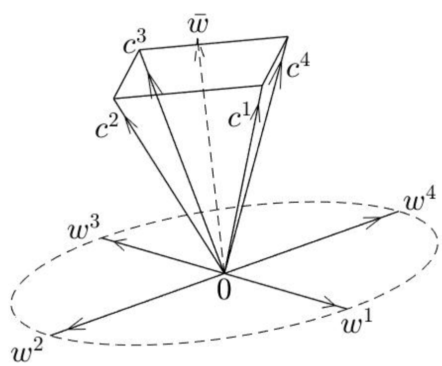


Figure 3.1.2: The cone constructed in $\mathbb{R}^3$ from the given data in (3.1.4).

We can generate a vector $\bar{w}$ and k vectors w^i satisfying Hypothesis 3.1.1 as follows. Our first goal is to generate a matrix $\widehat{W} \in \mathbb{R}^{(n-1)\times k}$ and a vector $\hat{\lambda} \in \mathbb{R}^k_{++}$ such that

1. each column of $\widehat{W}$ is of length 1,
2. $\operatorname{rank}(\widehat{W}) = n-1$,
3. $\widehat{W}\hat{\lambda} = 0$, and
4. $\sum_{i=1}^k \hat{\lambda}_i = 1$.

This means that 0 is in the relative interior of $\operatorname{conv}\left(\left\{\widehat{W}(:,i)\right\}_{i\in\mathcal{K}}\right)$. We generate such a matrix $\widehat{W}$ as follows: We first generate a vector $\lambda \in \mathbb{R}^k_{++}$. Then the null space of λ^T has

$k-1$ basis elements. We choose any $n-1$ basis elements of null(λ^T). We form an $(n-1)\times k$ matrix W_s, where each row is a basis element chosen above. We form a $k \times k$ diagonal matrix S such that each diagonal element is the norm of each column of $W_s \in \mathbb{R}^{(n-1)\times k}$. Define

$$\widehat{W} := W_s S^{-1} \quad \text{and} \quad \hat{\lambda} := \frac{S\lambda}{\sum_{i=1}^k (S\lambda)_i}.$$

Then, we have that each column of $\widehat{W}$ is of length 1, rank$(\widehat{W}) = n-1$, and 0 is the strict convex combination of columns of $\widehat{W}$ as desired.

We now lift the dimension of column vectors of $\widehat{W}$ to $\mathbb{R}^n$ by adding zero coordinates to each column of $\widehat{W}$, i.e.,

$$w^i := \begin{bmatrix} \widehat{W}(:,i) \\ 0 \end{bmatrix} \in \mathbb{R}^n, \ \forall i = 1,\dots,k.$$

Then, letting $\bar{w} = [0,\dots,0,1]^T \in \mathbb{R}^n$ yields $\langle w^i, \bar{w}\rangle = 0, \ \forall i = 1,\dots,k$. We generate an orthogonal matrix $O \in \mathbb{R}^{n\times n}$ and obtain

$$\bar{w} \leftarrow O\bar{w}, \text{ and } w^i \leftarrow Ow^i, \ \forall i = 1,\dots,k.$$

One may obtain an orthogonal matrix O by performing QR decomposition of an $n \times n$ matrix A, i.e., $O := Q$, where $A = QR$. Then we have a vector $\bar{w}$ and k vectors w^i as desired. Algorithm 3.1.1 shows the computational steps explained above.

Algorithm 3.1.1: Generate Vectors Satisfying Hypothesis 3.1.1 and $\bar{w}$

Input: n : dimension,
k : number of vectors in Hypothesis 3.1.1 we wish to generate
Generate a vector $\lambda \in \mathbb{R}^k_{++}$
Let W_0 be the matrix such that each column of W_0 is a basis element for null(λ^T)
Let $W_s := W_0(\mathcal{I},:)$, for some $|\mathcal{I}| = n-1$ and $\mathcal{I} \subset \{1,\dots,k\}$
Let $S \in \mathbb{R}^{k\times k}$ be a diagonal matrix such that $S(i,i) = \|W_s(i,:)\|, \forall i \in \{1,\dots,k\}$
Let $\widehat{W} = W_s S^{-1}$
$W = \begin{bmatrix} \widehat{W} \\ \mathbf{0}_k^T \end{bmatrix} \in \mathbb{R}^{n\times k}$, $\bar{w} = \begin{bmatrix} \mathbf{0}_{n-1} \\ 1 \end{bmatrix} \in \mathbb{R}^n$
Generate an orthogonal matrix $O \in \mathbb{R}^{n\times n}$
$W \leftarrow OW$, $\bar{w} \leftarrow O\bar{w}$
Return: a matrix $W \in \mathbb{R}^{n\times k}$ such that each column w^i corresponds to a vector in Hypothesis 3.1.1,
a vector $\bar{w}$ orthogonal to $w^i, \ \forall i \in \{1,\dots,k\}$

3.1.3 Properties of the Constructed Cone

In this section, we study some properties of the cone $K = \text{cone}\left(\{c^i\}_{i=1,\ldots,k}\right)$ constructed in Section 3.1.2.

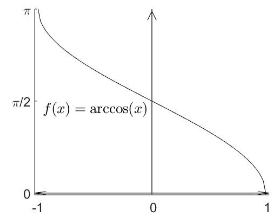


Figure 3.1.3: arccos function

The next theorem shows that the c^i's constructed in (3.1.2) are, in fact, extremal vectors of K.

Theorem 3.1.4. *Let $K = \text{cone}\left(\{c^i\}_{i=1,\ldots,k}\right)$ be the cone constructed in (3.1.3). Then each $c^i,\ i = 1, \ldots, k$, is an extremal vector of K.*

Proof. Let I be the subset of the indices $\{1, \ldots, k\}$. We proceed by induction on $|I|$. Suppose that $|I| = 1$. Then $K = \text{cone}(\{c^1\})$ and c^1 is clearly an extremal vector of K. Suppose that $K = \text{cone}(\{c^i\}_{i \in I})$ with $I = \{1, \ldots, j-1\}$ and each c^i, where $i \in I$, is an extremal vector of K. Let $K' = \text{cone}(\{c^i\}_{i \in I'})$ with $I' = I \cup \{j\}$. Suppose to the contrary that there is a member in $\{c^i\}_{i \in I'}$ that is not an extremal vector of K'. Without loss of generality, we let c^j be such a member, i.e., c^j is not an extremal vector of K'. By Lemma 2.1.23, we have

$$\gamma c^j = \sum_{i \in I} \lambda_i c^i \text{ for some } \gamma > 0, \text{ with } \lambda_i \geq 0 \text{ and } \sum_{i \in I} \lambda_i = 1.$$

Note that $\arccos\left(\frac{\langle c^i,\bar{w}\rangle}{\|c^i\|\|\bar{w}\|}\right) = \theta$, i.e., $\langle c^i, \bar{w}\rangle = \cos\theta, \forall i \in I'$. Thus we have

$$\begin{aligned}
\theta &= \arccos\left(\frac{\langle \gamma c^j, \bar{w}\rangle}{\|\gamma c^j\|\|\bar{w}\|}\right) \\
&= \arccos\left(\frac{\sum_{i\in I} \lambda_i \langle c^i, \bar{w}\rangle}{\|\gamma c^j\|}\right) \\
&= \arccos\left(\frac{\sum_{i\in I} \lambda_i \cos\theta}{\|\gamma c^j\|}\right) \\
&= \arccos\left(\frac{\sum_{i\in I} \lambda_i \|c^i\| \cos\theta}{\|\sum_{i\in I} \lambda_i c^i\|}\right) \\
&< \arccos\left(\frac{\sum_{i\in I} \lambda_i \|c^i\| \cos\theta}{\sum_{i\in I} \lambda_i \|c^i\|}\right) \\
&= \arccos\left(\cos\theta\right) \\
&= \theta.
\end{aligned} \tag{3.1.5}$$

The strict inequality holds, since arccos is strictly decreasing (See Figure 3.1.3.) and

$$\|\gamma c^j\| = \left\|\sum_{i\in I} \lambda_i c^i\right\| < \sum_{i\in I} \lambda_i \|c^i\|,$$

since c^i and c^ℓ are not collinear for all $i \neq \ell$. Therefore the strict inequality yields a contradiction and so c^j is an extremal vector of K'. □

We state a necessary condition for an element in K.

Lemma 3.1.5. *Let each c^i be constructed by (3.1.2). If $x \in K = \operatorname{cone}(\{c^i\}_{i=1,\ldots,k})$, then*

$$\arccos\left(\frac{\langle \bar{w}, x\rangle}{\|\bar{w}\|\|x\|}\right) \leq \theta.$$

Proof. We showed in Theorem 3.1.4 that $\{c^1, \ldots, c^k\}$ is the set of extremal vectors of K.

Then $x = \sum_{i=1}^{k} \lambda_i c^i$, for some $\lambda \geq 0$. Thus we have

$$\begin{aligned}
\frac{\langle w, x\rangle}{\|\bar{w}\|\|x\|} &= \frac{\sum_{i=1}^{k} \langle \bar{w}, \lambda_i c^i\rangle}{\|\bar{w}\|\|x\|} \\
&= \frac{\sum_{i=1}^{k} \lambda_i \langle \bar{w}, c^i\rangle}{\|x\|} \\
&= \frac{\sum_{i=1}^{k} \lambda_i \cos\theta}{\|x\|} \\
&= \frac{\sum_{i=1}^{k} \lambda_i \|c^i\| \cos\theta}{\|x\|} \\
&\geq \frac{\sum_{i=1}^{k} \lambda_i \|c^i\| \cos\theta}{\sum_{i=1}^{k} \lambda_i \|c^i\|} \\
&= \cos\theta.
\end{aligned}$$

Hence $\arccos\left(\frac{\langle \bar{w}, x\rangle}{\|\bar{w}\|\|x\|}\right) \leq \arccos(\cos\theta) = \theta$. $\square$

We can also show that K is a *full-dimensional* and *pointed* cone, i.e., a *proper cone*. We first show that K is full-dimensional.

Lemma 3.1.6 ([25, Theorem 2.4])**.** *The dimension of a convex set $C \subseteq \mathbb{R}^n$ is the maximum of the dimensions of the all simplices included in C.* $\square$

Lemma 3.1.7. *Let K be the cone constructed in (3.1.3). Then K is a full-dimensional cone.*

Proof. By Items (3) and (4) of Hypothesis 3.1.1, there exists a set $\{u^i\}_{i\in\mathcal{I}}$ of n affinely independent vectors u^i in $\operatorname{relint} \operatorname{conv}\left(\{w^i\}_{i\in\mathcal{K}}\right)$ (of dimension $n-1$). We can choose vectors u^i of the same length and let $\bar{u} = \gamma\bar{w}$ satisfy $\|\bar{u}\| = \|\gamma\bar{w}\| = \|u^i\|$, $\forall i \in \mathcal{I}$. Define

$$S := \left\{\cos\theta\bar{u} + \sin\theta u^i\right\}_{i\in\mathcal{I}}.$$

We show that S is a set of linearly independent vectors in $\mathbb{R}^n$. Since $\{u^i\}_{i\in\mathcal{I}}$ is the set of affinely independent vectors,

$$\sum_{i\in\mathcal{I}\setminus n} \alpha_i \left(u^i - u^n\right) = 0 \implies \alpha_i = 0,\ \forall i = 1, \ldots, n-1. \tag{3.1.6}$$

In order to show that S is a set of linearly independent vectors in $\mathbb{R}^n$, suppose that

$$\sum_{i\in\mathcal{I}} \beta_i \left(\cos\theta\bar{u} + \sin\theta u^i\right) = 0. \tag{3.1.7}$$

Then, we have

$$0 \begin{aligned}[t] &= \textstyle\sum_{i\in\mathcal{I}} \beta_i \left(\cos\theta \bar{u} + \sin\theta u^i - \sin\theta u^n + \sin\theta u^n\right) \\ &= \textstyle\sum_{i\in\mathcal{I}} \beta_i \Big(\left(\cos\theta \bar{u} + \sin\theta u^n\right) + \sin\theta \left(u^i - u^n\right) \Big). \end{aligned} \tag{3.1.8}$$

Then, by rearranging the terms in (3.1.8), we get

$$\sum_{i=1}^{n} \beta_i \sin\theta \left(u^i - u^n\right) = \sum_{i=1}^{n} -\beta_i \left(\cos\theta \bar{u} + \sin\theta u^n\right). \tag{3.1.9}$$

We note that the LHS of (3.1.9) is equal to $\sum_{i=1}^{n-1} \beta_i \sin\theta \left(u^i - u^n\right)$. We also note that

$$\left(\cos\theta \bar{u} + \sin\theta u^n\right) \notin \operatorname{span}\left(\{u^i - u^n\}_{i=1,\ldots,n-1}\right).$$

Therefore, (3.1.9) holds only when both sides of (3.1.9) are equal to zero. Thus we get

$$\sum_{i=1}^{n-1} \beta_i \left(u^i - u^n\right) = 0 \implies \beta_i = 0,\ \forall i = 1, \ldots, n-1,$$

by (3.1.6). It follows from (3.1.9) that $\beta_n = 0$. Therefore, S is a set of linearly independent vectors in $\mathbb{R}^n$.

We note that $\operatorname{conv}(S \cup \{0\})$ is a simplex in $\mathbb{R}^n$, and $\operatorname{conv}(S \cup \{0\}) \subset K$. Therefore, by Lemma 3.1.6, K is a full-dimensional cone. $\square$

Theorem 3.1.8. *Let $K = \operatorname{cone}\left(\{c^i\}_{i=1\ldots,k}\right)$ be constructed by (3.1.2). Then K is a pointed cone.*

Proof. Suppose to the contrary that there is a nonzero unit vector x such that $x \in K$ and $-x \in K$. Note that for all $y \in K$, we have $\frac{\langle y, \bar{w}\rangle}{\|y\|\|\bar{w}\|} \geq \cos\theta$ by Lemma 3.1.5. Then

$$\langle x, \bar{w}\rangle \geq \cos\theta \quad \text{and} \quad -\langle x, \bar{w}\rangle \geq \cos\theta,$$

so we have

$$\cos\theta \leq \langle x, \bar{w}\rangle \leq -\cos\theta.$$

This yields a contradiction since the inequalities hold only when $\langle x, \bar{w}\rangle = \cos\theta = 0$ but $\theta \in (0, \pi/2)$. $\square$

We also have a special element in the interior of K.

Lemma 3.1.9. *Let $K = \operatorname{cone}\left(\{c^i\}_{i=1,\ldots,k}\right)$ be constructed by (3.1.2). Then, $\bar{w} \in \operatorname{int}(K)$.*

Proof. By Items (3) and (4) of Hypothesis 3.1.1, there exists a set $\{u^i\}_{i\in\mathcal{I}}$ of n affinely independent vectors u^i in $\operatorname{relint}\operatorname{conv}\left(\{w^i\}_{i\in\mathcal{K}}\right)$. We can choose vectors u^i of the same length and let $\bar{u} = \gamma \bar{w}$ be satisfying $\|\bar{u}\| = \|\gamma \bar{w}\| = \|u^i\|, \ \forall i \in \mathcal{I}$. Define

$$s^i := \cos\theta \bar{u} + \sin\theta u^i, \ \forall i \in \mathcal{I}.$$

We can find $\lambda \in \mathbb{R}^n$ such that

$$0 = \sum_{i=1}^{n} \lambda_i u^i, \text{ for some } \lambda_i > 0 \text{ and } \sum_{i=1}^{n} \lambda_i = 1. \tag{3.1.10}$$

By summing $\lambda_i s^i = \lambda_i(\cos\theta \bar{u} + \sin\theta u^i), \forall i \in \mathcal{I}$, we have

$$\sum_{i=1}^{n} \lambda_i s^i = \sum_{i=1}^{n} \cos\theta \lambda_i \bar{u} + \sum_{i=1}^{n} \sin\theta \lambda_i u^i.$$

Then, with (3.1.10), we get

$$\cos\theta \bar{u} = \sum_{i=1}^{n} \lambda_i s^i. \tag{3.1.11}$$

Let z be any positive real number satisfying $1 - \sum_{i=1}^{n} \lambda_i / z > 0$. By dividing both sides of (3.1.11) by z, we have

$$\frac{\cos\theta}{z} \bar{u} = \sum_{i=1}^{n} \frac{\lambda_i}{z} s^i = \sum_{i=1}^{n} \frac{\lambda_i}{z} s^i + \left(1 - \sum_{i=1}^{n} \frac{\lambda_i}{z}\right) 0.$$

We note that the coefficients of s^i's and 0 lie in the interval $(0, 1)$. Hence $\frac{\cos\theta}{z}\bar{u}$ is in the interior of the simplex $\operatorname{conv}\left(0 \cup \left\{\frac{\lambda_i}{z} s^i\right\}_{i=1,\dots,n}\right)$. Therefore $\bar{w}$ is in the interior of the cone K. $\square$

Remark 3.1.10. *The properties of K shown in this section hold for every cone constructed by* (3.1.3) *with* any $\theta \in (0, \pi/2)$.

Given a cone K constructed by (3.1.2), we would like to construct a cone $K' \supsetneq K$, that is *bigger* than K. Let $\bar{\theta} \in (0, \pi/2)$ such that $\bar{\theta} > \theta$. We copy the construction (3.1.2) with θ replaced by $\bar{\theta}$, and then obtain d^i (See Figure 3.1.4.), i.e.,

$$\boxed{d^i = \cos\bar{\theta} \bar{w} + \sin\bar{\theta} w^i.} \tag{3.1.12}$$

We define two sets in $\mathbb{R}^n$ and two corresponding $n \times k$ matrices,

$$\boxed{\begin{array}{rcl} \mathcal{C} := \{c^1, \dots, c^k\} & \text{and} & \mathcal{D} := \{d^1, \dots, d^k\}, \\ C := \begin{bmatrix} c^1 & \cdots & c^k \end{bmatrix} & \text{and} & D := \begin{bmatrix} d^1 & \cdots & d^k \end{bmatrix}. \end{array}}$$

Throughout this thesis, we do not rearrange the orders of columns in C and D. That is, if c^i is the i-th column of C, we place d^i on i-th column of D. We always assume that

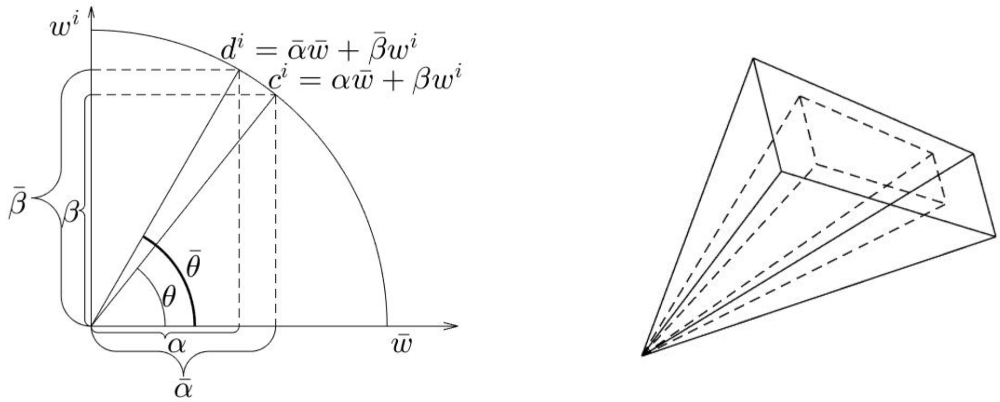


Figure 3.1.4: Construction of vector d^i moved further from c^i. The figure on the right is an example in $\mathbb{R}^3$, where the dashed line and the solid line represent cone($\mathcal{C}$) and cone($\mathcal{D}$), respectively.

cone($\mathcal{C}$) $\subsetneq$ cone($\mathcal{D}$), unless stated otherwise. We often denote

$$\text{cone}(\mathcal{C}) = \text{cone}\left(\{c^i\}_{i\in I}\right) = \{z \in \mathbb{R}^n : \exists\lambda \geq 0 \text{ such that } z = C\lambda\}, \tag{3.1.13}$$

$$\text{cone}(\mathcal{D}) = \text{cone}\left(\{d^i\}_{i\in I}\right) = \{z \in \mathbb{R}^n : \exists\lambda \geq 0 \text{ such that } z = D\lambda\}. \tag{3.1.14}$$

In addition, we interchangeably use the expressions in (3.1.13) and (3.1.14), when the meaning is clear.

Example 3.1.11. *Let the following be the given data:*

$$\bar{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, w^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, w^2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \theta = \frac{\pi}{4}, \bar{\theta} = \frac{\pi}{3}.$$

Then, by using (3.1.2) *and* (3.1.12), *we have*

$$c^1 = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}, c^2 = \begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}, d^1 = \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}, d^2 = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix},$$

and let

$$\mathcal{C} = \{c^1, c^2\}, \mathcal{D} = \{d^1, d^2\}, C = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}, D = \begin{bmatrix} 1/2 & -1/2 \\ \sqrt{3}/2 & \sqrt{3}/2 \end{bmatrix}.$$

With $\mathcal{C}, \mathcal{D} \subset \mathbb{R}^2$ and $C, D \in \mathbb{R}^{2\times 2}$ above, we interchangeably use the notations

$$\text{cone}(\mathcal{C}) = \text{cone}\left(\{c^i\}_{i=1,2}\right) = \{z \in \mathbb{R}^n : \exists\lambda \geq 0 \text{ such that } z = C\lambda\},$$

$$\text{cone}(\mathcal{D}) = \text{cone}\left(\{d^i\}_{i=1,2}\right) = \{z \in \mathbb{R}^n : \exists\lambda \geq 0 \text{ such that } z = D\lambda\}. \qquad \square$$

The following theorem guarantees that whenever we have $\bar{\theta} > \theta$, we can always construct a set of two cones of different sizes: one strictly contains the other. In other words, we can control the sizes (or containment) of the cones we wish to study.

Theorem 3.1.12. *Let* $\mathrm{cone}\,(\mathcal{C})$ *and* $\mathrm{cone}\,(\mathcal{D})$ *be the cones generated by the vectors constructed by* (3.1.2) *and* (3.1.12), *respectively. Then*

$$\mathrm{cone}\,(\mathcal{C}) \subsetneq \mathrm{cone}\,(\mathcal{D})\,.$$

Proof. By the construction of c^i and d^i, we have $\mathrm{ray}(c^i) \in \mathrm{cone}\,(\{\bar{w}, d^i\})$. Since $\bar{w} \in \mathrm{cone}\,(\mathcal{D})$, we have $\mathrm{cone}\,(\{\bar{w}\} \cup \mathcal{D}) = \mathrm{cone}\,(\mathcal{D})$. Hence, we get

$$\mathrm{ray}(c^i) \in \mathrm{cone}\left(\{\bar{w}, d^i\}\right) \subset \mathrm{cone}\,(\{\bar{w}\} \cup \mathcal{D}) = \mathrm{cone}\,(\mathcal{D})\,.$$

Thus each c^i can be written as

$$c^i = \sum_{j=1}^{k} \mu_j^i d^i, \text{ for some } \mu^i \geq 0.$$

Let $x \in \mathrm{cone}(\mathcal{C})$. Then

$$\begin{aligned} x &= \sum_{i=1}^{k} \lambda_i c^i, \text{ for some } \lambda_i \geq 0 \\ &= \sum_{i=1}^{k} \lambda_i \left(\sum_{j=1}^{k} \mu_j^i d^i \right), \text{ for some } \mu^i \geq 0 \\ &= \sum_{i=1}^{k} \sum_{j=1}^{k} \lambda_i \mu_j^i d^i \\ &= \sum_{s=1}^{k} \sigma_s d^s, \text{ for some } \sigma \geq 0. \end{aligned}$$

Therefore, $\mathrm{cone}\,(\mathcal{C}) \subset \mathrm{cone}\,(\mathcal{D})$.

