On Cutting Planes for Mixed-Integer Nonlinear Programming

vorgelegt von
M. Sc.
Felipe Serrano Musalem
ORCID: 0000-0002-7892-3951

an der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
– Dr. rer. nat. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Wolfgang König
Gutachter:	Prof. Dr. Thorsten Koch
	Prof. Dr. Juan Pablo Vielma

Tag der wissenschaftlichen Aussprache: 21. August 2020

Berlin 2021

Abstract

Mixed-integer nonlinear programming is a powerful technology that allows us to model and solve problems involving nonlinear functions, continuous, and discrete variables. The state-of-the-art solvers of mixed-integer nonlinear programs (MINLPs) use a combination of, among other techniques, branch-and-bound and cutting planes. In the late '90s, solvers for mixed-integer linear programs saw an increase in performance due to the incorporation of general-purpose cutting planes.

In this thesis, we deepen our understanding of a classical cutting planes algorithm, develop a strengthening technique, and two new cutting planes for MINLPs.

We first show that Veinott's supporting hyperplane algorithm is a particular case of Kelley's cutting plane algorithm. We further extend the applicability of Veinott's supporting hyperplane algorithm to solve convex problems represented by non-convex functions.

We then develop a technique to strengthen cutting planes for non-convex MINLPs. Many cuts for non-convex MINLPs strongly rely on the domain of the variables: tighter bounds produce tighter cuts. Using the point to be separated, we show that we can restrict the feasible region and still ensure the validity of the resulting cutting plane.

Finally, we develop two intersection cuts for non-convex MINLP. The first one is a technique to construct S-free sets for any factorable MINLP. For the second one, we show how to build maximal quadratic-free sets, from which we compute intersection cuts. These last cuts reduce the average running time of the solver SCIP by 20% on hard MINLPs.

Zusammenfassung

Die gemischt-ganzzahlige nichtlineare Programmierung ist eine leistungsstarke Technik, mit der wir Probleme modellieren und lösen können, die nichtlineare Funktionen und kontinuierliche und diskrete Variablen enthalten. Die hochmodernen Löser für gemischt-ganzzahlige nichtlineare Programme (MINLPs) verwenden unter anderem eine Kombination der Branch-and-Bound-Methode und Schnittebenengenerierung. In den späten 90er Jahren erfuhren die Löser für gemischt-ganzzahlige lineare Programme eine Leistungssteigerung durch die Einbeziehung von universell nutzbaren Schnittebenen.

In dieser Arbeit vertiefen wir unser Verständnis eines klassischen Schnittebenen-Algorithmus, wir entwickeln eine Verstärkungstechnik und zwei neue Schnittebenen für MINLPs.

Zunächst zeigen wir, dass der Stützhyperebenen-Algorithmus von Veinott ein Sonderfall des Kelley'schen Schnittebenen-Algorithmus ist. Darüber hinaus erweitern wir die Anwendbarkeit von Veinotts Stützhyperebenen-Algorithmus auf die Lösung konvexer Probleme, die durch nicht-konvexe Funktionen repräsentiert werden.

Anschließend entwickeln wir eine Technik zur Verstärkung der Schnittebenen für nicht-konvexe MINLPs. Viele Schnitte für nicht-konvexe MINLPs hängen stark vom Wertebereich der Variablen ab: Strengere Schranken erzeugen stärkere Schnitte. Anhand des zu separierenden Punktes zeigen wir, dass wir die zulässige Region einschränken können und dennoch die Gültigkeit der resultierenden Schnitte beibehalten.

Schließlich entwickeln wir zwei Überschneidungsschnittebenen für nicht-konvexe MINLPs. Der erste Schnitt ist eine Technik zur Konstruktion S-freier Mengen für beliebige faktorisierbare MINLPs. Für den zweiten Schnitt zeigen wir, wie man maximal quadratisch-freie Mengen bildet, aus denen wir Überschneidungsschnittebenen berechnen. Diese Schnitte reduzieren die durchschnittliche Laufzeit des Lösers SCIP um 20% bei schwierigen Problemen.

Contents

Chapter 1

Introduction

This thesis develops techniques for solving *mixed-integer nonlinear problems*, in particular, techniques related to *cutting planes*. A mixed-integer nonlinear problem (MINLP) belongs to the class of Mathematical Programming (MP).

In its simplest form, MP is concerned with finding the largest or smallest value that a function can attain in some domain. For example, finding the region of smallest surface that has a prescribed volume, or finding the path that a ball has to take so that it goes from point A to point B in the least amount of time under the influence of gravity. Already at this point one can suspect that MP has lots of applications, just imagine packing a given volume of liquid using the least amount of material. More modern examples of MP problems include finding the shortest path between two points in a city, or deciding where to open stores from a given set of possible locations such that customers' average shortest travel time is minimized, etc. One can find an impressive amount of applications in the survey of Boukouvala, Misener, and Floudas (2016).

The example problems mentioned above above have two distinct features. The first examples are continuous, that is, the solution can be any real number. In contrast, the last examples are discrete. Discrete structures appear, for example, when we can only choose from a finite set of possibilities.

One of the features of these type of problems is that they can be translated, with more or less work, to a mathematical model. That is, the set of feasible solutions can be described by equations and inequalities, called *constraints*, while the criterion we want to optimize over can be described as a function, called *objective function*. As a toy example, suppose we are interested in finding two non-negative integers such that the cube of one number is two units away from the square of the other and their sum is smallest. If x and y are the two integer numbers and v is the value of their sum, the problem above can be written as

$$\min\{v : v = x + y,\ x^3 - y^2 = 2,\ x, y \in \mathbb{Z}_+,\ v \in \mathbb{R}\}. \tag{1.1}$$

In (1.1) we encounter the constraints $v = x + y$, $x^3 - y^2 = 2$, $x, y \in \mathbb{Z}_+$ and $v \in \mathbb{R}$, and the objective function is just v, which is the quantity we want to minimize. The constraint $v = x + y$ is linear, while $x^3 - y^2 = 2$ is nonlinear. The variables x, y are restricted to be integers while v is continuous.

Such a model is an example of an MINLP problem. The "mixed-integer" comes from the fact that variables can be either discrete or continuous. The "nonlinear" makes reference to the possibility of having constraints represented by nonlinear functions.

More general, a generic MINLP can be written as

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_k(x) \leq 0 & \forall k \in [m], \\ & x_i \in \mathbb{Z} & \forall i \in I, \end{array}$$

where $m, n \in \mathbb{Z}_+$, $f, g_k : A \subseteq \mathbb{R}^n \to \mathbb{R}$, $[m] = \{1, \ldots, m\}$, $x \in \mathbb{R}^n$, and $I \subseteq [n]$. We note that assuming that the constraints are $g_k(x) \leq 0$ is without loss of generality, since $g_k(x) = 0$ is equivalent to $g_k(x) \leq 0$ and $-g_k(x) \leq 0$.

In practice, MINLP problems are difficult to solve. The best algorithm we currently have for trying to solve a general MINLP is the so-called *LP-based spatial branch and bound*. LP stands for *linear programming*, which is a subclass of MINLP concerned with optimization problems where all variables are continuous and all constraints are linear. In contrast to MINLPs, LPs are easy to solve in practice.

The basic idea of LP-based spatial branch and bound is to construct an *LP relaxation* of the MINLP, that is, an LP such that every feasible point of the MINLP is feasible for the LP. Solving this LP yields a bound on the optimal value of the MINLP. The solution of the LP, $\bar{x}$ is likely to be infeasible for the MINLP. Thus, the LP relaxation can, in principle, be refined by the introduction of cutting planes separating $\bar{x}$. These are linear inequalities that every point of the MINLP satisfies and $\bar{x}$ does not satisfy. By refining the LP relaxation, we obtain a better bound on the optimal value of the MINLP.

For example, it is not hard to see that $(x, y, v) = (3, 5, 8)$ is an optimal solution of (1.1) (just check that $(3, 5)$ is the only feasible point in $\{1, 2, 3\} \times \{1, 2, 3, 4, 5\}$). An LP relaxation of (1.1) is $\min\{v : v = x + y,\ x, y \geq 0\}$ for which an optimal solution is $(\bar{x}, \bar{y}, \bar{v}) = (0, 0, 0)$. The optimal value of the LP is 0, which is a (lower) bound on the optimal value of the MINLP, which is 8.

Now, since $x^3 = y^2 + 2$ and $y^2 \geq 0$, we can deduce that $x^3 \geq 2$. This implies that $x > 1$ and since x must be integral, we conclude that $x \geq 2$. Note that the LP solution does not satisfy $x \geq 2$. Thus, $\min\{v : v = x + y,\ x \geq 2,\ y \geq 0\}$ is a tighter LP relaxation. An optimal solution of this LP is $(\bar{x}, \bar{y}, \bar{v}) = (2, 0, 2)$ and yields a better lower bound. Cuts that involve a single variable are usually called *bound tightenings*.

Notice that the LP solution, $(\bar{x}, \bar{y}) = (2, 0)$ violates the constraint $x^3 = y^2 + 2$. In particular, if we interpret the equality as two inequalities, then the violated inequality is $x^3 - y^2 \leq 2$. Since $x \geq 2$ and $y \geq 0$, the above inequality is equivalent to $\sqrt{x^3 - 2} - y \leq 0$. The function $f(x) = \sqrt{x^3 - 2}$ is convex and differentiable at $x = 2$ and so $f(2) + f^{'}(2)(x - 2) \leq f(x)$, that is, $\sqrt{6}x - \sqrt{6} \leq \sqrt{x^3 - 2}$ for $x \geq 2$. Therefore, every feasible point must satisfy $\sqrt{6}x - \sqrt{6} - y \leq 0$. We see that $(\bar{x}, \bar{y}) = (2, 0)$ does not satisfy this inequality. Such an inequality is then cutting plane and its addition to the current LP relaxation makes it tighter. Indeed, by adding it and solving the corresponding LP we obtain the optimal point $(\bar{x}, \bar{y}, \bar{v}) = (2, \sqrt{6}, 2 + \sqrt{6})$ with value $2 + \sqrt{6}$, which is better than the one of the previous iteration.

However, at some point it might not be possible to compute a cutting plane and so the algorithm starts *branching*. In its most basic form, branching means to split the feasible region into two regions, in such a way that the union of both regions is the original feasible region. For example, in the last LP relaxation we obtained $\bar{y} = \sqrt{6}$. Branching on y at $\sqrt{6}$ produces two problems which are the same as the original one, except that in one the constraint $y \leq \sqrt{6}$ is added and in the other one, $y \geq \sqrt{6}$. Since y is restricted to be an integer we can further make these inequalities tighter. Thus, after branching on y we obtain the following problems

$$\min\{v : v = x + y,\ x^3 - y^2 = 2,\ y \leq 2,\ x, y \in \mathbb{Z}_+,\ v \in \mathbb{R}\} \text{ and}$$
$$\min\{v : v = x + y,\ x^3 - y^2 = 2,\ y \geq 3,\ x, y \in \mathbb{Z}_+,\ v \in \mathbb{R}\}.$$

The adjective *spatial* in spatial branch and bound means that the branching can also be done on continuous variables, for example, v. The adjective is added to distinguish the algorithm from the standard branch-and-bound algorithm for solving mixed-integer *linear* problems (MILPs). Via branching, the algorithm implicitly constructs a tree of problems.

By continuing the branching process the problem will eventually be solved. However, as can be seen from the example, cutting planes are an important tool for tightening the LP relaxation of the MINLP, whose purpose is to accelerate the solution process.

Let us look at another example to illustrate another important tool for solving MINLPs. Assume we are interested in buying some number of shirts and pants in such a way that the number of different outfits we can create is maximal. We enter a rather expensive shop where the cost of each shirt is 30 euros, while each pant is 70 euros, and we have 250 euros in our wallet. If s is the number of shirts and p the number of pants that we buy, then the number of outfits is $T = s \cdot p$. Then, the problem we try to solve is

$$\max\{T : T \leq s \cdot p,\ 3s + 7p \leq 25,\ s, p \in \mathbb{Z}_+\}.$$

Let us first notice that we do not have enough money to buy 9 shirts nor 4 pants, so $s \leq 8$ and $p \leq 3$. One way of obtaining a linear relaxation for this problem is to find a linear relaxation of the constraint $T \leq s \cdot p$. To obtain one, notice that for every feasible p and s we have that $s(3 - p) \geq 0$ and $(8 - s)p \geq 0$. Thus, $T \leq s \cdot p \leq \min\{3s, 8p\}$. These are the famous McCormick inequalities (McCormick, 1976). Our first linear relaxation then looks like

$$\max\{T : T \leq 3s,\ T \leq 8p,\ 3s + 7p \leq 25,\ s, p \in \mathbb{R}_+\}.$$

We could have added the bounds $s \leq 8$, $p \leq 3$, but less us keep it simple. The optimal solution of the linear relaxation is $(T, s, p) \approx (13.3, 4.4, 1.6)$. As this is an upper bound on the optimal value, we know that it is not possible to get 14 different outfits. Let us branch on $s \leq 4$ and $s \geq 4$. The first problem created is

$$\max\{T : T \leq s \cdot p,\ 3s + 7p \leq 25,\ s \leq 4,\ s, p \in \mathbb{Z}_+\}.$$

If we solve the linear relaxation

$$\max\{T : T \leq 3s,\ T \leq 8p,\ 3s + 7p \leq 25,\ s \leq 4,\ s, p \in \mathbb{R}_+\},$$

we obtain a value of $T = 12$. However, when branching on $s \leq 4$, the upper bound of s is reduced from 8 to 4. Thus, there is a chance that we can deduce a better linear relaxation of $T \leq s \cdot p$. Indeed, following the same reasoning as above we see that $T \leq s \cdot p \leq \min\{3s, 4p\}$. Now, solving the improved linear relaxation

$$\max\{T : T \leq 3s,\ T \leq 4p,\ 3s + 7p \leq 25,\ s \leq 4,\ s, p \in \mathbb{R}_+\},$$

yields $T \approx 9.09$, which is a much better upper bound. This shows that if we buy 4 or less shirts we can only hope for 9 outfits. The algorithm will continue either by branching or cutting. If anybody is interested, the maximum number of outfits is actually 6, far away from the possibility of 13 given by the first linear relaxation.

This example illustrates that the bounds of the variables are very important for building tight linear relaxations of MINLPs. Many details about branch-and-bound algorithms have not been dealt with in the previous explanation. For more details, including proofs of convergence, the reader is referred to Horst and Tuy (1990, Chapter IV).

The importance of bound propagation and cutting planes is...

Contributions and outline In Chapter 2, we investigate two classical algorithms for *convex MINLPs*, a subclass of MINLP in which all the functions appearing in nonlinear constraints are convex. These algorithms are Kelley's Cutting Plane algorithm and Veinott's Supporting Hyperplane algorithm. We show that the convergence of Veinott's algorithm follows from the convergence of Kelley's algorithm. The idea is to interpret Veinott's algorithm as Kelley's algorithm applied to a reformulation of the original problem. Such a reformulation only depends on the feasible region and not on functions used to represent it. Thus, we are able to extend the applicability of Veinott's algorithm to some problems with convex feasible region, but where constraint functions are not necessarily convex nor differentiable. Under a mild technical condition, Veinott's algorithm converges if the function are differentiable. To extend this result, we relax the differentiability assumption of the functions by introducing a notion of a generalized derivative which is enough to show the convergence of Veinott's algorithm.

In Chapter 3, we study in a more general setting the *separation problem*, namely, given a point $\bar{x}$ and a set S, find a valid linear cutting plane for S that separates $\bar{x}$, or show that none exists. In other words, if $A(S, \bar{x})$ is the set of all the answers of the separation problem, that is, all valid cuts for S that separate $\bar{x}$ from S, then the separation problem is to find an element of $A(S, \bar{x})$ or show that $A(S, \bar{x}) = \emptyset$. We show that given S and $\bar{x}$, there exists $\hat{S} \subseteq S$ such that $A(S, \bar{x}) = A(\hat{S}, \bar{x})$. The intuition of such a result is as follows. To ensure that a cutting plane is valid for a closed set S, it is enough to verify that it is valid for every vertex of S. However, in general, we want a cutting plane that separates a given point $\bar{x}$. Thus, to ensure validity of such a cut, it is enough to verify that it is valid for every vertex of S "near" $\bar{x}$. We use the concept of *visible points* of S from $\bar{x}$, $V_S(\bar{x})$, to formalize the meaning of "near" and show that $A(S, \bar{x}) = A(V_S(\bar{x}), \bar{x})$. We give a simple characterization of the visible points of S when S is the intersection of a quadratic constraint and a convex set. If S is the intersection of a polynomial constraint and a convex set, we provide an extended formulation for a relaxation of the visible points. As we will see, simple examples show that the visible points are not the smallest $\hat{S}$ such that $A(S, \bar{x}) = A(\hat{S}, \bar{x})$. Finally, we use the visible points to characterize the smallest $\hat{S}$ for different classes of sets.

Then, in Chapter 4, we focus on *intersection cuts*. Intersection cuts are an elegant technique to construct cutting planes that perfectly fits to LP-based approaches for MINLP. We show how to construct intersection cuts for general factorable MINLPs. The idea is to construct *concave underestimators* of a factorable function. Our approach is to mimic McCormick's procedure for building *convex underestimators*. Furthermore we propose a strengthening procedure for intersection cuts using *monoidal strengthening* in the presence of a single integer variable.

With the aid of the concave underestimators, we build so-called *S-free sets*, closed convex sets that do not contain any point of S in their interior, where S is normally the feasible region or a relaxation thereof. From an S-free set and a simplicial conic relaxation of the feasible region one can construct an intersection cut. As it turns out, the larger the S-free set the stronger the cut. Thus, it is natural to seek maximal S-free sets, that is, S-free sets that are not completely contained in any other S-free set. Although the constructions of Chapter 4 allow us to construct S-free sets, they are usually not maximal. In Chapter 5 we construct maximal S-free sets when S is given by a quadratic constraint.

In the remainder of this chapter we introduce our notation and general definitions that are used throughout the thesis. We explain, in a rather leisurely manner, more techniques in MINLP that are relevant for this thesis.

1.1 Mathematical Preliminaries

In this section, we introduce notation and some concepts that we use throughout the thesis. The reader is referred to the following references for some definitions and proofs of some of the claims made in this section without proof: Rockafellar (1970), Schrijver (1998) and Boyd and Vandenberghe (2004). We classify the concepts to make the reference easier.

Topology We will be working in $\mathbb{R}^n$. We denote its inner product between $x, y \in \mathbb{R}^n$ by $x^\mathsf{T} y$ and by $\|\cdot\|$ the euclidean norm. We denote by $B_r(x)$ and $D_r(x)$ the euclidean ball centered at x of radius r and its boundary, respectively. More precisely, $B_r(x) = \{y \in \mathbb{R}^n : \|y - x\| \leq r\}$ and $D_r(x) = \{y \in \mathbb{R}^n : \|y - x\| = r\}$.

Let $C \subseteq \mathbb{R}^n$. We denote the boundary, complement, closure, interior, and relative interior of C by ∂C, $(C)^c$, $\operatorname{cl} C$, $\operatorname{int} C$, and $\operatorname{ri} C$, respectively. Given $v \in \mathbb{R}^n$ and a set $C \subseteq \mathbb{R}^n$, we denote the distance between v and C by $\operatorname{dist}(v, C) = \inf_{x \in C} \|v - x\|$. Given two sets $A, B \subseteq \mathbb{R}^n$, the *Minkowski sum* of A and B is $\{a+b : a \in A, b \in B\}$ and we denote it by $A+B$. When A is a singleton, say $A = \{a\}$, we denote the sum by $a + B$. For a set of vectors $\{v^1, \dots, v^k\} \subseteq \mathbb{R}^n$, we denote by $\langle v^1, \dots, v^k \rangle$ the subspace generated by them.

Given some set $C \subseteq \mathbb{R}^n \times \mathbb{R}^m$, we denote by $\operatorname{proj}_x C$ the *projection* of C onto the x-space, that is, $\operatorname{proj}_x C = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^m, (x, y) \in C\}$. More generally, if H is a subspace of $\mathbb{R}^n$, we denote $\operatorname{proj}_H C$ the projection of C onto H.

Convex sets Given m points $x_1, \dots, x_m \in \mathbb{R}^n$ and given $\lambda_1, \dots, \lambda_m \in [0, 1]$ such that $\sum_{i=1}^m \lambda_i = 1$, the point $\sum_{i=1}^m \lambda_i x_i$ is said to be a *convex combination* of the points $x_1, \dots, x_m$. We say that C is *convex* if for every $x, y \in C$ and $\lambda \in [0, 1]$, $\lambda x + (1 - \lambda) y \in C$, that is, if for every pair of points in C every convex combination of them is in C. The *convex hull* of C is the smallest convex set that contains C, or equivalently the intersection of all convex sets containing C and is denoted by $\operatorname{conv} C$. The closure of the convex hull of C is denoted by $\overline{\operatorname{conv}}\, C$. The *extreme points* of a not necessarily convex set C are the points in C that cannot be written as convex combination of other points in C, and we denote them by $\operatorname{ext} C$. For example, if C is a square, then the extreme points are the vertices. If C is a disk, then the extreme points are all the points at the boundary. If C is this figure , then the two right vertices and all the points of the semi-circle at the left are extreme points. The beauty of the concept of *extreme points* is that those points are the only ones needed to describe the convex hull of a set.

A related concept is that of *exposed points*. When one optimizes a linear function over a set C, then an optimal solution, if one exists, is going to be at the boundary of C. The solution might be unique, for example, when optimizing in any direction over a circle. There might be multiple solutions, for example, when optimizing in the direction $(1, 0)$ over a square. Any $x_0 \in C$ such that there exists a linear function $\alpha^\mathsf{T} x$ for which x_0 is the *unique* solution of $\max_{x \in C} \alpha^\mathsf{T} x$ is called an *exposed point*. We denote the set of exposed points of C by $\exp C$. Every exposed point is an extreme point. However, not every extreme point is an exposed point. To see this, consider again $C =$. The two points where the semi-circle meets the straight part are extreme but not exposed.

The *gauge function* of a convex set C is $\phi_C(x) = \inf\{t : t > 0, \ \frac{x}{t} \in C\}$. The gauge function is a sort of distance measured by C. It measures what is the minimum that we have to scale C so that x is at its boundary.

Given a closed set S, a convex set C is said to be *S-free* if its interior does not contain any point of S. In other words, C is S-free if $S \cap \operatorname{int} C = \emptyset$. Let C be an S-free set. We say that C is *maximal* S-free if it holds that for every convex S-free set K, if $C \subseteq K$, then $C = K$.

Inequalities Let $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$. The set $\{x \in \mathbb{R}^n : \alpha^\mathsf{T} x = \beta\}$ is called an *affine subspace* and we say that α is its *normal*. The set $\{x \in \mathbb{R}^n : \alpha^\mathsf{T} x \leq \beta\}$ is a *half-space*. Both are convex. In general, a closed convex set can be written as the intersection of an arbitrary number of half-spaces. Usually, instead of writing the half-space as a set we just write $\alpha^\mathsf{T} x \leq \beta$. We say that $\alpha^\mathsf{T} x \leq \beta$ is *valid* or a *valid inequality* for C if $C \subseteq \{x \in \mathbb{R}^n : \alpha^\mathsf{T} x \leq \beta\}$. If $\alpha^\mathsf{T} x \leq \beta$ is a valid inequality for C and $\bar{x} \notin C$ is such that $\alpha^\mathsf{T} \bar{x} > \beta$, we say that $\alpha^\mathsf{T} x \leq \beta$ *separates* $\bar{x}$ from C. If $\alpha^\mathsf{T} x \leq \beta$ is a valid inequality for C and it is tight, that is, there exists a $y \in C$ such that $\alpha^\mathsf{T} y = \beta$, we say that $\alpha^\mathsf{T} x \leq \beta$ is a *supporting hyperplane* of C, or that it *supports* C. A closed convex set can be written as the intersection of its supporting hyperplanes. If the number of hyperplanes needed to describe a convex set is finite, then the convex set is called a *polyhedron*.

Cones A *cone* is a set $C \subseteq \mathbb{R}^n$ with the following property. If $x \in C$ and $\lambda \geq 0$, then $\lambda x \in C$. A cone is *pointed* if it has an extreme point, in which case this extreme point is called *apex*. Given m points $x_1, \ldots, x_m \in \mathbb{R}^n$ and $\lambda_1, \ldots, \lambda_m \geq 0$, the point $\sum_{i=1}^m \lambda_i x_i$ is said to be a *conic combination* of the points $x_1, \ldots, x_m$. In the context of cones, the *extreme rays* play the role of extreme points. A *ray* is a set of the form $\{\lambda x : \lambda \geq 0\}$ and we call it the *ray generated by* x. If C is a cone and $x \in C$, the ray generated by x is contained in C. We say that the ray generated by $x \in C$ is an extreme ray if x cannot be written as a conic combination of other points of C. Note that this is the same as saying that neither x nor any positive scaling of it can be written as a conic combination of other points of C. We say that a set $K \subseteq \mathbb{R}^n$ is a *translated cone* if there exist a cone C and $x \in \mathbb{R}^n$ such that $K = C + x$. A cone in $\mathbb{R}^n$ is said to be *simplicial* if it has exactly n extreme rays.

Every unbounded convex set contains a (translated) cone. The *recession cone* of a convex set C, denoted by $\operatorname{rec}(C)$, is the largest cone K such that $C + K = C$. In other words, $\operatorname{rec}(C)$ is the largest cone that can be translated to be completely contained in C. It is possible that a direction d and its opposite, $-d$, are both in the recession cone of C. The set of all such directions, that is, $\operatorname{rec}(C) \cap \operatorname{rec}(-C)$ is called the *lineality space* of C and is denoted by $\operatorname{lin}(C)$. It is the largest subspace, L, such that $L + C = C$. Note that a convex cone is pointed if and only if its lineality space is $\{0\}$.

Convex functions Let $g : X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function. The *epigraph* of g is the set of all points above the graph, $\operatorname{epi} g = \{(x, z) \in \mathbb{R}^{n+1} : z \geq g(x)\}$.

We say that g is *convex* in $C \subseteq X$ if C is convex and for every $x, y \in C$ and $\lambda \in [0, 1]$, $g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y)$. Equivalently, g is convex if its epigraph is convex. We say that g is *concave* when $-g$ is convex and every concept we define for convex functions has its counterpart for concave functions.

When g is differentiable and convex in C we have that $g(y) + \nabla g(y)^{\mathsf{T}}(x - y) \leq g(x)$ for every $x, y \in C$. For a given y, this inequality means that the tangent hyperplane at y of the graph of g, $g(y) + \nabla g(y)^{\mathsf{T}}(x - y)$, is always below the function. Equivalently, it means that the epigraph of $x \mapsto g(y) + \nabla g(y)^{\mathsf{T}}(x - y)$ is a valid inequality for the epigraph of g. Actually, since the inequality is tight when $x = y$, the inequality supports $\operatorname{epi} g$. In general, convex functions do not need to be differentiable, however, the epigraph is still convex and it still has supporting hyperplanes. A *subgradient* of a convex function is the normal of a supporting hyperplane, when the inequality is written in a similar form to the differentiable case. Specifically, a vector v is a subgradient of g at y if $g(y) + v^{\mathsf{T}}(x - y) \leq g(x)$ for every $x \in C$. The set of all subgradients of g at y is called the *subdifferential* of g at y and its denoted by $\partial g(y)$. Thus,

$$\partial g(y) = \{v \in \mathbb{R}^n \,:\, g(y) + v^{\mathsf{T}}(x - y) \leq g(x) \forall x \in C\}.$$

For example, $g(x) = |x|$ is convex, not differentiable at 0, and $\partial g(0) = [-1, 1]$.

A function g is *positively homogeneous* if $g(\lambda x) = \lambda g(x)$ for every $\lambda \geq 0$ and all x. A function g is *subadditive* if $g(x + y) \leq g(x) + g(y)$. A function is *sublinear* if it is positively homogeneous and subadditive. Equivalently, g is sublinear if it is positively homogeneous and convex. The epigraph of a sublinear function from $\mathbb{R}^n$ to $\mathbb{R}$ is a closed convex cone. We say that a convex set is *represented* by a sublinear function g if $C = \{x \,:\, g(x) \leq 1\}$.

Given a convex function $g : C \to \mathbb{R}$, $g(x) \leq 0$ is called a *convex constraint*. We have that for any $\bar{x} \in C$ and $v \in \partial g(\bar{x})$,

$$g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x}) \leq 0 \tag{1.2}$$

is a valid inequality for $g(x) \leq 0$. Thus, if $\bar{x} \in C$ violates the convex constraint, that is, $g(\bar{x}) > 0$, then (1.2) separates $\bar{x}$ from $g(x) \leq 0$. To see this, recall that $g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x}) \leq g(x)$ for every $x \in C$. In particular, if x satisfies the constraint, then $g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x}) \leq g(x) \leq 0$, which shows the validity of (1.2). Evaluating (1.2) at $\bar{x}$ yields $g(\bar{x}) \leq 0$ from where we conclude that $\bar{x}$ does not satisfy (1.2). We call such inequalities *gradient cutting planes*, or *gradient cuts* for short, because when g is differentiable v can only be the gradient $\nabla g(\bar{x})$.

If $g : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is a function and $C \subseteq X$ is convex, then we denote by g_C^{vex} a *convex underestimator* of g over C. This means that $g_C^{vex} : C \to \mathbb{R}$ is a convex function and underestimates g on C, that is, $g_C^{vex}(x) \leq g(x)$ for all $x \in C$. Similarly, we define a *concave overestimator*.

Matrices A matrix $M \in \mathbb{R}^{n\times n}$ is *symmetric* if $M = M^\mathsf{T}$. We say that a symmetric matrix M is *positive semi-definite* if $x^\mathsf{T} M x \geq 0$ for every $x \in \mathbb{R}^n$. Given an integer n, we denote by $\mathcal{S}_+^n$ the cone of positive semi-definite matrices of size $n \times n$. A matrix M is *copositive* if $x^\mathsf{T} M x \geq 0$ for every $x \in \mathbb{R}_+^n$. A $k \times k$ *submatrix* of a matrix M is a matrix formed by the deleting all but k columns and k rows of M. The *rank* of a matrix M is the number of linearly independent columns, which is the same as the number of linearly independent rows, and we denote it by $\operatorname{rk} M$.

General notation Given an interval $I \subseteq \mathbb{R}$ and an arbitrary set $A \subseteq \mathbb{R}^n$ we denote by IA the set $\{\lambda x \,:\, \lambda \in I, x \in A\}$. Likewise, for $x \in \mathbb{R}^n$, $Ix := \{\lambda x \,:\, \lambda \in I\}$.

Given $n \in \mathbb{N}$, we denote by $[n] = \{1, \ldots n\}$. If A and B are sets and A is finite, we denote by B^A the set $B^{|A|}$, where $|A|$ is the cardinality of A.

1.2 Intersection Cuts

Intersection cuts are the topic of chapters 4 and 5. In this section, we give a brief introduction to intersection cuts.

The history of intersection cuts and S-free sets dates back to the 60's. They were originally introduced in the nonlinear setting by Tuy (1964) for the problem of minimizing a concave function over a polytope. Later on, they were introduced in integer programming by Balas (1971) and have been largely studied since. The more modern form of intersection cuts deduced from an arbitrary convex S-free set is due to Glover (1973), although the term S-free was coined by Dey and Wolsey (2010).

We illustrate the idea with the following integer program

$$\max\{-12x+5y \,:\, x+4y \leq 17,\ -4x+y \leq -3,\ 5x-6y \leq 1,\ x, y \in \mathbb{Z}\}, \quad (1.3)$$

depicted in Figure 1.1. The LP relaxation solution is $\bar{x} = (\frac{29}{17}, \frac{65}{17})$. The nearest feasible point is at a distance of $\sqrt{13/17}$, and so there is no feasible point in the interior of the ball centered at $\bar{x}$ of radius $\sqrt{13/17}$. If $S = \{(x, y) \in \mathbb{Z}^2 \,:\, x + 4y \leq 17,\ -4x + y \leq -3,\ 5x - 6y \leq 1\}$, then this ball is an S-free set.

The LP solution is the apex of a cone whose extreme rays are the edges of the polyhedron adjacent to the LP solution. Now, consider the points where the extreme rays of the cone intersect the ball and build the hyperplane (in

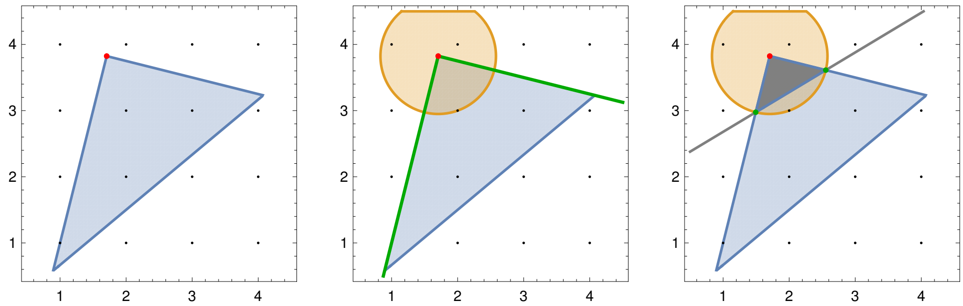


Figure 1.1: The left plot shows the integer points in black, the LP relaxation of (1.3) in blue, and the optimal LP solution in red. The middle plot highlights the ball centered at the optimal LP solution of radius equal to the distance between the optimal LP solution and the nearest feasible point in orange. It also shows the extreme rays of the conic relaxation starting at the optimal LP solution in green. The right plot shows the intersection points of the ball with the cone in green, the intersection cut in gray, and the region cutoff by the cut also in gray.

this case just a line) that goes through those points. This hyperplane defines a valid inequality that separates the LP solution from S. The reason why it is valid is that the region of the LP relaxation cutoff by the inequality is completely contained inside the ball. This happens because the ball is a convex set. As the ball does not contain *any* feasible point in its interior, the cut must be valid. Such a cutting plane is an *intersection cut*.

In general, there are three ingredients for the construction of intersection cuts. First, the set of (or a relaxation of the) feasible points S. Second, a simplicial cone that contains the feasible region and whose apex is the LP solution (or the point to separate). Third, an S-free set C that contains the LP solution in its interior. We ask for the cone to be simplicial so that the intersection of its extreme rays with C defines a unique hyperplane.

Note that the larger the S-free set, the better the intersection cut. The intuition is that if K and C are S-free and K is larger than C, then the intersection of an extreme ray of the cone with K will be farther away, and thus the cut will be deeper. This is illustrated in Figure 1.2 where we compare the cut obtained in the above example with the intersection cut deduced by using as S-free set the largest ball centered at the LP solution that does not include any integer point in its interior.

How can we build a simplicial cone whose apex is the LP solution and

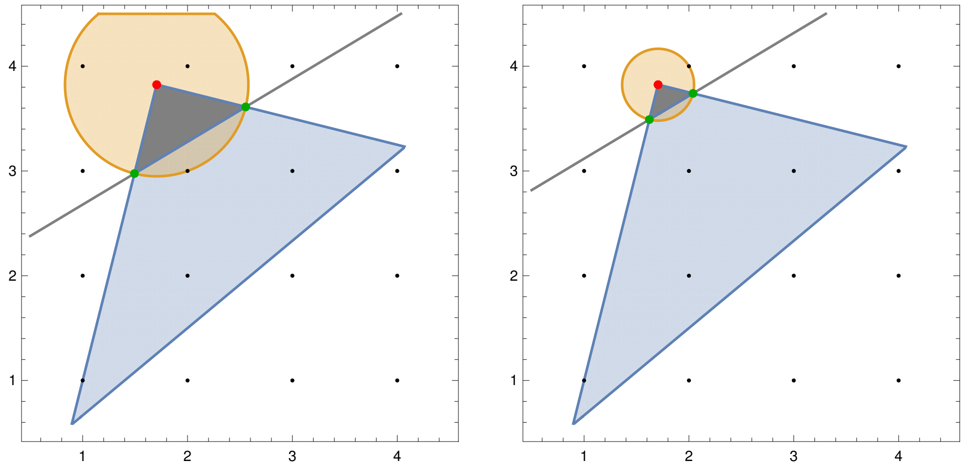


Figure 1.2: The left plot shows the intersection cut for (1.3) obtained above. The right plot shows the intersection cut obtained from the S-free set given by a $\mathbb{Z}^2$-free ball.

that contains the whole feasible region? Luckily, such a cone appears quite naturally when we solve the LP using the simplex algorithm. Consider a linear program $\max\{c^{\mathsf{T}}x \ : \ Ax \leq b\}$. The simplex algorithm starts at a vertex of $Ax \leq b$ and iteratively moves to a neighbor vertex with better objective value if there is one. If there is none, then the vertex is optimal. A vertex is a feasible point defined by the intersection of n independent hyperplanes among the m ones in $Ax \leq b$. Ignoring all but the n constraints that define a vertex, yields a simplicial cone whose apex is the vertex and contains the whole LP, see the middle plot in Figure 1.1. When an optimal solution is obtained, one can find out the n constraints that the simplex algorithm considered in order to define the solution. Therefore, intersection cuts are readily available in LP-based branch and bound algorithms if we are able to construct an S-free set that contains the LP solution in its interior.

We will now present a more algebraic deduction of intersection cuts whose advantage is that it admits a generalization of intersection cuts. As it turns out, this generalization is only relevant when the S-free set is unbounded. We will also give a geometric characterization of the generalization and show that in this case it no longer holds that larger S-free sets yield better cuts.

The simplex algorithm is usually presented using the so-called *standard from* of an LP, namely, $\max\{c^{\mathsf{T}}x \ : \ Ax = b, \ x \geq 0\}$. The advantage is that the algebraic description of the algorithm is simpler, but certainly the geometric intuition is obfuscated. But the story is the same. We have n variables and $m+n$ constraints, m from $Ax=b$ and n from $x \geq 0$. Since m of these constraints are equality, we simply need $n-m$ more to define a point, assuming, as we are, that the equality constraints are linearly independent. These $n-m$ can only come from $x \geq 0$. Thus, any vertex will have $n-m$ variables fixed to 0 and the others will be the unique solution to the remaining system of equations. As above, not every selection of $n-m$ constraints from $x \geq 0$ yields a vertex, but some do. In particular, if a selection does, then the matrix describing the remaining system is invertible. That is, the columns of A associated to the m variables *not* fixed to 0 after setting $n-m$ constraints from $x \geq 0$ to equality are linearly independent. These variables are called *basic* variables, their indices are called a *basis*, and the remaining variables are called *non-basic*.

Let B be a basis and let N be the indices of the non-basic variables. We can partition the system $Ax=b$ into basic and non-basic variables. For this we introduce the following notation: if $I \subseteq \{1,\dots,n\}$, then A_I represents the columns of A indexed by I, while x_I the subvector of variables indexed by I. Then $Ax=b$ is equivalent to $A_B x_B + A_N x_N = b$. From the above discussion A_B is an invertible matrix, thus $Ax=b$ is equivalent to $x_B = A_B^{-1}b - A_B^{-1}A_N x_N$. This is the so-called *tableau.*[1] There is a lot of important information in the tableau. In particular, the apex of the simplicial cone is $(x_B, x_N) = (A_B^{-1}b, 0)$, while its extreme rays are $(x_B, x_N) = (-A_B^{-1}A_N e_j, e_j)$ for $j \in N$. Note that although $x \in \mathbb{R}^n$, the feasible points are in an $n-m$ dimensional space, assuming A has full rank. So the cone is actually simplicial only in the solution space, as it has $n-m$ rays. Thus, it gets a bit more complicated to picture this, but the beauty is that we can deduce the intersection cuts directly from the tableau.

Consider an optimization problem $\mathcal{P}$ and assume that the tableau of an LP relaxation of it is $x = f + Rs$, where x are the basic and s the non-basic variables. Let S be a closed set such that for every feasible solution (x,s) of $\mathcal{P}$, it holds that $x \in S$. Furthermore, assume that $f \notin S$, that is, the optimal LP solution $(f,0)$ is not feasible. Let C be an S-free set such that $f \in \operatorname{int} C$. Let us assume that C is given by $C = \{x : \phi(x-f) \leq 1\}$, where ϕ is sublinear. Now, any $s \geq 0$ defines an $x = f + Rs$ and $\phi(x-f) = \phi(Rs)$. Thus, as long as $\phi(Rs) < 1$, $x \in C$ and x itself cannot be feasible. We conclude that if (x,s) is to be feasible, then $\phi(Rs) \geq 1$, that is, $\phi(Rs) \geq 1$ is a valid (nonlinear) inequality. To make it linear, we use the sublinearity of ϕ and the

[1] The tableau also has a row with the objective function, but we omit it as it is not relevant for our current discussion.

non-negativity of the variables. Indeed,

$$1 \leq \phi(Rs) = \phi(\sum_j R_j s_j) \leq \sum_j \phi(R_j s_j) = \sum_j \phi(R_j) s_j,$$

where the second inequality follows from the subadditivity of ϕ and the last equality follows from the positive homogeneity of ϕ and the non-negativity of s. Such a function ϕ is also called a *cut generating function*, since evaluating it at the given rays is sufficient enough to obtain the cut's coefficients.

When ϕ is the gauge of $C - f$, then the cut above corresponds to the intersection cut described geometrically above. Indeed, the points $s^i = \frac{1}{\phi(R_i)} e_i$, assuming $\phi(R_i) > 0$, satisfy the inequality $\sum_j \phi(R_j) s_j \geq 1$ with equality. These points define $x^i = f + \frac{1}{\phi(R_i)} R_i$ and satisfy $\phi(x^i - f) = 1$. This means that all x^i are on the boundary of C. In other words, the hyperplane $\sum_j \phi(R_j) s_j \geq 1$ passes through the $n - m$ points (x^i, s^i), which correspond to the intersection of the $n-m$ rays (R_i, e_i) with the boundary of the $S \times \mathbb{R}^{n-m}$-free set, $C \times \mathbb{R}^{n-m}$. As mentioned before, note that the LP is $Ax = b,\ x \geq 0$ so, even though $x \in \mathbb{R}^n$ and we need n points to define a hyperplane, the feasible region lives in the translated subspace $Ax = b$. Therefore, we are working on $\mathbb{R}^{n-m}$ embedded in $\mathbb{R}^n$ and only $n - m$ points define a unique hyperplane in the space that we are working on.

A sublinear function other than the gauge, if it exists, will yield better cut coefficients, thus, a better cut. As it turns out, if $C = \{x : \phi(x - f) \leq 1\}$ for some sublinear function ϕ and $f + R_i\mathbb{R}_+$ is a ray that is *not* in the interior of the recession cone of C, then $\phi(R_i)$ is equal to the gauge of $C - f$ at R_i. That is, the only way of improving on a coefficient is that $f + R_i\mathbb{R}_+ \in \operatorname{int}\operatorname{rec}(C)$. In other words, the possibility of improving the cut coefficients can only occur when C is unbounded and, furthermore, when a ray of the simplicial cone is in the interior of the recession cone of C. Note that when this occurs, then the gauge of $C - f$ at R_i is 0 and if an improvement is possible, then the coefficient must be negative. A negative coefficient can never be achieved with the gauge as the gauge is always non-negative.

This phenomenon was first observed by Glover (1974). Glover interpreted the negative coefficient as moving in the negative direction of the ray instead of the positive one.

Here we provide an interpretation of the negative edge extension. Consider the following set $S = \{(x, y) \in \mathbb{R}^2_+ : x - y \geq 2 \ \vee \ x - 5y \geq 1\}$, see Figure 1.3. Clearly, a maximal S-free set is $C = \{(x, y) \in \mathbb{R}^2 : x - y \leq 2,\ x - 5y \leq 1\}$. The cone with apex 0 and rays e_1 and e_2 is simplicial and contains the whole feasible region, so we use it to generate the intersection cut. The intersection

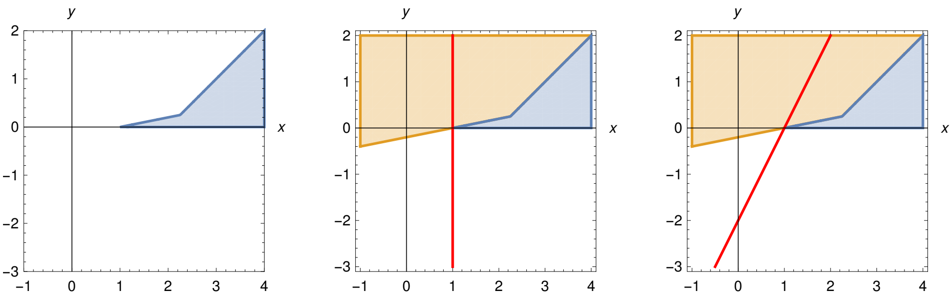


Figure 1.3: The left plot shows the set S in blue. The middle plot shows the set S in blue and C in orange with the intersection cut obtained by the gauge. The right plot shows S, C and the cut obtained with ϕ.

cut obtained from the simplicial cone and the C is $x \geq 1$. Indeed, the gauge of C, ϕ_C, satisfies $\phi_C(e_1) = 1$, since $e_1 \in C$, and $\phi_C(e_2) = 0$, as $\lambda e_2 \in C$ for every $\lambda \geq 0$. As it turns out, $C = \{(x, y) \in \mathbb{R}^2 \,:\, \phi(x, y) \leq 1\}$ for $\phi(x, y) = \max\{\frac{x-y}{2},\ x - 5y\}$. Note that $\phi(e_2) = \max\{-\frac{1}{2},\ -5y\} = -\frac{1}{2}$. Thus, ϕ is not the gauge and, more importantly, the cut $x - \frac{1}{2}y \geq 1$ is valid.

The interpretation of the coefficients of the intersection cut obtained by the gauge is as follows. If we move along the ray e_1, then we hit the boundary of C at $1e_1$, thus the cut coefficient is $\frac{1}{1}$. Instead, if we move along e_2, then we "hit" the boundary of C at "∞e_2" and the cut coefficient is $\frac{1}{\infty} = 0$.

However, we can actually tilt this cut to make it stronger. How much can we tilt it? Well, we can tilt as long at the cut off region is inside C. The tilted cut intersects the y axis at some negative point. The higher the point the stronger the cut, see Figure 1.4. The coefficient of the intersection cut obtained by the sublinear function ϕ corresponds to the tilting whose intersection with the y axis is the lowest point at which a supporting valid inequality for C intersects the y axis. In this case, such a point is $(0, -2)$ and so the cut coefficient is $-\frac{1}{2}$.

Something looks off, though, the cut is not the best possible. How can we achieve a better cut? Consider the weaker S-free set $K = \{(x, y) \in \mathbb{R}^2 \,:\, x - y \leq 1,\ x - 5y \leq 1\ \}$. We have that $K = \{(x, y) \in \mathbb{R}^2 \,:\, \psi(x, y) \leq 1\}$, where $\psi(x, y) = \max\{x - y,\ x - 5y\}$. Now the intersection cut is $x - y \geq 1$ and it cannot be strengthened anymore as it defines a facet of $\overline{\operatorname{conv}}(S)$.

What happened? By moving the facet $x - y \leq 2$ of C to the left until $x - y \leq 1$, we did not change the intersection point of the ray e_1. However, we did make the lowest point at which a valid inequality for K intersects the

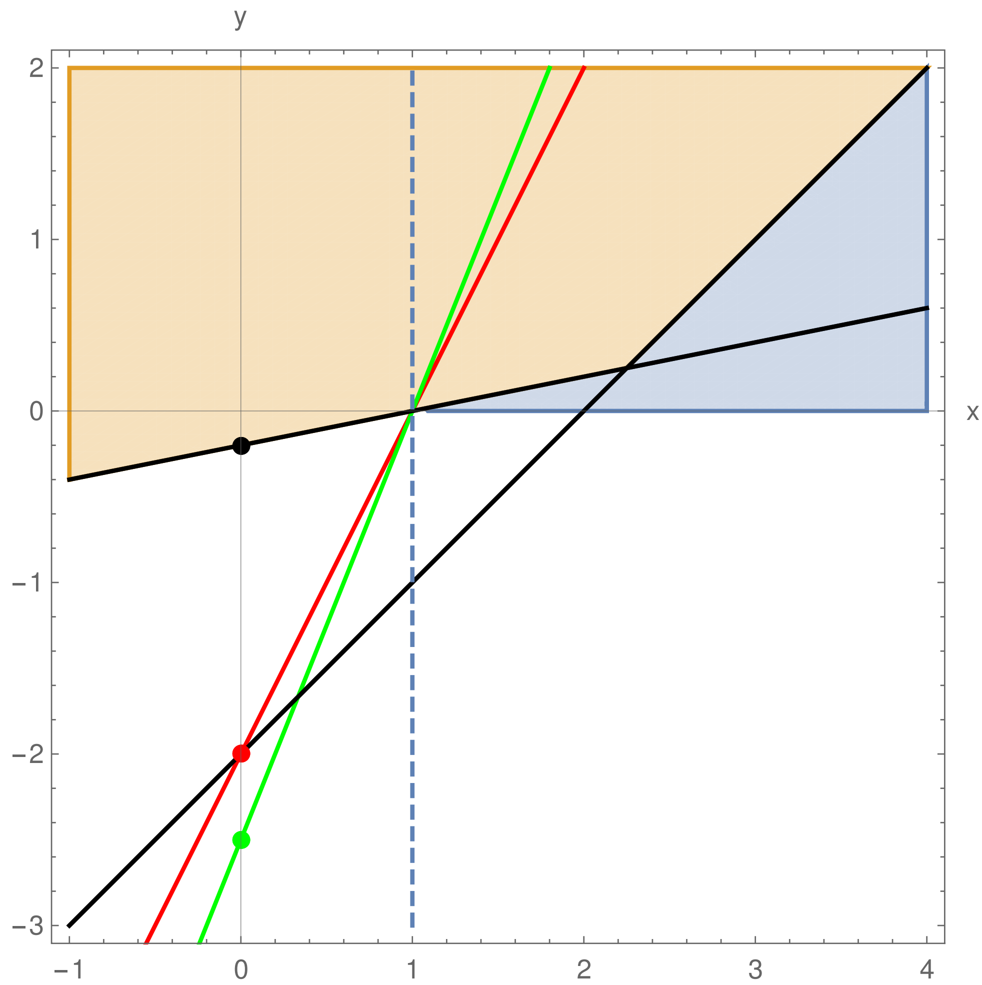


Figure 1.4: The plot shows the set S in blue and the set C in orange. We see the intersection cut obtained with the gauge (dashed), a better tilted cut that intersects the y axis at -2.5 (green), and the intersection cut obtained with $\phi(x, y)$ (red). The higher the intersection with the y axis, the better the cut. Also, the red cut intersects the y axis at the lowest point that a supporting valid inequality of C intersects the y axis. Supporting valid inequalities of C intersect the y axis between the black dot and the red dot.

