UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

Department of Industries
Valparaíso-Chile

On the polyhedral outer-approximation of convex functions with unbounded domains

Thesis submitted by

Nicolás Boyardi Alache

In partial fulfillment of the requirements for the degree of

Master of Science in Industrial Engineering

Thesis director:

Dr. Ing. Pablo Escalona

June 2026

EX UMBRA IN SOLEM

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☐ Memoria o trabajo de título ☒ Tesis de Postgrado

Título del trabajo: On the polyhedral outer-approximation of convex functions with unbounded domains.

Nombre del candidato(a): Nicolás Ignacio Boyardi Alache

Carrera / Grado: Magíster en ciencias de la Ingeniería Industrial

Campus: Casa Central **Departamento: Departamento de Ingeniería Industrial.**

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Pablo Escalona R., en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☑ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por **(marcar una opción)**:

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 30/06/2026 **Firma:**

Estudiante o Candidato(a):

Fecha: 30/06 **Firma:**

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

THESIS TITLE:

On the polyhedral outer-approximation of convex functions with unbounded domains

AUTOR:

NICOLÁS BOYARDI ALACHE

THESIS WORK, presented in partial fulfillment of the requirements for the degree of Master of science in industrial engineering of the Universidad Técnica Federico Santa María.

THESIS COMITEE:

MEMBERS	SIGNATURE
Dr. Sebastián Dávila Universidad de Santiago de Chile, Chile.	____________
Dr. Ing. Pablo Escalona Universidad Técnica Federico Santa María, Chile.	____________
Dr. Cristopher Hermosilla Universidad Técnica Federico Santa María, Chile.	____________
Dr. Rolando Rubilar Universidad Técnica Federico Santa María, Chile.	____________

Valparaíso, june 2025.

Abstract

Computing polyhedral outer-approximation (POA) of convex functions with unbounded domains and uniformly bounded pointwise error suffices to ensure the boundedness of convex MINLP linearizations. This thesis provides apparently new theoretical results that link functional and geometrical properties of a convex function and the finiteness of the error of a given approximation. Furthermore, a cutting-plane algorithm is developed to globally reduce the error of a POA as much as desired. The algorithm is proven correct, and in a series of numerical tests it is confirmed the convergence under reasonable times. Under mild assumptions, it is shown that the approximations developed in this work possesses sufficient properties to ensure that linearizations of convex MINLP are bounded.

Resumen

Para asegurar que las linealizaciones de problemas MINLP convexos sean acotados es suficiente con computar aproximaciones poliedrales externas (POA) de funciones convexas con dominios no acotados, con un error puntual uniformemente acotado. En esta tesis, se presentan resultados teóricos aparentemente nuevos, que conectan propiedades funcionales y geométricas de una función convexa con la finitud del error de alguna aproximación dada. Más aún, desarrollamos un algoritmo de planos cortantes para reducir, tanto como se desee, globalmente el error de una POA. Se prueba la correctitud del algoritmo, y se confirma su convergencia en una serie de experimentos numérico que terminaron en tiempos razonables. Bajo condiciones razonables, se demuestra que las aproximaciones desarrolladas en este trabajo poseen propiedades suficientes para garantizar que las linealizaciones de problemas MINLP convexos sean acotados.

Perhaps the forces that now menace freedom are too strong to be resisted for very long. It is still our duty to do whatever we can to resist them.
–Aldous Huxley (1958).

Contents

Abstract i

Resumen ii

1 Preliminaries 4
- Convex Analysis 4
- Polyhedral theory 12

2 On the existence of POAs of convex functions 16
- Duality of the error of approximation 19
- Existence of POA of convex functions 21
- Characterization of POAs 24
- POA of non-increasing convex functions over translated cones 25

3 The DOA algorithm 27
- The DOA algorithm 27
- Strict improvement of error 30

4 Computational study 31
- Error dynamic 32
- CPU Time dynamic 33

5 A brief discussion of applications 35
- POAs of convex functions with bounded domains 35
- Polyhedral Inner-Approximations (PIAs) of f 36
- PIAs and POAs of f^* 36
- On the linearization of Non-Linear optimization Problems 37

6 Conclusion and future work 39

A Test function 1 **41**

B Test functions 2.1 and $2.\infty$ **42**

C Test functions $3.\mathcal{A}$ and $3.\mathcal{B}$ **44**
Fill-rate function $\mathcal{A}$ 45
Backorder function $\mathcal{B}$ 46

D Test function 4 **49**

Bibliography **50**

Introduction

The property of compactness is among the most useful tools for proving theorems in mathematics. However, compactness is not easily guaranteed. For instance, any unbounded set is not compact. For example, a continuous function on a compact set always possesses a minima and maxima. The same problem over an unbounded set may not possess minima nor maxima. Nonetheless, they may still exist under additional conditions. In general, a result that assumed boundedness is considerably more difficult to prove if this assumption is discarded. Additional hypothesis or completely different techniques are required once a set or a domain is unbounded. An interesting instance of this phenomenon is the approximation of convex functions with unbounded domains. This problem arises naturally in the algorithmic solution of convex MINLP, where *"among the most successful algorithms for solving convex MINLP problems, we can find linear approximation algorithms"* [6]. One of these algorithms involves performing the approximation of convex functions before optimization (e.g., see [2, 3]). When unbounded domains are present, authors cope by approximating the function over a bounded subset of the domain. This thesis addresses the issue of having an unbounded domain by providing apparently new theoretical results and an algorithm to compute POAs of convex functions with unbounded domains.

Convex functions and convex sets are intimately related. In particular, one may construct convex sets using convex functions (e.g. level sets), and *vice versa* (e.g. functions defined by epigraphs). The problem of approximation convex sets is primarily separated considering the boundedness of the set, and secondarily by its dimension. Approximating bounded convex sets has been systematically studied for at least a century [5], and many algorithms are available. The study of bounded convex sets is also called the *convex body theory*. This area includes their approximation, algorithms, efficiency of algorithms, and so on. On the other hand, approximating unbounded convex sets is a problem with considerably less literature available. Necessary and sufficient conditions for approximating unbounded convex sets with finite Hausdorff distance were proved recently in [4]. A necessary condition is that the approximated set possesses a polyhedral recession cone. Recently in [1], they extend the notion of error to a wider class than the restrictive sets with polyhedral recession cones.

The problem of approximating convex functions is primarily separated considering the boundedness of the domain instead, and secondarily by the dimension of said domain. Most of the literature addresses the case of approximating convex functions defined on a bounded interval of $\mathbb{R}$ (e.g. [3]), due to the well-ordering of the real line. When the function is defined on a bounded set of $\mathbb{R}^n$, $n > 1$, fewer, and weaker, literature is available. For example, in [9] an optimization problem is designed to construct polyhedral approximations of non-linear functions that interpolate them over a finite lattice. In [10], neural networks are used to construct POAs of convex functions over bounded domains in $\mathbb{R}^n$, $n \geq 1$. This method relies on the boundedness of the domain, since the machine learning algorithm behind the scenes resorts to random sampling. To the best of our knowledge, the approximation of convex functions with unbounded n-dimensional domains remains unaddressed.

As stated earlier, this thesis addresses the POA of convex functions with unbounded domains. Some results in the present work are quite similar to those in [1, 4] regarding approximation of unbounded convex sets. More precisely, the existence of POAs of a function f is showed equivalent to the polyhedrality of the recession function f^∞ and the closedness of the domain of the conjugate. This result contrasts to the one in [4], which shows necessary the polyhedrality of the recession cone $\mathcal{C}^\infty$ to guarantee the existence of polyhedral approximations of the convex set $\mathcal{C}$. Regarding the existence of POAs of f, it is also proven necessary and sufficient the polyhedrality of *dom* f^*. These characterizations can be further detailed. If the approximations are constructed using supporting hyperplanes to $epi(f)$, then the POAs can be characterized as those polyhedral functions with at least some facets with particular slopes. These particular slopes may be obtained by the recession function f^∞ or analyzing the extreme points of the domain of the conjugate. A cutting-plane algorithm is proposed to globally reduce the error of a POA as small as desired, namely, the Dual-Outer Approximation algorithm (DOA algorithm). In essence, the algorithm first searches the points in the domain such that the pointwise difference between f and the POA is exactly the error of approximation. These points are characterized as the projection onto the domain of the vertices of the POA's epigraphs. Based on the aforementioned points, new hyperplanes are computed and incorporated to the approximation, until the error of approximation is arbitrarily small.

This thesis is structured as follows. In Chapter 1, definitions and notation regarding convex analysis and polyhedral theory are provided. The theoretical results regarding existence and characterizations for the POAs of convex functions are presented in Chapter 2. Next, the DOA algorithm is stated, and under mild assumptions, proved correct in Chapter 3. The computational performance and convergence of the DOA algorithm for some test functions is studied in Chapter 4. A brief discussion regarding some applications of the theoretical results in this thesis are covered in Chapter 5, specially the construction of bounded linear relaxations of convex MINLP. Chapter 6 concludes the work and presents some lines for future research. The appendix 6 contains the parameters computations required for the computational study of Chapter 4.

Chapter 1

Preliminaries

This chapter provides some definitions and results from Convex Analysis and Polyhedral Theory to guarantee a fluid reading of the present thesis. Most of the theoretical content in the present work is described mathematically in terms of the tools from these two subjects. Most of what it is present in this chapter comes from the famous book of R.T. Rockafellar *Convex Analysis* ([7]).

Additionally, it is assumed that the reader possess basic knowledge in Real Analysis, more precisely, of basic topology (open and closed sets), continuity of functions, and such.

Convex Analysis

In this section fundamental concepts such as convex set, convex function, conjugate function, recession function, subgradient, subdifferential, are defined, and the relationship between themselves are established. These concepts are key to understand the mathematical results regarding POAs of convex functions with unbounded domains. More precisely, the angle of duality that is relevant to the present thesis are alternative ways to compute explicitly the domain of the conjugate function, where the recession function appears naturally.

Definition 1.1 (Convex set)**:** A set $\mathcal{C} \subseteq \mathbb{R}^n$ is called a convex set if and only if for any pair $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ and for any scalar $\lambda \in [0, 1]$, then the convex combination of this pair is contained within $\mathcal{C}$, i.e., if the following inclusion:

$$\lambda \mathbf{x} + (1-\lambda)\mathbf{y} \in \mathcal{C},$$

holds.

Definition 1.2 (Convex hull)**:** For any (possibly non-convex) set $\mathcal{C} \subseteq \mathbb{R}^n$, the convex hull of $\mathcal{C}$, denoted $conv(\mathcal{C})$, is defined as the smallest convex set that contains the set $\mathcal{C}$. Equivalently, it is the set resulting from all the convex combinations that are possible using elements of $\mathcal{C}$, i.e.:

$$conv(\mathcal{C}) := \{\lambda \mathbf{x} + (1-\lambda)\mathbf{y} \mid \mathbf{x}, \mathbf{y} \in \mathcal{C}, \lambda \in [0,1]\}.$$

For the rest of the chapter, assume that $\mathcal{C} \subseteq \mathbb{R}^n$ is a closed convex set. One fundamental property of convex sets is that they are fully described by the intersection of half-spaces containing them. The definition of half-space, and the precise statement of their representation is given below.

Definition 1.3 (Hyperplanes, Half-spaces and Supporting Hyperplanes)**:** A set $\mathcal{H} \subseteq \mathbb{R}^n$ is said to be a hyperplane if there exists a vector $\boldsymbol{a} \in \mathbb{R}^n$ and constant $b \in \mathbb{R}$ such that $\mathcal{H} = \{\mathbf{x} \in \mathbb{R}^n \mid \boldsymbol{a}^\top \mathbf{x} = b\}$. Instead, one says that $\mathcal{H}$ is a half-space if $\mathcal{H} = \{\mathbf{x} \in \mathbb{R}^n \mid \boldsymbol{a}^\top \mathbf{x} \leq b\}$. Lastly, if $\mathcal{H}$ is a hyperplane, it is said that $\mathcal{H}$ is a supporting hyperplane of $\mathcal{C}$ if either one of the two possible half-spaces where $\mathcal{H}$ is their boundary contain $\mathcal{C}$.

Another relevant concept regarding the description of convex sets is that of extreme point.

Definition 1.4 (Extreme points)**:** A point $\mathbf{c} \in \mathcal{C}$ is an extreme point of $\mathcal{C}$ if $\mathcal{C} \setminus \{\mathbf{c}\}$ is still a convex set. The set $ext(\mathcal{C})$ is the set of all the extreme points of $\mathcal{C}$.

As an example, one use of the concept of extreme points is the Krein-Milmann theorem, which states that every compact convex set is the convex hull of its extreme points. However, for unbounded convex sets this representation is no longer valid. In the polyhedral case, a fix is known, considering the recession cone of the set. This shall be defined later, but it motivates defining the recession cone and cones themselves now.

Definition 1.5 (Cone)**:** A convex set $\mathcal{K}$ is said to be a cone if it is closed under non-negative scalar multiplication, i.e., if for all $\lambda \geq 0$, then $\lambda\mathcal{K} \subseteq \mathcal{K}$.

Definition 1.6 (Recession cone)**:** The recession cone of a convex set $\mathcal{C}$, denoted $\mathcal{C}^\infty$, is defined as the set of vectors $\mathbf{d} \in \mathbb{R}^n$ such that $\mathcal{C}$ is unbounded in the direction of $\mathbf{d}$, i.e., $\mathcal{C}^\infty := \{\mathbf{d} \in \mathbb{R}^n \mid \mathcal{C} + \mathbf{d} \subseteq \mathcal{C}\}$ (see [7, Theorem 8.1]). This set is always a cone.

From now on, assume that $\mathcal{K} \subseteq \mathbb{R}^n$ is a closed and convex cone. Observe that the only possible extreme point of $\mathcal{K}$ could be the point $\mathbf{0}$ itself. If it is the case, then it is said that $\mathcal{K}$ is **pointed**. Being pointed can be characterized as possessing lines, due to the conical structure of the set.

Proposition 1.1. *A cone $\mathcal{K}$ is pointed if and only if $\mathcal{K}$ does not contain any lines.*

Following the definition of cones, it is defined the dual concept of polar cones.

Definition 1.7 (Polar of a set)**:** The polar cone of $\mathcal{C}$ is defined as $\mathcal{C}^\circ := \{\mathbf{d} \in \mathbb{R}^n \mid \mathbf{d}^\top\mathbf{x} \leq 0, \ \ \forall \mathbf{x} \in \mathcal{C}\}$.