Since for each i, $\arccos\,(\langle \bar{w}, d^i \rangle) = \bar{\theta} > \theta$, so $d^i \notin \mathrm{cone}\,(\mathcal{C})$ by Lemma 3.1.5. Thus the strict containment follows. ☐

So far we have described cones such as $K = \mathrm{cone}(\mathcal{C})$ using the convex hull of its extreme rays. Below, we want to use the cone K as constraints for a feasible region of an **LP**. Each constraint of an **LP** is a halfspace. Therefore, we need to describe the cone K as an intersection of halfspaces. We first check if the dual cone K^* is a pointed cone. Given $K = \mathrm{cone}(\{c^1, \ldots, c^k\})$, Lemma 2.1.32 ensures that the dual cone K^* is given by

$$K^* = \{x \in \mathbb{R}^n : (c^i)^T x \geq 0, i = 1, \ldots, k\} = \{x \in \mathbb{R}^n : C^T x \geq 0\}.$$

Lemma 3.1.13. *Given* $K = \mathrm{cone}\,(\{c^i\}_{i \in I})$, K^* *is a pointed cone.*

Proof. Suppose that there is $\bar{x} \in \mathbb{R}^n$ such that $\pm\bar{x} \in K^*$. Then $C^T \bar{x} = 0$, where i-th row of C^T is $(c^i)^T$. Since $\{c^i\}_{i\in I}$ spans $\mathbb{R}^n$, there is a square $n \times n$ submatrix $\bar{C}^T$ of C^T such that $\bar{C}^T\bar{x} = 0$. Hence $\bar{x} = 0$. ☐

Having a pointed dual cone, we can convert $K = \operatorname{cone}(\{c^i\}_{i\in I})$ as an intersection of halfspaces using the procedure in Table 2.1.2. We include the table here for convenience:

$$\begin{array}{ccc}
K = \operatorname{cone}(\{c^i\}_{i\in I}) & \xrightarrow{\text{Lemma 2.1.32}} & K^* = \{x \in \mathbb{R}^n : \langle c^i, x\rangle \geq 0,\ i \in I\} \\
\| & & \downarrow \text{Algorithm 2.1.1} \\
K^{**} = \{x \in \mathbb{R}^n : \langle p^j, x\rangle \geq 0,\ j \in J\} & \xleftarrow{\text{Lemma 2.1.32}} & K^* = \operatorname{cone}(\{p^j\}_{j\in J})
\end{array}$$

We recall that a proper cone in $\mathbb{R}^n$ is nondegenerate, if exactly n distinct halfspaces are active at its vertex. We present a characterization of a nondegenerate cone K.

Proposition 3.1.14. *Given a proper cone $K = \operatorname{cone}(\{c^i\}_{i\in I})$ in $\mathbb{R}^n$, the following are equivalent.*

(1) K is nondegenerate,

(2) K has exactly n extreme rays, i.e., $|I| = n$,

(3) K^* is nondegenerate,

(4) K^* has exactly n extreme rays.

Proof. (1) $\Longrightarrow$ (2) holds by Lemma 2.1.36 and (2) $\Longrightarrow$ (1) holds by Corollary 2.1.37. (3) $\Longleftrightarrow$ (4) holds by replacing K with K^*. (2) $\Longleftrightarrow$ (3) holds by Corollary 2.1.38. ☐

So far, we constructed a proper cone given the vectors satisfying Hypothesis 3.1.1. The following remark states that given a proper nondegenerate cone, we can always find vectors satisfying Hypothesis 3.1.1.

Remark 3.1.15. [1] *Let P be any proper nondegenerate cone in $\mathbb{R}^n$ and let $a^i, \forall i = 1, \ldots, n$, be its extremal vectors with length 1. Then, there exists a unit vector $\bar{a} \in \mathbb{R}^n$, a set of unit vectors $\{\hat{a}^i\}_{i=1,\ldots,n} \subset \mathbb{R}^n$, and a scalar $\theta \in (0, \pi/2)$ such that $a^i = \cos\theta\bar{a} + \sin\theta\hat{a}^i$, $\forall i = 1, \ldots, n$.*

Proof. We note that $P = \operatorname{cone}(\{a^i\}_{i=1,\ldots,n})$. We define $A^T := \begin{bmatrix} a^1 & \cdots & a^n \end{bmatrix} \in \mathbb{R}^{n\times n}$. Consider the system $Ax = (1/\|A^{-1}e\|)e$. Since A is a nonsingular matrix, there is a unique $\bar{x}$ satisfying the system. Note that $\|\bar{x}\| = 1$ and let $\bar{a} = \bar{x}$. We also note that $\langle a^i, \bar{a}\rangle > 0$ and $\langle a^i, \bar{a}\rangle \leq \|a^i\|\|\bar{a}\| = 1$. If there exists $j \in \{1, \ldots, n\}$ such that $\langle a^j, \bar{a}\rangle = 1$, then $\langle a^i, \bar{a}\rangle = 1$,

[1] Remark 3.1.15 is not used later in this thesis.

$\forall i \in \{1, \ldots, n\}$. This yields $\bar{a} = a^i$, $\forall i \in \{1, \ldots, n\}$ and hence contradicts nonsingularity of A. Thus we have

$$0 < \langle a^i, \bar{a} \rangle < 1, \forall i \in \{1, \ldots, n\}.$$

In particular, we can find $\theta \in (0, \pi/2)$ such that $\cos\theta = \langle a^i, \bar{a} \rangle$, $\forall i \in \{1, \ldots, n\}$. For each $i = 1, \ldots, n$, we can also find a vector

$$\hat{a}^i \in \operatorname{span}\{\bar{a}, \hat{a}^i\} \text{ such that } \langle \bar{a}, \hat{a}^i \rangle = 0 \text{ and } \langle a^i, \hat{a}^i \rangle > 0.$$

Thus we have the result. $\square$

3.2 Properties of Two Sets of Interests

In this section, we investigate some properties of two sets we intend to study. In the later chapters, the sets we are about to study turn out to be the sets of dual optimal solutions of **LPs** we wish to study. The dual optimal solutions play an important role in sensitivity analysis.

3.2.1 Relations between $\{y : Cy = \bar{w}, y \geq 0\}$ and $\{z : Dz = \bar{w}, z \geq 0\}$

In this section, with the cones constructed in the previous section, we study special relations between the two polyhedral sets

$$\left\{y \in \mathbb{R}^k : Cy = \bar{w}, y \geq 0\right\} \text{ and } \left\{z \in \mathbb{R}^k : Dz = \bar{w}, z \geq 0\right\}.$$

Lemma 3.2.1. *If $\bar{w} = Cy$ for some $y \geq 0$, then there exists $z(y) \geq 0$ such that $\bar{w} = Dz$, $z \geq y$. In particular, we can choose z such that $z = \dfrac{\cos\bar{\theta}}{\cos\theta}\, y$.*

Proof. Using (3.1.2), we write

$$\begin{aligned}
\bar{w} &= Cy \\
&= \sum_{i=1}^{k} y_i c^i \\
&= \sum_{i=1}^{k} y_i (\cos\theta \bar{w} + \sin\theta w^i) \\
&= \sum_{i=1}^{k} y_i \cos\theta \bar{w} + \sum_{i=1}^{k} y_i \sin\theta w^i.
\end{aligned}$$

By rearranging the terms above (For simplicity, we use $\sum$ rather than $\sum_{i=1}^{k}$, when the meaning is clear.), we have

$$\sum y_i \sin\theta w^i \quad = \bar{w} - \sum y_i \cos\theta \bar{w}.$$

Thus $\sin\theta \sum y_i w^i = \eta\bar{w}$, for some $\eta \in \mathbb{R}$. Since $w^i \in \bar{w}^{\perp}, \forall i$, and $\sin\theta \neq 0$, the equality holds only when $\sum y_i w^i = 0$.

We also note that

$$\begin{aligned}
\frac{\cos\theta}{\cos\bar{\theta}} Dy &= Dy + \frac{\cos\theta - \cos\bar{\theta}}{\cos\bar{\theta}} Dy \\
&= Dy + \frac{\cos\theta - \cos\bar{\theta}}{\cos\bar{\theta}} \sum y_i \left(\cos\bar{\theta}\bar{w} + \sin\bar{\theta} w^i\right) \\
&= Dy + (\cos\theta - \cos\bar{\theta}) \sum y_i \bar{w} + \frac{(\cos\theta - \cos\bar{\theta})\sin\bar{\theta}}{\cos\bar{\theta}} \sum y_i w^i \\
&= Dy + (\cos\theta - \cos\bar{\theta}) \sum y_i \bar{w},
\end{aligned}$$

and

$$\begin{aligned}
\bar{w} &= Cy \\
&= \sum y_i (\cos\theta\bar{w} + \sin\theta w^i) \\
&= \cos\theta \sum y_i \bar{w} + \sin\theta \sum y_i w^i \\
&= \cos\bar{\theta} \sum y_i \bar{w} + \sin\bar{\theta} \sum y_i w^i + (\cos\theta - \cos\bar{\theta}) \sum y_i \bar{w} + (\sin\theta - \sin\bar{\theta}) \sum y_i w^i \\
&= Dy + (\cos\theta - \cos\bar{\theta}) \sum y_i \bar{w}.
\end{aligned}$$

Hence $Cy = \left(\cos\theta / \cos\bar{\theta}\right) Dy$. Letting $z = \left(\cos\theta / \cos\bar{\theta}\right) y$ gives the conclusion. $\square$

We note that $\left(\cos\theta / \cos\bar{\theta}\right) > 1$. This implies that if there is a vector $y \geq 0$ such that $Cy = \bar{w}$, there is always a vector z such that each nonzero coordinate of z is *strictly bigger* than the one of y satisfying $Dz = \bar{w}$. By observing the proof of Lemma 3.2.1, we see that the following also holds:

$$\text{If } z \geq 0 \text{ satisfies } \bar{w} = Dz, \text{ then } y = \frac{\cos\theta}{\cos\bar{\theta}} z \text{ satisfies } \bar{w} = Cy.$$

Example 3.2.2. *Let the following be the given data:*

$$\theta = \pi/6,\ \bar{\theta} = \pi/4,\ \bar{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, w^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, w^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, w^3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, w^4 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}.$$

Define the vectors

$$c^i := \cos\theta\bar{w} + \sin\theta w^i, \quad \text{and} \quad d^i := \cos\bar{\theta}\bar{w} + \sin\bar{\theta} w^i,\ \forall i = 1, \ldots, 4.$$

Define two 3×4 matrices C and D as follows:

$$C := \begin{bmatrix} 1/2 & 0 & -1/2 & 0 \\ 0 & 1/2 & 0 & -1/2 \\ \sqrt{3}/2 & \sqrt{3}/2 & \sqrt{3}/2 & \sqrt{3}/2 \end{bmatrix}, D := \begin{bmatrix} \sqrt{2}/2 & 0 & -\sqrt{2}/2 & 0 \\ 0 & \sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 & \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$$

Then

$$y^1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, y^2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, y^3 = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

are solutions to $\{y \in \mathbb{R}^4 : Cy = \bar{w},\ y \geq 0\}$. *Let*

$$z^i = \frac{\cos\theta}{\cos\bar{\theta}} y^i = \frac{\sqrt{3}}{\sqrt{2}} y^i,\ i = 1, 2, 3.$$

Then, by Lemma 3.2.1, each z^i *is a solution to* $\{z \in \mathbb{R}^4 : Dz = \bar{w},\ z \geq 0\}$. *We note that the solutions to* $\{y \in \mathbb{R}^4 : Cy = \bar{w},\ y \geq 0\}$ *and* $\{z \in \mathbb{R}^4 : Dz = \bar{w},\ z \geq 0\}$ *are not unique.* □

The rest of this section is not used later in this thesis. A reader may skip to Section 3.2.2. However we still present some interesting statements.

Remark 3.2.3. *Let* $\hat{\theta} = (\theta + \bar{\theta})/2$ *and define*

$$\hat{c}_i := \cos\hat{\theta}\bar{w} + \sin\hat{\theta}(-w^i),$$

with vectors $\bar{w}$ *and* w^i*'s given in Hypothesis 3.1.1. Let* $\hat{\mathcal{C}} = \{\hat{c}^1, \ldots, \hat{c}^k\}$ *and let* $\hat{C} = \begin{bmatrix} \hat{c}^1 & \cdots & \hat{c}^k \end{bmatrix} \in \mathbb{R}^{n \times k}$, *and define*

$$\operatorname{cone}(\hat{\mathcal{C}}) = \{x \in \mathbb{R}^n : x = \hat{C}z,\ z \geq 0\}.$$

Then, for all $x \in \operatorname{cone}(\mathcal{C})$ *with* $x = Cy$ *and* $y \geq 0$, *we have*

$$x = Cy = Dy + \hat{C}\gamma y,$$

where $\gamma = -2\sin\left(\frac{\theta - \bar{\theta}}{2}\right)$.

Proof. Suppose that $x \in \operatorname{cone}(\mathcal{C})$. Using the trigonometric relations

$$\cos\alpha - \cos\beta = -2\sin\left(\frac{\alpha - \beta}{2}\right)\sin\left(\frac{\alpha + \beta}{2}\right)$$

and

$$\sin\alpha - \sin\beta = 2\sin\left(\frac{\alpha - \beta}{2}\right)\cos\left(\frac{\alpha + \beta}{2}\right),$$

we obtain the following relation:

$$\begin{aligned}
x &= Cy \\
&= \cos\bar{\theta}\sum y_i\bar{w} + \sin\bar{\theta}\sum y_i w^i + \left(\cos\theta - \cos\bar{\theta}\right)\sum y_i\bar{w} + \left(\sin\theta - \sin\bar{\theta}\right)\sum y_i w^i \\
&= Dy + \left(\cos\theta - \cos\bar{\theta}\right)\sum y_i\bar{w} + \left(\sin\theta - \sin\bar{\theta}\right)\sum y_i w^i \\
&= Dy + \left(-2\sin\left(\tfrac{\theta-\bar{\theta}}{2}\right)\sin\left(\tfrac{\theta+\bar{\theta}}{2}\right)\right)\sum y_i\bar{w} + \left(2\sin\left(\tfrac{\theta-\bar{\theta}}{2}\right)\cos\left(\tfrac{\theta+\bar{\theta}}{2}\right)\right)\sum y_i w^i \\
&= Dy + \gamma\sin\hat{\theta}\sum y_i\bar{w} + \gamma\cos\hat{\theta}\sum y_i(-w^i) \\
&= Dy + \sin\hat{\theta}\sum(\gamma y_i)\bar{w} + \cos\hat{\theta}\sum(\gamma y_i)(-w^i) \\
&= Dy + \sum(\gamma y_i)\left(\sin\bar{\theta}\bar{w} + \cos\hat{\theta}(-w^i)\right) \\
&= Dy + \sum(\gamma y_i)\hat{c}^i \\
&= Dy + \hat{C}\gamma y,
\end{aligned}$$

where $\gamma = -2\sin\left(\frac{\theta-\bar{\theta}}{2}\right) > 0$. □

Remark 3.2.4. *The following are false statements:*

1. *If* $\mathrm{cone}(\mathcal{C}) \subset \mathrm{cone}(\mathcal{D})$, *then* $\mathrm{cone}(\{c^i + \epsilon e\}_{i=1,\ldots,k}) \subset \mathrm{cone}(\{d^i + \epsilon e\}_{i=1,\ldots,k})$.
2. $\mathrm{cone}(\{c^i + \epsilon e\}_{i=1,\ldots,k}) = \mathrm{cone}(\{c^1, \ldots, c^k\}) + \mathrm{cone}(\{\epsilon e, \ldots, \epsilon e\})$.
3. *For all* $y \geq 0$, *there exists* $\xi \geq 0$ *such that* $Cy = D(y + \xi)$ *(This means that the relation stated in Lemma 3.2.1 holds when we have* $\bar{w} = Cy$.*).*

3.2.2 Relations between $\{y : P^T y = \bar{w}, y \geq 0\}$ and $\{z : Q^T z = \bar{w}, z \geq 0\}$

We first note that using Lemma 2.1.32 and Algorithm 2.1.1, we can find vectors p^i such that

$$(\mathrm{cone}(\mathcal{C}))^* = \{x \in \mathbb{R}^n : C^T x \geq 0\} = \mathrm{cone}\left(\{p^1, \ldots, p^m\}\right).$$

Thus we have

$$\mathrm{cone}(\mathcal{C}) = \{x \in \mathbb{R}^n : Px \geq 0\},$$

where each $(p^i)^T$ is a row of matrix $P \in \mathbb{R}^{m\times n}$. Similarly, we can find vector q^i's such that

$$(\mathrm{cone}(\mathcal{D}))^* = \{x \in \mathbb{R}^n : D^T x \geq 0\} = \mathrm{cone}\left(\{q^1, \ldots, q^{m'}\}\right).$$

Thus we have

$$\mathrm{cone}(\mathcal{D}) = \{x \in \mathbb{R}^n : Qx \geq 0\},$$

where each $(q^i)^T$ is a row of matrix $Q \in \mathbb{R}^{m'\times n}$ (We show later in Lemma 3.2.6 and Corollary 3.2.8 that $m' = m$.).

In this section, we study a relation between $\{y : P^T y = \bar{w}, y \geq 0\}$ and $\{z : Q^T z = \bar{w}, z \geq 0\}$. In Section 3.2.1, we have studied the relation between $\{y : Cy = \bar{w}, y \geq 0\}$

and $\{z : Dz = \bar{w}, z \geq 0\}$. In the previous case, the columns of the data matrices C and D were explicitly given and hence obtaining the relation between two sets was easy. However, we do not know how the data matrices P and Q look like as each row of P and Q is obtained by computing the null spaces of certain matrices (See Algorithm 2.1.1.). In this section, we show that there is a similar relation, stated in Lemma 3.2.1, between the sets $\{y \in \mathbb{R}^m : P^T y = \bar{w}, y \geq 0\}$ and $\{z \in \mathbb{R}^m : Q^T z = \bar{w}, z \geq 0\}$. We first need a few lemmas in order to show the relation.

We recall that $\bar{w} \in \operatorname{int}(\operatorname{cone}(\mathcal{C}))$ by Lemma 3.1.9. The following lemma shows that $\bar{w}$ is also in the interior of $(\operatorname{cone}(\mathcal{C}))^*$.

Lemma 3.2.5. *Given $\bar{w} \in \mathbb{R}^n$ and $\{p^1, \ldots, p^m\} \subset \mathbb{R}^n$ as above, we have*

$$\bar{w} \in \operatorname{int}\left(\operatorname{cone}\left(\{p^1, \ldots, p^m\}\right)\right).$$

Proof. Note that $\operatorname{cone}(\{p^1, \ldots, p^m\}) = \{x \in \mathbb{R}^n : C^T x \geq 0\}$. Since $C = \begin{bmatrix} c^1 & \cdots & c^k \end{bmatrix}$ and $\langle c^i, \bar{w} \rangle > 0, \forall i$, we have $C^T \bar{w} > 0$. Hence $\bar{w} \in \operatorname{cone}(\{p^1, \ldots, p^m\})$. In particular, $\bar{w} \in \operatorname{int}(\operatorname{cone}(\{p^1, \ldots, p^m\}))$ due to the strict inequalities. ☐

Lemma 3.2.6. *Let p be an extremal vector of $\operatorname{cone}(\mathcal{C})^*$. Define the submatrix*

$$C_0 := C(:, \mathcal{I}) \in \mathbb{R}^{n \times (n-1)} \text{ of } C \text{ satisfying } C_0^T p = 0, \text{ where } |\mathcal{I}| = n - 1.$$

Then a vector q satisfying $D_0^T q = 0$, where $D_0 := D(:, \mathcal{I})$ is an extremal vector of $\operatorname{cone}(\mathcal{D})^*$.

We present an example before proving Lemma 3.2.6.

Example 3.2.7. *Let*

$$\bar{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, w^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, w^2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, w^3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } w^4 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}.$$

Given $\theta = \pi/6$ and $\bar{\theta} = \pi/4$, define

$$c^i = \cos\theta \bar{w} + \sin\theta w^i, \quad \text{and } d^i = \cos\bar{\theta} \bar{w} + \sin\bar{\theta} w^i, \ \forall i = 1, \ldots, 4.$$

Let $p = \left(1/2, 0, \sqrt{3}/2\right)^T$ be an extremal vector of the dual cone $\{x \in \mathbb{R}^3 : C^T x \geq 0\}$. We note that $p \in \operatorname{null}\left([c^1, c^2]^T\right)$. Let q be a vector satisfying $q \in \operatorname{null}\left([d^1, d^2]^T\right)$. Then $q = \frac{1}{\sqrt{3}}(1, 1, 1)^T$. We do not need to check $[d^1, d^2]^T$ determines an extremal vector of $(\operatorname{cone}(\mathcal{D}))^*$, because the statement of Lemma 3.2.6 guarantees that q is an extremal vector of $(\operatorname{cone}(\mathcal{D}))^*$. ☐

Now we prove Lemma 3.2.6.

Proof. We first show that

$$\dim\left(\operatorname{null}\left(D_0^T\right)\right) = 1.$$

Without loss of generality, we may assume that $\mathcal{I} = \{1, \dots, n-1\}$. Since $\{c^i\}_{i\in\mathcal{I}}$ are linearly independent, $\{d^i\}_{i\in\mathcal{I}}$ are linearly independent. Hence $D_0 \in \mathbb{R}^{n\times(n-1)}$ has rank $n-1$ and the dimension of null $\left(D_0^T\right)$ is 1. Thus there is only one vector satisfying $D_0^T q = 0$, up to scalar multiple. We choose $\|q\| = 1$ such that $\langle \bar{w}, q\rangle \geq 0$.

We define two hyperplanes:

$$\begin{aligned} h_C &= \{x \in \mathbb{R}^n : \langle p, x\rangle = 0, C_0^T p = 0\}, \\ h_D &= \{x \in \mathbb{R}^n : \langle q, x\rangle = 0, D_0^T q = 0\}. \end{aligned}$$

Now we show that $D_1^T q > 0$, where $D_1 := D(:, [k]\setminus\mathcal{I})$. Suppose to the contrary that there exists $j \in \{1, \dots, k\} \setminus \mathcal{I}$ such that $\langle d^j, q\rangle < 0$.

Let

$$T = \{\cos(\mu)\bar{w} + \sin(\mu)w^j : \mu \in [0, \bar{\theta}]\}$$

be the trajectory from $\bar{w}$ to d^j and define $T(\mu) := \cos\mu\bar{w} + \sin\mu w^j$. We show that T goes through the interior of cone$(\{d\}_{i\in\mathcal{I}})$ once. Let

$$l = \{\lambda\bar{w} + (1-\lambda)d^j : \lambda \in [0, 1]\}$$

be the line segment between $\bar{w}$ and d^j and define $l(\lambda) := \lambda\bar{w} + (1-\lambda)d^j$ (See Figure 3.2.1.).

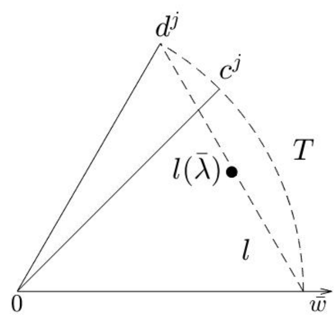


Figure 3.2.1: l is the line segment between $\bar{w}$ and d^j and T is the trajectory from $\bar{w}$ to d^j.

Since $\langle q, d^j\rangle < 0$ and $\langle q, \bar{w}\rangle \geq 0$, the hyperplane h_D separates d^j from $\bar{w}$. Hence l intersects h_D once. We show that l intersects cone$(\{d^i\}_{i\in\mathcal{I}}) \subset h_D$, in particular.

Let $l(\bar{\lambda})$ be the intersection of l and h_D. Suppose that the line segment l does not intersect cone$(\{d^i\}_{i\in\mathcal{I}})$, i.e., $l(\bar{\lambda}) \notin$ cone$(\{d^i\}_{i\in\mathcal{I}})$. Then

$$\text{cone}(\{d^i\}_{i\in\mathcal{I}}) \subsetneq \text{cone}\left(\{d^i\}_{i\in\mathcal{I}} \cup l(\bar{\lambda})\right) \subset h_D. \tag{3.2.1}$$

We note that cone $\left(\{d^i\}_{i\in\mathcal{I}} \cup l(\bar{\lambda})\right)$ is a pointed cone, since cone $\left(\{d^i\}_{i\in\mathcal{I}} \cup l(\bar{\lambda})\right) \subset$ cone$(\mathcal{D})$ and cone$(\mathcal{D})$ is a pointed cone by Theorem 3.1.8.

Since (3.2.1) holds, there exists $\ell \in \mathcal{I}$ such that

$$d^\ell = \sum_{i \in \mathcal{I} \setminus \ell}^{n-1} z_i d^i + z_j l(\bar{\lambda}) = \sum_{i \in \mathcal{I} \setminus \ell}^{n-1} z_i d^i + z_j (1 - \bar{\lambda}) d^j + z_j \bar{\lambda} \bar{w},$$

with $z_i \geq 0,\ \forall i \in \mathcal{I} \setminus \ell$ and $z_j \geq 0$ (See Figure 3.2.2.). We note that all the coefficients

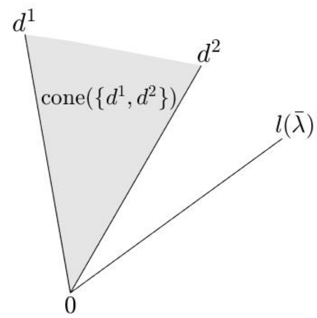


Figure 3.2.2: An illustration of a 2-dimensional subspace in $\mathbb{R}^3$. d^2 is in cone$(\{d^1, l(\bar{\lambda})\})$.