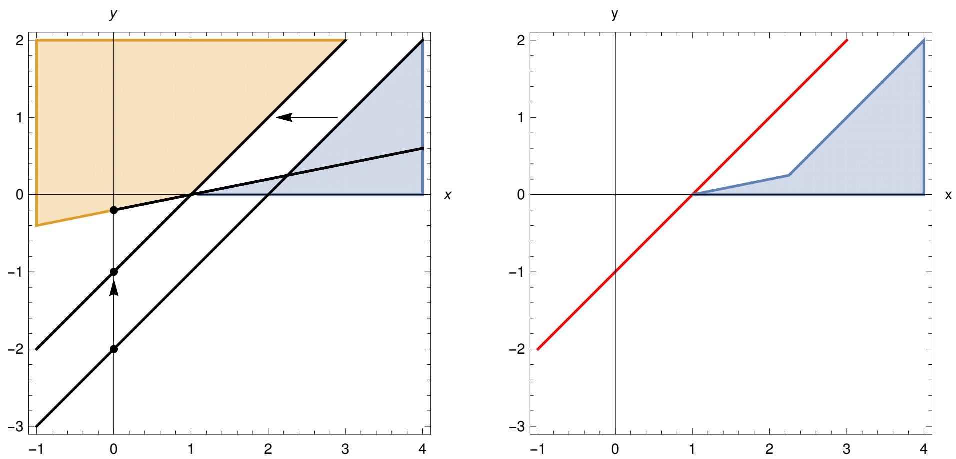


Figure 1.5: The left plot shows how shrinking the S-free set moves the lowest intersection with the y axis up. The right plot shows the final intersection cut, which defines the closure of the convex hull of S.

y axis higher, thus the cut is stronger. For an illustration see Figure 1.5.

The above is an example that larger S-free sets are *not* always better when one builds intersection cuts with sublinear functions other than the gauge. Let C be an S-free set. When the ray actually intersects the boundary of C, it is clear that if we extend C in that direction, then the intersection point is going to be farther away as we discussed above and illustrated in Figure 1.2. However, the interpretation of the cut coefficient with a sublinear function is a bit more involved and uses more global information. Indeed, making C larger in some direction will affect which inequalities are valid and so it can have a (negative) effect on the cut coefficient for rays that are contained inside C. This is what the above example illustrates.

We refer the reader to Conforti et al. (2011b) and Conforti et al. (2015) for more details on intersection cuts.

1.3 Duality

In chapters 2 and 5, we mention and use Slater's condition, respectively. This is a condition that ensure strong duality of convex problems. Here we give a brief introduction to duality aiming at explaining Slater's condition from a geometrical point of view.

In this section we give a brief introduction to intersection cuts. Consider a linear program $\max\{c^\mathsf{T}x : Ax \le b\}$. Suppose its optimal value is z. This means that $c^\mathsf{T}x \le z$ for every x such that $Ax \le b$. In fact, it is the tightest valid inequality for $Ax \le b$ with normal c. Thus, instead of solving $\max\{c^\mathsf{T}x : Ax \le b\}$ directly, one can try to find the tightest valid inequality for $Ax \le b$ with normal c. Alternatively, one can think of it as finding the best upper bound on the value that $c^\mathsf{T}x$ can achieve over $Ax \le b$. But how can we do this?

It should, of course, be possible to deduce the inequality $c^\mathsf{T}x \le z$ just from the information in $Ax \le b$. For example, consider $\max\{3x + y : 4x - y \le 2,\ -x + 3y \le 5\}$. The optimal solution is obtained at $(\bar{x}, \bar{y}) = (1, 2)$ and has a solution value of 5. Thus, the inequality $3x + y \le 5$ is valid for $\{(x, y) : 4x - y \le 2,\ -x + 3y \le 5\}$. Indeed, we can deduce it from $4x - y \le 2$ and $-x + 3y \le 5$ by multiplying the first inequality by 10, the second one by 7, and then adding them up. This yields $33x + 11y \le 55$, which is the same as $3x + y \le 5$.

It is a fundamental result in linear programming, called Farkas' lemma, that if $Ax \le b$ is non-empty, then every valid inequality can be deduced by considering a conic combination of the constraints (Ziegler, 1995). Why the non-empty assumption? The problem is that every inequality is valid when $Ax \le b$ is empty, but to be able to write every inequality as a conic combination of $Ax \le b$ one needs enough inequalities, more than the ones needed to describe an empty set. For example, $\{(x, y) \in \mathbb{R}^n : x \le 0,\ x \ge 1\}$ is clearly empty, thus the inequality $y \le 0$ is valid. However, there is no way of building that inequality by taking positive linear combinations of $x \le 0$ and $-x \le -1$.

With Farkas' lemma we can write the problem of finding the tightest valid inequality for $Ax \le b$ with normal c as follows. Every valid inequality is given by $\mu^\mathsf{T}Ax \le \mu^\mathsf{T}b$ for some $\mu \ge 0$. The normal of the inequality has to be c, thus we have the constraint $\mu^\mathsf{T}A = c$ and it has to be the tightest, that is, the right hand side, $\mu^\mathsf{T}b$ has to be the smallest. Thus, when $Ax \le b$ is feasible, we have

$$\min\{\mu^\mathsf{T}b : \mu^\mathsf{T}A = c,\ \mu \ge 0\} = \max\{c^\mathsf{T}x : Ax \le b\}.$$

The problem on the left hand side is called the *dual* problem and the one in the right hand side, the *primal*.

There are many ways of deducing the dual problem. A standard way is through *Lagrange duality*. The idea is as follows. The problem $\max\{c^\mathsf{T}x : Ax \le b\}$ can be written as an unconstrained problem using $I_{\mathbb{R}^m_-}$, the indicator function of $\mathbb{R}^m_-$,

$$I_{\mathbb{R}^m_-}(y) = \begin{cases} 0, & \text{if } y \leq 0 \\ +\infty, & \text{otherwise.} \end{cases}$$

We have $\max\{c^\mathsf{T} x \,:\, Ax \leq b\} = \max c^\mathsf{T} x - I_{\mathbb{R}^m_-}(Ax - b)$. The dual tries to bound the optimal value. One way to find a bound is to find an overestimator of the objective function. We have that $I_{\mathbb{R}^m_-}(y) \geq \mu^T y$ for any $\mu \in \mathbb{R}^m_+$. Indeed, if $y \not\leq 0$, then the left hand side is $+\infty$, so the inequality holds. Otherwise, the left hand side is 0, while the right one is non-positive, so the inequality holds. Therefore, for any $\mu \geq 0$,

$$\max\{c^\mathsf{T} x \,:\, Ax \leq b\} \leq \sup_x c^\mathsf{T} x - \mu^\mathsf{T}(Ax - b).$$

We can now take the best $\mu \geq 0$ to get

$$\max\{c^\mathsf{T} x \,:\, Ax \leq b\} \leq \inf_{\mu \geq 0} \sup_x c^\mathsf{T} x - \mu^\mathsf{T}(Ax - b).$$

The function $L(x, \mu) = c^\mathsf{T} x - \mu^\mathsf{T}(Ax - b)$ is called the *Lagrangian* function, $\theta(\mu) = \sup_x L(x, \mu)$ is the *Lagrangian dual function*, and $\inf_{\mu \geq 0} \theta(\mu)$ is the *(Lagrangian) dual* problem of $\max\{c^\mathsf{T} x \,:\, Ax \leq b\}$. We have that

$$\theta(\mu) = \sup_x c^\mathsf{T} x - \mu^\mathsf{T}(Ax - b) = \sup_x (c - A^\mathsf{T}\mu)^\mathsf{T} x + \mu^\mathsf{T} b = \begin{cases} \mu^\mathsf{T} b, & \text{if } c - A^\mathsf{T}\mu = 0 \\ \infty, & \text{otherwise.} \end{cases}$$

Thus, the Lagrangian dual is

$$\inf\{\mu^\mathsf{T} b \,:\, A^\mathsf{T}\mu = c,\ \mu \geq 0\},$$

which is the same as the linear programming dual.

The advantage of Lagrangian duality is that the deduction of the dual generalizes to other types of problems. For example, consider $\max\{e^x \,:\, x^2 \leq y,\ y \leq 1\}$. The reasoning in the linear case was to find valid inequalities that can be deduced from the constraints. Luckily, Farkas' lemma tells us how these valid inequalities look like and so we could write an optimization problem to find the tightest one. Here, it is not clear how the valid inequalities actually look like. However, Lagrangian duality still yields a dual.

The disadvantage, though, is that it will not be clear that the bound provided by the Lagrangian dual is equal to the optimal value of the primal. In fact, even if the primal is convex there can be a positive difference between the optimal values of the primal and dual problems. We refer to the optimal value of the primal as *primal value* and the optimal value of the dual es *dual* *value*. When the primal and dual values coincide, we say that *strong duality* holds. The difference between the primal and dual values is called *duality gap*.

To see that there are convex problems with positive duality gap, let us compute the Lagrangian dual of $\max\{-e^{-x} : \sqrt{x^2+y^2} \le y\}$. The Lagrangian function is $L(x,y,\mu) = -e^{-x} - \mu(\sqrt{x^2+y^2} - y)$. The Lagrangian dual function is $\theta(\mu) = \sup_{x,y} -e^{-x} - \mu(\sqrt{x^2+y^2} - y)$. By Cauchy-Schwarz inequality $y \le \sqrt{x^2+y^2}$ for all $x, y \in \mathbb{R}$, so $-\mu(\sqrt{x^2+y^2} - y) \le 0$ for every $x, y \in \mathbb{R}^2$ and $\mu \ge 0$. Thus, $\theta(\mu) \le \sup_{x,y} -e^{-x} = 0$.

Let us show that actually $\theta(\mu) = 0$ for all $\mu \ge 0$. Notice that

$$-e^{-x} - \mu(\sqrt{x^2+y^2} - y) = -e^{-x} - \mu\frac{x^2}{\sqrt{x^2+y^2}+y}.$$

Replacing y by e^x above and computing the limit as $x \to \infty$ we obtain

$$\lim_{x\to\infty} -e^{-x} - \mu\frac{x^2}{\sqrt{x^2+e^{2x}}+e^x} = 0.$$

Thus, $\theta(\mu) = 0$ for every $\mu \ge 0$.

However, the primal's feasible region is $\{0\} \times \mathbb{R}_+$ and its optimal value is, thus, $-e^0 = -1$.

To understand why this could happen, let us interpret the dual from a more geometric point of view. For this, let us abstract the problem a bit. Consider $\max\{f(x) : g_i(x) \le 0\}$. The Lagrangian dual function is then $\theta(\mu) = \sup_x f(x) - \sum_i \mu_i g_i(x)$. Thus, we have that $f(x) - \sum_i \mu_i g_i(x) \le \theta(\mu)$ for every x. An enlightening way of interpreting this inequality is to see it as a valid inequality of a set. Indeed, the inequality is saying that the hyperplane $y_0 - \sum_i \mu_i y_i \le \theta(\mu)$ is valid for the set $\Phi(\mathbb{R}^n) = \{(f(x), g_1(x), \ldots, g_m(x)) : x \in \mathbb{R}^n\}$, where $\Phi(x) = (f(x), g_1(x), \ldots, g_m(x))$. Thus, we can interpret the Lagrangian dual function as a function that given $\mu \ge 0$, finds the best right-hand side of a valid inequality with normal $(1, -\mu)$ for $\Phi(\mathbb{R}^n)$. Then, the Lagrangian dual problem seeks the normal $(1, -\mu)$ such that the valid inequality with that normal has the best (smallest in this case) right-hand side.

So, why do we have a positive duality gap for $\max\{-e^{-x} : \sqrt{x^2+y^2} \le y\}$? To answer this question we need to understand how $\Phi(\mathbb{R}^2)$ looks when $\Phi(x,y) = (-e^{-x}, \sqrt{x^2+y^2} - y)$. Figure 1.6 shows $\Phi([-\frac{1}{2},\frac{1}{2}] \times [-\frac{1}{2},\frac{1}{2}])$ and $\Phi([-\frac{1}{2},5] \times [-\frac{1}{2},150])$. One can prove that $\Phi(\mathbb{R}^2) = ((-\infty,0) \times (0,+\infty)) \cup \{(-1,0)\}$. From here we see that for every $\mu \ge 0$, the tightest valid inequality for $\Phi(\mathbb{R}^2)$ with normal $(1, -\mu)$ is $y_0 - \mu y_1 \le 0$. In other words, $\theta(\mu) = 0$ for every $\mu \ge 0$ as we saw above.

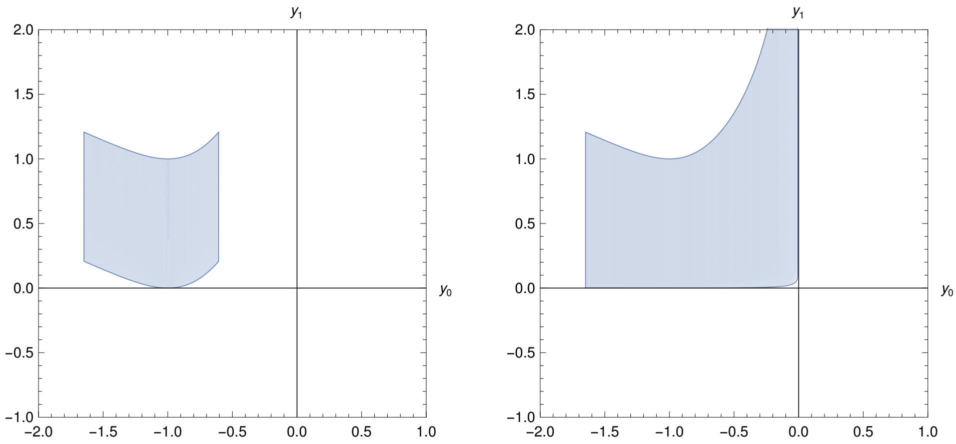


Figure 1.6: The left plot shows $\Phi([-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}])$ and the right one shows $\Phi([-\frac{1}{2}, 5] \times [-\frac{1}{2}, 150])$, where $\Phi(x, y) = (-e^{-x}, \sqrt{x^2 + y^2} - y)$.

When can we ensure that strong duality holds? Consider again $\max\{f(x) : g_i(x) \leq 0\}$ and let p^* be the optimal value. Assume that f is concave and the g_i are convex and notice that $y_0 - \sum_i \mu_i y_i \leq \theta$ is a valid inequality for $\Phi(\mathbb{R}^n)$ with $\mu \geq 0$, if and only if, it is valid for $\Phi(\mathbb{R}^n) + (\mathbb{R}_- \times \mathbb{R}^m_+)$. The advantage of $\Phi(\mathbb{R}^n) + (\mathbb{R}_- \times \mathbb{R}^m_+)$ over $\Phi(\mathbb{R}^n)$ is that it is convex. Now, as p^* is the optimal value, it follows that there cannot be any feasible point, x such that $g_i(x) \leq 0$ for all i, such that $f(x) < p^*$, that is,

$$(\Phi(\mathbb{R}^n) + (\mathbb{R}_- \times \mathbb{R}^m_+)) \cap ((p^*, +\infty) \times \mathbb{R}^m_-) = \emptyset.$$

We illustrate $\Phi(\mathbb{R}^n) + (\mathbb{R}_- \times \mathbb{R}^m_+)$ and $(p^*, +\infty) \times \mathbb{R}^m_-$ in Figure 1.7 for $\max\{-e^{-x} : \sqrt{x^2 + y^2} \leq y\}$.

Now, $\Phi(\mathbb{R}^n) + (\mathbb{R}_- \times \mathbb{R}^m_+)$ and $(p^*, +\infty) \times \mathbb{R}^m_-$ are two convex sets which do not intersect. Therefore, from separation theorems, we know that there must exist a hyperplane separating both sets. For our current example, $y_1 = 0$ is the only hyperplane that separates both sets, but remember that the dual tries to find a hyperplane with a nonzero coefficient for y_0 and contains $\Phi(\mathbb{R}^n)$ on one side, thus, $y_1 = 0$ is not feasible for the dual problem. So, how could we ensure that, first, such a hyperplane exists and, second, it actually separates $\Phi(\mathbb{R}^n)$ from $(p^*, +\infty) \times \mathbb{R}^m_-$? Note that the existence of such a hyperplane is related to the feasibility of the dual problem, while the separation of $\Phi(\mathbb{R}^n)$ from $(p^*, +\infty) \times \mathbb{R}^m_-$ ensures that the dual achieves the same value as the primal.

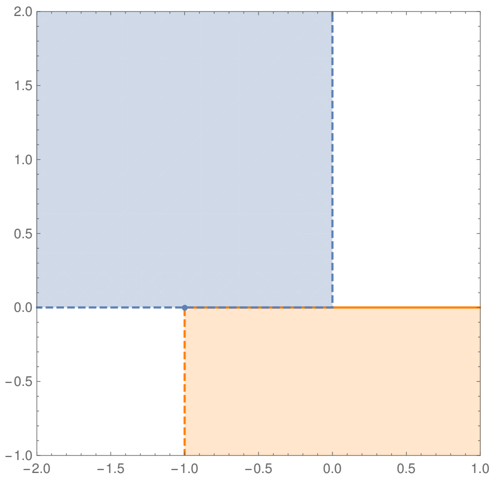


Figure 1.7: The set $\Phi(\mathbb{R}^2)$ is depicted in blue and $(-1,\infty)\times\mathbb{R}_-$ in orange, where $\Phi(x,y) = (-e^{-x}, \sqrt{x^2+y^2}-y)$.

We will now see that if $\Phi(\mathbb{R}^n)$ intersects the interior of $\mathbb{R}\times\mathbb{R}^m_-$, then we will have that the dual is feasible and equal to the primal. That is, if there exists an x_0 such that $g_i(x_0) < 0$ for all $i \in [m]$, then strong duality holds. Indeed, such a point forces every hyperplane separating $\Phi(\mathbb{R}^n)$ from $(p^*, +\infty)\times\mathbb{R}^m_-$ to have a nonzero coefficient for y_0. This should be fairly intuitive from the pictures. To see this algebraically, let $\mu_0 y_0 - \sum \mu_i y_i \leq \theta$ be a hyperplane that separates $\Phi(\mathbb{R}^n)$ from $(p^*, +\infty)\times\mathbb{R}^m_-$. In particular, $(\mu_0, \mu) \neq 0$ as otherwise $\mu_0 y_0 - \sum \mu_i y_i \leq \theta$ would not be a hyperplane. As $(f_0, g_1(x_0), \ldots, g_m(x_0)) \in \Phi(\mathbb{R}^n)$, it follows that $\mu_0 f(x_0) - \sum \mu_i g(x_0) \leq \theta$. As $(p, 0) \in (p^*, +\infty)\times\mathbb{R}^m_-$ for every $p > p^*$, it follows that $\theta \leq \mu_0 p$ for every $p > p^*$, which implies that $\theta \leq \mu_0 p^*$. Thus, $\mu_0 f(x_0) - \sum \mu_i g(x_0) \leq \mu_0 p^*$.

Now, if $\mu_0 = 0$, then $-\sum \mu_i g(x_0) \leq 0$, but $\mu \geq 0$ and $g(x_0) < 0$, which can only hold if $\mu = 0$. However, this contradicts $(\mu_0, \mu) \neq 0$. Therefore $\mu_0 > 0$ and we can normalize so that $\mu_0 = 1$. This shows that the dual is feasible and that its value is equal to the primal. Indeed, $f(x) - \sum \mu_i g(x) \leq p^*$ implies that $\theta(\mu) \leq p^*$, but by construction, $\theta(\mu) \geq p^*$.

If there exists an x_0 such that $g_i(x_0) < 0$ for $i \in [m]$, then we say that *Slater's condition* holds and x_0 is called an *Slater point*. Thus, we have proven that if the primal is feasible, bounded and, Slater's condition holds, then there is strong duality. The above result still holds when Slater's condition is weaken to ask that there exists a point x_0 such that $g_i(x_0) < 0$ for every g_i that is non-linear, see (Rockafellar, 1970, Theorem 28.2). The proof of such a result follows the same reasoning, but one needs a slightly stronger separation theorem that exploits the polyhedrality of $(p^*, +\infty) \times \mathbb{R}^m_-$, see (Rockafellar, 1970, Theorem 20.2).

More interpretations of duality among these lines can be found in Pourciau (1980).

1.4 Monoidal Strengthening

In Chapter 4, we apply a modification of monoidal strengthening to intersection cuts. In this section, we explain what monoidal strengthening is.

Monoidal strengthening is a technique introduced in 1980 by Balas and Jeroslow (1980). Our deduction of the monoidal strengthening technique applied to disjunctions is novel and is inspired by Wiese (2016, Section 4.2.3) and several conversations with Sven Wiese. We also present the general technique in the more modern framework of S-free sets.

Before we start, a *monoid* is the discrete analog of a convex convex. A monoid is a pair $(M, +)$ where M is a set and $+ : M \times M \to M$ such that $+$ is associative and there exist $0 \in M$ such that $+(0, \cdot)$ is the identity. The name monoidal strengthening comes from the use of a monoid to strengthen cuts.

As we discussed in Section 1.2, a simple way of generating cutting planes is through cut generating functions. In this setting, and for the rest of this section, we will assume that we have the following relaxation of the feasible region of our optimization problem

$$\{(x, y) \in \mathbb{R}^q_+ \times \mathbb{Z}^p_+ \,:\, \sum_i r^i x_i + \sum_j d^j y_j \in S\}, \tag{1.4}$$

where $S \subseteq \mathbb{R}^n$ is a closed set such that $0 \notin S$ and $r^i, d^j \in \mathbb{R}^n$. We also have a convex S-free set C such that $0 \in \operatorname{int} C$. The set C is represented by a sublinear function ϕ, that is,

$$C = \{z \in \mathbb{R}^n \,:\, \phi(z) \leq 1\}.$$

The intersection cut generated by ϕ that separates the point $(x, y) = (0, 0)$ from (1.4) is

$$\sum_i \phi(r^i) x_i + \sum_j \phi(d^j) y_j \geq 1$$

Probably the most intuitive way of understanding monoidal strengthening is to see it as a technique that takes a relaxation of the form

$$\{(x, y) \in \mathbb{R}^q_+ \times \mathbb{Z}^p_+ \,:\, \sum_i r^i x_i + \sum_j d^j y_j \in S\},$$

and builds new ones,

$$\{(x,y) \in \mathbb{R}^q_+ \times \mathbb{Z}^p_+ \,:\, \sum_i r'^i x_i + \sum_j d'^j y_j \in S'\}.$$

Each of them can generate a cut that separates $(0,0)$ and, of course, the idea is to select a "best" one. The construction of new relaxations exploits the fact that some variables are restricted to be integers and the structure of the set S. Let us see two examples before we present the general principle of monoidal strengthening.

1.4.1 One Row Relaxations: Gomory Cuts

Assume $f \notin \mathbb{Z}$, $n = 1$, and that (1.4) is

$$\sum_i r^i x_i + \sum_j d^j y_j \in S := \mathbb{Z} - f,$$

where $r^i, d^j \in \mathbb{R}$. As each $y_j \in \mathbb{Z}$, adding some integer multiple of y_j to the above relation does not change S. That is, if $m^j \in \mathbb{Z}$, then

$$\sum_i r^i x_i + \sum_j d^j y_j + \sum_j m^j y_j \in \mathbb{Z} + \sum_j m^j y_j - f = \mathbb{Z} - f.$$

Thus, if C is a convex S-free set represented by ϕ, such that $0 \in \operatorname{int} C$, then not only is $\sum_i \phi(r^i) x_i + \sum_j \phi(d^j) y_j \geq 1$ a valid inequality, but also

$$\sum_i \phi(r^i) x_i + \sum_j \phi(d^j + m^j) y_j \geq 1 \text{ for every } m^j \in \mathbb{Z}.$$

Note that in this particular case, the only maximal S-free set that contains 0 is $C = [-f, 1-f]$ and the only sublinear function ϕ such that $C = \{x \in \mathbb{R} \,:\, \phi(x) \leq 1\}$ is its gauge, $\phi(x) = \max\{\frac{x}{1-f}, -\frac{x}{f}\}$. Using ϕ and finding the best m^j for each d^j yields the Gomory cut (Gomory, 1960). By best here we mean the m^j that makes $\phi(d^j + m^j)$ the smallest.

1.4.2 Disjunctive Cuts

Let Q be an index set and consider an optimization problem $\mathcal{P}$ such that

$$S = \{(x,y) \in \mathbb{R}^q_+ \times \mathbb{Z}^p_+ \,:\, \bigvee_{k \in Q} a(k)^\mathsf{T} x + d(k)^\mathsf{T} y \geq 1\},$$

is a valid disjunction, that is, every feasible solution of $\mathcal{P}$ is in S. Here, we denote the vectors as $a(k) \in \mathbb{R}^q$ and $d(k) \in \mathbb{R}^p$ instead of the more usual notation a_k and d_k. As (1.4) we use

$$\sum_j e_j x_j + \sum_j e_{j+q} y_j = (x, y) \in S.$$

Consider the S-free set

$$C = \{(x, y) \in \mathbb{R}^{q+p} \,:\, a(k)^\mathsf{T} x + d(k)^\mathsf{T} y \leq 1 \text{ for } k \in Q\}$$

A sublinear function representing C, which may or may not be its gauge, is

$$\phi_C(x, y) = \max_{k \in Q} a(k)^\mathsf{T} x + d(k)^\mathsf{T} y. \tag{1.5}$$

Thus, we obtain the cut

$$1 \leq \sum_j \phi_C(e_j) x_j + \sum_j \phi_C(e_{q+j}) y_j = \sum_j \left(\max_{k \in Q} a(k)_j\right) x_j + \sum_j \left(\max_{k \in Q} d(k)_j\right) y_j$$

This cut is known as *disjunctive cut* (Balas, 1979) and the implication

$$\bigvee_{k \in Q} \sum_j a(k)_j x_j + \sum_j d(k)_j y_j \geq 1 \implies \sum_j x_j \max_{k \in Q} a(k)_j + \sum_j y_j \max_{k \in Q} d(k)_j \geq 1,$$

is known as the *maximum principle*.

Monoidal strengthening in this setting amounts to finding a new disjunction that every feasible point must satisfy. Balas and Jeroslow showed how to build new disjunctions. For their construction we need that if each disjunction is relaxed enough, then it is automatically satisfied. More formally, we need that for each k, there is a b_k such that every $x \in S$ satisfies $a(k)^\mathsf{T} x + d(k)^\mathsf{T} y \geq b_k$. In other words, we need that the expression $a(k)^\mathsf{T} x + d(k)^\mathsf{T} y$ is bounded from below in the feasible region of $\mathcal{P}$.

For example, consider $S = \{x \in \mathbb{R}^2_+ \,:\, \frac{x_1 + x_2}{3} \geq 1 \vee x_1 \geq 1\}$ and assume that the feasible region of $\mathcal{P}$ is S. Then, $\frac{x_1+x_2}{3} \geq \frac{1}{3}$ is a valid inequality for S. In other words, if we relax the first disjunctive term $\frac{x_1+x_2}{3} \geq 1$ by $\frac{2}{3}$, then we obtain an inequality satisfied by every element of S. Thus, $b^1 = \frac{1}{3}$. Similarly, $b^2 = 0$ is a lower bound for the second disjunctive term that makes it valid for every element of S.
On the other hand, consider $S = \{x \in \mathbb{R}^2_+ \,:\, \frac{x_1}{2} \geq 1 \vee x_1 - x_2 \geq 1\}$. While $b^1 = \frac{1}{2}$ is a valid bound for $\frac{x_1}{2}$, there is no b^2 such that $x_1 - x_2 \geq b^2$ is valid for S. One reason is that $x_1 = 2$, $x_2 \geq 0$ is in S and so $x_1 - x_2$ is unbounded from below.

Given the lower bounds b_k we have the following lemma which will allow us to build new disjunctions. Notice that we can, and will, assume that $b_k < 1$ as otherwise the disjunction is trivially satisfied.

Lemma 1.1. *Every $(x, y) \in S$ satisfies the disjunction*

$$\bigvee_{k \in Q} a(k)^T x + d(k)^\mathsf{T} y + (1 - b_k) z_k \geq 1 \tag{1.6}$$

whenever $z \in Z = \{z \in \mathbb{Z}^Q : z = 0 \vee \exists k, z_k \geq 1\}$.

Proof. If $z = 0$ there is nothing to prove. Let $z \neq 0$ and let $k_0 \in Q$ be such that $z_{k_0} \geq 1$. Then, the disjunction is satisfied because $a(k_0)^\mathsf{T} x + d(k_0)^\mathsf{T} y + (1 - b_{k_0}) z_{k_0} \geq 1$ is a relaxation of $a(k_0)^\mathsf{T} x + d(k_0)^\mathsf{T} y \geq b_{k_0}$ which is satisfied by hypothesis. To see that it is a relaxation just notice that $b_{k_0} \geq 1 - (1 - b_{k_0}) z_{k_0}$. □

As written in the lemma above, this disjunction is not interesting due to the fact that for any given $z \in Z$, (1.6) is either the original disjunction or is redundant. However, by making z depend on y, we obtain a non-trivial new disjunction.

Theorem 1.2 (Balas and Jeroslow (1980, Theorem 3))**.** *Let*

$$\mathcal{M} := \{m \in \mathbb{Z}^Q : \sum_{k \in Q} m_k \geq 0\} \tag{1.7}$$

and consider $m(k) \in \mathbb{R}^p$ for $k \in Q$ such that $(m(k)_j)_{k \in Q} \in \mathcal{M}$ for all $j \in [p]$.

Then,

$$\bigvee_{k \in Q} a(k)^\mathsf{T} x + (d(k) + (1 - b_k) m(k))^\mathsf{T} y \geq 1 \tag{1.8}$$

is a valid disjunction for $(x, y) \in S$.

Proof. Let $(\bar{x}, \bar{y}) \in S$ and let $z \in \mathbb{Z}^Q$ be defined by $z_k = m(k)^\mathsf{T} \bar{y}$. Since

$$\sum_k z_k = \sum_j \underbrace{\bar{y}_j}_{\in \mathbb{Z}_+} \underbrace{\sum_k m(k)_j}_{\geq 0} \geq 0$$

we conclude that $z \in Z$.

On the other hand, note that (1.6) is equivalent to

$$\bigvee_{k \in Q} a(k)^\mathsf{T} x + d(k)^\mathsf{T} y + (1 - b_k) m(k)^\mathsf{T} \bar{y} \geq 1.$$

As $z \in Z$, Lemma 1.1 implies that the previous disjunction is valid for every $(x, y) \in S$, in particular, for $(\bar{x}, \bar{y})$. Evaluating the disjunction at $(\bar{x}, \bar{y})$ yields

$$\bigvee_{k \in Q} a(k)^\mathsf{T} \bar{x} + d(k)^\mathsf{T} \bar{y} + (1 - b_k) m(k)^\mathsf{T} \bar{y} \geq 1.$$

which is equivalent to evaluating (1.8) at $(\bar{x}, \bar{y})$. Thus, $(\bar{x}, \bar{y})$ satisfies (1.8). It follows that every $(x, y) \in S$ satisfies (1.8) as we wanted to show. ☐

The theorem implies that each Q-tuple $M = (m(k) \in \mathbb{R}^p : k \in Q)$ such that $(m(k)_j)_{k \in Q} \in \mathcal{M}$ for all $j \in [p]$, yields a new valid disjunction, namely (1.8), which in turn yields a new S-free set

$$C_M = \{(x, y) \in \mathbb{R}^{q+p} : \phi_{C_M}(x, y) \leq 1\},$$

where $\phi_{C_M}(x, y) = \max_{k \in Q} a(k)^\mathsf{T} x + (d(k) + (1 - b_k) m(k))^\mathsf{T} y$. Therefore,

$$\sum_j \phi_{C_M}(e_j) x_j + \sum_j \phi_{C_M}(e_{q+j}) y_j \geq 1$$

is a valid inequality for S. This inequality reads

$$\sum_j \left(\max_{k \in Q} a(k)_j \right) x_j + \sum_j \left(\max_{k \in Q} (d(k)_j + (1 - b_k) m(k)_j) \right) y_j \geq 1.$$

Choosing the best possible tuple M yields

$$\sum_j \left(\max_{k \in Q} a(k)_j \right) x_j + \sum_j \left(\max_{k \in Q} \inf_{m \in \mathcal{M}} (d(k)_j + (1 - b_k) m_k) \right) y_j \geq 1. \quad (1.9)$$

1.4.3 Monoidal Strengthening

The general principle of monoidal strengthening is as follows. Assume we have a monoid M and an $(S + M)$-free set $C = \{(x, y) : \phi(x, y) \leq 1\}$ where ϕ is sublinear. If $\{(x, y) \in \mathbb{R}^n_+ \times \mathbb{Z}^p_+ : \sum_i r^i x_i + \sum_j d^j y_j \in S + M\}$ is a valid relaxation, then not only is $\sum_i \phi(r^i) x_i + \sum_j \phi(d^j) y_j \geq 1$ valid, but also

$$\sum_i \phi(r^i) x_i + \sum_j \inf_{m \in M} \phi(d^j + m) y_j \geq 1. \quad (1.10)$$

In particular, the previous cut is the strongest one that can be obtained with this technique.

The proof of validity follows from exploiting the integrality restrictions of y. Indeed, as y is a non-negative integer, $\sum m^j y_j \in M$ for every $m^j \in M$. Every feasible solution satisfies $\sum_i r^i x_i + \sum_j d^j y_j \in S + M$ and so they also satisfy

$$\sum_i r^i x_i + \sum_j (d^j + m^j) y_j \in S + M + \sum m^j y_j \subseteq S + M + M = S + M, \quad (1.11)$$

where the last equality holds because M is a monoid, in particular, closed under addition. Then, applying the cut generating function to (1.11), we obtain the valid inequality

$$\sum_i \phi(r^i) x_i + \sum_j \phi(d^j + m^j) y_j \geq 1.$$

As the $m^j \in M$ are arbitrary, we obtain (1.10).

For this technique to work, one actually needs a monoid. In the case of Gomory cuts, $S = \mathbb{Z} - f$. As $S + \mathbb{Z} = S$, one can use $M = \mathbb{Z}$ as the monoid for monoidal strengthening. One can also write the relaxation as $f + \sum_i r^i x_i + \sum_j d^j y_j \in S$ and consider S to be $\mathbb{Z}$, in which case the monoid is S itself. In the literature, it is rather common to use S or a subset of it as the monoid. For example, relaxations where $S = \mathbb{Z}^n$ or $S = \mathbb{Z}^n \cap P$, where P is a polyhedron, or even a convex set have been studied, see for example Andersen et al. (2007), Basu et al. (2010b), Morán and Dey (2011) and (Conforti et al., 2014, Chapter 6). When $S = \mathbb{Z}^n \cap P$ and P is a polyhedron, a typical monoid used for strengthening is $M = \mathbb{Z}^n \cap \operatorname{lin}(\operatorname{conv}(S))$, see for example Dey and Wolsey (2010), Conforti et al. (2011a) and Basu et al. (2012).

A more complicated setting is when $S + M \neq S$. The disjunctive case corresponds to this more complicated setting, but to be able to see this, we need to follow more closely the original derivation of Balas and Jeroslow (1980). Furthermore, note that our exposition of the disjunctive case in Section 1.4.2 is not an application of the general principle of monoidal strengthening as presented here, since we also modified the S-free set.

The setting for the disjunctive case is that we have an optimization problem on the variables $(x, y) \in \mathbb{R}^q_+ \times \mathbb{Z}^p_+$ such that

$$\bigvee_{k \in Q} a(k)^\mathsf{T} x + d(k)^\mathsf{T} y \geq 1$$

is a valid disjunction. In Section 1.4.2, we represented this by taking

$$S = \{(x, y) \in \mathbb{R}^q_+ \times \mathbb{Z}^p_+ \,:\, \bigvee_{k \in Q} a(k)^\mathsf{T} x + d(k)^\mathsf{T} y \geq 1\}$$

and $\sum_j e_j x_j + \sum_j e_{j+q} y_j = (x, y) \in S$ as our relaxation (1.4).

However, we can also represent the disjunction in a different way. Recall that b_k are the lower bounds on the disjunctive terms $a(k)^\mathsf{T} x + d(k)^\mathsf{T} y$, see Section 1.4.2. Let

$$S_b = \{w \in \mathbb{R}^Q \ : \ \bigvee_{k \in Q} w_k \geq 1, \ w \geq b\}.$$

We can model this disjunction via

$$\begin{pmatrix} a(1)^\mathsf{T} x + d(1)^\mathsf{T} y \\ \vdots \\ a(K)^\mathsf{T} x + d(K)^\mathsf{T} y \end{pmatrix} \in S_b \iff Ax + Dy \in S_b,$$

where $A = \begin{pmatrix} a(1)^\mathsf{T} \\ \vdots \\ a(K)^\mathsf{T} \end{pmatrix}$, $D = \begin{pmatrix} d(1)^\mathsf{T} \\ \vdots \\ d(K)^\mathsf{T} \end{pmatrix}$, and $Q = \{1, \dots, K\}$.

We have that $C_b = \{w \in \mathbb{R}^Q \ : \ w_k \leq 1\}$ is a convex S_b-free set. Note that $C_b = \{w \in \mathbb{R}^Q \ : \ \phi_{C_b}(w) \leq 1\}$, where $\phi_{C_b}(w) = \max_{k \in Q} w_k$, and it is sublinear. Note that $\phi_{C_b}(A_{\cdot j}) = \phi_C(e_j, 0)$ and $\phi_{C_b}(D_{\cdot j}) = \phi_C(0, e_j)$, where ϕ_C is defined in (1.5).

Now, consider the monoids $\mathcal{M}$ defined in (1.7) and $T = \{\tau \in \mathbb{R}^Q \ : \exists m \in \mathcal{M}, \ \tau_k = (1 - b_k) m_k, \text{ for } k \in Q\}$. Let us see that C_b is $(S_b + T)$-free. Let $\theta \in S_b + T$ so that $\theta = w + \tau$. If $\tau = 0$, then $\theta \in S_b$ and so $\theta \notin \operatorname{int} C_b$ as C_b is S_b-free. Otherwise, there exists $k \in Q$ such that $m_k > 0$ so $\tau_k \geq 1 - b_k$. Then, $\theta_k = w_k + \tau_k \geq b_k + 1 - b_k \geq 1$, thus, $\theta \notin \operatorname{int} C_b$.

Summarizing, C_b is not only S_b-free but also $(S_b + T)$-free. Thus, we can apply monoidal strengthening to obtain the cut (1.9). This argument is basically a modern rewrite of the argument in Balas and Jeroslow (1980). Note here that $S_b + T \neq S_b$.

In general, the challenge of monoidal strengthening is to find a monoid M such that given a closed set S and an S-free set C, C is also $(S + M)$-free, so that we can apply monoidal strengthening as above.

Chapter 2

On the Relation Between the Extended Supporting Hyperplane Algorithm and Kelley's Cutting Plane Algorithm

In this chapter we revisit two classical algorithms for convex mixed integer optimization, namely, Kelley's cutting plane algorithm and Veinott's supporting hyperplane algorithm. The motivation to look into these algorithm is the following. Some state-of-the-art LP-based MINLP solvers enforce convex constraint by adding gradient cutting planes. Simple examples show that these cuts do not necessarily support the feasible region, and so they are dominated. In order to build undominated cuts, or equivalently, supporting cuts, different separation procedures are needed such as the one proposed by Veinott.

However, it is not always the case that gradient cutting planes are not supporting. Thus, the purpose of this chapter is to understand when gradient cutting planes are supporting. Our findings naturally suggest a reformulation of the feasible region for which every gradient cut is supporting. As a consequence, we can show that Veinott's supporting hyperplane algorithm is just a special case of Kelley's cutting plane algorithm. As a result, we extend the applicability of the supporting hyperplane algorithm to convex problems represented by a class of general, not necessarily convex nor differentiable, functions.

The insights obtained in this chapter, together with an interpretation of gradient cutting planes as intersection cuts presented in Chapter 4 will motivate the basic construction of maximal quadratic-free sets presented in Chapter 5.

The chapter is organized as follows. In Section 2.1 we introduce the object of study of this chapter and review the literature on cutting plane approaches and efforts on obtaining supporting valid inequalities. In Section 2.2, we characterize functions whose linearizations are supporting hyperplanes to their

0-sublevel sets. Section 2.3 introduces the gauge function and shows how to use it for building supporting hyperplanes. We note that evaluating the gauge function is equivalent to the line search step of the supporting hyperplane algorithm. This equivalence provides the link between the supporting hyperplane and Kelley's cutting plane algorithm. In Section 2.4, we show that the cutting planes generated by the supporting hyperplane algorithm can also be generated by Kelley's algorithm when applied to a reformulation of the problem. This implies that the convergence of the supporting hyperplane algorithm follows from Kelley's. In Section 2.5, we show that we can apply the supporting hyperlane algorithm to problem whose feasible region is convex but represented via functions that are not necessarily convex nor differentiable. We introdue the concept of a *well-behaved* generalized directional derivative and show that if the functions have well-behaved generalized directional derivatives and 0 does not belong to the generalized subdifferential at points where the functions are zero, then the supporting hyperplane algorithm converges. Finally, Section 2.6 presents our concluding remarks.

This chapter is joint work with Ambros Gleixner and Robert Schwarz and has been submitted to the Journal of Global Optimization.

2.1 Background

A mixed integer convex program (MICP) is a problem of the form

$$\min\{c^{\mathsf{T}}x \,:\, x \in C \cap (\mathbb{Z}^p \times \mathbb{R}^{n-p})\}, \tag{2.1}$$

where C is a closed convex set, $c \in \mathbb{R}^n$, and p denotes the number of variables with integrality requirement. The use of a linear objective function is without loss of generality given that one can always transform a problem with a convex objective function into a problem of the form (2.1). We can represent the set C in different ways, one of the most common being as the intersection of sublevel sets of convex differentiable functions, that is,

$$C = \{x \in \mathbb{R}^n \,:\, g_j(x) \leq 0, j \in J\}. \tag{2.2}$$

Here, J is a finite index set and each g_j is convex and differentiable.

Several methods have been proposed for solving MICP. When the problem is continuous and represented as (2.2), one of the first proposed methods was the cutting plane algorithm by J. E. Kelley (1960). This algorithm exploits the convexity of a constraint function g to build gradient cuts.

The idea of Kelley's cutting plane (KCP) algorithm is to approximate the feasible region with a polytope, solve the resulting linear program (LP) and, if the LP solution is not feasible, separate it using gradient cuts to obtain a new polytope which is a better approximation of the feasible region and repeat, see Algorithm 2.1.

Algorithm 2.1: Kelley's cutting plane algorithm

$LP = \{x : x \in [l, u]\}, \bar{x} \leftarrow \arg\min_{x \in LP} c^\mathsf{T} x$
while $\max_{j \in J} g_j(\bar{x}) > \epsilon$ **do**
  forall *j such that* $g_j(\bar{x}) > 0$ **do**
    $LP \leftarrow LP \cap \{x : g_j(\bar{x}) + \nabla g_j(\bar{x})(x - \bar{x}) \leq 0\}$
  $\bar{x} \leftarrow \arg\min_{x \in LP} c^\mathsf{T} x$
return $\bar{x}$

Kelley shows that the algorithm converges to the optimum and it converges in finite time to a point close to the optimum. By solving integer programs (IP) using the cutting planes of Gomory (1958) instead of LP relaxations, Kelley shows that his cutting plane algorithm solves purely integer convex programs in finite time. The same algorithm works just as well for MICP. However, Kelley did not have access to a finite algorithm for solving mixed integer linear programs (MILP).

In an attempt to speed up Kelley's algorithm, Veinott (1967) proposes the *supporting hyperplane algorithm* (SH). A possible issue with Kelley's algorithm is that, in general, gradient cuts do not support the feasible region, see Figure 2.1. Therefore, it is expected that better relaxations can be achieved by using supporting cutting planes.

In order to construct supporting hyperplanes, Veinott suggests to build gradient cuts at boundary points of C. He uses an interior point of C to find the point on the boundary, $\hat{x}$, that intersects the segment joining the interior point and the solution of the current relaxation. These cuts are automatically supporting hyperplanes of C, at $\hat{x}$. However, since the cut is computed at $\hat{x}$ which is in C, it might happen that the gradient of the constraints active at $\hat{x}$ vanishes. For this reason, Veinott also requires that the functions representing C have non-vanishing gradients at the boundary. This is immediately implied by, e.g., Slater's condition (Section 1.3). Veinott also identifies that one can use his algorithm to solve (2.1) when representing C by quasi-convex functions, that is, functions whose sublevel sets are convex.

Recently, Kronqvist et al. (2016) rediscovered and implemented Veinott's algorithm (Veinott, 1967). They call their algorithm the *extended supporting hyperplane algorithm* (ESH). They discuss the practical importance of choosing a good interior point and propose some improvements over the original method, such as solving LP relaxations during the first iterations instead of the more expensive MILP relaxation. As a result, they present a computationally competitive solver implementation for MICPs defined by convex differentiable constraint functions (Kronqvist et al., 2018).

In this chapter, we would like to understand when, given a convex differentiable function g, gradient cuts of g are supporting to the convex set $C = \{x \in \mathbb{R}^n : g(x) \leq 0\}$. This question is motivated by the fact that in this case Kelley's algorithm automatically becomes a supporting hyperplane algorithm. In Theorem 2.3 we give a necessary and sufficient condition for a gradient cut of g at a given point to be a supporting hyperplane of C. In particular, this condition suggests to look at *sublinear functions*, i.e., convex and positively homogeneous functions. As it turns out, this naturally leads to Veinott's algorithm.

Sublinear functions and convex sets are deeply related. When the origin is in the interior of a convex set C, then we can represent C via its *gauge* function φ_C, which is sublinear (Rockafellar, 1970). We give the formal definition of the gauge function in Section 2.3, but for now it suffices to know that we can represent C as $C = \{x \in \mathbb{R}^n : \varphi_C(x) \leq 1\}$ and that, in particular, for every $\bar{x} \neq 0$ a gradient cut of φ_C at $\bar{x}$ supports all of its sublevel sets. The following example illustrates this.

Example 2.1. Consider the convex feasible region given by

$$C = \{(x, y) \in \mathbb{R}^2 : g(x, y) \leq 0\},$$

where $g(x, y) = x^2 + y^2 - 1$. We show through an example that gradient cuts of g are not necessarily supporting to C, explain why this happens, and show that changing the representation of C to use its gauge function solves the issue.

Separating the infeasible point $\bar{x} = (\frac{3}{2}, \frac{3}{2})$ by a gradient cut of g at $\bar{x}$ gives

$$\begin{aligned} g(\bar{x}) + \nabla g(\bar{x})(x - \bar{x}) \leq 0 & \\ \Leftrightarrow x + y \leq \frac{11}{6}&. \end{aligned}$$

This cut does not support C, see Figure 2.1. Alternatively, the gauge function of C is given by $\varphi_C(x, y) = \sqrt{x^2 + y^2}$ and $C = \{(x, y) : \sqrt{x^2 + y^2} \leq 1\}$. The gradient cut of φ_C at $\bar{x}$ is $x + y \leq \sqrt{2}$, which is supporting. $\square$

From the previous discussion it is a natural idea to represent C via its gauge function, namely, $C = \{x \in \mathbb{R}^n : \varphi_C(x) \leq 1\}$. However, as mentioned

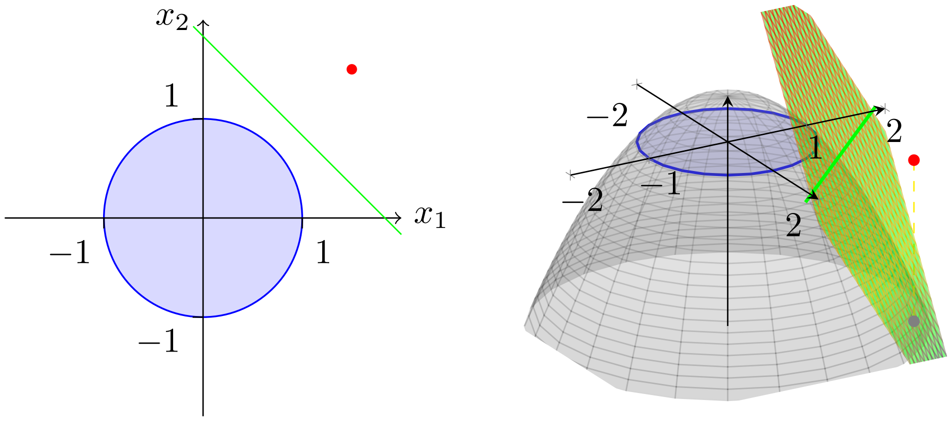


Figure 2.1: The feasible region C and the infeasible point $\bar{x} = (\frac{3}{2}, \frac{3}{2})$ to separate. On the left we see that the separating hyperplane is not supporting to C. On the right we see why this happens: the linearization of g at $\bar{x}$ is tangent to the epigraph of g (shown upside-down for clarity) at $(\bar{x}, g(\bar{x}))$. However, when this hyperplane intersects the x-y-plane, it is already far away from the epigraph, and in consequence, from the sublevel set. The intersection of the hyperplane with the x-y-plane is the gradient cut.

before, C is usually given by (2.2). Our main contribution is to show that reformulating (2.2) to the gauge representation will naturally lead to the ESH algorithm, see Section 2.3.2. As a consequence, the convergence proofs of Veinott (1967) and Kronqvist et al. (2016) follow directly from the convergence proof of Kelley's cutting plane algorithm (J. E. Kelley, 1960; Horst and Tuy, 1990), see Section 2.4. In other words, we show that the ESH algorithm is KCP algorithm applied to a different representation of the problem.[2]

Motivated by this approach of representing C by its gauge function, we are able to show that the ESH algorithm applied to (2.1) converges even when C is not represented by convex functions. This is related to recent work of Lasserre (2009) that tries to understand how different techniques behave when the convex set C is not represented via (2.2). Lasserre considers sets $C = \{x : g_j(x) \leq 0, j \in J\}$ where g_j are only differentiable, but not necessarily convex in the following setting:

[2] Strictly speaking, when the problem is mixed integer, the KCP algorithm only corresponds to the so-called LP-step (Kronqvist et al., 2016) of the ESH algorithm. However, given that the KCP algorithm allows for an straightforward extension to the mixed integer case, we will continue to compare the KCP algorithm to the ESH algorithm with respect to their technique of generating cutting planes.

Assumption 2.2. *For all $x \in C$ and all $j \in J$, if $g_j(x) = 0$, then $\nabla g_j(x) \neq 0$.*

Under this assumption, that is, if the gradients of active constraints do not vanish at the boundary of C, Lasserre shows that the KKT conditions are not only necessary but also sufficient for global optimality. In other words, every minimizer is a KKT point and every KKT point is a minimizer.

A series of generalizations follow the work of Lasserre. Dutta and Lalitha (2011) generalize the previous result to the case where C is represented by locally Lipschitz functions, not necessarily differentiable nor convex, but regular in the sense of Clarke (Clarke, 1990), see also Definition 2.15. Martínez-Legaz (2014) further generalize the result to the case where C is represented by *tangentially convex* functions (Lemaréchal, 1986; Pshenichnyi, 1971). Kabgani et al. (2017) generalize the result to the case where C is represented by functions that admit an *upper regular convexificator* URC (Jeyakumar and Luc, 1999), see also Definition 2.16. We note that regular functions in the sense of Clarke and tangentially convex functions admit a URC (Kabgani et al., 2017), thus the URC assumption is the most general among the ones considered in these works.

In terms of computations, Lasserre (2011, 2014) proposes an algorithm to find the KKT point via log-barrier functions. He shows that the algorithm converges to the KKT point if Assumption 2.2 holds.

For all these concepts of generalized derivative, there is a notion of directional derivative and a notion of subdifferential. For example, for functions that admit a URC, the notion of directional derivative is the upper Dini directional derivative and its subdifferential is the URC, see Definition 2.16. Let f be a function and let us denote by $f^{'}(x; d)$ a generalized directional derivative. We say that the directional derivative is *well-behaved* if $f^{'}(x; d) > 0$ implies that there exists $t_n \searrow 0$ such that $f(x + t_n d) > f(x)$.