The famous polarity theorem states that $\mathcal{K}^\circ$ is also a closed convex cone, and its polar is exactly $\mathcal{K}$, i.e., $\mathcal{K}^{\circ\circ} = (\mathcal{K}^\circ)^\circ = \mathcal{K}$. The definition of polar cone let us define that of normal cone.

Definition 1.8 (Normal cone)**:** The normal cone of $\mathcal{C}$ at $\mathbf{x} \in \mathcal{C}$, denoted as $\mathcal{N}_\mathcal{C}(\mathbf{x})$, is the polar set of $\mathcal{C} - x$, i.e., $\mathcal{N}_\mathcal{C}(\mathbf{x}) := (\mathcal{C} - \mathbf{x})^\circ = \{\mathbf{d} \in \mathbb{R}^n \mid \mathbf{d}^\top(\mathbf{y} - \mathbf{x}) \leq 0, \ \ \forall \mathbf{y} \in \mathcal{C}\}$.

The interest for normal cones in this work comes from the fact that subgradients of f can be characterized using normal cones to the epigraph of f. This will become useful much later in Chapter 3, where this characterization is needed to show some strict inequality. This *flash forward* serves its purpose by ending the review on convex sets and changing the topic to convex functions instead.

A fundamental tool around convex functions is the concept of epigraph[1] of a function. This is defined below.

Definition 1.9 (Epigraph of a function)**:** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. The epigraph of f, denoted $epi(f)$, is the set of all the points in $\mathbb{R}^n \times \mathbb{R}$ that are "above" the graph of the function, i.e.:

$$epi(f) := \{(\mathbf{x}, z) \in \mathbb{R}^n \times \mathbb{R} \mid f(\mathbf{x}) \leq z\}.$$

In this thesis, it shall be defined a convex set in the same manner as R.T. Rockafellar did, by linking the convexity of the function with the convexity of its epigraph.

Definition 1.10 (Generalized convex functions)**:** Consider the extended reals $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$, i.e., the real line plus the infinity as a valid value. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be a convex function if and only if $epi(f)$ is a convex set. The function f is said to be closed if $epi(f)$ is a closed set, or, equivalently, if f is lower-semicontinuous.

Definition 1.11 (Effective domain)**:** The effective domain, simply domain, of f, denoted $dom\ f := \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) < \infty\}$, is the projection on $\mathbb{R}^n$ of the epigraph. If $dom\ f \neq \emptyset$, then it is said that f is proper. If f is not negatively infinite, then it is said that f is finite at $dom\ f$.

Definition 1.12 (Indicator and support functions)**:** The indicator function of the set $\mathcal{C}$, denoted $\delta_{\mathcal{C}}$, is the function that is equal to 0 when evaluated in

[1] Literally *what is above the graph*.

any point in $\mathcal{C}$, and infinite otherwise, i.e.:

$$\delta_{\mathcal{C}}(\mathbf{x}) := \begin{cases} 0 & \text{if } \mathbf{x} \in \mathcal{C} \\ \infty^{+} & \text{otherwise.} \end{cases}$$

Likewise, the support function of $\mathcal{C}$ is the supremum over $\mathcal{C}$ of the dot product of the evaluation point with a point in $\mathcal{C}$, i.e.:

$$\sigma_{\mathcal{C}}(\mathbf{x}) := \sup_{\mathbf{y} \in \mathcal{C}} \mathbf{x}^{\top}\mathbf{y}.$$

The indicator and support functions are fundamental in this work. One of these reasons is remarked below.

Remark 1.1: For any convex set $\mathcal{D} \subseteq \mathbb{R}^n$ such that the real-valued function $f : \mathcal{D} \to \mathbb{R}^n$ is a convex function, there is a natural extension to the extended-real valued functions as $f + \delta_{\mathcal{D}}$. Referring to the latter may be more suitable in classical convex analysis, and the former may be more natural in other settings, such as optimization problems.

The rest of the reasons hide behind the concept of conjugacy. For instance, for closed and convex sets, the indicator and support functions are *conjugate* to each other. Below is defined what it is meant by being the conjugate of some function.

Definition 1.13 (Conjugate function)**:** The convex conjugate of f is defined as follows:

$$f^{*}(\mathbf{x}^{*}) := \sup_{\mathbf{x} \in dom\ f} \mathbf{x}^{*\top}\mathbf{x} - f(\mathbf{x}).$$

The conjugate operator $(\cdot)^{*}$ is the operator that maps f into f^{*}. The operation of applying the conjugate operator to f is referred as *conjugating*.

Proposition 1.2. *The indicator and support function (over a non-empty closed and convex set) are conjugate to each other, i.e.,* $\delta_{\mathcal{C}}^{*} \equiv \sigma_{\mathcal{C}}$ *and* $\sigma_{\mathcal{C}}^{*} = \delta_{\mathcal{C}}$.

Proposition 1.3 (Properties of the conjugacy)**.** *For any convex function f, the conjugate f^* is always a closed and convex function. Moreover, if f is proper, then f^* is proper, and if f is also closed, then f and f^* are conjugate to each other, i.e., $f^{**} = (f^*)^* = f$.*

Proposition 1.4 (Algebraic properties of the conjugate)**.** *For any $c \in \mathbb{R}, \mathbf{a} \in \mathbb{R}^n$ and $\alpha > 0$, the following relations hold:*

- $(f + c)^* = f^* - c$.
- $(f(\cdot) + \mathbf{a}^\top(\cdot))^*(\mathbf{x}^*) = f^*(\mathbf{x}^* - \mathbf{a})$.
- *For any $\mathbf{x}, \mathbf{x}^* \in \mathbb{R}^n$, $f(\mathbf{x}) + f^*(\mathbf{x}^*) \geq \mathbf{x}^\top \mathbf{x}^*$ holds*[2].
- *If g is a convex function satisfying the inequality $g \leq f$, then the reversed inequality holds for their conjugates, i.e., $g^* \geq f^*$.*

Due to the last property mentioned above, the conjugacy operation is said to be order-reversing.

The definition of conjugate function, due to Fenchel and Rockafellar, encapsulates a deep meaning in geometrical terms, particularly regarding supporting hyperplanes. In functional terms, supporting hyperplanes of *epi*(f) are mostly characterized using the concept of subgradients, which is defined below.

Definition 1.14 (Subgradients and subdifferentials)**:** A vector $\mathbf{x}^* \in \mathbb{R}^n$ is said to be a subgradient of f at a point $\mathbf{x} \in \mathbb{R}^n$ if and only if the following inequality holds:

$$f(\mathbf{x}) + \mathbf{x}^{*\top}(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{x}), \qquad \forall \mathbf{y} \in \mathbb{R}^n.$$

The set of all subgradients of f at $\mathbf{x}$ is denoted as $\partial f(\mathbf{x})$. The multivalued mapping ∂f is called the subdifferential. The set of all points $\mathbf{x} \in$ *dom* f such that $\partial f(\mathbf{x})$ is non-empty is denoted by *dom* ∂f.

Remark 1.2: An important geometrical interpretation, highlighted by R.T. Rockafellar [7, p.214], is that when f is finite at some point $\mathbf{x} \in \mathbb{R}^n$, then the graph of the function $h(\mathbf{y}) := f(\mathbf{x}) + \mathbf{x}^{*\top}(\mathbf{y} - \mathbf{x})$ is a non-vertical supporting-hyperplane to *epi*(f) at the point $(\mathbf{x}, f(\mathbf{x}))$.

[2]This inequality is usually referred as the Fenchel Inequality.

The Fenchel Inequality can become the *Fenchel Equality* for particular choices of $\mathbf{x}$ and $\mathbf{x}^*$. This is stated next.

Proposition 1.5. *The Fenchel Equality $f(\mathbf{x}) + f(\mathbf{x}^*) = \mathbf{x}^\top \mathbf{x}^*$ holds if and only if $\mathbf{x}^* \in \partial f(\mathbf{x})$, or, equivalently, if $\mathbf{x} \in \partial f^*(\mathbf{x}^*)$.*

Remark 1.3: A consequence of the Fenchel (in)equality is that it allows to characterize the minimum of a convex function in terms of the subdifferential. In particular, the equality $f^*(\mathbf{0}) = -min\ f$ holds in the extended-real-valued sense.

Another interesting property of subgradients is that they can be characterized using normal cones to the epigraph.

Proposition 1.6 (Epigraphical description of subgradients)**.** *Let f be any closed, proper and convex function. Then $\mathbf{x}^* \in \partial f(\mathbf{x})$ if and only if $(\mathbf{x}^*, -1) \in \mathcal{N}_{epi(f)}(\mathbf{x}, f(\mathbf{x}))$.*

Proof. See [8, §8]. □

It should be noted that the points in the domain of f constitutes the possible subgradients of f^*. Likewise, the points in the domain of f^* are the possible subgradients of f. This seemingly fortuitous relation is more than a mere coincidence, and in fact, a full duality relation. To establish this result, we need to discuss the so-called recession function.

Definition 1.15 (Recession Function)**:** The recession function of f, denoted f^∞, is defined as the function such that its epigraph is the recession cone of f's epigraph, i.e., $epi(f^\infty) := epi(f)^\infty$. The recession operator $(\cdot)^\infty$ is the operator that maps f to f^∞. The operation of applying the conjugate operator to f is referred as *receding*.

Some basic properties are directly deduced from the definition.

Proposition 1.7. *If f is a proper and convex function, then the following formula for computing f^∞ holds:*

$$f^\infty(\mathbf{d}) = \sup\{f(\mathbf{x} + \mathbf{d}) - f(\mathbf{x}),\ \ \mathbf{x} \in dom\ f\}, \forall \mathbf{d} \in \mathbb{R}^n.$$

Additionally, if f is proper, then f^∞ is a positively homogeneous proper convex function, i.e., $f^\infty(\lambda \mathbf{d}) = \lambda f^\infty(\mathbf{d})$ for all $\lambda \geq 0$. Moreover, if f is closed, then f^∞ is closed too.

Some algebraic properties of the recession function are stated below.

Proposition 1.8 (Algebraic properties of the recession function)**.** *If f is closed, then the following formulas holds [7, Theorem 8.5]:*

$$f^\infty(\mathbf{d}) = \sup_{\lambda>0} \frac{f(\mathbf{x}+\lambda\mathbf{d}) - f(\mathbf{x})}{\lambda} = \lim_{\lambda\to\infty^+} \frac{f(\mathbf{x}+\lambda\mathbf{d}) - f(\mathbf{x})}{\lambda}, \qquad \forall \mathbf{d} \in \mathbb{R}^n.$$

Moreover, for any convex function g, constant $c \in \mathbb{R}$, it holds that:

- $(f+g)^\infty \equiv f^\infty + g^\infty$.
- $(f+c)^\infty \equiv f^\infty$.
- $(\sup_i g_i(\cdot))^\infty(\mathbf{d}) = \sup_i g_i^\infty(\mathbf{d})$.
- *For any two closed, proper and convex functions $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ such that $g \leq f$, then $g^\infty \leq f^\infty$.*
- *The recession function is involutory, i.e., $(f^\infty)^\infty = f^\infty$.*

Due to the last two properties mentioned above, the recession operation is said to be order-preserving and involutory.

In practice, definitions using the supremum and limit operations are useful for the manual computation of the recession function.

R.T. Rockafellar proved the aforementioned duality between the subgradients of f (f^*), and the domain of f^* (f). This result makes use of the support function.

Proposition 1.9. *If f is proper and convex, then $f^{*\infty} = \sigma_{dom\ f}$, and if it is also closed, then $f^\infty \equiv \sigma_{dom\ f^*}$.*

Proof. See [7, Theorem 13.3]. ☐

One of the main consequences of the result above is that f^∞ is polyhedral if and only if $\overline{dom\ f^*}$ is a polyhedral set. Another interesting, and maybe stronger, relation is deduced from the above result. More precisely, a strong geometrical tie exists between *dom* f and *dom* f^*. This is stated next.

Proposition 1.10. *The recession cones of dom f and dom f^* are polar to each other, i.e., (dom $f)^{\infty\circ} = ($dom $f^*)^{\infty}$.*

Proof. See [7, Theorem 14.2]. □

Corolary 1.1. *It is true that $\delta_{\mathcal{K}} \equiv \sigma_{\mathcal{K}^\circ}$ and $\delta_{\mathcal{K}^\circ} \equiv \sigma_{\mathcal{K}}$.*

Polyhedral theory

Polyhedral theory can be regarded as a branch of convex analysis where the convex sets or convex functions considered are finitely generated, i.e., that the set or function can be fully described using only a finite amount of data and operations. This section is devoted to providing these finite characterizations.

Definition 1.16 (Polyhedral set)**:** A set $\mathcal{P} \subseteq \mathbb{R}^n$ is called a polyhedral set if and only if it is the finite intersection of half-spaces, i.e., $\mathcal{P} = \cap_{i\in I}\{\mathbf{x} \in \mathbb{R}^n \mid \alpha_i^\top \mathbf{x} - \beta_i \leq 0\}$, for some $\alpha_i \in \mathbb{R}^n$ and $\beta_i \in \mathbb{R}$, for all $i \in I$, where $I \subset \mathbb{N}$ is finite [7, §19].

Remark 1.4: Every polyhedral set is a closed and convex set. The set of extreme points of a polyhedral set is at most finite. Thus, they are referred as vertices instead. Extreme directions are the minimal directions which their conical hull constitutes the recession cone of a particular polyhedral set. The famous Minkowski-Weyl theorem states that every polyhedral set, regardless of its boundedness, is the Minkowski sum between the convex hull of some finite set of points and its recession cone (which itself is the conical hull of a finite set of directions). The Minkowski-Weyl theorem is the analog of the Krein-Milmann theorem in the polyhedral case, but considering the unbounded case too. Any polyhedral set contains at least one vertex if and only if it does not contain any lines.

Convex functions where defined as the functions which their epigraph is a convex set. In the same manner, polyhedral (convex) functions are defined as those functions which their epigraph is a polyhedral set.

Definition 1.17 (Polyhedral function)**:** A function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be a polyhedral function if and only if $epi(h)$ is a polyhedral set.

Proposition 1.11. *For any polyhedral function h, it is true that h^* and h^∞ are polyhedral functions. Furthermore, dom h and dom h^* are polyhedral sets (and so their recession cones). $\partial h(\mathbf{x})$ is a polyhedral set for every $\mathbf{x} \in \mathbb{R}^n$.*

Previously, it was indirectly stated that convex sets, and in particular polyhedral sets, can be represented in two different but equivalent ways. The first is known as the H-representation, which is characterized by stating explicitly the parameters of every supporting hyperplane of the set. The second is the V-representation, which explicitly states the vertices and extreme directions of the polyhedron. From them, the set is reconstructed as the sum between the convex hull of the vertices and the conical hull of the extreme directions (Minkowski-Weyl theorem).