$z_i,\ z_j(1-\bar{\lambda})$ and $z_j\bar{\lambda}$ are nonnegative. This implies that d^ℓ is a conic combination of $\{d^i\}_{i \in \mathcal{I} \setminus l} \cup \{d^j\} \cup \{\bar{w}\}$ and hence this contradicts d^ℓ being an extremal vector of cone$(\mathcal{D})$. Therefore, l must intersect cone$(\{d^i\}_{i \in \mathcal{I}})$.

Since cone$(\{\bar{w}\} \cup \{c^i\}_{i \in \mathcal{I}}) \subsetneq$ cone$(\{\bar{w}\} \cup \{d^i\}_{i \in \mathcal{I}})$, l also intersects cone$(\{c^i\}_{i \in \mathcal{I}})$ and hence T intersects cone$(\{c^i\}_{i \in \mathcal{I}})$.

Each element x in cone$(\{c^i\}_{i \in \mathcal{I}} \cup \{\bar{w}\})$ satisfies $\arccos\left(\frac{\langle x, \bar{w} \rangle}{\|x\|\|\bar{w}\|}\right) \leq \theta$ by Lemma 3.1.5. In particular, $\arccos\left(\frac{\langle x, \bar{w} \rangle}{\|x\|\|\bar{w}\|}\right) < \theta$, if $x \neq c^i, \forall i \in \mathcal{I}$. Since $\langle c^j, \bar{w} \rangle = \cos\theta$ and $c^j \neq c^i, \forall i \in \mathcal{I}$, we have $c^j \notin$ cone$(\{c^i\}_{i \in \mathcal{I}} \cup \{\bar{w}\})$.

Since $\langle p, T(0) \rangle = \langle p, \bar{w} \rangle \geq 0$, the trajectory T starts from $\{x \in \mathbb{R}^n : \langle p, x \rangle \geq 0\}$. Since T intersects cone$(\{c^i\}_{i \in \mathcal{I}}) \subset h_C$ and

$$\text{cone}(\{c^i\}_{i \in \mathcal{I}} \cup \{\bar{w}\}) \subset \{x \in \mathbb{R}^n : \langle p, x \rangle \geq 0\},$$

the hyperplane h_C must separate c^j from $\bar{w}$. Hence, we have $\langle c^j, p \rangle < 0$ and this contradicts the hypothesis $C_1^T p > 0$. Thus we have $D_1^T q > 0$. □

Corollary 3.2.8. *Let q be an extremal vector of* cone$(\mathcal{D})^*$. *Define the submatrix $D_0 \in \mathbb{R}^{n \times (n-1)}$ of D satisfying $D_0^T p = 0$, where $|\mathcal{I}| = n-1$. Then a vector p satisfying $C_0^T q = 0$, where $C_0 := C(:, \mathcal{I})$ is an extremal vector of* cone$(\mathcal{C})^*$. □

The consequence of Lemma 3.2.6 and Corollary 3.2.8 is that the matrices P and Q have the same size. In other words, $\{x \in \mathbb{R}^n : Px \geq 0\}$ and $\{x \in \mathbb{R}^n : Qx \geq 0\}$ have the same number of inequalities. We also note that if cone$(\mathcal{C})$ is a nondegenerate cone, then $n = m = k$ by Proposition 3.1.14.

Throughout the thesis, if the i-th row $(p^i)^T$ of P is determined by null $\left(C_0^T\right)$, then we place the vector determined by null $\left(D_0^T\right)$ to the i-th row of Q.

Lemma 3.2.9. *Let p^T and q^T be i-th row of matrices P and Q, respectively. Let $C_{\mathcal{I}}$ be the maximal submatrix of C such that $C_{\mathcal{I}}^T p = 0$ and let $D_{\mathcal{I}}$ be the maximal submatrix of D such that $D_{\mathcal{I}}^T q = 0$, for some $\mathcal{I}$. Then, $\operatorname{rank}([\bar{w}, p, q]) = 2$. In particular, $q \in \operatorname{cone}(\{\bar{w}, p\})$.*

Proof. We define

$$\alpha := \cos\theta,\ \ \beta := \sin\theta,\ \ \bar{\alpha} := \cos\bar{\theta},\ \text{ and } \bar{\beta} := \sin\bar{\theta}.$$

Then we get

$$\begin{aligned} c^i &= \cos\theta\bar{w} + \sin\theta w^i = \alpha\bar{w} + \beta w^i, \text{ and} \\ d^i &= \cos\bar{\theta}\bar{w} + \sin\bar{\theta} w^i = \bar{\alpha}\bar{w} + \bar{\beta} w^i. \end{aligned}$$

Since p is an extremal vector of $\{x \in \mathbb{R}^n : C^T x \geq 0\}$, exactly $n-1$ halfspaces are active at p among $\langle c^i, x\rangle \geq 0,\ \forall i \in \{1, \dots, k\}$. If also follows that $|\mathcal{I}| = n-1$. Then

$$C_{\mathcal{I}} = \alpha W + \beta \overline{W} \text{ and } D_{\mathcal{I}} = \bar{\alpha} W + \bar{\beta}\overline{W},$$

where $W = \begin{bmatrix}\bar{w} & \cdots & \bar{w}\end{bmatrix} \in \mathbb{R}^{n\times(n-1)}$ and $\overline{W} = \begin{bmatrix}w^1 & \cdots & w^{n-1}\end{bmatrix} \in \mathbb{R}^{n\times(n-1)}$ (We may assume that the first $n-1$ vectors of $\{w^1, \dots, w^k\}$ form the matrix $\overline{W}$ for simplicity.).

Since

$$\bar{w} \in \operatorname{int}\left(\operatorname{cone}(\mathcal{C})\right) = \operatorname{int}\left(\{x \in \mathbb{R}^n : Px \geq 0\}\right),$$

we get $\langle p, \bar{w}\rangle > 0$. Let z be an orthonormal vector to $\bar{w}$, lying in $\operatorname{span}(\{p, \bar{w}\})$ and satisfying $\langle p, z\rangle > 0$ (See Figure 3.2.3.). Then, we can write

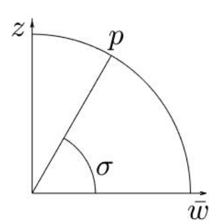


Figure 3.2.3: z satisfies $\langle z, \bar{w}\rangle = 0$, $z \in \operatorname{span}(\{p, \bar{w}\})$ and $\langle p, z\rangle > 0$.

$$p = \cos\sigma\bar{w} + \sin\sigma z,\ \text{ for some } \sigma \in (0, \pi/2).$$

Define

$$q' := \cos\sigma'\bar{w} + \sin\sigma' z,\ \text{ for some } \sigma' \in (0, \pi/2).$$

We want to show that $q = q'$. We note that

$$\begin{aligned} 0 = C_{\mathcal{I}}^T p &= \left(\alpha W^T + \beta \overline{W}^T\right) p \\ &= \alpha W^T p + \beta \overline{W}^T p \\ &= \alpha W^T (\cos\sigma \bar{w} + \sin\sigma z) + \beta \overline{W}^T (\cos\sigma \bar{w} + \sin\sigma z) \\ &= \alpha \cos\sigma W^T \bar{w} + \alpha \sin\sigma W^T z + \beta \cos\sigma \overline{W}^T \bar{w} + \beta \sin\sigma \overline{W}^T z. \end{aligned} \tag{3.2.2}$$

Since $W^T \bar{w} = e$ and $W^T z = \overline{W}^T \bar{w} = 0$, it follows that

$$0 = \alpha \cos\sigma e + \beta \sin\sigma \overline{W}^T z.$$

Thus we have

$$e = -\frac{\beta \sin\sigma}{\alpha \cos\sigma} \overline{W}^T z. \tag{3.2.3}$$

Similar expansion of $D_{\mathcal{I}}^T q'$ used in (3.2.2) gives

$$D_{\mathcal{I}}^T q' = D_{\mathcal{I}}^T (\cos\sigma' \bar{w} + \sin\sigma' z) = \bar{\alpha} \cos\sigma' e + \bar{\beta} \sin\sigma' \overline{W}^T z. \tag{3.2.4}$$

Plugging (3.2.3) into (3.2.4) yields

$$D_{\mathcal{I}}^T q' = \left[-\bar{\alpha} \cos\sigma' \frac{\beta \sin\sigma}{\alpha \cos\sigma} + \bar{\beta} \sin\sigma'\right] \overline{W}^T z. \tag{3.2.5}$$

We make an observation on the coefficient of the RHS in (3.2.5):

$$\begin{aligned} -\bar{\alpha} \cos\sigma' \frac{\beta \sin\sigma}{\alpha \cos\sigma} + \bar{\beta} \sin\sigma' &= -\cos\bar{\theta} \cos\sigma' \frac{\sin\theta \sin\sigma}{\cos\theta \cos\sigma} + \sin\bar{\theta} \sin\sigma' \\ &= \sin\bar{\theta} \sin\sigma' \left(-\frac{\cos\bar{\theta}}{\sin\bar{\theta}} \frac{\cos\sigma'}{\sin\sigma'} \frac{\sin\theta \sin\sigma}{\cos\theta \cos\sigma} + 1\right) \\ &= \sin\bar{\theta} \sin\sigma' \left(-\frac{\tan\theta \tan\sigma}{\tan\bar{\theta} \tan\sigma'} + 1\right). \end{aligned}$$

Since the range of the tangent function is $(-\infty, \infty)$, there exists $\sigma' \in (-\pi/2, \pi/2)$ such that

$$\tan\sigma' = (\tan\sigma \tan\theta) / \tan\bar{\theta}.$$

Since $\tan\sigma, \tan\theta, \tan\bar{\theta} > 0$, we must have $\tan\sigma' > 0$. Thus $\sigma' \in (0, \pi/2)$. We may take such σ' and we have $D_{\mathcal{I}}^T q' = 0$ and $\|q'\| = 1$.

Since rank $\left(D_{\mathcal{I}}^T\right) = n - 1$, there exists only one vector x that satisfies $D_{\mathcal{I}}^T x = 0$ up to scalar multiple. If there are two vectors x^1, x^2 such that $\|x^1\| = \|x^2\| = 1$ and $D_{\mathcal{I}}^T x^1 = D_{\mathcal{I}}^T x^2 = 0$, x^1 and x^2 can hold only one of the following two cases: $x^1 = -x^2$ or $x^1 = x^2$.

We note that

$$\langle \bar{w}, q' \rangle = \cos\sigma' \langle \bar{w}, \bar{w} \rangle + \sin\sigma' \langle \bar{w}, z \rangle = \cos\sigma' > 0,$$

and

$$\bar{w} \in \text{cone}(\mathcal{D}) = \{x \in \mathbb{R}^n : Qx \geq 0\} \implies \langle \bar{w}, q \rangle \geq 0.$$

Therefore, q and q' cannot have different signs and so we have

$$q = q' = \cos\sigma' \bar{w} + \sin\sigma' z.$$

Thus $\text{rank}([\bar{w}, p, q]) = 2$. Since polarization is order-reversing (See Lemma 2.1.25.), we have

$$\text{cone}(\mathcal{C}) \subset \text{cone}(\mathcal{D}) \implies \text{cone}\left(\{p^i\}_{i \in \{1,\ldots,m\}}\right) \supset \text{cone}\left(\{q^i\}_{i \in \{1,\ldots,m\}}\right).$$

Therefore, $q \in \text{cone}(\{\bar{w}, p\})$. $\square$

We can represent Lemma 3.2.9 pictorially (See Figure 3.2.4.). For each 2-dimensioanl subspace of $\mathbb{R}^n$, $\hat{w}^i$ plays the role of z in the proof of Lemma 3.2.9. We pay attention to the subscript i of θ_i. We may have different θ_i's for each two-dimensional subspace (Note that in Figure 3.1.4, same θ is used in each two-dimensional subspace in $\mathbb{R}^n$.).

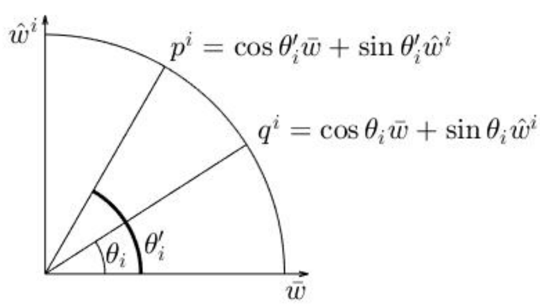


Figure 3.2.4: p^i is rotated further from $\bar{w}$ than q^i.

Now we are ready to show the special property we mentioned at the beginning of this section.

Lemma 3.2.10. *For all $y \in \{y : P^T y = \bar{w}, y \geq 0\}$, there exists $z \in \{z : Q^T z = \bar{w}, z \geq 0\}$ such that $y \geq z$.*

Proof. By Lemma 3.2.9, we have

$$\begin{aligned} q^i &= \cos\theta_i \bar{w} + \sin\theta_i \hat{w}^i, \quad \text{and} \\ p^i &= \cos\theta_i' \bar{w} + \sin\theta_i' \hat{w}^i, \end{aligned}$$

where each $\hat{w}^i$ is an orthonormal vector to $\bar{w}$ and lying in $\text{span}(\{\bar{w}, p^i\})$.

Suppose that $\bar{w} = P^T y$. Then,

$$\bar{w} = P^T y = \sum y_i \cos\theta_i' \bar{w} + \sum y_i \sin\theta_i' \hat{w}^i.$$

Then, for some $\eta \in \mathbb{R}$,

$$\eta \bar{w} = \sum y_i \sin\theta_i' \hat{w}^i. \tag{3.2.6}$$

Since $\hat{w}_i \in \bar{w}^\perp$, $\forall i = 1, \dots, m$, the equality (3.2.6) holds only when $\eta = 0$ and hence

$$\sum y_i \sin\theta_i' \hat{w}^i = 0. \tag{3.2.7}$$

We note that

$$Q^T y = \sum y_i \cos\theta_i \bar{w} + \sum y_i \sin\theta_i \hat{w}^i.$$

Then, we have

$$\begin{aligned} P^T y &= \sum y_i \cos\theta_i' \bar{w} + \sum y_i \sin\theta_i' \hat{w}^i \\ &= \sum y_i \cos\theta_i' \bar{w} + \sum y_i \sin\theta_i' \hat{w}^i + \sum y_i(\cos\theta_i - \cos\theta_i)\bar{w} + \sum y_i(\sin\theta_i - \sin\theta_i)\hat{w}^i \\ &= \sum y_i \cos\theta_i \bar{w} + \sum y_i \sin\theta_i \hat{w}^i + \sum y_i(\cos\theta_i' - \cos\theta_i)\bar{w} + \sum y_i(\sin\theta_i' - \sin\theta_i)\hat{w}^i \\ &= Q^T y + \sum y_i(\cos\theta_i' - \cos\theta_i)\bar{w} + \sum y_i(\sin\theta_i' - \sin\theta_i)\hat{w}^i \\ &= Q^T y + \sum y_i(\cos\theta_i' - \cos\theta_i)\bar{w} - \sum y_i \sin\theta_i \hat{w}^i. \end{aligned} \tag{3.2.8}$$

The last equality holds by (3.2.7).

Let $y_i' = \frac{\cos\theta_i'}{\cos\theta_i} y_i$, $\forall i = 1, \dots, m$. We note that $\frac{\cos\theta_i'}{\cos\theta_i} < 1$. Then, we have

$$\begin{aligned} Q^T y' &= \sum y_i' q^i \\ &= \sum \frac{\cos\theta_i'}{\cos\theta_i} y_i q^i \\ &= \sum \frac{\cos\theta_i'}{\cos\theta_i} y_i q^i + \sum y_i(\cos\theta_i - \cos\theta_i)\bar{w} + \sum y_i(\sin\theta_i - \sin\theta_i)\hat{w}^i \\ &= \sum y_i \cos\theta_i \bar{w} + \sum y_i \sin\theta_i \hat{w}^i + \sum \frac{\cos\theta_i'}{\cos\theta_i} \cos\theta_i y_i \bar{w} + \sum \frac{\cos\theta_i'}{\cos\theta_i} \sin\theta_i y_i \hat{w}^i \\ &\quad - \sum y_i \cos\theta_i \bar{w} - \sum y_i \sin\theta_i \hat{w}^i \\ &= Q^T y + \sum y_i(\cos\theta_i' - \cos\theta_i)\bar{w} + \sum y_i \left(\frac{\cos\theta_i' \sin\theta_i}{\cos\theta_i} - \sin\theta_i \right) \hat{w}^i. \end{aligned} \tag{3.2.9}$$

We focus on the last term of (3.2.9). If we show

$$\sum y_i \frac{\cos\theta_i' \sin\theta_i}{\cos\theta_i} \hat{w}^i = 0, \tag{3.2.10}$$

then the last line of (3.2.8) and the last line of (3.2.9) are the same, and hence we have $P^T y = Q^T y'$. Thus, if we show (3.2.10), we are done.

Let p^j be an extremal vector formed by $\left\{(c^i)^T\right\}_{i \in \mathcal{I}}$, for some $\mathcal{I} \subset \{1, \dots, k\}$ and $|\mathcal{I}| =$

$n-1$. Then, $\forall i \in \mathcal{I}$, we have $\langle p^j, c^i \rangle = 0$, since p^j is in the orthogonal complement of $\operatorname{span}(\{c^i\}_{i\in\mathcal{I}})$. Hence, for $i \in \mathcal{I}$,

$$\begin{aligned} 0 &= \langle p^j, c^i \rangle \\ &= \left\langle \cos\theta'_j \bar{w} + \sin\theta'_j \hat{w}^j \ , \ \cos\theta \bar{w} + \sin\theta w^i \right\rangle \\ &= \cos\theta'_j \cos\theta \bar{w}^T \bar{w} + \cos\theta'_j \sin\theta \bar{w}^T w^i + \sin\theta'_j \cos\theta \bar{w}^T \hat{w}^j + \sin\theta'_j \sin\theta (\hat{w}^j)^T w^i \\ &= \cos\theta'_j \cos\theta \bar{w}^T \bar{w} + \sin\theta'_j \sin\theta (\hat{w}^j)^T w^i. \end{aligned}$$

Therefore, we have

$$\cos\theta'_j \cos\theta = -\sin\theta'_j \sin\theta (w^i)^T \hat{w}^j \implies (w^i)^T \hat{w}^j = -\frac{\cos\theta'_j \cos\theta}{\sin\theta'_j \sin\theta} = -\frac{1}{\tan\theta'_j \tan\theta}. \tag{3.2.11}$$

Similarly, let q^j be an extremal vector formed by $\left\{(d^i)^T\right\}_{i\in\mathcal{I}}$. Then, $\forall i \in \mathcal{I}$, we have $\langle q^j, d^i \rangle = 0$, since q^j is in the orthogonal complement of $\operatorname{span}(\{d^i\}_{i\in\mathcal{I}})$. Thus, for $i \in \mathcal{I}$,

$$\begin{aligned} 0 &= \langle q^j, d^i \rangle \\ &= \left\langle \cos\theta_j \bar{w} + \sin\theta_j \hat{w}^j \ , \ \cos\bar{\theta} \bar{w} + \sin\bar{\theta} w^i \right\rangle \\ &= \cos\theta_j \cos\bar{\theta} \bar{w}^T \bar{w} + \cos\theta_j \sin\bar{\theta} \bar{w}^T w^i + \sin\theta_j \cos\bar{\theta} \bar{w}^T \hat{w}^j + \sin\theta_j \sin\bar{\theta} (\hat{w}^j)^T w^i \\ &= \cos\theta_j \cos\bar{\theta} \bar{w}^T \bar{w} + \sin\theta_j \sin\bar{\theta} (\hat{w}^j)^T w^i. \end{aligned}$$

Therefore,

$$\cos\theta_j \cos\bar{\theta} = -\sin\theta_j \sin\bar{\theta} (w^i)^T \hat{w}^j \implies (w^i)^T \hat{w}^j = -\frac{\cos\theta_j \cos\bar{\theta}}{\sin\theta_j \sin\bar{\theta}} = -\frac{1}{\tan\theta_j \tan\bar{\theta}}. \tag{3.2.12}$$

Thus, (3.2.11) and (3.2.12) imply that

$$\frac{1}{\tan\theta'_j \tan\theta} = \frac{1}{\tan\theta_j \tan\bar{\theta}} \iff \frac{\tan\bar{\theta}}{\tan\theta} = \frac{\tan\theta'_j}{\tan\theta_j}, \quad \forall j \in \{1, \ldots, m\}.$$

We note that $\tan\theta_j / \tan\theta'_j$ is the same constant, $\forall j \in \{1, \ldots, m\}$. Let $\gamma := \tan\theta / \tan\bar{\theta}$ be the constant.

We observe the following:

$$\begin{aligned} \sum y_i \left(\frac{\cos\theta'_i \sin\theta_i}{\cos\theta_i} \right) \hat{w}^i &= \sum y_i \left(\frac{\cos\theta'_i \sin\theta_i}{\sin\theta'_i \cos\theta_i} \right) \sin\theta'_i \hat{w}^i \\ &= \sum y_i \frac{\tan\theta_i}{\tan\theta'_i} \sin\theta'_i \hat{w}^i \\ &= \gamma \sum y_i \sin\theta'_i \hat{w}^i \\ &= 0. \end{aligned} \tag{3.2.13}$$

The last equality of (3.2.13) holds by (3.2.7). Therefore, (3.2.13) verifies (3.2.10), and we are done. $\square$

By observing the proof of Lemma 3.2.10, we see that the following also holds:

$$\text{If } z \geq 0 \text{ satisfies } Q^T z = \bar{w}, \text{ then there exists } y \text{ satisfying } P^T y = \bar{w} \text{ and } y \geq z.$$

We note that Lemma 3.2.10 looks similar to Lemma 3.2.1. Lemma 3.2.1 tells us that

$$\bar{w} = Cy, \text{ for some } y \geq 0 \implies \exists\ z = \frac{\cos\bar{\theta}}{\cos\theta} y \text{ such that } \bar{w} = Dz.$$

In other words, Lemma 3.2.1 tells us that given $y \geq 0$ satisfying $\bar{w} = Cy$, there is $z \geq 0$ with $\bar{w} = Dz$ and all nonzero coordinates of z are *proportionally bigger* than the nonzero coordinates of y. However, Lemma 3.2.10 tells us slightly different property. It tells us that given $\bar{y} \geq 0$ satisfying $P^T\bar{y} = \bar{w}$, there is a $\bar{z} \geq 0$ with $Q^T\bar{z} = \bar{w}$ and all nonzero coordinates of $\bar{z}$ are *just bigger* than the nonzero coordinates of $\bar{y}$. Hence given such $\bar{y}$, we cannot clearly specify each coordinate of $\bar{z}$, unless we perform further computations.

Chapter 4

Sensitivity of Optimal Value Function

In this chapter, we present the main results for this thesis. Specifically, we perform postoptimal analysis to show that the different sizes of the cones (or, feasible regions) affect robustness of the optimal value. The sensitivity analysis in this thesis has some different aspects from the classical sensitivity analysis of linear programming. While the classical sensitivity analysis focuses on obtaining allowable perturbation ranges, this thesis focuses on the change of the optimal value under the assumption that some reasonable bounds for the perturbations are given. We also view the problem geometrically, i.e., we relate the dual optimal solutions to the geometrical structures constructed in Chapter 3. We can perform the postoptimal analysis relatively easier, because of our knowledge of the primal-dual optimal solutions for our models. This is shown in Sections 4.1 and 4.2 below.

This chapter is organized as follows: in Section 4.1, we define two families of **LPs**, **LP**$(\mathcal{P})$ and **LP**$(\mathcal{Q})$, in order to study their sensitivities. In Section 4.2, we study the optimal solutions to instances $(\mathcal{P})$ and $(\mathcal{Q})$ in the two families, as well as their dual optimal solutions. In Section 4.3, we study some properties of the optimal value functions of $(\mathcal{P})$ and $(\mathcal{Q})$. We then consider the **LPs** divided into two classes: nondegenerate (Section 4.4) and degenerate (Section 4.5). In Section 4.4, we study the strict monotonicity and the sensitivity of the optimal value function, when the given **LP** is nondegenerate. In Section 4.5, we study the sensitivity of the optimal value function via directional differentiability, when the given **LP** is degenerate. We also further study a sufficient condition for local differentiability of the optimal value functions of $(\mathcal{P})$ and $(\mathcal{Q})$.