In this sense we show that if C is represented by functions whose generalized directional derivatives are well-behaved, then the ESH converges to the global optimum, under the equivalent of Assumption 2.2 (see (2.8)) for the corresponding subdifferential. The upper Dini directional derivative is certainly well-behaved and, thus, our result shows that the ESH converges when C is represented by functions that admit a URC. We also show that for ∂°-pseudoconvex (see Definition 2.19) constraints, the Clarke directional derivative (see Definition 2.15) is well-behaved. Therefore, our result generalizes the result of Eronen et al. (2017) that the ESH converges when C is represented by ∂°-pseudoconvex functions.

We also show, via an example, that if we use Clarke's subdifferential (Clarke, 1990), the ESH does not need to converge when the functions are only Lipschitz continuous but not regular in the sense of Clarke.

Finally, we provide a characterization of convex functions whose linearizations are supporting to their sublevel sets. Although elementary, the authors are not aware of its presence in the literature. In particular, this result allows us to identify some families of functions for which gradient cuts are never supporting (see Example 2.7) and some for which they are always supporting (see Corollary 2.5 and Example 2.6).

2.1.1 Literature Review

We can think of the algorithms of J. E. Kelley (1960) and Veinott (1967) as a mixture of two ingredients: which relaxation to solve and where to compute the cutting plane. Indeed, at each iteration we have a point x^k that we would like to separate with a linear inequality $\beta + \alpha^{\mathsf{T}}(x - x_0) \leq 0$. For Kelley's algorithm, $x_0 = x^k$, while for Veinott's algorithm, $x_0 \in \partial C$, and for both $\alpha \in \partial g(x_0)$ and $\beta = g(x_0)$. Choosing different relaxations and different points where to compute the cutting planes yields different algorithms. This framework is developed in Horst and Tuy (1990).

Following the previous framework, Duran and Grossmann (1986) propose the, so-called, *outer approximation algorithm* for MICP. The idea is to solve an MILP relaxation, but instead of computing a cutting plane at the MILP optimum, or at the boundary point on the segment between the MILP optimum and some interior point, they suggest to compute cutting planes at a solution of the nonlinear program (NLP) obtained after fixing the integer variables to the integer values given by the MILP optimal solution. This is a much more expensive algorithm but has the advantage of finite convergence. Of course, this does not work in complete generality and we need some assumptions, for example, requiring some constraint qualifications. Moreover, when obtaining an infeasible NLP after fixing the integer variables, care must be taken to prevent the same integer assignment in future iterations. To handle such cases, Duran and Grossmann propose the use of integer cuts. However, Fletcher and Leyffer (1994) point out that this is not necessary. They show that the gradient cuts at the solution of a slack NLP separates the integer assignment. Eronen et al. (2012) show that a naive generalization of the outer approximation algorithm to the non-differentiable case will not work. They provide a generalization for a particular class of function. Wei and Ali (2015a,b) provide further generalizations to the non-differentiable case.

A related algorithm to the outer approximation method is the so-called generalized Benders decomposition (Geoffrion, 1972). We refer to Duran and Grossmann (1986); Fletcher and Leyffer (1994); Quesada and Grossmann

(1992) for discussions about the relation between these two algorithms. Wei and Ali (2015c) extend the generalized Benders decomposition to Banach spaces.

Westerlund and Pettersson (1995) propose the so-called extended cutting plane algorithm. This algorithm is the extension of Kelley's cutting plane to MICP and they show that the algorithm convergences. Further extensions and convergence proofs of cutting plane and outer approximation algorithms for non-smooth problems are given in Eronen et al. (2012). An interesting generalization of the extended cutting plane algorithm to solve a class of non-convex problems is the so-called α extended cutting plane algorithm introduced by Westerlund et al. (1998). They consider problem (2.1) where C is represented by differentiable pseudoconvex constraints. The idea is that, even though a gradient cut might not be valid, one can tilt the cut in order to make it valid. The tilting is done by multiplying the gradient by some α, hence the name. We refer to Westerlund et al. (1998) for more details.

As mentioned at the beginning, the assumption that the objective function is linear is without loss of generality, provided that the original objective function is convex. However, some classes of problems cannot be encompassed by (2.1), for example, when the objective function is quasi-convex. An extension of the KCP algorithm, the (α) extended cutting plane algorithm, and the ESH to convex problems with a class of quasi-convex objectives were developed by Plastria (1985), Eronen et al. (2013), and Westerlund et al. (2018), respectively.

Yet another technique for producing tight cuts is to project the point to be separated onto C (Horst and Tuy, 1990). Using the projected point and the difference between the point and its projection, one can build a supporting hyperplane that separates the point. In the same reference, Horst and Tuy show that this algorithm converges.

There have been attempts at building tighter relaxations by ensuring that gradient cuts are supporting, in a more general context than convex mixed integer nonlinear programming. Belotti et al. (2009) consider bivariate convex constraints of the form $f(x) - y \leq 0$, where f is a univariate convex function. They propose projecting the point to be separated onto the curve $y = f(x)$ and building a gradient cut at the projection. However, their motivation is not to find supporting hyperplanes, but to find the most violated cut. Indeed, as we will see, gradient cuts for these types of constraints are always supporting (Example 2.6). Other work along these lines includes the one by Lubin et al. (2015), where the authors derive an efficient procedure to project onto a two dimensional constraint derived from a Gaussian linear chance constraint, thus building supporting valid inequalities.

Another algorithm for solving non-smooth convex optimization problems is the so-called bundle method (Hiriart-Urruty and Lemaréchal, 1993). This method has also been extended to consider the mixed integer case by de Oliveira (2016).

Finally, in terms of applications, we would like to point out that the supporting hyperplane algorithm is very popular in stochastic optimization (van Ackooij et al., 2018, 2013; van Ackooij and de Oliveira, 2016; Arnold et al., 2013; Prékopa, 1995; Prékopa and Szántai, 1978; Szántai, 1988).

2.2 Characterization of Functions with Supporting Linearizations

We now give necessary and sufficient conditions for the linearization of a convex, not necessarily differentiable, function g at a point $\bar{x}$ to support the region $C = \{x \in \mathbb{R}^n : g(x) \leq 0\}$. In order for this to happen, the supporting hyperplane has to support the epigraph on the whole segment joining the point of C where it supports and $(\bar{x}, g(\bar{x}))$. In other words, the function must be affine on the segment joining the set C and $\bar{x}$. This is due to the convexity of g.

Theorem 2.3. *Let $g\colon \mathbb{R}^n \to \mathbb{R}$ be a convex function, $C = \{x \in \mathbb{R}^n : g(x) \leq 0\} \neq \emptyset$, and $\bar{x} \notin C$. There exists a subgradient $v \in \partial g(\bar{x})$ such that the valid inequality*

$$g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x}) \leq 0 \tag{2.3}$$

supports C, if and only if, there exists $x_0 \in C$ such that $\lambda \mapsto g(x_0 + \lambda(\bar{x} - x_0))$ is affine in $[0, 1]$.

Proof. ($\Rightarrow$) Let $x_0 \in \partial C$ be the point where (2.3) supports C. The idea is to show that the affine function $x \mapsto g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x})$ coincides g at two points, $\bar{x}$ and x_0. Then, by the convexity of g, it must coincide with g on the segment joining both points.

In more detail, by definition of x_0 we have,

$$g(\bar{x}) + v^{\mathsf{T}}(x_0 - \bar{x}) = 0. \tag{2.4}$$

For $\lambda \in [0, 1]$, let $l(\lambda) = x_0 + \lambda(\bar{x} - x_0)$ and $\rho(\lambda) = g(l(\lambda))$. Since g is convex and l affine, ρ is convex.

Since v is a subgradient,

$$g(\bar{x}) + v^{\mathsf{T}}(l(\lambda) - \bar{x}) \leq \rho(\lambda) \text{ for every } \lambda \in [0, 1].$$

After some algebraic manipulation and using that $\rho(1) = g(\bar{x}) = v^{\mathsf{T}}(\bar{x} - x_0)$, we obtain

$$\rho(1)\lambda \leq \rho(\lambda).$$

On the other hand, $\rho(0) = 0$ and $\rho(\lambda)$ is convex, thus we have $\rho(\lambda) \leq \lambda\rho(1) + (1-\lambda)\rho(0) = \lambda\rho(1)$ for $\lambda \in [0,1]$. Therefore, $\rho(\lambda) = \rho(1)\lambda$, hence $g(l(\lambda))$ is affine in $[0,1]$.

($\Leftarrow$) The idea is to show that there is a supporting hyperplane H of $\operatorname{epi} g \subseteq \mathbb{R}^n \times \mathbb{R}$ which contains the graph of g restricted to the segment joining x_0 and $\bar{x}$, that is, $A = \{(x_0 + \lambda(\bar{x} - x_0), g(x_0 + \lambda(\bar{x} - x_0))) \,:\, \lambda \in [0,1]\}$. Then the intersection of such H with $\mathbb{R}^n \times \{0\}$ will give us (2.3).

The set A is a convex nonempty subset of $\operatorname{epi} g$ that does not intersect the relative interior of $\operatorname{epi} g$. Hence, there exists a supporting hyperplane,

$$H = \{(x,z) \in \mathbb{R}^n \times \mathbb{R} \,:\, v^{\mathsf{T}}x + az = b\},$$

to $\operatorname{epi} g$ containing A (Rockafellar, 1970, Theorem 11.6).

Since $g(x_0) \leq 0$ and $g(\bar{x}) > 0$, it follows that A is not parallel to the x-space. Therefore, H is also not parallel to the x-space and so $v \neq 0$. Since A is not parallel to the z-axis, it follows that $a \neq 0$. We assume, without loss of generality, that $a = -1$.

The point $(\bar{x}, g(\bar{x}))$ belongs to $A \subseteq H$, thus $v^{\mathsf{T}}\bar{x} - g(\bar{x}) = b$ and $H = \{(x, g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x})) \,:\, x \in \mathbb{R}^n\}$. Given that H supports the epigraph, then v is a subgradient of g, in particular,

$$g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x}) \leq g(x) \text{ for every } x \in \mathbb{R}^n.$$

Let $z(x)$ be the affine function whose graph is H, that is, $z(x) = g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x})$. We now need to show that $g(\bar{x}) + v^{\mathsf{T}}(x - \bar{x}) \leq 0$ supports C by exhibiting an $\hat{x} \in C$ such that $g(\bar{x}) + v^{\mathsf{T}}(\hat{x} - \bar{x}) = 0$. By construction, $z(x_0 + \lambda(\bar{x} - x_0)) = g(x_0 + \lambda(\bar{x} - x_0))$. Since $z(x_0 + \lambda(\bar{x} - x_0))$ is non-positive for $\lambda = 0$ and positive for $\lambda = 1$, it has to be zero for some λ_0. Let $\hat{x} = x_0 + \lambda_0(\bar{x} - x_0)$. Then $g(\hat{x}) = z(\hat{x}) = 0$ and we conclude that $\hat{x} \in C$ and $g(\bar{x}) + v^{\mathsf{T}}(\hat{x} - \bar{x}) = 0$. □

Specializing the theorem to differentiable functions directly leads to the following:

Corollary 2.4. *Let $g\colon \mathbb{R}^n \to \mathbb{R}$ be a convex differentiable function, $C = \{x \in \mathbb{R}^n \,:\, g(x) \leq 0\}$, and $\bar{x} \notin C$. Then the valid inequality*

$$g(\bar{x}) + \nabla g(\bar{x})^{\mathsf{T}}(x - \bar{x}) \leq 0,$$

supports C, if and only if, there exists $x_0 \in C$ such that $\lambda \mapsto g(x_0 + \lambda(\bar{x} - x_0))$ is affine in $[0,1]$.

Proof. Since g is differentiable, the subdifferential of g consists only of the gradient of g. $\square$

A natural candidate for functions with supporting gradient cuts at every point are functions whose epigraph is a translation of a convex cone.

Corollary 2.5 (Sublinear functions). *Let $h(x)$ be a sublinear function. For this type of function, gradient cuts always support $C = \{x : h(x) \leq c\}$, for any $c \geq 0$.*

Proof. This follows directly from Theorem 2.3, since $0 \in C$ and $\lambda \mapsto h(\lambda\bar{x})$ is affine in $\mathbb{R}_+$ for any $\bar{x}$. $\square$

However, these are not the only functions that satisfy the conditions of Theorem 2.3 for every point. The previous theorem implies that linearizations always support the constraint set if a convex constraint $g(x) \leq 0$ is linear in one of its arguments.

Example 2.6 (Functions with linear variables). Let $f \colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be a convex function of the form $f(x, y) = g(x) + a^\mathsf{T} y + c$, with $a \neq 0$ and $g \colon \mathbb{R}^m \to \mathbb{R}$ convex. Then gradient cuts support $C = \{(x, y) : f(x, y) \leq 0\}$. Indeed, assume without loss of generality that $a_1 > 0$ and let $(\bar{x}, \bar{y}) \notin C$. Then there exists a $\lambda > 0$ such that $f(\bar{x}, \bar{y} - \lambda e_1) = g(\bar{x}) + a^\mathsf{T}\bar{y} + c - a_1\lambda = 0$. The statement follows from Theorem 2.3.

Consider separating a point (x_0, z_0) from a constraint of the form $z = g(x)$ with $g \colon \mathbb{R} \to \mathbb{R}$ and convex, with $z_0 < g(x_0)$ (that is, separating on the convex constraint $g(x) \leq z$). As mentioned earlier, Belotti et al. (2009) suggest projecting (x_0, z_0) to the graph $z = g(x)$ and computing a gradient cut there. This example shows that this step is unnecessary when the sole purpose is to obtain a cut that is supporting to the graph. $\square$

By contrast, if $g(x)$ is strictly convex, linearizations at points x such that $g(x) \neq 0$ are never supporting to $g(x) \leq 0$. This follows directly from Theorem 2.3 since $\lambda \mapsto g(x + \lambda v)$ is not affine for any v. We can also characterize convex quadratic functions with supporting linearizations.

Example 2.7 (Convex quadratic functions). Let $g(x) = x^\mathsf{T} A x + b^\mathsf{T} x + c$ be a convex quadratic function, i.e., A is an n by n symmetric and positive semidefinite matrix. We show that gradient cuts support $C = \{x \in \mathbb{R}^n : g(x) \leq 0\}$, if and only if, b is not in the range of A, i.e., $b \notin R(A) = \{Ax : x \in \mathbb{R}^n\}$.

First notice that $l_v(\lambda) = g(x + \lambda v)$ is affine linear, if and only if, $v \in \ker(A)$.

Let $v \in \ker(A)$ and $\bar{x} \notin C$. Then there is a $\lambda \in \mathbb{R}$ such that $\bar{x} + \lambda v \in C$ if and only if l_v is not constant. Thus, gradient cuts are *not* supporting, if and only if, l_v is constant for every $v \in \ker(A)$. But l_v is constant for every $v \in \ker(A)$, if and only if, $b^\mathsf{T} v = 0$ for every $v \in \ker(A)$, which is equivalent to $b \in \ker(A)^\perp = R(A^\mathsf{T}) = R(A)$, since A is symmetric. Hence, gradient cuts support C, if and only if, $b \notin R(A)$.

In particular, if $b = 0$, i.e., there are no linear terms in the quadratic function, then gradient cuts are never supporting hyperplanes. Also, if A is invertible, $b \in R(A)$ and gradient cuts are not supporting. This is to be expected since in this case g is strictly convex. □

2.3 The Gauge Function

Any MICP of form (2.1) can be reformulated to an equivalent MICP with a single constraint for which every linearization supports the continuous relaxation of the feasible region. To this end, we can use any sublinear function whose 1-sublevel set is C. Each convex set C has at least one sublinear function that represents it, namely, the *gauge function* (Rockafellar, 1970) of C.

Definition 2.8. *Let $C \subseteq \mathbb{R}^n$ be a convex set such that $0 \in \operatorname{int} C$. The* gauge *of C is*

$$\varphi_C(x) = \inf \{ \ t > 0 \ : \ x \in tC \ \} \, .$$

Proposition 2.9 (Tuy (2016, Proposition 1.11))**.** *Let $C \subseteq \mathbb{R}^n$ be a convex set such that $0 \in \operatorname{int} C$, then $\varphi_C(x)$ is sublinear. If, in addition, C is closed, then it holds that*

$$C = \{x \in \mathbb{R}^n \ : \ \varphi_C(x) \leq 1\}$$

and

$$\partial C = \{x \in \mathbb{R}^n \ : \ \varphi_C(x) = 1\}.$$

Combining Proposition 2.9 with Corollary 2.5, we can see that the gauge function is appealing for separation, because it always generates supporting hyperplanes.

2.3.1 Using the Gauge Function for Separation

Even though the gauge function is exactly what we need to ensure supporting gradient cuts, in general, there is no closed-form formula for it. Therefore, it is not always possible to explicitly reformulate C as $\varphi_C(x) \leq 1$.

Furthermore, if one is interested in solving mathematical programs with a numerical solver, performing such a reformulation might introduce some numerical issues one would have to take care of. Solvers usually solve up to a given tolerance, that is, they accept points that satisfy $g_j(x) \leq \varepsilon$ for some $\varepsilon > 0$. Then, even though $C = \{x : \varphi_C(x) \leq 1\}$, it might be that $\{x \in \mathbb{R}^n : \varphi_C(x) \leq 1 + \varepsilon\} \nsubseteq \{x \in \mathbb{R}^n : g_j(x) \leq \varepsilon\}$. In fact, even simple constraints show this behavior. Consider $C = \{x : x^2 - 1 \leq 0\}$. In this case, $\varphi_C(x) = |x|$ and for $x_0 = 1 + \varepsilon$, we have $\varphi_C(x_0) = 1 + \varepsilon$. Then x_0 would be ε-feasible for $\varphi_C(x) \leq 1$, although it would be infeasible for $x^2 - 1 \leq 0$, since $2\varepsilon + \varepsilon^2 > \varepsilon$.

Luckily, one does not need to reformulate in order to take advantage of the gauge function for tighter separation. The next propositions show how to use the gauge function and a point $\bar{x} \notin C$ to obtain a boundary point of C and that linearizing at that boundary point gives a supporting valid inequality that actually separates $\bar{x}$. For ensuring the existence of a supporting hyperplane we need Assumption 2.2. For example, Assumption 2.2 is satisfied whenever Slater's condition (Section 1.3) is satisfied for (2.1) with C represented by (2.2), that is, when there exists x_0 such that $g_j(x_0) < 0$ for every $j \in J$.

Before we state the propositions we start with a simple lemma.

Lemma 2.10. *Let $C \subseteq \mathbb{R}^n$ be a closed convex set such that $0 \in \operatorname{int} C$, let $\hat{x} \in \partial C$ and $\bar{x} \notin C$. Let $\alpha \in \mathbb{R}^n, \beta \in \mathbb{R}$ such that $\alpha \neq 0$ and $\alpha^\mathsf{T} x \leq \beta$ is a valid inequality for C that supports C at $\hat{x}$. If the segment joining 0 and $\bar{x}$ contains $\hat{x}$, then the inequality separates $\bar{x}$ from C.*

Proof. Consider $l(\lambda) = \alpha^\mathsf{T}(\lambda \bar{x}) - \beta$ and let $\lambda_0 \in (0,1)$ be such that $\lambda_0 \bar{x} = \hat{x}$. The function l is a strictly increasing affine linear function. Indeed, $0 \in \operatorname{int} C$ implies that $l(0) < 0$, while $l(\lambda_0) = 0$. Thus, $l(1) > 0$, i.e., $\alpha^\mathsf{T} \bar{x} > \beta$. □

Proposition 2.11. *Let $C \subseteq \mathbb{R}^n$ be a closed convex set such that $0 \in \operatorname{int} C$ and let $\bar{x} \notin C$. Then $\frac{\bar{x}}{\varphi_C(\bar{x})} \in \partial C$.*

Proof. First, $\varphi_C(\bar{x}) \neq 0$ since $\bar{x} \notin C$. The positive homogeneity of φ_C implies that $\varphi_C\left(\frac{\bar{x}}{\varphi_C(\bar{x})}\right) = \frac{\varphi_C(\bar{x})}{\varphi_C(\bar{x})} = 1$. Proposition 2.9 implies $\frac{\bar{x}}{\varphi_C(\bar{x})} \in \partial C$. □

Let $J_0(x)$ be the set of indices of the active constraints at x, i.e., $J_0(x) = \{j \in J : g_j(x) = 0\}$.

Proposition 2.12. *Let $C = \{x : g_j(x) \leq 0, j \in J\}$ be such that $0 \in \operatorname{int} C$ and let φ_C be its gauge function. Assume that Assumption 2.2 holds. Given $\bar{x} \notin C$, define $\hat{x} = \frac{\bar{x}}{\varphi_C(\bar{x})}$. Then, for any $j \in J_0(\hat{x})$, the gradient cut of g_j at $\hat{x}$ yields a valid supporting inequality for C that separates $\bar{x}$.*

Proof. By the previous proposition, we have that $\hat{x} \in \partial C$. Let $j \in J_0(\hat{x})$. Then the gradient cut of g_j at $\hat{x}$ yields a valid supporting inequality. The fact that it separates follows from Lemma 2.10. Note that Lemma 2.10 is applicable since Assumption 2.2 ensures that the normal of the gradient cut is nonzero. □

Hence, we can get supporting valid inequalities separating a given point $\bar{x} \notin C$ by using the gauge function to find the point $\hat{x} = \frac{\bar{x}}{\varphi_C(\bar{x})} \in \partial C$. Then Proposition 2.12 ensures that the gradient cut of any active constraint at $\hat{x}$ will separate $\bar{x}$ from C. But how do we compute $\varphi_C(\bar{x})$?

2.3.2 Evaluating the Gauge Function

Let $C = \{x : g_j(x) \leq 0, j \in J\}$ be a closed convex set such that $0 \in \operatorname{int} C$ and consider

$$f(x) = \max_{j \in J} g_j(x). \tag{2.5}$$

In general, evaluating the gauge function of C at $\bar{x} \notin C$ is equivalent to solving the following one dimensional equation

$$f(\lambda \bar{x}) = 0, \;\; \lambda \in (0, 1). \tag{2.6}$$

If λ^* is the solution, then $\varphi_C(\bar{x}) = \frac{1}{\lambda^*}$.

One can solve such an equation using a line search. Note that the line search is looking for a point $\hat{x} \in \partial C$ on the segment between 0 and $\bar{x}$. This is exactly what the (extended) supporting hyperplane algorithm performs when it uses 0 as its interior point.

We would also like to remark that a closed-form formula expression for the gauge function of C is equivalent to a closed-form formula for the solution of (2.6). It is possible to find such a formula for some functions, e.g., when f is a convex quadratic function.

Next, we briefly discuss what happens when 0 is not in the interior of C and when C has no interior. In the next section we discuss the implications of the fact that evaluating the gauge function is equivalent to the line search step of the supporting hyperplane algorithm.

2.3.3 Handling Sets with Empty Interior

When $\operatorname{int} C = \emptyset$, we can still use the methods discussed above by applying a trick from Kronqvist et al. (2016). Assuming $C = \{x \in \mathbb{R}^n : g_j(x) \leq 0, j \in J\} \neq \emptyset$, consider the set $C_\epsilon = \{x \in \mathbb{R}^n : g_j(x) \leq \epsilon, j \in J\}$. This set satisfies $\operatorname{int} C_\epsilon \neq \emptyset$ and optimizing over C_ϵ provides an ϵ-optimal solution.

2.3.4 Using a Nonzero Interior Point

If $x_0 \in \operatorname{int} C$ and $x_0 \neq 0$, we can translate C so that 0 is in its interior. Equivalently, we can build a gauge function centered on x_0. This is given by

$$\varphi_{x_0,C}(x) = \varphi_{C-x_0}(x - x_0).$$

Then, given $\bar{x} \notin C$, the point

$$\hat{x} = \frac{\bar{x} - x_0}{\varphi_{C-x_0}(\bar{x} - x_0)} + x_0 \tag{2.7}$$

belongs to the boundary of C. Equivalently, $\hat{x} = x_0 + \lambda^*(\bar{x} - x_0)$, where λ^* solves

$$f(x_0 + \lambda(\bar{x} - x_0)) = 0,\ \ \lambda \in (0,1),$$

with $f(x) = \max_{j\in J} g_j(x)$ as in (2.5).

2.4 Convergence Proofs

Consider an MICP given by (2.1) with C represented as (2.2). Let f be defined as in (2.5). As mentioned above, the ESH algorithm computes an interior point of C (which we will assume to be 0) and performs a line search between $\bar{x} \notin C$ and 0 in order to find a point on the boundary. It computes a gradient cut at the boundary point, solves the relaxation again, and repeats the process. From our previous discussion, computing a gradient cut at the boundary point is equivalent to computing a gradient cut at $\frac{\bar{x}}{\varphi_C(\bar{x})}$. Therefore, the generated cuts are $f(\frac{\bar{x}}{\varphi_C(\bar{x})}) + v^\mathsf{T}(x - \frac{\bar{x}}{\varphi_C(\bar{x})}) \leq 0$, where $v \in \partial f(\frac{\bar{x}}{\varphi_C(\bar{x})})$.

To prove the convergence of the ESH algorithm, Veinott and Kronqvist et al. use tailored arguments. Here we show that the convergence of the algorithm follows from the convergence of KCP. We note that the KCP algorithm still converges when C is represented by a convex non-differentiable function. One needs to replace gradients by subgradients and one can use any subgradient (Horst and Tuy, 1990). Therefore, given that $\varphi_C(x)$ is a convex function, we know that KCP converges when applied to $\min\{c^\mathsf{T}x : \varphi_C(x) \leq 1\}$. Thus, in order to prove that ESH converges, it is sufficient to show that the cutting planes generated by ESH can also be generated by KCP.

We first prove that the normals of (normalized) supporting valid inequalities are subgradients of the gauge function at the supporting point.

Lemma 2.13. *Let $\alpha^\mathsf{T}x \leq 1$ be a valid and supporting inequality for C. Let $\hat{x} \in \partial C$ be a point where it supports C, i.e., $\alpha^\mathsf{T}\hat{x} = 1$. Then $\alpha \in \partial\varphi_C(\hat{x})$.*

Proof. We need to show that $\varphi_C(\hat{x}) + \alpha^\mathsf{T}(x - \hat{x}) \leq \varphi_C(x)$ for every x. Note that since $\hat{x} \in \partial C$, we have that $\varphi_C(\hat{x}) = 1$ and we just have to prove that $\alpha^\mathsf{T} x \leq \varphi_C(x)$.

When x is such that $\varphi_C(x) > 0$, we have $\frac{x}{\varphi_C(x)} \in C$. Due to the validity of $\alpha^\mathsf{T} x \leq 1$, it follows that $\alpha^\mathsf{T} \frac{x}{\varphi_C(x)} \leq 1$.

Now let x be such that $\varphi_C(x) = 0$. Then $\varphi_C(\lambda x) = 0$ for every $\lambda > 0$, i.e., $\lambda x \in C$ for every $\lambda > 0$. Hence, $\alpha^\mathsf{T}(\lambda x) \leq 1$ for every $\lambda > 0$ which implies that $\alpha^\mathsf{T} x \leq 0 = \varphi_C(x)$. □

Now we prove that the inequalities generated by the ESH algorithm can also be generated by the KCP algorithm. Given that the KCP algorithm converges even for non-smooth convex function (Horst and Tuy, 1990), the next theorem implies the convergence of the ESH algorithm.

Theorem 2.14. *Consider an MICP given by* (2.1) *with C represented as* (2.2) *such that $0 \in \operatorname{int} C$ and Assumption 2.2 holds. Let f be defined as in* (2.5) *and let $\bar{x} \notin C$ be the current relaxation solution to separate. Let $f(\frac{\bar{x}}{\varphi_C(\bar{x})}) + v^\mathsf{T}(x - \frac{\bar{x}}{\varphi_C(\bar{x})}) \leq 0$, with $v \in \partial f(\frac{\bar{x}}{\varphi_C(\bar{x})})$, be the inequality generated by the ESH algorithm using 0 as the interior point. Then KCP applied to $\min\{c^\mathsf{T} x : \varphi_C(x) \leq 1\}$ can generate the same inequality.*

Proof. Let $\hat{x} = \frac{\bar{x}}{\varphi_C(\bar{x})}$. First, let us show that Assumption 2.2 implies $v \neq 0$. Indeed, if $v = 0$, then $f(\hat{x}) + v^\mathsf{T}(x - \hat{x}) \leq f(x)$ and $0 \in C$ imply that $0 \geq f(0) \geq f(\hat{x}) + v^\mathsf{T}(0 - \hat{x}) = 0$. Let $j \in J$ be such that $g_j(0) = f(0) = 0$. Then $\lambda \mapsto g_j(\lambda \hat{x})$ is constant in $[0, 1]$. Thus, its derivative at 1 is 0, i.e., $\nabla g_j(\hat{x})^\mathsf{T} \hat{x} = 0$. This implies that $\nabla g_j(\hat{x})^\mathsf{T} \bar{x} = 0$. Furthermore, $\nabla g_j(\hat{x}) \neq 0$ by Assumption 2.2 and so Lemma 2.10 implies that $\nabla g_j(\hat{x})^\mathsf{T}(x - \hat{x}) \leq 0$ separates $\bar{x}$ from C. But this contradicts the equality $\nabla g_j(\hat{x})^\mathsf{T} \bar{x} = 0$.

Let us manipulate the inequality obtained by the ESH algorithm. Notice that $f(\hat{x}) = 0$ and so the inequality reads as $v^\mathsf{T} x \leq v^\mathsf{T} \hat{x}$. By Lemma 2.10, $\bar{x}$ is cut off by $v^\mathsf{T} x \leq v^\mathsf{T} \hat{x}$, i.e., $v^\mathsf{T} \bar{x} > v^\mathsf{T} \hat{x}$. This, together with $\varphi_C(\bar{x}) > 1$, implies that $v^\mathsf{T} \bar{x} > 0$. Summarizing, the inequality obtained by the ESH algorithm can be rewritten as

$$\left(\frac{\varphi_C(\bar{x})}{v^\mathsf{T} \bar{x}} v \right)^\mathsf{T} x \leq 1.$$

Lemma 2.13 implies that $\frac{\varphi_C(\bar{x})}{v^\mathsf{T} \bar{x}} v \in \partial \varphi_C(\hat{x})$. Since φ_C is positively homogeneous, $\partial \varphi_C(\hat{x}) = \partial \varphi_C(\bar{x})$. Hence, if the KCP algorithm applied to $\min\{c^\mathsf{T} x : \varphi_C(x) \leq 1\}$ separates $\bar{x}$ using $\frac{\varphi_C(\bar{x})}{v^\mathsf{T} \bar{x}} v \in \partial \varphi_C(\bar{x})$, then it would generate the gradient cut

$$\varphi_C(\bar{x}) - 1 + \tfrac{\varphi_C(\bar{x})}{v^\mathsf{T} \bar{x}} v^\mathsf{T}(x - \bar{x}) \leq 0.$$

The left hand side of the above inequality is equivalent to $-1 + \frac{\varphi_C(\bar{x})}{v^\mathsf{T}\bar{x}} v^\mathsf{T} x$. This shows that the gradient cut constructed by the KCP algorithm is the same as the one construction by the ESH algorithm. □

2.5 Convex Programs Represented by Non-Convex Non-Smooth Functions

In this section we consider problem (2.1) with C represented as

$$C = \{x : g_j(x) \leq 0, j \in J\},$$

where the functions g_j are not necessarily convex. As mentioned in the introduction, convex problems represented by non-convex functions have been considered in Dutta and Lalitha (2011); Kabgani et al. (2017); Lasserre (2009, 2011, 2014); Martínez-Legaz (2014). These different works have generalized each other by considering more general classes of non-smooth functions.

2.5.1 The ESH Algorithm in the Context of Generalized Differentiability

When a function is non-smooth there are many ways of extending the notion of differentiability. Informally, it is common to first define a notion of directional derivative and then a generalization of the gradient. As the directional derivative of g at x in the direction d is given by $\nabla g(x)^\mathsf{T} d$, the notion of generalized gradient tries to capture this relation.

A classic notion of generalized derivative is Clarke's subdifferential.

Definition 2.15 (Clarke (1990); Clarke et al. (1998))**.** *The* Clarke directional derivative *of a function $g : \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}$ in the direction $d \in \mathbb{R}^n$ is defined as*

$$g^\circ(\bar{x}; d) = \limsup_{x \to \bar{x}, t \searrow 0} \frac{g(x + td) - g(x)}{t}.$$

The Clarke subdifferential *of g at $\bar{x}$ is*

$$\partial^\circ g(\bar{x}) = \{\eta \in \mathbb{R}^n : \eta^\mathsf{T} d \leq g^\circ(\bar{x}; d) \, \forall d \in \mathbb{R}^n\}.$$

We say that g is directionally differentiable *at $\bar{x}$ if directional derivatives of g at $\bar{x}$ exist, that is,*

$$g'(\bar{x}; d) = \lim_{t \searrow 0} \frac{g(\bar{x} + td) - g(\bar{x})}{t},$$

exists for every $d \in \mathbb{R}^n$. Finally, g is regular in the sense of Clarke *at $\bar{x}$ if the g is directional differentiable at $\bar{x}$ and $g'(\bar{x}; d) = g^\circ(\bar{x}; d)$ for every $d \in \mathbb{R}$.*

Another interesting class is the following.

Definition 2.16 (Jeyakumar and Luc (1999))**.** *Let* $g : \mathbb{R}^n \to \mathbb{R}$*. The* upper Dini directional derivative *of* g *at* $\bar{x}$ *in the direction* $d \in \mathbb{R}^n$ *is*

$$g^+(\bar{x}; d) = \limsup_{t \searrow 0} \frac{g(\bar{x} + td) - g(\bar{x})}{t}.$$

The function g *has an* upper regular convexificator *(URC) at* $\bar{x}$ *if there exists a closed set* $\partial^+ g(\bar{x}) \subseteq \mathbb{R}^n$ *such that for each* $d \in \mathbb{R}^n$*,*

$$g^+(\bar{x}; d) = \sup_{\alpha \in \partial^+ g(\bar{x})} \alpha^\mathsf{T} d.$$

We abstract the notion of directional derivative and subdifferential as follows.

Definition 2.17. *Let* $g : \mathbb{R}^n \to \mathbb{R}$ *be a function. A* generalized directional derivative *of* g *is a function* $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$*, and the generalized directional derivative of* g *at* x *in the direction* d *is* $h(x; d)$*. We say that* g admits a generalized subdifferential *at* x *if there exists* $A = A(x) \subseteq \mathbb{R}^n$ *such that* $h(x; d) = \sup_{v \in A(x)} v^\mathsf{T} d$ *for all* $d \in \mathbb{R}^n$*.*

For example, if g is locally Lipschitz, then Clarke's directional derivative is a generalized directional derivative and $\partial^\circ g(x)$ is a generalized subdifferential as $g^\circ(x; d) = \sup\{v^\mathsf{T} d : v \in \partial^\circ g(x)\}$ (Clarke et al., 1998, Proposition 2.1.5). Or, if g admits a URC, then Dini's directional derivative is a generalized directional derivative that admits a generalized subdifferential.

However, the above definition of generalized directional derivative and subdifferential is so general, that any support function of a set yields a generalized directional derivative that admits a generalized subdifferential. The following definition adds a further requirement in order to make this general notion useful.

Definition 2.18. *Let* h *be a generalized directional derivative of* g*. We say that the generalized directional derivative is* well-behaved *if* $h(x; d) > 0$ *implies that there exists* $t_n \searrow 0$ *such that* $g(x + t_n d) > g(x)$*.*

As we will see, this is the key property to show that the ESH algorithm converges.

Clearly, if g is differentiable, then the directional derivative is well-behaved. Also, Dini's directional derivative is well-behaved. As we will see in the next section, Clarke's directional derivative is not well-behaved in general. However, if the function is regular in the sense of Clarke, then it is well-behaved. Another important class of functions for which Clarke's directional derivative is well-behaved is the class of ∂°-pseudoconvex functions.

Definition 2.19. *A function $g : \mathbb{R}^n \to \mathbb{R}$ is* ∂°-pseudoconvex *if*

– *it is locally Lipschitz and,*

– *for every $x, y \in \mathbb{R}^n$, if $g(y) < g(x)$, then $g^\circ(x; y - x) < 0$*

To show that it is well-behaved, we need to following result.

Lemma 2.20 (Bagirov et al. (2014, Lemma 5.3))**.** *If a function g is ∂°-pseudoconvex, then for every $x, y \in \mathbb{R}^n$, if $g(y) = g(x)$, then $g^\circ(x; y - x) \leq 0$. In particular, if $g(y) \leq g(x)$, then $g^\circ(x; y - x) \leq 0$.*

The contrapositive of the last statement is if $g^\circ(x; y - x) > 0$, then $g(y) > g(x)$. As $g^\circ(x; \cdot)$ is positively homogeneous (Clarke et al., 1998, Proposition 2.1.1), we conclude that if g is ∂°-pseudoconvex, $g^\circ(x; d) > 0$ for some $d \in \mathbb{R}^n$, and $t > 0$, then $g(x + td) > g(x)$. Thus, if g is ∂°-pseudoconvex, then Clarke's directional derivative is well-behaved.

Now we are ready to prove the main result of this section. Recall that $J_0(x) = \{j \in J \,:\, g_j(x) = 0\}$.

Theorem 2.21. *Let $C = \{x \,:\, g_j(x) \leq 0, j \in J\}$ be such that C is convex, closed, and $0 \in \operatorname{int} C$. Assume that for each $x \in C$ and $j \in J_0(x)$, the function g_j has a well-behaved generalized directional derivative at x denoted by h_j, and that it admits a generalized subdifferential, $\partial^* g_j(x)$. Furthermore, assume that*

$$\partial^* g_j(x) \setminus \{0\} \neq \emptyset \quad \text{for all } x \in C \text{ and } j \in J_0(x). \tag{2.8}$$

Let φ_C be the gauge function of C. For $\bar{x} \notin C$, define $\hat{x} = \frac{\bar{x}}{\varphi_C(\bar{x})}$. Then, for every $j \in J_0(\hat{x})$ and every $v \in \partial^* g_j(\hat{x}) \setminus \{0\}$, the gradient cut, $g_j(\hat{x}) + v^\mathsf{T}(x - \hat{x}) \leq 0$, is a valid supporting inequality for C that separates $\bar{x}$.

Proof. By Proposition 2.11 we have that $\hat{x} \in \partial C$. Let $j \in J_0(\hat{x})$ and let us a consider an arbitrary $v \in \partial^* g_j(\hat{x}) \setminus \{0\}$. The gradient cut of g_j at $\hat{x}$ is $v^\mathsf{T}(x - \hat{x}) \leq 0$.

We first show that the gradient cut is valid, that is, $v^\mathsf{T}(y - \hat{x}) \leq 0$ for all $y \in C$. If this is not the case, then there exists $y_0 \in C$ for which $v^\mathsf{T}(y_0 - \hat{x}) > 0$.

Since g_j admits a generalized subdifferential at $\hat{x}$, we have that

$$h_j(\hat{x}; y_0 - \hat{x}) = \sup_{\eta \in \partial^* g_j(\hat{x})} \eta^\mathsf{T}(y_0 - \hat{x}).$$

As $v \in \partial^* g_j(\hat{x})$, it follows that $h_j(\hat{x}; y_0 - \hat{x}) > 0$. Since h_j is well-behaved, there is a sufficiently small $t \in (0, 1)$ such that $g_j(\hat{x} + t(y_0 - \hat{x})) > 0$. Thus, $\hat{x} + t(y_0 - \hat{x}) \notin C$. However, the convexity of C implies that $\hat{x} + \lambda(y_0 - \hat{x}) \in C$ for $\lambda \in [0, 1]$, which is a contradiction.

The fact that the gradient cut separates $\bar{x}$ follows from Lemma 2.10. Note that $v \neq 0$ by hypothesis. □

Theorem 2.21 extends the algorithm of Veinott to further representations of the set C. In particular, it implies that the ESH converges (via an argument similar to Theorem 2.14's proof) when the constraints admit a URC or are ∂°-pseudoconvex. Thus, it generalizes the result of Eronen et al. (2017).

Remark 2.22. Any representation of a convex set C as $\{x \in \mathbb{R}^n : g_j(x) \leq 0, j \in J\}$ yields a way to evaluate its gauge function, namely,

$$\varphi_C(x) = \inf \left\{ t > 0 \,:\, \max_j g_j\left(\frac{x}{t}\right) = 0 \right\}.$$

This infimum can be computed using a line search procedure.

However, what is more important is the ability to compute subgradients. Given any method to compute subgradients of the gauge function, we can apply the KCP algorithm using the implicitly defined gauge function. This allows us, for example, to drop (2.8). This algorithm is more general than the one proposed by Lasserre (2011), but it will not necessarily converge to a KKT point of the original problem. □

2.5.2 Limits to the Applicability of the ESH Algorithm

The idea of the proof of Theorem 2.21 is that since C is convex, $\hat{x}+\lambda(y-\hat{x}) \in C$ for every $y \in C$ and $\lambda \in [0, 1]$. Hence, the functions g_j do not increase when moving in the direction $y - \hat{x}$ from $\hat{x}$. Thus, a notion of subdifferential that characterizes a well-behaved directional derivative yields valid gradient cuts. The abstract definitions introduced above try to capture this line of reasoning.

Note that this is also how the proofs of the 'only if' parts of (Lasserre, 2009, Lemma 2.2), (Kabgani et al., 2017, Theorem 1), (Dutta and Lalitha, 2011, Proposition 2.2), and the $\subseteq$ inclusion of (Martínez-Legaz, 2014, Proposition 6) work. For example, Lasserre (2009) assumes that the g_j is differentiable, in which case the generalized subdifferential is just the singleton given by the gradient and the generalized directional derivative is the classic directional derivative. Dutta and Lalitha (2011) assume that the functions are locally Lipschitz and regular in the sense of Clarke.

It is a natural question to wonder how important the regularity assumption is. As the following example shows, the ESH algorithm can produce invalid cutting planes when using Clarke's subdifferential and the constraints are not regular in the sense of Clarke. In particular, this shows that, without the assumption of regularity, Clarke's directional derivative is not well-behaved, in general.

Example 2.23. Consider the function $g(x_1, x_2) = \max\{\min\{3x_1 + x_2, 2x_1 + 3x_2\}, x_1\}$. The set $C = \{(x_1, x_2) : g(x_1, x_2) \leq 0\}$ is convex, closed and its interior is nonempty as shown in Figure 2.2. Note that as g is piecewise linear, it is globally Lipschitz continuous (Scholtes, 2012, Proposition 2.2.7). Using Clarke et al. (1998, Theorem 2.8.1), it follows that $\partial^\circ g(0) = \operatorname{conv}\{(3,1), (2,3), (1,0)\}$. Then $2x_1 + 3x_2 \leq 0$ is a gradient cut of g at 0. However, it is not valid as $(-1, 3)$ is feasible but $-2 + 9 > 0$.

In particular, it must be that g is not regular in the sense of Clarke and that g° is not well-behaved. To see that g is not well-behaved, consider the direction $d = (-1, 1)$. Notice that $g((0,0) + td) = tg(-1, 1) = -t$, and so g is strictly decreasing in the direction d. However, $g^\circ(0; d) = \max_{v \in \partial^\circ g(0)} -v_1 + v_2 = 1$. This also shows that g is not regular. The directional derivative of g at 0 in the direction d is $-1 \neq 1$. □

2.6 Concluding Remarks

In this chapter, we have shown that the extended supporting hyperplane algorithm introduced by Veinott (1967) and rediscovered by Kronqvist et al. (2016) is identical to Kelley's classic cutting plane algorithm applied to a suitable reformulation of the problem. We used this new perspective in order to prove the convergence of the method for the larger class of problems with convex feasible regions represented by non-convex non-smooth constraints which admit a generalized subdifferential and whose generalized directional derivative is well-behaved. This class includes ∂°-pseudoconvex functions and functions that admit a URC. Functions that admit a URC include differentiable functions and locally Lipschitz functions that are regular in the sense of Clarke. More generally, the algorithm extends to any representation of a convex set that allows to compute subgradients of its gauge function. These theoretical results bear relevance in practice, as the experimental results in Kronqvist et al.

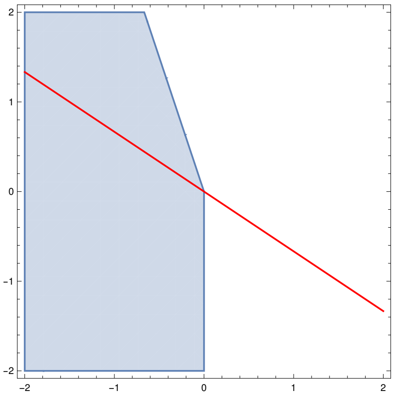


Figure 2.2: Counterexample showing that, in general, the ESH algorithm can generate invalid cutting planes if the constraints are just Lipschitz continuous. The convex feasible region $\max\{\min\{3x_1 + x_2, 2x_1 + 3x_2\}, x_1\} \leq 0$ in blue and the boundary of the invalid gradient cut $2x_1 + 3x_2 \leq 0$ in red.

(2016, 2018) have already demonstrated the computational benefits of the supporting hyperplane algorithm in comparison to alternative state-of-the-art solving methods.

Another intuition gain from this chapter, which we will use in Chapter 5, is that if we want the gradient cuts to be supporting, then the constraint function cannot be "too" convex. Indeed, as we saw, gradient cuts from strictly convex functions will never be supporting.

Chapter 3

Visible Points, the Separation Problem, and Applications to Mixed-Integer Nonlinear Programming

From now on we move away from convex mixed-interger non-linear programs and consider non-convex mixed-integer linear programs. In this chapter we introduce a technique to produce tighter cutting planes for mixed-integer non-linear programs. Usually, a cutting plane is generated to cut off a specific infeasible point. The underlying idea is to use the infeasible point to restrict the feasible region in order to obtain a tighter domain. To ensure validity, we require that every valid cut separating the infeasible point from the restricted feasible region is still valid for the original feasible region. We translate this requirement in terms of the separation problem and the reverse polar. In particular, if the reverse polar of the restricted feasible region is the same as the reverse polar of the original feasible region, then any cut valid for the restricted feasible region that *separates* the infeasible point, is also valid for the original feasible region.

We show that the reverse polar of the so-called visible points of the feasible region from the infeasible point coincides with the reverse polar of the feasible region. In the special case where the feasible region is described by a single non-convex constraint intersected with a convex set we provide a characterization of the visible points. Furthermore, when the non-convex constraint is quadratic the characterization is particularly simple. We also provide an extended formulation for a relaxation of the visible points when the non-convex constraint is a general polynomial.

Finally, we give some conditions under which for a given set there is an inclusion-wise smallest set, in some predefined family of sets, whose reverse polars coincide.

3.1 Introduction

The separation problem is a fundamental problem in optimization (Grötschel et al., 1993). Given a set $S \subseteq \mathbb{R}^n$ and a point $\bar{x} \in \mathbb{R}^n$, the separation problem is

> Decide if $\bar{x}$ is in the closure of convex hull of S or find a valid for S that separates $\bar{x}$.

Algorithms to solve optimization problems, especially those based on solving relaxations, such as branch and bound, need to deal with the separation problem. Consider, for example, solving a mixed integer linear problem via branch and bound (Conforti et al., 2014, Section 9.2). The solution to the linear relaxation plays the role of $\bar{x}$, while a relaxation based on a subset of the constraints is used as S for the separation problem, see (Conforti et al., 2014, Chapter 6).

The separation problem can be rephrased in terms of the *reverse polar* (Balas, 1998; Zaffaroni, 2008) of S at $\bar{x}$, defined as

$$S^{\bar{x}} = \{\alpha \in \mathbb{R}^n : \alpha^{\mathsf{T}}(x - \bar{x}) \geq 1, \forall x \in S\}.$$

The elements of $S^{\bar{x}}$ are the normals of the hyperplanes that separate $\bar{x}$ from $\overline{\operatorname{conv}}\, S$. Hence, the separation problem can be stated equivalently as

> Decide if $S^{\bar{x}}$ is empty or find an element from it.

The point of departure of the present work is the following observation.

Observation 3.1. If there is a set V such that $(S \cap V)^{\bar{x}} = S^{\bar{x}}$, then, as far as the separation problem is concerned, the feasible region can be regarded as $S \cap V$ instead of S.

A set V such that $V^{\bar{x}} = S^{\bar{x}}$ will be called a *generator* of $S^{\bar{x}}$. Intuitively, if a set V is such that $V \cap S$ generates $S^{\bar{x}}$, that is, if we can ensure that a cut valid for $V \cap S$ that separates $\bar{x}$ is also valid for S, then V should at least contain the points of S that are "near" $\bar{x}$. To formalize the meaning of "near" we use the concept of *visible points* (Deutsch et al., 2013) of S from $\bar{x}$, which are the points $x \in S$ for which the segment joining x with $\bar{x}$ only intersects S at x, see Definition 3.5. In other words, they are the points of S that can be "seen" from $\bar{x}$. In Proposition 3.9 we show that the visible points are a generator of $S^{\bar{x}}$.

As a motivation, we present an application of our results in the context of nonlinear programming, which is treated in more detail in Section 3.4.

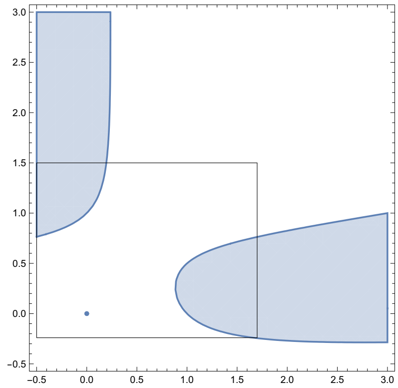


Figure 3.1: The feasible region $g(x) \leq 0$ and $\bar{x} = (0, 0)$ together with the box V.

Example 3.2. Consider the separation problem of $\bar{x} = (0, 0)$ from $S = \{x \in B : g(x) \leq 0\}$ where

$$B = [-\tfrac{1}{2}, 3] \times [-\tfrac{1}{2}, 3],$$
$$g(x_1, x_2) = -x_1^2 x_2 + 5x_1 x_2^2 - x_2^2 - x_2 - 2x_1 + 2,$$

as depicted in Figure 3.1. A standard technique for solving the separation problem for S and $\bar{x}$ is to construct a convex underestimator of g over B (Vigerske, 2013, Sections 6.1.2 and 7.5.1). The quality of a convex underestimator depends on the bounds of the variables and tighter bounds yield tighter underestimators. As we will see (Proposition 3.9 and Theorem 3.27), $R^{\bar{x}} = S^{\bar{x}}$ where

$$R = \{x \in B : g(x) = 0, \nabla g(x)^\mathsf{T} x \leq 0\}.$$

It is possible to show that $R \subseteq V$, where $V = [-\frac{1}{2}, \frac{17}{10}] \times [-\frac{6}{25}, \frac{3}{2}]$. Hence, by Corollary 3.25, $(V \cap S)^{\bar{x}} = S^{\bar{x}}$. This means that we can solve the separation problem over $\{x \in V : g(x) \leq 0\}$ instead of S. Therefore, if we were to compute an underestimator of g, it could be computed over $V \subsetneq B$. $\square$

Methods for obtaining tighter bounds for mixed integer nonlinear programming (MINLP) are of paramount importance. Indeed, not only bound tightening procedures enhance the performance of MINLP solvers, but also many algorithms for solving MINLPs require that all variables are bounded (Hamed and McCormick, 1993). We refer to the recent survey of Puranik and Sahinidis (2017) for more information on bound tightening procedures and its impact on MINLP solvers.