In some situations, working with H-representations of convex sets may be more convenient than V-representations, and *vice versa*. For example, H-representations are convenient for stating linear constraints in optimization problems. On the other hand, V-representations are of great use in the enumeration of solutions to optimization problems over polyhedra, where its vertices tend to be the only candidates for optimal solution.

Since polyhedral functions are defined through their epigraphs (which are polyhedral sets), then there are (at least) two representations for polyhedral functions as well, namely, their H and V-representations. In fact, one representation imply the over due to a duality relation between the representations of a polyhedral function h and its conjugate h^*: H-representations are dualized into V-representations and *vice versa*. The precise definition of these representations and the aforementioned duality result for the case of functions are stated below.

Definition 1.18 (Representations of convex functions)**:** Let the finite sets $\Pi \subseteq \mathbb{R}^n$, $C_\Pi = (c_\pi)_{\pi\in\Pi}$, $\Pi^* \subseteq \mathbb{R}^n$, and $C_{\Pi^*} = (c_{\pi^*})_{\pi^*\in\Pi^*}$. The polyhedral function h is said to be in an H-representation if and only if:

$$epi(h) = \left\{ (\mathbf{x}, z) \in \mathbb{R}^n \;\middle|\; \begin{array}{rr} {\pi^*}^\top \mathbf{x} - z \le c_{\pi^*}, & \forall \pi^* \in \Pi^* \\ \pi^\top \mathbf{x} \le c_\pi, & \forall \pi \in \Pi \end{array} \right\}, \tag{1.1}$$

Alternatively, h is in V-representation if and only if:

$$epi(h) = conv\{(\pi^*, c_{\pi^*})\}_{\Pi^*} + cone\{(\pi, c_\pi)\}_{\Pi} + epi(\delta_{\mathbf{0}})$$

These definitions are essentially geometrical, i.e., in terms of sets. The representations purely in terms of functional terms is shown below.

Proposition 1.12. *A function h is in H-representation if and only if there exists some sets $\Pi \subseteq \mathbb{R}^n$, $C_\Pi = (c_\pi)_{\pi\in\Pi}$, $\Pi^* \subseteq \mathbb{R}^n$, and $C_{\Pi^*} = (c_{\pi^*})_{\pi^*\in\Pi^*}$ such that*

$$h(\mathbf{x}) = \delta_{\mathcal{P}}(\mathbf{x}) + \sup_{\pi^*\in\Pi^*} {\pi^*}^\top \mathbf{x} - c_{\pi^*}, \tag{1.2}$$

where $\mathcal{P} := \{\mathbf{x} \in \mathbb{R}^n \mid \pi^\top \mathbf{x} \leq c_\Pi, \ \ \forall \pi \in \Pi\}$.

Proof. It is trivial to show that from (1.2), the epigraph of h matches exactly (1.1). □

Proposition 1.13 (Duality of representations)**.** *Assume h is a polyhedral function with the following H-representation:*

$$h(\mathbf{x}) = \delta_{\mathcal{P}}(\mathbf{x}) + \sup_{\pi^*\in\Pi^*} {\pi^*}^\top \mathbf{x} - c_{\pi^*},$$

where $\mathcal{P} := \{\mathbf{x} \in \mathbb{R}^n \mid \pi^\top \mathbf{x} \leq c_\pi, \ \ \forall \pi \in \Pi\}$. *Then the V-representation of h^* is naturally given by:*

$$epi(h^*) = conv\{(\pi^*, c_{\pi^*})\}_{\Pi^*} + cone\{(\pi, c_\pi)\}_{\Pi} + epi(\delta_{\mathbf{0}})$$

The same is true by interchanging the roles of h and h^*.

Proof. Using properties of the recession function, support function, and polarity, it is deduced that

$$\begin{aligned} h^\infty(\mathbf{d}) &= \delta_{\mathcal{P}^\infty}(\mathbf{d}) + \sigma_{conv(\Pi^*)}(\mathbf{d}), \\ &= \sigma_{\mathcal{P}^{\infty\circ}}(\mathbf{d}) + \sigma_{conv(\Pi^*)}(\mathbf{d}), \\ &= \sigma_{\mathcal{P}^{\infty\circ}+conv(\Pi^*)}(\mathbf{d}). \end{aligned} \tag{1.3}$$

Observe that, since $\mathcal{P}^\infty = \{\mathbf{d} \in \mathbb{R}^n \mid \pi^\top \mathbf{d} \leq 0,\ \forall \pi \in \Pi\}$, then $\mathcal{P}^{\infty^\circ} = cone(\Pi)$. Thus, by Proposition 1.9, equation (1.3) implies that $dom\ h^* = conv(\Pi^*) + cone(\Pi)$. Using that $h(\mathbf{x}) = \min\{\mu \mid (\mathbf{x}, \mu) \in epi(h)\}$, then:

$$\begin{aligned} h^*(\mathbf{x}^*) &= \sup_{\mathbf{x}} \mathbf{x}^\top \mathbf{x}^* - h(\mathbf{x}), \\ &= \sup_{\mathbf{x}} \mathbf{x}^\top \mathbf{x}^* - \min_{(\mathbf{x},\mu)\in epi(h)} \mu, \\ &= \sup_{(\mathbf{x},\mu)\in epi(h)} \mathbf{x}^\top \mathbf{x}^* - \mu. \end{aligned} \tag{1.4}$$

It should be noted that, for any $\mathbf{x}^* \in dom\ h^* = conv(\Pi^*) + cone(\Pi)$, $h^*(\mathbf{x}^*)$ is a finite value, therefore the formula in (1.4) defines a bounded linear program. Using strong duality, it is deduced that:

$$h^*(\mathbf{x}^*) = \min_{\substack{\mathbf{a},\mathbf{b}\geq \mathbf{0} \\ \mathbf{x}^* = \sum_i \mathbf{a}_i \pi_i^* + \sum_j \mathbf{b}_j \pi_j \\ \sum_i \mathbf{a}_i = 1}} \sum_i \mathbf{a}_i c_{\pi_i^*} + \sum_j \mathbf{b}_j c_{\pi_j},$$

which exactly the form for finitely generated polyhedral functions [7, see §19]. Finitely generated functions are exactly those in V-representation, therefore proving what it was desired.

The converse is trivial and it can be found in the proof of [7, Theorem 19.2].

□

Chapter 2

On the existence of POAs of convex functions

To give some intuition to the reader, it shall be briefly discussed the existence of POAs for the bounded case. Suppose f is a convex function with a bounded polyhedral domain. If f is a closed, proper, and convex function, its minimum always exists. This implies that the function $g(\mathbf{x}) := \delta_{dom\ f} + \min\ f$ is a POA of f. It is "polyhedral" because it is sum of polyhedral functions. By "outer" it is meant that $g(\mathbf{x}) \leq f(\mathbf{x})$ for all $\mathbf{x} \in dom\ f$. And by "approximation", that $\max_{\mathbf{x} \in dom\ f} f(\mathbf{x}) - g(\mathbf{x}) < \infty^+$. The latter condition is true because $dom\ f$ is bounded and $f - g$ is continuous in $dom\ f$. It should be noted that if $min\ f$ where not finite, then it would be necessary that $f \equiv -\infty$ in $dom\ f$. Therefore, the existence of POAs in the bounded case is equivalent to the finiteness of $\min\ f$. The latter is true for any non-trivial function f, thus the existence of POAs is basically always true.

In contrast, when $dom\ f$ is unbounded, the finiteness of $\min\ f$ is not sufficient, nor necessary, to show that $\|f - g\|_\infty$ is finite. The counter example for both conditions is the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x$, which is its own POA despite not having a finite minimum. Other counter examples includes the function $f(x) = x^2$ (sufficiency) and $f(x) = \sqrt{1 + x^2} - 2x$ (necessity; consider the POA $g(x) := \sup\{-3x, -x\}$).

Due to the above discussion, one concludes that the POA of convex functions with unbounded domains is an essentially different problem from the bounded case. In this chapter it is linked the existence of POAs to algebraic and functional properties of the function that is being approximated. The first step towards this goal is establishing a measure of the error of the approximation, for which the natural choice is the metric for continuous functions. It should be noted that since we are dealing with generalized convex functions, the error of approximation will always be infinite if the domains of the two functions does not coincide. Under this consideration, we define the ∞-norm as a measure of error and give some notation below.

Definition 2.1 (∞-norm)**:** Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be two convex functions with the same convex domain $\mathcal{D}$. The ∞-norm of $|f - g|$ is defined as follows:

$$\|f - g\|_\infty := \sup_{\mathbf{x} \in \mathcal{D}} |f(\mathbf{x}) - g(\mathbf{x})|.$$

Moreover, if $g \leq f$, then the ∞-norm of $|f - g|$ is referred as the ∞-norm of $f - g$ instead, or the error of approximation of g with respect to f.

From the above definition, an intriguing duality between the ∞-norm of $f - g$ and $g^* - f^*$ is proven below.

Lemma 2.1 (Duality of the ∞-norm)**.** *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be two closed, proper and convex functions, such that $g \leq f$. Then the following statements are equivalent:*

- *dom f = dom g and $\|f - g\|_\infty < \infty^+$.*
- *dom g^* = dom f^* and $\|g^* - f^*\|_\infty < \infty^+$.*

Moreover, if both statements are true, then the norms coincide.

Proof. Let us start by proving that *dom* g^* = *dom* f^* and $\|g^* - f^*\|_\infty < \infty^+$. Define $\varepsilon := \|f - g\|_\infty$. Since *dom* f = *dom* g, and using that $g \leq f$, then the finiteness of ε implies that $g \leq f \leq g + \varepsilon$. Conjugating once, it is deduced that $g^* \geq f^* \geq g^* - \varepsilon$. This implies that *dom* g^* = *dom* f^*. Thus, $\|g^* - f^*\|_\infty$ is well-defined and finite, because:

$$\begin{aligned}
\|g^* - f^*\|_\infty &:= \sup_{\mathbf{y} \in dom\ g^*} g^*(\mathbf{y}) - f^*(\mathbf{y}), \\
&\leq \sup_{\mathbf{y} \in dom\ g^*} (f^*(\mathbf{y}) + \varepsilon) - f^*(\mathbf{y}), \\
&= \varepsilon < \infty^+.
\end{aligned}$$

Conversely, define $\varepsilon^* := \|g^* - f^*\|_\infty$. Applying *verbatim* the same procedure with g^*, f^* and ε^* as it was done with f, g and ε, it is concluded that $\|f^{**} - g^{**}\|_\infty$ is finite, because it is not greater than ε^*. Furthermore, since both f and g are closed, proper, and convex functions, then $f = f^{**}$ and $g = g^{**}$. Thus, $\|f - g\|_\infty = \|f^{**} - g^{**}\|_\infty \leq \varepsilon^* < \infty^+$.

At last, assume that both statements are true. Then, the steps to prove the precedent implications are valid now, and thus provide a chain of inequalities that allows to conclude the desired result, i.e.:

$$\|f - g\|_\infty = \|f^{**} - g^{**}\|_\infty \leq \varepsilon^* = \|g^* - f^*\|_\infty \leq \varepsilon = \|f - g\|_\infty.$$

□

If the function g is polyhedral, a closed formula for $\|f - g\|_\infty$ in terms of the V-representation of g is available. Moreover, if said representation were unique, then so is the formula. Establishing a unique V-representation of g requires defining the notion of lineality space.

Definition 2.2 (Lineality space)**:** Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a non-empty closed convex set. The *lineality space*, *Lin* $\mathcal{C}$, is defined as the maximal subspace contained in the recession cone of $\mathcal{C}$, i.e., $Lin\ \mathcal{C} = \mathcal{C}^\infty \cap -\mathcal{C}^\infty$. Similarly, the *lineality space*, *Lin* f, of a closed convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is the subspace of vectors $\mathbf{d} \in \mathbb{R}^n$ such that f is linear in $\mathbf{d}$, i.e., $f(\mathbf{x} + \alpha\mathbf{d}) = f(\mathbf{x}) + \alpha f^\infty(\mathbf{d})$ for any $\mathbf{x} \in dom\ f$ and $\alpha \in \mathbb{R}$. The *lineality* of a lineality space is its dimension. If the lineality space is of dimension 0, then it is said that it has no lineality. The orthogonal to the lineality space is named the *alineality space*.

R.T. Rockafellar [7, §8] observed that any convex set may be decomposed as the direct sum of its lineality space and a closed convex set with no lineality, i.e., for any closed and convex set $\mathcal{C} \subseteq \mathbb{R}^n$, $\mathcal{C} = Lin\ \mathcal{C} \oplus \mathcal{C} \cap (Lin\ \mathcal{C})^\perp$. However, he did not consider that convex functions can be decomposed similarly. The domain of f can be orthogonally decomposed using *Lin* f instead of $Lin(dom\ f)$. This means that *dom* f now is partitioned considering where f is affine and where it is not. Doing so let one ignore the directions where f is affine. This is very useful, because it allows constructing a canonical V-representation of the POAs of f in terms of the alineality space, due to POAs having the same lineality spaces as their approximated function (this is based upon the fact that POAs of f and f itself have the same recession functions). The definition of the alinear component, and the relation between linear spaces of POAs are shown below.

Definition 2.3 (Alinear component)**:** The alinear component of f is defined as the function f constrained to the alineality space of $Lin\ f$:

$$f^{\perp}(\mathbf{x}) := f(\mathbf{x}) + \delta_{(Lin\ f)^{\perp}}(\mathbf{x}).$$

Proposition 2.1. *Let $g : \mathbb{R}^n \to \mathbb{R}$ be a closed, proper and convex function such that $g^{\infty} \equiv f^{\infty}$. Then $Lin\ f = Lin\ g$.*

Proof. See [7, Theorem 8.8]. Here, it is proven the characterization that $Lin\ f = \{\mathbf{d} \in \mathbb{R}^n \mid -f^{\infty}(-\mathbf{d}) = f^{\infty}(\mathbf{d})\}$. Thus, it is direct that $Lin\ f = Lin\ g$ whenever $f^{\infty} \equiv g^{\infty}$. □

The above decomposition of the domain uniquely determines a closed and finite formula for $\|f - g\|_{\infty}$ in terms of the V-representation of g. This is stated below.