4.1 LP Models

With the data from Section 3.1, we consider a vector $\bar{w}$ for the objective function and two classes of cones using the set and matrix $\mathcal{C}, C$, respectively. (See Figure 4.1.1.)

1. $\text{cone}(\mathcal{C}) = \{x \in \mathbb{R}^n : Px \geq 0\}$,
2. $\text{cone}(\mathcal{C})^\circ = \{x \in \mathbb{R}^n : C^T x \leq 0\}$, i.e., the negative polar of $\text{cone}(\mathcal{C})$.

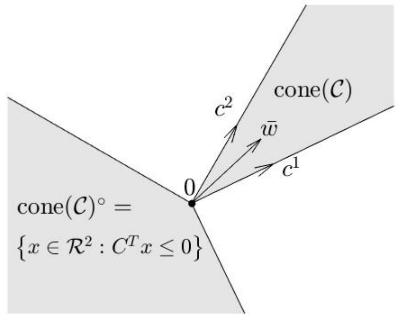


Figure 4.1.1: An illustration of cone($\mathcal{C}$) and its negative polar cone in $\mathbb{R}^2$.

Using these two cones, we define two **LPs** with nonnegativity constraints on the variables:

1. $\min\{\langle \bar{w}, x\rangle : Px \geq 0, x \geq 0\}$,
2. $\min\{\langle -\bar{w}, x\rangle : C^T x \leq 0, x \geq 0\}$.

In the presence of data uncertainty, we wish to form the robust counterparts of the above **LPs**. Before forming their robust counterparts, we consider a general form of an **LP**.

Given

$$\textbf{LP} \qquad \min\{\langle c, x\rangle : Ax \geq b, x \geq 0\}, \tag{4.1.1}$$

with $A \in \mathbb{R}^{m\times n}$, suppose that each entry of the data matrix A is uncertain and each uncertain entry is known to have perturbation range $[-\epsilon, \epsilon]$, with $\epsilon > 0$. It is shown in (2.4.9) that the robust counterpart of each constraint $a_i^T x \geq b_i$ of (4.1.1) is

$$a_i^T x - \epsilon e^T x \geq b_i.$$

Hence, the robust counterpart of (4.1.1) is

$$\min\{\langle c, x\rangle : Ax - \epsilon Ex \geq b,\ x \geq 0\},$$

where we recall that $E = e^m(e^n)^T$ is the matrix of ones.

If b above is the zero vector, then the constraint system $Ax \geq b$ of (4.1.1) forms a homogeneous system. If we further assume that 0 is the unique optimal solution to (4.1.1), and we consider the robust counterpart of the homogeneous system $Ax \geq 0$, i.e.,

$$\min\{\langle c, x\rangle : Ax - \epsilon Ex \geq 0,\ x \geq 0\}, \tag{4.1.2}$$

then the uncertainty does not do anything to the problem as 0 remains the optimal solution to (4.1.2). Hence, having $b = 0$ trivializes the problem even after perturbations. This motivates us to translate the cones away from the origin.

In our problems, we translate the two cones cone($\mathcal{C}$) and $(\text{cone}(\mathcal{C}))^\circ$ by e to obtain a cone with vertex at e, (see Definition 2.1.12) and optimal solution at e. We thus have a nonzero RHS in our constraint systems (See Figure 4.1.2.):

$$A(x-e) \geq 0 \iff Ax \geq Ae. \tag{4.1.3}$$

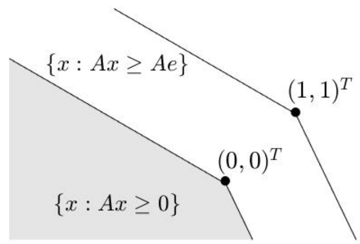


Figure 4.1.2: An illustration in $\mathbb{R}^2$: The shaded region is the feasible region of the original cone.

Hence with (4.1.3), its robust counterpart becomes

$$\min\{\langle c, x\rangle : Ax - \epsilon Ex \geq Ae,\ x \geq 0\}. \tag{4.1.4}$$

Now we are ready to define the **LPs** for sensitivity analysis. We define an **LP**

$$(\mathcal{P}) \qquad \begin{array}{cl} \min & \langle \bar{w}, x\rangle \\ \text{subject to} & Px \geq Pe \\ & x \geq 0. \end{array}$$

Suppose that the data matrix P on the LHS of the constraints $Px \geq Pe$ of $(\mathcal{P})$ is uncertain, and each uncertain entry of P is known to take perturbations in the range $[-\epsilon, \epsilon]$, with $\epsilon > 0$. Then the robust counterpart of $(\mathcal{P})$ is

$$(\mathcal{R}(\epsilon)) \qquad \begin{array}{cl} \min & \langle \bar{w}, x\rangle \\ \text{subject to} & Px - \epsilon Ex \geq Pe \\ & x \geq 0. \end{array} \tag{4.1.5}$$

Similarly, we define an **LP**

$$(\mathcal{Q}) \qquad \begin{array}{cl} \min & \langle -\bar{w}, x\rangle \\ \text{subject to} & C^T x \leq C^T e \\ & x \geq 0. \end{array}$$

Suppose that the data matrix C^T on the LHS of the constraints $C^T x \leq C^T e$ is uncertain, and each entry of C^T is known to take perturbations in the range $[-\epsilon, \epsilon]$. Then the robust counterpart of $(\mathcal{Q})$ is

$$(\mathcal{S}(\epsilon)) \qquad \begin{array}{rl} \min & \langle -\bar{w}, x\rangle \\ \text{subject to} & C^T x + \epsilon E x \leq C^T e \\ & x \geq 0. \end{array} \tag{4.1.6}$$

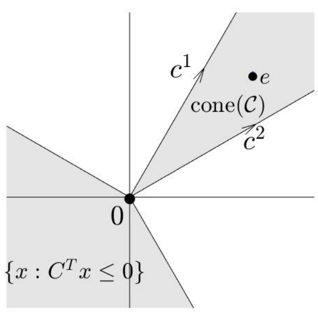
Example 4.1.1. *Consider the cone generated by* $c^1 = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix}, c^2 = \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}$. *Then* $\text{cone}(\{c^1, c^2\}) = \{x \in \mathbb{R}^2 : Px \geq 0\}$, *where* $P = \begin{bmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix}$. *We translate* $\text{cone}(\{c^1, c^2\})$ *by* e *and obtain the system*

$$\{x \in \mathbb{R}^2 : Px \geq Pe\} = \left\{ x \in \mathbb{R}^2 : \begin{bmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix} x \geq \begin{bmatrix} (\sqrt{3}-1)/2 \\ (\sqrt{3}-1)/2 \end{bmatrix} \right\}.$$

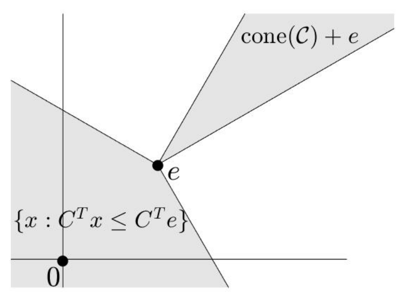
The negative polar cone of $\text{cone}(\{c^1, c^2\})$ *is* $\{x \in \mathbb{R}^2 : C^T x \leq 0\}$, *where* $C = [c^1, c^2]$. *We translate the vertex of* $\{x \in \mathbb{R}^2 : C^T x \leq 0\}$ *to* e *and obtain*

$$\{x \in \mathbb{R}^2 : C^T x \leq C^T e\} = \left\{ x \in \mathbb{R}^2 : \begin{bmatrix} \sqrt{3}/2 & 1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix} x \leq \begin{bmatrix} (\sqrt{3}+1)/2 \\ (\sqrt{3}+1)/2 \end{bmatrix} \right\}.$$

See Figure 4.1.3. □



(a) cone($\mathcal{C}$) and (cone($\mathcal{C}$))$^\circ$.



(b) cone($\mathcal{C}$) + e and (cone($\mathcal{C}$))$^\circ$ + e.

Figure 4.1.3: Two cones in (b) are translations of the cones in (a).

4.2 Primal and Dual Optimal Solutions to $(\mathcal{P})$ and $(\mathcal{Q})$

In this section, we find the primal and dual optimal solutions to $(\mathcal{P})$ (Section 4.2.1) and $(\mathcal{Q})$ (Section 4.2.2). We also show that the uniqueness of dual optimal solutions to $(\mathcal{P})$ and $(\mathcal{Q})$ depends on degeneracy/nondegeneracy of cone($\mathcal{C}$).

4.2.1 The Optimal Solutions to $(\mathcal{P})$ and its Dual

We recall the primal problem

$$(\mathcal{P}) \qquad \begin{array}{rl} \min & \langle \bar{w}, x\rangle \\ \text{subject to} & Px \geq Pe \in \mathbb{R}^m \\ & x \in \mathbb{R}^n_+. \end{array}$$

Then, the dual $(\mathcal{DP})$ of the primal $(\mathcal{P})$ is

$$(\mathcal{DP}) \qquad \begin{array}{rl} \max & \langle Pe, y\rangle \\ \text{subject to} & P^T y \leq \bar{w} \\ & y \geq 0. \end{array}$$

Recall that the homogeneous inequalities $Px \geq 0$ define cone$(\mathcal{C})$. We now see that the shift to $Px \geq Pe$ moves the vertex of the cone to e, which is the unique optimal solution to $(\mathcal{P})$. In addition, we see that degeneracy of cone$(\mathcal{C})$ directly corresponds to degeneracy of the dual optimal solutions. In other words, the number of extremal vectors of cone$(\mathcal{C})$ directly relates to the number of dual optimal solutions.

Lemma 4.2.1. *Given the primal-dual pair $(\mathcal{P})$ and $(\mathcal{DP})$ above, we have the following:*

1. *The set of optimal solutions to $(\mathcal{DP})$ satisfies*

$$\text{Argmax}\left\{\langle Pe, y\rangle : P^T y \leq \bar{w},\ y \geq 0\right\} = \left\{y \in \mathbb{R}^m : P^T y = \bar{w},\ y \geq 0\right\}.$$

2. *The unique optimal solution to $(\mathcal{P})$ is $x = e$.*

3. *If* cone$(\mathcal{C})$ *is a nondegenerate cone, then $(\mathcal{DP})$ has a unique optimal solution.*

4. *If* cone$(\mathcal{C})$ *is a degenerate cone, then $(\mathcal{DP})$ does not have a unique optimal solution.*

Proof. 1. We note that e is a feasible solution to $(\mathcal{P})$. We observe that

$$\langle \bar{w}, e\rangle = \langle P^T y, e\rangle = \langle Pe, y\rangle,$$

for all $y \geq 0$ satisfying $P^T y = \bar{w}$. In other words, for each nonnegative vector y such that $P^T y = \bar{w}$, y achieves a primal objective value $\langle \bar{w}, e\rangle$. Hence, $\langle \bar{w}, e\rangle$ is the optimal value of $(\mathcal{DP})$ and

$$\{y \in \mathbb{R}^m : P^T y = w,\ y \geq 0\} \subset \text{Argmax}\left\{\langle Pe, y\rangle : P^T y \leq \bar{w},\ y \geq 0\right\}.$$

Now suppose to the contrary that

$$\exists y' \in \text{Argmax}\left\{\langle Pe, y\rangle : P^T y \leq \bar{w},\ y \geq 0\right\} \setminus \{y \in \mathbb{R}^m : P^T y = w,\ y \geq 0\}.$$

Then, there exists $i \in \{1, \dots, n\}$ such that $e_i^T P^T y' < e_i^T \bar{w}$, i.e., one of the inequalities in $P^T y' \leq \bar{w}$ must be strict. Then, we get

$$\langle \bar{w}, e \rangle > \langle P^T y', e \rangle = \langle Pe, y' \rangle,$$

and y' does not achieve the primal objective value, which means that y' is not an optimal solution. Therefore, $\{y \in \mathbb{R}^m : P^T y = \bar{w},\ y \geq 0\}$ is the set of optimal solutions to $(\mathcal{DP})$.

2. By observing the proof in Item 1 of Lemma 4.2.1 directly above, it also follows that e is an optimal solution to $(\mathcal{P})$. Here, we show that e is not only an optimal solution to $(\mathcal{P})$, but also the unique optimal solution to $(\mathcal{P})$.

   Suppose that x' is an optimal solution to $(\mathcal{P})$. By Lemma 3.2.5 we get

   $$\bar{w} \in \operatorname{int}\big(\operatorname{cone}(\{p^1, \dots, p^m\})\big).$$

   Therefore we know that there exists a dual optimal solution $y' > 0$ such that $P^T y' = \bar{w}$. Hence, we have

   $$0 \leq \langle Px' - Pe, y' \rangle = \langle P^T y', x' \rangle - \langle P^T y', e \rangle = \langle \bar{w}, x' \rangle - \langle P^T y', e \rangle = 0.$$

   The last inequality holds by strong duality. Thus we get

   $$\langle Px' - Pe, y' \rangle = 0,\ y' > 0 \implies Px' = Pe.$$

   Since $P \in \mathbb{R}^{m \times n}$ with $m \geq n$ and $\operatorname{rank}(P) = n$, we have $\operatorname{rank}(P^T P) = n$. Hence,

   $$P^T P x' = P^T P e \implies x' = e.$$

   Therefore, $x' = e$ is the only optimal solution to $(\mathcal{P})$.

3. Since $\operatorname{cone}(\mathcal{C})$ is nondegenerate, the dual cone has exactly n extremal vectors by Proposition 3.1.14. By Lemma 3.2.5, we have $\bar{w} \in \operatorname{cone}(\{p^1, \dots, p^n\})$. Hence, there exists a nonnegative vector y such that $\bar{w} = P^T y$. Since P is a nonsingular matrix, the system $P^T y = \bar{w}$ has the unique solution.

4. Since $\operatorname{cone}(\mathcal{C})$ is degenerate, $\operatorname{cone}(\{p^1, \dots, p^m\})$ has more than n extremal vectors by Proposition 3.1.14. Since $\operatorname{cone}(\{p^1, \dots, p^m\})$ is a proper cone, it has a compact cone base B, say $B = \{b^1, \dots, b^m\}$. Let $\bar{w}_B \in B$, where $\bar{w}_B = \gamma \bar{w}$, for some $\gamma > 0$.

   By Minkowski's theorem, we need at most n extreme points in B to represent $\bar{w}_B$ as a convex combination of extreme points, since $\dim(\operatorname{conv}(B)) = n - 1$. Hence, we have

   $$\bar{w}_B = \sum_{i \in \mathcal{I}} \lambda_i b^i,\ \lambda \geq 0 \text{ and } \sum_{i \in \mathcal{I}} \lambda_i = 1, \quad \text{for some } \lambda \in \mathbb{R}^n \text{ and } |\mathcal{I}| = n.$$

Let $b^j \in B$, where $j \in [m] \setminus \mathcal{I}$. Then, by Lemma 2.1.5, we can write $\bar{w}_B$ as a convex combination that involves b^j. Since there are at least two distinct convex combinations for $\bar{w}_B$, $(\mathcal{DP})$ does not have a unique optimal solution. $\square$

Remark 4.2.2. *We note that previously in Section 3.2.2 we studied the relations between the two sets*

$$\{y \in \mathbb{R}^m : P^T y = \bar{w},\ y \geq 0\},\ \{y \in \mathbb{R}^m : Q^T y = \bar{w},\ y \geq 0\}.$$

The above results hold for both sets, i.e., for both P, Q.

Remark 4.2.3. *Given a feasible primal-dual* ***LP*** *pair, it is well-known that there is a pair of primal-dual optimal solutions that satisfies the strict complementarity. We note that there is a vector $y' > 0$ satisfying $P^T y' = \bar{w}$ by Lemma 3.2.5. Hence, with $x' = e$, we see that (x', y') is a pair of optimal solutions of $(\mathcal{P})$ and $(\mathcal{DP})$ that holds the strict complementarity.*

4.2.2 The Optimal Solutions to $(\mathcal{Q})$ and its Dual

The argument in this section parallels the argument used in Section 4.2.1. A reader may skip the proof given in this section. We recall that

$$(\mathcal{Q}) \qquad \begin{array}{rl} \min & \langle -\bar{w}, x \rangle \\ \text{subject to} & C^T x \leq C^T e \\ & x \geq 0. \end{array}$$

Then, the dual $(\mathcal{DQ})$ of $(\mathcal{Q})$ is

$$(\mathcal{DQ}) \qquad \begin{array}{rl} \max & \langle -C^T e, y \rangle \\ \text{subject to} & Cy \geq \bar{w} \\ & y \geq 0. \end{array}$$

We study the set of the primal-dual optimal solutions of the primal-dual **LPs** given above and show that the uniqueness of optimal solutions to $(\mathcal{DQ})$ depends on the degeneracy/nondegeneracy of $\operatorname{cone}(\mathcal{C})$.

Lemma 4.2.4. *Given the primal-dual pair $(\mathcal{Q})$ and $(\mathcal{DQ})$ above, we have the following:*

1. $\operatorname{Argmax}\left\{\langle -C^T e, y \rangle : Cy \geq \bar{w},\ y \geq 0\right\} = \{y \in \mathbb{R}^n : Cy = \bar{w},\ y \geq 0\}$,*i.e.*, $\{y \in \mathbb{R}^n : Cy = \bar{w},\ y \geq 0\}$ *is the set of optimal solutions to $(\mathcal{DQ})$.*

2. *$x = e$ is the unique optimal solution to $(\mathcal{Q})$.*

3. *If $\operatorname{cone}(\mathcal{C})$ is a nondegenerate cone, then $(\mathcal{DQ})$ has a unique optimal solution.*

4. *If $\operatorname{cone}(\mathcal{C})$ is a degenerate cone, then $(\mathcal{DQ})$ does not have a unique optimal solution (See Example 3.2.2.).*

Proof. 1. We use strong duality (See Proposition 2.3.5.) to find the optimal value of $(\mathcal{Q})$ and $(\mathcal{DQ})$. Since $x = e$ satisfies the constraints of $(\mathcal{Q})$, $(\mathcal{Q})$ is feasible. We also know that $\text{cone}(\mathcal{C})$ that $\bar{w} \in \text{int}(\text{cone}(\mathcal{C}))$ by Lemma 3.1.9. Hence, there exists a nonnegative vector y such that $Cy = \bar{w}$. Thus, $(\mathcal{DQ})$ is feasible and strong duality holds, i.e., the optimal values of $(\mathcal{Q})$ and $(\mathcal{DQ})$ coincide. For $y \geq 0$ such that $Cy = \bar{w}$, we observe that

$$\langle -\bar{w}, e\rangle = \langle -Cy, e\rangle = \langle -C^T e, y\rangle.$$

Hence, the optimal value of $(\mathcal{Q})$ and $(\mathcal{DQ})$ is $\langle -\bar{w}, e\rangle$ and

$$\{y \in \mathbb{R}^n : Cy = \bar{w},\ y \geq 0\} \subset \text{Argmax}\left\{\langle -C^T e, y\rangle : Cy \geq \bar{w},\ y \geq 0\right\}.$$

Suppose to the contrary that

$$\exists y' \in \text{Argmax}\left\{\langle -C^T e, y\rangle : Cy \geq \bar{w},\ y \geq 0\right\} \setminus \{y \in \mathbb{R}^n : Cy = \bar{w},\ y \geq 0\}.$$

Then, there exists $i \in \{1, \ldots, m\}$ such that $e_i^T Cy > e_i^T \bar{w}$, i.e., one of the inequalities of $Cy \geq \bar{w}$ must be strict. Thus we get

$$\langle Cy', e\rangle > \langle \bar{w}, e\rangle \implies \langle -\bar{w}, e\rangle > \langle -Cy', e\rangle = \langle -C^T e, y'\rangle.$$

Since y' does not achieve the optimal value, y' is not an optimal solution. Therefore, $\{y \in \mathbb{R}^n : Cy = \bar{w},\ y \geq 0\}$ is the set of optimal solutions to $(\mathcal{DQ})$.

2. By observing the proof in Item 1 of Lemma 4.2.4, it also follows that e is an optimal solution to $(\mathcal{Q})$. Here, we show that e is not only an optimal solution to $(\mathcal{Q})$, but also the unique optimal solution to $(\mathcal{Q})$.

   Suppose that x' is an optimal solution to $(\mathcal{Q})$. We know that there exists a dual optimal solution $y' > 0$ such that $Cy' = \bar{w}$, since $\bar{w} \in \text{int}(\text{cone}(\mathcal{C}))$ by Lemma 3.1.9. Hence, we have

$$0 \geq \langle C^T x' - C^T e, y'\rangle = \langle Cy', x'\rangle - \langle Cy', e\rangle = \langle \bar{w}, x'\rangle - \langle Cy', e\rangle = 0.$$

   The last inequality holds by the strong duality. Thus we get

$$\langle C^T x' - C^T e, y'\rangle = 0,\ y' > 0 \implies C^T x' = C^T e.$$

   Since $C \in \mathbb{R}^{n\times k}$ with $k \geq n$ and $\text{rank}(C) = n$, we have $\text{rank}(CC^T) = n$. Hence,

$$CC^T x' = CC^T e \implies x' = e.$$

   Therefore, $x' = e$ is the only optimal solution to $(\mathcal{Q})$. $\square$

Remark 4.2.5. *We note that previously in Section 3.2.1 we studied the relations between the two sets*

$$\{y \in \mathbb{R}^k : Cy = \bar{w},\ y \geq 0\},\ \{y \in \mathbb{R}^k : Dy = \bar{w},\ y \geq 0\}.$$

The above results hold for both sets, i.e., for both C, D.

Remark 4.2.6. *Given a feasible primal-dual **LP** pair, it is well-known that there is a pair of primal-dual optimal solutions that satisfies the strict complementarity. We note that there is a vector $y' > 0$ satisfying $Cy' = \bar{w}$ by Lemma 3.1.9. Hence, with $x' = e$, we see that (x', y') is a pair of optimal solutions of $(\mathcal{Q})$ and $(\mathcal{DQ})$ that holds the strict complementarity.*

4.3 Properties of the Optimal Value Functions

In this section, we study some basic properties of the optimal value functions of $(\mathcal{P})$ and $(\mathcal{Q})$ and its robust counterparts.

We define the optimal value function $\psi(\epsilon)$ of the robust counterpart $(\mathcal{R}(\epsilon))$ in (4.1.5):

$$(\mathcal{R}(\epsilon)) \qquad \psi(\epsilon) := \min\{\langle \bar{w}, x\rangle : (P - \epsilon E)x \geq Pe,\ x \geq 0\}. \tag{4.3.1}$$

We interchangeably use the notations $(\mathcal{P})$ and $(\mathcal{R}(0))$, since $(\mathcal{R}(0))$ and $(\mathcal{P})$ are the same **LP**. If we want to emphasize the matrix P in (4.3.1), we write

$$(\mathcal{R}(\epsilon, P)) \qquad \psi_P(\epsilon) := \min\{\langle \bar{w}, x\rangle : (P - \epsilon E)x \geq Pe,\ x \geq 0\}.$$

Similarly, we define the optimal value function $\phi(\epsilon)$ of the robust counterpart $(\mathcal{S}(\epsilon))$ in (4.1.6):

$$(\mathcal{S}(\epsilon)) \qquad \phi(\epsilon) := \min\left\{\langle -\bar{w}, x\rangle : (C^T + \epsilon E)x \leq C^T e,\ x \geq 0\right\}. \tag{4.3.2}$$

We note that $(\mathcal{S}(0))$ and $(\mathcal{Q})$ are the same **LP**. If we want to put an emphasis on the matrix C in (4.3.2), we write

$$(\mathcal{S}(\epsilon, C)) \qquad \phi_C(\epsilon) := \min\left\{\langle -\bar{w}, x\rangle : (C^T + \epsilon E)x \leq C^T e,\ x \geq 0\right\}.$$

We study some properties of the optimal value function $\psi(\epsilon)$. By making necessary changes to the arguments below, we can show that the same properties that hold for $\psi(\epsilon)$ also hold for $\phi(\epsilon)$.