However, the technique that we introduce in this chapter is *not* a bound tightening technique in the classic sense, i.e., the tighter bounds that might be learned from V are not valid for the original problem, but only for the separation problem at hand.

We would like to point out that Venkatachalam and Ntaimo (2016) discusses a similar idea — to modify the separation problem — is used in the context of stochastic mixed integer programming. Their objective is to speed-up the solution of the separation problem. In contrast, our objective is to produce tighter cutting planes for MINLP.

Contributions We show that for every closed set S, there exists an inclusion-wise smallest closed convex set that generates $S^{\bar{x}}$ (Theorem 3.21). When S is compact, there is an inclusion-wise smallest closed set that generates $S^{\bar{x}}$ (Theorem 3.23). Furthermore, under some mild assumptions on S, we show that there is an inclusion-wise smallest closed convex set C such that $C \cap S$ generates $S^{\bar{x}}$ (Theorem 3.22). We also show the existence of a generator, $V_S(\bar{x})$, of $S^{\bar{x}}$ which is more suitable for computations.

We apply our results to MINLP and give an explicit description of $V_S(\bar{x})$ when $S = \{x \in C : g(x) \leq 0\}$, where C is a closed convex set containing $\bar{x}$, and g is continuous (Section 3.4.1). For the important case of quadratic constraints, i.e., when g is a quadratic function, we show that $V_S(\bar{x})$ has a particularly simple expression (Theorem 3.29).
For the case when g is a general polynomial, we provide an extended formulation for a relaxation of $V_S(\bar{x})$ based on the theory of non-negative univariate polynomials (Theorem 3.34).

3.2 Visible Points and the Reverse Polar

In this section we introduce the concept of visible points and reverse polar, and state some basic properties about them. The main result in this section is that the reverse polar of the visible points of a set is the reverse polar of the set (Proposition 3.9).

Unless stated otherwise, we will assume $\bar{x} = 0$. This is without loss of generality, since we can always translate the set S to $S - \bar{x}$. We start by restating the definition of reverse polar.

Definition 3.3. *Let $S \subseteq \mathbb{R}^n$ and $\bar{x} \in \mathbb{R}^n$. The* reverse polar *of S at $\bar{x}$ is*

$$S^{\bar{x}} = \{\alpha \in \mathbb{R}^n : \alpha^\mathsf{T}(x - \bar{x}) \geq 1, \text{ for all } x \in S\}.$$

As stated in the introduction, the reverse polar contains all valid inequalities for S that separate $\bar{x}$ from S.

Definition 3.4. *Let $S, V \subseteq \mathbb{R}^n$ and $\bar{x} \in \mathbb{R}^n$. We say that V is a* generator *of $S^{\bar{x}}$ if and only if*

$$V^{\bar{x}} = S^{\bar{x}}$$

Definition 3.5. *Let $S \subseteq \mathbb{R}^n$ be closed and $\bar{x} \notin S$. The set of* visible points *of S from $\bar{x}$ is*

$$\begin{aligned} V_S(\bar{x}) &= \{x \in S : (x + [0,1](\bar{x} - x)) \cap S = \{x\}\} \\ &= \{x \in S : (x + (0,1](\bar{x} - x)) \cap S = \emptyset\}. \end{aligned}$$

We denote $V_S(0)$ by V_S and note that

$$V_S = \{x \in S : [0,1]x \cap S = \{x\}\} = \{x \in S : [0,1)x \cap S = \emptyset\}.$$

The following concept is, in some sense, the opposite of the visible points.

Definition 3.6. *Let $S \subseteq \mathbb{R}^n$ be closed. The* shadow *of S from 0 is*

$$\operatorname{shw} S = [1, \infty)S.$$

The concept of shadow has also been called *penumbra* (Rockafellar, 1970, p. 22),(Tind and Wolsey, 1982; Conforti and Wolsey, 2018) and *aureole closure* (Ruys, 1974). The followings are some basic properties of the reverse polar.

Lemma 3.7. *Ruys (1974, Property 9.2.2) Let $S, T \subseteq \mathbb{R}^n$. Then,*

1. $S^{\mathbf{0}} = (\operatorname{shw} S)^{\mathbf{0}} = (\operatorname{conv} S)^{\mathbf{0}} = (\operatorname{cl} S)^{\mathbf{0}} = (\overline{\operatorname{conv}}\, S)^{\mathbf{0}}$.

2. $S^{\mathbf{0}} = \emptyset$ *if and only if* $0 \in \overline{\operatorname{conv}}\, S$.

3. $S \subseteq T$ *implies* $T^{\mathbf{0}} \subseteq S^{\mathbf{0}}$.

4. *If* $0 \notin \overline{\operatorname{conv}}\, S$, *then* $(S^{\mathbf{0}})^{\mathbf{0}} = \operatorname{shw} \overline{\operatorname{conv}}\, S$.

We will now show that V_S is a generator of $S^{\mathbf{0}}$. To this end, we need the following lemma, which says that the shadow of what can be seen of a set is the same as the shadow of the whole set. Likewise, what can be seen of a set is the same as what can be seen of the shadow of the set.

Lemma 3.8. *Let $S \subseteq \mathbb{R}^n$ be a closed set such that $0 \notin S$. Then,* $\operatorname{shw} V_S = \operatorname{shw} S$ *and* $V_{\operatorname{shw} S} = V_S$.

Proof. First we prove that $\operatorname{shw} V_S = \operatorname{shw} S$. Clearly, $\operatorname{shw} V_S \subseteq \operatorname{shw} S$.

Let $y \in \operatorname{shw} S$, then $y = \lambda x$ with $x \in S$, $\lambda \geq 1$. Let $I = \{\mu \geq 0 : \mu x \in S\}$ and $\mu_0 = \min I$. The minimum exists since I is closed and not empty as S is closed and $1 \in I$, respectively. From $1 \in I$, we deduce $\mu_0 \leq 1$, and from $0 \notin S$, $\mu_0 > 0$. Hence, $\mu_0 x \in V_S$ and $y = \frac{\lambda}{\mu_0}(\mu_0 x) \in \operatorname{shw} V_S$, since $\frac{\lambda}{\mu_0} \geq 1$.

Now we prove that $V_{\operatorname{shw} S} = V_S$. Clearly, $S \subseteq \operatorname{shw} S$ implies that $V_S \subseteq V_{\operatorname{shw} S}$.

Let $x_0 \in V_{\operatorname{shw} S}$. Then, $[0,1)x_0 \notin \operatorname{shw} S$. As $S \subseteq \operatorname{shw} S$ it follows that $[0,1)x_0 \notin S$. Hence, if we manage to show that $x_0 \in S$, then $x_0 \in V_S$ which is what we want to prove.

As $x_0 \in V_{\operatorname{shw} S}$, it must be that $x_0 \in \operatorname{shw} S$. This means that there exists $\lambda \geq 1$ and $x \in S$ such that $x_0 = \lambda x$. Note that $\frac{1}{\lambda} x_0 = x \in S \subseteq \operatorname{shw} S$. In other words, $\frac{1}{\lambda} x_0 \in \operatorname{shw} S$ but, as we mentioned above, $[0,1)x_0 \notin \operatorname{shw} S$. This implies that $\frac{1}{\lambda} \geq 1$. Therefore, $\lambda = 1$, which means that $x_0 = x \in S$. □

Proposition 3.9. *Let $S \subseteq \mathbb{R}^n$ be a closed set. Then,*

$$(S \cap V_S)^{\mathbf{0}} = V_S^{\mathbf{0}} = S^{\mathbf{0}}.$$

Proof. The first equality just comes from the fact that $V_S \subseteq S$.

If $0 \in S$, then the equality holds as all the sets are empty. Otherwise, the equality follows from $V_S^{\mathbf{0}} = (\operatorname{shw} V_S)^{\mathbf{0}} = (\operatorname{shw} S)^{\mathbf{0}} = S^{\mathbf{0}}$, where the first and last equalities are by Lemma 3.7 and the middle one, by Lemma 3.8. □

3.3 The Smallest Generators

3.3.1 Motivation

In the previous section we showed that there is a set $U \subseteq S$ such that $(U \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$, namely, $U = V_S$. This set can be used to improve separation routines as was shown already in Example 3.2. We will come back to applications of the visible points to separation in the next section.

The topic of this section is motivated by the following example, where the set V_S is much larger than the smallest generator.

Example 3.10. Consider the constrained set $S = \{(x_1, x_2) \in \mathbb{R}^2_+ : (x_1 - x_2)^2 \geq 1\}$ depicted in Figure 3.2. The visible points are the lines $x_2 = x_1 + 1$

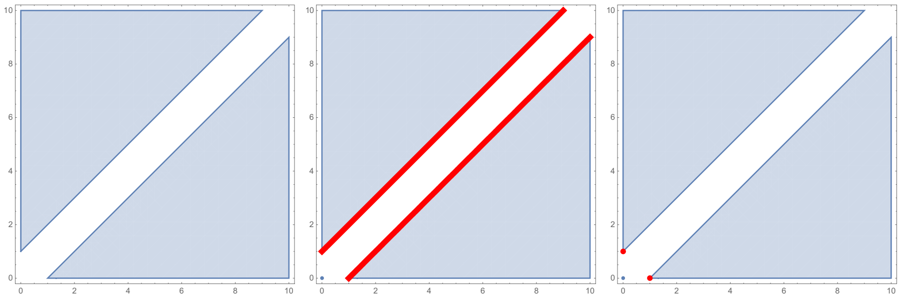


Figure 3.2: The region S. In the middle picture V_S are the points described by the thick red line. In the right picture the red points form the smallest set V such that $V^{\mathbf{0}} = S^{\mathbf{0}}$.

and $x_2 = x_1 - 1$ intersected with the first orthant. However, it is not hard to see that $V = \{(0,1),(1,0)\}$ is the smallest *closed* generator of $S^{\mathbf{0}}$. □

This example motivates the following question.

Question 3.11. What is, if any, the smallest closed set U such that $U^{\mathbf{0}} = S^{\mathbf{0}}$?

The reason we restrict to generators that are closed sets is to avoid representation issues. For example, if S is the ball of radius 1 centered at $(2,0)$, then Theorem 3.29 implies that the left arc joining $(2,1)$ and $(2,-1)$ generates $S^{\mathbf{0}}$. However, the rational points on this arc also generate $S^{\mathbf{0}}$ and the smallest set generating $S^{\mathbf{0}}$ does not exist. In order to avoid such issues, we concentrate on closed generators.

As can be seen from simple examples, such as $S = \mathbb{R}_+ \times \{1\}$ for which every $a \geq 0$ defines the generator $(\{0\} \cup [a,\infty)) \times \{1\}$, the smallest closed generator must not exist. However, a smallest closed convex generator might exist and so we ask the following question.

Question 3.12. What is, if any, the smallest closed convex generator of $S^{\mathbf{0}}$?

We are mainly interested in applying our results to the separation problem, as already explained in the introduction. In that case, the set S usually looks like $S = C \cap F$, where C is a convex set and F is the sublevel set of some non-convex function, see the next section. In this context, replacing C by a smaller convex set might be beneficial for the separation problem (see Example 3.32). Thus, it is also natural to consider the following question.

Question 3.13. What is, if any, the smallest closed convex set U such that $S \cap U$ generates $S^{\mathbf{0}}$?

The last two questions are not the same. Informally, S is only used to define $S^{\mathbf{0}}$ in Question 3.12, and so any other set T such that $T^{\mathbf{0}} = S^{\mathbf{0}}$ can be used to formulate the question. For instance, we can assume without loss of generality that S is closed and convex, since Lemma 3.7 implies that $(\overline{\operatorname{conv}} S)^{\mathbf{0}} = S^{\mathbf{0}}$. In contrast, in Question 3.13 we are asking for the smallest generator contained in S.

As we will see, the answer to Question 3.12 is that $\overline{\operatorname{conv}}\, V_{\overline{\operatorname{conv}} S}$ is the smallest closed convex generator of $S^{\mathbf{0}}$. However, the next two examples show that Question 3.13 is a bit more delicate.

The first example shows that, in general, there is no unique smallest closed convex set U such that $(S \cap U)^{\mathbf{0}} = S^{\mathbf{0}}$.

Example 3.14. Let $S = \{(1,0), (0,1), (-1,0), (0,-1)\}$. Since $0 \in \operatorname{conv} S$, $S^{\mathbf{0}} = \emptyset$.

Clearly $V = \{0\} = V_{\operatorname{conv} S}$ is the smallest closed convex set such that $V^{\mathbf{0}} = \emptyset$. However, $S \cap V = \emptyset$, which implies that $(S \cap V)^{\mathbf{0}} = \mathbb{R}^2 \neq S^{\mathbf{0}}$. Furthermore, $U_1 = \{(\lambda, 0) : \lambda \in [-1,1]\}$ and $U_2 = \{(0,\lambda) : \lambda \in [-1,1]\}$ are both closed convex and $(U_i \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$. Since $U_1 \not\subseteq U_2$ and $U_2 \not\subseteq U_1$ we conclude that there is no smallest closed convex set U such that $(U \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$. □

However, we cannot even expect to find a minimal closed convex set U such that $(S \cap U)^{\mathbf{0}} = S^{\mathbf{0}}$.

Example 3.15. Let $S = \{(0,1)\} \cup \{(\lambda, 2) : \lambda \geq 0\}$. We have $S^{\mathbf{0}} = \{\alpha : \alpha_1 \geq 0, \alpha_2 \geq 1\}$.

Indeed, $(0,1) \in S$ implies that $\alpha_2 \geq 1$. If $\alpha_1 < 0$ for some $\alpha \in S^{\mathbf{0}}$, then there is a large enough λ such that $\lambda\alpha_1 + 2\alpha_2 < 1$ and $(\lambda, 2) \in S$. On the other hand, if $\alpha_1 \geq 0$ and $\alpha_2 \geq 1$, then $\alpha_1 x_1 + \alpha_2 x_2 \geq 1$ for every $(x,y) \in S$.

Let $T_M = \{(0,1)\} \cup \{(\lambda, 2) : \lambda \geq M\}$ and $U_M = \overline{\operatorname{conv}}\, T_M$. The same argument as above shows that $(U_M \cap S)^{\mathbf{0}} = T_M^{\mathbf{0}} = S^{\mathbf{0}}$. Notice that any U with $(U \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$ must contain a sequence $\lambda_n \to \infty$ such that $(\lambda_n, 2) \in S$. Thus, any minimal U, if it exists, must be of the form U_M for some $M \geq 0$.

It is clear that $U_{M_1} \subseteq U_{M_2}$ if and only if $M_1 > M_2$ and $\bigcap_{M>0} U_M = \{(\lambda, 1) : \lambda \geq 0\}$. However, $S \cap \{(\lambda,1) : \lambda \geq 0\} = \{(0,1)\}$ and $\{(0,1)\}^{\mathbf{0}} \neq S^{\mathbf{0}}$. Therefore, there is no minimal U.

On the other hand, $V = \{(\lambda, 1) : \lambda \geq 0\} = V_{\overline{\operatorname{conv}} S}$ is the smallest closed convex set such that $V^{\mathbf{0}} = S^{\mathbf{0}}$. □

However, these are the only "pathological cases". Indeed, as we will see, if $\operatorname{conv} S$ is closed (e.g. when S is compact) and $0 \notin \operatorname{conv} S$, (i.e., $S^{\mathbf{0}} \neq \emptyset$), then $\overline{\operatorname{conv}}\, V_{\operatorname{conv} S}$ is the smallest closed convex set such that $\overline{\operatorname{conv}}\, V_{\operatorname{conv} S} \cap S$ generates $S^{\mathbf{0}}$.

Remark 3.16. The closure operations are needed because, in general, V_S and $\operatorname{conv} V_S$ are not closed, even when S is convex and compact. Indeed, it is shown in Deutsch et al. (2013, Example 15.5) that for

$$S := (1,0,0) + \operatorname{cone}\{(1,\alpha,\beta) \,:\, \alpha^2 + (\beta-1)^2 \leq 1\},$$

V_S is open. The authors show that the points $(2, \sin(t), 1+\cos(t))$ are visible for $t \in (0,\pi)$, but the limit when t approaches π, (2,0,0), is not. The remark follows from a modification of this example so that S is compact, e.g., by intersecting it with $[0,3] \times \mathbb{R}^2$. □

3.3.2 Preliminaries

Here we collect a few lemmata that we are going to need in order to answer Questions 3.11, 3.12 and 3.13.

Lemma 3.17. *Deutsch et al. (2013, Proposition 15.19) Let S be a closed convex set such that $0 \notin S$. If $x \in V_S$ is a strict convex combination of $x_1, \ldots x_m \in S$, then $x_1, \ldots x_m \in V_S$.*

This result immediately implies the following two lemmata.

Lemma 3.18. *Let $S \subseteq \mathbb{R}^n$ be a closed convex set such that $0 \notin S$. Then,* $\operatorname{ext} V_S = V_S \cap \operatorname{ext} S$.

Proof. We start by proving $\operatorname{ext} V_S \subseteq V_S \cap \operatorname{ext} S$. Let $x \in \operatorname{ext} V_S$. Clearly, $x \in V_S$. If $x \notin \operatorname{ext} S$, then there are $x_1, \ldots, x_m \in S$ such that x is a strict convex combination of $x_1, \ldots, x_m$. Lemma 3.17 implies that $x_i \in V_S$ for every $i = 1, \ldots, m$. Thus, x is not an extreme point of V_S. This contradiction proves that $x \in \operatorname{ext} S$.

If $x \in V_S \cap \operatorname{ext} S$ but $x \notin \operatorname{ext} V_S$, then x is a strict convex combination of some elements of V_S. Since $V_S \subseteq S$, x is a strict convex combination of some element of S. This is a contradiction with $x \in \operatorname{ext} S$. □

Lemma 3.19. *Let $S \subseteq \mathbb{R}^n$ be closed set such that* $\operatorname{conv} S$ *is closed and* $0 \notin \operatorname{conv} S$. *Then,*

$$\operatorname{conv} V_{\operatorname{conv} S} = \operatorname{conv}(S \cap V_{\operatorname{conv} S}).$$

Proof. From $S \cap V_{\operatorname{conv} S} \subseteq V_{\operatorname{conv} S}$, it follows that $\operatorname{conv}(S \cap V_{\operatorname{conv} S}) \subseteq \operatorname{conv} V_{\operatorname{conv} S}$.

To prove the other inclusion it is enough to show that $V_{\operatorname{conv} S} \subseteq \operatorname{conv}(S \cap V_{\operatorname{conv} S})$. Let $x \in V_{\operatorname{conv} S}$. Then, $x \in \operatorname{conv} S$ and so x is a strict convex combination of some points of $x_1, \ldots, x_m \in S$. Then, by Lemma 3.17, $x_1, \ldots, x_m \in S \cap V_{\operatorname{conv} S}$. Thus, $x \in \operatorname{conv}(S \cap V_{\operatorname{conv} S})$. □

We remark that the previous lemma does not follow from Lemma 3.18 by just taking the convex hull operation to the equality, since $\operatorname{conv} S$ may not have extreme points.

The following is a slight extension of (Rockafellar, 1970, Corollary 18.3.1).

Lemma 3.20. *Let $S \subseteq \mathbb{R}^n$ be a closed set. Then,* $\operatorname{ext} \overline{\operatorname{conv}}\, S \subseteq S$.

Proof. Recall that x_0 is an exposed point of a closed convex set C if and only if there exists an α such that $\{x_0\} = \arg\max_{x \in C} \alpha^{\mathsf{T}} x_0$.

We will show that the exposed points of $\overline{\operatorname{conv}}\, S$ is a subset of S. Then, by Straszewicz's Theorem (Rockafellar, 1970, Theorem 18.6) and the closedness of S, it follows that $\operatorname{ext} \overline{\operatorname{conv}}\, S \subseteq S$. Note that when the set of exposed points is empty, the result follows trivially. Thus, we assume that the set of exposed points is non-empty.

Let x_0 be an exposed point of $\overline{\operatorname{conv}}\, S$ and let α be a direction that exposes it. Then, $\sup_{x \in S} \alpha^{\mathsf{T}} x = \alpha^{\mathsf{T}} x_0$. Since S is closed, there exists $x_1 \in S$ such that $\alpha^{\mathsf{T}} x_1 = \alpha^{\mathsf{T}} x_0$. However, since $x_1 \in S \subseteq \overline{\operatorname{conv}}\, S$ and α exposes x_0, we must have $x_1 = x_0$. Thus, $x_0 \in S$. □

3.3.3 Results

Let us start by answering Question 3.12.

Theorem 3.21. *Let $S \subseteq \mathbb{R}^n$ be closed. Then,*

$$(\overline{\operatorname{conv}}\, V_{\overline{\operatorname{conv}} S})^{\mathbf{o}} = S^{\mathbf{o}}.$$

Furthermore, if $C \subseteq \mathbb{R}^n$ is a closed convex generator of $S^{\mathbf{o}}$, then

$$\overline{\operatorname{conv}}\, V_{\overline{\operatorname{conv}} S} \subseteq C.$$

Proof. Note that if $S^{\mathbf{o}} = \emptyset$, then $0 \in \overline{\operatorname{conv}}\, S$ and $V_{\overline{\operatorname{conv}} S} = \{0\}$, from which the theorem clearly follows. Thus, we assume $S^{\mathbf{o}} \neq \emptyset$.

Lemma 3.7 implies that $(\overline{\operatorname{conv}}\, V_{\overline{\operatorname{conv}}\, S})^{\mathbf{0}} = (V_{\overline{\operatorname{conv}}\, S})^{\mathbf{0}}$ and $S^{\mathbf{0}} = (\overline{\operatorname{conv}}\, S)^{\mathbf{0}}$. Proposition 3.9 implies $(\overline{\operatorname{conv}}\, S)^{\mathbf{0}} = (V_{\overline{\operatorname{conv}}\, S})^{\mathbf{0}}$.

To show the second statement of the theorem, let C be closed and convex such that $C^{\mathbf{0}} = S^{\mathbf{0}}$. Since C is closed and convex, it is enough to prove that $V_{\overline{\operatorname{conv}}\, S} \subseteq C$. Suppose, by contradiction, that this is not the case, i.e., there is an $\bar{x} \in V_{\overline{\operatorname{conv}}\, S}$ such that $\bar{x} \notin C$. There are two cases, either $[0,1]\bar{x} \cap C = \emptyset$ or $[0,1]\bar{x} \cap C \neq \emptyset$. We will deduce a contradiction from each of them.

First, suppose $[0,1]\bar{x} \cap C = \emptyset$. Both sets are closed and $[0,1]\bar{x}$ is bounded, thus, they can be separated. Indeed, as $0 \in [0,1]\bar{x}$, Rockafellar (1970, Corollary 11.4.1) ensures the existence of α such that $\alpha x \geq 1$ for every $x \in C$ and $\alpha\bar{x} < 1$. This means that $\alpha \in C^{\mathbf{0}}$. However, $\alpha \notin (\overline{\operatorname{conv}}\, S)^{\mathbf{0}} = S^{\mathbf{0}}$, since $\bar{x} \in \overline{\operatorname{conv}}\, S$. This contradicts $S^{\mathbf{0}} = C^{\mathbf{0}}$.

Now, suppose $[0,1]\bar{x} \cap C \neq \emptyset$. Since $0 \notin C$ (as $C^{\mathbf{0}} = S^{\mathbf{0}} \neq \emptyset$) and $\bar{x} \notin C$, there must be $\mu \in (0,1)$ such that $\mu\bar{x} \in C$. However, $\bar{x} \in V_{\overline{\operatorname{conv}}\, S}$ implies that $\mu\bar{x} \notin \overline{\operatorname{conv}}\, S$. Thus, the same argument as above ensures that $\mu\bar{x}$ can be separated from $\overline{\operatorname{conv}}\, S$. Therefore, there is an α such that $\alpha^{\mathsf{T}} x \geq 1$ for every $x \in \overline{\operatorname{conv}}\, S$ while $\alpha^{\mathsf{T}}\mu\bar{x} < 1$. Hence, $\alpha \in S^{\mathbf{0}}$ and the contradiction follows from the fact that $\mu\bar{x} \in C$ implies $\alpha \notin C^{\mathbf{0}}$.

Therefore, we conclude that $\overline{\operatorname{conv}}\, V_{\overline{\operatorname{conv}}\, S} \subseteq C$. □

Now we show that if $\operatorname{conv} S$ is closed and $0 \notin \operatorname{conv} S$, then $\overline{\operatorname{conv}}\, V_{\operatorname{conv} S}$ is the answer to Question 3.13, i.e., is *the* smallest closed convex U such that $(U \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$.

Theorem 3.22. *Let $S \subseteq \mathbb{R}^n$ be a closed set such that* $\operatorname{conv} S$ *is closed and* $0 \notin \operatorname{conv} S$*, i.e., $S^{\mathbf{0}} \neq \emptyset$. Then,*

$$(\overline{\operatorname{conv}}(V_{\operatorname{conv} S}) \cap S)^{\mathbf{0}} = S^{\mathbf{0}}.$$

Furthermore, if C is closed and convex such that $(C \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$, then

$$\overline{\operatorname{conv}}\, V_{\operatorname{conv} S} \subseteq C.$$

Proof. We first show that $(\overline{\operatorname{conv}}(V_{\operatorname{conv} S}) \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$.
Clearly,

$$V_{\operatorname{conv} S} \cap S \subseteq \overline{\operatorname{conv}}(V_{\operatorname{conv} S}) \cap S \subseteq S.$$

Lemma 3.7 implies that

$$S^{\mathbf{0}} \subseteq (\overline{\operatorname{conv}}(V_{\operatorname{conv} S}) \cap S)^{\mathbf{0}} \subseteq (V_{\operatorname{conv} S} \cap S)^{\mathbf{0}}.$$

Thus, it is enough to show that $(V_{\operatorname{conv} S} \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$. This follows from

$$
\begin{aligned}
(S \cap V_{\operatorname{conv} S})^{\mathbf{0}} &= (\operatorname{conv}(S \cap V_{\operatorname{conv} S}))^{\mathbf{0}} && \text{Lemma 3.7} \\
&= (\operatorname{conv} V_{\operatorname{conv} S})^{\mathbf{0}} && \text{Lemma 3.19} \\
&= (V_{\operatorname{conv} S})^{\mathbf{0}} && \text{Lemma 3.7} \\
&= (\operatorname{conv} S)^{\mathbf{0}} && \text{Proposition 3.9} \\
&= S^{\mathbf{0}}. && \text{Lemma 3.7}
\end{aligned}
$$

To show the second statement of the theorem, let C be a closed convex set such that $(C \cap S)^{\mathbf{0}} = S^{\mathbf{0}}$. Lemma 3.7 implies that $(C \cap S)^{\mathbf{0}} = (\overline{\operatorname{conv}}(C \cap S))^{\mathbf{0}}$. Theorem 3.21 implies that $\overline{\operatorname{conv}}\, V_{\operatorname{conv} S} \subseteq \overline{\operatorname{conv}}(C \cap S)$. Clearly, $V_{\operatorname{conv} S} \subseteq \overline{\operatorname{conv}}\, V_{\operatorname{conv} S}$ and $\overline{\operatorname{conv}}(C \cap S) \subseteq C \cap \operatorname{conv} S$. Therefore, $V_{\operatorname{conv} S} \subseteq C \cap \operatorname{conv} S$ which implies $V_{\operatorname{conv} S} \subseteq C$ as we wanted. $\square$

Finally, we answer Question 3.11 in the case where S is compact.

Theorem 3.23. *Let S be any closed set such that $0 \notin \overline{\operatorname{conv}}\, S$. If D is any closed generator of $S^{\mathbf{0}}$, then*

$$\overline{\operatorname{ext}}\, V_{\overline{\operatorname{conv}} S} \subseteq D.$$

If, in addition, S is compact, then $\overline{\operatorname{ext}}\, V_{\operatorname{conv} S}$ is the smallest closed generator of $S^{\mathbf{0}}$.

Proof. First, by Lemma 3.7 and $D^{\mathbf{0}} = S^{\mathbf{0}}$, we have $\operatorname{shw} \overline{\operatorname{conv}}\, D = \operatorname{shw} \overline{\operatorname{conv}}\, S$. Then, Lemma 3.8 implies that $V_{\overline{\operatorname{conv}} D} = V_{\overline{\operatorname{conv}} S}$. Hence, $\operatorname{ext} V_{\overline{\operatorname{conv}} D} = \operatorname{ext} V_{\overline{\operatorname{conv}} S}$. Therefore, $\operatorname{ext} V_{\overline{\operatorname{conv}} S} = \operatorname{ext} V_{\overline{\operatorname{conv}} D} \subseteq \operatorname{ext} \overline{\operatorname{conv}}\, D \subseteq D$, where the first and second containments are due to Lemma 3.18 and Lemma 3.20, respectively.

To prove the second statement, by Lemma 3.7, it is enough to show that $(\operatorname{ext} V_{\operatorname{conv} S})^{\mathbf{0}} = S^{\mathbf{0}}$. First, as $\operatorname{ext} V_{\operatorname{conv} S} \subseteq \operatorname{conv} S$, we have $S^{\mathbf{0}} \subseteq (\operatorname{ext} V_{\operatorname{conv} S})^{\mathbf{0}}$ by Lemma 3.7.

To prove the other containment take any $\alpha \in (\operatorname{ext} V_{\operatorname{conv} S})^{\mathbf{0}}$. Let $x \in \operatorname{conv} S$ be arbitrary. We will prove that $\alpha^{\mathsf{T}} x \geq 1$. This will imply that $\alpha \in (\operatorname{conv} S)^{\mathbf{0}} = S^{\mathbf{0}}$ and, therefore, that $(\operatorname{ext} V_{\operatorname{conv} S})^{\mathbf{0}} \subseteq S^{\mathbf{0}}$.
Let $\lambda \in (0, 1]$ be such that $\lambda x \in V_{\operatorname{conv} S}$. If $\lambda x \in \operatorname{ext} V_{\operatorname{conv} S}$, then $\alpha^{\mathsf{T}} \lambda x \geq 1$, which implies that $\alpha^{\mathsf{T}} x \geq \frac{1}{\lambda} \geq 1$.
Now, assume $\lambda x \notin \operatorname{ext} V_{\operatorname{conv} S}$. Since S is compact, $\operatorname{conv} S$ is closed and we can use Lemma 3.18 to obtain that $\operatorname{ext} V_{\operatorname{conv} S} = V_{\operatorname{conv} S} \cap \operatorname{ext} \operatorname{conv} S$. Thus, $\lambda x \notin \operatorname{ext} \operatorname{conv} S$. Also by the compactness of S, Rockafellar (1970, Theorem

18.5.1) implies that λx is a strict convex combination of some $x_1, \ldots, x_m \in \operatorname{ext}\operatorname{conv} S$.
Lemma 3.17 implies that $x_1, \ldots, x_m \in V_{\operatorname{conv} S}$ and so Lemma 3.18 implies that $x_1, \ldots, x_m \in \operatorname{ext} V_{\operatorname{conv} S}$. Since $\alpha \in (\operatorname{ext} V_{\operatorname{conv} S})^{\mathbf{0}}$, it follows $\alpha^{\mathsf{T}} x_i \geq 1$ for every $i = 1, \ldots, m$. Hence, $\alpha^{\mathsf{T}} \lambda x \geq 1$ and, as before, $\alpha^{\mathsf{T}} x \geq \frac{1}{\lambda} \geq 1$. □

We remark that the closure operation is needed since the extreme points of a set, in general, do not form a closed set, see Rockafellar (1970, p. 167).

3.4 Applications to MINLP

Here we apply the results from Section 3.2 to MINLP.

In this section, unless specified otherwise, $\bar{x} \in \mathbb{R}^n$, C is a closed convex set that contains $\bar{x}$, and $S := \{x \in C \,:\, g(x) \leq 0\}$, where $g : C \to \mathbb{R}$ is continuous and $g(\bar{x}) > 0$. The idea is that C represents a convex relaxation of our MINLP and $\bar{x} \in C$ is the current relaxation solution that is infeasible for a constraint $g(x) \leq 0$.

The basic scheme for applying our results is the following translation of Observation 3.1.

Proposition 3.24. *Let $D \subseteq C$ be such that $(D \cap S)^{\bar{x}} = S^{\bar{x}}$, and $T = \{x \in D \,:\, g(x) \leq 0\}$. If $\alpha^{\mathsf{T}}(x - \bar{x}) \geq 1$ is a valid inequality for T, then it is valid for S.*

Proof. Directly from $\alpha \in T^{\bar{x}} = (D \cap S)^{\bar{x}} = S^{\bar{x}}$. □

Of course, the applicability of the previous proposition relies on our ability to obtain an easy-to-compute set D that satisfies the hypothesis. As shown in Section 3.3, $D = \overline{\operatorname{ext}}\, \overline{\operatorname{conv}}\, V_{\operatorname{conv} S}(\bar{x})$ is the smallest we can hope for, but it is useless from a practical point of view. Instead, the set of visible points of S (or a set enclosing them) is, computationally, a better candidate as we will see in Section 3.4.1.

Corollary 3.25. *Let $D \subseteq C$ be such that $V_S(\bar{x}) \subseteq D$, and $T = \{x \in D \,:\, g(x) \leq 0\}$. If $\alpha^{\mathsf{T}}(x - \bar{x}) \geq 1$ is a valid inequality for T, then it is valid for S.*

Proof. Clearly, $V_S(\bar{x}) \subseteq T = D \cap S \subseteq S$. The inclusion-reversing property of the reverse polar implies that $S^{\bar{x}} \subseteq (D \cap S)^{\bar{x}} \subseteq V_S(\bar{x})^{\bar{x}} = S^{\bar{x}}$, where the last equality follows from Proposition 3.9. The statement follows from Proposition 3.24. □

In the context of separation via convex underestimators Corollary 3.25 reads as follows.

Corollary 3.26. *Let $D \subseteq C$ be a closed convex set such that $V_S(\bar{x}) \subseteq D$, and let $T = \{x \in D \,:\, g(x) \le 0\}$. If $g_D^{vex}(\bar{x}) > 0$ and $\partial g_D^{vex}(\bar{x}) \neq \emptyset$, then a gradient cut of g_D^{vex} at $\bar{x}$ is valid for S.*

Proof. Let $T_r = \{x \in D \,:\, g_D^{vex}(x) \le 0\}$ and $v \in \partial g_D^{vex}(\bar{x})$. The cut $g_D^{vex}(\bar{x}) + v^\mathsf{T}(x - \bar{x}) \le 0$ is valid for T_r, and separates $\bar{x}$ from T_r. Since T_r is a relaxation, i.e. $T \subseteq T_r$, it follows that the cut is also valid for T, and Corollary 3.25 implies its validity for S. □

The previous result tells us that if we find a box, tighter than the bounds, that contains the visible points, then we might be able to construct tighter underestimators. However, to compute a box containing $V_S(\bar{x})$ we need to know how $V_S(\bar{x})$ looks like. That is the topic of the next section.

3.4.1 Characterizing the Visible Points

From the definition of visible points we have:

Theorem 3.27. *Let $g : \mathbb{R}^n \to \mathbb{R}$ be a continuous function, $C \subseteq \mathbb{R}^n$ a closed convex set, and $S = \{x \in C : g(x) \le 0\}$. If $\bar{x} \in C$ and $g(\bar{x}) > 0$, then*

$$V_S(\bar{x}) = \{x \in C \,:\, g(x) = 0,\ g(x + \lambda(\bar{x} - x)) > 0 \text{ for every } \lambda \in (0,1]\}\,. \tag{3.1}$$

Furthermore, if g is differentiable, then every $x \in V_S(\bar{x})$ satisfies

$$\nabla g(x)^\mathsf{T}(\bar{x} - x) \ge 0.$$

Proof. Given that $\bar{x} \notin S$, by definition we have $x \in V_S(\bar{x})$ if and only if $x \in S$ and for every $\lambda \in (0,1]$, $x + \lambda(\bar{x} - x) \notin C$ or $g(x + \lambda(\bar{x} - x)) > 0$. However, the convexity of C and $\bar{x} \in C$ imply that for $x \in S$, $x + \lambda(\bar{x} - x) \in C$. Hence,

$$V_S(\bar{x}) = \{x \in C \,:\, g(x) \le 0,\ g(x + \lambda(\bar{x} - x)) > 0 \text{ for every } \lambda \in (0,1]\}.$$

Since g is continuous, it follows that for $x \in V_S(\bar{x})$,

$$0 \ge g(x) = \lim_{\lambda \to 0^+} g(x + \lambda(x - \bar{x})) \ge 0.$$

Thus, $g(x) = 0$ which proves (3.1).

Now, assume that g is differentiable and let $x \in V_S(\bar{x})$. Then,

$$0 \le \lim_{\lambda \to 0^+} \frac{g(x + \lambda(\bar{x} - x))}{\lambda} = \lim_{\lambda \to 0^+} \frac{g(x + \lambda(\bar{x} - x)) - g(x)}{\lambda} = \nabla g(x)^\mathsf{T}(\bar{x} - x).$$

This concludes the claim. □

Remark 3.28. Note that if we drop the hypothesis that $\bar{x}$ is in C, then there might be visible points for which g is strictly negative, and there does not seem to be a nice description of the visible points. In such a case, $V_S(\bar{x})$ would be a disjunctive set and we would even lose the valid (non-linear) inequality $\nabla g(x)^\mathsf{T}(\bar{x} - x) \geq 0$. Likewise, if C was not convex, or if we had more than one non-convex constraint, e.g., some variable has to be binary, then there does not seem to be a nice description of the visible points. This last point is rather unfortunate, it means that it might not be easy to generalize the technique to relaxations that involve more than one non-convex constraint. In particular, since a mixed-integer set usually consists of multiple non-convex constraints, the techniques presented here might not be applicable to MILPs. On the other hand, considering more constraints might allow us to see more of the feasible region. Therefore, in such cases one might have to try to use stronger generators such as $\overline{\operatorname{conv}}\, V_{\operatorname{conv} S}$, see also Venkatachalam and Ntaimo (2016). □

Quadratic constraints

For quadratic constraints, the visible points have a particularly simple description.

Theorem 3.29. *Let C be a closed, convex set that contains $\bar{x}$. Let $g(x) = x^\mathsf{T}Qx + b^\mathsf{T}x + c$ and $S = \{x \in C : g(x) \leq 0\}$. If $g(\bar{x}) > 0$, then*

$$V_S(\bar{x}) = \left\{x \in C \,:\, g(x) = 0,\ \nabla g(\bar{x})^\mathsf{T}x + b^\mathsf{T}\bar{x} + 2c \geq 0\right\}$$

Proof. ($\subseteq$) Let $x \in V_S(\bar{x})$. By Theorem 3.27, we have $g(x) = 0$ and $\nabla g(x)^\mathsf{T}(\bar{x} - x) \geq 0$. Equivalently,

$$\begin{aligned} x^\mathsf{T}Qx + b^\mathsf{T}x + c &= 0, \\ 2x^\mathsf{T}Q(\bar{x} - x) + b^\mathsf{T}(\bar{x} - x) &\geq 0. \end{aligned}$$

By multiplying the equation by 2, adding it to the inequality, and re-arranging terms we obtain the result.

($\supseteq$) Let x satisfy $g(x) = 0$ and $\nabla g(\bar{x})^\mathsf{T}x + b^\mathsf{T}\bar{x} + 2c \geq 0$. Then, subtracting $2g(x)$ from $\nabla g(\bar{x})^\mathsf{T}x + b^\mathsf{T}\bar{x} + 2c \geq 0$ yields $\nabla g(x)^\mathsf{T}(\bar{x} - x) \geq 0$. Let

$$q(\lambda) = g(x + \lambda(\bar{x} - x)), \text{ for } \lambda \in \mathbb{R}.$$

The derivative is given by $q'(\lambda) = \nabla g(x + \lambda(\bar{x} - x))^\mathsf{T}(\bar{x} - x)$, and $q'(0) = \nabla g(x)^\mathsf{T}(\bar{x} - x) \geq 0$. Since q is quadratic, $q(1) = g(\bar{x}) > 0$, $q(0) = g(x) = 0$, and

$q'(0) \geq 0$, we have that q has no roots in $(0,1]$. Thus, $g(x+\lambda(\bar{x}-x)) = q(\lambda) > 0$ for every $\lambda \in (0,1]$ and, from Theorem 3.27, we conclude that $x \in V_S(\bar{x})$ as we wanted. $\square$

Remark 3.30. Theorem 3.29 implies in particular that the visible points of a closed convex set intersected with a quadratic constraint, from a point in the convex set, is always closed. This does not contradict Deutsch et al. (2013, Example 15.5) mentioned in Remark 3.16. Indeed, if one represents the cone as a quadratic constraint $g(x) \leq 0$, then the origin must be feasible for the quadratic constraint. This follows from the fact that the ray $[1,\infty)(1,0,0)$ is in the boundary of the cone, which implies that $g(\lambda,0,0) = 0$ for $\lambda \geq 0$. But $g(\lambda,0,0)$ is a univariate quadratic function and as such can have at most two roots if it is nonzero. Hence, $g(\lambda,0,0) = 0$ and, in particular, $g(0,0,0) = 0$. $\square$

Remark 3.31. The hyperplane $\nabla g(\bar{x})^\mathsf{T} x + b^\mathsf{T}\bar{x} + 2c = 0$ is known as the *polar hyperplane* (Fasano and Pesenti, 2017) of the point $\bar{x}$ with respect to the quadratic g in projective geometry. In fact, homogenizing the quadratic g yields the quadric

$$g_h(x,x_0) = x^\mathsf{T} Q x + b^\mathsf{T} x x_0 + c x_0^2 = \begin{pmatrix} x \\ x_0 \end{pmatrix}^\mathsf{T} \begin{pmatrix} Q & \frac{b}{2} \\ \frac{b^\mathsf{T}}{2} & c \end{pmatrix} \begin{pmatrix} x \\ x_0 \end{pmatrix}.$$

The polar hyperplane of $\begin{pmatrix} \bar{x} \\ 1 \end{pmatrix}$ with respect to $g_h(x,x_0) = 0$ is then given by

$$\begin{aligned} &\nabla g_h(x,x_0)^\mathsf{T}(\bar{x},1) = 0 \\ \iff &2\bar{x}^\mathsf{T} Q x + b^\mathsf{T}\bar{x} x_0 + b^\mathsf{T} x + 2c x_0 = 0. \end{aligned}$$

Intersecting with $x_0 = 1$ yields $\nabla g(\bar{x})^\mathsf{T} x + b^\mathsf{T}\bar{x} + 2c = 0$. $\square$

Example 3.32. Consider the function

$$g(x_1,x_2,x_3) = -x_1x_2 + x_1x_3 + x_2x_3 - x_1 - x_2 - x_3 + 1,$$

the boxed domain $B = [-\frac{1}{10}, 2] \times [0,2]^2$, the constrained set

$$S = \{x \in B \,:\, g(x) \leq 0\},$$

and the infeasible point $\bar{x} = (0,0,0)$. By Theorem 3.29, the visible points from $\bar{x}$ are given by

$$V_S(\bar{x}) = \{(x_1,x_2,x_3) \in B \,:\, g(x) = 0,\; x_1 + x_2 + x_3 \geq 0\},$$

as shown in Figure 3.3.

The tightest box bounding $V_S(\bar{x})$ is

$$R = [-\frac{1}{10}, 1] \times [0, \frac{1}{20}(23 + 3\sqrt{5})] \times [0, \frac{1}{20}(19 + 3\sqrt{5})].$$

The linear underestimators of g obtained by using McCormick inequalities (McCormick, 1976) for each term over B and R are

$$1 \leq x_1 + 3x_2 + \frac{11}{10}x_3 \text{ and } 1 \leq x_1 + 2x_2 + \frac{11}{10}x_3,$$

respectively. Since $0 \leq x_2$, it follows that the underestimator over R dominates the underestimator over B. We remark that the improvement in this particular cut is only due to the improvement on the upper bound of x_1.

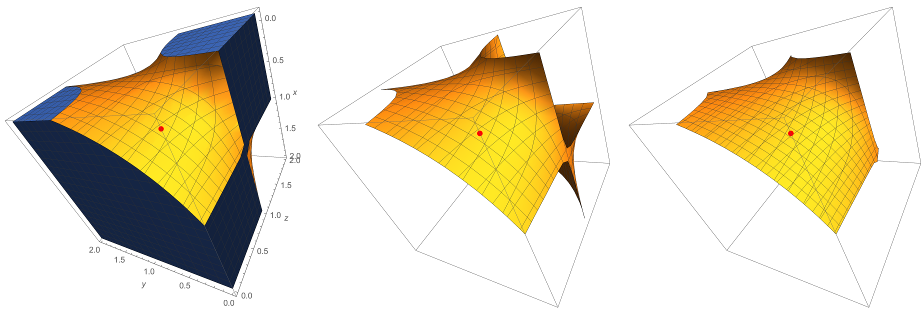


Figure 3.3: The left plot shows the feasible region S and $\bar{x}$. The set $\{x \in B : g(x) = 0\}$ appears in the middle plot. Finally, the visible points, $V_S(\bar{x})$, are plotted on the right.

$\square$

Polynomial constraints

For a general polynomial g, the condition

$$g(x + \lambda(\bar{x} - x)) > 0 \text{ for every } \lambda \in (0, 1] \tag{3.2}$$

of (3.1) asks for the univariate polynomial $p_x(\lambda) = g(x + \lambda(\bar{x} - x))$ to be positive on $(0, 1]$. We can then use the theory of non-negative polynomials to translate a relaxation of the infinitely many constraints (3.2) to a finite number of constraints. From the following classic characterization of univariate non-negative polynomials on intervals, see for instance Powers and Reznick (2000), we can derive an extended formulation for the relaxation of (3.1),

$$R_S(\bar{x}) := \{x \in C : g(x) = 0, \; g(x + \lambda(\bar{x} - x)) \geq 0 \text{ for every } \lambda \in [0, 1]\}.$$

Theorem 3.33. *Let $p \in \mathbb{R}[\lambda]$ be a polynomial. Then p is non-negative on $[0,1]$ if and only if*

1. *the degree of p is $2d$ and there exist $s_1, s_2 \in \mathbb{R}[\lambda]$ of degree d and $d-1$, respectively, such that*
$$p(\lambda) = s_1(\lambda)^2 + \lambda(1-\lambda)s_2(\lambda)^2.$$
2. *the degree of p is $2d+1$ and there exist $s_1, s_2 \in \mathbb{R}[\lambda]$ of degree d, such that*
$$p(\lambda) = \lambda s_1(\lambda)^2 + (1-\lambda)s_2(\lambda)^2.$$

Theorem 3.34. *Let C be a closed convex set that contains $\bar{x}$. Let $g(x)$ be a polynomial such that $g(\bar{x}) > 0$ and $S = \{x \in C : g(x) \leq 0\}$. Let $p_x(\lambda) = g(x + \lambda(\bar{x} - x))$.*

1. *If the degree of g is $2d$, then*
$$R_S(\bar{x}) = \operatorname{proj}_x E,$$
where E is
$$\begin{aligned} \{(x, A, B) \in C \times \mathcal{S}^d_+ \times \mathcal{S}^d_+ :& \\ g(x) &= 0, \\ p'_x(0) &= B_{00}, \\ \frac{p_x^{(k+2)}(0)}{(k+2)!} &= \sum_{\substack{i+j=k \\ 0 \leq i,j \leq d-1}} A_{ij} - B_{ij} + \sum_{\substack{i+j=k+1 \\ 0 \leq i,j \leq d-1}} B_{ij}, \text{ for } 0 \leq k \leq 2d-2\}. \end{aligned}$$
2. *If the degree of g is $2d+1$, then*
$$R_S(\bar{x}) = \operatorname{proj}_x E,$$
where E is
$$\begin{aligned} \{(x, A, B) \in C \times \mathcal{S}^{d+1}_+ \times \mathcal{S}^d_+ :& \\ g(x) &= 0, \\ p'_x(0) &= A_{00}, \\ \frac{p''_x(0)}{2} &= 2A_{01} + B_{00}, \\ \frac{p_x^{(k+3)}(0)}{(k+3)!} &= \sum_{\substack{i+j=k+2 \\ 0 \leq i,j \leq d}} A_{ij} + \sum_{\substack{i+j=k+1 \\ 0 \leq i,j \leq d-1}} B_{ij} - \sum_{\substack{i+j=k \\ 0 \leq i,j \leq d-1}} B_{ij}, \text{ for } 0 \leq k \leq 2d-2\}. \end{aligned}$$

Proof. We just prove the case of even degree as the proof for the odd degree case is similar. We have $x \in R_S(\bar{x})$ if and only if $p_x(0) = 0$ and $p_x(\lambda)$ is non-negative on $[0, 1]$. By Theorem 3.33, this is equivalent to $p_x(0) = 0$ and there exist polynomials s_1, s_2 of degree d and $d - 1$, respectively, such that

$$p_x(\lambda) = s_1(\lambda)^2 + \lambda(1 - \lambda)s_2(\lambda)^2.$$

Given that $0 = p_x(0) = s_1(0)^2$, the polynomial s_1 has a root at 0 and we can write it as $s_1(\lambda) = \lambda r_1(\lambda)$ where r_1 is a polynomial of degree $d - 1$. Thus, $x \in R_S(\bar{x})$ if and only if $p_x(0) = 0$ and there exist polynomials r_1, r_2 of degree $d - 1$ such that

$$p_x(\lambda) = \lambda^2 r_1(\lambda)^2 + \lambda(1 - \lambda)r_2(\lambda)^2.$$

Let $\Lambda = (1, \lambda, \ldots, \lambda^{d-1})^\mathsf{T}$. The polynomials r_i can be written as $r_i = c_i^\mathsf{T}\Lambda$ for some $c_i \in \mathbb{R}^d$. Then, $r_1(\lambda)^2 = \Lambda^\mathsf{T} A\Lambda$ and $r_2(\lambda)^2 = \Lambda^\mathsf{T} B\Lambda$ for some $A, B \in \mathcal{S}_+^d$.

Thus, $x \in R_S(\bar{x})$ if and only if $p_x(0) = 0$ and there exist $A, B \in \mathcal{S}_+^d$ such that

$$p_x(\lambda) = \lambda^2\Lambda^\mathsf{T} A\Lambda + \lambda(1 - \lambda)\Lambda^\mathsf{T} B\Lambda.$$

Since $p_x(\lambda)$ is a polynomial of degree $2d$, its Taylor expansion at 0 yields

$$p_x(\lambda) = \sum_{k=1}^{2d} \frac{p_x^{(k)}(0)}{k!}\lambda^k.$$

Identifying coefficients, we conclude the theorem. □

Remark 3.35. One could also add the constraints $\mathrm{rk}(A) = \mathrm{rk}(B) = 1$ to E in the statement of Theorem 3.34. The correctness can easily be seen from the proof since $A = c_1c_1^\mathsf{T}$ and $B = c_2c_2^\mathsf{T}$. Although it makes the set more restricted, the rank constraint is non-convex and does not change the projection. Thus, we decided to leave it out. □

We can recover Theorem 3.29 from Theorem 3.34. The set E of Theorem 3.34 for the quadratic case ($d = 1$) is described by $g(x) = 0$, $p'_x(0) = B_{00}$ and $p''_x(0)/2 = A_{00} - B_{00}$, where $A_{00}, B_{00} \geq 0$. This implies that $0 < g(\bar{x}) = p_x(1) = p'_x(0) + p''_x(0)/2 = A_{00}$. Therefore, $R_S(\bar{x})$ consists of the x such that $p_x(0) = 0$ and $p'_x(0) \geq 0$. This last constraint is equivalent to $\nabla g(x)^\mathsf{T}(\bar{x} - x) \geq 0$ which is the only constraint needed, apart from $g(x) = 0$, to prove Theorem 3.29.