Proposition 2.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed, proper, and convex function, and g be a polyhedral function. If $dom\ f = dom\ g$, $g \leq f$, and $\|f - g\|_{\infty} < \infty^{+}$, then it is true that $Lin\ g = Lin\ f$, and the following formula holds:*

$$\|f - g\|_{\infty} = \|f^{\perp} - g^{\perp}\|_{\infty} = \sup_{\pi^* \in \Pi^*} f(\pi^*) - g(\pi^*), \tag{2.1}$$

where $\Pi^* := \{\pi^* \in \mathbb{R}^n \mid (\pi^*, g(\pi^*)) \in ext(epi(g^{\perp}))\}$.

Proof. To start, it shall be proven that the equality $\|f - g\|_{\infty} = \|f^{\perp} - g^{\perp}\|_{\infty}$ holds. Since $dom\ f = dom\ g$ and $\|f - g\|_{\infty} < \infty^{+}$, it follows that $f^{\infty} \equiv g^{\infty}$ (Lemma 2.1). Therefore, $Lin\ f = Lin\ g$ (Proposition 2.1). Also, due to the fact that $dom\ f$ can be decomposed as $dom\ f \cap (Lin\ f)^{\perp} + Lin\ f$, observe that:

$$\begin{aligned}
\|f-g\|_\infty &:= \sup_{\mathbf{x}\in dom\ f} f(\mathbf{x})-g(\mathbf{x}),\\
&= \sup_{\substack{\mathbf{x}\in dom\ f\cap(Lin\ f)^\perp\\ \mathbf{d}\in Lin\ f}} f(\mathbf{x}+\mathbf{d})-g(\mathbf{x}+\mathbf{d}),\\
&= \sup_{\substack{\mathbf{x}\in dom\ f\cap(Lin\ f)^\perp\\ \mathbf{d}\in Lin\ f}} (f^\perp(\mathbf{x})+f^\infty(\mathbf{d}))-(g^\perp(\mathbf{x})+g^\infty(\mathbf{d})),\\
&= \sup_{\mathbf{x}\in dom\ f\cap(Lin\ f)^\perp} f^\perp(\mathbf{x})-g^\perp(\mathbf{x}),\\
&=: \|f^\perp-g^\perp\|_\infty.
\end{aligned}$$

It follows to show the formula regarding Π^*. Since $\|f^\perp-g^\perp\|_\infty=\|f-g\|_\infty$, then the former is finite, and by Lemma 2.1 it is also equal to $\|(g^\perp)^*-(f^\perp)^*\|_\infty$, with $dom\ (g^\perp)^*=dom\ (f^\perp)^*$. By definition, $g^\perp$ has no lineality. Therefore, Π^* is well-defined, because the lack of lineality of $g^\perp$ implies that $epi(g^\perp)$ possesses at least one vertex. Recall that, under these circumstances, Π^* represents the set of all the non-vertical slopes of $g^{\perp*}$. In consequence, the desired formula is deduced:

$$\begin{aligned}
\|f^\perp-g^\perp\|_\infty &= \|g^{\perp*}-f^{\perp^*}\|_\infty,\\
&= \sup_{\mathbf{x}^*\in dom\ f^{\perp*}} g^{\perp*}(\mathbf{x}^*)-f^{\perp*}(\mathbf{x}^*),\\
&= \sup_{\mathbf{x}^*\in dom\ f^{\perp*}}\ \sup_{\pi^*\in\Pi^*} {\pi^*}^\top\mathbf{x}^*-g^\perp(\pi^*)-f^{\perp*}(\mathbf{x}^*),\\
&= \sup_{\pi^*\in\Pi^*}\ \sup_{\mathbf{x}^*\in dom\ f^{\perp*}} {\pi^*}^\top\mathbf{x}^*-g^\perp(\pi^*)-f^{\perp*}(\mathbf{x}^*),\\
&= \sup_{\pi^*\in\Pi^*} f^\perp(\pi^*)-g^\perp(\pi^*),\\
&= \sup_{\pi^*\in\Pi^*} f(\pi^*)-g(\pi^*).
\end{aligned}$$

□

Corolary 2.1. *If dom f = dom g, $g\le f$, and $f^\infty\equiv g^\infty$, then $\|f-g\|_\infty=\|f^\perp-g^\perp\|_\infty$, even if they are not finite.*

Proof. The proof of the equality has been already established in the proof of Proposition 2.2. Such equality was proven only assuming that $f^\infty\equiv g^\infty$. □

The results stated before are preliminary to show which convex functions admits POAs. This characterization is proven below, and represents the main result of this thesis.

Theorem 2.1 (Existence of POA of convex functions)**.** *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed, proper, and convex function, and g a polyhedral function such that dom f = dom g and satisfies the inequality $g \leq f$. Consider the statements:*

(i) $\|f - g\|_\infty < \infty^+$.

(ii) *dom* $f^* =$ *dom* g^*.

(iii) $f^\infty \equiv g^\infty$ *and dom* f^* *is closed.*

Then (i)—(ii)—(iii) are equivalent.

Proof. The proof of $(i) \Rightarrow (ii)$ is a direct implication of Lemma 2.1. To show that $(ii) \Rightarrow (i)$, recall that for closed proper convex functions it is true that $f^\infty \equiv \sigma_{dom\ f^*}$. Thus, $Lin\ f = Lin\ g$, since $f^\infty \equiv g^\infty$. By Corollary 2.1, it holds that $\|f - g\|_\infty = \|f^\perp - g^\perp\|_\infty$. It remains to show that $\|g^{\perp *} - f^{\perp *}\|_\infty$ is finite. Since $g^\perp$ is a polyhedral function with no lineality, then there exists two finite sets $\Pi, \Pi^* \subset \mathbb{R}^n$ such that the V-representation of $g^\perp$ is given by $epi(g^\perp) = conv(\{(\pi^*, g^\perp(\pi^*))\}_{\pi^* \in \Pi^*}) + cone(\{\pi, g^{\perp\infty}(\pi)\}_{\pi \in \Pi}) + epi(\delta_{\mathbf{0}})$. Thus, $g^{\perp *}$ is a polyhedral function with an H-representation given by $g^{\perp *}(\mathbf{x}^*) = \delta_{\mathcal{P}^*}(\mathbf{x}) + \sup_{\pi^* \in \Pi^*} \pi^{*\top}\mathbf{x} - g(\pi^*)$, where $\mathcal{P}^* := \{\mathbf{x} \in \mathbb{R}^n \mid \pi^\top \mathbf{x} \leq g^{\perp\infty}(\pi),\ \ \forall \pi \in \Pi\}$. Observe that $dom\ g^{\perp *} = dom\ f^{\perp *}$, since the former is equal to $dom\ g^* + Lin\ g$. As $f^{\perp *} \leq g^{\perp *}$, note that:

$$
\begin{aligned}
\|g^{\perp *} - f^{\perp *}\|_\infty &= \sup_{\mathbf{x}^* \in dom\ f^{\perp *}} g^{\perp *}(\mathbf{x}^*) - f^{\perp *}(\mathbf{x}^*), \\
&= \sup_{\substack{\mathbf{x}^* \in dom\ f^{\perp *} \\ \pi^* \in \Pi^*}} \pi^{*\top}\mathbf{x}^* - g^\perp(\pi^*) - f^{\perp *}(\mathbf{x}^*), \\
&= \sup_{\substack{\mathbf{x}^* \in dom\ f^{\perp *} \\ \pi^* \in \Pi^*}} (\mathbf{x}^{*\top}\pi^* - f^{\perp *}(\mathbf{x}^*)) - g^\perp(\pi^*), \\
&= \sup_{\pi^* \in \Pi^*} f^\perp(\pi^*) - g^\perp(\pi^*).
\end{aligned}
$$

Thus, it follows that $\|g^{\perp *} - f^{\perp *}\|_\infty \equiv \sup_{\pi^* \in \Pi^*} f^\perp(\pi^*) - g^\perp(\pi^*)$ is finite, since it is the supremum of finite values, and therefore equal to $\|f^\perp - g^\perp\|_\infty$ and $\|f - g\|_\infty$ (see Lemma 2.1 and Proposition 2.2).

Regarding $(ii) \Rightarrow (iii)$, observe that $f^\infty = \sigma_{dom\ f^*} = \sigma_{dom\ g^*} = g^\infty$. Also, $dom\ f^*$ is closed, since $dom\ g^*$ is polyhedral, and $dom\ g^* = dom\ f^*$. The proof of $(iii) \Rightarrow (ii)$ follows from the equality $f^\infty \equiv g^\infty$, which is equivalent to $\sigma_{dom\ f^*} \equiv \sigma_{dom\ g^*}$. This implies that $\overline{conv(dom\ f^*)} = \overline{conv(dom\ g^*)}$. In this case, both $dom\ f^*$ and $dom\ g^*$ are closed convex sets, thus $dom\ g^* = dom\ f^*$, as desired. □

In Theorem 2.1 two characterizations are given for the approximation of a convex function f by g. The first characterization, statement (ii), should be viewed as an implicit characterization, because $dom\ f^*$ is not immediately available just from the definition of f. On the other hand, the second characterization, statement (iii), should be viewed as an explicit characterization, because verifying these conditions becomes rather algorithmic. First is verifying whether f^∞ is polyhedral or not. In the affirmative case, it would imply that the closure of $dom\ f^*$ is polyhedral. Second is verifying if f^* is finite at the vertices of its domain, which are at most finite. Checking the vertices are sufficient, due to convexity of the domain. The necessity of verifying the closedness of $dom\ f^*$ is illustrated in the following counter-example.

Example 2.1: Consider the function $f(x) := \delta_{[0,\infty)}(x) - \sqrt{x}$. It is a closed, proper, and convex function, since both $\delta_{[0,\infty)}(x)$ and $-\sqrt{x}$ are too. Moreover, observe that $f^\infty(x) \equiv \delta_{[0,\infty)}(x)$ is a polyhedral function. Suppose that g is an outer-approximation of f. This implies that $g^\infty \equiv f^\infty \equiv \delta_{[0,\infty)}(x)$. Because g is polyhedral, for a sufficiently large $a > 0$, it must be that $g(x) = -c$, for some constant $c \in \mathbb{R}$, for any $x \geq a$. Otherwise, g^∞ would have positive or negative values to the right. If g is assumed to be an outer-approximation of f, then it would be that $-c = g(x) \leq f(x)$ and any $x > a$. The contradiction comes after the fact that f diverges to $-\infty$ to the right. This counter-example is due to $dom\ f^*$ not being closed. In fact, it is easy to show that $f^*(x^*) = \delta_{(-\infty,0]}(x^*) - \frac{1}{4x^*}$. This function diverges at $x^* = 0$, therefore its domain is the open set $(-\infty, 0)$. The graphs of f and f^* are illustrated in Figure 2.1.

It is proven in Theorem 2.1 that the convex functions f that can be ap-

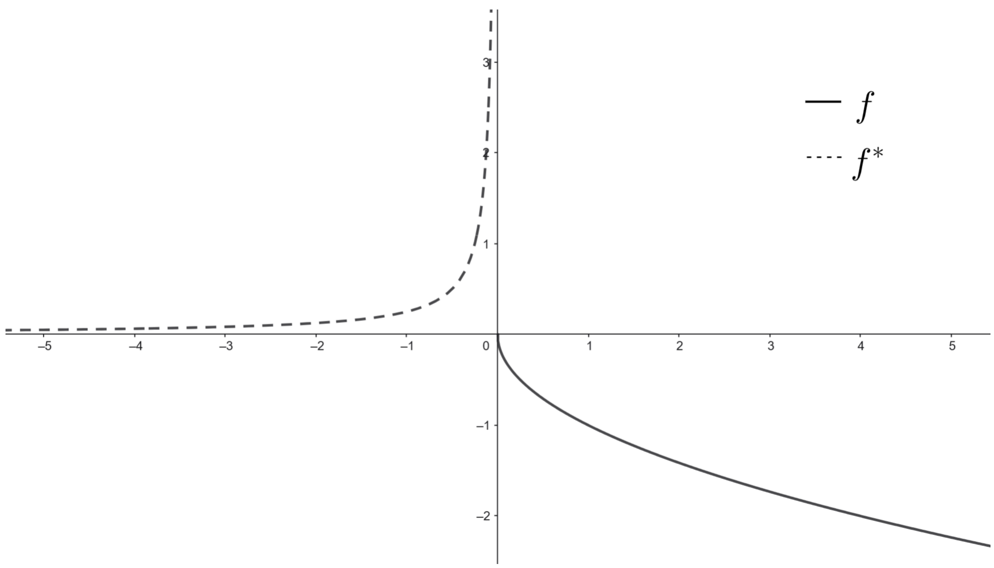


Figure 2.1: The graph of the functions f and f^*.

proximated by polyhedral functions g are those where *dom* f and *dom* f^* are polyhedral. However, the theorem only let us check whether a polyhedral function does approximate f or not, i.e., it is not constructive. However, this can be addressed in a more constructive manner if the H-representation of g is known. Among all the polyhedral functions g, the best in terms of approximation error are the ones constructed by using supporting hyperplanes to *epi*(f). This class of functions are defined below.

Definition 2.4: Let $\mathcal{P}$ be a polyhedron, and $f : \mathcal{P} \to \overline{\mathbb{R}}$ be a closed, proper, and convex function, and $\Gamma \subset$ *dom* f^* be a finite set. The POA of f with respect to Γ, given by its H-representation, is defined as follows:

$$f_\Gamma(\mathbf{x}) := \delta_{\mathcal{P}}(\mathbf{x}) + \sup_{\gamma \in \Gamma} \gamma^\top \mathbf{x} - f^*(\gamma).$$

Clearly, f_Γ is a polyhedral function, since it is the sum of polyhedral functions. Moreover, it is “outer” with respect to f, in the sense that $f_\Gamma \leq f$. This is a consequence of the Fenchel Inequality, which implies that $\gamma^\top \mathbf{x} - f^*(\gamma) \leq f(\mathbf{x})$, $\forall \gamma \in$ *dom* f^*, $\mathbf{x} \in$ *dom* f. Conditions over Γ in order that f_Γ approximates f are stated below.