1. For some $\bar{\epsilon} > 0$, we have

$$\psi(\epsilon) \text{ is a non-decreasing monotone function on } [0, \bar{\epsilon}], \tag{4.3.3}$$

by the definition of robust counterpart. We recall from the first line of (2.4.3) that

$$\begin{gathered} x \text{ is robust feasible to } \tilde{a}^T x \leq \tilde{\beta},\ \forall (\tilde{a}, \tilde{\beta}) \in \mathcal{U} \\ \Updownarrow \\ a^T x \leq \beta,\ \forall [a; \beta] \in \left\{[a^0; \beta^0] + \textstyle\sum_{\ell=1}^{L} \zeta_\ell \left[a^\ell; \beta^\ell\right],\ \zeta \in \mathcal{Z}\right\},\ \text{for some } \mathcal{Z}. \end{gathered}$$

In our case,

$$x \text{ is robust feasible to } \tilde{a}^T x \leq \beta,\ \forall \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon]$$
$$\Updownarrow$$
$$a^T x \leq \beta,\ \forall [a;\beta] \in \left\{ [a^0;\beta^0] + \textstyle\sum_{\ell=1}^{L} \zeta_\ell [e;0],\ \zeta \in \{u : \|u\|_\infty \leq \epsilon\} \right\}.$$

This implies that if the perturbation set gets bigger (i.e., ϵ increases), then the uncertainty set gets bigger. Hence, the robust optimal value cannot decrease as the feasible region of the robust **LP** gets smaller.

2. We observe that
$$\psi(0) < \psi(\epsilon),\ \forall \epsilon > 0. \tag{4.3.4}$$
We recall that e is the only optimal solution to $(\mathcal{P})$ (See Item 2 of Lemma 4.2.1.), and e does not satisfy the inequality $Px - \epsilon Ex \geq Pe$, $\forall \epsilon > 0$. We note by the above that the feasible region of $(\mathcal{R}(\epsilon))$ gets smaller as ϵ increases. Thus (4.3.4) holds, since the feasible region of $(\mathcal{R}(0))$ contains the feasible region of $(\mathcal{R}(\epsilon))$ and e is no longer feasible for $(\mathcal{R}(\epsilon)), \forall \epsilon > 0$.

3. Given two cones $\operatorname{cone}(\mathcal{C})$ and $\operatorname{cone}(\mathcal{D})$ constructed by the vectors from (3.1.2) and (3.1.12), respectively, we have
$$\{x \in \mathbb{R}^n : Px \geq 0\} = \operatorname{cone}(\mathcal{C}) \subset \operatorname{cone}(\mathcal{D}) = \{x \in \mathbb{R}^n : Qx \geq 0\}.$$
We observe the following:
$$\psi_P(0) = \psi_Q(0). \tag{4.3.5}$$
The equality (4.3.5) holds because Item 2 of Lemma 4.2.1 holds for any cone constructed by (3.1.3). In other words, we have constructed a family of **LPs** that have the same optimal value and the same optimal solution.

Item 1 above states that $\psi(\epsilon)$ is a nondecreasing function. Furthermore, if the optimal solution of $(\mathcal{P})$ is nondegenerate (i.e., exactly n constraints are active at the optimal solution), we can show that $\psi(\epsilon)$ is a *strictly increasing* function and we show this in Section 4.4.2. Given $(\mathcal{R}(\epsilon, P))$ and $(\mathcal{R}(\epsilon, Q))$, Item 2 above states that $\psi_P(0) < \psi_P(\epsilon)$ and $\psi_Q(0) < \psi_Q(\epsilon)$, $\forall \epsilon > 0$. Item 3 above states that $\psi_P(0) = \psi_Q(0)$. In other words, both $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ have the same function value at $\epsilon = 0$ and both function values increase for $\epsilon > 0$. In Section 4.4.2 (in nondegenerate cases) and Section 4.5.3 (in degenerate cases), we show that $\psi_P(\epsilon)$ always increases more than $\psi_Q(\epsilon)$ for small $\epsilon > 0$. That is, $\psi_P(\epsilon)$ is more sensitive than $\psi_Q(\epsilon)$ due to its local geometric structure.

4.4 The Strict Monotonicity and Sensitivity of the Optimal Value Function: Nondegenerate Cases

In this section we focus on ($\mathcal{P}$) where exactly n halfspaces are active at the optimal solution e. In this case, we say ($\mathcal{P}$) is *nondegenerate*.

4.4.1 Known Results on Parametric LP: Nondegenerate Cases

We first present some known results on a parametric **LP** under linear, scalar θ perturbations of the coefficient matrix. We define a family of **LPs** in **SEF**

$$\boldsymbol{LP}(\theta) \qquad z(\theta) := \min\left\{\langle c, x\rangle : A^\theta x = b, x \geq 0\right\}, \tag{4.4.1}$$

where $A^\theta = F + \theta G \in \mathbb{R}^{m\times n}$, for some $F, G \in \mathbb{R}^{m\times n}$ and $\theta \in \mathbb{R}$. For completeness, we include the proof of the following theorem on parametric **LP** in the nondegenerate case. In the following theorem, *nondegenerate basis* refers to a basis where the basic variables associated with the basis are nonzero.

Theorem 4.4.1 ([12, Theorem 1]). [1] *Let $P(\theta)$ be an instance of the parametric **LP** as given in (4.4.1), and let $\bar\theta \in \mathbb{R}$ be given. Suppose that $\mathcal{B}$ is a unique nondegenerate optimal basis for $P(\bar\theta)$, and $\bar x$ and $\bar\pi$ are unique optimal solutions to $P(\bar\theta)$ and its dual, respectively. Then for all θ near (in a neighbourhood of) $\bar\theta$, $\mathcal{B}$ is a nondegenerate optimal basis for $P(\theta)$, and the optimal value function $z(\theta)$ and the optimal solution $x(\theta)$ of $P(\theta)$ hold the following:*

(1) $z(\theta) = \sum_{i=0}^{\infty} c_{\mathcal{B}}^T(\theta-\bar\theta)^i(-B^{-1}G_{\mathcal{B}})^i\bar x_{\mathcal{B}}$,

(2) $z^k(\theta) = \sum_{i=k}^{\infty} \frac{i!}{(i-k)!} c_{\mathcal{B}}^T(\theta-\bar\theta)^{(i-k)}(-B^{-1}G_{\mathcal{B}})^i\bar x_{\mathcal{B}}$ *, for $k \in \mathbb{N}$, where $z^k(\theta)$ is the k-th derivative of $z(\theta)$,*

(3) $x(\theta) = (x_{\mathcal{B}}(\theta), x_N(\theta)) = \left(\sum_{i=0}^{\infty}(\theta-\bar\theta)^i(-B^{-1}G_{\mathcal{B}})^i\bar x_{\mathcal{B}}\ ,\ 0\right)$,

(4) $z^k(\bar\theta) = (k!)c_{\mathcal{B}}^T(-B^{-1}G_{\mathcal{B}})^k\bar x_{\mathcal{B}}$,

where $B = A_{\mathcal{B}}^{\bar\theta}$.

Proof. We show (3) $\implies$ (1) $\implies$ (2) $\implies$ (4). Let $B = A_{\mathcal{B}}^{\bar\theta}$ be the optimal basis matrix of $P(\bar\theta)$. We observe that

$$A_{\mathcal{B}}^\theta = F_{\mathcal{B}} + \theta G_{\mathcal{B}},\ \ B = A_{\mathcal{B}}^{\bar\theta} = F_{\mathcal{B}} + \bar\theta G_{\mathcal{B}}.$$

[1] The original statement in [12, Theorem 1] was written with a maximization **LP**.

Hence, we have

$$A_{\mathcal{B}}^{\theta} - B = (\theta - \bar{\theta})G_{\mathcal{B}}.$$

For all θ near $\bar{\theta}$, $(A_{\mathcal{B}}^{\theta})^{-1}$ exists and so

$$I = A_{\mathcal{B}}^{\theta}(A_{\mathcal{B}}^{\theta})^{-1} = (B + (\theta - \bar{\theta})G_{\mathcal{B}})(A_{\mathcal{B}}^{\theta})^{-1}. \tag{4.4.2}$$

Premultiplying B^{-1} to (4.4.2) yields

$$B^{-1} = (I + (\theta - \bar{\theta})B^{-1}G_{\mathcal{B}})(A_{\mathcal{B}}^{\theta})^{-1} \implies (A_{\mathcal{B}}^{\theta})^{-1} = B^{-1} - (\theta - \bar{\theta})(B^{-1}G_{\mathcal{B}})(A_{\mathcal{B}}^{\theta})^{-1}. \tag{4.4.3}$$

By recursively substituting for $(A_{\mathcal{B}}^{\theta})^{-1}$ in (4.4.3), we obtain

$$(A_{\mathcal{B}}^{\theta})^{-1} = \sum_{i=0}^{\infty}(\theta - \bar{\theta})^i(-B^{-1}G_{\mathcal{B}})^i B^{-1}.$$

(This series converges for all θ such that $|\theta - \bar{\theta}| < 1/\rho(-B^{-1}G_{\mathcal{B}})$. See Lemma 4.4.4.) Item (3) follows from $x_{\mathcal{B}}(\theta) = (A_{\mathcal{B}}^{\theta})^{-1}b$. Item (1) follows from $z(\theta) = c^T x(\theta) = c_{\mathcal{B}}^T x_{\mathcal{B}}(\theta)$. Taking derivatives from (1) gives (2). Plugging $\bar{\theta}$ into θ in (2) gives (4). □

We can obtain a simple formula by considering a special case of Item (4) in Theorem 4.4.1. We introduce the following lemma first.

Lemma 4.4.2. *Consider the* ***LP*** *in* ***SEF***

$$\textbf{\textit{LP}} \qquad \max\{\langle c, x\rangle : Ax = b, x \geq 0\}.$$

Suppose that x^* is a nondegenerate optimal basic feasible solution with basis $\mathcal{B}$. Then, a corresponding dual optimal solution is $y^* = (A_{\mathcal{B}}^T)^{-1}c_{\mathcal{B}}$. In addition, if there is a unique nondegenerate optimal solution, then there is a unique dual optimal solution.

Proof. Let x^* be a nondegenerate optimal feasible solution and y^* a dual optimal solution. Then complementary slackness implies

$$\left(c_j - A(:,j)^T y^*\right) x_j^* = 0,\ \forall j \in \mathcal{B} \implies A(:,j)^T y^* = c_j,\ \forall j \in \mathcal{B}.$$

Since $A_{\mathcal{B}}$ is nonsingular,

$$y^* = \left(A_{\mathcal{B}}^T\right)^{-1} c_{\mathcal{B}}. \tag{4.4.4}$$

The uniqueness is clear. □

With Lemma 4.4.2, we have the following corollary.

Corollary 4.4.3. *The derivative of the optimal value function of* ***LP***(θ) *at $\bar{\theta}$ exists and is*

$$z'\left(\bar{\theta}\right) = -\bar{\pi}^T G\bar{x}.$$

Proof. By pugging $k = 1$ into Item (4) in Lemma 4.4.1, we have

$$\begin{aligned} z'(\bar{\theta}) &= (1!)c_{\mathcal{B}}^T(-B^{-1}G_{\mathcal{B}})^1\bar{x}_{\mathcal{B}} \\ &= -c_{\mathcal{B}}^T B^{-1} G_{\mathcal{B}}\bar{x}_{\mathcal{B}} \\ &= -c_{\mathcal{B}}^T(A_{\mathcal{B}}^{\bar{\theta}})^{-1}G\bar{x} \\ &= -\bar{\pi}^T G\bar{x}, \end{aligned}$$

where dual solution is given by $\bar{\pi}^T = c_{\mathcal{B}}^T(A_{\mathcal{B}}^{\bar{\theta}})^{-1}$ in (4.4.4). □

4.4.2 The Monotonicity and Sensitivity in Nondegenerate Cases

In this section, we confine ourselves to the case where our data matrix P of

$$(\mathcal{P}) \qquad \min\{\langle \bar{w}, x\rangle : Px \geq Pe, x \geq 0\}$$

is a *square* matrix. In other words, there are exactly n active halfspaces at the optimal solution e. In this case, we recall that there are exactly n extreme rays in $\operatorname{cone}(\mathcal{C}) = \{x \in \mathbb{R}^n : Px \geq 0\}$ by Proposition 3.1.14. We also recall that $x^* = e$ is the unique optimal solution to $(\mathcal{P})$ by Item 2 of Lemma 4.2.1. The uniqueness of optimal solution does not change after transforming $(\mathcal{P})$ into **SEF**. Therefore,

$$\min\{\langle \bar{w}, x\rangle : Px - Is = Pe, x \geq 0, s \geq 0\} \tag{4.4.5}$$

has a unique *nondegenerate* optimal solution, which is given by $\begin{pmatrix} x^* \\ s^* \end{pmatrix} = \begin{pmatrix} e \\ 0 \end{pmatrix}$, with the optimal basis $\mathcal{B} = \{1, \ldots, n\}$. We showed in Lemma 4.4.2 that the dual of **LP** (4.4.5) has a unique optimal solution as well.

Our aim in this section is to show the following:

1. $\psi(\epsilon)$ is a strictly increasing function on $[0, \bar{\epsilon}]$, for some $\bar{\epsilon} > 0$.

   We showed in (4.3.3) that $\psi(\epsilon)$ is a non-decreasing function. When we have nondegeneracy, we can guarantee that $\psi(\epsilon)$ is a strictly increasing function.

2. Given two cones $\operatorname{cone}(\mathcal{C})$ and $\operatorname{cone}(\mathcal{D})$ constructed by the vectors in (3.1.2) and (3.1.12), respectively, we have

   $$\{x \in \mathbb{R}^n : Px \geq 0\} = \operatorname{cone}(\mathcal{C}) \subset \operatorname{cone}(\mathcal{D}) = \{x \in \mathbb{R}^n : Qx \geq 0\}.$$

   We define two **LPs**

   $$\psi_P(\epsilon) = \min\left\{\langle \bar{w}, x\rangle : (Pe - \epsilon E)x \geq Pe, x \geq 0\right\},$$

   $$\psi_Q(\epsilon) = \min\left\{\langle \bar{w}, x\rangle : (Qe - \epsilon E)x \geq Qe, x \geq 0\right\}.$$

Then, there is $\bar{\mu} > 0$ such that

$$\psi_P(\epsilon) > \psi_Q(\epsilon), \ \forall \epsilon \in (0, \bar{\mu}). \tag{4.4.6}$$

The strict inequality in (4.4.6) implies that

$$\psi_P(\epsilon) - \psi_P(0) > \psi_Q(\epsilon) - \psi_Q(0), \ \forall \epsilon \in (0, \bar{\mu}).$$

This means that when $(\mathcal{P})$ has a locally *smaller* feasible region, the difference between its nominal optimal value and robust optimal value is *larger* than when $(\mathcal{P})$ has a locally *bigger* feasible region. In other words, a locally *small* feasible region of an **LP** drives the problem more sensitive under data uncertainty. Throughout this thesis, we often omit the word 'locally' and state 'smaller feasible region' instead of 'locally smaller feasible region'. Similarly, we often state 'bigger feasible region' instead of 'locally bigger feasible region'.

The Strict Monotonicity of Robust Optimal Value Function $\psi(\epsilon)$

In this section, we show that $\psi(\epsilon)$ is a differentiable strictly increasing function, for all positive ϵ near $\bar{\epsilon} = 0$. That is,

$$\psi'(\epsilon) > 0, \ \text{for } \epsilon > 0 \text{ near } 0. \tag{4.4.7}$$

We recall from (4.3.3) that $\psi(\epsilon)$ is a non-decreasing function due the the definition of robust counterpart.

To show (4.4.7), we make use of Item (2) of Theorem 4.4.1:

$$z'(\theta) = \sum_{i=1}^{\infty} i c_{\mathcal{B}}^T (\theta - \bar{\theta})^{(i-1)} (-B^{-1} G_{\mathcal{B}})^i \bar{x}_{\mathcal{B}}. \tag{4.4.8}$$

We note that Theorem 4.4.1 is written with **LPs** in **SEF**. Hence, we first write $(\mathcal{R}(\epsilon))$ in **SEF** :

$$\begin{array}{rll} \psi(\epsilon) = & \min & \langle \bar{w}, x \rangle \\ & \text{subject to} & \begin{bmatrix} P - \epsilon E, & -I_n \end{bmatrix} \begin{pmatrix} x \\ s \end{pmatrix} = Pe \\ & & x, s \geq 0. \end{array} \tag{4.4.9}$$

We can write the constraint coefficient matrix of (4.4.9) in an explicit parametric form:

$$\begin{bmatrix} P - \epsilon E, & -I \end{bmatrix} = \begin{bmatrix} P & -I \end{bmatrix} + \epsilon \begin{bmatrix} -E & O \end{bmatrix}. \tag{4.4.10}$$

In our case, the data we should consider in (4.4.8) is the following:

$$\theta = \epsilon, \ \bar{\theta} = 0, \ B = P \in \mathbb{R}^{n \times n}, \ G_{\mathcal{B}} = -E \in \mathbb{R}^{n \times n} \text{ and } \bar{x}_{\mathcal{B}} = e.$$

The i-th term of the series in (4.4.8) becomes

$$ic_{\mathcal{B}}^T(\theta - \bar{\theta})^{(i-1)}(-B^{-1}G_{\mathcal{B}})^i\bar{x}_{\mathcal{B}} = i\epsilon^{i-1}c_{\mathcal{B}}^T B^{-1}(EB^{-1})^{i-1}Ee = in\epsilon^{i-1}y^T(EB^{-1})^{i-1}e,$$

where y is the dual optimal solution to $(\mathcal{P})$ by Lemma 4.4.2. Hence, we have

$$\psi'(\epsilon) = ny^T\left(\sum_{i=1}^{\infty} i\epsilon^{i-1}(EB^{-1})^{i-1}\right)e. \tag{4.4.11}$$

We first check the convergence for small ϵ of the *matrix* power series in (4.4.11):

$$\sum_{i=1}^{\infty} i\epsilon^{i-1}(EB^{-1})^{i-1}. \tag{4.4.12}$$

The radius of convergence of the *scalar* power series $\sum_{k=0}^{\infty} a_k z^k$ is

$$R = 1/(\limsup_{k\to\infty} \sqrt[k]{|a_k|}),$$

and is equal to $\lim_{k\to\infty}\left|\frac{a_k}{a_{k+1}}\right|$, if the limit exists. The spectral radius of a matrix $A \in \mathbb{C}^{n\times n}$ is defined as

$$\rho(A) := \max\{|\lambda_i| : \lambda_i \text{ is an eigenvalue of } A\}. \tag{4.4.13}$$

Lemma 4.4.4 ([19, Theorem 5.6.15])**.** *Let R be the radius of convergence of a scalar power series $\sum_{k=0}^{\infty} a_k z^k$, and let $A \in \mathbb{C}^{n\times n}$ be given. The matrix power series $\sum_{k=0}^{\infty} a_k A^k$ converges, if $\rho(A) < R$.* □

Now we show that $\psi'(\epsilon) > 0$, for small $\epsilon > 0$. Let $a_k = k\epsilon^{k-1}$. Then, we have

$$R = \lim_{k\to\infty}\left|\frac{a_k}{a_{k+1}}\right| = \lim_{k\to\infty}\left|\frac{k}{k+1}\frac{1}{\epsilon}\right| = \frac{1}{\epsilon}.$$

Hence, as long as we have $\rho(EB^{-1}) < 1/\epsilon$, the series in (4.4.12) converges. Since EB^{-1} is a rank one matrix, there is exactly one nonzero eigenvalue of EB^{-1}. We note that the nonzero eigenvalue of EB^{-1} is $e^TB^{-1}e$ and the corresponding eigenvector is e. Thus the spectral radius of EB^{-1} is

$$\rho(EB^{-1}) = |e^TB^{-1}e|.$$

Hence, Lemma 4.4.4 reads, if $|e^TB^{-1}e| < 1/\epsilon$, the series (4.4.12) converges. Thus we assume that ϵ is given small enough so that $|e^TB^{-1}e| < 1/\epsilon$ is satisfied.

Let $\lambda := e^TB^{-1}e$. Since $EB^{-1}e = \lambda e$, (4.4.11) becomes

$$\psi'(\epsilon) = ny^T\left(\sum_{i=1}^{\infty} i\epsilon^{i-1}(EB^{-1})^{i-1}e\right) = ny^T\left(\sum_{i=1}^{\infty} i(\epsilon\lambda)^{i-1}e\right). \tag{4.4.14}$$

Each coordinate of the vector $\sum_{i=1}^{\infty} i(\epsilon\lambda)^{i-1}e$ in (4.4.14) has the same value. Hence, in order to study the convergence of vector series (4.4.14), we may consider a series on $\mathbb{R}$, that is of the form $\sum_{k=0}^{\infty} c_k$, where $c_k = k(\epsilon\lambda)^{k-1}$.

We first recall the formula for a general geometric series and its derivative. We define a function $f(z) := \sum_{k=0}^{\infty} z^k$, with the domain $\{z \in \mathbb{R} : |z| < 1\}$. Then, the geometric series converges to $1/(1-z)$. $f(z)$ is well-defined and differentiable on $z \in (-1, 1)$. Hence, we have

$$f'(z) = \sum_{k=0}^{\infty} kz^{k-1} = \frac{1}{(1-z)^2} > 0. \tag{4.4.15}$$

Thus the series $\sum_{k=0}^{\infty} c_k$ with $c_k = k(\epsilon\lambda)^{k-1}$, with $|\epsilon\lambda| < 1$, converges to a positive number by (4.4.15). Therefore, if $\epsilon > 0$ satisfies $|e^T B^{-1} e| < 1/\epsilon$, then

$$\left(\sum_{i=1}^{\infty} i\epsilon^{i-1}(EB^{-1})^{i-1}\right) e > 0.$$

We recall that $ny^T \geq 0$ and $y \neq 0$. Thus we have $\psi'(\epsilon) > 0$ (See (4.4.11).), and (4.4.7) is verified.

The Strict Inequality $\psi_P'(0) > \psi_Q'(0)$

We recall that the constraint coefficient matrix of **LP** (4.4.9) in a parametric form is

$$\begin{bmatrix} P - \epsilon E, & -I \end{bmatrix} = \begin{bmatrix} P & -I \end{bmatrix} + \epsilon \begin{bmatrix} -E & O \end{bmatrix}.$$

Let y be the dual variable associated with the equality constraints at $\epsilon = 0$. Then, by Corollary 4.4.3, we have

$$\psi'(0) = -y^{*T} \begin{bmatrix} -E & O \end{bmatrix} \begin{pmatrix} x^* \\ s^* \end{pmatrix} = -y^{*T} \begin{bmatrix} -Ee & 0 \end{bmatrix} = y^{*T} Ee, \tag{4.4.16}$$

where (x^*, s^*) and y^* are the primal-dual optimal solutions. We note that $y^{*T}Ee$ does not involve the slack variable s^*. Hence, we may consider the optimal solution of $(\mathcal{R}(\epsilon))$ without transforming into **SEF** to study the sensitivity.

Suppose that we have two robust counterparts,

$$\psi_P(\epsilon) = \min\{\langle \bar{w}, x\rangle : (P - \epsilon E)x \geq Pe, x \geq 0\}, \tag{4.4.17}$$

$$\psi_Q(\epsilon) = \min\{\langle \bar{w}, x\rangle : (Q - \epsilon E)x \geq Qe, x \geq 0\}, \tag{4.4.18}$$

where $\text{cone}(\mathcal{C}) = \{x : Px \geq 0\}$, $\text{cone}(\mathcal{D}) = \{x : Qx \geq 0\}$ and $\text{cone}(\mathcal{C}) \subsetneq \text{cone}(\mathcal{D})$. Let x_P^* and x_Q^* be the optimal solutions to (4.4.17) and (4.4.18), respectively, when $\epsilon = 0$. We know that $x_P^* = x_Q^* = e$ by Item 2 of Lemma 4.2.1. Let y_P^* and y_Q^* be the optimal solutions to the duals of (4.4.17) and (4.4.18), when $\epsilon = 0$. We recall from Item 1 of Lemma 4.2.1

that the set of dual optimal solutions to $(\mathcal{R}(0, P))$ is

$$\left\{y_P \in \mathbb{R}^m : P^T y_P = \bar{w},\ y_p \geq 0\right\}.$$

Similarly, the set of dual optimal solutions of $(\mathcal{R}(0, Q))$ is

$$\left\{y_Q \in \mathbb{R}^m : Q^T y_Q = \bar{w},\ y_Q \geq 0\right\}.$$

We recall that, for all nonnegative y_P such that $P^T y_P = \bar{w}$, there exists $y_Q \geq 0$ such that $Q^T y_Q = \bar{w}$ and $y_P \geq y_Q$, by Lemma 3.2.10. Therefore, using Corollary 4.4.3, we obtain

$$\psi_P'(0) = -(y_P^*)^T(-E)x_P^* = (y_P^*)^T Ee = ne^T y_P^*,$$

$$\psi_Q'(0) = -(y_Q^*)^T(-E)x_Q^* = (y_Q^*)^T Ee = ne^T y_Q^*.$$

Since $y_P \geq y_Q$ and $y_P \neq y_Q$, we conclude that

$$\psi_P'(0) > \psi_Q'(0). \tag{4.4.19}$$

That is, when the feasible region is smaller, its robust optimal value is more sensitive at $\epsilon = 0$. We state the result (4.4.19) in the proposition below. The above argument gives the proof for the following proposition.