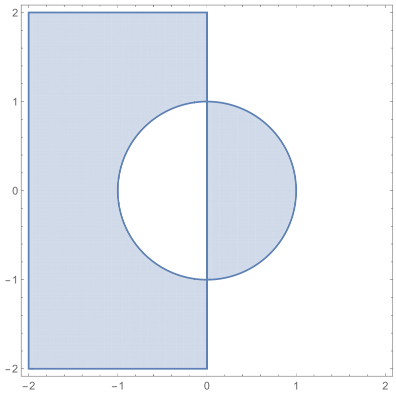


Figure 3.4: Feasible region $g(x) \leq 0$ of Example 3.36 that shows that $\operatorname{cl} V_S(\bar{x}) \neq R_S(\bar{x})$ when the degree of g is greater than 2.

The previous deduction is only possible because $V_S(\bar{x}) = R_S(\bar{x})$ holds for a quadratic constraint. This equality does not hold as soon as the degree is greater than 2, even after replacing $V_S(\bar{x})$ by its closure, as shown in the following example.

Example 3.36. Consider $g(x_1, x_2) = (x_1^2 + x_2^2 - 1)x_1$, $S = \{(x_1, x_2) : g(x_1, x_2) \leq 0\}$, and $\bar{x} = (1, -2)$. The set S consists of the right half of the unit ball and the half space $x_1 \leq 0$ without the interior of the left half of the unit ball, see Figure 3.4. The point $z = (-1, 0)$ is not visible from $\bar{x}$, because $g(z+\lambda(\bar{x}-z)) = g(-1+2\lambda, -2\lambda) = ((2\lambda-1)^2+4\lambda^2-1)(2\lambda-1) = 4\lambda(2\lambda-1)^2$ is zero at $\lambda = \frac{1}{2}$. On the other hand, $z \in R_S(\bar{x})$ since $4\lambda(2\lambda - 1)^2 \geq 0$ for every $\lambda \in [0, 1]$. In this example $V_S(\bar{x})$ is closed, so we conclude that $\operatorname{cl} V_S(\bar{x}) \neq R_S(\bar{x})$. $\square$

3.5 Conclusions and Outlook

Using the concept of visible points, we introduced a technique that allows to reduce the domains in separation problems. Such a result is particularly interesting for MINLP, since the tightness of the domain directly affects the quality of underestimators, from which cuts are obtained.

Some questions that could be interesting to look at in the future are the followings. Is there a tighter domain other than V_S that can be efficiently exploited? Is there a useful characterizations of V_S when S contains more than one non-convex constraint, in particular, if some variables are restricted to be integer?

Chapter 4

Intersection Cuts for Factorable Mixed-Integer Nonlinear Programming

We now move to our final stop, intersection cuts (see Section 1.2). In this chapter we develop a technique for constructing S-free sets where $S = \{x : f(x) \leq 0\}$ and f is an arbitrary *factorable* function. In the next chapter we specialized to the case where f is quadratic and we construct *maximal* S-free sets.

In order to build an S-free for the case that f is factorable, we develop a procedure that constructs a *concave underestimator* of f that is tight at a given point. A peculiarity of these underestimators is that they do not rely on a bounded domain. We propose a strengthening procedure for the intersection cuts that exploits the bounds of the domain. Finally, we propose an extension of monoidal strengthening to take advantage of the integrality of non-basic variables.

In Section 4.1 we introduce our setting, motivate intersection cuts for MINLP by making a parallel between branch and bound for MILP and MINLP, and describe the contributions of the chapter. In Section 4.2 we review some literature and related works. Then we jump right into the construction of concave underestimators in Section 4.3. The improvement using bound information is presented in Section 4.4, while our application of monoidal strengthening appears in Section 4.5. We offer a summary of the chapter in Section 4.6.

This chapter is based on the publication Serrano (2019).

4.1 Motivation

In this chapter we propose a procedure for generating intersection cuts for MINLP. We consider MINLP of the following form

$$\begin{aligned} \max\ & c^\mathsf{T} x \\ \text{s.t. } & g_j(x) \le 0, j \in J \\ & Ax = b \\ & x_i \in \mathbb{Z}, i \in I \\ & x \ge 0, \end{aligned} \tag{4.1}$$

where $J = \{1, \dots, l\}$ denotes the indices of the nonlinear constraints, $g_j \colon \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuous and factorable (see Definition 4.1), $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $I \subseteq \{1, \dots, n\}$ are the indices of the integer variables. We denote the set of feasible solutions by S and a generic relaxation of S by R, that is, $S \subseteq R$.

The current state of the art for solving MINLP to global optimality is via linear programming (LP), convex nonlinear programming and (MILP) relaxations of S, together with spatial branch and bound (Belotti et al., 2009; Kılınç and Sahinidis, 2017; Lin and Schrage, 2009; Misener and Floudas, 2014; Tawarmalani and Sahinidis, 2005; Vigerske and Gleixner, 2017). Let us recall, roughly, how LP-based spatial branch and bound works. The initial polyhedral relaxation is solved and yields $\bar{x}$. If the solution $\bar{x}$ is feasible for (4.1), we obtain an optimal solution. If not, we try to separate the solution from the feasible region. This is usually done by considering each violated constraint separately. Let $g(x) \le 0$ be a violated constraint of (4.1). If $g(\bar{x}) > 0$ and g is convex, then $g(\bar{x}) + v^\mathsf{T}(x - \bar{x}) \le 0$, where $v \in \partial g(\bar{x})$, is a valid cut. If g_j is non-convex, then a convex underestimator g_{vex}, that is, a convex function such that $g_{vex}(x) \le g(x)$ over the feasible region, is constructed and if $g_{vex}(\bar{x}) > 0$ the previous cut is constructed for g_{vex}. If the point cannot be separated, then we branch, that is, we select a variable x_k in a violated constraint and split the problem into two problems, one with $x_k \le \bar{x}_k$ and the other one with $x_k \ge \bar{x}_k$.

Applying the previous procedure to the MILP case, that is (4.1) with $J = \emptyset$, reveals a problem with this approach. In this case, the polyhedral relaxation is just the linear programming (LP) relaxation. Assuming that $\bar{x}$ is not feasible for the MILP, then there is an $i \in I$ such that $x_i \notin \mathbb{Z}$. Let us treat the constraint $x_i \in \mathbb{Z}$ as a nonlinear non-convex constraint represented by some function as $g(x_i) \le 0$. Then, $g(\bar{x}_i) > 0$. A convex underestimator $\bar{g}$ of g must satisfy that $g_{vex}(z) \le 0$ for every $z \in \mathbb{R}$, since $g_{vex}(z) \le g(z) \le 0$ for every $z \in \mathbb{Z}$ and $g_{vex}(z)$ is convex. Thus, separation is not possible and we need to branch. However, for the current state-of-the-art algorithms for MILP, cutting planes are a fundamental component (Achterberg and Wunderling, 2013).

Recall, from Section 1.2, that when solving the LP relaxation, we obtain $x_B = \bar{x}_B + Rx_N$, where B and N are the indices of the basic and non-basic variables, respectively. Since $\bar{x}$ is infeasible for the MILP, there must be some $k \in B \cap I$ such that $\bar{x}_k \notin \mathbb{Z}$. Now, even though $\bar{x}$ cannot be separated from the violated constraint $x_k \in \mathbb{Z}$, the equivalent constraint, $\bar{x}_k + \sum_{j \in N} r_{kj} x_j \in \mathbb{Z}$ can be used to separate $\bar{x}$.

In the MINLP case, this framework generates equivalent non-linear constraints with some appealing properties, in particular, violated points can *always* be separated. The change of variables $x_k = \bar{x}_k + \sum_{j \in N} r_{kj} x_j$ for the basic variables present in a violated nonlinear constraint $g(x) \leq 0$, produces the non-linear constraint $h(x_N) \leq 0$ for which $h(0) > 0$ and $x_N \geq 0$. Assuming that the convex envelope of h exists in $x_N \geq 0$, then we can always construct a valid inequality. Indeed, by Tawarmalani and Sahinidis (2002, Corollary 3), the convex envelope of h is tight at 0. Since an ϵ-subgradient[3] always exists for any $\epsilon > 0$ and $x \in \operatorname{dom} h$ (Brondsted and Rockafellar, 1965), an $\frac{h(0)}{2}$-subgradient, for instance, at 0 will separate it.

Even when there is no convex underestimator for h, a valid cutting plane does exist. Continuity of h implies that $X = \{x_N \geq 0 : h(x_N) \leq 0\}$ is closed and Conforti et al. (2015, Lemma 2.1) ensures that $0 \notin \overline{conv} X$, thus, a valid inequality exists. We introduce a technique to construct such a valid inequality. The idea is to build a *concave underestimator* of h, h_{ave}, such that $h_{ave}(0) = h(0) > 0$. Then, $C = \{x_N : h_{ave}(x_N) \geq 0\}$ is an *S-free set*, that is, a convex set that does not contain any feasible point in its interior, and as such can be used to build an intersection cut (IC) (Tuy, 1964; Balas, 1971; Glover, 1973).

First contribution In Section 4.3, we present a procedure to build concave underestimators for factorable functions that are tight at a given point. The procedure is similar to McCormick's method for constructing convex underestimators, and generalizes Proposition 3.2 and improves Proposition 3.3 of Khamisov (1999). A simple way to build a concave underestimator of a function is to write the function as a difference of convex (d.c.), then, by linearizing the convex part a concave underestimator is obtained. However, even if a function is known to have a d.c. representation, it is not always clear

[3] An ϵ-subgradient of a convex function f at $y \in \operatorname{dom} f$ is v such that $f(x) \geq f(y) - \epsilon + v^{\mathsf{T}}(x - y)$ for all $x \in \operatorname{dom} f$

how to construct it.

These underestimators can be used to build intersection cuts. We note that IC from a concave underestimator can generate cuts that cannot be generated by using the convex envelope. This should not be surprising, given that intersection cuts work at the feasible region level, while convex underestimators depend on the graph of the function. A simple example is $\{x \in [0,2] : -x^2+1 \leq 0\}$. When separating 0, the intersection cut gives $x \geq 1$, while using the convex envelope over $[0,2]$ yields $x \geq 1/2$.

There are many differences between concave underestimators and convex ones. Maybe the most interesting one is that concave underestimators do not need bounded domains to exist. As an extreme example, $-x^2$ is a concave underestimator of itself, but a convex underestimator only exists if the domain of x is bounded. Even though this might be regarded as an advantage, it is also a problem. If concave underestimators are independent of the domain, then we cannot improve them when the domain shrinks.

Second contribution In Section 4.4, we propose a strengthening procedure that uses the bounds of the variables to enlarge the S-free set. Our procedure improves on the one used by Tuy (1964).

Other techniques for strengthening IC have been proposed, such as, exploiting the integrality of the non-basic variables (Balas and Jeroslow, 1980; Conforti et al., 2011a; Dey and Wolsey, 2010), improving the relaxation R (Balas and Margot, 2011; Porembski, 1999, 2001) and computing the convex hull of $R \setminus C$ (Basu et al., 2011; Conforti et al., 2015; Glover, 1974; Sen and Sherali, 1986, 1987).

Third contribution By interpreting IC as disjunctive cuts (Balas, 1979), we extend the monoidal strengthening technique of Balas and Jeroslow (1980) to our setting in Section 4.5. Although its applicability seems to be limited, we think it is of independent interest, especially for MILP.

4.2 Literature Review and Related Work

There have been many efforts on generalizing cutting planes from MILP to MINLP, we refer the reader to Modaresi et al. (2015) and the references therein. Modaresi et al. (2015) study how to compute $\mathrm{conv}(R \setminus C)$ where R is not polyhedral, but C is a k-branch split. In practice, such sets C usually come from the integrality of the variables. Works that build sets C which do not come from integrality considerations include Belotti (2011); Bienstock et al. (2019); Fischetti et al. (2016, 2017); Fischetti and Monaci (2019); Saxena et al. (2010a,b). We refer to Bonami et al. (2011) and the references therein for more details.

Fischetti et al. (2016) applied intersection cuts to bilevel optimization. Bienstock et al. (2016, 2019) studied outer-product-free sets; these can be used for generating intersection cuts for polynomial optimization when using an extended formulation. Fischetti and Monaci (2019) constructed bilinear-free sets through a bound disjunction and, in each term of the disjunction, underestimating the bilinear term with McCormick inequalities (McCormick, 1976). The complement of this disjunction is the bilinear-free set.

We would like to point out that the disjunctions built in Belotti (2011); Fischetti and Monaci (2019); Saxena et al. (2010b,a) can be interpreted as piecewise linear concave underestimators. However, our approach is not suitable for disjunctive cuts built through cut generating LPs (Balas et al., 1993), since we generate infinite disjunctions, see Section 4.5, so we rely on the classic concept of intersection cuts where R is a translated simplicial cone.

Khamisov (1999) studies functions $f : \mathbb{R}^n \to \mathbb{R}$, representable as $f(x) = \max_{y \in R} \varphi(x, y)$ where φ is continuous and concave on x. These functions allow for a concave underestimator at *every* point. He shows that this class of functions is very general, in particular, the class of functions representable as difference of convex functions is a strict subset of this class. He then proposes a procedure to build concave underestimators of composition of functions which is a special case of Theorem 4.4 below. He also suggests how to build an underestimator for the product of two functions over a compact domain. The construction is based on writing the product as a difference of convex and then using a construction for the square of a function. The construction of a convex overestimator of f^2 is based on a piecewise linear overestimation of the function x^2 over the range of f, which is why Khamisov needs a compact domain for f. We simplify the construction for the product and no longer need a compact domain. We still write the product as a d.c. but we use Theorem 4.4 instead of a piecewise linear overestimator, allowing us to drop the compactness assumption.

Although not directly related to our work, other papers that use underestimators other than convex are Buchheim and D'Ambrosio (2016); Buchheim and Traversi (2013); Hasan (2018). We would also like to mention here the work of Towle and Luedtke (2019) that proposes a method for constructing valid cutting planes with a similar approach to intersection cuts, but allowing $\bar{x}$ to not be in the S-free set. The S-free sets developed in this chapter could also be used in their framework.

4.3 Concave Underestimators

In his seminal paper, McCormick (1976) proposed a method to build convex underestimators of *factorable* functions.

Definition 4.1. *Given a set of univariate functions $\mathcal{L}$, e.g.,$\mathcal{L} = \{\cos, \cdot^n, \exp, \log, ...\}$, the set of* factorable functions $\mathcal{F}$ *is the smallest set that contains $\mathcal{L}$, the constant functions, and is closed under addition, product and composition.*

As an example, $e^{-(\cos(x^2)+xy/4)^2}$ is a factorable function for $\mathcal{L} = \{\cos, \exp\}$.

Given the inductive definition of factorable functions, to show a property about them one just needs to show that said property holds for all the functions in $\mathcal{L}$, constant functions, and that it is preserved by the product, addition and composition. For instance, McCormick (1976) proves, constructively, that every factorable function admits a convex underestimator and a concave overestimator, by showing how to construct estimators for the sum, product and composition of two functions for which estimators are known.

An estimator for the sum of two functions is the sum of the estimators. For the product, McCormick uses the well-known McCormick inequalities. Less known is the way McCormick handles the composition $f(g(x))$. Let f_{vex} be a convex underestimator of f and $z_{\min} = \arg\min f_{vex}(z)$. Let g_{vex} be a convex underestimator of g and g^{ave} a concave overestimator. McCormick shows[4] that $f_{vex}(\mathrm{mid}\{g_{vex}(x), g^{ave}(x), z_{\min}\})$ is a convex underestimator of $f(g(x))$, where $\mathrm{mid}\{x, y, z\}$ is the median between x, y and z. It is well known that the optimum of a convex function over a closed interval is given by such a formula, thus

$$f_{vex}(\mathrm{mid}\{g_{vex}(x), g^{ave}(x), z_{\min}\}) = \min\{f_{vex}(z) : z \in [g_{vex}(x), g^{ave}(x)]\},$$

see also Tsoukalas and Mitsos (2014).

Definition 4.2. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be convex, and $f : \mathcal{X} \to \mathbb{R}$ be a function. We say that $f_{ave} : \mathcal{X} \to \mathbb{R}$ is a* concave underestimator of f at $\bar{x} \in \mathcal{X}$ *if f_{ave} is concave, $f_{ave}(x) \leq f(x)$ for every $x \in \mathcal{X}$ and $f_{ave}(\bar{x}) = f(\bar{x})$. Similarly we define a* convex overestimator of f at $\bar{x} \in \mathcal{X}$.

Remark 4.3. For simplicity, we will consider only the case where $\mathcal{X} = \mathbb{R}^n$. This restriction leaves out some common functions like log. One possibility

[4] He actually leaves it as an exercise for the reader.

to include these function is to let the range of the function to be $\mathbb{R} \cup \{\pm\infty\}$. Then, $\log(x) = -\infty$ for $x \in \mathbb{R}_-$. Note that other functions like $\sqrt{x}$ can be handled by replacing them by a concave underestimator defined on all $\mathbb{R}$. $\square$

We now show that every factorable function admits a concave underestimator at a given point. Since the case for the addition is easy, we just need to specify how to build concave underestimators and convex overestimators for

- the product of two functions for which estimators are known,
- the composition $f(g(x))$ where estimators of f and g are known and f is univariate.

Theorem 4.4. *Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. Let g_{ave}, f_{ave} be, respectively, a concave underestimator of g at $\bar{x}$ and of f at $g(\bar{x})$. Further, let g^{vex} be a convex overestimator of g at $\bar{x}$. Then, $h : \mathbb{R}^n \to \mathbb{R}$ given by*

$$h(x) := \min\{f_{ave}(g_{ave}(x)), f_{ave}(g^{vex}(x))\},$$

is a concave underestimator of $f \circ g$ at $\bar{x}$.

Proof. Clearly, $h(\bar{x}) = f(g(\bar{x}))$.

To establish $h(x) \leq f(g(x))$, notice that

$$h(x) = \min\{f_{ave}(z) : g_{ave}(x) \leq z \leq g^{vex}(x)\}. \tag{4.2}$$

Since $z = g(x)$ is a feasible solution and f_{ave} is an underestimator of f, we obtain that $h(x) \leq f(g(x))$.

Now, let us prove that h is concave. To this end, we again use the representation (4.2). To simplify notation, we write g_1, g_2 for g_{ave}, g^{vex}, respectively. We prove concavity by definition, that is,

$$h(\lambda x_1 + (1-\lambda)x_2) \geq \lambda h(x_1) + (1-\lambda)h(x_2), \text{ for } \lambda \in [0,1].$$

Let

$$\begin{aligned} I &= [g_1(\lambda x_1 + (1-\lambda)x_2), g_2(\lambda x_1 + (1-\lambda)x_2)] \\ J &= [\lambda g_1(x_1) + (1-\lambda)g_1(x_2), \lambda g_2(x_1) + (1-\lambda)g_2(x_2)]. \end{aligned}$$

By the concavity of g_1 and convexity of g_2 we have $I \subseteq J$. Therefore,

$$h(\lambda x_1 + (1-\lambda)x_2) = \min\{f_{ave}(z) : z \in I\} \geq \min\{f_{ave}(z) : z \in J\}.$$

Since f_{ave} is concave, the minimum is achieved at the boundary,

$$\min\{f_{ave}(z) : z \in J\} = \min_{i\in\{1,2\}} f_{ave}(\lambda g_i(x_1) + (1-\lambda)g_i(x_2)).$$

Furthermore, $f_{ave}(\lambda g_i(x_1)+(1-\lambda)g_i(x_2)) \geq \lambda f_{ave}(g_i(x_1))+(1-\lambda)f_{ave}(g_i(x_2))$ which implies that

$$\begin{aligned} h(\lambda x_1 + (1-\lambda)x_2) &\geq \min_{i\in\{1,2\}} \lambda f_{ave}(g_i(x_1)) + (1-\lambda)f_{ave}(g_i(x_2)) \\ &\geq \min_{i\in\{1,2\}} \lambda f_{ave}(g_i(x_1)) + \min_{i\in\{1,2\}} (1-\lambda)f_{ave}(g_i(x_2)) \\ &= \lambda h(x_1) + (1-\lambda)h(x_2), \end{aligned}$$

as we wanted to show. □

Remark 4.5. The generalization of Theorem 4.4 to the case where f is multivariate in the spirit of Tsoukalas and Mitsos (2014) is straightforward. □

The computation of a concave underestimator and convex overestimator of the product of two functions reduces to the computation of estimators for the square of a function through the polarization identity

$$f(x)g(x) = \frac{1}{4}(f(x)+g(x))^2 - \frac{1}{4}(f(x)-g(x))^2.$$

This identity is based on writing the product x_1x_2 as a difference of convex. In particular, it can be proven by doing an eigenvalue decomposition of the Hessian of x_1x_2 Let $h : \mathbb{R}^n \to \mathbb{R}$ for which we know estimators $h_{vex} \leq h \leq h^{ave}$ at $\bar{x}$. From Theorem 4.4, a convex overestimator of h^2 at $\bar{x}$ is given by $\max\{{h_{vex}}^2, h^{ave2}\}$. On the other hand, a concave underestimator of h^2 at $\bar{x}$ can be constructed from the underestimator $h^2(x) \geq h^2(\bar{x}) + 2h(\bar{x})(h(x) - h(\bar{x}))$. From here we obtain

$$\begin{cases} 2h(\bar{x})h^{ave}(x) - h^2(\bar{x}), & \text{if } h(\bar{x}) \leq 0 \\ 2h(\bar{x})h_{vex}(x) - h^2(\bar{x}), & \text{if } h(\bar{x}) > 0. \end{cases} \tag{4.3}$$

Example 4.6. Let us compute a concave underestimator of $f(x) = e^{-(\cos(x^2)+x/4)^2}$ at 0. Estimators of x^2 are given by $0 \leq x^2 \leq x^2$. For $\cos(x)$, estimators are $\cos(x) - x^2/2 \leq \cos(x) \leq 1$. Then, a concave underestimator of $\cos(x^2)$ is, according to Theorem 4.4, $\min\{\cos(0)-0^2/2, \cos(x^2)-x^4/2\} = \cos(x^2)-x^4/2$.

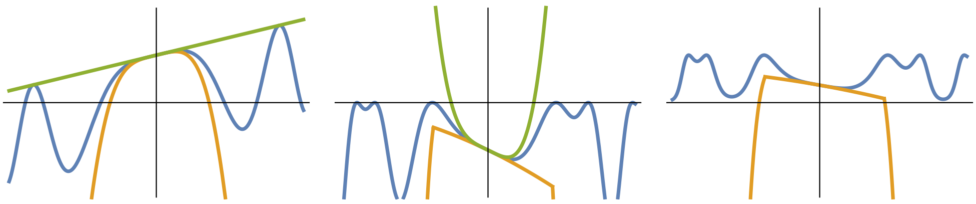


Figure 4.1: Concave underestimator (orange) and convex overestimator (green) of $\cos(x^2) + x/4$ (left), $-(\cos(x^2) + x/4)^2$ (middle) and $f(x)$ (right) at $x = 0$.

A convex overestimator is 1. Hence, $\cos(x^2) - x^4/2 + x/4 \leq \cos(x^2) + x/4 \leq 1 + x/4$.

Given that $-x^2$ is concave, a concave underestimator of $-(\cos(x^2)+x/4)^2$ is $\min\{-(\cos(x^2) - x^4/2 + x/4)^2, -(1 + x/4)^2\}$. To compute a convex overestimator of $-(\cos(x^2)+x/4)^2$, we compute a concave underestimator of $(\cos(x^2)+x/4)^2$. Since, $\cos(x^2) + x/4$ at 0 is 1, (4.3) yields $2(\cos(x^2) - x^4/2 + x/4) - 1$.

Finally, a concave underestimator of e^x at $x = -1$ is just its linearization, $e^{-1}+e^{-1}(x+1)$ and so $e^{-1}+e^{-1}(1+\min\{-(\cos(x^2)-x^4/2+x/4)^2, -(1+x/4)^2\})$ is a concave underestimator of $f(x)$. The intermediate estimators as well as the final concave underestimator are illustrated in Figure 4.1. □

For ease of exposition, in the rest of the chapter we assume that the concave underestimator is differentiable. All results can be extended to the case where the functions are only sub- or super-differentiable.

4.3.1 Concave Underestimators and Intersection Cuts for Convex Constraints

Here we show that if we apply our procedure to construct an S-free set from a violated convex constraint and compute an intersection cut using the smallest representation (see Section 1.2), we just recover the gradient cut. Even more this gradient cut is the same that we would have computed in the original space. In particular, the point is separable in the original space if and only if it is separable in the non-basic space. If one recalls that gradient cuts do not use bounds information, then this might not be surprising.

Let $g(x)$ be a differentiable convex function and consider the constraints $g(x) \leq 0$. Suppose $x_B = f + Rx_N$ is the current optimal tableau and $(x_B, x_N) = (f, 0)$ the optimal LP solution. Further, assume that $g(f, 0) > 0$.

Let $h(x_N) = g(f + Rx_N, x_N)$ and note that this function is still convex since it is the composition of a convex function with an affine map. A concave underestimator at 0 is just the linearization of h at 0, that is,

$$h(0) + \nabla h(0)^\mathsf{T} x_N.$$

Then, the S-free set is $C = \{x_N : h(0)+\nabla h(0)^\mathsf{T} x_N \geq 0\} = \{x_N : -\frac{1}{h(0)}\nabla h(0)^\mathsf{T} x_N \leq 1\}$. Thus the smallest representation is given by the sublinear function (actually, linear) $\rho(x_N) = -\frac{1}{h(0)}\nabla h(0)^\mathsf{T} x_N$. In the space of the non-basic variables the rays are just e_i for $i \in N$. Thus, the intersection cut is $\sum_{i\in N} \rho(e_i) x_i \geq 1$, that is, $-\frac{1}{h(0)}\nabla h(0)^\mathsf{T} x_N \geq 1$. Manipulating the last expression we arrive at $h(0) + \nabla h(0)^\mathsf{T} x_N \leq 0$. This is the same as the gradient cut of h at 0.

Furthermore,

$$\begin{aligned} h(0) + \nabla h(0)^\mathsf{T} x_N &= g(f,0) + \nabla g(f,0)^\mathsf{T} \begin{pmatrix} R \\ I \end{pmatrix} x_N \\ &= g(f,0) + \nabla g(f,0)^\mathsf{T} \begin{pmatrix} x_B - f \\ x_N \end{pmatrix} \end{aligned}$$

This last expression is the gradient cut of g at $(f,0)$.

Thus, there is nothing to be gain from this approach for convex constraints. An interesting observation, in connection to Chapter 2, is that the S-free set, either $\{x_N : h(0) + \nabla h(0)^\mathsf{T} x_N \geq 0\}$ in the non-basic space or $\{x : g(f,0) + \nabla g(f,0)^\mathsf{T} \begin{pmatrix} x_B - f \\ x_N \end{pmatrix} \geq 0\}$, is not going to be maximal if it does not support the constraint. In particular, if g is strictly convex the S-free set is not maximal. This will be important in the next chapter.

Remark 4.7. Also, this already provides evidence that the S-free sets constructed by our approach will not be maximal in general. Assume we have a function g, $S = \{x : g(x) \leq 0\}$, and we write g as a difference of convex $g = f - h$. Say we linearize the function f at a point $\bar{x}$ such that $g(\bar{x}) > 0$ to obtain $f \geq l$. Then, the concave underestimator is $l - h \leq g$ and the S-free set is $l - h \geq 0$. If f is strictly convex, we would have $l(x) < f(x)$ for every $x \neq \bar{x}$. This S-free set will not touch S. If it did, that is, if there is a point x both in S and the S-free set, then $f(x) - h(x) \geq l(x) - h(x) \geq 0 \geq g(x) = f(x) - h(x)$, thus $x = \bar{x}$ and $g(\bar{x}) = 0$, which contradicts our assumption.

This argument is very far from a proof since, first, our procedure does not really construct a d.c. decomposition, but rather use a d.c. as an intermediate step for the product. Second, an S-free does not need to touch S in order to be maximal (see Chapter 5). $\square$

4.4 Enlarging the S-free Sets by Using Bound Information

In Section 4.3, we showed how to build concave underestimators which give us S-free sets. Note that the construction does not make use of the bounds of the domain. We can exploit the bounds of the domain by the observation that the concave underestimator only needs to underestimate within the feasible region. However, to preserve the convexity of the S-free set, we must ensure that the underestimator is still concave.

Let $h(x) \leq 0$ be a constraint of (4.1), assume $x \in [l, u]$ and let h_{ave} be a concave underestimator of h. Throughout this section, $S = \{x \in [l, u] : h(x) \leq 0\}$. In order to construct a concave function $\hat{h}$ such that $\{x : \hat{h}(x) \geq 0\}$ contains $\{x : h_{ave}(x) \geq 0\}$, consider the following function

$$\hat{h}(x) = \min\{h_{ave}(z) + \nabla h_{ave}(z)^{\mathsf{T}}(x - z) : \; z \in [l, u], \; h_{ave}(z) \geq 0\}. \tag{4.4}$$

A similar function was already considered by Tuy (1964). The only difference is that Tuy's strengthening does not use the restriction $h_{ave}(z) \geq 0$, see Figure 4.2.

Proposition 4.8. *Let h_{ave} be a concave underestimator of h at $\bar{x} \in [l, u]$, such that $h(\bar{x}) > 0$. Define $\hat{h}$ as in (4.4). Then, the set $C = \{x : \hat{h}(x) \geq 0\}$ is a convex S-free set and $C \supseteq \{x : h_{ave}(x) \geq 0\}$.*

Proof. The function $\hat{h}$ is concave since it is the minimum of linear functions. This establishes the convexity of C.

To show that $C \supseteq \{x : h_{ave}(x) \geq 0\}$, notice that $h_{ave}(x) = \min_z h_{ave}(z) + \nabla h_{ave}(z)^{\mathsf{T}}(x - z)$. The inclusion follows from observing that the objective function in the definition of $\hat{h}(x)$ is the same as above, but over a smaller domain.

To show that it is S-free, we will show that for every $x \in [l, u]$ such that $h(x) \leq 0$, $\hat{h}(x) \leq 0$.

Let $x_0 \in [l, u]$ such that $h(x_0) \leq 0$. Since h_{ave} is a concave underestimator at $\bar{x}$, $h_{ave}(\bar{x}) > 0$ and $h_{ave}(x_0) \leq 0$. If $h_{ave}(x_0) = 0$, then, by definition, $\hat{h}(x_0) \leq h_{ave}(x_0) = 0$ and we are done. We assume, therefore, that $h_{ave}(x_0) < 0$.

Consider $g(\lambda) = h_{ave}(\bar{x} + \lambda(x_0 - \bar{x}))$ and let $\lambda_1 \in (0, 1)$ be such that $g(\lambda_1) = 0$. The existence of λ_1 is justified by the continuity of g, $g(0) > 0$ and $g(1) < 0$. Equivalently, $x_1 = \bar{x} + \lambda_1(x_0 - \bar{x})$ is the intersection point between the segment joining x_0 with $\bar{x}$ and $\{x : h_{ave}(x) = 0\}$. The linearization of g at λ_1 evaluated at $\lambda = 1$ is negative, because g is concave, and equals $h_{ave}(x_1) + \nabla h_{ave}(x_1)^T(x_0 - x_1)$. Finally, given that $x_1 \in [l, u]$ and $h_{ave}(x_1) = 0$, x_1 is feasible for (4.4) and we conclude that $\hat{h}(x_0) < 0$. $\square$

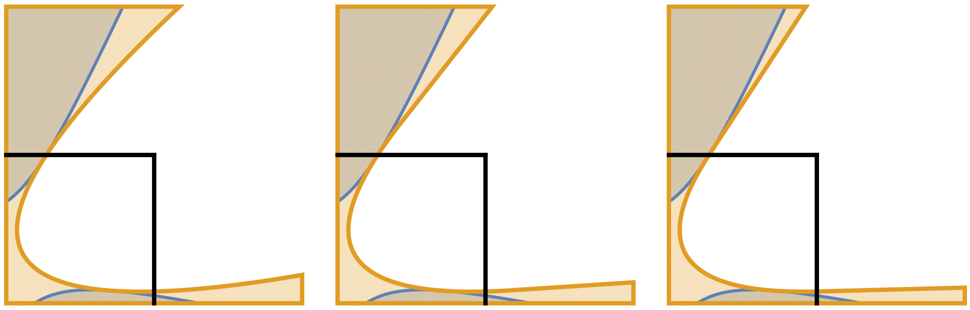


Figure 4.2: Feasible region $\{x, y \in [0,2] : h(x,y) \leq 0\}$, where $h = x^2 - 2y^2 + 4xy - 3x + 2y + 1$, in blue together with $h_{ave}(x,y) \leq 0$ at $\bar{x} = (1,1)$ (left), Tuy's strengthening (middle) and $\hat{h} \leq 0$ (right) in orange. Region shown is $[0,4]^2$, $[0,2]^2$ is bounded by black lines. The difference between the S-free sets can be seen on the top of the picture.

In general, evaluating $\hat{h}$ is a difficult problem and there is no closed form formula. However, when h_{ave} is quadratic, the problem in the right hand side of (4.4) is convex and a cut could be strengthen in polynomial time.

4.5 "Monoidal" Strengthening

We show how to strengthen cuts from reverse convex constraints when exactly *one* non-basic variable is integer. Our technique is based on monoidal strengthening applied to disjunctive cuts, see Lemma 4.10 and the discussion following it. If more than one variable is integer, we can generate one cut per integer variable, relaxing the integrality of all but one variable at a time. However, under some conditions (see Remark 4.12), we can exploit the integrality of several variables at the same time. For an introduction to monoidal strengthening see Section 1.4.

Throughout this section, we assume that we already have a concave underestimator, and that we have performed the change of variables described in the introduction. Therefore, we consider the constraint $\{x \in [0,u] : h(x) \leq 0\}$ where $h : \mathbb{R}^n \to \mathbb{R}$ is concave and $h(0) > 0$. Let $Y = \{y \in [0,u] : h(y) = 0\}$. The convex S-free set $C = \{x \in [0,u] : h(x) \geq 0\}$ can be written as

$$C = \bigcap_{y \in Y} \{x \in [0,u] : \nabla h(y)^\mathsf{T} x \geq \nabla h(y)^\mathsf{T} y\}.$$

The concavity of h implies that $h(0) \leq h(y) - \nabla h(y)^\mathsf{T} y$ for all y in the domain of h. In particular, if $y \in Y$, then $\nabla h(y)^\mathsf{T} y \leq -h(0) < 0$. Since all feasible points satisfy $h(x) \le 0$, they must satisfy the infinite disjunction

$$\bigvee_{y \in Y} \frac{\nabla h(y)^\mathsf{T}}{\nabla h(y)^\mathsf{T} y} x \ge 1. \tag{4.5}$$

The maximum principle (see Section 1.4) implies that with

$$\alpha_j = \max_{y \in Y} \frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y}, \tag{4.6}$$

the cut $\sum_j \alpha_j x_j \ge 1$ is valid. We remark that the maximum exists, since the concavity of h implies that for $y \in Y$, $h(e_j) \le \partial_j h(y) - \nabla h(y)^\mathsf{T} y$. This implies, together with $\nabla h(y)^\mathsf{T} y \le -h(0) < 0$, that $\frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y} \le 1 + \frac{h(e_j)}{\nabla h(y)^\mathsf{T} y}$. If $h(e_j) \ge 0$, then $\frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y} \le 1$. Otherwise, $\frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y} \le 1 - \frac{h(e_j)}{h(0)}$.

The application of monoidal strengthening (Balas and Jeroslow, 1980, Theorem 3) to a valid disjunction $\bigvee_i \alpha^i x \ge 1$ requires the existence of bounds β_i such that $\alpha^i x \ge \beta_i$ is valid for every feasible point. Let $\beta(y)$ be such a bound for (4.5). An example of $\beta(y)$ is

$$\beta(y) = \min_{x \in [0,u]} \frac{\nabla h(y)^\mathsf{T} x}{\nabla h(y)^\mathsf{T} y}.$$

Remark 4.9. If $\beta(y) \ge 1$, then $\nabla h(y)^\mathsf{T} x / \nabla h(y)^\mathsf{T} y \ge 1$ is redundant and can be removed from (4.5). Therefore, we can assume without loss of generality that $\beta(y) < 1$. $\square$

The following lemma is just a restatement of Lemma 1.1 in Section 1.4.

Lemma 4.10. *Every $x \ge 0$ that satisfies* (4.5)*, also satisfies*

$$\bigvee_{y \in Y} \frac{\nabla h(y)^\mathsf{T} x}{\nabla h(y)^\mathsf{T} y} + z(y)(1 - \beta(y)) \ge 1, \tag{4.7}$$

where $z \colon Y \to \mathbb{Z}$ is such that $z \equiv 0$ or there is a $y_0 \in Y$ for which $z(y_0) > 0$.

Proof. If $z \equiv 0$, then (4.7) reduces to (4.5).

Otherwise, let $y_0 \in Y$ such that $z(y_0) > 0$, that is, $z(y_0) \ge 1$. By Remark 4.9, for every $y \in Y$, it holds $1 - \beta(y) > 0$, and so

$$z(y_0)(1 - \beta(y_0)) \ge 1 - \beta(y_0).$$

Therefore, $\beta(y_0) \ge 1 - z(y_0)(1 - \beta(y_0))$. Since every $x \ge 0$ satisfying (4.5) satisfies $\frac{\nabla h(y_0)^\mathsf{T} x}{\nabla h(y_0)^\mathsf{T} y_0} \ge \beta(y_0)$, we conclude that $\frac{\nabla h(y_0)^\mathsf{T} x}{\nabla h(y_0)^\mathsf{T} y_0} + z(y_0)(1 - \beta(y_0)) \ge 1$ holds. $\square$

Remark 4.11. Even if some disjunctive terms have no lower bound, that is, $\beta(y) = -\infty$ for $y \in Y' \subseteq Y$, Lemma 4.10 still holds if, additionally, $z(y) = 0$ for all $y \in Y'$. This means that we are not using that disjunction for the strengthening. In particular, if for some variable x_j, α_j is defined by some $y \in Y'$, then this cut coefficient cannot be improved. □

Assume now that $x_k \in \mathbb{Z}$ for every $k \in K \subseteq \{1, \ldots, n\}$. One way of constructing a new disjunction is to find a set of functions M such that for any choice of $m^k \in M$ and any feasible assignment of x_k, $z(y) := \sum_{k \in K} x_k m^k(y)$ satisfies the conditions of Lemma 4.10, that is, z is in

$$Z = \{z \colon Y \to \mathbb{Z} : z \equiv 0 \vee \exists y \in Y, z(y) > 0\}.$$

Once such a family of functions has been identified, the cut $\sum_j \gamma_j x_j \geq 1$ with $\gamma_j = \alpha_j$ if $j \notin K$, and

$$\gamma_k = \inf_{m \in M} \max_{y \in Y} \frac{\partial_k h(y)}{\nabla h(y)^\mathsf{T} y} + m(y)(1 - \beta(y)) \quad \text{for } k \in K, \tag{4.8}$$

is valid and at least as strong as (4.6). Any $M \subseteq Z$ such that $(M, +)$ is a monoid, that is, $0 \in M$ and M is closed under addition can be used in (4.8).

The question that remains is how to choose M. For example, the monoid $M = \{m \colon Y \to \mathbb{Z} : m$ has finite support and $\sum_{y \in Y} m(y) \geq 0\}$ is an obvious candidate for M. However, the problem is how to optimize over such an M, see (4.8).

We circumvent this problem by considering only one integer variable at a time. Fix $k \in K$. In this setting we can use Z as M, which is *not* a monoid. Indeed, if $z \in Z$, then $x_k z \in Z$ for any $x_k \in \mathbb{Z}_+$. The advantage of using Z is that the solution of (4.8) is easy to characterize.

With $M = Z$, the cut coefficients (4.8) of all variables are the same as (4.6) except for x_k. The cut coefficient of x_k is given by

$$\inf_{z \in Z} \max_{y \in Y} \frac{\partial_k h(y)}{\nabla h(y)^\mathsf{T} y} + z(y)(1 - \beta(y)).$$

To compute this coefficient, observe that one would like to have $z(y) < 0$ for points y such that the objective function of (4.6) is large. However, z must be positive for at least one point. Therefore,

$$\min_{y \in Y} \frac{\partial_k h(y)}{\nabla h(y)^\mathsf{T} y} + (1 - \beta(y))$$

is the best coefficient we can hope for if $z \not\equiv 0$. This coefficient can be achieved by

$$z(y) = \begin{cases} 1, & \text{if } y \in \arg\min_{y \in Y} \frac{\partial_k h(y)}{\nabla h(y)^\mathsf{T} y} + (1 - \beta(y)), \\ -L, & \text{otherwise} \end{cases} \tag{4.9}$$

where $L > 0$ is sufficiently large.

Summarizing, we can obtain the following cut:

$$\alpha_j = \begin{cases} \max_{y \in Y} \frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y} & \text{if } j \neq k \\ \min\{\max_{y \in Y} \frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y}, \min_{y \in Y} \frac{\partial_j h(y)}{\nabla h(y)^\mathsf{T} y} + (1 - \beta(y))\} & \text{if } j = k \end{cases} \tag{4.10}$$

Remark 4.12. Let $z^k \in Z$ be given by (4.9) for each $k \in K$. Assume there is a subset $K_0 \subseteq K$ and a monoid $M \subseteq Z$ such that $z^k \in M$ for every $k \in K_0$. Then, the strengthening can be applied to all x_k for $k \in K_0$.

Alternatively, if there is a constraint enforcing that at most one of the x_k can be non-zero for $k \in K_0$, e.g., $\sum_{k \in K} x_k \leq 1$, then the strengthening can be applied to all x_k for $k \in K_0$.

In the finite case, our application of monoidal strengthening would be dominated by the original technique of Balas and Jeroslow (1980) by using an appropriate monoid. However, in the presence of extra constraint, such as the one described above, our technique can dominate vanilla monoidal strengthening. □

Example 4.13. Consider the constraint $\{x \in \{0, 1, 2\} \times [0, 5] : h(x) \leq 0\}$, where $h(x_1, x_2) = -10x_1^2 - 1/2x_2^2 + 2x_1x_2 + 4$, see Figure 4.3. The IC is given by $\sqrt{5/2}x_1 + 1/(2\sqrt{2})x_2 \geq 1$. Note that $(1/\sqrt{10}, \sqrt{10}) \in Y$ and yields the term $1/\sqrt{10}x_2 \geq 1$ in (4.5). Since $x_2 \geq 0$, $\beta(1/\sqrt{10}, \sqrt{10}) = 0$. Hence, (4.10) yields $\alpha_1 \leq \min\{\sqrt{5/2}, 1\} = 1$ and the strengthened inequality is $x_1 + 1/(2\sqrt{2})x_2 \geq 1$. □

4.6 Conclusions

We have introduced a procedure to generate concave underestimators of factorable functions, which can be used to generate intersection cuts, together with two strengthening procedures.

It remains to be seen the practical performance of these intersection cuts. We expect that its generation is cheaper than the generation of disjunctive cuts, given that there is no need to solve an LP. As for the strengthening procedures, they might be too expensive to be of practical use. An alternative is to

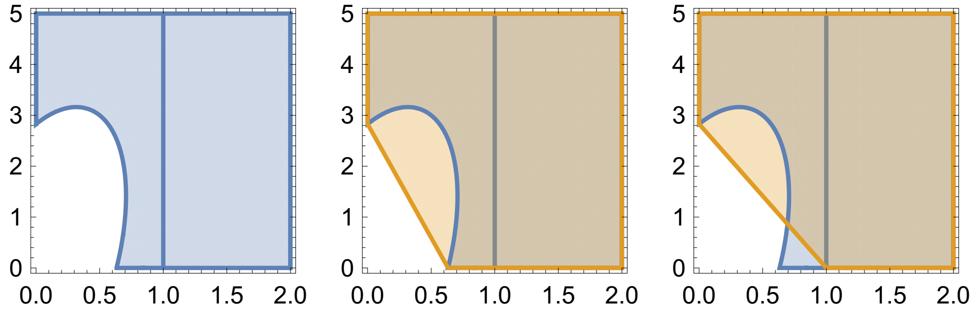


Figure 4.3: The feasible region $\{x \in \{0,1,2\} \times [0,5] : h(x) \leq 0\}$ from Example 4.13 (left), the IC (middle), and the strengthened cut (right).

construct a polyhedral inner approximation of the S-free set and use monoidal strengthening in the finite setting. However, in this case, the strengthening proposed in Section 4.4 has no effect. Nonetheless, as far as the author knows, this has been the first application of monoidal strengthening that is able to exploit further problem structure such as demonstrated in Remark 4.12 and it might be interesting to investigate further.

With respect to maximality, we cannot expect, in principle, that the S-free sets constructed via the techniques presented here is maximal. In the next chapter we show how to construct maximal S-free sets when S is described by a single quadratic constraint.

Chapter 5

Maximal Quadratic-Free Sets

As we discussed in Section 1.2, classic intersection cuts are undominated when they are generated from maximal S-free sets. However, maximality can be a challenging goal in general. In this chapter, we show how to construct maximal S-free sets when S is defined as a general quadratic inequality.

The chapter is organized as follows. In Section 4.1 we introduce our setting, review some related work and describe the contributions of the chapter. In Section 5.2 we introduce some definitions and necessary conditions to prove maximality of S-free sets. In particular, we define *exposing points* and *exposing point at infinity* and show that if C is an S-free set whose defining inequalities are exposed or exposed at infinity, then C is maximal. In Section 5.3 we show how to construct maximal S-free sets when S is defined by a *homogeneous* quadratic function. Section 5.4 presents the construction of maximal S-free sets when S is defined by a *homogeneous* quadratic function and a homogeneous linear inequality constraints. The construction of a maximal S-free set when S is the sublevel set of any quadratic function is presented in Section 5.5. Our constructions depend on a "canonical" representation of the set S. The effects of this representation are discussed in Section 5.6. In Section 5.6 we collect some generalizations and remarks. In particular, we generalize the construction of Section 5.3 to show how to construct construct maximal S-free set when S is the 0-sublevel set of a difference of sublinear functions. We also show how to handle more than one homogeneous linear inequality, extending the result of Section 5.4. We discuss how our results can extend the work of Bienstock et al. (2019) by constructing maximal outer-product-free sets when the considered 2 by 2 minor contains entries to the diagonal. We show, via an example, that our construction does *not* capture every possible maximal quadratic-free set, even in the homogeneous case.

The cuts developed in this section have been implemented in Chmiela (2020). We briefly discuss their computational impact on Section 5.8. We offer a summary and some directions for further research on Section 5.9. Finally, we present some omitted proofs in Section 5.10.

This chapter is joint work with Gonzalo Muñoz. An extended abstract based on this chapter has been accepted on the proceedings of Integer Programming and Combinatorial Optimization Muñoz and Serrano (2020).

5.1 Background

Consider a generic optimization problem,

$$\min \quad c^\mathsf{T} x \tag{5.1a}$$
$$\text{s.t.} \quad x \in S \subseteq \mathbb{R}^n. \tag{5.1b}$$

A particularly important case is obtained when (5.1) is a quadratic problem, that is,

$$S = \{x \in \mathbb{R}^n \,:\, x^\mathsf{T} Q_i x + b_i^\mathsf{T} x + c_i \leq 0,\, i = 1, \dots, m\}$$

for certain $n \times n$ matrices Q_i, not necessarily positive semi-definite. Note that if $\bar{x} \notin S$, there exists $i \in \{1, \dots, m\}$ such that

$$\bar{x} \notin S_i := \{x \in \mathbb{R}^n \,:\, x^\mathsf{T} Q_i x + b_i^\mathsf{T} x + c_i \leq 0\},$$

and constructing an S_i-free set containing $\bar{x}$ would suffice to ensure separation. Thus, slightly abusing notation, given $\bar{x}$ we focus on a systematic way of constructing S-free sets containing $\bar{x}$, where S is defined using a single quadratic inequality:

$$S = \{x \in \mathbb{R}^n \,:\, x^\mathsf{T} Q x + b^\mathsf{T} x + c \leq 0\}.$$

As a final note, if we consider the simplest form of intersection cuts, where the cuts are computed using the intersection points of the S-free set and the extreme rays of the simplicial conic relaxation of S (i.e., using the gauge), then the largest the S-free set the better. In other words, if two S-free sets C_1, C_2 are such that $C_1 \subsetneq C_2$, the intersection cut derived from C_2 is stronger than the one derived from C_1 Conforti et al. (2015). Therefore, we aim at computing *maximal* S-free sets.

5.1.1 Related Work

From all the works that construct intersection cuts in a non-linear setting reviewed in Section 4.2, the only one that ensures maximality of the corresponding S-free sets is the work of Bienstock et al. (2016, 2019). While their approach can also be used to generate cutting planes in our setting (general quadratic inequalities), the definition of S differs: Bienstock et al. use a moment-based extended formulation of polynomial optimization problems (Shor, 1987; Lasserre, 2001; Laurent, 2009) and from there define S as the set of matrices which are positive semi-definite and of rank 1, which the authors refer to as *outer-products*. Maximality is computed with respect to this notion. It is unclear if a maximal *outer-product*-free set can be converted into a maximal quadratic-free set. There is an even more fundamental difference that makes these approaches incomparable at this point: in a quadratic setting, the approach of Bienstock et al. would compute a cutting plane in extended space of dimension proportional to n^2, whereas our approach can construct a maximal S-free set in the original space. The quadratic dimension increase can be a drawback in some applications, however stronger cuts can be derived from extended formulations in some cases (Bodur et al., 2017). A thorough comparison of these approaches is subject of future work.

5.1.2 Contribution

The main contribution of this chapter is an explicit construction of maximal S-free sets, when S is defined using a non-convex quadratic inequality (Theorem 5.36 and Theorem 5.46). We achieve this by relying on the fact that any quadratic inequality can be represented using a homogeneous quadratic inequality intersected with a linear *equality*. While these maximal S-free sets are constructed using semi-infinite representations, we show equivalent closed-form representations of them.

In order to construct these sets, we also derive maximal S-free sets for sets S defined as the intersection of a homogeneous quadratic inequality intersected with a linear *homogeneous inequality*. These are an important intermediate step in our construction, but they are of independent interest as well.

In order to show our results, we state and prove a criterion for maximality of S-free sets which generalizes a criterion proven by Dey and Wolsey (the 'only if' of (Dey and Wolsey, 2010, Proposition A.4)) in the case of maximal *lattice-free* sets (Definition 5.2 and Theorem 5.6). We also develop a new criterion that can handle a special phenomenon that arises in our setting and also in non-linear integer programming: the boundary of a maximal S-free set may not even intersect S. Instead, the intersection might be "at infinity". We formalize this in Definition 5.9 and show the criterion in Theorem 5.11.

5.1.3 Notation

Perhaps the least standard notation we use is denoting an inequality $\alpha^\mathsf{T} x \leq \beta$ by (α, β). If $\beta = 0$ we denote it as well as α. This is based on the fact that in the polar of a convex set —roughly, the set of all valid inequalities— the inequalities are points and, although we do not use any polarity results, many of the ideas in this chapter were originally developed from looking at the polar.

5.2 Preliminaries

In this section we collect definitions and results that are going to be useful later on. As we mentioned above, our main object of study is the set $S = \{x \in \mathbb{R}^p : q(x) \leq 0\} \subseteq \mathbb{R}^p$, where q is a quadratic function. To make the analysis easier, we can work on $\mathbb{R}^{p+1}$ and consider the cone generated by $S \times \{1\}$, namely, $\{(x, z) \in \mathbb{R}^{p+1} : z^2 q(\frac{x}{z}) \leq 0,\ z \geq 0\}$. To recover the original S, however, we must intersect the cone with $z = 1$. Since we are interested in maximal S-free sets, this motivates the following definition, see also Basu et al. (2010a).