Lemma 2.2. *Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a closed, proper, and convex function such that dom f^∞ is a full dimensional polyhedron and dom f^* is polyhedral. Then ext(dom f^*) is non-empty.*

Proof. Using [7, Theorem 13.4], and the fact that $Lin\ \mathcal{C} = Lin\ \delta_{\mathcal{C}}$, it is deduced the desired condition:

$$\begin{aligned} Lineality\ dom\ f^* &= Lineality\ \delta_{dom\ f^*}, \\ &= n - dim(\sigma_{dom\ f^*}), \\ &= n - dim(f^\infty), \\ &= 0. \end{aligned}$$

Because *dom* f^* is a polyhedral set with no lineality, then it must possess at least one vertex. □

Proposition 2.3 (Characterization of POAs)**.** *Let $\mathcal{P} \subseteq \mathbb{R}^n$ be a polyhedron such that $\mathcal{P}^\infty$ is full-dimensional. Assume that $f : \mathcal{P} \to \mathbb{R}$ is a closed, proper, and convex function such that dom$(f^\infty) = \mathcal{P}^\infty$ and dom f^* is polyhedral. Define $\Pi :=$ ext(dom f^*), and let $\Gamma \subset$ dom f^* be a finite set. Then, the following equivalence holds:*

$$\|f - f_\Gamma\|_\infty < \infty^+ \iff \Pi \subseteq \Gamma.$$

Proof. Before proving both implications, two statements should be remarked. First, *ext*(*dom* f^*) is not empty, due to Lemma 2.2. Thus, Π is well-defined. Second, it should be noted that Π is the set of non-vertical supporting hyperplanes of f^∞. To see that, first recall that *dom* f^* has a recession cone polar to the recession cone of *dom* f. Thus, *dom* $f^* = conv(\Pi) + \mathcal{P}^{\infty\circ}$. In consequence, $f^\infty \equiv \sigma_{dom\ f^*} \equiv \sigma_{\mathcal{P}^{\infty\circ}+\Pi} \equiv \delta_{\mathcal{P}} + \sigma_\Pi$.

After these remarks, it follows to prove the first implication. Assuming that $\|f - f_\Gamma\|_\infty < \infty^+$, then using Theorem 2.1 it must be that $conv(\Pi) + \mathcal{P}^{\infty\circ} = conv(\Gamma) + \mathcal{P}^{\infty\circ}$. Observe that the extreme points of a set that is sum of two sets is the sum of the extreme points of each set, whenever the sets in the sum possess vertices at all. Indeed, in this case $\mathcal{P}^{\infty\circ}$ has the $\mathbf{0}$ vector as it only vertex, since *dom* f^* has no lineality. Consequently, it is deduced that $ext(conv(\Pi)) = ext(conv(\Gamma))$. Note that the former is just equal to Π, since, by definition, it is the set of extreme points of a set, so the extreme points of its convex hull are the same points. The latter is contained in Γ by definition. Thus, $\Pi \subseteq \Gamma$.

Conversely, using the definition of supremum it is easy to show that $f_\Pi \leq f_\Gamma$. This, and the fact that $f_\Gamma \leq f$, implies that $\|f - f_\Gamma\|_\infty \leq \|f - f_\Pi\|_\infty$. Furthermore, it is true that the latter must be finite due to the following. From one side, it is true that $f^\infty \equiv \delta_{\mathcal{P}^\infty} + \sigma_\Pi$. From another side, it is easily computed that $f_\Pi^\infty \equiv \delta_{\mathcal{P}^\infty} + \sigma_\Pi$ as well, since the recession operator nulls out any constant. Therefore, by Theorem 2.1, $\|f - f_\Pi\|_\infty < \infty^+$. □

The assumption that *dom* f^∞ is full-dimensional allows proving the necessity of the finite inclusion $\Pi \subseteq \Gamma$. This condition is usually met. Otherwise, only the characterization in Theorem 2.1 remains valid.

In general, Theorem 2.1 allows determining which convex functions possess POAs. If such approximations exist, Proposition 2.3 shows that, under extra assumptions, a "minimal" (with respect to the set inclusion) approximation exists, namely, f_Π. This result can be applied to compute this minimal approximation for whole classes of functions. One of those classes is addressed below.

Corolary 2.2. *Let $\mathcal{K} \subset \mathbb{R}^n$ be a full-dimensional polyhedral pointed cone, $\mathbf{p} \in \mathbb{R}^n$ and $f : \mathcal{K} + \mathbf{p} \to \mathbb{R}$ be a closed, proper, and convex function. Assume that $\min_{\mathcal{K}+\mathbf{p}} f$ is finite, and f is non-increasing with respect to $\mathcal{K}$, i.e., for any $\mathbf{x}, \mathbf{y} \in \mathcal{K}+\mathbf{p}$ such that $\mathbf{x} - \mathbf{y} \in \mathcal{K}$, then $f(\mathbf{x}) \leq f(\mathbf{y})$ holds. Consequently, it is true that dom f^* is closed, $\Pi := ext(dom\ f^*) = \{\mathbf{0}\}$, $f_\Pi \equiv \delta_{\mathcal{P}} + min\ f$, $f_\Pi \leq f$ and $\|f - f_\Pi\|_\infty < \infty^+$.*

Proof. First, it shall be proven that f^∞ is exactly $\delta_{\mathcal{K}}$. This will be done by bounding the limit form of f^∞. Consider any $\mathbf{x} \in \mathcal{K}$ and $\mathbf{d} \in \mathbb{R}^n$. Then, due to f having a minimum, it follows that:

$$\begin{aligned} f^\infty(\mathbf{d}) &= \lim_{\lambda \to \infty^+} \frac{f(\mathbf{x} + \mathbf{p} + \lambda \mathbf{d}) - f(\mathbf{x})}{\lambda}, \\ &\geq \lim_{\lambda \to \infty^+} \frac{\min\ f + \delta_{\mathcal{K}+\mathbf{p}}(\mathbf{x} + \mathbf{p} + \lambda \mathbf{d}) - f(\mathbf{x})}{\lambda}, \\ &= \delta_{\mathcal{K}}(\mathbf{d}). \end{aligned}$$

To show the converse inequality, consider additionally a constant $\lambda > 0$. Since f is non-increasing with respect to $\mathcal{K}$, it must hold that

$$f(\mathbf{x} + \mathbf{p} + \lambda \mathbf{d}) - f(\mathbf{x} + \mathbf{p}) \leq \delta_{\mathcal{K}}(\lambda \mathbf{d}).$$

Dividing both sides of the inequality by λ, and taking $\lambda \to \infty^+$, it is deduced that $f^\infty(\mathbf{d}) \leq \delta_{\mathcal{K}}(\mathbf{d})$. Therefore, $f^\infty(\mathbf{d}) \equiv \delta_{\mathcal{K}}(\mathbf{d})$. This implies that $\overline{dom\ f^*} =$

$\mathcal{K}^{\circ}$, since $\delta_{\mathcal{K}} \equiv \sigma_{\mathcal{K}^{\circ}}$. Moreover, by assumption $min\ f = -f^{*}(\mathbf{0})$ is finite, thereby implying that $\mathbf{0} \in dom\ f^{*}$. Since $\mathcal{K}$ is polyhedral, so is $\mathcal{K}^{\circ}$ and $dom\ f^{*}$. So it must be that $dom\ f^{*}$ is closed.

Since all conditions of Proposition 2.3 are met, it follows that the polyhedral function $f_{\Pi} \equiv \delta_{\mathcal{K}+\mathbf{p}} + \min_{\mathcal{K}+\mathbf{p}} f$ is in fact the minimal POA of f. □

This result allows, for example, to approximate the exponential function defined at the preimage of $(0, 1]$.

Example 2.2: Let $f(x) = e^{x} + \delta_{(-\infty,0]}(x)$. In this case, $\mathcal{K} = (-\infty, 0]$ is full-dimensional polyhedral pointed cone, and $p = 0$. The function $f(x)$ is non-decreasing, since e^{x} and $\delta_{(-\infty,0]}(x)$ are non-decreasing. This implies that for any $x, y \in (-\infty, 0]$ such that $x \leq y$, then $f(x) \leq f(y)$. The condition that $x \leq y$ is equivalent to the inclusion $x - y \in \mathcal{K}$. Thus, f is non-increasing with respect to $\mathcal{K}$. Lastly, observe that $min\ f = 0$. Therefore, using Corollary 2.2, it is concluded that the minimal approximation of f is $f_{\Pi}(x) = \delta_{(-\infty,0]}(x)$, where $\Pi = \{0\}$. Moreover, $f - f_{\Pi}$ is increasing, thus its maximum is 1, attained at $x = 0$.

Chapter 3

The DOA algorithm

In this chapter, the DOA algorithm is defined. The purpose of the DOA algorithm is to receive as input the set Γ and return the set $\overline{\Gamma}$. The output is an augmentation of the input, i.e., $\overline{\Gamma} \supseteq \Gamma$. It is the result of iteratively adding valid subgradients of f to the set Γ. This is done until an arbitrarily small error $\varepsilon > 0$ is achieved. More precisely, until the inequality $\|f - f_{\overline{\Gamma}}\|_\infty \leq \varepsilon < \|f - f_\Gamma\|_\infty$ holds.

Let f be any function satisfying the hypothesis of Proposition 2.3. Let $\varepsilon > 0$ and Γ such that $\Pi \subseteq \Gamma \subseteq \mathit{dom}\ f^*$. Equation (2.1) provides a finite set of candidates Γ^*, for which at least one of them must attain the maximum of $f - f_\Gamma$. Let $\chi^* := \arg\max_{\gamma^* \in \Gamma^*} f(\gamma^*) - f_\Gamma(\gamma^*)$ be the set of maxima in Γ^*. For every point $\mathbf{x} \in \chi^*$, compute a subgradient as follows. If $\mathbf{x} \in \mathit{dom}\ \partial f$, compute any subgradient $\mathbf{x}^*$ of $\partial f^*(\mathbf{x})$. Otherwise, compute $\overline{\mathbf{x}}^* = -\frac{\mathbf{x}-\overline{\mathbf{x}}}{f_\Gamma(\mathbf{x})-f(\overline{\mathbf{x}})}$, where $\overline{\mathbf{x}}$ using the projection of $(\mathbf{x}, f_\Gamma(\mathbf{x}))$ onto $\mathit{epi}(f)$. Augment Γ into $\Gamma \cup \chi$, where χ is defined as the union of the subgradients computed for each $\mathbf{x} \in \chi^*$. Compute the error $\|f - f_{\Gamma\cup\chi}\|_\infty$, and if it is not less than ε, then repeat all the steps in this paragraph, using $\Gamma \cup \chi$ instead of Γ. The DOA algorithm is summarized in Algorithm 1.

To show the correctness of Algorithm 1 is the same as to show that the augmenting of Γ into $\Gamma \cup \chi$ does improve the error of approximation, i.e., $\|f - f_{\Gamma\cup\chi}\|_\infty < \|f - f_\Gamma\|_\infty$. In order to show this, first it is proven the following lemma regarding various steps of the algorithm.

Lemma 3.1 (Refination of the POAs)**.** *Assume the same hypothesis of Proposition 2.3. Choose any* $\mathbf{x} \in \chi^* := \arg\max_{\gamma^* \in \Gamma^*} f(\gamma^*) - f_\Gamma(\gamma^*)$ *and let* $(\overline{\mathbf{x}}, \overline{z}) := \mathit{Proj}_{\mathit{epi}(f)}(\mathbf{x}, f_\Gamma(\mathbf{x}))$*. Then, the following statements hold true:*

Algorithm 1 Dual outer-approximation algorithm

$\Gamma = \Pi$
$\varepsilon = \infty^+$
while $\varepsilon > \overline{\varepsilon}$ **do**
 $\Gamma^* = ext(epi(f_\Gamma^\perp))$
 $\chi^* = \arg\max_{\mathbf{y}\in\Gamma^*} f^\perp(\mathbf{y}) - f_\Gamma^\perp(\mathbf{y})$
 $\varepsilon = \max_{\mathbf{y}\in\Gamma^*} f^\perp(\mathbf{y}) - f_\Gamma^\perp(\mathbf{y})$
 $\chi = \emptyset$
 for $\mathbf{x} \in \chi^*$ **do**
  if $\partial f(\mathbf{x}^*) \neq \emptyset$ **then**
   $\mathbf{x}^* \leftarrow Choose(\partial f(\mathbf{x}^*))$.
  else
   $\overline{\mathbf{x}} \leftarrow \arg\min_{\mathbf{y}\in\mathcal{P}} \|(\mathbf{y}, f(\mathbf{y})) - (\mathbf{x}, f_\Gamma(\mathbf{x}))\|_2^2$
   $\overline{\mathbf{x}}^* = -\frac{\mathbf{x}^* - \overline{\mathbf{x}}^*}{f_\Gamma(\mathbf{x}^*) - f(\overline{\mathbf{x}}^*)}$
  end if
  $\chi \leftarrow \chi \cup \{\overline{\mathbf{x}}^*\}$
 end for
 if $\varepsilon > \overline{\varepsilon}$ **then**
  $\Gamma \leftarrow \Gamma \cup \chi$
 end if
end while
return Γ

(1) $\overline{z} = f(\overline{\mathbf{x}})$, *and* $\overline{\mathbf{x}} \in \arg\min_{\mathbf{y}\in\mathcal{P}} \|(\mathbf{y}, f(\mathbf{y})) - (\mathbf{x}, f(\mathbf{x}))\|_2^2$.

(2) $\overline{\mathbf{x}}^* := -\frac{\mathbf{x}-\overline{\mathbf{x}}}{f_\Gamma(\mathbf{x})-f(\overline{\mathbf{x}})} \in \partial f(\overline{\mathbf{x}})$, *and* $f(\mathbf{x}) - f_{\Gamma\cup\{\overline{\mathbf{x}}^*\}}(\mathbf{x}) < \|f - f_\Gamma\|_\infty$.

(3) If $\mathbf{x} \in$ *dom* ∂f*, then* $f(\mathbf{x}) - f_{\Gamma\cup\{\mathbf{x}^*\}}(\mathbf{x}) = 0$ *for any* $\mathbf{x}^* \in \partial f(\mathbf{x})$.

Proof.