Proposition 4.4.5. *For the two robust optimal value functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ with nondegenerate optimal solutions, we have*

$$\psi_P'(0) > \psi_Q'(0).$$ □

The Strict inequality $\psi_P(\epsilon) > \psi_Q(\epsilon)$

Now we want to show that

$$\psi_P(\epsilon) > \psi_Q(\epsilon),\ \forall \epsilon \in (0, \bar{\epsilon}),\ \text{for some } \bar{\epsilon} > 0.$$

We summarize the results we have so far:

(1) $\psi_P(0) = \psi_Q(0)$ (See (4.3.5) on page 64.),

(2) $\psi_P'(0) > \psi_Q'(0)$ (See Proposition 4.4.5.).

We observe the following: By Item (1) and Item (2) above, we have

$$0 < \lim_{h \to 0} \frac{\psi_P(h) - \psi_p(0) - \psi_Q(h) + \psi_Q(0)}{h} = \lim_{h \to 0} \frac{\psi_P(h) - \psi_Q(h)}{h}.$$

Therefore, we conclude the following:

$$\boxed{\text{There is an interval } I = (0, \bar{\epsilon}) \text{ such that } \psi_P(\epsilon) > \psi_Q(\epsilon),\ \forall \epsilon \in I.} \tag{4.4.20}$$

This implies that

$$\psi_P(\epsilon) - \psi_P(0) > \psi_Q(\epsilon) - \psi_Q(0),\ \forall \epsilon \in I.$$

That is, there is an interval such that the robust optimal value of a *smaller* feasible region is always worse than the robust optimal value of a *bigger* feasible region.

4.4.3 The Robust Optimal Value Function $\phi(\epsilon)$

In this section, we make a brief note on the robust optimal value function $\phi(\epsilon)$ of the robust counterpart $(\mathcal{S}(\epsilon))$. A reader who wishes to recall the definition of $\phi(\epsilon)$ and $(\mathcal{S}(\epsilon))$ may refer to Section 4.2.2 and Section 4.3.

In the previous sections, we have studied the optimal value function $\psi(\epsilon)$ of $(\mathcal{R}(\epsilon))$ under the nondegenerate case. We note that the argument on the optimal value function $\phi(\epsilon)$ of $(\mathcal{S}(\epsilon))$ parallels the argument in the previous sections owing to Lemma 3.2.1 (Note that Lemma 3.2.10 played an important role on getting the sensitivity result in Section 4.4.2.). Thus, by making necessary changes in the previous section, we have the following result:

1. $\phi(\epsilon)$ is a strictly increasing function for all ϵ near 0,
2. There is an interval $I = (0, \bar{\epsilon})$ such that $\phi_C(\epsilon) < \phi_D(\epsilon)$, $\forall \epsilon \in I$.

We recall that $\text{cone}(\mathcal{C})^\circ \supset \text{cone}(\mathcal{D})^\circ$ since polarization is order-reversing (See Lemma 2.1.25.). Hence, the above argument also shows that there is an interval such that the robust optimal value of a *smaller* feasible region is always worse than the robust optimal value of a *bigger* feasible region.

4.5 The Sensitivity of the Optimal Value Function: Degenerate Cases

In this section, we focus on $(\mathcal{P})$, where more than n halfspaces are active at the optimal solution e. When an **LP** in **SEF** has a degenerate basic optimal solution, we obtain more than one optimal basis. Hence, the approach used in Theorem 4.4.1 is no longer valid.

A reader may question that we can use the approach introduced by Fiacco [11] as $(\mathcal{R}(\epsilon))$ is in a nice parametric form. Fiacco in [11] contains a comprehensive sensitivity analysis on parametric nonlinear programming (hence applicable to **LPs**) and nice formulae for the derivatives of optimal solutions and optimal values with respect to the parameters. However, most of the arguments are presented under the linear independence constraint qualification. This triggers problems when the number of active constraints of an **LP** at its optimal solution is bigger than its dimension. In this case, the Jacobian matrix contains the gradients of the active linear constraints and hence it is not invertible. Thus, instead of using the approach from [11], we perform our analysis using *directional differentiability*. A reader interested in non-linear parametric programming may refer to [3, 11].

The rest of this section is organized as follows: In Section 4.5.1, we present *why* the optimal value function $\psi(\epsilon)$ might not be differentiable at $\epsilon = 0$ in the presence of degeneracy. In Section 4.5.2, we present that the optimal value function $\psi(\epsilon)$ is directionally differentiable. In Section 4.5.3, with directional differentiability obtained in Secion 4.5.2, we conclude the same result stated in (4.4.20) in degenerate cases. In Section 4.5.4, we suggest a sufficient condition for differentiability of $\psi(\epsilon)$ at $\epsilon = 0$. Finally, in Section 4.5.5, we show that the properties hold for the optimal value function $\psi(\epsilon)$ for **LP** $(\mathcal{P})$ also hold for the optimal value function $\phi(\epsilon)$ for **LP** $(\mathcal{Q})$.

4.5.1 The Generalized Directional Derivative

We recall that the optimal value function $\psi(\epsilon)$ is differentiable by Theorem 4.4.1, given that the **LP** is nondegenerate. In this section, we show *where* the non-differentiability of $\psi(\epsilon)$ comes from in degenerate cases.

Known Results on Parametric LP: Degenerate Cases

Theorem 4.5.1 ([32, Theorem 1]). *Given $c \in \mathbb{R}^n, b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$, consider the following **LP***

$$\tilde{z}(A) = \max\{\langle c, x\rangle : Ax = b, x \geq 0\}.$$

Let $\operatorname{dom}(\tilde{z})$, i.e., the domain of $\tilde{z}$ be $\{A \in \mathbb{R}^{m\times n} : \tilde{z}(A)$ is finite$\}$. If

(1) A is an interior point of $\operatorname{dom}(\tilde{z})$, and

(2) $\tilde{z}(A)$ is locally Lipschitz in a neighbourhood of A,

then

$$\partial\tilde{z}(A) = \operatorname{conv}\{-ux^T : \begin{array}{l} u \text{ is any optimal dual solution,} \\ x \text{ is any primal optimal solution}\}. \end{array} \qquad \square$$

Remark 4.5.2. *We note that Theorem 4.5.1 is also applicable to nondegenerate cases. In the nondegenerate case, optimal solutions x and u of a primal-dual pair are unique. Then, each entry of $\partial\tilde{z}(A)$ is a real number, rather than an interval (under the assumption that the hypotheses of Theorem 4.5.1 were satisfied.).*

The Generalized Directional Derivative of the Optimal Value Function $\hat{z}$

In this section, we want to show where the non-differentiability of the optimal value function might be coming from, by utilizing Theorem 4.5.1. We also study the generalized directional derivative of the optimal value function of $(\mathcal{P})$ in **SEF**. Hence, we check that the hypotheses of Theorem 4.5.1 are satisfied with our **LP** :

1. We first introduce Proposition 4.5.3 to relate the first hypothesis (1) of Theorem 4.5.1. Proposition 4.5.3 states that for a primal-dual pair of **LPs** satisfying the regularity condition, any small perturbations around given data do not render the pair unsolvable.

2. We mention [8, Theorem 4.3] to relate the second hypothesis (2) of Theorem 4.5.1. [8, Theorem 4.3] states the Lipschitz property of optimal value function under small perturbations of given data.

We note that the **LP** in Theorem 4.5.1 is written in **SEF**. Hence, we need to transform $(\mathcal{P})$ in **SEF** :

$$(\mathcal{P}_{\text{SEF}}) \qquad \begin{array}{rll} \hat{z}([P, -I_m]) = & \min & \langle \bar{w}, x \rangle \\ & \text{subject to} & \begin{bmatrix} P & -I_m \end{bmatrix} \begin{pmatrix} x \\ s \end{pmatrix} = Pe \\ & & x, s \geq 0. \end{array}$$

We explain the notation used in Proposition 4.5.3. Given $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ and a matrix $A \in \mathbb{R}^{m \times n}$, we define

$$A(X) - Y := \{Ax - y : x \in X,\ y \in Y\}.$$

Proposition 4.5.3 ([24, Theorem 1])**.** *Let K_1, K_2 be non-empty polyhedral convex cones in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let $A \in \mathbb{R}^{m \times n}$. Let $(\overline{P}), (\overline{D})$ be the following **LPs** :*

$$(\overline{P}) \quad \begin{array}{rl} \min & \langle c, x \rangle \\ \text{subject to} & Ax - b \in K_2^* \\ & x \in K_1, \end{array} \qquad\qquad (\overline{D}) \quad \begin{array}{rl} \max & \langle u, b \rangle \\ \text{subject to} & c - A^T u \in K_1^* \\ & u \in K_2. \end{array}$$

Then, the following are equivalent.

1. *The constraints of $(\overline{P})$ and of $(\overline{D})$ are regular, i.e.,*

$$b \in \operatorname{int}\{A(K_1) - K_2^*\} \ \text{ and } \ c \in \operatorname{int}\left\{A^T(K_2) + K_1^*\right\}.$$

2. *The sets of optimal solutions of $(\overline{P})$ and of $(\overline{D})$ are nonempty and bounded.*

3. *There exists $\epsilon_0 > 0$ such that for any A', b' and c' with*

$$\epsilon' \equiv \max\{\|A' - A\|, \|b' - b\|, \|c' - c\|\} < \epsilon_0,$$

the two dual problems

$$(P') \quad \begin{array}{rl} \min & \langle c', x \rangle \\ \text{subject to} & A'x - b' \in K_2^* \\ & x \in K_1, \end{array} \qquad\qquad (D') \quad \begin{array}{rl} \max & \langle u, b' \rangle \\ \text{subject to} & c' - (A')^T u \in K_1^* \\ & u \in K_2 \end{array}$$

are solvable. $\square$

The following lemma shows that regularity of the system $\{x \in \mathbb{R}^n : Ax \geq b, x \in \mathbb{R}^n_+\}$ is equivalent to Slater condition.

Lemma 4.5.4. *Let $\{x \in \mathbb{R}^n : Ax \geq b, x \in \mathbb{R}^n_+\}$ be a given system of inequalities with $A \in \mathbb{R}^{m\times n}$. Assume that there exists $\hat{x}$ such that $A\hat{x} > b, \hat{x} \geq 0$. Then, the following are equivalent.*

(1) $b \in \operatorname{int}\left\{A(\mathbb{R}^n_+) - \mathbb{R}^m_+\right\}$,

(2) there exists x' such that $Ax' > b, x' \in \mathbb{R}^n_+$.

Proof. Suppose that Item (1) holds. Then, $\forall d$ with small enough $\epsilon > 0$, we have $b + \epsilon d = A\bar{x} - \bar{s}$, for some $\bar{x} \in \mathbb{R}^n_+$, $\bar{s} \in \mathbb{R}^m_+$. Taking d such that $-\bar{s} < \epsilon d$ yields $b = A\bar{x} - \bar{s} - \epsilon d < A\bar{x}$. Hence, Item (2) holds.

Conversely, suppose that Item (2) holds. Then $Ax' - s = b$, for some $s \in \mathbb{R}^m_{++}$. Let $d \in \mathbb{R}^m$ with $\|d\| = 1$ and let $\delta = \min_i\{s_i\}$. Define $\bar{s} := s - \delta d$. Then we get

$$b + \delta d = Ax' - s + \delta d = Ax' - \bar{s}.$$

We note that $\bar{s} \geq 0$, since $\delta d_i \leq \delta \leq s_i$, $\forall i$. Therefore, $\exists \delta > 0$ such that $\forall d \in \mathbb{R}^m$ with $\|d\| = 1$, we have

$$b + \delta d = Ax' - \bar{s},\ x \geq 0,\ \bar{s} \geq 0.$$

Hence, $b \in \operatorname{int}\left\{A(\mathbb{R}^n_+) - \mathbb{R}^m_+\right\}$, so Item (1) holds. $\square$

Remark 4.5.5 states that a regular system (defined in Item 1 of Proposition 4.5.3) of inequalities $\{x \in \mathbb{R}^n : Ax \geq b, x \in \mathbb{R}^n_+\}$ remains regular after being transformed into constraints of **SEF**.

Remark 4.5.5. *Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, we define two systems*

$$\begin{aligned} S_1 &:= \{x \in \mathbb{R}^n : Ax \geq b, x \in \mathbb{R}^n_+\}, \text{ and} \\ S_2 &:= \left\{\begin{pmatrix} x \\ s \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^m : Ax - s = b, x \in \mathbb{R}^n_+, s \in \mathbb{R}^m_+\right\}. \end{aligned}$$

System S_1 is regular if, and only if, system S_2 is regular.

Proof. If system S_2 is regular, we have

$$b \in \operatorname{int}\left\{\begin{bmatrix} A, & -I \end{bmatrix}\begin{bmatrix} \mathbb{R}^n_+ \\ \mathbb{R}^m_+ \end{bmatrix} - (\mathbb{R}^m)^*\right\} = \operatorname{int}\left\{A(\mathbb{R}^n_+) - \mathbb{R}^m_+\right\}. \tag{4.5.1}$$

We note that (4.5.1) is equivalent to Item (1) stated in Lemma 4.5.4. $\square$

We note that our systems $\{x \in \mathbb{R}^n : Px \geq Pe, x \geq 0\}$ and $\{y \in \mathbb{R}^m : P^T y \leq \bar{w}, y \geq 0\}$ have points $\hat{x}, \hat{y}$ such that $P\hat{x} > Pe, \hat{x} \geq 0$ and $P^T\hat{y} < \bar{w}, \hat{y} \geq 0$ by Lemma 3.2.5 and the construction of cone($\mathcal{C}$). Hence, by Lemma 4.5.4,

$$Pe \in \operatorname{int}\left\{P(\mathbb{R}^n_+) - \mathbb{R}^m_+\right\} \text{ and } \bar{w} \in \operatorname{int}\left\{P^T(\mathbb{R}^m_+) + \mathbb{R}^n_+\right\}$$

are satisfied. By Remark 4.5.5, we have that the constraint system of ($\mathcal{P}_{\text{SEF}}$) and its dual are regular. Thus, by Proposition 4.5.3, $\exists \epsilon_0 > 0$ such that $\forall [P', J'] \in \text{Ball}([P, -I_m], \epsilon_0)$, the primal ($\mathcal{P}_{\text{SEF}}$) and its dual are solvable (Note that we do not perturb the RHS nor the objective.). Therefore we have

$$[P', J'] \in \operatorname{dom}(\hat{z}), \ \forall [P', J'] \in \text{Ball}([P, -I_m], \epsilon_0),$$

and the hypothesis (1) of Theorem 4.5.1 holds with ($\mathcal{P}$).

Now we check that the hypothesis (2) of Theorem 4.5.1 holds with ($\mathcal{P}$). It is shown in [8, Theorem 4.3][2] that given an **LP** satisfying the regularity condition, the optimal value function is locally Lipschitz near the given data. Thus the hypotheses of Theorem 4.5.1 are satisfied and we are ready to apply Theorem 4.5.1 to ($\mathcal{R}(0)$).

Applying Theorem 4.5.1 to ($\mathcal{R}(0)$) in **SEF** gives

$$\partial\hat{z}([P, -I_m]) = \operatorname{conv}\left\{-y\left(x^T, s^T\right) : (x; s), y \text{ optimal solutions to primal-dual pair}\right\}.$$

Since ($\mathcal{P}$) has the unique optimal solution $x^* = e$, the optimal solution to ($\mathcal{P}_{\text{SEF}}$) is $(x^*; s^*) = (e; 0)$. Then, $\partial\hat{z}([P, -I_m])$ reduces to

$$\partial\hat{z}([P, -I_m]) = \operatorname{conv}\left\{-[ye^T, O_{m\times m}] : y \text{ optimal solution to } (\mathcal{DP})\right\}.$$

We recall that there are more than one optimal solutions to dual ($\mathcal{DP}$) by Item 4 of Lemma 4.2.1. That is, $\partial\hat{z}([P, -I_m])$ is *not a singleton*. We also recall that if $\partial f(x)$ is not a singleton, then f is not continuously differentiable, by the contrapositive of Lemma 2.2.8. Hence, $\hat{z}$ is not differentiable in regards to matrix perturbations, in general. However, with the direction of perturbations we are interested, we may obtain some partial knowledge on the sensitivity of the optimal value function $\hat{z}$ of ($\mathcal{P}_{\text{SEF}}$).

We note that the constraint matrix of ($\mathcal{R}(\epsilon)$) in **SEF** is given by

$$\begin{bmatrix} P - \epsilon E, & -I \end{bmatrix} = \begin{bmatrix} P & -I \end{bmatrix} + \epsilon \begin{bmatrix} -E & O \end{bmatrix} \in \mathbb{R}^{m\times(n+m)}.$$

That is, we want to perturb the matrix $[P, -I_m]$ in the direction $[-E, O] \in \mathbb{R}^{m\times(n+m)}$.

We recall from Proposition 2.2.7 that the generalized gradient of a function f at x in direction v is given by

$$f^\circ(x; v) = \max\left\{\langle \xi, v\rangle : \xi \in \partial f(x)\right\}.$$

[2] [8, Theorem 4.3] was stated with linear semi-infinite constraints but the statement holds with ordinary **LPs** as well.

Hence, the generalized gradient of a function $\hat{z}$ at $[P, -I_m]$ in the direction $[-E, O]$ is given by

$$\begin{aligned}
\hat{z}^\circ([P, -I_m]; [-E, O]) &= \max\left\{\left\langle -[ye^T, O_{m\times m}], [-E, O]\right\rangle : y \text{ optimal solution to } (\mathcal{DP})\right\} \\
&= \max\left\{y^T E e^n : y \text{ optimal solution to } (\mathcal{DP})\right\} \\
&= \max\left\{n\langle y, e^m\rangle : y \text{ optimal solution to } (\mathcal{DP})\right\}.
\end{aligned} \tag{4.5.2}$$

The first equality in (4.5.2) holds by Theorem 4.5.1 and the uniqueness of the primal optimal solution. By observing (4.5.2), we see that the sensitivity of the optimal value function $\hat{z}$ highly depends on the optimal solutions of dual $(\mathcal{DP})$. In Section 4.5.2, we show that the directional derivative of $\psi(\epsilon)$ at $\epsilon = 0$ in the direction $1 \in \mathbb{R}$ (positive direction) coincides with (4.5.2).

Remark 4.5.6. *In the nondegenerate case, we may use Theorem 4.5.1 and get the total derivative of $\hat{z}$. We note that the total derivative of $\hat{z}(P')$, where $P' = [P, -I_n]$, can be written as*

$$d\hat{z}(P') = \sum_{i,j} \frac{\partial \hat{z}(P')}{\partial p'_{ij}} dp'_{ij}.$$

Since

$$dp'_{ij} = \begin{cases} -1, & \forall i \in \{1, \dots, n\},\ j \in \{1, \dots, n\} \\ 0, & \forall i \in \{1, \dots, n\},\ j \in \{n+1, \dots 2n\} \end{cases},$$

the result coincides with the nondegenerate case.

4.5.2 The Directional Differentiability of $\psi(\epsilon)$

We have shown, in Section 4.5.1, that the optimal value function $\hat{z}$ might not be differentiable at $[P, -I_m]$. That is, the optimal value function $\psi(\epsilon)$ might not be differentiable at $\epsilon = 0$. However, we note that we are interested in the domain of $\psi(\epsilon)$ that is positive real number, i.e., $\epsilon \in (0, \bar{\epsilon})$, for some $\bar{\epsilon} > 0$. Hence, we turn our attention to the directional differentiability of $\psi(\epsilon)$, rather than the differentiability of $\psi(\epsilon)$ at $\epsilon = 0$. In this section, we show that ψ is directionally differentiable in direction of 1, that is,

$$\lim_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \quad \text{exsits.} \tag{4.5.3}$$

Fiacco [11, Section 2.3] and Still [29, Chapter 7] contain the related result presented in this section. We want to show (4.5.3) by getting an upper bound (lim sup) and a lower bound (lim inf), and claim that the upper bound and the lower bound are equal.

We recall that the robust counterpart of $(\mathcal{P})$ is given by

$$(\mathcal{R}(\epsilon)) \qquad \psi(\epsilon) = \min\{\langle \bar{w}, x\rangle : (P - \epsilon E)x \geq Pe, x \geq 0\}.$$

We may write $(\mathcal{R}(\epsilon))$ as

$$(\mathcal{R}(\epsilon)) \qquad \begin{array}{rcl} \psi(\epsilon) = & \min & \langle \bar{w}, x\rangle \\ & \text{subject to} & g_i(x,\epsilon) \le 0, \quad \forall i = 1,\dots,m \\ & & g_i(x,\epsilon) \le 0, \quad \forall i = m+1,\dots,m+n, \end{array}$$

where

$$\begin{array}{lll} g_i(x,\epsilon) & := -(p^i)^T x + \epsilon e^T x + (p^i)^T e, & \forall i = 1,\dots m, \\ g_i(x,\epsilon) & := -x_i, & \forall i = m+1,\dots,m+n, \end{array}$$

(We may omit ϵ in the input of $g_i(x,\epsilon)$, $\forall i = m+1,\dots,m+n$.). Let

$$S(\epsilon) := \operatorname{Argmin}\left\{\langle \bar{w}, x\rangle : (P - \epsilon E)x \ge Pe, x \ge 0\right\},$$

i.e., the set of optimal solutions of $(\mathcal{R}(\epsilon))$. We recall that $S(0) = \{e\}$ by Item 2 of Lemma 4.2.1. Let $I(x)$ be the set of indices such that $g_i(x,\epsilon)$ is active at x. We note that at the optimal solution x^* of $(\mathcal{R}(0))$, the set of active indices is $I(x^*) = \{1,\dots,m\}$. Then the Lagrangian of $(\mathcal{R}(\epsilon))$ *near* $(x^*,\epsilon) = (e,0)$ is given by

$$L(x,\epsilon,y) = \langle \bar{w}, x\rangle + \sum_{i=1}^{m} y_i g_i(x,\epsilon). \tag{4.5.4}$$

Theorem 4.5.7 gives an upper bound for (4.5.3).