Definition 5.1. *Given $S, C, H \subseteq \mathbb{R}^n$ where S is closed, C is closed and convex and H is an affine hyperplane, we say that C is* S-free with respect to H *if $C \cap H$ is $S \cap H$-free w.r.t the induced topology in H. We say C is* maximal S-free with respect to H*, if for any $C' \supseteq C$ that is S-free with respect to H it holds that $C' \cap H \subseteq C \cap H$.*

5.2.1 Techniques for Proving Maximality

In this section we describe some sufficient conditions to prove that a convex set C is maximal S-free which will be used in the chapter.

A sufficient (and necessary) condition for a full dimensional convex C lattice-free (that is, $S = \mathbb{Z}^n$) set to be maximal is that C is a polyhedron and there is a point of $\mathbb{Z}^n$ in the relative interior of each of its facets (Conforti et al., 2014, Theorem 6.18). More generally, if C is a full dimensional S-free polyhedron such that there is a point of S in the relative interior of each facet, then C is maximal. The problem with extending this property to non-polyhedral maximal S-free sets is that they might not even have facets, e.g., if S is the complement of $\operatorname{int} B_1(0)$ and C is $B_1(0)$ in dimension 3 or higher. The motivation of the next definition is to capture the property of a facet that is key for proving maximality.

Definition 5.2. *Given a convex set $C \subseteq \mathbb{R}^n$ and a valid inequality $\alpha^\mathsf{T} x \leq \beta$, we say that a point $x_0 \in \mathbb{R}^n$* exposes (α, β) *with respect to C or that (α, β) is* exposed *by x_0 if*

– $\alpha^\mathsf{T} x_0 = \beta$ *and,*

– *if* $\gamma^\mathsf{T} x \leq \delta$ *is any other non-trivial valid inequality for* C *such that* $\gamma^\mathsf{T} x_0 = \delta$*, then there exists a* $\mu > 0$ *such that* $\gamma = \mu\alpha$ *and* $\beta = \mu\delta$*.*

In some cases we omit saying "with respect to C*" if it is clear from context.*

To get some intuition, if C is a polyhedron and $x \in C$ exposes an inequality, then that inequality is a facet and x is in the relative interior of the facet.

Remark 5.3. It is very important to note that if there exists a point exposing a valid inequality of C, then C is full dimensional. The reader should keep this in mind throughout the whole chapter. □

Remark 5.4. For some convex C, a point $x \notin C$ can expose a valid inequality of C. For instance, consider $C = \{x \in \mathbb{R}^2 \,:\, x_1 + x_2 \geq 1\}$. Then $(0,0) \notin C$ and exposes $x_1 + x_2 \geq 0$.

The name "exposed inequality" comes from the concept of *exposed* point, see Section 1.1. Actually, from the standard duality between points and hyperplanes (a hyperplane can be characterized by its normal which is a point), one can interpret a exposed inequality just as the dual of an exposed point. In more details and to simplify ideas, let us assume that $0 \in \operatorname{int}(C)$. Recall that a point $x_0 \in C$ is exposed if there exists a valid inequality of C, $\alpha^\mathsf{T} x \leq 1$, such that $\{x \in C \,:\, \alpha^\mathsf{T} x = 1\} = \{x_0\}$. If α_0 is an exposed point of the polar of C, $C^\circ = \{\alpha \,:\, \alpha^\mathsf{T} x \leq 1,\ \forall x \in C\}$, then there is a valid inequality, $x_0^\mathsf{T}\alpha \leq 1$, such that $\{\alpha \in C^\circ \,:\, x_0^\mathsf{T}\alpha = 1\} = \{\alpha_0\}$. In other words, if $\alpha^\mathsf{T} x \leq 1$ is valid for C (i.e. $\alpha \in C^\circ$) and $\alpha^\mathsf{T} x_0 = 1$, then $\alpha = \alpha_0$. We see that x_0 is a point (direction) that shows that α_0 is an exposed inequality, or, that x_0 *exposes* α_0. See also Lemma 5.15. □

We now show that our definition is indeed helpful to show maximality.

Theorem 5.5. *Let* $K, K' \subseteq \mathbb{R}^n$ *be convex sets such that* $K \subseteq K'$*. If* $\alpha^\mathsf{T} x \leq \beta$ *is*

– *valid for* K*,*

– *not valid for* K'*, and*

– *exposed by* $x_0 \in K$ *with respect to* K*,*

then $x_0 \in \operatorname{int}(K')$*.*

Proof. As $x_0 \in K$ exposes $\alpha^\mathsf{T} x \leq \beta$, it holds that $\alpha^\mathsf{T} x_0 = \beta$ and, thus, x_0 is in the boundary of K. Suppose x_0 is not in the interior of K'. Then it must be in the boundary of K' and there is a valid inequality for K', $\gamma^\mathsf{T} x \leq \delta$, such that $\gamma^\mathsf{T} x_0 = \delta$.

As $K \subsetneq K'$, $\gamma^\mathsf{T} x \leq \delta$ is also valid for K. Given that (γ, δ) is tight at x_0 and x_0 exposes (α, β), we conclude that there is a $\mu > 0$ such that $\gamma = \mu\alpha$ and $\beta = \mu\delta$. However, since $\alpha^\mathsf{T} x \leq \beta$ is not valid for K', it follows that $\gamma^\mathsf{T} x \leq \delta$ cannot be valid for K'. This contradiction proves the claim. $\square$

Theorem 5.6. *Let $S \subseteq \mathbb{R}^n$ be a closed set and $C \subseteq \mathbb{R}^n$ a convex S-free set. Assume that $C = \{x \in \mathbb{R}^n : \alpha^T x \leq \beta, \forall (\alpha, \beta) \in \Gamma\}$ and that for every (α, β) there is an $x \in S \cap C$ that exposes (α, β). Then, C is maximal S-free.*

Proof. To show that C is maximal we are going to show that for every $\bar{x} \notin C$, $S \cap \operatorname{int}(\operatorname{conv}(C \cup \{\bar{x}\}))$ is nonempty.

Let $\bar{x} \notin C$ and let $(\alpha, \beta) \in \Gamma$ be a separating inequality, i.e., $\alpha^\mathsf{T} \bar{x} > \beta$. Let $C' = \operatorname{conv}(C \cup \{\bar{x}\})$.

By hypothesis, there is an $x_0 \in S \cap C$ that exposes (α, β). Since (α, β) is valid for C and not for C', Theorem 5.5 implies that $x_0 \in \operatorname{int}(C')$. $\square$

With minor modifications one can also get the following sufficient condition for maximality with respect to a hyperplane.

Theorem 5.7. *Let $S \subseteq \mathbb{R}^n$ be a closed set, H be an affine hyperplane, and $C \subseteq \mathbb{R}^n$ be a convex S-free set. Assume that $C = \{x \in \mathbb{R}^n : \alpha^T x \leq \beta, \forall (\alpha, \beta) \in \Gamma\}$ and that for every (α, β) there is an $x \in S \cap C \cap H$ that exposes (α, β). Then, C is maximal S-free with respect to H.*

Remark 5.8. Points that expose inequalities are also called *smooth* points. A *smooth* point of C is a point for which there exists a unique supporting hyperplane to C at it Goberna et al. (2010). Therefore, if $x_0 \in C$, then x_0 exposes some valid inequality of C, if and only if, x_0 is a smooth point of C.

A related concept is that of *blocking points* Basu et al. (2019). However, blocking points need not to be smooth points in general, that is, they do not need to expose any inequality. As seen in Theorem 5.6 we use exposing points to determine maximality of a convex S-free set. Similarly, in the context of lifting Conforti et al. (2011a), blocking points are used to determine maximality of a translated convex cone $S \times \mathbb{Z}_+$-free set. $\square$

There is another phenomenon that does not occur when $S = \mathbb{Z}^n$. If S is a quadratic set, the inequalities of a maximal S-free set might not be exposed by any point of S. For instance, consider $S = \{(x, y) \in \mathbb{R}^2 : x^2 + 1 \leq y^2\}$. The boundary of S is a hyperbola with asymptotes $x = \pm y$. Thus, $C = \{(x, y) \in \mathbb{R}^2 : x \geq |y|\}$ is a maximal S-free set, because its inequalities are asymptotes of S, but they are not exposed by points of S. This phenomenon also occurs when $S = \mathbb{Z}^n \cap K$, with K convex Morán and Dey (2011). However, in that case, it also turns out that maximal S-free sets are polyhedral and their constructions rely on the concept of a facet (see for instance (Morán and Dey, 2011, Theorem 3.2)) which we do not have access to in the general case. In our case, we extend the definition of what it means for an inequality to be exposed in order to handle a situation like the one above. We do this by interpreting that asymptotes are exposed "at infinity".

Definition 5.9. *Given a convex set $C \subseteq \mathbb{R}^n$ with non-empty recession cone and a valid inequality $\alpha^\mathsf{T} x \leq \beta$, we say that a sequence $(x_n)_n \subseteq \mathbb{R}^n$* exposes (α, β) at infinity *with respect to C if*

- $\|x_n\| \to \infty$,
- $\frac{x_n}{\|x_n\|} \to d \in \operatorname{rec}(C)$,
- *d exposes $\alpha^\mathsf{T} x \leq 0$ with respect to* $\operatorname{rec}(C)$, *and*
- *there exists y such that $\alpha^\mathsf{T} y = \beta$ such that* $\operatorname{dist}(x_n, y + \langle d \rangle) \to 0$.

As before, we omit saying "with respect to C" if it is clear from context.

Using this definition, we can prove an analogous result to Theorem 5.5 for inequalities exposed at infinity.

Theorem 5.10. *Let $K, K' \subseteq \mathbb{R}^n$ be convex sets such that $K \subseteq K'$. If $\alpha^\mathsf{T} x \leq \beta$ is*

- *valid for K,*
- *not valid for K', and*
- *exposed at infinity by $(x_n)_n$ with respect to K,*

then there exists a k such that $x_k \in \operatorname{int}(K')$.

Proof. Suppose that for all k, x_k is not in the interior of K'. Then, for each k there exists a non-trivial valid inequality for K', $\gamma_k^\mathsf{T} x \le \delta_k$, such that $\gamma_k^\mathsf{T} x_k \ge \delta_k$. We can assume without loss of generality that $\|(\gamma_k, \delta_k)\| = 1$. Hence, going through a subsequence if necessary, there exist $\gamma \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ such that $\gamma_k \to \gamma$ and $\delta_k \to \delta$ when $k \to \infty$ and $\|(\gamma, \delta)\| = 1$. Note that the inequality (γ, δ) is valid for K'. The idea is to show that (γ, δ) defines the same inequality as (α, β).

As $d = \lim_{k\to\infty} \frac{x_k}{\|x_k\|} \in \operatorname{rec}(K)$ (see Definition 5.9) and (γ, δ) is valid for $K' \supseteq K$, then $\gamma^\mathsf{T} x \le 0$ is valid for $\operatorname{rec}(K)$. In particular, $\gamma^\mathsf{T} d \le 0$. On the other hand,

$$\frac{\delta_k}{\|x_k\|} \le \gamma_k^\mathsf{T} \frac{x_k}{\|x_k\|} \text{ implies } 0 \le \gamma^\mathsf{T} d,$$

We conclude that $\gamma^\mathsf{T} d = 0$. As d exposes $\alpha^\mathsf{T} x \le 0$ with respect to $\operatorname{rec}(K)$, there exists a $\mu \ge 0$ such that $\gamma = \mu\alpha$. Note that we cannot conclude that $\mu > 0$ since, at this point, we do not know that (γ, δ) is a non-trivial inequality (e.g. it could be $0^\mathsf{T} x \le 1$).

Let y be such that $\alpha^\mathsf{T} y = \beta$ and $\operatorname{dist}(x_k, y + \langle d \rangle) \to 0$, which exists by Definition 5.9. Let $w_k = x_k - d^\mathsf{T} x_k d$. We have that

$$\operatorname{dist}(x_k, y + \langle d \rangle) = \operatorname{dist}(x_k - y, \langle d \rangle) = \|x_k - y - d^\mathsf{T}(x_k - y)d\| = \|w_k - (y - d^\mathsf{T} y d)\|.$$

Thus, $w_k \to y - d^\mathsf{T} y d$ as $k \to \infty$.

Since each (γ_k, δ_k) is valid for K', $\gamma_k^\mathsf{T} d \le 0$. Additionally, for large enough k it must hold that $d^\mathsf{T} x_k > 0$. Therefore,

$$\delta_k \le \gamma_k^\mathsf{T} x_k = \gamma_k^\mathsf{T}(d^\mathsf{T} x_k d + w_k) \le \gamma_k^\mathsf{T} w_k.$$

Computing the limit when $k \to \infty$ we get,

$$\delta \le \mu\alpha^\mathsf{T}(y - d^\mathsf{T} y d) = \mu\alpha^\mathsf{T} y = \mu\beta.$$

If $\mu = 0$, then $\gamma = 0$ and $\delta \le 0$. As $\|(\gamma, \delta)\| = 1$, it follows that $\delta = -1$, which cannot be since (γ, δ) is a valid inequality for K' and K' is, by hypothesis, non-empty. We conclude that $\mu > 0$ and that $\mu\alpha^\mathsf{T} x \le \mu\beta$ is valid for K', which implies that $\alpha^\mathsf{T} x \le \beta$ is valid for K', contradicting the hypothesis of the theorem. $\square$

With the previous results it is straightforward to prove the following generalization of Theorem 5.7.

Theorem 5.11. *Let $S \subseteq \mathbb{R}^n$ be a closed set, H be an affine hyperplane, and $C \subseteq \mathbb{R}^n$ be a convex S-free set. Assume that $C = \{x \in \mathbb{R}^n : \alpha^T x \le$*

$\beta, \forall(\alpha,\beta) \in \Gamma\}$ and that for every (α,β) there is, either, an $x \in S \cap C \cap H$ that exposes (α,β), or sequence $(x_n)_n \subseteq S \cap H$ that exposes (α,β) at infinity. Then, C is maximal S-free with respect to H.

Another useful result for studying maximal S-free sets is the following (see also (Conforti et al., 2014, Lemma 6.17)). It states that in some cases we can project S into a lower dimensional space and find maximal sets that are free for the projection. This result is also useful for visualizing higher dimensional S-free sets.

Theorem 5.12. *Let C be a full dimensional closed convex cone with lineality space L. Let $S \subseteq \mathbb{R}^n$ be closed. Then, C is maximal S-free if and only if $(C \cap L^\perp)$ is maximal $\mathrm{cl}(\mathrm{proj}_{L^\perp} S)$-free.*

Proof. ($\Rightarrow$) If $C \cap L^\perp$ is not maximal, let $K \subseteq L^\perp$ be a $\mathrm{cl}(\mathrm{proj}_{L^\perp} S)$-free set that contains it. Then, $K + L \supsetneq C$. Since C is maximal S-free, there exists an $x \in S$ such that $x \in \mathrm{int}(K+L) = \mathrm{int}(K) + \mathrm{int}(L)$ ((Rockafellar, 1970, Corollary 6.6.2)). That is, $x = k + \ell$ with $k \in \mathrm{int}(K)$ and $\ell \in L$. Thus, $x - \ell \in K \subseteq L^\perp$ which implies that $x - \ell \in \mathrm{proj}_{L^\perp} S$ and contradicts the fact that K is $\mathrm{cl}(\mathrm{proj}_{L^\perp} S)$-free.
($\Leftarrow$) By contradiction, suppose that C is not maximal S-free and let $K \supsetneq C$ be a closed convex S-free set. Then $K \cap L^\perp \supsetneq C \cap L^\perp$, which implies that $K \cap L^\perp$ is not $\mathrm{cl}(\mathrm{proj}_{L^\perp} S)$-free. This implies that $\exists \tilde{s} \in \mathrm{cl}(\mathrm{proj}_{L^\perp} S) \cap \mathrm{int}(K \cap L^\perp)$. Moreover, we can further assume $\tilde{s} \in \mathrm{proj}_{L^\perp} S \cap \mathrm{int}(K \cap L^\perp)$, as any sequence contained in $\mathrm{proj}_{L^\perp} S$ converging to an element of $\mathrm{cl}(\mathrm{proj}_{L^\perp} S) \cap \mathrm{int}(K \cap L^\perp)$ must have an element in $\mathrm{proj}_{L^\perp} S \cap \mathrm{int}(K \cap L^\perp)$.

By the definition of orthogonal projection, there must exist $s \in S$ and $\ell \in L$ such that $\tilde{s} = s - \ell$. Thus, we obtain $s - \ell \in \mathrm{int}(K \cap L^\perp)$, i.e.

$$s \in \mathrm{int}(K \cap L^\perp) + L.$$

Since the lineality space of K must contain L, we conclude $s \in \mathrm{int}(K)$; a contradiction with K being S-free. $\square$

5.3 Maximal Quadratic-Free Sets for Homogeneous Quadratics

In this section we construct maximal S^h-free sets that contain a vector $\bar{x} \notin S^h$ for $S^h = \{x \in \mathbb{R}^p : x^\mathsf{T} Q x \leq 0\}$. This is our building block towards maximality in the general case. After a change of variable, we can assume that

$$S^h = \{(x,y,z) \in \mathbb{R}^{n+m+l} \ : \ \sum_{i=i}^{n} x_i^2 - \sum_{i=i}^{m} y_i^2 \leq 0\}$$
$$= \{(x,y) \in \mathbb{R}^{n+m} \ : \ \sum_{i=i}^{n} x_i^2 - \sum_{i=i}^{m} y_i^2 \leq 0\} \times \mathbb{R}^l.$$

Thus, we will only focus on $S^h = \{(x,y) \in \mathbb{R}^{n+m} \ : \ \sum_{i=i}^{n} x_i^2 - \sum_{i=i}^{m} y_i^2 \leq 0\}$ and assume we are given $(\bar{x}, \bar{y})$ such that $\|\bar{x}\|^2 > \|\bar{y}\|^2$.

Remark 5.13. The transformation used to bring S^h to the last "diagonal" form is, in general, not unique. Nonetheless, maximality of the S^h-free sets is preserved, as there always is such transformation that is one-to-one. In Section 5.6 we discuss the effect different choices of this transformation have. □

5.3.1 Removing Strict Convexity Matters

A simple way of obtaining an S^h-free set is via a concave underestimator of $f(x,y) = \sum_{i=i}^{n} x_i^2 - \sum_{i=i}^{m} y_i^2 = \|x\|^2 - \|y\|^2$ directly. A concave underestimator tight at $(\bar{x}, \bar{y})$ is obtained after linearizing the convex function $\|x\|^2$ at $\bar{x}$, that is, $\|\bar{x}\|^2 + 2\|\bar{x}\|(x - \bar{x}) - \|y\|^2$. The concave underestimator yields the S^h-free set $\{(x,y) \in \mathbb{R}^{n+m} \ : \ \|\bar{x}\|^2 + 2\|\bar{x}\|(x - \bar{x}) - \|y\|^2 \geq 0\}$. However, simple examples show that such an S^h-free set is not maximal.

Example 5.14. The case $n = m = 1$ with $\bar{x} = 3$ yields the S^h-free set

$$C = \{(x,y) \in \mathbb{R}^2 \ : \ -9 + 6x - y^2 \geq 0\}$$

In Figure 5.1 we can see that the set is not maximal S^h-free. □

As discussed in Section 4.3.1 the problem is that $\|x\|^2$ is a strictly convex function. Indeed, suppose $S = \{x \in \mathbb{R}^n \ : \ f(x) \leq 0\}$ where f is strictly convex. The S-free set obtained via a concave underestimator at $\bar{x}$ is $C = \{x \in \mathbb{R}^n \ : \ f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) \geq 0\}$. It is not hard to see that the strict convexity of f implies that C is not maximal S-free. The reason is that, as we saw in Chapter 2, linearizations of f at $\bar{x} \notin S$ will not support S. On the other hand, if f is instead sublinear, then any linearization of f supports S, thus it yields a maximal S free set.

The previous observation motivates the following. The set S^h can be equivalently be described by $S^h = \{(x,y) \in \mathbb{R}^{n+m} \ : \ \|x\| - \|y\| \leq 0\}$. Now,

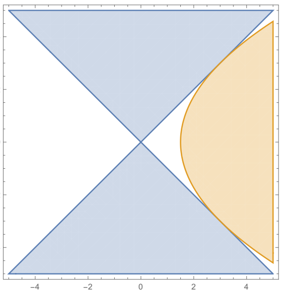


Figure 5.1: S^h in Example 5.14 (blue) and the S^h-free set constructed using a concave underestimator of $\|x\|^2 - \|y\|^2$ (orange).

the function $f(x, y) = \|x\| - \|y\|$ has the following concave underestimator at $\bar{x} \neq 0$, $\frac{\bar{x}^\mathsf{T} x}{\|\bar{x}\|} - \|y\|$, which yields the S^h-free set

$$C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \lambda^\mathsf{T} x \geq \|y\|\}, \tag{5.2}$$

where $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. This set turns out to be maximal, even if we consider any other $\lambda \in D_1(0)$ We note that in Bienstock et al. (2016), the authors use a similar technique and reformulate a 4-variable homogeneous quadratic condition of *outer-product*-free sets in the form $\|x\| \leq \|y\|$. This allows them to construct maximal outer-product-free sets that are of the form C_λ.

5.3.2 Maximal S^h-free Sets

We now prove that C_λ is maximal S^h-free. The main idea is to exploit that every inequality describing C_λ has a point in $S^h \cap C_\lambda$ exposing it and use Theorem 5.6. We begin with a Lemma whose proof we present in Section 5.10. We recall that a function is sublinear if and only if it is convex and positive homogeneous.

Lemma 5.15. *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a sublinear function, $\lambda \in D_1(0)$, and let*

$$C = \{(x, y) \,:\, \phi(y) \leq \lambda^\mathsf{T} x\}.$$

If $(\bar{x}, \bar{y}) \in C$ is such that ϕ is differentiable at $\bar{y}$ and $\phi(\bar{y}) = \lambda^\mathsf{T} \bar{x}$, then $(\bar{x}, \bar{y})$ exposes the valid inequality $-\lambda^\mathsf{T} x + \nabla \phi(\bar{y})^\mathsf{T} y \leq 0$.

In particular, if $\beta_0 \in \partial\phi(0)$ is an exposed point of $\partial\phi(0)$, exposed by $\bar{y}$, and $\phi(\bar{y}) = \lambda^\mathsf{T} \bar{x}$, then $(\bar{x}, \bar{y})$ exposes the valid inequality $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$.

Theorem 5.16. *Let $S^h = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|\}$ and $C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \lambda^\mathsf{T} x \geq \|y\|\}$ for $\lambda \in D_1(0)$. Then, C_λ is a maximal S^h-free set. Furthermore, if $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$, C_λ contains $(\bar{x}, \bar{y})$ in its interior.*

Proof. The S^h-freeness follows by construction. To show that C_λ is maximal, we first notice that

$$C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} \,:\, -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0,\ \forall \beta \in D_1(0)\}.$$

We just need to show that every inequality $(-\lambda, \beta)$ is exposed by a point $(x, y) \in S^h \cap C_\lambda$.

Since the norm function $\|\cdot\|$ is sublinear, differentiable everywhere but in the origin, and $\|\beta\| = 1 = \lambda^\mathsf{T}\lambda$, Lemma 5.15 shows that $(\lambda, \beta) \in S^h \cap C_\lambda$ exposes $(-\lambda, \beta)$. From Theorem 5.6 we conclude that C_λ is maximal S^h-free.

The fact that $(\bar{x}, \bar{y}) \in \operatorname{int}(C_\lambda)$ when $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$, can be verified directly. □

5.4 Homogeneous Quadratics With a Single Homogeneous Linear Constraint

Finding maximal S-free sets for S defined using a non-homogeneous quadratic function is much more challenging than the previous case. In general, using a homogenization and diagonalization, any such S can be described as

$$\{(x, y, z) \in \mathbb{R}^{n+m+l} \,:\, \|x\| \leq \|y\|,\ a^\mathsf{T} x + d^\mathsf{T} y + h^\mathsf{T} z = -1\}. \tag{5.3}$$

Remark 5.17. Similarly to our discussion in Remark 5.13, the choice of transformation to bring a non-homogenous quadratic to the form (5.3) is not unique. Different choices can produce different vectors a, d, h. Nonetheless, maximality of S-free sets is preserved through these transformations if they are one-to-one. We discuss the effect of the different choices of such transformations in Section 5.6. □

First of all, we note that the case $h \neq 0$ can be tackled directly using Section 5.3. Indeed, if this is the case it is not hard to see that $C \times \mathbb{R}^l$ is maximal S-free (with respect to the corresponding hyperplane), where C is any maximal S^h-free. This follows from Theorem 5.12. Thus, in what follows we consider

$$S = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|, a^\mathsf{T} x + d^\mathsf{T} y = -1\}.$$

Also note that using transformations that yield the latter form of S allow us to assume that the given point $(\bar{x}, \bar{y}) \notin S$ satisfies

$$\|\bar{x}\| > \|\bar{y}\|,\ a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} = -1.$$

We elaborate on this point in Section 5.6.

The set S above is our final goal. However, at this point, a simpler set to study is

$$S_{\leq 0} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|, a^\mathsf{T} x + d^\mathsf{T} y \leq 0\}.$$

In this section we construct maximal $S_{\leq 0}$-free sets that contain $(\bar{x}, \bar{y})$ satisfying

$$\|\bar{x}\| > \|\bar{y}\|,\, a^\mathsf{T} \bar{x} + d^\mathsf{T} \bar{y} \leq 0.$$

While this set is interesting on its own, it provides an important intermediate step into our construction of maximal S-free sets.

As it turns out, the construction of maximal $S_{\leq 0}$-free sets depends on whether $\|a\| < \|d\|$ or $\|a\| \geq \|d\|$ and on the value of m. Unfortunately, each case requires different ideas. The following remark dismisses a simple case:

Remark 5.18. If $m = 1$ and $\|a\| < \|d\|$ then $S_{\leq 0}$ is convex. To see this, assume that $d > 0$ and let $(x, y) \in S_{\leq 0}$ with $y \neq 0$. Then, $dy \leq -a^\mathsf{T} x \leq \|a\|\|x\| \leq \|a\||y| < d|y|$. This can only happen if $y < 0$. Therefore, $S_{\leq 0}$ is the second order cone $\{(x, y) \,:\, \|x\| \leq -y\}$. The case $d < 0$ is analogous. We remark that the assumption $\|a\| < |d|$ is fundamental for the argument. As we show in Example 5.25, $S_{\leq 0}$ is not necessarily convex if $\|a\| = |d|$. □

We divide the remaining cases in the following:

Case 1 $\|a\| \leq \|d\| \wedge m > 1$.

Case 2 $\|a\| \geq \|d\|$.

Note that both our strategies allow us to handle the overlapping case $\|a\| = \|d\| \wedge m > 1$. We start with the more natural idea that follows from our previous discussions. This yields the proof of Case 1 and motivates our case distinction.

5.4.1 Case 1: $\|a\| \leq \|d\| \wedge m > 1$

The strategy for proving maximality of C_λ was to write C_λ as

$$C_\lambda = \{(x, y) \in \mathbb{R}^{n+m} \,:\, -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0,\ \forall \beta \in D_1(0)\},$$

and to find an exposing point in $S^h \cap C_\lambda$ for each of the inequalities defining C_λ. As $S_{\leq 0} \subseteq S^h$, C_λ is clearly $S_{\leq 0}$-free. However, if we try to prove it is maximal following the same technique, we find that it is not clear that some inequalities have exposing points in $S_{\leq 0} \cap C_\lambda$. The exposing point of the inequality $(-\lambda, \beta)$, (λ, β) is in $S_{\leq 0}$ if and only if $a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0$. Let

$$G(\lambda) = \{\beta \,:\, \|\beta\| = 1,\ a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0\}.$$

It is natural to ask, then, if

$$C_{G(\lambda)} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0,\ \forall \beta \in G(\lambda)\}$$

is maximal $S_{\leq 0}$-free. Intuitively, $C_{G(\lambda)}$ is obtained from C_λ by removing from its description all inequalities that do not have an exposing point in $a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0$. It is reasonable to expect maximality, as, by construction, every inequality has a point exposing it. Indeed,

Proposition 5.19. *If $C_{G(\lambda)} \neq \emptyset$ and C is any $S_{\leq 0}$-free set such that $C_\lambda \subseteq C$, then $C \subseteq C_{G(\lambda)}$.*

Proof. Suppose, by contradiction, that $C \not\subseteq C_{G(\lambda)}$. This implies that there must exist $\beta_0 \in G(\lambda)$ such that $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$ is not valid for C. As $C_\lambda \subseteq C_{G(\lambda)}$, $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$ is valid for C_λ.

As we saw in Theorem 5.16, $(\lambda, \beta_0) \in C_\lambda$ exposes $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$, and since $C_\lambda \subseteq C$, Theorem 5.5 implies that $(\lambda, \beta_0) \in \operatorname{int}(C)$. However, since $\beta_0 \in G(\lambda)$, we have $(\lambda, \beta_0) \in S_{\leq 0}$. This contradicts the $S_{\leq 0}$-freeness of C. □

This result shows that $C_{G(\lambda)}$ is the largest (inclusion-wise) set that one can aspire to obtain from C_λ. However, it is unclear if $C_{G(\lambda)}$ is $S_{\leq 0}$-free. Even more, it is unclear whether $G(\lambda)$ is non-empty or not. In the following we study when $C_{G(\lambda)}$ is $S_{\leq 0}$-free

We start by showing that when $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$, $G(\lambda)$ is non-empty.

Proposition 5.20. *Let $(\bar{x}, \bar{y}) \notin S_{\leq 0}$ such that $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} \leq 0$ and let $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. Then,*

$$G(\lambda) \neq \emptyset.$$

If, in addition, $d = 0$, then $G(\lambda) = D_1(0)$ and $C_{G(\lambda)} = C_\lambda$ is maximal $S_{\leq 0}$-free.

Proof. As $(\bar{x}, \bar{y}) \notin S_{\leq 0}$, we have that $\|\bar{y}\| < \|\bar{x}\|$. Since $m > 1$, then we can find $z \in \mathbb{R}^m \setminus \{0\}$ such that $d^\mathsf{T} z = 0$ and $\|\frac{\bar{y}}{\|\bar{x}\|} + z\| = 1$. Also, $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} \leq 0$ and $d^\mathsf{T} z = 0$ imply that $a^\mathsf{T}\lambda + d^\mathsf{T}(\frac{\bar{y}}{\|\bar{x}\|} + z) \leq 0$. Thus, $\frac{\bar{y}}{\|\bar{x}\|} + z \in G(\lambda)$.

Regarding the second statement of the proposition, if $d = 0$ then clearly either $G(\lambda) = D_1(0)$ or $G(\lambda) = \emptyset$. Since we are in the case $G(\lambda) \neq \emptyset$, this immediately implies $C_{G(\lambda)} = C_\lambda$. Thus, Proposition 5.19 implies its maximality. $\square$

In light of Proposition 5.19, we just need for $C_{G(\lambda)}$ to be $S_{\leq 0}$-free for it to be maximal. Note that

$$C_{G(\lambda)} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \max_{\beta \in G(\lambda)} y^\mathsf{T}\beta \leq \lambda^\mathsf{T} x\}, \tag{5.4}$$

and so to prove $S_{\leq 0}$-freeness, it is enough to show that for every $(x, y) \in S_{\leq 0}$, $\max_{\beta \in G(\lambda)} y^\mathsf{T}\beta \geq \lambda^\mathsf{T} x$. In trying to prove this inequality is where the conditions of this case naturally arise.

Proposition 5.21. *Let $(\bar{x}, \bar{y}) \notin S_{\leq 0}$ such that $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} \leq 0$ and $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. If $\|d\| \geq \|a\|$ and $m > 1$, then $C_{G(\lambda)}$ is maximal $S_{\leq 0}$-free and contains $(\bar{x}, \bar{y})$ in its interior.*

Proof. As discussed above, it is enough to show that

$$\max_{\beta \in G(\lambda)} y^\mathsf{T}\beta \geq \lambda^\mathsf{T} x \text{ for every } (x, y) \in S_{\leq 0}. \tag{5.5}$$

Informally, the strategy is to find a dual of $\max_{\beta \in G(\lambda)} y^\mathsf{T}\beta$ so that the inequality we have to prove is of the form "minimum of something greater or equal than $\lambda^\mathsf{T} x$", which often times is easier to reason about. As the objective function of $\max_{\beta \in G(\lambda)} y^\mathsf{T}\beta$ is linear and $m > 1$, we can replace the $\|\beta\| = 1$ constraint with an inequality and obtain

$$\max_{\beta \in G(\lambda)} y^\mathsf{T}\beta = \max\{y^\mathsf{T}\beta \,:\, \|\beta\| \leq 1, a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0\}. \tag{5.6}$$

As $G(\lambda)$ is constructed from an infeasible point $(\bar{x}, \bar{y}) \notin S_{\leq 0}$ such that $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} \leq 0$, i.e., $\|\bar{y}\| < \|\bar{x}\|$, we have $\|\bar{y}/\|\bar{x}\|\| < 1$. Moreover, perturbing the latter we can argue that the rightmost optimization problem in (5.6) has a strictly feasible point. Thus, Slater's condition holds and we have that

$$\max\{y^\mathsf{T}\beta \,:\, \|\beta\| \leq 1, a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0\} = \inf_{\theta \geq 0} \|y - d\theta\| - \lambda^\mathsf{T} a\theta. \tag{5.7}$$

Using (5.7), (5.5) is equivalent to

$$\inf_{\theta \geq 0} \|y - d\theta\| - \lambda^\mathsf{T} a\theta \geq \lambda^\mathsf{T} x \text{ for every } (x, y) \in S_{\leq 0}. \tag{5.8}$$

We now prove that if $(x, y) \in S_{\leq 0}$, then $\lambda^\mathsf{T}(x + a\theta) \leq \|y - d\theta\|$, which implies the result.

By Cauchy-Schwarz and $\|\lambda\| = 1$, we have that $\lambda^\mathsf{T}(x + a\theta) \leq \|x + a\theta\|$. Furthermore, $\|x + a\theta\|^2 = \|x\|^2 + 2\theta a^T x + \|a\theta\|^2$. Since $\theta \geq 0$, $\theta a^\mathsf{T} x \leq -\theta d^\mathsf{T} y$. Together with $\|x\|^2 \leq \|y\|^2$ they imply

$$\begin{aligned}\|x + a\theta\|^2 &\leq \|y\|^2 - 2\theta d^\mathsf{T} y + \|a\|^2\theta^2 \\ &= \|y - d\theta\|^2 + (\|a\|^2 - \|d\|^2)\theta^2 \\ &\leq \|y - d\theta\|^2,\end{aligned}$$

where the last inequality follows since $\|d\| \geq \|a\|$.

We have shown that $\|x+a\theta\| \leq \|y-d\theta\|$. Hence, $\lambda^\mathsf{T}(x+a\theta) \leq \|y-d\theta\|$ as we wanted to show, which implies that $C_{G(\lambda)}$ is $S_{\leq 0}$-free. Finally, Proposition 5.19 implies the maximality of $C_{G(\lambda)}$, and $(\bar{x}, \bar{y}) \in \operatorname{int}(C_{G(\lambda)})$ since $C_\lambda \subseteq C_{G(\lambda)}$. $\square$

Remark 5.22. Using Proposition 5.54 one can show that $\max_\beta\{y^\mathsf{T}\beta \,:\, \|\beta\| \leq 1, a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0\}$ is

$$\begin{cases} \|y\|, & \text{if } a^\mathsf{T}\lambda\|y\| + y^\mathsf{T} d \leq 0 \\ \sqrt{(1 - (\frac{a^\mathsf{T}\lambda}{\|d\|})^2)(\|y\|^2 - (\frac{y^\mathsf{T} d}{\|d\|^2})^2)} - \frac{a^\mathsf{T}\lambda y^\mathsf{T} d}{\|d\|^2}, & \text{otherwise.} \end{cases} \tag{5.9}$$

Note that this is well defined since if $\|d\| = 0$, then $\|a\| = 0$ and so (5.9) $= \|y\|$. This yields a closed-form expression for $C_{G(\lambda)}$ of the form

$$C_{G(\lambda)} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, (5.9) \leq \lambda^\mathsf{T} x\}. \tag{5.10}$$

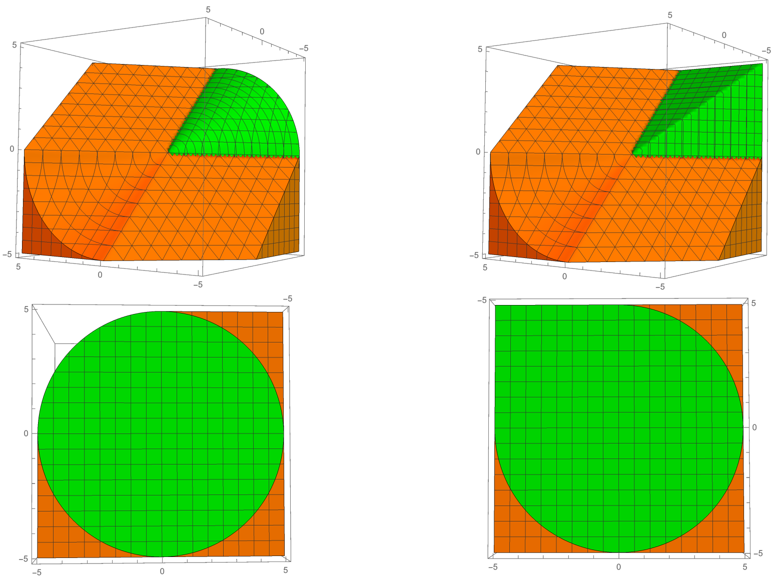
$\square$

The last proposition provides certain guarantees of when a simple modification of C_λ yields maximal $S_{\leq 0}$-free sets. Our proof heavily relies on our assumptions $\|a\| \leq \|d\|$ (to show (5.8)) and $m > 1$ (to show (5.6)), so the natural question is whether these conditions are actually necessary for our statement to be true. Thus, before moving on to the next case, we argue why these conditions are indeed necessary in our statements. The following examples motivate our case distinction and illustrate all cases we have covered.

Example 5.23. Consider the following set of the type $S_{\leq 0}$, which we denote $S^1_{\leq 0}$:

$$S^1_{\leq 0} = \{(x, y_1, y_2) \in \mathbb{R}^3 \,:\, |x| \leq \|y\|, \quad ax + d^\mathsf{T} y \leq 0\}$$

with $a = 1$ and $d = (1, -1)^\mathsf{T}$. Let us consider the point $(\bar{x}, \bar{y}) = (-1, 0, 0)^\mathsf{T}$, clearly satisfying the linear inequality, but not in $S^1_{\leq 0}$. In Figure 5.2 we show $S^1_{\leq 0}$, the $S^1_{\leq 0}$-free set given by C_λ and the set $C_{G(\lambda)}$ for $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. Since in this case $|a| = 1 \leq \sqrt{2} = \|d\|$ and $m > 1$, we know $C_{G(\lambda)}$ is maximal $S^1_{\leq 0}$-free. $\square$



(a) $S^1_{\leq 0}$ in Example 5.23 (orange) and the corresponding C_λ set (green). The latter is $S^1_{\leq 0}$-free but not maximal.

(b) $S^1_{\leq 0}$ in Example 5.23 (orange) and the corresponding $C_{G(\lambda)}$ set (green). The latter is maximal $S^1_{\leq 0}$-free.

Figure 5.2: Sets C_λ and $C_{G(\lambda)}$ in Example 5.23 for the case $\|a\| \leq \|d\|$.

Example 5.24. Consider the set $S^2_{\leq 0}$, defined as

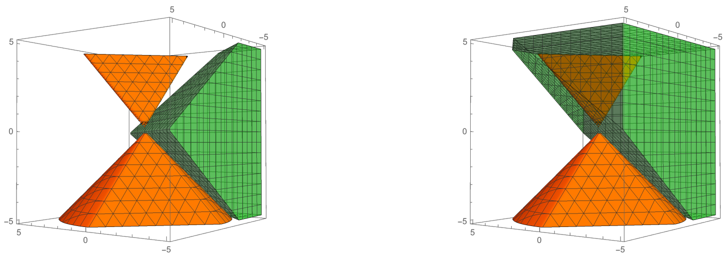
$$S^2_{\leq 0} = \{(x_1, x_2, y) \in \mathbb{R}^3 \,:\, \|x\| \leq |y|, \quad a^\mathsf{T} x + dy \leq 0\}$$

with $a = (-1/\sqrt{2}, 1/\sqrt{2})^\mathsf{T}$ and $d = 1/\sqrt{2}$ (the $1/\sqrt{2}$ terms are not really important now as we can scale the inequality, but we reuse this example in subsequent sections where they do matter), and $(\bar{x}, \bar{y}) = (-1, -1, 0)^\mathsf{T}$. This point satisfies the linear inequality in $S^2_{\leq 0}$, but it is not in $S^2_{\leq 0}$. Let $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$.

In this case $a^\mathsf{T}\lambda = 0$, and as a consequence the corresponding set $G(\lambda)$ is given by the singleton $\{-1\}$. In Figure 5.3 we show $S^2_{\leq 0}$, the $S^2_{\leq 0}$-free set given by C_λ and the set $C_{G(\lambda)}$. In this case $\|a\| = 1 > 1/\sqrt{2} = |d|$, so we have no guarantee on the $S^2_{\leq 0}$-freeness of $C_{G(\lambda)}$. Even more, it is not $S^2_{\leq 0}$-free. $\square$

Example 5.25. Let us consider the following example with $n = 2$, $m = 1$ and $\|d\| = \|a\|$. Let $a = (-3, 4)^\mathsf{T}, d = 5$ and consider $(\bar{x}, \bar{y}) = (-4, -3, -1)$ and $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. Clearly $(\bar{x}, \bar{y}) \notin S_{\leq 0}$, but satisfies the linear constraint. In this case, $\beta \in G(\lambda)$ must satisfy

$$5 \cdot \beta \leq 0, |\beta| = 1$$



(a) $S^2_{\leq 0}$ in Example 5.24 (orange) and the corresponding C_λ set (green). The latter is $S^2_{\leq 0}$-free but not maximal.

(b) $S^2_{\leq 0}$ in Example 5.24 (orange) and the corresponding $C_{G(\lambda)}$ set (green). The latter is not $S^2_{\leq 0}$-free.

Figure 5.3: Sets C_λ and $C_{G(\lambda)}$ in Example 5.24 for the case $\|a\| > \|d\|$.

thus $G(\lambda) = \{-1\}$. Nonetheless, $(x, y) = (3, -4, 5) \in S_{\leq 0}$, and

$$\lambda^\mathsf{T} x + y = 0 + 5 > 0$$

This means $(x, y) \in \operatorname{int}(C_{G(\lambda)})$. Thus, $C_{G(\lambda)}$ is not $S_{\leq 0}$-free. □

Remark 5.26. The situation in Example 5.25 is similar to the one depicted in Figure 5.3b. Roughly speaking, when $\|a\| = \|d\|$ the upper region becomes a single line and this line intersects the interior of $C_{G(\lambda)}$. Intuitively, when we consider S where $a^\mathsf{T} x + d^\mathsf{T} y = -1$, this line should not appear. Even more, S should be convex. We will see that this is the case in the Section 5.5.1. □

5.4.2 Case 2: $\|a\| \geq \|d\|$

As we have seen in Example 5.24, when $\|a\| \leq \|d\|$ does not hold, $C_{G(\lambda)}$ is not necessarily $S_{\leq 0}$-free. On the other hand, C_λ is $S_{\leq 0}$-free but not necessarily maximal. As before, we are looking for a convex set C that is maximal $S_{\leq 0}$-free set that *contains* C_λ. We point out that in not all statements of this section we require $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$.

Projecting-out the lineality space

The lineality space of C_λ is $L = \{(x, y) \,:\, \lambda^\mathsf{T} x = 0, y = 0\}$ and as $C_\lambda \subseteq C$, it must be that L is contained in the lineality space of C. By Theorem 5.12, $\operatorname{proj}_{L^\perp} C$ is maximal $\operatorname{proj}_{L^\perp} S_{\leq 0}$-free, thus, it might be possible (and we show it is) to find C by studying maximal $\operatorname{proj}_{L^\perp} S_{\leq 0}$-free sets. We note that $L^\perp = \langle \lambda \rangle \times \mathbb{R}^m$ and

$$\operatorname{proj}_{L^\perp} S_{\leq 0} = \{(\lambda^\mathsf{T} x, y) \,:\, \|x\| \leq \|y\|,\; a^\mathsf{T} x + d^\mathsf{T} y \leq 0\}.$$

After analyzing low dimensional instances of $\operatorname{proj}_{L^\perp} S_{\leq 0}$ we conjecture that $(\operatorname{proj}_{L^\perp} S_{\leq 0})^c$ is formed by the union of two disjoint convex sets. If this is true, it would directly provide maximal $\operatorname{proj}_{L^\perp} S_{\leq 0}$-free sets.

In order to show that this is actually true, we follow the following strategy. For each point $y \in \mathbb{R}^m$, the points $(\lambda^T x, y) \in \operatorname{proj}_{L^\perp} S_{\leq 0}$ lie on an interval, namely, $\{\lambda^\mathsf{T} x \,:\, \|x\| \leq \|y\|,\ a^\mathsf{T} x + d^\mathsf{T} y \leq 0\}$. Thus, we define the functions

$$\begin{aligned}
&y \mapsto \max\{\lambda^\mathsf{T} x \,:\, \|x\| \leq \|y\|, a^\mathsf{T} x + d^\mathsf{T} y \leq 0\} \text{ and}\\
&y \mapsto \min\{\lambda^\mathsf{T} x \,:\, \|x\| \leq \|y\|, a^\mathsf{T} x + d^\mathsf{T} y \leq 0\}.
\end{aligned}$$

If the first function is convex and the second is concave, then the closure of $(\operatorname{proj}_{L^\perp} S_{\leq 0})^c$ is the union of the epigraph of the first one and the hypograph of the second one. Thus, it suffices to show that

$$\phi_\lambda(y) = \max_x\{\lambda^\mathsf{T} x \,:\, \|x\| \leq \|y\|, a^\mathsf{T} x + d^\mathsf{T} y \leq 0\} \tag{5.11}$$

is convex for every $\lambda \in D_1(0)$, as the second function is $-\phi_{-\lambda}$.

We first show that ϕ_λ is defined over all $\mathbb{R}^m$.

Proposition 5.27. *If $\|d\| \leq \|a\|$, then for every y the set $\{(x, y) \,:\, \|x\| \leq \|y\|, a^\mathsf{T} x \leq -d^\mathsf{T} y\}$ is not empty.*

Proof. Note that $x = -d^\mathsf{T} y \frac{a}{\|a\|^2}$ belongs to the set. Indeed, $a^\mathsf{T} x = -d^\mathsf{T} y$, in particular, $a^\mathsf{T} x \leq -d^\mathsf{T} y$. Also, $\|d\| \leq \|a\|$ implies that $\|x\| \leq \frac{\|d\|}{\|a\|}\|y\| \leq \|y\|$. $\square$

We now show that ϕ_λ is convex. Furthermore, we prove that ϕ_λ is sublinear, that is, convex and positive homogeneous. The proof is basically to find ϕ_λ explicitly and then verify its properties. Note that in this case $\|a\| = 0$ implies that the linear inequality in $S_{\leq 0}$ is trivial. Thus,we assume without loss of generality, that $\|a\| = 1$.

Proposition 5.28. *Let $\lambda, a \in D_1(0) \subseteq \mathbb{R}^n$ and $d \in \mathbb{R}^m$ such that $\|d\| \leq 1$. Then,*

$$\phi_\lambda(y) = \begin{cases} \|y\|, & \textit{if } \lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \leq 0 \\ \sqrt{(\|y\|^2 - (d^\mathsf{T} y)^2)(1 - (\lambda^\mathsf{T} a)^2)} - d^\mathsf{T} y \lambda^\mathsf{T} a, & \textit{otherwise.} \end{cases} \tag{5.12}$$

Furthermore, ϕ_λ is sublinear and

- *if $\|d\| = 1 \wedge m > 1$, then ϕ_λ is differentiable $\mathbb{R}^m \setminus d\mathbb{R}_+$,*

– *otherwise* ϕ_λ *is differentiable in* $\mathbb{R}^m \setminus \{0\}$.

Proof. The fact that ϕ_λ is positive homogeneous can be easily verified. We leave the proof that ϕ_λ is of the form (5.12) to Section 5.10, see Proposition 5.54. Thus convexity and differentiability remains.

First, note that if $\lambda = a$, then $\phi_\lambda(y) = -d^\mathsf{T} y$. This function is clearly sublinear and differentiable everywhere. On the other hand, if $\lambda = -a$, then $\phi_\lambda(y) = \|y\|$. This function is clearly sublinear and differentiable everywhere but the origin.

We now consider $\lambda \neq \pm a$. Let

$$\begin{aligned} A_1 &= \{y \,:\, \lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \leq 0\}, \\ A_2 &= \{y \,:\, \lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \geq 0\}, \end{aligned} \tag{5.13}$$

and let ϕ_λ^1 and ϕ_λ^2 be the restriction of ϕ_λ to A_1 and A_2, respectively.

To show that ϕ_λ is convex we are going to use (Solovev, 1983, Theorem 3). In our particular case, since ϕ_λ is positively homogeneous, this theorem implies that we just need to check that ϕ_λ is convex on each convex subset of A_1 and A_2, $\phi_\lambda^1 = \phi_\lambda^2$ on $A_1 \cap A_2$, and that

$$\phi_\lambda'(y;\rho) + \phi_\lambda'(y;-\rho) \geq 0, \text{ for all } \rho \in \mathbb{R}^m \setminus \{0\}, y \in A_1 \cap A_2. \tag{5.14}$$

Here, $\phi_\lambda'(y;\rho)$ is the directional derivative of ϕ_λ at y in the direction of ρ.

Clearly, ϕ_λ is convex in each convex subset of A_1. The function ϕ_λ^2 is of the form $c_1\|y\|_W - c_2 d^\mathsf{T} y$, where $W = I - dd^\mathsf{T} \succeq 0$ and c_1, c_2 are constants. Thus, ϕ_λ is convex on each convex subset of A_2.

It is not hard to see that $\phi_\lambda^1(y) = \phi_\lambda^2(y)$ for $y \in A_1 \cap A_2$.

Let us verify (5.14) for $y \neq 0$. For this, first notice that $\phi_\lambda^1(y)$ is differentiable whenever $y \neq 0$. Likewise, $\phi_\lambda^2(y)$ is differentiable whenever $y \neq 0$ if $\|d\| < 1$ or whenever $y \notin d\mathbb{R}_+$ if $\|d\| = 1$. However, if $y \in A_1 \cap A_2 \setminus \{0\}$ and $\|d\| = 1$, then $y \notin d\mathbb{R}_+$, thus ϕ_λ^2 is differentiable in a neighborhood of y. Furthermore,

$$\begin{aligned} \nabla\phi_\lambda^2(y) &= \frac{(1-(\lambda^\mathsf{T} a)^2)(I - dd^\mathsf{T})y}{\sqrt{(\|y\|^2 - (d^\mathsf{T} y)^2)(1-(\lambda^\mathsf{T} a)^2)}} - \lambda^\mathsf{T} ad \\ &= \frac{1}{\|y\|}(I - dd^\mathsf{T})y - \lambda^\mathsf{T} ad \\ &= \frac{y}{\|y\|} \\ &= \nabla\phi_\lambda^1(y). \end{aligned}$$

Therefore, ϕ_λ is differentiable in whenever $y \neq 0$ if $\|d\| < 1$ or whenever $y \notin d\mathbb{R}_+$ if $\|d\| = 1$. Thus, (5.14) holds with equality for $y \in A_1 \cap A_2 \setminus \{0\}$.