Regarding (1), observe that $(\overline{\mathbf{x}}, \overline{z}) := Proj_{epi(f)}(\mathbf{x}, f_\Gamma(\mathbf{x}))$ is well-defined because $epi(f)$ is a non-empty closed and convex set. By definition, $\overline{\mathbf{x}} \in \mathcal{P}$. To show that $\overline{z} = f(\overline{\mathbf{x}})$, recall that $(\overline{\mathbf{x}}, \overline{z})$ is the projection onto $epi(f)$ of $(\mathbf{x}, f_\Gamma(\mathbf{x}))$ if and only if $(\mathbf{x} - \overline{\mathbf{x}}, f_\Gamma(\mathbf{x}) - \overline{z}) \in \mathcal{N}_{epi(f)}(\overline{\mathbf{x}}, \overline{z})$. Since $(\mathbf{x}, f_\Gamma(\mathbf{x}))$ is not an element of $epi(f)$, then this normal vector is non-zero. Moreover, $f_\Gamma(\mathbf{x}) < \overline{z}$ because f is convex, so the last component of any normal vector to the epigraph must have a negative last component. This implies that $|f_\Gamma(\mathbf{x}) - z|^2$, is decreasing with respect to z, at $z = \overline{z}$. With this, and using the definition of the projection as a minimization problem, observe that:

$$\begin{aligned}(\overline{\mathbf{x}}, \overline{z}) &:= \arg\min_{\substack{\mathbf{y}\in\mathcal{P}\\ z\geq f(\mathbf{y})}} \|\mathbf{x} - \mathbf{y}\|_2^2 + |f_\Gamma(\mathbf{x}) - z|^2,\\ &= \arg\min_{\substack{\mathbf{y}\in\mathcal{P}\\ z= f(\mathbf{y})}} \|\mathbf{x} - \mathbf{y}\|_2^2 + |f_\Gamma(\mathbf{x}) - f(\mathbf{y})|^2.\end{aligned}$$

Thus, $\overline{z} = f(\overline{\mathbf{x}})$. Moreover, $\overline{\mathbf{x}}$ is independent of $\overline{z}$, therefore it is just equal to $\arg\min_{\mathbf{y}\in\mathcal{P}} \|(\mathbf{y}, f(\mathbf{y})) - (\mathbf{x}, f(\mathbf{x}))\|_2^2$.

With respect to (2), the vector $\overline{\mathbf{x}}^* := -\frac{\mathbf{x}-\overline{\mathbf{x}}}{f_\Gamma(\mathbf{x})-f(\overline{\mathbf{x}})}$ must be an element of $\partial f(\overline{\mathbf{x}})$, due to Proposition 1.6.

To show the strict inequality of (2), observe that the non-zero vector $(\overline{\mathbf{x}}^*, -1)$ is parallel to $(\mathbf{x} - \overline{\mathbf{x}}, f_\Gamma(\mathbf{x}) - f(\overline{\mathbf{x}}))$. Thus:

$$\begin{aligned}0 &< (\mathbf{x} - \overline{\mathbf{x}}, f_\Gamma(\mathbf{x}) - f(\overline{\mathbf{x}}))^\top (\overline{\mathbf{x}}^*, -1),\\ &= \mathbf{x}^\top \overline{\mathbf{x}}^* - \overline{\mathbf{x}}^\top \overline{\mathbf{x}}^* + f(\overline{\mathbf{x}}) - f_\Gamma(\mathbf{x}),\\ &= \mathbf{x}^\top \overline{\mathbf{x}}^* - f^*(\overline{\mathbf{x}}^*) - f(\mathbf{x}) + \|f - f_\Gamma\|_\infty,\\ &\leq f_{\Gamma\cup\{\overline{\mathbf{x}}^*\}}(\mathbf{x}) - f(\mathbf{x}) + \|f - f_\Gamma\|_\infty.\end{aligned}$$

To show (3), let $\mathbf{x} \in$ *dom* ∂f and choose $\mathbf{x}^* \in \partial f(\mathbf{x})$. By the Fenchel Equality, $f(\mathbf{x}) = \mathbf{x}^{*\top}\mathbf{x} - f^*(\mathbf{x}^*)$. Moreover, it holds that $f(\mathbf{x}) = f_{\Gamma\cup\{\mathbf{x}^*\}}(\mathbf{x})$, since $f \geq f_{\Gamma\cup\{\mathbf{x}^*\}} \geq f_{\{\mathbf{x}^*\}}$. □

Theorem 3.1. *Under the assumptions of Lemma 3.1, and assuming that $f^\perp$ is strictly convex on dom $f^\perp$, then* $\arg\max_{dom\ f^\perp} f - f_\Gamma \subseteq \chi^*$.

Proof. Suppose by contradiction that there exists some $\mathbf{x} \in \arg\max_{dom\ f^\perp} f - f_\Gamma$ such that $\mathbf{x} \notin \chi^*$. Thus, there exists a segment $[\boldsymbol{a}, \boldsymbol{b}] \subseteq dom\ f$ that strictly contains $\mathbf{x}$ such that $f_\Gamma^\perp|_{[\boldsymbol{a},\boldsymbol{b}]}$ is affine. Therefore, by construction $(f^\perp - f_\Gamma^\perp)|_{[\boldsymbol{a},\boldsymbol{b}]}$ is a strictly convex function. The contradiction comes from [7, Theorem 32.1], which implies that $(f^\perp - f_\Gamma^\perp)|_{[\boldsymbol{a},\boldsymbol{b}]}$ must also be constant, since it attains its maximum at $\mathbf{x} \in (\boldsymbol{a}, \boldsymbol{b})$. □

The main consequence of Proposition 2.3 is that under the strict convexity of $f^\perp$, the equality $\chi^* = \arg\max_{\mathbf{x}\in dom\ f^\perp} f(\mathbf{x}) - f_\Gamma(\mathbf{x})$ holds. This equality means that the DOA algorithm, in this case, decreases the error strictly at each iteration, thus proving it is correct.

Chapter 4

Computational study

In this chapter, numerical results are presented. The main objectives of the computational study are to measure the performance of the DOA algorithm and verify its convergence.

The algorithm was applied to 6 test functions, which are defined in the Appendix. The parameters to correctly apply the algorithm per function are: the function f; the H-representation of $\mathcal{P} := dom\ f$ and $Lin(f)$; the set $\Pi := ext(dom\ f^*)$; the values $f^*(\pi)$, for every $\pi \in \Pi$; a function $Choose(\partial f(\mathbf{x}))$, that for any $\mathbf{x} \in dom\ \partial f$, it chooses any element of $\partial f(\mathbf{x})$, and otherwise returns the normal exterior to $epi(f)$, $\overline{\mathbf{x}}^*$ (see point 2 of Lemma 3.1). A quick summary of the functions is as follows. Test function 1 is the exponential function; 2.1 and 2.∞ have hyperboloids as epigraph; 3.$\mathcal{A}$ and 3.$\mathcal{B}$ are the fill-rate and back-order functions from the theory of inventory control; 4 is a non-trivial function with non-trivial lineality.

In order to gauge the performance of the DOA algorithm, for each test function it is measured the total CPU time and the total amount of supporting hyperplanes ($|\Gamma|$) at termination. The termination criteria are the following. $\|f - f_\Gamma\|_\infty \leq \varepsilon := 10^{-4}$, total CPU time of 7200 seconds, or if the vertex enumeration algorithm breaks due to numerical inaccuracies. It is considered that the algorithm converged if the termination is due to the error criteria. Each test function was approximated once, since the DOA algorithm is deterministic.

The algorithm was coded in Python 3.12.2. The main non-conventional library used was *pcdd*, which enumerates the vertices of polyhedra. All the tests were conducted using a personal computer of 64 bits with Windows 11, an AMD processor Ryzen 5 240 (4.30 GHz) of 6 nuclei and 12 threads, and

16 Gb of RAM.

Error dynamic

The error at the termination of the algorithm, among other measures such as the total amount of hyperplanes $|\Gamma|$, the total CPU Time, and the termination criteria are registered in Table 1. Every test function terminated by the error criteria, except for function 2.1, which terminated by numerical inaccuracies. The error achieved by every function is under the predefined value of $\varepsilon = 10^{-4}$, except for function 2.1. However, its error was close enough. The total CPU time spent shows that either the algorithm converged in no more than 3 seconds, or it took over 100 seconds. The total hyperplanes at termination, $|\Gamma|$ follows a similar trend with total CPU time. That is, functions with high $|\Gamma|$ register high CPU times. For instance, the polyhedra generated by functions 3.$\mathcal{A}$ and 3.$\mathcal{B}$ are less symmetric and tend to generate more vertices than the rest of functions. On the other hand, functions 1 and 4 possess 2-dimentional epigraphs, and therefore their polyhedrons are simpler. The former because it is defined over $\mathbb{R}$, and the latter because it possesses a lineality of 1.

Table 4.1: Performance of test functions at termination.

Test function	f	$\|\Gamma\|$	CPU time (s)	$\|f - f_\Gamma\|_\infty$	Convergence
1	$e^x + \delta_{(-\infty,0]}(x)$	99	0.156	$9.949 \cdot 10^{-5}$	yes
2.1	$\sqrt{1 + \|\mathbf{x}\|_1^2}$	419	1.297	$1.495 \cdot 10^{-4}$	no
2.∞	$\sqrt{1 + \|\mathbf{x}\|_\infty^2}$	518	2.875	$9.957 \cdot 10^{-5}$	yes
3.$\mathcal{A}$	$\mathcal{A}(x, y)$	1120	159.6	$9.992 \cdot 10^{-5}$	yes
3.$\mathcal{B}$	$\mathcal{B}(x, y)$	987	118.0	$9.983 \cdot 10^{-5}$	yes
4	$\log(e^x + e^y)$	153	0.266	$9.973 \cdot 10^{-5}$	yes

In Figure 4.1 it is shown the error of approximation per iteration in a $\log_{10}$ scale. The error decreases strictly and asymptotically to 0, as expected by Theorem 3.1. However, they do not decrease at the same rate, neither with a hierarchy between them. More precisely, Functions 1 and 4 converged faster than the other functions at any iteration, meanwhile functions 3.$\mathcal{A}$ and 3.$\mathcal{B}$, eventually, become the slowest functions. It is interesting that functions 2.1 and 2.∞ are slower than the rest of functions until about iteration 200, where the roles are inverted with functions 3.$\mathcal{A}$ and 3.$\mathcal{B}$. Almost every remark in this paragraph is supported by the ideas stated earlier with respect to symmetry and dimension of the epigraphs. However, it is not clear that those arguments could explain the different initial rates of convergence of each function.

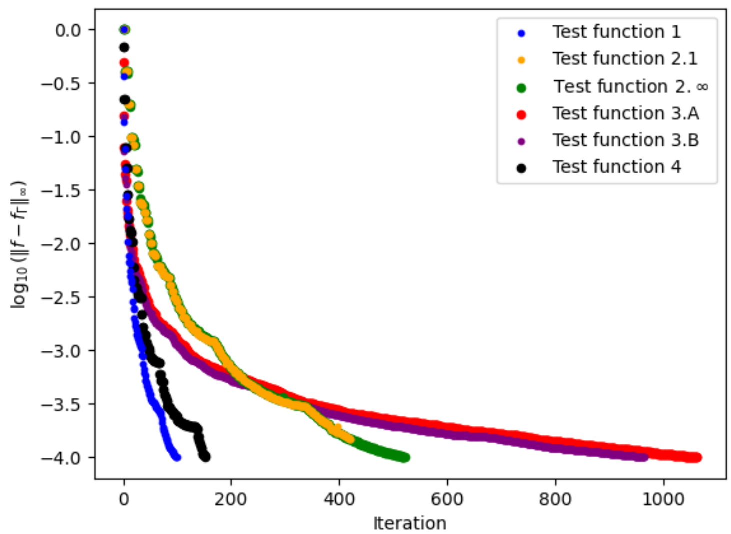
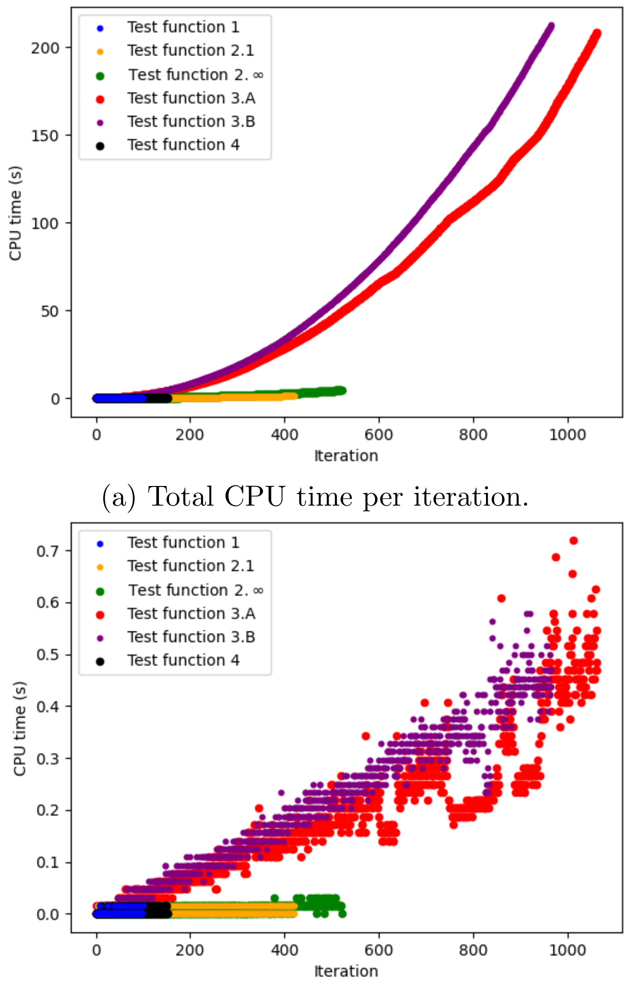


Figure 4.1: Error of approximation of f_Γ per iteration.

CPU Time dynamic

Figure 4.2a shows the total CPU time until at the end of each iteration. The CPU times exhibit a quadratic growth rate in terms of the iterations. This is more noticeable in the cases of functions 3.$\mathcal{A}$ and 3.$\mathcal{B}$. This claim is further supported by Figure 4.2b, which shows that the CPU times between iterations growth rate is at most linear. Thus, the accumulated CPU times must behave at most quadratic. It is clear that functions 3.$\mathcal{A}$ and 3.$\mathcal{B}$ possess the higher growth rates. Furthermore, it should be noted that the increase of time per iteration is due mostly to the vertex enumeration, which is carried out from scratch at each iteration and grows in complexity as more hyperplanes are added. However, the different growth ratios can be explained considering each function. In particular, functions 3.$\mathcal{A}$ and 3.$\mathcal{B}$ are highly non-linear, more costly computationally, and their polyhedra tend to generate more vertices than the other functions.



(a) Total CPU time per iteration.

(b) CPU time between iterations.

Figure 4.2: CPU time performance per function per iteration.

Chapter 5

A brief discussion of applications

The present thesis is mainly a theoretical work that provides the tools for generating POAs of a class of convex functions (those such that their conjugate has polyhedral domain). This specific setting may seem quite specific under the light of the theory of Numerical Analysis. This is in fact true, because POAs does not emerge randomly but in specific applications, to say, in the numerical resolution of convex MINLP. At the end of this chapter it shall be discussed how the approximations constructed in this thesis are a very useful tool in the context of linearizing convex MINLP. However, before that, other general applications of the POAs discussed in the present work are presented.