Theorem 4.5.7. *Given a family of $(\mathcal{R}(\epsilon))$, for small $\epsilon > 0$, let $\psi(\epsilon)$ be the optimal value function of $(\mathcal{R}(\epsilon))$. Let M be the set of optimal solutions of $(\mathcal{DP})$. Then, we have*

$$\limsup_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \le \max\left\{n\langle e, y\rangle : y \in M\right\}.$$

Proof. Consider the **LP**

$$(P^+) \qquad \begin{array}{rcl} v^+ = & \displaystyle\min_{\xi} & \langle \bar{w}, \xi\rangle \\ & \text{subject to} & (p^i)^T \xi \ge n, \forall i = 1,\dots,m. \end{array}$$

Then, the dual (D^+) of (P^+) is

$$(D^+) \qquad \begin{array}{rl} \displaystyle\max_{y} & n\langle e, y\rangle \\ \text{subject to} & P^T y = \bar{w} \\ & y \ge 0. \end{array}$$

We note that the set of feasible solutions to (D^+) is M, i.e., the set of optimal solutions to $(\mathcal{DP})$ (See Item 1 of Lemma 4.2.1.). Since cone$(\mathcal{C}) = \{x \in \mathbb{R}^n : Px \ge 0\}$ is a full-dimensional cone, there exists a vector $\bar{\xi}$ such that $P\bar{\xi} > 0$. Hence, we know that there exist feasible solutions to (P^+) and (D^+). Thus, by strong duality (See Proposition 2.3.5.), we have

$$v^+ = \max\{n\langle e, y\rangle : y \in M\}.$$

Let ξ^* be a minimizer of (P^+). Let $\bar{x}$ be the optimal solution to $(\mathcal{R}(0))$. Define, for small $\delta > 0$ and $\epsilon > 0$,

$$x_\epsilon := \bar{x} + \epsilon(\xi^* + \delta\bar{\xi}).$$

Then, for $i \in I(\bar{x}) = \{1, \dots, m\}$ and for small $\epsilon > 0$, we have

$$\begin{aligned} g_i(x_\epsilon, \epsilon) &= g_i(x_\epsilon, \epsilon) - g_i(\bar{x}, 0) \\ &= \left\langle -p^i, \bar{x} + \epsilon\left(\xi^* + \delta\bar{\xi}\right)\right\rangle + \epsilon\left\langle e, \bar{x} + \epsilon\left(\xi^* + \delta\bar{\xi}\right)\right\rangle + (p^i)^T e + (p^i)^T\bar{x} - (p^i)^T e \\ &= \epsilon\langle -p^i, \xi^* + \delta\bar{\xi}\rangle + \epsilon\langle e, \bar{x}\rangle + \epsilon^2\langle e, \xi^* + \delta\bar{\xi}\rangle \\ &= \epsilon\left[\langle -p^i, \xi^* + \delta\bar{\xi}\rangle + \langle e, \bar{x}\rangle\right] + o(\epsilon). \end{aligned} \tag{4.5.5}$$

The first equality holds since $g_i(\bar{x}, 0) = 0, \forall i \in I(\bar{x})$. Then, by dividing both sides of (4.5.5) by ϵ, we get

$$\begin{aligned} g_i(x_\epsilon, \epsilon)/\epsilon &= \langle -p^i, \xi^* + \delta\bar{\xi}\rangle + \langle e, \bar{x}\rangle + o(\epsilon)/\epsilon \\ &= \langle -p^i, \xi^*\rangle + \langle e, \bar{x}\rangle + \langle -p^i, \delta\bar{\xi}\rangle + o(\epsilon)/\epsilon \\ &= \langle -p^i, \xi^*\rangle + n + \delta\langle -p^i, \bar{\xi}\rangle + o(\epsilon)/\epsilon \\ &\le \delta\langle -p^i, \bar{\xi}\rangle + o(\epsilon)/\epsilon \qquad \text{by the feasibility of } \xi^* \text{ to } (P^+) \\ &< 0 \qquad \text{since } P\bar{\xi} > 0. \end{aligned}$$

Hence, x_ϵ is a feasible solution to $(\mathcal{R}(\epsilon))$, and this implies that $\psi(\epsilon) \le \langle \bar{w}, x_\epsilon\rangle$. Thus we have

$$\begin{aligned} \psi(\epsilon) - \psi(0) &\le \langle \bar{w}, x_\epsilon\rangle - \langle \bar{w}, \bar{x}\rangle \qquad \text{since } \psi(0) = \langle \bar{w}, \bar{x}\rangle \\ &= \langle \bar{w}, \bar{x} + \epsilon(\xi^* + \delta\bar{\xi})\rangle - \langle \bar{w}, \bar{x}\rangle \\ &= \langle \bar{w}, \epsilon(\xi^* + \delta\bar{\xi})\rangle \\ &= \epsilon\langle \bar{w}, \xi^*\rangle + \epsilon\delta\langle \bar{w}, \bar{\xi}\rangle. \end{aligned}$$

Hence,

$$\begin{aligned} \tfrac{\psi(\epsilon) - \psi(0)}{\epsilon} &\le \langle \bar{w}, \xi^*\rangle + \delta\langle \bar{w}, \bar{\xi}\rangle \\ &= v^+ + \delta\langle \bar{w}, \bar{\xi}\rangle. \end{aligned}$$

Since $\delta > 0$ can be chosen arbitrarily small, we have

$$\limsup_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \le \max\{n\langle e, \mu\rangle : \mu \in M\}. \qquad \square$$

We observe in the proof of Theorem 4.5.7 that getting an upper bound of (4.5.3) does not require any argument on convexity (other than strong duality of (P^+) and (D^+).). However, obtaining a lower bound of (4.5.3) requires the convexity status.

Theorem 4.5.8. *Given a family of $(\mathcal{R}(\epsilon))$, for small $\epsilon > 0$, let $\psi(\epsilon)$ be the optimal value* *function of* $(\mathcal{R}(\epsilon))$. *Let* M *be the set of optimal solutions of* $(\mathcal{DP})$. *Then, we have*

$$\liminf_{\epsilon\downarrow 0} \frac{\psi(\epsilon)-\psi(0)}{\epsilon} \geq \max\left\{n\langle e,\mu\rangle : \mu \in M\right\}.$$

Proof. Let x_ℓ be an optimal solution to $(\mathcal{R}(\epsilon_\ell))$ for small $\epsilon_\ell > 0$. Such optimal solutions exist, since the constraints systems of $(\mathcal{P})$ and $(\mathcal{DP})$ are regular, and for any small perturbations $(\mathcal{R}(\epsilon))$ is solvable by Proposition 4.5.3. Let $\{x_\ell\} \subset \mathbb{R}^n$ and $\{\epsilon_\ell\} \subset \mathbb{R}$ be sequences such that $x_\ell \in S(\epsilon_\ell)$, converging to $\bar{x} \in S(0)$ and $\epsilon_\ell \downarrow 0$.

We recall the Lagrangian near $(e, 0)$ defined in (4.5.4). Since $\bar{x}$ is the optimal solution to $(\mathcal{R}(0))$, we have $\nabla_x L(\bar{x}, 0, y) = 0,\ \forall y \in M$. Since $L(\bar{x}, 0, y)$ is linear (convex) with respect to x, $\nabla_x L(\bar{x}, 0, y) = 0$ implies that $\bar{x}$ is a global minimizer of $L(x, 0, y)$, for each $y \in M$. Hence, we have

$$\begin{aligned} L(x_\ell, \epsilon_\ell, y) - L(x_\ell, 0, y) &\leq L(x_\ell, \epsilon_\ell, y) - L(\bar{x}, 0, y) \\ &= \langle w, x_\ell\rangle - \langle w, \bar{x}\rangle + \textstyle\sum_{i\in I(\bar{x})} y_i \left(g_i(x_\ell, \epsilon_\ell) - g_i(\bar{x}, 0)\right) \\ &\leq \langle w, x_\ell\rangle - \langle w, \bar{x}\rangle \\ &= \psi(\epsilon_\ell) - \psi(0). \end{aligned} \tag{4.5.6}$$

The first inequality in (4.5.6) holds since $\bar{x}$ is a global minimizer. The second inequality in (4.5.6) holds since $g_i(x_\ell, \epsilon_\ell) \leq 0$ and $g_i(\bar{x}, 0) = 0,\ \forall i \in I(\bar{x})$. Thus we get

$$\begin{aligned} \psi(\epsilon_\ell) - \psi(0) &\geq L(x_\ell, \epsilon_\ell, y) - L(x_\ell, 0, y) \\ &= \sum_{i\in I(\bar{x})} y_i \left(-(p^i)^T x_\ell + \epsilon_\ell e^T x_\ell + (p^i)^T e\right) - \sum_{i\in I(\bar{x})} y_i \left(-(p^i)^T x_\ell + (p^i)^T e\right) \\ &= \sum_{i\in I(\bar{x})} y_i \epsilon_\ell e^T x_\ell \\ &= \epsilon_\ell e^T x_\ell \sum_{i\in I(\bar{x})} y_i. \end{aligned}$$

Dividing the above by $\epsilon_\ell > 0$ gives

$$\frac{\psi(\epsilon_\ell) - \psi(0)}{\epsilon_\ell} \geq e^T x_\ell \sum_{i\in I(\bar{x})} y_i. \tag{4.5.7}$$

Letting $l \to \infty$ in (4.5.7) yields

$$\lim_{l\to\infty} \frac{\psi(\epsilon_\ell) - \psi(0)}{\epsilon_\ell} \geq \lim_{l\to\infty} e^T x_\ell \sum_{i\in I(\bar{x})} y_i,$$

which implies

$$\liminf_{\epsilon\downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \geq e^T \bar{x} \sum_{i\in I(\bar{x})} y_i = n\langle e, y\rangle. \tag{4.5.8}$$

Since (4.5.8) holds for each $y \in M$, we have

$$\liminf_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \geq \max\{n\langle e, y\rangle : y \in M\}. \qquad \square$$

We are now ready to show directional differentiability of $\psi(\epsilon)$. With Theorem 4.5.7 and Theorem 4.5.8, we have

$$\max_{y \in M} n\langle e, y\rangle \leq \liminf_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \leq \limsup_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} \leq \max_{y \in M} n\langle e, y\rangle,$$

where M is the set of optimal solutions of $(\mathcal{DP})$. Therefore, $\lim_{\epsilon \downarrow 0} \frac{\psi(\epsilon)-\psi(0)}{\epsilon}$ exists and is equal to

$$\psi'(0; 1) = \lim_{\epsilon \downarrow 0} \frac{\psi(\epsilon) - \psi(0)}{\epsilon} = \max_{y \in M} n\langle e, y\rangle. \tag{4.5.9}$$

We note that (4.5.9) coincides with (4.5.2).

4.5.3 The Sensitivity of the Optimal Value Functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$

In this section, we show the sensitivity of the robust optimal value functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ using the result obtained in Section 4.5.2, precisely the equality (4.5.9).

Given two $(\mathcal{R}(\epsilon, P))$ and $(\mathcal{R}(\epsilon, Q))$, we now compare the values of directional derivatives of $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$, as $\epsilon \downarrow 0$.

Theorem 4.5.9. *Given two **LPs** $(\mathcal{R}(\epsilon, P))$ and $(\mathcal{R}(\epsilon, Q))$ and their optimal value functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$, respectively, we have*

$$\psi_P(\epsilon) > \psi_Q(\epsilon), \ \forall \text{ small } \epsilon > 0.$$

Proof. Let M_P and M_Q be the set of optimal solutions:

$$M_P := \operatorname{Argmax}\{\langle Pe, y\rangle : P^T y \leq \bar{w}, y \geq 0\}, \text{ and } M_Q := \operatorname{Argmax}\{\langle Qe, y\rangle : Q^T y \leq \bar{w}, y \geq 0\}.$$

Then, by Item 1 of Lemma 4.2.1, we have

$$\begin{aligned} M_P &= \{y \in \mathbb{R}^m : P^T y = \bar{w}, y \geq 0\}, \\ M_Q &= \{y \in \mathbb{R}^m : Q^T y = \bar{w}, y \geq 0\}. \end{aligned}$$

Then, (4.5.9) leads to

$$\begin{aligned} \psi_P'(0; 1) &= \max\{n\langle e, y\rangle : y \in M_P\}, \\ \psi_Q'(0; 1) &= \max\{n\langle e, y\rangle : y \in M_Q\}. \end{aligned}$$

By Lemma 3.2.10, we have

$$\psi_P'(0;1) = \max\{n\langle e,y\rangle : y \in M_P\} > \max\{n\langle e,y\rangle : y \in M_Q\} = \psi_Q'(0;1).$$

This yields

$$0 < c := \psi_P'(0;1) - \psi_Q'(0;1) = \lim_{\epsilon\downarrow 0}\frac{\psi_P(\epsilon)-\psi_P(0)}{\epsilon} - \lim_{\epsilon\downarrow 0}\frac{\psi_Q(\epsilon)-\psi_Q(0)}{\epsilon}.$$

Then, we have

$$\lim_{\epsilon\downarrow 0}\frac{\psi_P(\epsilon)-\psi_P(0)-\psi_Q(\epsilon)+\psi_Q(0)}{\epsilon} = c.$$

Thus, with $\psi_P(0) = \psi_Q(0)$, we obtain

$$\forall\bar{\epsilon} > 0,\ \exists\delta > 0 \text{ such that } 0 < \epsilon < \delta \implies \left|\frac{\psi_P(\epsilon)-\psi_Q(\epsilon)}{\epsilon} - c\right| < \bar{\epsilon}.$$

This implies that for $\bar{\epsilon}$ satisfying $0 < \bar{\epsilon} < c$, we have

$$-\bar{\epsilon} + c < \frac{\psi_P(\epsilon)-\psi_Q(\epsilon)}{\epsilon} < \bar{\epsilon} + c \implies 0 < \epsilon(-\bar{\epsilon}+c) < \psi_P(\epsilon) - \psi_Q(\epsilon).$$

Therefore, we have

$$\boxed{\psi_P(\epsilon) > \psi_Q(\epsilon),\ \forall \text{ small } \epsilon > 0,} \tag{4.5.10}$$

as desired. □

This implies that the robust optimal value is more *sensitive* if the given **LP** has a *smaller* feasible region near the nominal optimal solution. We note that the result given in (4.5.10) is the same as the one in (4.4.20).

4.5.4 A Sufficient Condition for Differentiability of $\psi(\epsilon)$ at $\epsilon = 0$

We showed that $\psi(\epsilon)$ is differentiable near 0, if $(\mathcal{P})$ is nondegenerate (See Section 4.4.2.). The uniqueness of dual optimal solution played a big role in terms of getting the differentiability of ψ. As we noted in Item 4 of Lemma 4.2.1, we do not have the uniqueness of dual optimal solutions in degenerate cases. In degenerate cases, differnetiability of $\psi(\epsilon)$ at $\epsilon = 0$ may or may not hold depending on some properties of the set of optimal solutions of $(\mathcal{DP})$. In this section, we show a sufficient condition for differentiability of $\psi(\epsilon)$ at $\epsilon = 0$.

In Section 4.5.2, we only considered the directional differentiability of ψ in the *positive* direction and it is given by

$$\psi'(0;1) = \max_{y\in M} n\langle e,y\rangle, \tag{4.5.11}$$

where M is the set of optimal solutions of $(\mathcal{DP})$. By observing the proofs and making necessary changes in the previous section, we can derive the directional differential of ψ in the *negative* direction, that is,

$$\psi'(0;-1) = \max_{y \in M} -n\langle e, y\rangle. \tag{4.5.12}$$

From (4.5.11) and (4.5.12), we can show that degeneracy of the optimal solution of $(\mathcal{P})$ does *not* always imply non-differentiability of $\psi(\epsilon)$ at $\epsilon = 0$. We first observe a special case in Example 4.5.10.

Example 4.5.10. *Let $w^1 = e_1, w^2 = e_2, w^3 = -e_1, w^4 = -e_2$ and $\bar{w} = e_3$. Construct the cone*

$$\operatorname{cone}\{c^i\}_{i=1,\ldots,4}, \text{ where } c^i = \cos\theta w^i + \sin\theta \bar{w} \text{ with } \theta = \pi/4.$$

Then, we have

$$\operatorname{cone}(\mathcal{C}) = \{x \in \mathbb{R}^3 : Px \geq 0\}, \text{ where } P = \frac{1}{\sqrt{7}}\begin{bmatrix} -\sqrt{3} & -\sqrt{3} & 1 \\ \sqrt{3} & -\sqrt{3} & 1 \\ \sqrt{3} & \sqrt{3} & 1 \\ -\sqrt{3} & \sqrt{3} & 1 \end{bmatrix}.$$

The set of dual optimal solutions to $(\mathcal{DP})$ is given by $\overline{M} := \{y \in \mathbb{R}^4 : P^T y = \bar{w},\ y \geq 0\}$. We note that for all $y \in \overline{M}$, we must have $1 = \bar{w}_3 = \langle P(:,3), y\rangle = (1/\sqrt{7})\langle e, y\rangle$. □

From Example 4.5.10, we observe that for all $y \geq 0$ such that $P^T y = \bar{w}$, we must have $\langle e, y\rangle = \sum_{i=1}^4 y_i = \sqrt{7}$. This implies that

$$\max_{y \in \overline{M}} n\langle e, y\rangle = -\max_{y \in \overline{M}} -n\langle e, y\rangle \implies \psi'(0;1) = -\psi'(0;-1). \tag{4.5.13}$$

Hence, the change of optimal value of $\psi(\epsilon)$ from $\epsilon = 0$ in both directions yields the *same* magnitude of change. Thus, when (4.5.13) occurs, we must have the differentiability of $\psi(\epsilon)$ at $\epsilon = 0$. From this observation, we conclude a sufficient condition for the differentiability of $\psi(\epsilon)$ at $\epsilon = 0$.

Theorem 4.5.11. *Given $(\mathcal{P})$ and its dual $(\mathcal{DP})$, let M be the set of optimal solutions of $(\mathcal{DP})$. If*

$$\sum_{i=1}^m y_i = \gamma,\ \forall y \in M, \text{ for some constant } \gamma \in \mathbb{R},$$

then the robust optimal value function $\psi(\epsilon)$ is differentiable at 0. □

Remark 4.5.12. *We note that, in the nondegenerate case, there is only one optimal solution to $(\mathcal{DP})$ (See Item 3 of Remark 4.2.1.). Therefore, Theorem 4.5.11 also applies to the case where given $(\mathcal{P})$ is nondegenerate.*

The Case of Non-differentiability

In this section, we observe an interesting phenomenon in the absence of differentiability of $\psi(\epsilon)$ at $\epsilon = 0$. Given two **LPs** $(\mathcal{R}(\epsilon, P))$ and $(\mathcal{R}(\epsilon, Q))$, where P is the data matrix of $\text{cone}(\mathcal{C})$ and Q is the data matrix of $\text{cone}(\mathcal{D})$, we have the following:

$$\psi'_P(0;1) = \max_{y \in M_P} n\langle e, y\rangle > \max_{y \in M_Q} n\langle e, y\rangle = \psi'_Q(0;1) > 0, \text{ and} \tag{4.5.14}$$

$$\psi'_P(0;-1) = \max_{y \in M_P} -n\langle e, y\rangle < \max_{y \in M_Q} -n\langle e, y\rangle = \psi'_Q(0;-1) < 0. \tag{4.5.15}$$

The strict inequality (4.5.14) implies that the change of $\psi_P(\epsilon)$ at $\epsilon = 0$ to the *positive* direction is *greater* than the change of $\psi_Q(\epsilon)$ at $\epsilon = 0$. The strict inequality (4.5.15) implies that the change of $\psi_P(\epsilon)$ at $\epsilon = 0$ to the *negative* direction is also *greater* than the change of $\psi_Q(\epsilon)$ at $\epsilon = 0$. Therefore, the above implies that at $\epsilon = 0$, the functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ do not kiss; they must *cross*. Figure 4.5.1 shows an illustration of this phenomenon (Figure 4.5.1 was drawn with piece-wise linear functions for illustrative purposes. There is no guarantee that $\psi_P(\epsilon)$ is a piece-wise linear function.).

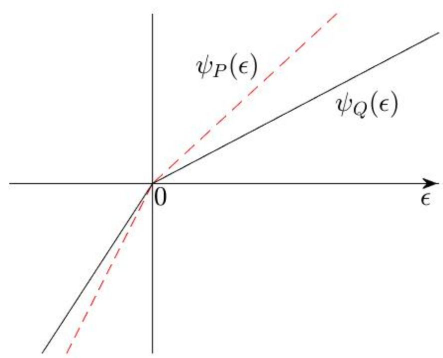


Figure 4.5.1: An illustration of two functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$: Note that two functions *cross* at the origin.

We make an interesting observation on $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ on domain $(-\bar{\epsilon}, 0)$, for some $\bar{\epsilon} > 0$. We introduced the notion of *optimistic counterpart* in Section 2.4.2. Using (2.4.14), we can show that the optimistic counterpart of $(\mathcal{P})$ is

$$\min\{\langle \bar{w}, x\rangle : Px + \epsilon Ex \geq Pe,\ x \geq 0\}. \tag{4.5.16}$$

We pay attention to the coefficient of the matrix E in (4.5.16). We note that the optimal value function of **LP** (4.5.16) can be expressed using the optimal value function of $(\mathcal{R}(\epsilon, P))$, that is,

$$\psi_P(-\boldsymbol{\epsilon}) = \min\{\langle \bar{w}, x\rangle : Px \boldsymbol{+\epsilon}\, Ex \geq Pe,\ x \geq 0\}.$$

Similarly, we may write

$$\psi_Q(-\boldsymbol{\epsilon}) = \min\{\langle \bar{w}, x\rangle : Qx + \boldsymbol{\epsilon}\, Ex \geq Qe,\ x \geq 0\}.$$

Hence, we observe that the optimal value functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ on the *positive* domain are associated with the *robust* counterparts while the optimal value functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ on the *negative* domain are associated with the *optimistic* counterparts.

With (4.5.14) and (4.5.15), we conclude as follows:

1. Given the robust counterparts of two **LPs**, $(\mathcal{R}(\epsilon, Q))$ (with the *bigger* feasible region) enjoys its property more than $(\mathcal{R}(\epsilon, P))$ (with the *smaller* feasible region).
2. Given the optimistic counterparts of two **LPs**, $(\mathcal{R}(-\epsilon, P))$ (with the *smaller* feasible region) enjoys its property more than $(\mathcal{R}(-\epsilon, Q))$ (with the *bigger* feasible region).

4.5.5 The Robust Optimal Value Function $\phi(\epsilon)$

In this section, we make a brief note on the robust optimal value function $\phi(\epsilon)$ of the robust counterpart $(\mathcal{S}(\epsilon))$. A reader who wishes to recall the definition of $\phi(\epsilon)$ and $(\mathcal{S}(\epsilon))$ may refer to Section 4.2.2 and Section 4.3. We recall that we state the similar argument in Section 4.4.3 under nondegenerate cases.

From Section 4.5.1 to Section 4.5.4, we have studied the robust optimal value function $\psi(\epsilon)$ of $(\mathcal{R}(\epsilon))$ under the degenerate case. We note that the argument on the optimal value function $\phi(\epsilon)$ of $(\mathcal{S}(\epsilon))$ parallels the argument in the previous sections owing to Lemma 3.2.1 (Note that Lemma 3.2.10 played an important role on getting the sensitivity result in Section 4.5.2.). Thus, by making necessary changes in the previous section, we have the following result: given two **LPs** $(\mathcal{S}(\epsilon, C))$ and $(\mathcal{S}(\epsilon, D))$ and their optimal value functions $\phi_C(\epsilon)$ and $\phi_D(\epsilon)$, respectively, we have

$$\phi_C(\epsilon) < \phi_D(\epsilon),\ \forall \text{ small } \epsilon > 0.$$

We recall that $\text{cone}(\mathcal{C})^\circ \supset \text{cone}(\mathcal{D})^\circ$ since polarization is order-reversing (See Lemma 2.1.25.). Hence, the above argument also shows that there is an interval such that the robust optimal value of a *smaller* feasible region is always worse than the robust optimal value of a *bigger* feasible region.

We also conclude the following: given $(\mathcal{Q})$ and its dual $(\mathcal{DQ})$, let M be the set of optimal solutions of $(\mathcal{DQ})$. If

$$\sum_{i=1}^{m} y_i = \gamma,\ \forall y \in M,\ \text{for some constant } \gamma,$$

then $\phi(\epsilon)$ is differentiable at $\epsilon = 0$.

We also have

$$\phi'_C(0;1) = \max_{y \in M_C} n\langle e, y\rangle < \max_{y \in M_D} n\langle e, y\rangle = \phi'_D(0;1), \text{ and}$$

$$\phi'_C(0;-1) = \max_{y \in M_C} -n\langle e, y\rangle > \max_{y \in M_D} -n\langle e, y\rangle = \phi'_D(0;-1).$$

Therefore, $\phi_C(\epsilon)$ and $\phi_D(\epsilon)$ cross at $\epsilon = 0$ (See Figure 4.5.1 for an illustration after replacing $\psi_P(\epsilon)$ to $\phi_D(\epsilon)$ and replacing $\psi_Q(\epsilon)$ to $\phi_C(\epsilon)$.).

4.6 Interpretations to the Robust Optimization Problem

In this section, we briefly summarize the arguments of the previous sections as well as how the result applies to the original robust optimization problem.

Given the vectors $\bar{w} \in \mathbb{R}^n$ and $\{w^i\}_{i=1,\ldots,k} \subset \mathbb{R}^n$ satisfying Hypothesis 3.1.1, we construct a family of cones so that we can control their sizes using $\theta \in (0, \pi/2)$ (Section 3.1). We fix two distinct $\theta, \bar{\theta}$ with $\theta < \bar{\theta}$ and construct two cones:

$$\begin{array}{ll} \text{cone}(\mathcal{C}) := \text{cone}\left(\{c^i\}_{i \in \{1,\ldots,l\}}\right), & \text{where } c^i := \cos\theta\bar{w} + \sin\theta w^i, \\ \text{cone}(\mathcal{D}) := \text{cone}\left(\{d^i\}_{i \in \{1,\ldots,l\}}\right), & \text{where } d^i := \cos\bar{\theta}\bar{w} + \sin\bar{\theta} w^i. \end{array}$$

We note that $\text{cone}(\mathcal{C}) \subsetneq \text{cone}(\mathcal{D})$.