It remains to verify (5.14) for $y = 0$. Let ρ be such that $\rho \in A_1$ and $-\rho \in A_2$. As ϕ_λ is positively homogeneous, $\phi'_\lambda(0; \cdot) = \phi_\lambda(\cdot)$. Hence,

$$\phi'_\lambda(0;\rho) = \|\rho\| \text{ and } \phi'_\lambda(0;-\rho) = \sqrt{1-(\lambda^T a)^2}\sqrt{\|\rho\|^2 - (d^\mathsf{T}\rho)^2} + d^\mathsf{T}\rho\lambda^T a.$$

We need to prove that

$$\sqrt{1-(\lambda^T a)^2}\sqrt{\|\rho\|^2 - (d^\mathsf{T}\rho)^2} + d^\mathsf{T}\rho\lambda^T a + \|\rho\| \geq 0.$$

By Cauchy-Schwarz, $|d^\mathsf{T}\rho\lambda^T a| \leq \|d\|\|\rho\| < \|\rho\|$. Thus, $d^\mathsf{T}\rho\lambda^T a + \|\rho\| > 0$. Since $\sqrt{1-(\lambda^T a)^2}\sqrt{\|\rho\|^2 - (d^\mathsf{T}\rho)^2} \geq 0$, the inequality follows. Therefore, ϕ_λ is convex.

We have proved that ϕ_λ is convex and differentiable in $\mathbb{R}^m \setminus \{0\}$ if $\|d\| < 1$ and in $\mathbb{R}^m \setminus d\mathbb{R}_+$ if $\|d\| = 1$. It remains to show that if $m = 1$ and $\|d\| = 1$, then ϕ_λ is differentiable in $\mathbb{R}^m \setminus \{0\}$. This follows from (5.12) since $\phi^2_\lambda(y) = -dy\lambda^\mathsf{T} a$ in this case. This concludes the proof. $\square$

With this, we have completed the proof of sublinearity of ϕ_λ. Moreover, we have explicitly described the function. As a corollary:

Corollary 5.29. *The epigraph of ϕ_λ and the hypograph of $-\phi_{-\lambda}$ are maximal* $\operatorname{proj}_{L^\perp} S_{\leq 0}$*-free sets.*

While this result provides two convex sets, it is not clear which one to chose. This means, which of these two constructed $\operatorname{proj}_{L^\perp} S_{\leq 0}$-free sets will yield an $S_{\leq 0}$-free containing the given solution $(\bar{x}, \bar{y})$. We answer this next.

Lemma 5.30. *Consider $(\bar{x}, \bar{y})$ such that $\|\bar{x}\| > \|\bar{y}\|$ and $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} \leq 0$ and $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. Then, the projection of $(\bar{x}, \bar{y})$ onto $L^\perp$ is in the interior of the epigraph of ϕ_λ.*

Proof. The projection of $(\bar{x}, \bar{y})$ onto $L^\perp$ is given by $(\lambda^\mathsf{T}\bar{x}, \bar{y})$. Then, $\phi_\lambda(\bar{y}) = \max_x\{\lambda^\mathsf{T} x \,:\, \|x\| \leq \|\bar{y}\|, a^\mathsf{T} x + d^\mathsf{T}\bar{y} \leq 0\} \leq \lambda^\mathsf{T}\lambda\|\bar{y}\| = \|\bar{y}\|$. Thus, $\lambda^\mathsf{T}\bar{x} = \|\bar{x}\| > \|\bar{y}\| \geq \phi_\lambda(\bar{y})$. $\square$

Back to the original space

Finally, we use the above to construct $S_{\leq 0}$-free sets, i.e., in the original space. Embedded in $\mathbb{R}^{n+m}$, the epigraph of ϕ_λ is $\{(t\lambda, y) \,:\, y \in \mathbb{R}^m,\ \phi_\lambda(y) \leq t\}$. Thus,

$$\begin{aligned} C_{\phi_\lambda} &= \{(t\lambda, y) \,:\, y \in \mathbb{R}^m,\ \phi_\lambda(y) \leq t\} + L \\ &= \{(t\lambda + z, y) \,:\, y \in \mathbb{R}^m,\ \lambda^T z = 0,\ \phi_\lambda(y) \leq t\} \\ &= \{(x, y) \,:\, \phi_\lambda(y) \leq \lambda^\mathsf{T} x\}. \end{aligned} \tag{5.15}$$

As a summary we prove that C_{ϕ_λ} is maximal $S_{\leq 0}$-free without going through the projection.

Proposition 5.31. *Let $\lambda \in D_1(0)$ and $\phi_\lambda(y) = \max_x\{\lambda^\mathsf{T} x \,:\, (x, y) \in S_{\leq 0}\}$. If $\|a\| = 1 \geq \|d\|$, then $C_{\phi_\lambda} = \{(x, y) \,:\, \phi_\lambda(y) \leq \lambda^\mathsf{T} x\}$ is maximal $S_{\leq 0}$-free.*

Additionally, if $(\bar{x}, \bar{y}) \notin S_{\leq 0}$ is such that $a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} \leq 0$, letting $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ ensures $(\bar{x}, \bar{y}) \in \operatorname{int}(C_{\phi_\lambda})$.

Proof. We will prove that C_{ϕ_λ} is convex, free and maximal.

The convexity of C_{ϕ_λ} follows directly from *Proposition* 5.28. Also, C_{ϕ_λ} is $S_{\leq 0}$-free since if $(x, y) \in S_{\leq 0}$, then $\phi_\lambda(y) \geq \lambda^\mathsf{T} x$. Therefore, (x, y) is not in the interior of C_{ϕ_λ}.

We now focus on proving maximality. In the cases where ϕ_λ is differentiable in $\mathbb{R}^m \setminus \{0\}$ we can directly write

$$C_{\phi_\lambda} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq \lambda^\mathsf{T} x,\ \forall \beta \in D_1(0)\}.$$

Let $\beta \in D_1(0)$ and let x_β be the optimal solution of the problem (5.11) which defines $\phi_\lambda(\beta)$. That is, $\lambda^\mathsf{T} x_\beta = \phi_\lambda(\beta)$. By Lemma 5.15, the inequality $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq 0$ is exposed by (x_β, β).

The only remaining case is $\|d\| = 1 \wedge m > 1$, where ϕ_λ is only differentiable in $D_1(0)\setminus\{d\}$. Since in this case $m > 1$ we can safely remove a single inequality from the outer-description of C_{ϕ_λ} without affecting it, i.e.,

$$C_{\phi_\lambda} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq \lambda^\mathsf{T} x,\ \forall \beta \in D_1(0) \setminus \{d\}\}.$$

Using the same argument as above we can find an exposing point of each inequality $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq 0$ for $\beta \in D_1(0) \setminus \{d\}$.

The fact that $(\bar{x}, \bar{y}) \in \operatorname{int}(C_{\phi_\lambda})$ when $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ follows directly since $C_\lambda \subseteq C_{\phi_\lambda}$.

□

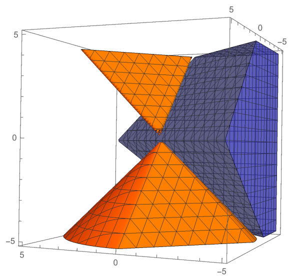


Figure 5.4: $S^2_{\leq 0}$ in Example 5.24 (orange) and C_{ϕ_λ} set (blue). The latter is maximal $S^2_{\leq 0}$-free.

Example 5.32. Let us recall the set $S^2_{\leq 0}$ in Example 5.24.

$$S^2_{\leq 0} = \{(x_1, x_2, y) \in \mathbb{R}^3 \,:\, \|x\| \leq |y|, \quad a^\mathsf{T} x + dy \leq 0\}$$

with $a = (-1/\sqrt{2}, 1/\sqrt{2})^\mathsf{T}$, $d = 1/\sqrt{2}$, and $(\bar{x}, \bar{y}) = (-1, -1, 0)^\mathsf{T}$. In Figure 5.3 we showed that the set C_λ is $S^2_{\leq 0}$-free but not maximal, and $C_{G(\lambda)}$ is not $S^2_{\leq 0}$-free. In Figure 5.4 we show the set C_{ϕ_λ}, which is maximal $S^2_{\leq 0}$-free. For this example, we know $\lambda^\mathsf{T} a = 0$, thus

$$\lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \leq 0 \iff y \leq 0.$$

A simple calculation using (5.12) yields

$$\phi_\lambda(y) = \begin{cases} -y, & \text{if } y \leq 0 \\ \frac{y}{\sqrt{2}} & \text{if } y > 0 \end{cases}$$

$\square$

Remark 5.33. As we saw in the proof of Proposition 5.28 if $\lambda = a$, then $\phi_\lambda(y) = -d^\mathsf{T} y$. This implies that $C_{\phi_\lambda} = \{(x, y) \,:\, a^\mathsf{T} x + d^\mathsf{T} y \geq 0\}$. By definition, this set does not contain any point from $a^\mathsf{T} x + d^\mathsf{T} y \leq 0$ in its interior, thus, it is a very uninteresting maximal $S_{\leq 0}$-free set. One is usually interested in constructing a maximal $S_{\leq 0}$-free set that contain a point $(\bar{x}, \bar{y})$ that satisfies $a^\mathsf{T} x + d^\mathsf{T} y \leq 0$. Hence, by Lemma 5.30, whenever we assume that $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ where $a^\mathsf{T} \bar{x} + d^\mathsf{T} \bar{y} \leq 0$ and $\|\bar{x}\| > \|\bar{y}\|$, it will automatically hold that $\lambda \neq a$. $\square$

Remark 5.34. At this point we would like to show some relations between C_λ, C_{ϕ_λ} and $C_{G(\lambda)}$. The inequalities defining C_λ are $(-\lambda, \beta)$ for $\beta \in D_1(0)$.

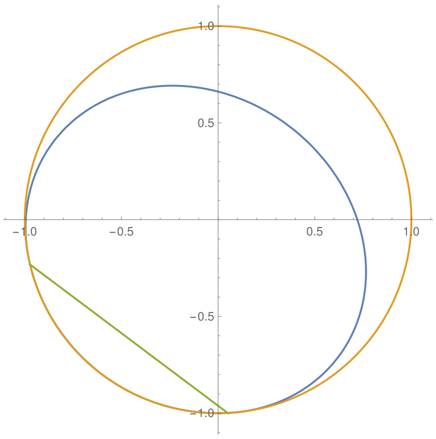


Figure 5.5: Let $a = (\frac{3}{5}, -\frac{4}{5}), d = (\frac{3}{10}, \frac{2}{5})$, and $\lambda = (\frac{63}{65}, \frac{16}{65})$. The boundary of the y coordinates of the polars of C_λ, $C_{G(\lambda)}$, and C_{ϕ_λ} are depicted in orange, green, and blue, respectively. They all coincide below the green line.

Equivalently, the polar of C_λ is the cone generated by $\{-\lambda\} \times \operatorname{conv} D_1(0) = \{-\lambda\} \times B_1(0)$.

The inequalities defining $C_{G(\lambda)}$ are $(-\lambda, \beta)$ for $\beta \in G(\lambda) = \{\beta \in D_1(0) : \beta\lambda^\mathsf{T} a + d^\mathsf{T}\beta \leq 0\}$. Equivalently, the polar of $C_{G(\lambda)}$ is the cone generated by $\{-\lambda\} \times \operatorname{conv} G(\lambda)$.

The inequalities defining C_{ϕ_λ} are $(-\lambda, \nabla\phi_\lambda(\beta))$ for $\beta \in D_1(0)$. When $\beta \in G(\lambda)$, then $\phi_\lambda(y) = \|y\|$ and so the inequalities are $(-\lambda, \beta)$. In other words, some inequalities defining C_{ϕ_λ} coincide with the inequalities defining $C_{G(\lambda)}$ and C_λ. Thus, when C_{ϕ_λ} is convex (i.e., when $\|a\| \geq \|d\|$), there is a region where all three convex sets look the same. In terms of the polars, when $\|a\| \geq \|d\|$, the polar of C_{ϕ_λ} is between the polars of $C_{G(\lambda)}$ and C_λ. This is depicted in Figure 5.5. $\square$

5.5 Non-Homogeneous Quadratics

As discussed at the beginning of the previous section, we now study a general non-homogeneous quadratic which can be written as

$$S = \{(x, y) \in \mathbb{R}^{n+m} : \|x\| \leq \|y\|,\ a^\mathsf{T} x + d^\mathsf{T} y = -1\}.$$

We assume we are given $(\bar{x}, \bar{y})$ such that

$$\|\bar{x}\| > \|\bar{y}\|,\ a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} = -1.$$

Much like in Section 5.4, we begin by dismissing a simple case.

Remark 5.35. The case $\|a\| \leq \|d\| \wedge m = 1$ can be treated separately. Note that, as opposed to the analogous analysis at the beginning of Section 5.4, here we include the case where the norms are equal. As already noted in Remark 5.26, we should expect S to be convex in this case. Indeed, as $d \neq 0$ (if not, then $a = 0$ and $S = \emptyset$) we can write $y = \frac{1}{d}(-1 - a^\mathsf{T} x)$ and consequently

$$\begin{aligned}
S &= \{(x, y) \in \mathbb{R}^{n+1} \,:\, \|x\|^2 \leq \frac{1}{d^2}(1 + 2a^\mathsf{T} x + (a^\mathsf{T} x)^2), a^\mathsf{T} x + d^\mathsf{T} y = -1\} \\
&= \{(x, y) \in \mathbb{R}^{n+1} \,:\, x^\mathsf{T} \left(I - \frac{1}{d^2} aa^\mathsf{T}\right) x - \frac{1}{d^2}(1 + 2a^\mathsf{T} x) \leq 0, a^\mathsf{T} x + d^\mathsf{T} y = -1\}.
\end{aligned}$$

Since $I - \frac{1}{d^2} aa^\mathsf{T}$ is positive semi-definite whenever $|d| \geq \|a\|$, the set S is convex. Thus, a maximal S-free set, or even directly a cutting plane, can be obtained using a supporting hyperplane. □

Similarly to Section 5.4, we distinguish the following cases:

Case 1 $\|a\| \leq \|d\| \wedge m > 1$.

Case 2 $\|a\| > \|d\|$.

Since $S \subsetneq S_{\leq 0}$, then $C_{G(\lambda)}$ (C_{ϕ_λ}) is S-free in Case 1 (Case 2) as per Section 5.4. It is natural to wonder whether these sets are maximal already.

5.5.1 Case 1: $\|a\| \leq \|d\| \wedge m > 1$

The technique we used to prove maximality of $C_{G(\lambda)}$ with respect to $S_{\leq 0}$ is to exploit that $C_{G(\lambda)}$ is defined by the inequalities of C_λ exposed by elements in $S_{\leq 0}$. Following this approach, we study which inequalities of $C_{G(\lambda)}$ are exposed by a point of S. Recall that

$$C_{G(\lambda)} = \{(x, y) \in \mathbb{R}^{n+m} \,:\, -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0,\ \forall \beta \in G(\lambda)\},$$

where

$$G(\lambda) = \{\beta \in \mathbb{R}^m \,:\, \|\beta\| = 1,\ a^\mathsf{T} \lambda + d^\mathsf{T} \beta \leq 0\}.$$

Consider an inequality in the definition of $C_{G(\lambda)}$ given by $(-\lambda, \beta)$ such that $a^\mathsf{T} \lambda + d^\mathsf{T} \beta < 0$. Then, the point $(\lambda, \beta) \in S_{\leq 0}$ can be scaled by $\mu = \frac{-1}{a^\mathsf{T} \lambda + d^\mathsf{T} \beta}$ to the exposing point $\mu(\lambda, \beta) \in S$. Thus, almost every inequality describing $C_{G(\lambda)}$ is exposed by points of S. Furthermore, we can simply remove the inequalities that are not exposed by points of S from $C_{G(\lambda)}$ without changing the set $C_{G(\lambda)}$. We specify this next.

Theorem 5.36. *Let* $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$,

$$H = \{(x,y) \in \mathbb{R}^{n+m} \,:\, a^\mathsf{T} x + d^\mathsf{T} y = -1\}$$

and

$$S_{\leq 0} = \{(x,y)\mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|,\ a^\mathsf{T} x + d^\mathsf{T} y \leq 0\},$$

where $\|a\| \leq \|d\| \wedge m > 1$*. Then,* $C_{G(\lambda)}$ *is maximal* $S_{\leq 0}$*-free with respect to* H *and contains* $(\bar{x}, \bar{y})$ *in its interior.*

Proof. By Proposition 5.21, we know that $C_{G(\lambda)}$ is maximal $S_{\leq 0}$-free. Thus, $C_{G(\lambda)}$ is $S_{\leq 0}$-free with respect to H. To prove maximality, we note that thanks to $m > 1$:

$$C_{G(\lambda)} = \{(x,y) \in \mathbb{R}^{n+m} \,:\, -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0,\ \forall \beta \in \operatorname{ri}(G(\lambda))\},$$

where

$$\operatorname{ri}(G(\lambda)) = \{\beta \in \mathbb{R}^m \,:\, \|\beta\| = 1, a^\mathsf{T}\lambda + d^\mathsf{T}\beta < 0\}$$

is the relative interior of $G(\lambda)$. Consider $\beta_0 \in \operatorname{ri}(G(\lambda))$. As we saw in Proposition 5.19, $(\lambda, \beta_0) \in C_{G(\lambda)} \cap S_{\leq 0}$ exposes the inequality $(-\lambda, \beta_0)$. As $C_{G(\lambda)} \cap S_{\leq 0}$ is a (non-convex) cone, we have that for any $\mu > 0$, $\mu(\lambda, \beta_0) \in C_{G(\lambda)} \cap S_{\leq 0}$ exposes the inequality $(-\lambda, \beta_0)$. Since $a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0 < 0$, $\mu = -\frac{1}{a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0} > 0$ and so

$$-\frac{(\lambda, \beta_0)}{a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0} \in S_{\leq 0} \cap H \cap C_{G(\lambda)}, \tag{5.16}$$

exposes the inequality $(-\lambda, \beta_0)$. The claim now follows from Theorem 5.7. □

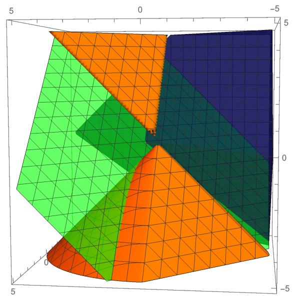
The above theorem states that obtaining a maximal S-free set in this case amounts to simply using the maximal $S_{\leq 0}$-free set $C_{G(\lambda)}$, and then intersecting with H. Recall that $S = S_{\leq 0} \cap H$. The next case is considerably different.

5.5.2 Case 2: $\|a\| > \|d\|$

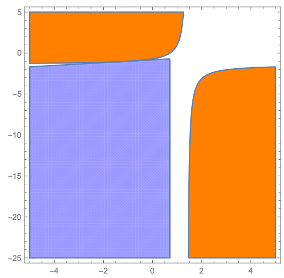
We begin with an important remark regarding an assumption made in the analogous case of the previous section.

Remark 5.37. Since in this case $\|a\| > 0$, we can, again, assume that $\|a\| = 1$. Indeed, we can always rescale the variables (x, y) by $\|a\|$ to obtain such requirement.

Also note that since $\|d\| < \|a\| = 1$, then ϕ_λ is differentiable in $D_1(0)$. See Proposition 5.28. □



(a) $S^2_{\leq 0}$ (orange), H (green) and C_{ϕ_λ} (blue).



(b) Projection onto (x_1, x_2) of $S^2_{\leq 0} \cap H$ (orange) and $C_{\phi_\lambda} \cap H$ (blue). One of the facets of $C_{\phi_\lambda} \cap H$ has a gap with the boundary of $S^2_{\leq 0} \cap H$.

Figure 5.6: Plots of $S^2_{\leq 0}$, H and C_{ϕ_λ} as defined in Example 5.38 showing that C_{ϕ_λ} is not necessarily maximal $S^2_{\leq 0}$-free with respect to H in the case $\|a\| > \|d\|$.

Unfortunately, in this case the maximality of C_{ϕ_λ} with respect to $S_{\leq 0}$ does not carry over to S, as the following example shows.

Example 5.38. We continue with $S^2_{\leq 0}$ defined in Example 5.24. In Figure 5.4 we showed how C_{ϕ_λ} gives us a maximal $S^2_{\leq 0}$-free set. If we now consider

$$H = \{(x, y) \in \mathbb{R}^{n+m} \,:\, a^\mathsf{T} x + d^\mathsf{T} y = -1\}$$

with $a = (-1/\sqrt{2}, 1/\sqrt{2})^\mathsf{T}$ and $d = 1/\sqrt{2}$, we do not necessarily obtain that $C_{\phi_\lambda} \cap H$ is maximal $S^2_{\leq 0} \cap H$-free. In Figure 5.6 we illustrate this issue.

□

Figure 5.6 of the previous example displays an interesting feature though: the inequalities defining C_{ϕ_λ} seem to have the correct "slope" and just need to be translated. We conjecture, then, that in order to find a maximal S-free set, we only need to adequately relax the inequalities of C_{ϕ_λ}.

Set-up

Recall that

$$\begin{aligned} C_{\phi_\lambda} &= \{(x, y) \,:\, \phi_\lambda(y) \leq \lambda^\mathsf{T} x\} \\ &= \{(x, y) \,:\, -\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq 0, \forall \beta \in D_1(0)\}. \end{aligned}$$

We denote by $r(\beta)$ the amount by which we need to relax each inequality of C_{ϕ_λ} such that

$$C = \{(x, y) \,:\, -\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq r(\beta), \forall \beta \in D_1(0)\}, \tag{5.17}$$

is S-free. Note that when β satisfies $\lambda^\mathsf{T} a + d^\mathsf{T}\beta < 0$, the inequalities of C_{ϕ_λ} are the same as the ones of $C_{G(\lambda)}$ (see also Remark 5.34) and, just like in Section 5.5.1, they have exposing points in S. An inequality of this type can be seen in Figure 5.6b: it is the inequality of C_{ϕ_λ} tangent to S at one of its exposing points. Thus, we expect that $r(\beta) = 0$ when $\lambda^\mathsf{T} a + d^\mathsf{T}\beta < 0$. In the following we find $r(\beta)$ when $\lambda^\mathsf{T} a + d^\mathsf{T}\beta \geq 0$ and show maximality of the resulting set.

Following the spirit of Section 5.4.2, not all statement in this section require $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. However, we assume $\lambda \neq \pm a$. This assumption, however, is not restrictive when constructing maximal S-free sets, as the following remark shows.

Remark 5.39. If $\lambda = -a$, then for every $\beta \in D_1(0)$ it holds that $\lambda^\mathsf{T} a + d^\mathsf{T}\beta < 0$. In this case $r(\beta)$ will be simply defined as 0 everywhere and $C = C_{\phi_\lambda}$. This means all inequalities defining C have an exposing point in S and maximality follows directly.

On the other hand, if we take $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ with $(\bar{x}, \bar{y}) \in H$ and $\|\bar{x}\| > \|\bar{y}\|$, we have that if additionally $\lambda = a$

$$\begin{aligned} a^\mathsf{T}\bar{x} + d^\mathsf{T}\bar{y} = -1 &\iff \|\bar{x}\| + d^\mathsf{T}\bar{y} = -1 \\ &\implies \|\bar{y}\| + d^\mathsf{T}\bar{y} < -1. \end{aligned}$$

The latter cannot be, as $\|d\| < 1$. □

Remark 5.40. The assumption $\lambda \neq \pm a$ has an unexpected consequence: as $\lambda \neq \pm a$ and $\|a\| = \|\lambda\| = 1$, it must hold that $n \geq 2$. This implicit assumption, however, does not present an issue: whenever $n = 1$ either $\lambda = a$ or $\lambda = -a$. By Remark 5.39, if we use $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$, then $\lambda = -a$. Thus, $C = C_{\phi_\lambda}$ and maximality holds. □

Construction of $r(\beta)$

Let $\beta \in D_1(0)$ be such that $\lambda^\mathsf{T} a + d^\mathsf{T}\beta \geq 0$. Then, the face of C_{ϕ_λ} defined by the valid inequality $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq 0$ does not intersect S. See Lemma 5.55 for a proof of this statement.

In particular, the inequality is not exposed by any point in $S \cap C_{\phi_\lambda}$. However, it is exposed by $(x_\beta, \beta) \in S_{\leq 0}$, where x_β is given by (5.27) (see the proof of Proposition 5.31). Note that $(x_\beta, \beta) \in H_0 = \{(x, y) : a^\mathsf{T} x + d^\mathsf{T} y = 0\}$, as otherwise we can scale it so that it belongs to S.

The quantity $r(\beta)$ is the amount we need to relax the inequality in order to be an "asymptote", and we compute it as follows. We first find a sequence of points, $(x_n, y_n)_{n\in\mathbb{N}}$, in $S_{\leq 0}$ that converge to (x_β, β), enforcing that no element of the sequence belongs to H_0. If we find such sequence, then every $(x_n, y_n) \in S_{\leq 0}$ can be scaled to be in S:

$$z_n = -\frac{(x_n, y_n)}{a^\mathsf{T} x_n + d^\mathsf{T} y_n} \in S.$$

This last scaled sequence diverges, as the denominator goes to 0 due to $(x_n, y_n) \to (x_\beta, \beta) \in H_0$. The idea is that the violation $(-\lambda, \nabla\phi_\lambda(\beta))^\mathsf{T} z_n$ given by this sequence will give us, in the limit, the maximum relaxation that will ensure S-freeness (see Figure 5.7). Then, we would define

$$r(\beta) = \lim_{n\to\infty} (-\lambda, \nabla\phi_\lambda(\beta))^\mathsf{T} z_n = -\lim_{n\to\infty} \frac{-\lambda^\mathsf{T} x_n + \nabla\phi_\lambda(\beta)^\mathsf{T} y_n}{a^\mathsf{T} x_n + d^\mathsf{T} y_n}.$$

We remark that this limit is what we intuitively aim for, but it might not even be well defined in general. In what follows, we construct a sequence that yields a closed-form expression for the above limit. Additionally, we show that such definition of $r(\beta)$ yields the desired maximal S-free set.

The sequence. Our goal is to find a sequence $(x_n, y_n)_n$ such that $(x_n, y_n) \in S_{\leq 0}$, $a^\mathsf{T} x_n + d^\mathsf{T} y_n < 0$ and $(x_n, y_n) \to (x_\beta, \beta)$. We take $y_n = \beta$ and x_n such that $\|x_n\| = \|\beta\| = 1$, $a^\mathsf{T} x_n + d^\mathsf{T} \beta < 0$ and $x_n \to x_\beta$. Note that these always exists as $\|a\| = 1$ and $\|d\| < 1$. We illustrate such a sequence with our running example.

Example 5.41. We continue with Example 5.38. As we mentioned in Example 5.32, in this case

$$\phi_\lambda(y) = \begin{cases} -y, & \text{if } y \leq 0 \\ \frac{y}{\sqrt{2}} & \text{if } y > 0 \end{cases}$$

and since $\lambda = \frac{1}{\sqrt{2}}(-1, -1)^\mathsf{T}$, we see that

$$\begin{aligned} C_{\phi_\lambda} = \{(x, y) : \frac{1}{\sqrt{2}}(x_1 + x_2) - y \leq 0, \quad &\text{(5.18a)} \\ \frac{1}{\sqrt{2}}(x_1 + x_2) + \frac{1}{\sqrt{2}} y \leq 0\}. \quad &\text{(5.18b)} \end{aligned}$$

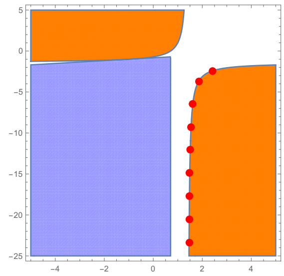


Figure 5.7: Projection onto (x_1, x_2) of $S^2_{\leq 0} \cap H$ (orange) and C_{ϕ_λ} (blue), along with the first two coordinates of the sequence $(z_n)_{n\in\mathbb{N}}$ defined in Example 5.41 for several values of n (red). The sequence is diverging "downwards".

It is not hard to check that $-(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \sqrt{2}) \in S^2_{\leq 0} \cap H \cap C_{\phi_\lambda}$ exposes inequality (5.18a). This is the tangent point in Figure 5.6b we discussed above.

On the other hand, (5.18b), which is obtained from $\beta = 1$, does not have an exposing point in $S^2_{\leq 0} \cap H \cap C_{\phi_\lambda}$, and corresponds to an inequality we should relax as per our discussion. This inequality, however, is exposed by $(x_\beta, \beta) = (0, -1, 1) \in S^2_{\leq 0} \cap C_{\phi_\lambda}$. Consider now the sequence defined as

$$(x_n, y_n) = \left(\frac{1}{\sqrt{n^2+1}}, -\frac{n}{\sqrt{n^2+1}}, 1 \right) \in S^2_{\leq 0}.$$

Clearly the limit of this sequence is $(0, -1, 1)$ and

$$a^\mathsf{T} x_n + d^\mathsf{T} y_n = \frac{1}{\sqrt{2}} \left(-\frac{1}{\sqrt{n^2+1}} - \frac{n}{\sqrt{n^2+1}} + 1 \right) < 0.$$

Now we let

$$z_n = -\frac{(x_n, y_n)}{a^\mathsf{T} x_n + d^\mathsf{T} y_n} \in S^2_{\leq 0} \cap H.$$

As we mention above, this sequence diverges. Continuing with Figure 5.6, in Figure 5.7, we plot the first two components of the sequence $(z_n)_{n\in\mathbb{N}}$ along with $S^2_{\leq 0} \cap H$ and $C_{\phi_\lambda} \cap H$. From this figure we can anticipate where our argument is going: the sequence $(z_n)_{n\in\mathbb{N}}$ moves along the boundary of $S^2_{\leq 0} \cap H$ towards an "asymptote" from where we can deduce $r(\beta)$. The latter is given by the gap between inequality (5.18b) and the asymptote. □

Computing the limit. Here we compute

$$r(\beta) = -\lim_{n\to\infty} \frac{-\lambda^\mathsf{T} x_n + \nabla\phi_\lambda(\beta)^\mathsf{T} y_n}{a^\mathsf{T} x_n + d^\mathsf{T} y_n}.$$

We proceed to rewrite the limit.

Since $y_n = \beta$ and x_β is the optimal solution of (5.11), we have:

$$\begin{aligned}\nabla\phi_\lambda(\beta)^\mathsf{T} y_n &= \phi_\lambda(\beta) = \lambda^\mathsf{T} x_\beta\\ d^\mathsf{T} y_n &= -a^\mathsf{T} x_\beta.\end{aligned}$$

Thus,

$$\begin{aligned}r(\beta) &= -\lim_{n\to\infty} \frac{-\lambda^\mathsf{T} x_n + \nabla\phi_\lambda(\beta)^\mathsf{T} y_n}{a^\mathsf{T} x_n + d^\mathsf{T} y_n}\\ &= -\lim_{n\to\infty} \frac{-\lambda^\mathsf{T} x_n + \lambda^\mathsf{T} x_\beta}{a^\mathsf{T} x_n - a^\mathsf{T} x_\beta}\\ &= \lim_{n\to\infty} \frac{\lambda^\mathsf{T}(x_n - x_\beta)}{a^\mathsf{T}(x_n - x_\beta)}.\end{aligned}$$

Notice that x_β belongs to the 2 dimensional space generated by λ and a, which we denote by Λ. Note that it is indeed 2 dimensional, since $\lambda \neq \pm a$, see Remark 5.39. Furthermore, we can assume that x_n also belongs to Λ as any other component of x_n is irrelevant for the value of the limit. Indeed, as $\mathbb{R}^n = \Lambda \oplus \Lambda^\perp$, then $x_n = x_n^\parallel + x_n^\perp$, where $x_n^\parallel \in \Lambda$ and $x_n^\perp \in \Lambda^\perp$, and

$$\frac{\lambda^\mathsf{T}(x_n - x_\beta)}{a^\mathsf{T}(x_n - x_\beta)} = \frac{\lambda^\mathsf{T}(x_n^\parallel - x_\beta)}{a^\mathsf{T}(x_n^\parallel - x_\beta)}.$$

To compute the limit observe that

$$\frac{\lambda^\mathsf{T}(x_n - x_\beta)}{a^\mathsf{T}(x_n - x_\beta)} = \frac{\lambda^\mathsf{T} \frac{x_n - x_\beta}{\|x_n - x_\beta\|}}{a^\mathsf{T} \frac{x_n - x_\beta}{\|x_n - x_\beta\|}}.$$

Notice that $\frac{x_n - x_\beta}{\|x_n - x_\beta\|}$ converges, as $x_n \in \Lambda$, $\|x_n\| = 1$, and $x_n \to x_\beta$. Let $\hat{x}$ be the limit and note that $\hat{x}$ is orthogonal to x_β. Indeed,

$$\begin{aligned}{x_\beta}^\mathsf{T}\hat{x} &= \lim_{n\to\infty} {x_\beta}^\mathsf{T} \frac{x_n - x_\beta}{\|x_n - x_\beta\|}\\ &= \lim_{n\to\infty} \frac{{x_\beta}^\mathsf{T} x_n - 1}{\|x_n - x_\beta\|}\\ &= \lim_{n\to\infty} -\frac{\|x_n - x_\beta\|^2}{2\|x_n - x_\beta\|}\\ &= 0.\end{aligned}$$

Hence,

$$r(\beta) = \lim_{n\to\infty} \frac{\lambda^\mathsf{T}(x_n - x_\beta)}{a^\mathsf{T}(x_n - x_\beta)} = \frac{\lambda^\mathsf{T}\hat{x}}{a^\mathsf{T}\hat{x}}.$$

Since we are interested in the quotient of $\lambda^\mathsf{T}\hat{x}$ and $a^\mathsf{T}\hat{x}$, any multiple of $\hat{x}$ can be used, that is, any vector orthogonal to x_β in Λ. Using λ and a as basis for Λ, we have that for $x \in \Lambda$ with coordinates x_λ and x_a, the vector y with coordinates $y_\lambda = -(x_a + x_\lambda\lambda^\mathsf{T}a)$ and $y_a = x_\lambda + x_a\lambda^\mathsf{T}a$ is orthogonal to x. Indeed,

$$\begin{aligned}
x^\mathsf{T}y &= (x_\lambda\lambda + x_a a)^\mathsf{T}(y_\lambda\lambda + y_a a)\\
&= x_\lambda y_\lambda + x_a y_a + (x_\lambda y_a + x_a y_\lambda)\lambda^\mathsf{T}a\\
&= (x_\lambda + x_a\lambda^\mathsf{T}a)y_\lambda + (x_a + x_\lambda\lambda^\mathsf{T}a)y_a\\
&= 0.
\end{aligned}$$

Thus, let $\tilde{x} = -(x_{\beta_a} + x_{\beta_\lambda}\lambda^\mathsf{T}a)\lambda + (x_{\beta_\lambda} + x_{\beta_a}\lambda^\mathsf{T}a)a$. Given that $\lambda^\mathsf{T}a + d^\mathsf{T}\beta \geq 0$, from (5.27) (see Section 5.10) we have

$$x_\beta = \sqrt{\frac{1-(d^\mathsf{T}\beta)^2}{1-(\lambda^T a)^2}}\lambda - \left(d^\mathsf{T}\beta + \lambda^T a\sqrt{\frac{1-(d^\mathsf{T}\beta)^2}{1-(\lambda^T a)^2}}\right)a. \tag{5.19}$$

Note that while this last explicit formula for x_β is the one stated for the case $\lambda^\mathsf{T}a + d^\mathsf{T}\beta > 0$, it also holds when $\lambda^\mathsf{T}a + d^\mathsf{T}\beta = 0$. Therefore,

$$\begin{aligned}
\tilde{x} &= (d^\mathsf{T}\beta)\lambda + \left(\sqrt{\frac{1-(d^\mathsf{T}\beta)^2}{1-(\lambda^T a)^2}} - \left(d^\mathsf{T}\beta + \lambda^T a\sqrt{\frac{1-(d^\mathsf{T}\beta)^2}{1-(\lambda^T a)^2}}\right)\lambda^\mathsf{T}a\right)a\\
&= (d^\mathsf{T}\beta)\lambda + \phi_\lambda(\beta)a.
\end{aligned}$$

All together, we obtain

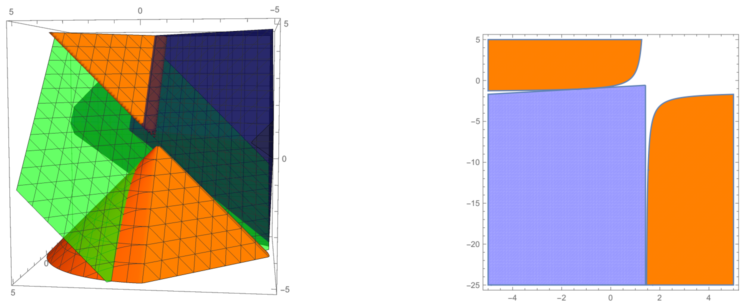
$$r(\beta) = \frac{\lambda^\mathsf{T}\tilde{x}}{a^\mathsf{T}\tilde{x}} = \frac{d^\mathsf{T}\beta + \lambda^\mathsf{T}a\phi_\lambda(\beta)}{\phi_\lambda(\beta) + d^\mathsf{T}\beta\lambda^\mathsf{T}a}.$$

Note that if $\lambda^\mathsf{T}a + d^\mathsf{T}\beta = 0$, then $r(\beta) = 0$. We summarize the above discussion in the following result.

Lemma 5.42. *Let $a, \lambda, \beta \in D_1(0)$, $d \in B_1(0)$, and $\lambda \neq \pm a$ be such that $\|d\| < \|a\|$ and $\lambda^\mathsf{T}a + d^\mathsf{T}\beta \geq 0$. Then, every sequence $(x_n)_{n\in\mathbb{N}} \subseteq \langle\lambda, a\rangle$ converging to x_β such that $\|x_n\| = 1$ and $a^\mathsf{T}x_n + d^\mathsf{T}\beta < 0$, satisfies*

$$r(\beta) = \lim_{n\to\infty} \frac{\lambda^\mathsf{T}(x_n - x_\beta)}{a^\mathsf{T}(x_n - x_\beta)} = \frac{d^\mathsf{T}\beta + \lambda^\mathsf{T}a\phi_\lambda(\beta)}{\phi_\lambda(\beta) + d^\mathsf{T}\beta\lambda^\mathsf{T}a}.$$

Such sequences are always guaranteed to exist.



(a) $S^2_{\leq 0}$ (orange), H (green) and C_1 (blue). In this case C_1 is no longer $S^2_{\leq 0}$-free.

(b) Projection onto (x_1, x_2) of $S^2_{\leq 0} \cap H$ (orange) and $C_1 \cap H$ (blue).

Figure 5.8: Plots of $S^2_{\leq 0}$, H and C_1 as defined in Example 5.43 showing that C_1 is maximal $S^2_{\leq 0}$-free with respect to H.

Therefore, for $\beta \in D_1(0)$, we define

$$r(\beta) = \begin{cases} 0, & \text{if } \lambda^\mathsf{T} a + d^\mathsf{T}\beta \leq 0 \\ \frac{d^\mathsf{T}\beta + \lambda^\mathsf{T} a \phi_\lambda(\beta)}{\phi_\lambda(\beta) + d^\mathsf{T}\beta\lambda^\mathsf{T} a}, & \text{otherwise.} \end{cases}$$

We extend r to $y \in \mathbb{R}^m \setminus \{0\}$ by $r(y) = r(\frac{y}{\|y\|})$ and leave it undefined at 0.

Example 5.43. We continue with our running example in Example 5.41. In this case $r(-1) = 0$, and since $\phi_\lambda(\beta) = 1/\sqrt{2}$, $\lambda^\mathsf{T} a = 0$ and $d = 1/\sqrt{2}$ it can be checked that

$$r(1) = 1.$$

Now, let

$$\begin{aligned} C_1 &= \{(x, y) \, : \, -\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq r(\beta), \text{ for all } \beta \in D_1(0)\} \\ &= \{(x, y) \, : \, \frac{1}{\sqrt{2}}(x_1 + x_2) - y \leq 0, \frac{1}{\sqrt{2}}(x_1 + x_2) + \frac{1}{\sqrt{2}} y \leq 1\}. \end{aligned}$$

Figure 5.8 shows the same plots as Figure 5.6 with C_1 instead of C_{ϕ_λ}.

□

As we see below, the characterization of r as a limit is going to be useful to prove maximality of C. However, to show that C is free, we need a different interpretation of r.

Lemma 5.44. *For every $\beta \in D_1(0)$, $r(\beta) = \theta(\beta)$, where $\theta(\beta)$ is defined in* (5.28) *and corresponds to the optimal dual solution of the optimization problem defining $\phi_\lambda(\beta)$.*

Proof. If $\lambda^\mathsf{T} a + d^\mathsf{T}\beta \le 0$, $r(\beta) = 0 = \theta(\beta)$. Let $\beta \in D_1(0)$ be such that $\lambda^\mathsf{T} a + d^\mathsf{T}\beta > 0$. Then,

$$\begin{aligned} r(\beta) &= \frac{d^\mathsf{T}\beta + \lambda^\mathsf{T} a\phi_\lambda(\beta)}{\phi_\lambda(\beta) + d^\mathsf{T}\beta\lambda^\mathsf{T} a} \\ &= \frac{d^\mathsf{T}\beta + \lambda^\mathsf{T} a\sqrt{1-(\lambda^\mathsf{T} a)^2}\sqrt{1-(d^\mathsf{T}\beta)^2} - d^\mathsf{T}\beta(\lambda^\mathsf{T} a)^2}{\sqrt{1-(\lambda^\mathsf{T} a)^2}\sqrt{1-(d^\mathsf{T}\beta)^2}} \\ &= \frac{d^\mathsf{T}\beta\sqrt{1-(\lambda^\mathsf{T} a)^2}}{\sqrt{1-(d^\mathsf{T}\beta)^2}} + \lambda^\mathsf{T} a \\ &= \theta(\beta). \end{aligned}$$

□

S-freeness and maximality proofs

We now show that C is S-free and then that it is maximal.

Theorem 5.45. *Let $\lambda \in D_1(0)$ such that $\lambda \neq \pm a$,*

$$C = \{(x, y) : -\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \le r(\beta),\ \textit{for all}\ \beta, \|\beta\| = 1\}.$$

and $S = \{(x, y) : \|x\| \le \|y\|, a^\mathsf{T} x + d^\mathsf{T} y = -1\}$, with $\|d\| < \|a\| = 1$. Then, C is S-free.

Proof. Let $(x_0, y_0) \in S$ and let $\beta_0 = \frac{y_0}{\|y_0\|}$. The claim will follow if we are able to show that $-\lambda^\mathsf{T} x_0 + \nabla\phi_\lambda(\beta_0)^\mathsf{T} y_0 \ge r(\beta_0)$.

Since x_0 satisfies $\|x_0\| \le \|y_0\|$ and $a^\mathsf{T} x_0 + d^\mathsf{T} y_0 = -1$, it follows that

$$\lambda^\mathsf{T} x_0 \le \max_x\{\lambda^\mathsf{T} x : \|x\| \le \|y_0\|, a^\mathsf{T} x + d^\mathsf{T} y_0 \le -1\}.$$

By weak duality we have

$$\max_x\{\lambda^\mathsf{T} x : \|x\| \le \|y_0\|, a^\mathsf{T} x + d^\mathsf{T} y_0 \le -1\} \le \inf_{\theta\ge 0} \|y_0\|\|\lambda - a\theta\| - (d^\mathsf{T} y_0 + 1)\theta.$$

Recall that $\theta(y_0)$ is the optimal dual solution to the optimization problem defining $\phi_\lambda(y_0)$. Thus, it holds that $\theta(y_0) \in \mathbb{R}_+$ and $\theta(y_0) < +\infty$ because $\|d\| < 1$. Consequently,

$$\inf_{\theta\ge 0} \|y_0\|\|\lambda - a\theta\| - (d^\mathsf{T} y_0 + 1)\theta \le \|y_0\|\|\lambda - a\theta(y_0)\| - (d^\mathsf{T} y_0 + 1)\theta(y_0) = \phi_\lambda(y_0) - \theta(y_0),$$

where the last equality follows from the strong duality between the optimization problem that defines ϕ_λ and its dual, see Proposition 5.54. All the inequalities together show that

$$\lambda^\mathsf{T} x_0 \leq \phi_\lambda(y_0) - \theta(y_0).$$

From (5.28) and Lemma 5.44 it follow $\theta(y_0) = \theta(\beta_0) = r(\beta_0)$. Thus,

$$-\lambda^\mathsf{T} x_0 + \phi_\lambda(y_0) \geq r(\beta_0),$$

as we wanted to establish. □

Theorem 5.46. *Let $\lambda \in D_1(0)$ such that $\lambda \neq \pm a$,*

$$H = \{(x, y) \in \mathbb{R}^{n+m} \,:\, a^\mathsf{T} x + d^\mathsf{T} y = -1\},$$

$$S_{\leq 0} = \{(x, y)\mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|,\ a^\mathsf{T} x + d^\mathsf{T} y \leq 0\},$$

and

$$C = \{(x, y) \,:\, -\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq r(\beta),\ \textit{for all } \beta \in D_1(0)\}.$$

where $\|d\| < \|a\| = 1$. Then, C is maximal $S_{\leq 0}$-free with respect to H.

Additionally, if $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$ with $(\bar{x}, \bar{y}) \in H$ and $\|\bar{x}\| > \|\bar{y}\|$, then $(\bar{x}, \bar{y}) \in \operatorname{int}(C)$.

Proof. Let $S = S_{\leq 0} \cap H$. By Theorem 5.45, C is S-free.

To show maximality we will use Theorem 5.11, that is, we will show that every inequality of C is either exposed by a point in $S \cap C$ or exposed at infinity by a sequence in S.

Let $\beta_0 \in D_1(0)$ and consider the valid inequality $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta_0)^\mathsf{T} y \leq r(\beta_0)$. Assume, first, that $a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0 < 0$ As $a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0 < 0$, we have that $r(\beta_0) = 0$, $\phi_\lambda(\beta_0) = \|\beta_0\| = 1$, and $\nabla\phi_\lambda(\beta_0) = \beta_0$. Hence, the inequality is $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$. It is exposed by

$$\frac{-1}{a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0}(\lambda, \beta_0) \in S \cap C_{\phi_\lambda} \subseteq S \cap C.$$

Now, let us assume that $a^\mathsf{T}\lambda + d^\mathsf{T}\beta_0 \geq 0$. We will show that there is a sequence in S that exposes $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta_0)^\mathsf{T} y \leq r(\beta_0)$ at infinity. Let $(x_n)_n \subseteq \langle\lambda, a\rangle$ be a sequence converging to x_{β_0} such that $\|x_n\| = 1$, $a^\mathsf{T} x_n + d^\mathsf{T}\beta_0 < 0$ (Lemma 5.42).

$$r(\beta_0) = \lim_{n\to\infty} \frac{\lambda^\mathsf{T}(x_n - x_{\beta_0})}{a^\mathsf{T}(x_n - x_{\beta_0})}.$$

Consider the sequence conformed by

$$z_n = -\frac{(x_n, \beta_0)}{a^\mathsf{T} x_n + d^\mathsf{T} \beta_0} = \frac{(x_n, \beta_0)}{a^\mathsf{T}(x_{\beta_0} - x_n)} \in S,$$

where the equality above follows from $a^\mathsf{T} x_{\beta_0} + d^\mathsf{T} \beta_0 = 0$. We proceed to verify that z_n exposes $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta_0)^\mathsf{T} y \leq r(\beta_0)$ at infinity.

As $x_n \to x_{\beta_0}$, we have that $\|z_n\| \to \infty$. Also, $\frac{z_n}{\|z_n\|} = \frac{1}{\sqrt{2}}(x_n, \beta_0)$ converges to $v = \frac{1}{\sqrt{2}}(x_{\beta_0}, \beta_0) \in C_{\phi_\lambda} = \operatorname{rec}(C)$ and exposes $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta_0)^\mathsf{T} y \leq 0$.

Finally, we have to show that there exists a w such that $(-\lambda, \nabla\phi_\lambda(\beta_0))^\mathsf{T} w = r(\beta_0)$ and $\operatorname{dist}(z_n, w + \langle v \rangle) \to 0$. Let $\hat{x} = \lim_{n\to\infty} \frac{x_n - x_{\beta_0}}{\|x_n - x_{\beta_0}\|}$ and let $w = (-\frac{\hat{x}}{a^\mathsf{T}\hat{x}}, 0)$. We have that $(-\lambda, \nabla\phi_\lambda(\beta_0))^\mathsf{T} w = r(\beta_0)$. Also,

$$z_n - \frac{\sqrt{2}}{a^\mathsf{T}(x_{\beta_0} - x_n)} v = \frac{1}{a^\mathsf{T}(x_{\beta_0} - x_n)}(x_n - x_{\beta_0}, 0) \to -(\frac{\hat{x}}{a^\mathsf{T}\hat{x}}, 0) = w.$$

Thus, $\operatorname{dist}(z_n, w + \langle v \rangle) \to 0$. □

A closed-form formula for C

Since the construction of C involves translating some of the inequalities of C_{ϕ_λ} of its outer-description, it is natural to ask if this translation yields a translation of the whole function ϕ_λ. This would yield a closed-form formula for C which is much more appealing from a computational standpoint.