POAs of convex functions with bounded domains

Generating POAs of a function f where $dom\ f$ is bounded is a particular case of the theory developed in this thesis. The idea is to use Proposition 2.3 but with a slight modification to the hypothesis. More precisely, the assumption that $\mathcal{P}^\infty := (dom\ f)^\infty$ is full-dimensional must be discarded, since, in this case, this cone is no other than $\{\mathbf{0}\}$. However, by the same argument, it is true that $dom\ f^* = \mathbb{R}^n$. Therefore, a V-representation of $dom\ f^*$ is $\mathbf{0} + \mathbb{R}^n$. Thus, $\Pi = \{\mathbf{0}\}$, and the rest of the proof carries over. Thus, we deduce that

$f_\Pi(\mathbf{x}) = \delta_{dom\ f}(\mathbf{x}) + \min\ f$ is a[1] minimal POA of f.

Polyhedral Inner-Approximations (PIAs) of f

Assuming the conditions of Proposition 2.3, recall that the function f_Γ, which satisfies the inequality $f_\Gamma \leq f$, is a POA of f. Since said inequality is equivalent to the inclusion $epi(f) \subseteq epi(f_\Gamma)$, then being "outer" means that the epigraph f is covered by the epigraph of the POA. Likewise, a polyhedral inner-approximation (PIA) would occur if the inequality/inclusion were reversed. Indeed, for any $\Gamma \supseteq \Pi$, the inequality $f \leq f_\Gamma + \|f - f_\Gamma\|_\infty$ holds. Equivalently, $epi(f_\Gamma + \|f - f_\Gamma\|_\infty) \subseteq epi(f)$, thus we say that the function $f_\Gamma + \|f - f_\Gamma\|_\infty$ is a PIA f.

PIAs and POAs of f^*

Recall that the conjugate of a POA of f is a PIA of f^*, and *vice versa*. More precisely, the inequalities $f_\Gamma + \|f - f_\Gamma\|_\infty$ implies that $(f_\Gamma)^* \geq f^* \geq (f_\Gamma)^* - \|f - f_\Gamma\|_\infty$. Thus, $(f_\Gamma)^*$ and $(f_\Gamma)^* - \|f - f_\Gamma\|_\infty$ are PIAs and POAs of f^*, respectively. Moreover, the function $(f_\Gamma)^*$ may be described in terms of f^* and $\Gamma^* := \{\gamma^* \in \mathbb{R}^n \mid (\gamma^*, f_\Gamma(\gamma^*)) \in ext(epi(f_\Gamma^\perp))\}$. By construction, the set Γ^* contains all the non-vertical supporting hyperplanes of the H-representation of $(f_\Gamma)^*$. Additionally, $dom\ f^* = dom\ (f_\Gamma)^*$. Therefore, $(f_\Gamma)^* = f^*_{\Gamma^*}$.

[1] It should be noted that the minimal POA is uniquely determined by Π. Thus, any other choice for Π would lead to a slightly different approximation. However, the choice of $\Pi = \mathbf{0}$ is the most canonical.

On the linearization of Non-Linear optimization Problems

A general Convex Mixed Integer Non-Linear Problem (Convex MINLP), denoted the master problem, is a problem of the form:

$$\begin{aligned}(M_{f,\mathbf{g}}) : \ \min_{\mathbf{x}} \ & f(\mathbf{x}) \\ s.t. \ & \mathbf{g}(\mathbf{x}) \leq \mathbf{0} \\ & \mathbf{x} \in X\end{aligned}$$

where $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $g_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ are closed proper convex functions, for $i \in \{1, ..., m\}$, $\mathbf{g} := (g_i)_{i=1}^m$ is a vector-valued function, and where the set X establishes that the coordinates of $\mathbf{x}$ are either in $\mathbb{Z}$ or in $\mathbb{R}$.

To solve $M_{f,\mathbf{g}}$, common solution approaches involve, partially, dealing with its continuous relaxation denoted as $P_{f,\mathbf{g}}$. By definition, the optimal value of the relaxation, which we assume bounded, is a lower bound for the master problem, i.e., $P^*_{f,\mathbf{g}} \leq M^*_{f,\mathbf{g}}$. Variants of this NLP are expected to be solved many times. Thus, sometimes it is preferred to linearize the NLP into an LP, hoping the LP can be solved quick enough. However, some linearizations may be unbounded if the domain is unbounded and the approximation error of the linearization is not finite uniformly. This is show below.

Example 5.1: Consider the problem $P^*_f := \min f(x)$, where $f(x) := e^x + \delta_{(-\infty,0]}(x)$. Let $a < 0$, and define $g(x) := e^a x - e^a(a+1) + \delta_{(-\infty,0]}(x)$. This function is a POA of f constrained to $[a, 0]$. This approximation is naturally extended to *dom* f, but losing the finite uniform error of the approximation. Indeed, g is not an approximation of f, i.e., $\|f - g\|_\infty = \infty^+$, since $\lim\limits_{x\to-\infty} f(x) - g(x) = \infty^+$. Moreover, the approximated problem $P^*_g = \min g(x)$ differs from P^*_f, since the former is unbounded, but the latter have a finite optimal value equal to 0.

In Example 5.1, a convex NLP was linearized without ensuring the error is uniformly bounded, and it was shown that the linearized problem was unbounded despite the NLP being not. The present thesis presents necessary and sufficient conditions to ensure the approximation error is finite and uniformly bounded across the domain. This is sufficient, assuming strong duality of the NLP, to ensure the linearization is bounded as well, which is proven below.

Proposition 5.1. *Assume that $\underline{f}$ and $\underline{g_i}$ are outer-approximations of f and g_i, with a uniform bound of $\varepsilon > 0$ and $\varepsilon_i > \mathbf{0}$, for $i \in \{1, ..., m\}$, and define $\underline{\mathbf{g}} := (\underline{g_i})_{i=1}^m$ and $\boldsymbol{\varepsilon} := (\varepsilon_i)_{i=1}^m$. That is to say, $\underline{f} \leq f \leq \underline{f} + \varepsilon$ and $\underline{\mathbf{g}} \leq \mathbf{g} \leq \underline{\mathbf{g}} + \boldsymbol{\varepsilon}$. If strong duality holds for $P_{f,\mathbf{g}}$, and if λ^* is the optimal dual variable associated to the inequality $\mathbf{g} \leq \mathbf{0}$, then it is true that:*

$$0 \leq P^*_{f,\mathbf{g}} - P^*_{\underline{f},\underline{\mathbf{g}}} \leq \varepsilon + \lambda^{*\top}\boldsymbol{\varepsilon}.$$

Proof. Since it is true that $\underline{\mathbf{g}} \leq \mathbf{g}$, then the inclusion $\{\mathbf{x} \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\} \subseteq \{\mathbf{x} \mid \underline{\mathbf{g}}(\mathbf{x}) \leq \mathbf{0}\}$ holds. Using the fact that $f \geq \underline{f}$, the lower bound is deduced:

$$P^*_{f,\mathbf{g}} = \min_{\mathbf{g}(\mathbf{x}) \leq \mathbf{0}} f(\mathbf{x}) \geq \min_{\mathbf{g}(\mathbf{x}) \leq \mathbf{0}} \underline{f}(\mathbf{x}) \geq \min_{\underline{\mathbf{g}}(\mathbf{x}) \leq \mathbf{0}} \underline{f}(\mathbf{x}) = P^*_{\underline{f},\underline{\mathbf{g}}}.$$

The upper bound is deduced using Lagrangian Duality. In particular, observe that:

$$\begin{aligned}
P^*_{f,\mathbf{g}} &= \max_{\lambda \geq \mathbf{0}} \min_{\mathbf{x}} f(\mathbf{x}) + \lambda^\top \mathbf{g}(\mathbf{x}), \\
&= \min_{\mathbf{x}} f(\mathbf{x}) + \lambda^{*\top} \mathbf{g}(\mathbf{x}), \\
&\leq \min_{\mathbf{x}} (\underline{f}(\mathbf{x}) + \varepsilon) + \lambda^{*\top} (\underline{\mathbf{g}}(\mathbf{x}) + \boldsymbol{\varepsilon}), \\
&= \min_{\mathbf{x}} \left(\underline{f}(\mathbf{x}) + \lambda^{*\top} \underline{\mathbf{g}}(\mathbf{x}) \right) \varepsilon + \lambda^{*\top} \boldsymbol{\varepsilon}, \\
&\leq \max_{\lambda \geq \mathbf{0}} \min_{\mathbf{x}} \left(\underline{f}(\mathbf{x}) + \lambda^\top \underline{\mathbf{g}}(\mathbf{x}) \right) + \varepsilon + \lambda^{*\top} \boldsymbol{\varepsilon}, \\
&\leq P^*_{\underline{f},\underline{\mathbf{g}}} + \varepsilon + \lambda^{*\top} \boldsymbol{\varepsilon}.
\end{aligned}$$

□

Chapter 6

Conclusion and future work

In this thesis, it was addressed the POA of convex functions which their domain and of their conjugate are polyhedrons. Necessary and sufficient conditions are stated in order to achieve a finite uniform approximation error across the domain. Moreover, these conditions imply the existence of a minimal approximation with respect to the supporting hyperplanes required to describe it. This is used to show that a subclass of functions possesses the same minimal approximation. However, in general, this minimal approximation must be computed manually. An algorithm is proposed to reduce the error of a POA until a predefined error $\varepsilon > 0$ is reached. Our method relies on augmenting the set of hyperplanes that define the approximation using valid hyperplanes at the points where the pointwise error is maximized, i.e., is equal to the approximation error. These points are characterized as the projection onto the domain of the vertices of the POA's epigraph. Under mild assumptions, the algorithm was shown correct, i.e., the error has to decrease at each iteration.

The computational performance of the algorithm was illustrated by a set of 6 test functions. The algorithm converged for all tests, with one exception due to numerical inaccuracies. Proving the convergence of the algorithm poses a challenging subject for future research, due to the lack of boundedness of the domain. Additionally, we observed that the major bottleneck of the algorithm, regarding speed, comes from performing the vertex enumeration from scratch at every iteration. This could become an issue, since for some of the applications that the present thesis aims to (see section 5) it would be mandatory that the algorithm is as quick as possible. This could be improved exploiting the fact that each new facet added to the polyhedron at each iteration prunes a small subset of its vertices. That is to say, the polyhedron is mathematically updated locally, and that should be the case algorithmically. Designing a specialized data structure that updates the polyhedron instead of reconstructing it will improve the CPU times dramatically.

Appendix A

Test function 1

Let $f(x) := e^x + \delta_{(-\infty,b]}(x)$, for any $b \in \mathbb{R}$. Since this is just the same as in Example 2.2 but with $p = b$, then this function is approximable by f_Π, where $\Pi = \{0\}$ and $min\ f = 0$. Moreover, regarding the subdifferential, note that for any $x \in (-\infty, b)$, $f^{'}(x) = f(x)$. For the case $x = b$, it is considered the continuous extension of $f^{'}(x) = f(x)$.

Appendix B

Test functions 2.1 and 2.∞

The next case considers a parametrized function, where 2 cases are of interest. Let $p, q \in [1, \infty^+]$, such that the equality[1] $\frac{1}{p} + \frac{1}{q} = 1$ holds in the extended real-valued sense. In the case that either p or q are ∞^+, then the other is exactly 1. Consider the function $f(\mathbf{x}) := \sqrt{1 + \|\mathbf{x}\|_p^2}$. Since $\|x\|_p$ is a norm, then it is a closed, proper, and convex function. Thus, f is a closed, proper, and convex function as well. Using the limit formula, it is computed the recession function:

$$\begin{aligned} f^\infty(\mathbf{d}) &= \lim_{\lambda \to \infty^+} \frac{\sqrt{1 + \|\mathbf{x} + \lambda \mathbf{d}\|_p^2} - \sqrt{1 + \|\mathbf{x}\|_p^2}}{\lambda}, \\ &= \lim_{\lambda \to \infty^+} \sqrt{\frac{1}{\lambda^2} + \left\| \frac{\mathbf{x}}{\lambda} + \mathbf{d} \right\|_p^2} - \frac{\sqrt{1 + \|\mathbf{x}\|_p^2}}{\lambda}, \\ &= \|\mathbf{d}\|_p. \end{aligned}$$

Recall that any p-norm may be computed by its dual norm, i.e., $\|\mathbf{d}\|_p = \sup_{\mathbf{y} \in \{\|\mathbf{y}\|_q \leq 1\}} \mathbf{y}^\top \mathbf{d}$. The latter is just equal to $\sigma_{\mathbb{B}_q}(\mathbf{d})$, where $\mathbb{B}_q := \mathbb{B}_q(0, 1) = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y}\|_q \leq 1\}$, i.e., is the unitary ball in the q norm. Thus, $f^\infty(\mathbf{d}) = \sigma_{\mathbb{B}_q}(\mathbf{d})$. In which case, f^∞ is polyhedral if and only if $\mathbb{B}_q$ is polyhedral. This is the case if q is either 1 or ∞, which is in bijection to p being ∞^+ or 1, respectively. Thus, assume that $p \in \{1, \infty^+\}$. Let $\Pi = ext(\mathbb{B}_q)$. Since $\mathbb{B}_q$ is bounded, it is deduced that $\overline{dom\ f^*} = \mathbb{B}_q = conv(\Pi)$. It follows to show that $dom\ f^*$ is closed, which will be proven by showing that $f^*(\pi) = 0$, for

[1] This equality is a necessary condition to formulate the famous Hölder Inequality.

every $\pi \in \Pi$. Recall that for any $\pi \in \Pi$, $\|\pi\|_q = \sigma_{\mathbb{B}_p}(\pi) = 1$, and since $\mathbb{B}_q$ is a compact set, there exists a vector $\overline{\mathbf{x}}_\pi \in \mathbb{B}_p$ such that $\|\pi\|_q = \pi^\top \overline{\mathbf{x}}_\pi = 1$. Therefore, the following lower bound is computed:

$$\begin{aligned} f^*(\pi) &= \sup_{\mathbf{x}\in\mathbb{R}^n} \pi^\top \mathbf{x} - f(\mathbf{x}), \\ &\geq \sup_{\lambda\geq 0} \lambda\pi^\top \overline{\mathbf{x}}_\pi - \sqrt{1 + \|\lambda\overline{\mathbf{x}}_\pi\|_p^2}, \\ &= \lim_{\lambda\to\infty^+} \lambda - \sqrt{1+\lambda^2} = 0. \end{aligned}$$

Conversely, the Hölder inequality implies that $\pi^\top \mathbf{x} \leq \|\pi\|_q \|\mathbf{x}\|_p \leq \sqrt{1 + \|\mathbf{x}\|_p^2}$, since $\|\pi\|_q = 1$. Thus, $f^*(\pi) \leq 0$ by definition. Therefore, we conclude that $f^*(\pi) = 0$ for all π in Π, as desired. The approximation is $f_\Pi(\mathbf{x}) = \sigma_\Pi(\mathbf{x})$, with $\Pi = ext(\mathbb{B}_q)$.