We first focus on $\text{cone}(\mathcal{C})$. With Algorithm 2.1.1 and Table 2.1.2, we find the halfspaces defining the cone:

$$\text{cone}(\mathcal{C}) := \{x \in \mathbb{R}^n : Px \geq 0\}.$$

We translate the cone by e and obtain the following system:

$$\{x \in \mathbb{R}^n : Px \geq 0\} + e = \{x \in \mathbb{R}^n : Px \geq Pe\}.$$

We then construct **LPs** using the cones constructed above along with the nonnegativity on the variables (Section 4.2):

$$(\mathcal{P}) \qquad \min\left\{\langle \bar{w}, x\rangle : Px \geq Pe,\ x \geq 0\right\}.$$

We note that the translation of the cone yields $(\mathcal{P})$ to have its unique optimal solution at $x^* = e$. We also note that the constraint $x \geq 0$ of $(\mathcal{P})$ may change the feasible region of $(\mathcal{P})$ but is redundant in the sense that it does not affect the optimal solution at all. The constraint $x \geq 0$ seems redundant knowing that $x^* = e$ is the unique optimal solution, but it plays an important role on forming a nice parametric form of robust counterpart; see (2.4.9).

Now we suppose that each entry of the LHS coefficient matrix P of the constraint $Px \geq Pe$ is *uncertain* and we are given small perturbation range $[-\epsilon, \epsilon]$ for each entry of

P. We then form the following robust counterpart (Section 4.3):

$$(\mathcal{R}(\epsilon, P)) \qquad \psi_P(\epsilon) = \min\left\{\langle \bar{w}, x\rangle : Px - \epsilon Ex \geq Pe,\ x \geq 0\right\}.$$

Similarly, we do the same procedure described above for $\text{cone}(\mathcal{D}) = \{x \in \mathbb{R}^n : Qx \geq 0\}$ and obtain the following robust counterpart:

$$(\mathcal{R}(\epsilon, Q)) \qquad \psi_Q(\epsilon) = \min\left\{\langle \bar{w}, x\rangle : Qx - \epsilon Ex \geq Qe,\ x \geq 0\right\}.$$

With the robust counterparts $(\mathcal{R}(\epsilon, P))$ and $(\mathcal{R}(\epsilon, Q))$, we show in Section 4.4 (in nondegenerate cases) and Section 4.5 (in degenerate cases) that the optimal value function of a *smaller* feasible region yields a *worse* robust optimal value than the robust optimal value function of a *bigger* feasible region:

$$\psi_P(\epsilon) > \psi_Q(\epsilon), \quad \text{for small } \epsilon > 0.$$

That is, the robust optimal value of a smaller feasible region is *more sensitive* than the robust optimal value of a bigger feasible region.

Though this thesis exploits a certain class of **LPs**, the arguments in this thesis give us insights on why some robust optimization problems yield very pessimistic robust optimal values. Essentially, the sensitivity of the robust optimal value function is related to the magnitude of dual optimal solutions. And the magnitude of dual optimal solutions is strongly related to the geometry near the primal optimal solution. To our knowledge, there are no existing results in robust optimization problems that involve geometric structures near the nominal optimal solutions and try to study the robustness. Again, this gives us a moment to think about what may drive the robust optimal values to become conservative.

Chapter 5

Numerical Result

In this chapter, we test the results presented in Chapter 4 numerically. We generated our data as presented in Section 3.1.2.

Figure 5.0.1 shows the changes in the optimal value $\psi \cong \psi(\theta)$ in (4.3.1), with respect to changes in θ, i.e., with respect to changes in the size of cone($\mathcal{C}$). We keep $\epsilon = 10^{-3}$ fixed for all instances. (Note that we can change the size of cone($\mathcal{C}$) by making changes to θ.). As we showed in Section 4.5.2, a *smaller* feasible region always yields a *larger* change in the optimal value under data uncertainty.

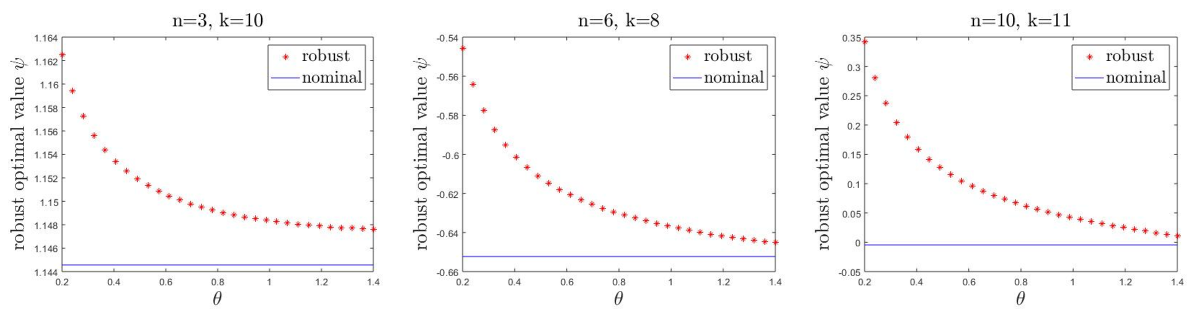


Figure 5.0.1: Fix $\epsilon = 10^{-3}$; changes in robust optimal value $\psi(\theta)$ w.r.t. changes in θ.

Figure 5.0.2 shows the changes in the optimal value $\phi \cong \phi(\theta)$ with respect to changes in θ, i.e., with respect to the changes in the size of cone($\mathcal{C}$). We keep $\epsilon = 10^{-3}$ fixed again. We note that a *smaller* θ means that the negative polar cone of the constructed cone is *bigger* (See Lemma 2.1.25.). Therefore, we see that a *smaller* feasible region yields a *larger* change in the optimal value under data uncertainty, as was shown in Section 4.5.5.

Let cone($\mathcal{C}$) $= \{x \in \mathbb{R}^n : Px \geq 0\}$ be constructed with $\theta = \pi/6$, and cone($\mathcal{D}$) $= \{x \in \mathbb{R}^n : Qx \geq 0\}$ be constructed with $\theta = \pi/4$. Then Figure 5.0.3 shows the change in the optimal value functions $\psi_P(\epsilon)$ and $\psi_Q(\epsilon)$ with respect to the changes of ϵ from 10^{-10} to 10^{-2}. We note that $\psi_P(\epsilon) > \psi_Q(\epsilon)$ from the given instances and this means that: *a larger given uncertainty set corresponding to a larger ϵ, yields a larger change in the robust optimal value.*

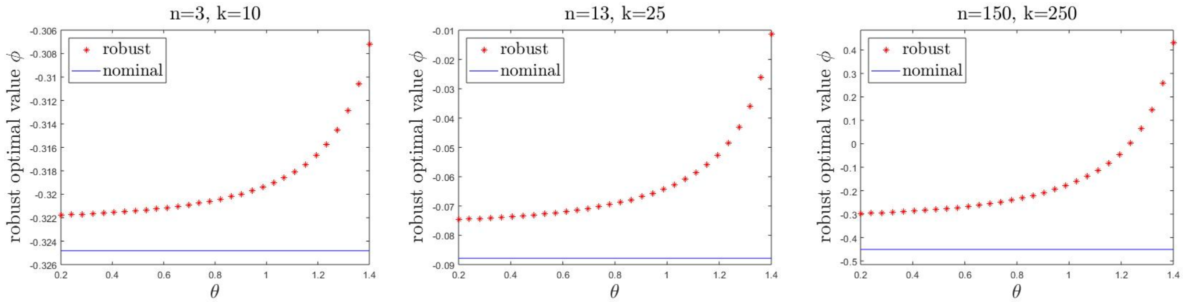


Figure 5.0.2: Fix $\epsilon = 10^{-3}$; changes in robust optimal value $\phi(\theta)$.

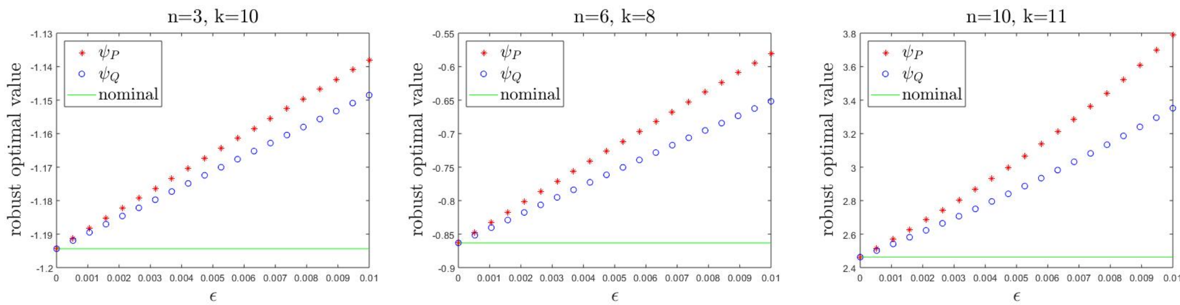


Figure 5.0.3: Two cones cone($\mathcal{C}$) $\subset$ cone($\mathcal{D}$); changes in $\psi_P(\epsilon)$, $\psi_Q(\epsilon)$ with $\epsilon \in [10^{-10}, 10^{-2}]$.

Given cone($\mathcal{C}$) constructed with $\theta = \pi/6$, and cone($\mathcal{D}$) constructed with $\theta = \pi/4$, we have cone$(\mathcal{C})^\circ \supset$ cone$(\mathcal{D})^\circ$ (See Lemma 2.1.25.). Figure 5.0.4 shows the changes in the optimal value functions $\phi_C(\epsilon)$ and $\phi_D(\epsilon)$ for the problems over the polar cones, with respect to the changes in ϵ varying from 10^{-10} to 10^{-2}. We note that in the given intervals, $\phi_C(\epsilon) < \phi_D(\epsilon)$. As above, this means that a *smaller* feasible region in the nominal problem implies: *a larger given uncertainty set corresponding to a larger ϵ, yields a larger change in the robust optimal value.*

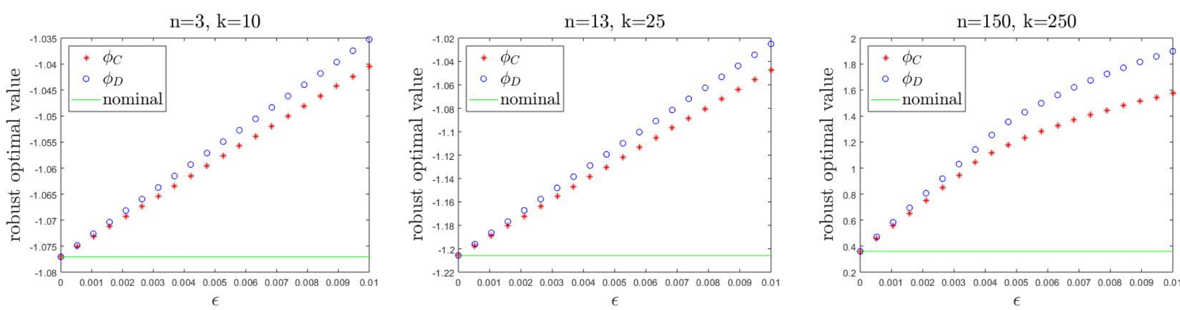


Figure 5.0.4: two polar cones cone$(\mathcal{C})^\circ \supset$ cone$(\mathcal{D})^\circ$; changes in $\phi_C(\epsilon)$, $\phi_D(\epsilon)$ with $\epsilon \in [10^{-10}, 10^{-2}]$.

Chapter 6

Conclusions and Further Notes

In this thesis, we study the special case of linear optimization to show what may affect the sensitivity of a robust optimization reformulation. In this special case, we show that the robust optimization problem with a locally *smaller* feasible region yields a more conservative robust optimal value than the one with a locally *bigger* feasible region. Following is a brief summary of the results presented throughout this thesis.

Given the vectors $\bar{w} \in \mathbb{R}^n$ and $\{w^i\}_{i=1,\ldots,k} \subset \mathbb{R}^n$ satisfying Hypothesis 3.1.1, we construct a family of proper cones so that we can control their sizes using $\theta \in (0, \pi/2)$ in Section 3.1. We fix two distinct $\theta, \bar{\theta} \in (0, \pi/2)$ with $\theta < \bar{\theta}$ and construct two cones:

$$\begin{aligned} \operatorname{cone}(\mathcal{C}) &:= \operatorname{cone}\left(\{c^i\}_{i\in\{1,\ldots,l\}}\right), \quad \text{where } c^i := \cos\theta\bar{w} + \sin\theta w^i, \\ \operatorname{cone}(\mathcal{D}) &:= \operatorname{cone}\left(\{d^i\}_{i\in\{1,\ldots,l\}}\right), \quad \text{where } d^i := \cos\bar{\theta}\bar{w} + \sin\bar{\theta} w^i. \end{aligned}$$

The cones satisfy strict containment, $\operatorname{cone}(\mathcal{C}) \subsetneq \operatorname{cone}(\mathcal{D})$.

We first focus on $\operatorname{cone}(\mathcal{C})$. With Algorithm 2.1.1 and Table 2.1.2, we find the halfspaces defining the cone:

$$\operatorname{cone}(\mathcal{C}) := \{x \in \mathbb{R}^n : Px \geq 0\}.$$

We translate $\operatorname{cone}(\mathcal{C})$ by $e \in \mathbb{R}^n$ and obtain the following system:

$$\{x \in \mathbb{R}^n : Px \geq 0\} + e = \{x \in \mathbb{R}^n : Px \geq Pe\}.$$

We then construct **LPs** using the cones constructed above along with the nonnegativity on the variables in Section 4.2:

$$(\mathcal{P}) \qquad \min\left\{\langle \bar{w}, x\rangle : Px \geq Pe,\ x \geq 0\right\}.$$

Now suppose that each entry of the LHS coefficient matrix P of the constraint $Px \geq Pe$ is *uncertain* in the given perturbation range $[-\epsilon, \epsilon]$. We form the following robust counterpart $(\mathcal{R}(\epsilon, P))$ as in Section 4.3:

$$(\mathcal{R}(\epsilon, P)) \qquad \psi_P(\epsilon) := \min\left\{\langle \bar{w}, x\rangle : Px - \epsilon Ex \geq Pe,\ x \geq 0\right\}.$$

Similarly, we form the following robust counterpart $(\mathcal{R}(\epsilon, Q))$ using $\text{cone}(\mathcal{D}) = \{x \in \mathbb{R}^n : Qx \geq 0\}$:

$$(\mathcal{R}(\epsilon, Q)) \qquad \psi_Q(\epsilon) := \min \{\langle \bar{w}, x \rangle : Qx - \epsilon Ex \geq Qe,\ x \geq 0\}.$$

With the robust counterparts $(\mathcal{R}(\epsilon, P))$ and $(\mathcal{R}(\epsilon, Q))$, we show in Section 4.4 (in nondegenerate cases) and Section 4.5 (in degenerate cases) that the robust optimal value function of a locally *smaller* feasible region yields a *worse* robust optimal value than the robust optimal value function of a locally *bigger* feasible region:

$$\psi_P(\epsilon) > \psi_Q(\epsilon), \quad \text{for small } \epsilon > 0.$$

That is, the robust optimal value of a locally smaller feasible region is *more sensitive* than the robust optimal value of a locally bigger feasible region.

Further Notes

The motivation of this thesis is to see how the sharpness of a vertex (an optimal solution) of a polyhedron impacts the sensitivity of the robust optimal value. We wish to utilize the notion of solid angle to measure the sharpness of a vertex. However there is no easy way to compute the solid angle of an arbitrary vertex of a polyhedron (For the definition of the solid angle, see [5, Section 11.1].). Given a point $\bar{x}$ in the polyhedron, computing the solid angle at $\bar{x}$ generally requires the knowledge on solid angles of all faces of the polyhedron ([5, Example 11.1] illustrates an example of the 3-simplex.). Hence, we generate data so that we can always control the *sizes* of the vertex neighbourhoods of a polyhedron.

After constructing the cone, we translate the vertex of the cone to e so that we have a non-homogeneous system of linear inequalities. We then impose nonnegativity on the variables. Nonnegativity of variables played an important role in terms of converting the robust counterpart into a simple parametric form. We note that the optimal solution e is in the interior of the nonnegative orthant and we perform *local* analysis. Therefore the robust optimal solution is obtained near e under small perturbations. Hence, imposing nonnegativity on the variables is not restrictive.

Throughout this thesis, we assumed that *all* the coefficients of the data matrix are uncertain with the perturbation range $[-\epsilon, \epsilon]$. We can also derive a similar result in the cases where *only some* of the coefficients of the data matrix are uncertain with the perturbation range $[-\epsilon, \epsilon]$. By making necessary changes to (2.4.4) - (2.4.9), we can show that the inequality with uncertainty

$$\langle \tilde{a}, x \rangle \leq \beta,\ \tilde{a}_i \in [a_i - \epsilon, a_i + \epsilon],\ i \in \mathcal{I} \subset \{1, \ldots, n\}$$

can be reformulated as follows:

$$\langle a, x \rangle + \epsilon \langle u, x \rangle \leq \beta, \text{ where } u \text{ is a 0-1 vector in } \mathbb{R}^n.$$

It follows that the robust counterpart of $(\mathcal{P})$ with *some* uncertain data with perturbation range $[-\epsilon, \epsilon]$ becomes

$$\widehat{\psi}_P(\epsilon) = \min\{\langle \bar{w}, x\rangle : Px - \epsilon Ux \geq Pe, x \geq 0\},$$

where U is a 0-1 matrix. We can define the corresponding robust counterpart using Q with the robust optimal value function $\widehat{\psi}_Q(\epsilon)$. By making a similar argument given in Chapter 4, we can show that $\widehat{\psi}_P(\epsilon) = \bar{y}^T Ue$, for some dual optimal solution $\bar{y}$. It also follows that $\widehat{\psi}_Q(\epsilon) = \bar{z}^T Ue$, for some dual optimal solution $\bar{z}$. Then, by Lemma 3.2.10, we have that

$$\widehat{\psi}_P(\epsilon) \geq \widehat{\psi}_Q(\epsilon), \text{ for small } \epsilon > 0.$$

We note that we might lose the strict inequality above due to some 0 entries in the matrix U.

Similarly, if we are given an $(\mathcal{P})$ where each entry $p_{i,j}$ of the LHS data matrix P is uncertain with perturbation range $[-\epsilon_{i,j}, \epsilon_{i,j}]$ (i.e., we are given a different perturbation range for each entry of P), then we may formulate the robust counterpart as follows:

$$\widetilde{\psi}_P(\epsilon) = \min\{\langle \bar{w}, x\rangle : Px - \epsilon Vx \geq Pe, x \geq 0\},$$

where

$$V_{i,j} = \begin{cases} \epsilon_{i,j}/\epsilon & , \text{ if } P_{i,j} \text{ is uncertain with perturbation range } [-\epsilon_{i,j}, \epsilon_{i,j}] \\ 0 & , \text{ otherwise.} \end{cases}$$

We can define the corresponding robust counterpart using Q with the robust optimal value function $\widetilde{\psi}_Q(\epsilon)$. Then, a similar argument above gives

$$\widetilde{\psi}_P(\epsilon) \geq \widetilde{\psi}_Q(\epsilon), \text{ for small } \epsilon > 0.$$

Limitations and Further Work

We have only considered a special objective vector $\bar{w}$ of an **LP**. One may ask 'What about any objective vector $w' \in \mathbb{R}^n$ satisfying $w' \in \text{int}(\text{cone}(\mathcal{C}))$?'. In the case of such objective vectors, a difficulty arises by Item 3 of Remark 3.2.4, i.e., we cannot guarantee that Lemma 3.2.1 and Lemma 3.2.10 hold with an arbitrary vector instead of $\bar{w}$. It will be interesting to find a condition for an objective vector that guarantees the arguments in this thesis.

In the presence of degeneracy, we do not know if the strict monotonicity of the optimal value functions $\psi(\epsilon)$ and $\phi(\epsilon)$ holds. The numerical result strongly shows that the optimal value functions are strictly increasing. However there is no guarantee that the strict monotonicity holds for nondegenerate cases.

References

[1] C.D. Aliprantis and R. Tourky. *Cones and duality*, volume 84 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2007. 3, 7, 8, 10, 11, 12

[2] H. Arsham and M. Oblak. Perturbation analysis of general lp models: A unified approach to sensitivity, parametric, tolerance, and more-for-less analysis. *Mathematical and Computer Modelling*, 13(8):79 – 102, 1990. vii, 27, 28

[3] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Nonlinear parametric optimization*. Birkhäuser Verlag, Basel-Boston, Mass., 1983. 73

[4] A. Beck and A. Ben-Tal. Duality in robust optimization: primal worst equals dual best. *Oper. Res. Lett.*, 37(1):1–6, 2009. 26

[5] M. Beck and S. Robins. *Computing the continuous discretely*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2015. Integer-point enumeration in polyhedra, With illustrations by David Austin. 91

[6] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2009. 22

[7] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004. 21

[8] M.J. Cánovas, M.A. López, J. Parra, and F.J. Toledo. Lipschitz continuity of the optimal value via bounds on the optimal set in linear semi-infinite optimization. *Math. Oper. Res.*, 31(3):478–489, 2006. 74, 76

[9] F.H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990. 16, 19

[10] M. Conforti, G. Cornuéjols, and G. Zambelli. *Integer programming*, volume 271 of *Graduate Texts in Mathematics*. Springer, Cham, 2014. 3, 10, 12, 13

[11] A.V. Fiacco. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, volume 165 of *Mathematics in Science and Engineering*. Academic Press, 1983. 72, 73, 77

[12] R.M. Freund. Postoptimal analysis of a linear program under simultaneous changes in matrix coefficients. Number 24, pages 1–13, 1985. Mathematical programming, I. 29, 65

[13] T. Gal. Shadow prices and sensitivity analysis in linear programming under degeneracy: state-of-the-art survey. *OR Spektrum*, 8(2):59–71, 1986. 27

[14] Tomá ˇ s Gál. *Postoptimal analyses, parametric programming, and related topics*. McGraw-Hill International Book Co., New York, 1979. Translated from the German by Geoffrey V. Davis, With a foreword by H.-J. Zimmermann. 27, 28

[15] M.J. Gisbert, M.J. Cánovas, J. Parra, and F.J. Toledo. Calmness of the optimal value in linear programming. *SIAM J. Optim.*, 28(3):2201–2221, 2018. 28

[16] B. L. Gorissen, I. Yanikoglu, and D. den Hertog. A practical guide to robust optimization. *Omega*, 53:124 – 137, 2015. 24

[17] B. Guenin, J. Könemann, and L. Tunçel. *A Gentle Introduction to Optimization*. Cambridge University Press, 2014. 20, 21

[18] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*. Grundlehren Text Editions. Springer-Verlag, Berlin, 2001. Abridged version of ıt Convex analysis and minimization algorithms. I [Springer, Berlin, 1993; MR1261420 (95m:90001)] and ıt II [ibid.; MR1295240 (95m:90002)]. 18

[19] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013. 69

[20] B. Jansen, J.J. de Jong, C. Roos, and T. Terlaky. Sensitivity analysis in linear programming: just be careful! *European Journal of Operational Research*, 101(1):15 – 28, 1997. 27

[21] B. Jansen, C. Roos, and Tamás Terlaky. An interior point approach to postoptimal and parametric analysis in linear programming report 92-21. 1993. 27

[22] D.G. Luenberger. *Linear and Nonlinear Programming*. Reading. Addison-Wesley, Massachusetts, second edition, 1984. 21

[23] H.D. Mills. Marginal values of matrix games and linear programs. In *Linear inequalities and related systems*, Annals of Mathematics Studies, no. 38, pages 183–193. Princeton University Press, Princeton, N. J., 1956. 27

[24] S. M. Robinson. A characterization of stability in linear programming. *Operations Res.*, 25(3):435–447, 1977. 74

[25] R.T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970. 3, 9, 36

[26] R.T. Rockafellar and R.J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998. 28

[27] T. Saaty and S. Gass. Parametric objective function. I. *J. Operations Res. Soc. Amer.*, 2:316–319, 1954. 27

[28] B. Simon. *Convexity*, volume 187 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2011. An analytic viewpoint. 5

[29] G. Still. Lectures on parametric optimization: An introduction. Technical report, University of Twente, Netherlands, 2018. 77

[30] J.E. Strum. Note on "two-sided shadow prices". *Journal of Accounting Research*, 7(1):160–162, 1969. 27

[31] J.E. Ward and R.E. Wendell. Approaches to sensitivity analysis in linear programming. *Ann. Oper. Res.*, 27(1-4):3–38, 1990. 27

[32] D. De Wolf and Y. Smeers. Generalized derivatives of the optimal value of a linear program with respect to matrix coefficients. Technical report, Universite libre de Bruxelles, Belgium and Universite Catholique de Louvain, Belgium, 2000. 28, 29, 73

Index