Lastly, it should be noted that f is not differentiable, but subdifferentiable. Using subdifferential rules, one deduces that:

$$\partial f(\mathbf{x}) = \frac{\|\mathbf{x}\|_p}{\sqrt{1 + \|\mathbf{x}\|_p^2}} \partial \|\mathbf{x}\|_p.$$

Recall that $\|\mathbf{x}\|_p = \sigma_{\mathcal{B}_q}(\mathbf{x})$. Thus, $\partial f(\mathbf{x}) = \partial\sigma_{\mathcal{B}_q}(\mathbf{x})$. The latter equals $\arg\max_{\pi\in\Pi} \pi^\top \mathbf{x}$ (see [7, corollary 23.5.3]). This is always a non-empty finite set. If it is not single-valued, consider the barycenter of $\partial f(\mathbf{x})$, which in this case coincides with the mean of the set Π.

Appendix C

Test functions $3.\mathcal{A}$ and $3.\mathcal{B}$

In Inventory Control Theory, particularly, in the subject of the calculus of optimal reorder points under uncertainty, two famous convex functions with unbounded domain appear naturally. These are the fill-rate function $\mathcal{A}$ and the backorder function $\mathcal{B}$. Both functions are inside the easy-to-approximate class of functions described in Corollary 2.2. Before showing their inclusion to said class of functions, the precise definition of the fill-rate and backorder function shall be given below. To achieve this, it is necessary to define auxiliary concepts before.

Let D be a continuous random variable[1] with finite mean μ and variance σ. Let F_D be the cumulative probability function (CDF) of D, and f_D the probability density function (PDF) of D. Define the functions $G, H : [\mu, \infty^+) \to \mathbb{R}$ as follows:

$$
\begin{aligned}
G(x) &:= \int_x^{\infty^+} 1 - F_D(t)dt, \\
H(x) &:= \int_x^{\infty^+} G(t)dt.
\end{aligned}
$$

Recall that $F_D(x) \in [0, 1]$ for any $x \in \mathbb{R}$, thus G is non-negative. In consequence H is non-negative as well. What's more, $H^{'}(x) = -G(x) \leq 0$. Also, observe that $G^{'}(x) = F_D(x) - 1 \leq 0$ and $G^{''}(x) = f_D(x) \geq 0$. Furthermore, $H^{''}(x) = 1 - F_D(x) \geq 0$. That is to say that both G and H are non-increasing convex functions. Moreover, they are closed, and proper functions. Closed-

[1] In the theory of Inventory control, D denotes the demand of a product.

ness is directly deduced from the integrability of $1 - F_D$ and G. Properness is deduced from the Chebyshev inequality.

Fill-rate function $\mathcal{A}$

An additional hypothesis over D is necessary in order to guarantee the convexity of $\mathcal{A}$. Suppose that there exists a constant $\nu \in \mathbb{R}$ such that for any $x \geq \nu$, the distribution $f_D(x)$ is non-increasing. Let $c := \max\{\nu, \mu\}$. The fill-rate function is defined as follows:

$$\mathcal{A} : [c, \infty^+) \times [0, \infty^+) \to \mathbb{R}$$
$$(x, y) \mapsto \frac{1}{y} \int_x^{x+y} 1 - F_D(t) dt.$$

As the integrand is non-negative, $\mathcal{A}$ is too. To show that this function is well-defined, it follows to show that it is finite at every point of the domain. By the consideration of the constant c, then $\mathcal{A}$ must be convex (see [11]). Moreover, it is clear that $\mathcal{A}(x, y) \leq 1$, by upper-bounding the term $-F_D$ by 0. Therefore, it is well-defined, and moreover, it is continuous at *dom* $\mathcal{A}$. Thus, it should be noted that *dom* $\mathcal{A} = \mathcal{K} + \mathbf{p}$, where $\mathcal{K} = [0, \infty^+)^2$ and $\mathbf{p} = (c, 0)$.

It follows to show that $\mathcal{A}$ is non-increasing with respect to $\mathcal{K}$. First, it will be shown that this is true for $\mathcal{A}$ restricted to *int*(*dom* $\mathcal{A}$). Consider any $(x, y) \in$ *int*(*dom* $\mathcal{A}$). The gradient of $\mathcal{A}$ is:

$$\nabla \mathcal{A}(x, y) = \left(\frac{F_D(x) - F_D(x + y)}{y}, \frac{1}{y} \int_x^{x+y} \frac{F_D(t) - F_D(x + y)}{y} dt \right).$$

Recall that F_D is always non-decreasing, therefore the terms $F_D(x) - F_D(x + y)$ and $F_D(t) - F_D(x + y)$ are both non-positive. Thus, $\nabla \mathcal{A}(x, y) \leq (0, 0)$. This directly implies that $\mathcal{A}$ is non-increasing with respect to $\mathcal{K}$ in *int*(*dom* $\mathcal{A}$).

From the formula derived for $\nabla \mathcal{A}$ at *int*(*dom* $\mathcal{A}$), it is easy to see that it has a continuous extension to the boundary of *dom* $\mathcal{A}$. In particular, for any

$x \in [c, \infty^+)$:

$$\begin{aligned}
\lim_{y\to 0} \nabla \mathcal{A}(x,y) &= \lim_{y\to 0} \left(\frac{F_D(x) - F_D(x+y)}{y}, \frac{1}{y} \int_x^{x+y} \frac{F_D(t) - F_D(x+y)}{y} dt \right), \\
&= \left(-f_D(x), \lim_{y\to 0} \int_0^1 \frac{F_D(ty+x) - F_D(x+y)}{y} dt \right), \\
&= \left(-f_D(x), \int_0^1 (t-1) f_D(x) dt \right), \\
&= f_D(x) \left(-1, -\frac{1}{2} \right), \\
&\leq \mathbf{0}.
\end{aligned}$$

Likewise with the other edge of the boundary, consider any $y \in [0, \infty^+)$:

$$\begin{aligned}
\lim_{x\to c} \nabla \mathcal{A}(x,y) &= \lim_{x\to c} \left(\frac{F_D(x) - F_D(x+y)}{y}, \frac{1}{y} \int_x^{x+y} \frac{F_D(t) - F_D(x+y)}{y} dt \right), \\
&= \left(\frac{F_D(c) - F_D(c+y)}{y}, \lim_{x\to c} \frac{1}{y} \int_0^y \frac{F_D(t+x) - F_D(x+y)}{y} dt \right), \\
&= \left(\frac{F_D(c) - F_D(c+y)}{y}, \frac{1}{y} \int_c^{c+y} \frac{F_D(t) - F_D(c+y)}{y} dt \right), \\
&\leq \mathbf{0}.
\end{aligned}$$

In summary, since the partial derivatives of $\mathcal{A}$ are non-positive all over $\mathcal{K} + \mathbf{p}$, then $\mathcal{A}$ is non-increasing with respect to $\mathcal{K} = [0, \infty^+)^2$.

It remains to show that min $\mathcal{A}$ is exactly 0. As stated earlier, $\mathcal{A}$ is non-negative. Additionally, the bound $\mathcal{A}(x,y) \leq \frac{G(x)}{y}$ is true as a consequence of the non-negativity of G. Taking the limit of y towards ∞^+, one concludes that the minimum is not greater than 0, because this limit constitutes a minimizing succession of values of $\mathcal{A}$ along a ray. But as $\mathcal{A}$ is non-negative, the conclusion is that the minimum is exactly 0. This ends the proof to the claim that $\mathcal{A}$ satisfy the conditions of Corollary 2.2, for $\mathcal{K} = [0, \infty^+)^2$ and $\mathbf{p} = (c, 0)$, and min $\mathcal{A} = 0$.

Backorder function $\mathcal{B}$

The backorder function is convex without the assumption of the constant c in the case of the fill-rate function (see [11]). It only needs to be defined considering μ instead. The backorder function is defined as follows:

$$\mathcal{B} : [\mu, \infty^+) \times [0, \infty^+) \to \mathbb{R}$$
$$(x, y) \mapsto \frac{1}{y} \int_x^{x+y} G(t)dt.$$

In this case, the domain is the translated cone $\mathcal{K} + \mathbf{p}$, where $\mathcal{K} = [0, \infty^+)^2$ and $\mathbf{y} = (\mu, 0)$. Since G is non-increasing, it is true that $\mathcal{B}(x, y) \leq G(x)$. Moreover, clearly $\mathcal{B}$ is non-negative. Therefore, $\mathcal{B}$ is well-defined, closed, and proper.

It follows to show that $\mathcal{B}$ is non-increasing with respect to $\mathcal{K}$. The argument here is similar to the one used to show that $\mathcal{A}$ was non-increasing. The gradient of $\mathcal{B}$ might be computed easily at $int(dom\ \mathcal{B})$, obtaining that:

$$\nabla\mathcal{B}(x, y) = \left(\frac{G(x + y) - G(x)}{y}, \frac{1}{y} \int_x^{x+y} \frac{G(x + y) - G(t)}{y} dt\right).$$

Clearly, as G is non-increasing, then both terms $G(x+y)-G(x)$ and $G(x+y) - G(t)$ are non-positive, thus concluding that $\nabla B(x, y) \leq (0, 0)$. That is to say, $\mathcal{B}$ is non-increasing with respect to $\mathcal{K}$ at $dom\ \mathcal{B}$.

From the formula derived for $\nabla\mathcal{B}$, a continuous extension to the boundary of $dom\ \mathcal{B}$ is possible. In particular, observe that for any $x \in [\mu, \infty^+)$:

$$\begin{aligned}
\lim_{y\to 0} \nabla\mathcal{B}(x, y) &= \lim_{y\to 0} \left(\frac{G(x + y) - G(x)}{y}, \frac{1}{y} \int_x^{x+y} \frac{G(x + y) - G(t)}{y} dt\right), \\
&= \left(G^{'}(x), \lim_{y\to 0} \int_0^1 \frac{G(x + y) - G(ty + x)}{y} dt\right), \\
&= \left(G^{'}(x), \int_0^1 (1 - t)G^{'}(x)dt\right), \\
&= (F_D(x) - 1)\left(1, \frac{1}{2}\right), \\
&\leq \mathbf{0}.
\end{aligned}$$

Likewise with the other edge of the boundary, for any $y \in [0, \infty^+)$:

$$\begin{aligned}
\lim_{x\to \mu} \nabla\mathcal{B}(x, y) &= \lim_{x\to \mu} \left(\frac{G(x + y) - G(x)}{y}, \frac{1}{y} \int_x^{x+y} \frac{G(x + y) - G(t)}{y} dt\right), \\
&= \left(\frac{G(\mu + y) - G(\mu)}{y}, \frac{1}{y} \int_\mu^{\mu+y} \frac{G(\mu + y) - G(t)}{y} dt\right), \\
&\leq \mathbf{0}.
\end{aligned}$$

Finally, it remains to show that min $\mathcal{B}$ is finite and, in fact, equal to 0. Noticing that $\lim_{x\to\infty^+} G(x) = 0$, and that $0 \le \mathcal{B}(x, y) \le G(x)$, then clearly the minimum is 0, as this limit provides a minimizing succession.

Appendix D

Test function 4

This case considers a function with lineality. Let $f(\mathbf{x}) := \log(\sum_i e^{x_i})$. It has already been proven closed, proper, and convex (see [7, §16]). Moreover, *dom* $f^* =$ *conv*$\{\mathbf{e}_i\}_{i=1}^n$, and $f^*(\mathbf{e}_i) = 0$ for all i. Thus, $f_\Pi \equiv \sigma_\Pi$, where $\Pi = \{\mathbf{e}_i\}_i$. Numerically, $\partial f(\mathbf{x}) = \{(e^{x_j}/(\sum_i e^{x_i}))_{j=1}^n\}$. It is easy to show that $(Lin\ f)^\perp = \mathbb{1}^\perp$.

Bibliography

[1] Dörfler, D. On the Approximation of Unbounded Convex Sets by Polyhedra. *Journal Of Optimization Theory And Applications*. **194**, 265-287 (2022,3)

[2] Escalona, P., Brotcorne, L., Angulo, A., Wolf, N. and Mora, C. (2026), The uncapacitated joint location-inventory problem with shortage costs: a selective extended cutting plane algorithm. Intl. Trans. in Op. Res. https://doi.org/10.1111/itor.70160

[3] Hsieh, C. On solving robust log-optimal portfolio: A supporting hyperplane approximation approach. *European Journal Of Operational Research*. **313**, 1129-1139 (2024), https://www.sciencedirect.com/science/article/pii/S0377221723007579

[4] Ney Polyhedral approximation of convex sets with an application to large deviation probability theory. *Journal Of Convex Analysis*. **2**, 229-240 (1995), http://eudml.org/doc/233358

[5] Minkowski, Hermann. "Volumen und Oberfläche." Mathematische Annalen 57 (1903): 447-495. http://eudml.org/doc/158108.

[6] Melo, Fampa & Raupp Two linear approximation algorithms for convex mixed integer nonlinear programming. *Annals Of Operations Research*. **316**, 1471-1491 (2022)

[7] Rockafellar, R. T. Convex Analysis. (Princeton University Press,1970),http://www.jstor.org/stable/j.ctt14bs1ff

[8] Rockafellar, R.T. and Wets, R.J.-B. (1997) Variational Analysis. 1st edn. Berlin, Germany: Springer (Grundlehren der mathematischen Wissenschaften, 317). https://doi.org/10.1007/978-3-642-02431-3.

[9] Toriello, A. & Vielma, J. Fitting piecewise linear continuous functions. *European Journal Of Operational Research*. **219**, 86-95 (2012), https://www.sciencedirect.com/science/article/pii/S0377221711011246

[10] Warin, X. The GroupMax Neural Network Approximation of Convex Functions. *IEEE Transactions On Neural Networks And Learning Systems*. **35**, 11608-11612 (2024)

[11] Zipkin, P. Inventory Service-Level Measures: Convexity and Approximation. *Management Science*. **32(8)** pp. 975-981 (1986)