In what follows, we ask whether there exists an (x_0, y_0) such that for every β such that

$$\begin{aligned} &\{(x,y) \,:\, -\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq r(\beta), \text{ for all } \beta, \lambda^\mathsf{T} a + d^\mathsf{T}\beta \geq 0\} \\ =&\{(x,y) \,:\, -\lambda^\mathsf{T}(x - x_0) + \nabla\phi_\lambda(\beta)^\mathsf{T}(y - y_0) \leq 0, \text{ for all } \beta, \lambda^\mathsf{T} a + d^\mathsf{T}\beta \geq 0\}. \end{aligned}$$

In order to reach this equality it would suffice to satisfy

$$\lambda^\mathsf{T} x_0 - \nabla\phi_\lambda(\beta)^\mathsf{T} y_0 = -r(\beta). \tag{5.20}$$

Note that since $\lambda^\mathsf{T} a + d^\mathsf{T}\beta \geq 0$

$$\begin{aligned} \nabla\phi_\lambda(\beta) &= \sqrt{1 - (\lambda^T a)^2}\frac{W\beta}{\|\beta\|_W} - \lambda^\mathsf{T} a d \\ r(\beta) &= \lambda^T a + d^\mathsf{T}\beta\frac{\sqrt{1 - (\lambda^T a)^2}}{\|\beta\|_W}. \end{aligned} \tag{5.21}$$

where $W = I - dd^\mathsf{T}$. Thus (5.20) becomes

$$\lambda^\mathsf{T}(x_0 + ad^\mathsf{T} y_0) - \sqrt{1-(\lambda^T a)^2}\frac{\beta^\mathsf{T} W y_0}{\|\beta\|_W} = -\lambda^T a - d^\mathsf{T}\beta\frac{\sqrt{1-(\lambda^T a)^2}}{\|\beta\|_W}.$$

From the last expression, we see that if we are able to find (x_0, y_0) such that

$$x_0 + ad^\mathsf{T} y_0 = -\,a \tag{5.22a}$$

$$d^\mathsf{T}\beta = \beta^\mathsf{T} W y_0 \tag{5.22b}$$

then (5.20) would hold. Note that d is an eigenvector of $W = I - dd^\mathsf{T}$ with eigenvalue $1 - \|d\|^2$. Thus, with $y_0 = \frac{d}{1-\|d\|^2}$ we can easily check that (5.22b) holds. With y_0 defined, in order to satisfy (5.22a) it suffices to set

$$x_0 = -a(d^\mathsf{T} y_0 + 1) = -\frac{a}{1-\|d\|^2}.$$

In summary, we arrive to the following expression for C,

$$C = \left\{ (x,y) \,:\, \begin{array}{lr} \phi_\lambda(y) \le \lambda^\mathsf{T} x & \text{if } \lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \le 0 \\ \phi_\lambda\left(y - \dfrac{d}{1-\|d\|^2}\right) \le \lambda^\mathsf{T}\left(x + \dfrac{a}{1-\|d\|^2}\right) & \text{otherwise} \end{array} \right\}. \tag{5.23}$$

5.6 On the Diagonalization and Homogenization of Quadratics

Consider an arbitrary quadratic set

$$\mathcal{Q} = \{s \in \mathbb{R}^p \,:\, s^\mathsf{T} Q s + b^\mathsf{T} s + c \le 0\}.$$

Given a point $\bar{s} \notin \mathcal{Q}$ we can construct a maximal $\mathcal{Q}$-free set that contains $\bar{s}$ using the techniques developed in the previous sections. The idea to do this is first to find a one-to-one map T such that

$$T(\mathcal{Q}) = S_{\le 0} \cap H = \{(x,y,z) \in \mathbb{R}^{n+m+l} \,:\, \|x\| \le \|y\|,\, a^\mathsf{T} x + d^\mathsf{T} y + h^\mathsf{T} z = -1\}$$
$$T(\bar{s}) \in H \setminus S_{\le 0},$$

for some hyperplane H, that is, for some a, d and h.

Then, we construct a maximal $\mathcal{Q}$-free set using the following fact which can be easily verified: if C is a maximal $S_{\le 0}$-free set with respect to H that contains $T(\bar{s})$, then $T^{-1}(C)$ is a maximal $\mathcal{Q}$-free set containing $\bar{s}$.

Here we show a surprising fact: *which* maximal $\mathcal{Q}$-free set is obtained heavily depends on the choice of T. We illustrate this interesting feature with our running example.

Example 5.47. Let

$$\mathcal{Q} = \{s \in \mathbb{R}^2 \,:\, -2 + 2\sqrt{2}s_1 - 2\sqrt{2}s_2 + 2s_1s_2 \leq 0\}$$

and $\bar{s} = (-2,-2) \notin \mathcal{Q}$. The following map

$$\tau_1(s_1,s_2) = (s_1, s_2, \sqrt{2} + s_1 - s_2)$$

is one-to-one and satisfies

$$\tau_1(\mathcal{Q}) = S^2_{\leq 0} \cap H_1,$$

where $S^2_{\leq 0} \cap H_1$ is defined in Example 5.38 and is given by

$$S^2_{\leq 0} \cap H_1 = \{(x_1,x_2,y) \in \mathbb{R}^3 \,:\, \|x_1,x_2\| \leq |y|,\ -x_1 + x_2 + y = -\sqrt{2}\}.$$

Computing a maximal $S^2_{\leq 0}$-free set with respect to H_1 containing $\tau_1(\bar{s}) = (-2,-2,\sqrt{2})$ yields the same maximal $S^2 \cap H_1$-free set we compute in Example 5.43, that is

$$\begin{aligned} C_1 \cap H_1 = \{(x,y) \,:\, & \frac{1}{\sqrt{2}}(x_1+x_2) - y \leq 0, \\ & \frac{1}{\sqrt{2}}(x_1+x_2) + \frac{1}{\sqrt{2}}y \leq 1 \\ & -x_1 + x_2 + y = -\sqrt{2}\}. \end{aligned}$$

As τ_1^{-1} is simply the projection onto the first two coordinates, we have that

$$\tau_1^{-1}(C_1) = \left\{ s \in \mathbb{R}^2 \,:\, \left(\frac{1}{\sqrt{2}} - 1\right) s_1 + \left(\frac{1}{\sqrt{2}} + 1\right) s_2 + \sqrt{2} \leq 0,\ \sqrt{2}s_1 - 2 \leq 0 \right\}$$

is our maximal $\mathcal{Q}$-free set. This is exactly the set we show in Figure 5.8b.

Now we consider a different transformation for $\mathcal{Q}$. Let

$$\begin{aligned} T_1(s_1,s_2) &= \frac{1}{2}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} s_1 - \sqrt{2} \\ s_2 + \sqrt{2} \end{bmatrix}, \\ T_2(s_1,s_2) &= (-1, s_1, s_2), \text{ and} \\ \tau_2 &= T_2 \circ T_1. \end{aligned}$$

For the curious reader, T_1 is obtained from an eigen-decomposition of the quadratic form. After some algebraic manipulation we can see that

$$\begin{aligned} T_1(\mathcal{Q}) &= \{w \in \mathbb{R}^2 \,:\, T_1^{-1}(w_1,w_2) \in \mathcal{Q}\} \\ &= \{w \in \mathbb{R}^2 \,:\, 1 - w_1^2 + w_2^2 \leq 0\}. \end{aligned}$$

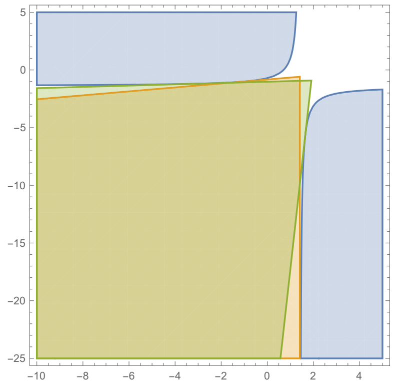


Figure 5.9: Different maximal S-free sets obtained from different transformations, as discussed in Example 5.47. The quadratic set $\mathcal{Q}$ (blue), a maximal $\mathcal{Q}$-free set obtained from τ_1 (orange), and another such set obtained from τ_2 (green).

Thus, τ_2 is one-to-one and

$$\tau_2(\mathcal{Q}) = \{(x_1, x_2, y)\mathbb{R}^2 \,:\, \|x_1, x_2\| \leq |y|,\, x_1 = -1\}.$$

Letting $S^3_{\leq 0} = \{(x_1, x_2, y)\mathbb{R}^3 \,:\, \|x_1, x_2\| \leq |y|,\, x_1 \leq 0\}$ and $H_2 = \{(x_1, x_2, y)\mathbb{R}^3 \,:\, x_1 = -1\}$, we have that $\tau_2(\mathcal{Q}) = S^3_{\leq 0} \cap H_2$. We can now construct a maximal $S^3_{\leq 0}$-free set with respect to H_2. For this, note that in this case $a = (1, 0)$ and $d = 0$. Also, $\tau_2(\bar{s}) = (-1, -2, \sqrt{2})$ and so $\lambda = \frac{1}{\sqrt{5}}(-1, -2)$. As $a^\mathsf{T}\lambda|y| + dy < 0$ for every $y \in \mathbb{R}$, we have that $r(y) = 0$ and $\phi_\lambda(y) = |y|$. By Theorem 5.46,

$$C_2 = \{(x_1, x_2, y) \in \mathbb{R}^3 \,:\, |y| \leq \lambda^\mathsf{T} x\}$$

is maximal $S^3_{\leq 0}$-free set with respect to H_2. Therefore, $\tau_2^{-1}(C_2)$ is maximal $\mathcal{Q}$-free. In Figure 5.9 we show the sets $\mathcal{Q}$ and both maximal $\mathcal{Q}$-free sets given by $\tau_1^{-1}(C_1)$ and $\tau_2^{-1}(C_2)$. Note that in this case, the set $\tau_2^{-1}(C_2)$ does not have an asymptote, and both its facets have an exposing point. $\square$

This example shows the important role of the transformation used to bring the quadratic set to the form S. The resulting maximal S-free set can significantly change. This opens an array of interesting questions regarding the role of transformations in our approach: Can we distinguish the transformations that generate S-free sets with asymptotes? Is there a benefit/downside from using the latter sets? These an other questions are left for future work.

5.7 Further Remarks and Generalizations

In this section we collect some further remarks and generalizations. We start by generalizing Theorem 5.16 to the case where S is represented as the difference of two sublinear functions in independent variables. Then we generalize Proposition 5.21 to the case of several homogeneous linear inequalities. After this we show that we can use Proposition 5.21 to extend the work of Bienstock et al. (2016) by constructing further outer-product-free sets. We also present simpler proofs of some of the outer-products-free sets developed there. Finally, we present an example that shows that there are more quadratic-free sets than the ones that we construct on this chapter.

5.7.1 Generalizing Theorem 5.16

We can generalize Theorem 5.16 to the case when S can be written as the sublevel set of a difference of sublinear functions in *independent* variables.

Theorem 5.48. *Let $\sigma : \mathbb{R}^n \to \mathbb{R}$ be a sublinear function and let $\rho : \mathbb{R}^m \to \mathbb{R}$ be a sublinear function that is positive except at 0. Let*

$$S = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,:\, \sigma(x) \leq \rho(y)\}$$

and let $\bar{x} \neq 0$ be such that there exists a $\bar{y}$ such that $(\bar{x}, \bar{y}) \notin S$. Let $\lambda \in \partial\sigma(\bar{x})$ and

$$C_\lambda = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,:\, \lambda^\mathsf{T} x \geq \rho(y)\}.$$

Then C_λ is maximal S-free.

Proof. First note that $\sigma(\bar{x}) > 0$ since otherwise, due to the positivity of ρ, $(\bar{x}, y) \in S$ for any $y \in \mathbb{R}^m$. Therefore, $0 \notin \partial\sigma(\bar{x})$, in particular $\lambda \neq 0$, and we can assume without loss of generality that $\|\lambda\| = 1$.

We are going to prove maximality via Theorem 5.6. For this notice that

$$C_\lambda = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,:\, \lambda^\mathsf{T} x \geq \beta^\mathsf{T} y \text{ for all } \beta \in \exp \partial\rho(0)\}.$$

Thus, we just need to show that every inequality is exposed by a point of $S \cap C_\lambda$. We show this now. Let $\beta \in \exp \partial\rho(0)$, let y_0 be such that it exposes β, and let $x_0 = \frac{\rho(y_0)}{\sigma(\bar{x})}\bar{x}$. By Proposition 5.53 (see Section 5.10), $\lambda^\mathsf{T}\bar{x} = \sigma(\bar{x})$, which implies that $\lambda^\mathsf{T} x_0 = \rho(y_0)$. Then, Lemma 5.15 implies that (x_0, y_0) exposes $\lambda^\mathsf{T} x \geq \beta^\mathsf{T} y$.

We need to show that $(x_0, y_0) \in S \cap C_\lambda$. As we saw, $\lambda^\mathsf{T} x_0 = \rho(y_0)$, which implies that $(x_0, y_0) \in C_\lambda$. Finally,

$$\sigma(x_0) = \sigma(\frac{\rho(y_0)}{\sigma(\bar{x})}\bar{x}) = \rho(y_0)$$

implies that $(x_0, y_0) \in S$. Notice that the second equality holds because $\rho(y_0) > 0$ and $\sigma(\bar{x}) > 0$. $\square$

The next example shows that the positivity of ρ is necessary.

Example 5.49. Consider $\sigma(x, y) = |x+y| + \|(x, y)\|$ and $\rho(z) = 2z + |z|$. Both functions are positively homogeneous. Let $S = \{(x, y, z) \in \mathbb{R}^3 : \sigma(x, y) \leq \rho(z)\}$, $\bar{x} = (1, 1)$. Then, $\nabla\sigma(\bar{x}) = (1 + \frac{1}{\sqrt{2}})\bar{x}$ and

$$C_\lambda = \{(x, y, z) \in \mathbb{R}^3 : (1 + \frac{1}{\sqrt{2}})(x + y) \geq 2z + |z|\}.$$

We now show that C_λ is not maximal S-free, see also Figure 5.10.

Consider

$$K = \{(x, y, z) \in \mathbb{R}^3 : (1 + \frac{1}{\sqrt{2}})(x + y) \geq 3z\}.$$

Since $z \leq |z|$, then $C_\lambda \subsetneq K$. Furthermore, K is S-free. Indeed, given that $\sigma(x, y) \geq 0$, then any element of S satisfies $z \geq 0$. Thus if (x, y, z) is in S and the interior of K, it must satisfy $z \geq 0$ and

$$\sigma(x, y) \leq 3z < (1 + \frac{1}{\sqrt{2}})(x + y).$$

This is impossible since $(1 + \frac{1}{\sqrt{2}})(x + y) = \nabla\sigma(\bar{x})^\mathsf{T}(x, y) \leq \sigma(x, y)$ for every x, y. $\square$

The next example shows that it is necessary that each sublinear function is in a different set of variables.

Example 5.50. Consider $\sigma(x, y) = |x + y| + x$ and $\rho(x, y) = \|(x, y)\| + y$. Both functions are positively homogeneous. Let $S = \{(x, y) \in \mathbb{R}^2 : \sigma(x, y) \leq \rho(x, y)\}$ and $(\bar{x}, \bar{y}) = (1, 1)$. We have that $(\bar{x}, \bar{y}) \notin S$ and $\nabla\sigma(\bar{x}, \bar{y}) = (2, 1)$. From Figure 5.11 it is easy to see that

$$C = \{(x, y) \in \mathbb{R}^2 : 2x + y \geq \rho(x, y)\}$$

is not maximal S-free. $\square$

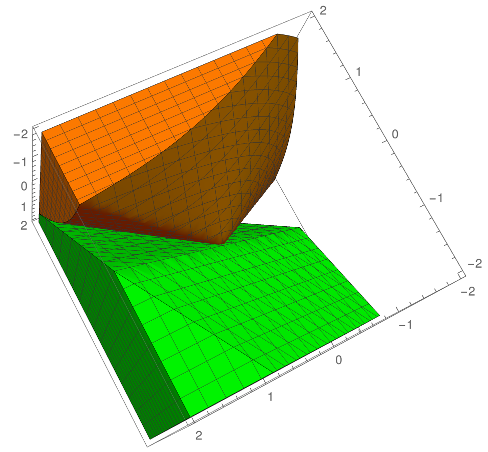


Figure 5.10: The set S (orange) and C_λ (green) from Example 5.49.

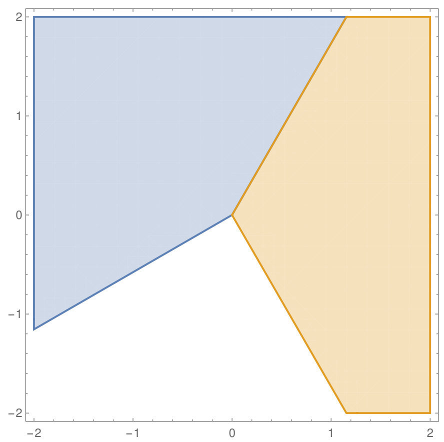


Figure 5.11: The set S (blue) and C (orange) from Example 5.50.

5.7.2 Generalizing Proposition 5.21

Let $P = \{(x, y) : Ax + Dy \leq 0\}$ and let $S_P = \{(x, y) \in P : \|x\| \leq \|y\|\}$. Here we construct maximal S_P-free sets under some conditions on P. The construction generalizes Proposition 5.21.

The construction follows basically the same steps, but there is one extra issue. Just like $G(\lambda)$, we define $G^P(\lambda) = \{\beta : \|\beta\| = 1, A\lambda + D\beta \leq 0\}$. Also, just like $C_{G(\lambda)}$ we define

$$C_{G^P(\lambda)} = \{(x, y) \in \mathbb{R}^{n+m} : -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0, \ \forall \beta \in G^P(\lambda)\}.$$

The extension of Proposition 5.19 presents the extra hypothesis needed. In the proof of Proposition 5.21 it was key to write the non-convex problem $\max_{\beta \in G(\lambda)} y^\mathsf{T} \beta$ as the convex problem $\max\{y^\mathsf{T}\beta : \|\beta\| \leq 1, \ a^\mathsf{T}\lambda + d^\mathsf{T}\beta \leq 0\}$, see (5.6). However, in general, the same does not work using $G^P(\lambda)$ instead of $G(\lambda)$. Indeed,

$$\max\{y^\mathsf{T}\beta : \|\beta\| \leq 1, \ A\lambda + D\beta \leq 0\} \tag{5.24}$$

can have optimal solutions for which $\|\beta\| < 1$. This can never happen when we have a single inequality and $\beta \in \mathbb{R}^m$ with $m > 1$. To force that every optimal solution of (5.24) satisfies $\|\beta\| = 1$ we are going to ask that P has no vertex of the form (λ, β) with $\|\beta\| < 1$.

Alternatively, we could define $G^P(\lambda) = \{\beta : \|\beta\| \leq 1, A\lambda + D\beta \leq 0\}$ and $C_{G^P(\lambda)}$ with the new $G^P(\lambda)$. However, it would not be clear if there is a point in S_P exposing an inequality with $\|\beta\| < 1$. Indeed, it must no happen. This can be seen from modifying Example 5.25. Consider $a = (-2, 4)$ instead of $a = (-3, 4)$. The modification discussed here yields $C_{G(\lambda)} = \{(x, y) : \lambda^\mathsf{T} x + y \geq 0, \ \lambda^\mathsf{T} x - \frac{4}{25} y \geq 0\}$. The second inequality comes from $\beta = \frac{4}{25}$. However, with the new a, $\{(x, y) : \lambda^\mathsf{T} x + y \geq 0\}$ is already maximal.

Finally, we need to generalize the condition $\|d\| \geq \|a\|$. This generalizes to the condition $DD^\mathsf{T} - AA^\mathsf{T}$ is copositive. All together, we have the following.

Proposition 5.51. *Let $(\bar{x}, \bar{y}) \in P \setminus S_P$ and $\lambda = \frac{\bar{x}}{\|\bar{x}\|}$. Assume that P has no vertex (λ, β) with $\|\beta\| < 1$. If $DD^\mathsf{T} - AA^\mathsf{T}$ is copositive, then*

$$C_{G^P(\lambda)} = \{(x, y) \in \mathbb{R}^{n+m} : -\lambda^\mathsf{T} x + \beta^\mathsf{T} y \leq 0, \ \forall \beta \in G^P(\lambda)\}.$$

is maximal S_P-free.

Proof. As P has no vertex (λ, β) with $\|\beta\| < 1$, we have that $G^P(\lambda) \neq \emptyset$ if and only if $\{\|\beta\| \leq 1, \ A\lambda + D\beta \leq 0\} \neq \emptyset$. Since $A\bar{x} + D\bar{y} \leq 0$ and

$\|\bar{x}\| > \|\bar{y}\|$, it holds that $A\lambda + D\frac{\bar{y}}{\|\bar{x}\|} \leq 0$ and $\frac{\|\bar{y}\|}{\|\bar{x}\|} < 1$. In other words, $(-\lambda, \frac{\bar{y}}{\|\bar{x}\|}) \in \{\|\beta\| \leq 1,\ A\lambda + D\beta \leq 0\}$. Note that $(-\lambda, \frac{\bar{y}}{\|\bar{x}\|})$ is a Slater point, see Section 1.3.

To show that $C_{G^P(\lambda)}$ is S_P-free, it is enough to show that

$$\max_{\beta \in G(\lambda)} y^\mathsf{T}\beta \geq \lambda^\mathsf{T} x \text{ for every } (x,y) \in S_P.$$

Since P has no vertex (λ, β) with $\|\beta\| < 1$ the maximum above is equivalent to (5.24). By strong duality (5.24) is equal to

$$\min_{\theta \geq 0} \|y - D^\mathsf{T}\theta\| - \lambda^\mathsf{T} A^\mathsf{T}\theta.$$

Now we just need to show that for any $\theta \geq 0$ and every $(x,y) \in S_P$, the expression $-\lambda^T x + \|y - D^\mathsf{T}\theta\| - \lambda^\mathsf{T} A^\mathsf{T}\theta$ is non-negative. We will now prove that $\lambda^T(x + A^\mathsf{T}\theta) \leq \|y - D^\mathsf{T}\theta\|$, which implies the freeness.

By Cauchy-Schwarz and $\|\lambda\| = 1$, we have that $\lambda^\mathsf{T}(x + A^\mathsf{T}\theta) \leq \|x + A^\mathsf{T}\theta\|$. Furthermore, $\|x + A^\mathsf{T}\theta\|^2 = \|x\|^2 + 2\theta^\mathsf{T} Ax + \|A^\mathsf{T}\theta\|^2$. Since $\theta \geq 0$, $\theta^\mathsf{T} Ax \leq -\theta^\mathsf{T} Dy$. In addition, $\|x\|^2 \leq \|y\|^2$. Thus,

$$\begin{aligned} \|x + A^\mathsf{T}\theta\|^2 &\leq \|y\|^2 - 2\theta^\mathsf{T} Dy + \|A^\mathsf{T}\theta\|^2 \\ &= \|y - D\theta^\mathsf{T}\|^2 + \|A^\mathsf{T}\theta\|^2 - \|D^\mathsf{T}\theta\|^2 \\ &\leq \|y - D\theta^\mathsf{T}\|^2, \end{aligned}$$

where the last inequality is due to the copositivity of $DD^\mathsf{T} - AA^\mathsf{T}$.

We have shown that $\|x + A^\mathsf{T}\theta\| \leq \|y - D\theta^\mathsf{T}\|$. Hence, $\lambda^\mathsf{T}(x + A^\mathsf{T}\theta) \leq \|y - D\theta^\mathsf{T}\|$ and we conclude that $C_{G^P(\lambda)}$ is S_P-free.

Finally, we have to prove maximality. Suppose there exists an S_P-free, C, such that $C_{G^P(\lambda)} \subsetneq C$. This implies that there must exist $\beta_0 \in G^P(\lambda)$ such that $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$ is not valid for C. As $C_\lambda \subseteq C_{G^P(\lambda)}$, $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$ is valid for C_λ.

As we saw in Theorem 5.16, $(\lambda, \beta_0) \in C_\lambda$ exposes $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$, and since $C_\lambda \subseteq C$, Theorem 5.5 implies that $(\lambda, \beta_0) \in \operatorname{int}(C)$. However, since $\beta_0 \in G^P(\lambda)$, we have that $(\lambda, \beta_0) \in S_P$. This contradicts the S_P-freeness of C. □

5.7.3 Extensions to the Work of Bienstock et al. (2016)

Bienstock et al. (2019) construct maximal S-free sets for $S = \{X \in \mathcal{S}^n_+ : \operatorname{rk}(X) = 1\}$. They show that

$$C_{ijkl} = \{X \in \mathcal{S}^n_+ : \lambda_1(x_{ij} + x_{lk}) + \lambda_2(x_{ik} - x_{lj}) \geq \|(x_{ik} + x_{lj}, x_{ij} - x_{lk})\|\}$$

is maximal S-free under some conditions of λ_1 and λ_2 depending on i, j, k, l, see Bienstock et al. (2019, Theorem 4). In other words, the matrices for which the entries of a given 2×2 submatrix satisfies the condition above. To simplify notation we will denote the entries of the submatrix by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and C_{ijkl} by C. For example, if the submatrix is taken from the columns i, j and rows k, l such that $k = j$, that is, b is in the diagonal, then $\lambda_1 = 1, \lambda_2 = 0$ yields a maximal S-free set according to (Bienstock et al., 2019, Theorem 4). Or if none of a, b, c, d corresponds to an entry in the diagonal, then any $(\lambda_1, \lambda_2) \in D_1(0)$ yields a maximal S-free set.

This last result can be deduced as follows. By using the projection theorem Theorem 5.12 we can reduce finding maximal S-free sets to finding the maximal S_0-free sets, where

$$S_0 = \{(a, b, c, d) \in \mathbb{R}^4 \,:\, ad = bc\}.$$

The set

$$S_1 = \{(a, b, c, d) \in \mathbb{R}^4 \in ad \leq bc\}$$

is a non-convex S_0-free set. Using the eigenvalue decomposition we obtain C as a maximal S_1-free set. Theorem 5.16 tells us that C is going to be maximal for any $(\lambda_1, \lambda_2) \in D_1(0)$.

The difficulty when some entries belong to the diagonal is that if $X \in S$, then its diagonal entries are non-negative. Thus, if, say, b is in the diagonal, then $S_0 = \{(a, b, c, d) \in \mathbb{R}^4 \,:\, ad = bc,\ b \geq 0\}$. Thus, $S_1 = \{(a, b, c, d) \in \mathbb{R}^4 \in ad \leq bc,\ b \geq 0\}$ and we can use the techniques from Section 5.4 to construct maximal S_1-free sets.

5.7.4 There Are More Quadratic-Free Sets

It is an interesting question whether every quadratic-free set can be obtain via the construction presented in this chapter. In this section we show that the answer is no. Even for the homogeneous case we can find S^h-free sets that are not given by our construction.

The S^h-free sets C_λ have the following property.

Proposition 5.52. *If* $(x, y) \in S^h \cap C_\lambda \setminus \{(0, 0)\}$ *then* $\frac{x}{\|x\|} = \lambda$.

Proof. If $(x, y) \in C_\lambda$, then $\lambda^T x \geq \|y\|$. If $(x, y) \in S^h$, then $\|y\| \geq \|x\|$. By Cauchy-Schwarz, $\|x\| \geq \lambda^T x$. All together imply that $\|x\| = \lambda^T x$, which implies that $\frac{x}{\|x\|} = \lambda$. $\square$

Consider now $S = \{(x_1, x_2, y_1, y_2) \in \mathbb{R}^4 : \|x\| \leq \|y\|\}$ and let

$$C = \operatorname{conv}\{(1,0,1,0),(0,1,0,1),(1,1,0,0),(1,1,-1,0),(0,0,0,0)\}.$$

C is full dimensional and S-free. To see this, note that the points in the interior of C are of the form $(\lambda_1 + \lambda_3 + \lambda_4, \lambda_2 + \lambda_3 + \lambda_4, \lambda_1 - \lambda_4, \lambda_2)$ for which $\lambda_1 + \ldots + \lambda_5 = 1$ and $\lambda_i > 0$ for $i = 1, \ldots, 5$. For such a point to be in S it must hold satisfy

$$(\lambda_1 + \lambda_3 + \lambda_4)^2 + (\lambda_2 + \lambda_3 + \lambda_4)^2 \leq (\lambda_1 - \lambda_4)^2 + \lambda_2^2.$$

But subtracting the right hand side and factorizing, this is the same as

$$(2\lambda_1 + \lambda_3)(\lambda_3 + 2\lambda_4) + (2\lambda_2 + \lambda_3 + \lambda_4)(\lambda_3 + \lambda_4) \leq 0.$$

No $\lambda_i > 0$ satisfy the above inequality.

Notice that $(1,0,1,0),(0,1,0,1) \in S \cap C$, but the property in Proposition 5.52 does not hold since $(1,0) \neq (0,1)$. Therefore, C can be extended to a maximal S-free set such that $C \neq C_\lambda$ for every $\lambda \in D_1(0)$.

5.8 Computational Experiments

These cuts, among others, are studied computationally in the Master's thesis of Antonia Chmiela (Chmiela, 2020). In her work, a transformation similar to τ_2 of Section 5.6 is used to transform a general quadratic into the form needed to construct the S-free set. Specifically, the idea is to write the quadratic part of a quadratic function as a d.c. using the eigenvalues and eigenvectors and then, to homogenize it.

Two experiments are performed in Chmiela (2020) using the MINLP solver SCIP (Gamrath et al., 2020; Vigerske and Gleixner, 2018; Achterberg, 2009)

The first one consists of testing how much gap can be closed in the root node, when as many cuts as possible are added. This means the following. SCIP creates an initial linear relaxation of the optimization problem at hand. After solving this relaxation we obtain a first lower bound d_1. Then, the root node is processed by tightening bound, adding cutting planes, resolving the LP relaxation, etc. (for more details consult Achterberg (2009)). Just before branching starts, a last lower bound d_2 is obtained. The *gap closed* is then $\frac{d_2 - d_1}{p - d_1}$, where p is the value of the optimal solution. Note that this measure only makes sense when $d_1 \neq p$, thus, in particular, feasibility problems are not considered.

The second experiment consists of an assessment of the performance of SCIP with the cuts included. That is, how much faster (or slower) SCIP is when the intersection cuts are used.

		MAX			DEFAULT			relative	
subset	instances	solved	time	nodes	solved	time	nodes	time	nodes
clean	2689	1625	97.73	1647	1619	99.83	1691	1.02	1.03
affected	1188	716	138.72	3790	710	145.58	4030	1.04	1.06
$[0, 3600]$	1652	1625	9.46	342	1619	9.81	362	1.04	1.06
$[1, 3600]$	965	938	40.19	1112	932	42.64	1198	1.06	1.08
$[10, 3600]$	650	623	135.17	2533	617	146.12	2798	1.08	1.10
$[100, 3600]$	359	332	462.07	8278	326	516.28	9558	1.12	1.16
$[1000, 3600]$	135	108	1226.30	31104	102	1493.00	40748	1.22	1.31
all-optimal	1598	1598	7.91	271	1598	8.17	284	1.03	1.05
diff-timeout	54	27	1202.26	-	21	1418.06	-	1.18	-

Table 5.1: Comparison of running time (in seconds) and number of nodes when using SCIP with the settings MAX and DEFAULT, respectively. The columns "relative" denote the corresponding relative shifted geometric mean of the results obtained by DEFAULT with respect to the results of MAX.

The results of the first experiments are as follows. Out of 690 instances from the MINLPLIB MINLPLIB for which there was a difference in the gap closed, 512 closed more gap. In average, a 3% more gap can be closed in the root node. However, solvers do not add as many cuts as they can and at some point they decide to start branching. Thus, although this result is positive, the empirical performance still needs to be assessed.

For the second experiment, we reproduce Table 4.5 from Chmiela (2020) as Table 5.1. We can observe that, as expected, less nodes are needed. Three reasons why a slowdown might be expected from intersection cuts are the following. First, to compute them, one needs access to the LP tableau, which is not a cheap operation if performed often. Second, intersection cuts are generally dense, which might render the LP harder to solve. Third, the numerics of these cuts can be really bad, in the sense that they might have large and small coefficients, again making the LP harder to solve. Despite all this, we do see a speed-up in the solving time. For example, in the instances for which either SCIP with or without the cut took at least 1000 seconds to solve, the hard instances, we see a 20% speed-up. When considering all instances that did not fail, the speed-up is a modest 2%. These results were obtained by adding at most 20 intersection cuts only in the root node. Although this might sound as a small number, the performance is sensible to this limits. Experiments performed by the author adding additionally at most 2 cuts per node in the tree, led to a 10% slowdown.

These results show that there is potential in this type of cuts for nonlinear problems. For more details, we refer the reader to Chmiela (2020).

5.9 Summary and Future Work

In this chapter we have shown how to construct *maximal* quadratic-free sets, i.e., convex sets whose interior does not intersect the sublevel set of a quadratic function. Using the long-studied *intersection cut* framework, these sets can be used in order to generate deep cutting planes for quadratically constrained problems. We strongly believe that, by carefully laying a theoretical framework for quadratic-free sets, this chapter provides an important contribution to the understanding and future computational development of non-convex quadratically constrained optimization problems.

The maximal quadratic-free sets we construct in this chapter allow for an *efficient* computation of the corresponding intersection cuts. Computing such cutting planes amount to solving a simple one-dimensional convex optimization problem using the quadratic-free sets we show here. Moreover, even if in our constructions and maximality proofs we use semi-infinite outer-descriptions of S-free sets such as (5.17), all of them have closed-form expressions that are more adequate for computational purposes: see (5.2), (5.10), (5.15), (5.23) for these expressions for the sets C_λ, $C_{G(\lambda)}$, C_{ϕ_λ} and C, respectively, and (5.12) for the explicit description of the ϕ_λ function. This ensures efficient separation in LP-based methods for quadratically constrained optimization problems.

The empirical performance of these intersection cuts is promising. The development of a cut strengthening procedure is likely to be important for obtaining an even better empirical performance. Other important open questions involve the better understanding of the role different transformations of quadratic inequalities have (Section 5.6), a theoretical and empirical comparison with the method proposed by Bienstock et al. Bienstock et al. (2016, 2019), and devising new methods for producing other families of quadratic-free sets. All this is subject of ongoing work.

5.10 Missing Proofs

The following is a useful identity that sublinear functions satisfies. For positively homogeneous and differentiable functions the result is implied by the well-known Euler homogeneous function theorem.

Proposition 5.53. *If $\phi : \mathbb{R}^n \to \mathbb{R}$ is sublinear, then $\phi(x) = \beta^\mathsf{T} x$ for every $\beta \in \partial\phi(x)$.*

Proof. Let $\beta \in \partial\phi(x)$. It holds that $\phi(x) + \beta^\mathsf{T}(y - x) \leq \phi(y)$ for every $y \in \mathbb{R}^n$. Taking $y = 2x$ and $y = \frac{1}{2}x$, we conclude that $\phi(x) = \beta^\mathsf{T} x$. $\square$

Lemma 5.15. *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a sublinear function, $\lambda \in D_1(0)$, and let*

$$C = \{(x, y) \,:\, \phi(y) \leq \lambda^\mathsf{T} x\}.$$

If $(\bar{x}, \bar{y}) \in C$ is such that ϕ is differentiable at $\bar{y}$ and $\phi(\bar{y}) = \lambda^\mathsf{T}\bar{x}$, then $(\bar{x}, \bar{y})$ exposes the valid inequality $-\lambda^\mathsf{T} x + \nabla\phi(\bar{y})^\mathsf{T} y \leq 0$.

In particular, if $\beta_0 \in \partial\phi(0)$ is an exposed point of $\partial\phi(0)$, exposed by $\bar{y}$, and $\phi(\bar{y}) = \lambda^\mathsf{T}\bar{x}$, then $(\bar{x}, \bar{y})$ exposes the valid inequality $-\lambda^\mathsf{T} x + \beta_0^\mathsf{T} y \leq 0$.

Proof. We need to verify both conditions of Definition 5.2. As ϕ is positively homogeneous and differentiable at $\bar{y}$, then $\phi(\bar{y}) = \nabla\phi(\bar{y})\bar{y}$. Thus, evaluating $-\lambda^\mathsf{T} x + \nabla\phi(\bar{y})^\mathsf{T} y$ at $(\bar{x}, \bar{y})$ yields $-\lambda^\mathsf{T}\bar{x} + \phi(\bar{y})$, which is 0 by hypothesis. This shows that the inequality is tight at $(\bar{x}, \bar{y})$.

Now, let $\alpha^\mathsf{T} x + \gamma^\mathsf{T} y \leq \delta$ be a non-trivial valid inequality tight at $(\bar{x}, \bar{y})$. Then, $\delta = \alpha^\mathsf{T}\bar{x} + \gamma^\mathsf{T}\bar{y}$ and we can rewrite the inequality as $\alpha^\mathsf{T}(x - \bar{x}) + \gamma^\mathsf{T}(y - \bar{y}) \leq 0$. Notice that $(\phi(y)\lambda, y) \in C$, thus, $\alpha^\mathsf{T}\lambda(\phi(y) - \phi(\bar{y})) + \gamma^\mathsf{T}(y - \bar{y}) \leq 0$ for every $y \in \mathbb{R}^m$. Subtracting $\alpha^\mathsf{T}\lambda\nabla\phi(\bar{y})^\mathsf{T}(y - \bar{y})$ and dividing by $\|y - \bar{y}\|$ we obtain the equivalent expression

$$\alpha^\mathsf{T}\lambda\frac{\phi(y) - \phi(\bar{y}) - \nabla\phi(\bar{y})^\mathsf{T}(y - \bar{y})}{\|y - \bar{y}\|} \leq (-\gamma - \alpha^\mathsf{T}\lambda\nabla\phi(\bar{y}))^\mathsf{T}\frac{y - \bar{y}}{\|y - \bar{y}\|}.$$

Since ϕ is differentiable at $\bar{y}$, the limit when y approaches $\bar{y}$ of the left hand side of the above expression is 0. However, one can make the expression $\frac{y - \bar{y}}{\|y - \bar{y}\|}$ converge to any point of $D_1(0)$. Therefore,

$$0 \leq (-\gamma - \alpha^\mathsf{T}\lambda\nabla\phi(\bar{y}))^\mathsf{T}\beta$$

for every $\beta \in D_1(0)$. This implies that $\gamma = -\alpha^\mathsf{T}\lambda\nabla\phi(\bar{y})$. From here we see that $\alpha \neq 0$ as otherwise $\alpha = \gamma = 0$ and the inequality would be trivial.

Given that any $(x, 0)$ such that $\lambda^T x = 0$ belongs to C, it follows that α is parallel to λ, i.e., there exists $\nu \in \mathbb{R}$ such that $\alpha = \nu\lambda$. Furthermore, $(\mu\lambda, 0) \in C$ for every $\mu \geq 0$, implies that $0 > \alpha^\mathsf{T}\lambda = \nu$. Therefore, $\gamma = -\nu\nabla\phi(\bar{y})$ and the inequality reads $\nu\lambda^\mathsf{T}(x - \bar{x}) - \nu\nabla\phi(\bar{y})^\mathsf{T}(y - \bar{y}) \leq 0$. Dividing by $|\nu|$ and using that $-\lambda^\mathsf{T}x + \nabla\phi(\bar{y})^\mathsf{T}y \leq 0$ is tight at $(\bar{x}, \bar{y})$, we conclude that the inequality can be written as

$$-\lambda^\mathsf{T}x + \nabla\phi(\bar{y})^\mathsf{T}y \leq 0.$$

The second claims follows from the first part of the lemma and the fact that if β_0 is an exposed point of $\partial\phi(0)$ and $\bar{y}$ exposes it, then ϕ is differentiable at $\bar{y}$ and $\nabla\phi(\bar{y}) = \beta_0$. To show this last statement, it is enough to prove that $\partial\phi(\bar{y}) = \{\beta_0\}$, as then (Rockafellar, 1970, Theorem 25.1) implies that $\beta_0 = \nabla\phi(\bar{y})$.

We first show that $\beta_0 \in \partial\phi(\bar{y})$. We have that $\phi(y) = \max\{\beta^\mathsf{T}y \,:\, \beta \in \partial\phi(0)\}$. Since $\bar{y}$ exposes β_0, we have that $\phi(\bar{y}) = \beta_0^\mathsf{T}\bar{y}$. Given that $\beta_0 \in \partial\phi(0)$, we have that $\beta_0^\mathsf{T}y \leq \phi(y)$. Thus, $\phi(\bar{y}) + \beta_0^\mathsf{T}(y - \bar{y}) \leq \phi(y)$ and we conclude that $\beta_0 \in \partial\phi(\bar{y})$.

Now, let $\beta \in \partial\phi(\bar{y})$. Then, $\phi(\bar{y}) + \beta^\mathsf{T}(y - \bar{y}) \leq \phi(y)$ for all y. Proposition 5.53 implies that $\beta^\mathsf{T}y \leq \phi(y)$ and we conclude that $\beta \in \partial\phi(0)$. But $\bar{y}$ exposes β_0, which means that β_0 is the *only* solution to $\phi(\bar{y}) = \max\{\beta^\mathsf{T}\bar{y} \,:\, \beta \in \partial\phi(0)\}$. This implies that $\beta = \beta_0$. Hence, $\partial\phi(\bar{y}) = \{\beta_0\}$ as we wanted to show. □

Proposition 5.54. *Let $a, \lambda \in D_1(0)$, $\lambda \neq \pm a$ and let $d \in \mathbb{R}^m$ be such that $\|d\| \leq 1$. The (Lagrangian) dual problem of*

$$\max_x\{\lambda^\mathsf{T}x \,:\, \|x\| \leq \|y\|, a^\mathsf{T}x + d^\mathsf{T}y \leq 0\} \tag{5.25}$$

is

$$\inf_\theta\{\|\lambda - \theta a\|\|y\| - \theta d^\mathsf{T}y \,:\, \theta \geq 0\}. \tag{5.26}$$

The optimal solution to (5.25) *is $x : \mathbb{R}^m \to \mathbb{R}^n$,*

$$x(y) = \begin{cases} \lambda\|y\|, & \text{if } \lambda^\mathsf{T}a\|y\| + d^\mathsf{T}y \leq 0 \\ \sqrt{\frac{\|y\|^2 - (d^\mathsf{T}y)^2}{1 - (\lambda^T a)^2}}\lambda - \left(d^\mathsf{T}y + \lambda^T a\sqrt{\frac{\|y\|^2 - (d^\mathsf{T}y)^2}{1 - (\lambda^T a)^2}}\right)a, & \text{otherwise.} \end{cases} \tag{5.27}$$

The optimal dual solution is $\theta : \mathbb{R}^m \to \mathbb{R}_+ \cup \{+\infty\}$,

$$\theta(y) = \begin{cases} 0, & \text{if } \lambda^\mathsf{T}a\|y\| + d^\mathsf{T}y \leq 0 \\ \lambda^T a + d^\mathsf{T}y\frac{\sqrt{1 - (\lambda^T a)^2}}{\sqrt{\|y\|^2 - (d^\mathsf{T}y)^2}}, & \text{otherwise.} \end{cases} \tag{5.28}$$

Here, $\frac{1}{0} = +\infty$ *and* $r + (+\infty) = +\infty$ *for every* $r \in \mathbb{R}$. *Moreover, strong duality holds, that is,* $(5.25) = (5.26)$, *and*

$$(5.25) = \begin{cases} \|y\|, & \text{if } \lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \leq 0 \\ \sqrt{(\|y\|^2 - (d^\mathsf{T} y)^2)(1 - (\lambda^T a)^2)} - d^\mathsf{T} y \lambda^\mathsf{T} a, & \text{otherwise.} \end{cases} \tag{5.29}$$

Finally, (5.29) *holds even if* $\lambda = \pm a$.

Proof. First, note that since $\lambda \neq \pm a$ and $\|d\| \leq 1$, $x(y)$ and $\theta(y)$ are defined for every $y \in \mathbb{R}^m$. Second, to make some of the calculations that follow more amenable, let $S(y) = \sqrt{\frac{\|y\|^2 - (d^\mathsf{T} y)^2}{1 - (\lambda^T a)^2}}$.

The Lagrangian of (5.25) is $L : \mathbb{R}^n \times \mathbb{R}^2_+ \to \mathbb{R}$,

$$L(x, \mu, \theta) = \lambda^\mathsf{T} x - \mu(\|x\| - \|y\|) - \theta(a^\mathsf{T} x + d^\mathsf{T} y).$$

Thus, the dual function is

$$d(\mu, \theta) = \max_x L(x, \mu, \theta).$$

We have that $d(\mu, \theta)$ is infinity whenever $\mu < \|\lambda - a\theta\|$, and $\mu\|y\| - \theta d^\mathsf{T} y$ otherwise. Hence, the dual problem, $\min_{\theta,\mu \geq 0} d(\mu, \theta)$, is $\min\{\mu\|y\| - \theta d^\mathsf{T} y \theta : \theta \geq 0, \mu \geq \|\lambda - a\theta\|\}$ which is (5.26).

Let us assume that $\lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y \leq 0$. Clearly, $x(y) = \lambda\|y\|$ is feasible for (5.25). Its objective value is $\|y\|$. On the other hand, $\theta(y) = 0$ is always feasible for (5.25). Its objective value is also $\|y\|$, therefore, $x(y)$ is the primal optimal solution and $\theta(y)$ the dual optimal solution.

Now let us consider the case $\lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y > 0$. Let us check that $\theta(y)$ is dual feasible, that is, $\theta(y) \geq 0$. Note that, due to the positive homogeneity of $\theta(y)$ and the condition $\lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y > 0$ with respect to y, we can assume without loss of generality that $\|y\| = 1$.

Let $\alpha = \lambda^\mathsf{T} a$ and $\beta = d^\mathsf{T} y$. Since $\theta(d) = +\infty \geq 0$ when $\|d\| = 1$, we can assume that $y \neq d$ when $\|d\| = 1$. Note that the same does not occur when $y = -d$ since we are assuming $\lambda^\mathsf{T} a\|y\| + d^\mathsf{T} y > 0$. Thus, $\alpha, \beta \in (-1, 1)$.

We will prove that $\theta(y)\sqrt{1 - \beta^2} = \alpha\sqrt{1 - \beta^2} + \beta\sqrt{1 - \alpha^2} \geq 0$, which implies that $\theta(y) \geq 0$. If $\alpha, \beta \geq 0$, then we are done. As $\alpha + \beta > 0$, at least one of them must be positive. Let us assume $\alpha > 0$ and $\beta < 0$, the other case is analogous. Then, $\alpha > -\beta \geq 0$. This implies that $\alpha^2 > \beta^2$. Subtracting $\alpha^2\beta^2$, factorizing and taking square roots we obtain the desired inequality.

Let us compute the value of the dual solution $\theta(y)$. First, $y = d$ and $\|d\| = 1$, $\theta(y) = +\infty$, which means that the optimal value is

$$\lim_{\theta \to +\infty} \|\lambda - \theta a\| - \theta = -\lambda^\mathsf{T} a.$$

One way of computing this limit is to multiply and divide the expression by $\frac{\|\lambda-\theta a\|+\theta}{\theta}$, expand, and simplify the numerator and denominator until one obtains something simple enough.

Now assume $y \neq d$ if $\|d\| = 1$. Observe that $\|\lambda - \theta(y)a\|\|y\| - \theta(y)d^\mathsf{T} y = \sqrt{\|\lambda - \theta(y)a\|^2}\|y\| - \theta(y)d^\mathsf{T} y$. We have that

$$\begin{aligned}
\|\lambda - \theta(y)a\|^2 &= 1 + \theta(y)(\theta(y) - 2\lambda^\mathsf{T} a) \\
&= 1 + (\theta(y) - \lambda^\mathsf{T} a + \lambda^\mathsf{T} a)(\theta(y) - \lambda^\mathsf{T} a - \lambda^\mathsf{T} a) \\
&= 1 + (\theta(y) - \lambda^\mathsf{T} a)^2 - (\lambda^\mathsf{T} a)^2.
\end{aligned}$$

Replacing $\theta(y)$, we obtain

$$\begin{aligned}
\|\lambda - \theta(y)a\|^2 &= 1 + \frac{(d^\mathsf{T} y)^2}{S(y)} - (\lambda^\mathsf{T} a)^2 \\
&= \frac{1}{S(y)}(S^2(y)(1 - (\lambda^\mathsf{T} a)^2) + (d^\mathsf{T} y)^2) \\
&= \frac{\|y\|^2}{S^2(y)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\lambda - \theta(y)a\|\|y\| - \theta(y)d^\mathsf{T} y &= \frac{\|y\|^2}{S(y)} - d^\mathsf{T} y\lambda^\mathsf{T} a - \frac{(d^\mathsf{T} y)^2}{S(y)} \\
&= \frac{\|y\|^2 - (d^\mathsf{T} y)^2}{S(y)} - d^\mathsf{T} y\lambda^\mathsf{T} a \\
&= \sqrt{(\|y\|^2 - (d^\mathsf{T} y)^2)(1 - (\lambda^T a)^2)} - d^\mathsf{T} y\lambda^\mathsf{T} a.
\end{aligned}$$

Let us now check the feasibility of $x(y)$. Let us first check that $\|x(y)\|^2 \leq \|y\|^2$. We have $\|x(y)\|^2 = S^2(y) - 2S(y)(d^\mathsf{T} y + S(y)\lambda^\mathsf{T} a)\lambda^\mathsf{T} a + (d^\mathsf{T} y + \lambda^\mathsf{T} a S(y))^2$. Expanding and removing common terms yields $\|x(y)\|^2 = S^2(y)(1 - (\lambda^\mathsf{T} a)^2) + (d^\mathsf{T} y)^2 = \|y\|^2$. Thus, the first constraint is satisfied.
To check the second constraint just notice that, as $\|a\| = 1$, $a^\mathsf{T} x(y) = -d^\mathsf{T} y$.

The primal value of $x(y)$ is

$$\lambda^\mathsf{T} x(y) = S(y)(1 - (\lambda^\mathsf{T} a)^2) - d^\mathsf{T} y\lambda^\mathsf{T} a = \sqrt{(\|y\|^2 - (d^\mathsf{T} y)^2)(1 - (\lambda^T a)^2)} - d^\mathsf{T} y\lambda^\mathsf{T} a.$$

As it coincides with the value of the dual solution, even when $y = d$ and $\|d\| = 1$, we conclude that both are optimal.

It only remains to check (5.29) for $\lambda = \pm a$. If $\lambda = -a$, then the linear constraint becomes $\lambda^\mathsf{T} x \geq d^\mathsf{T} y$ and the optimal solution is $x = \lambda\|y\|$. If $\lambda = a$, then the linear constraint becomes $\lambda^\mathsf{T} x \leq -d^\mathsf{T} y$ and $x = -d^\mathsf{T} y \lambda$ is then optimal. In both cases (5.29) holds. $\square$

Lemma 5.55. *Consider the set*

$$S = \{(x, y)\mathbb{R}^{n+m} \,:\, \|x\| \leq \|y\|,\ a^\mathsf{T} x + d^\mathsf{T} y = -1\}$$

with a, d such that $\|a\| > \|d\|$. Let $\lambda, \beta \in D_1(0)$ be two vectors satisfying $\lambda^\mathsf{T} a + d^\mathsf{T} \beta \geq 0$ and consider C_{ϕ_λ} defined in (5.15)*.*

Then, the face of C_{ϕ_λ} defined by the valid inequality $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\beta)^\mathsf{T} y \leq 0$ does not intersect S.

Proof. By contradiction, suppose that $(\bar{x}, \bar{y}) \in C_{\phi_\lambda}$ is such that

$$(\bar{x}, \bar{y}) \in S \quad \wedge \quad -\lambda^\mathsf{T} \bar{x} + \nabla\phi_\lambda(\beta)^\mathsf{T} \bar{y} = 0.$$

The latter equality and the fact that ϕ_λ is sublinear implies $\phi_\lambda(\bar{y}) = \lambda^\mathsf{T} \bar{x}$. Moreover, $\bar{x}$ is a feasible solution of the optimization problem $\phi_\lambda(\bar{y})$, which implies it is an optimal solution.

By Lemma 5.15 we know $(\bar{x}, \bar{y})$ exposes the valid inequality of C_{ϕ_λ} given by $-\lambda^\mathsf{T} x + \nabla\phi_\lambda(\bar{y})^\mathsf{T} y \leq 0$. By definition of exposing point this means

$$\nabla\phi_\lambda(\bar{y}) = \nabla\phi_\lambda(\beta).$$

From (5.21), since W is invertible, we can see that this implies $\beta = \frac{\bar{y}}{\|\bar{y}\|}$. However, as $\lambda^\mathsf{T} a + d^\mathsf{T} \beta \geq 0$, the optimal solution of in the definition of $\phi_\lambda(\bar{y})$, x_0, must satisfy $a^\mathsf{T} x_0 + d^\mathsf{T} \bar{y} = 0$. This contradicts $\phi_\lambda(\bar{y}) = \lambda^\mathsf{T} \bar{x}$, since $\bar{x}$ is an optimal solution but $a^\mathsf{T} \bar{x} + d^\mathsf{T} \bar{y} = -1$. $\square$

Chapter 6

Conclusion

In this thesis, we have mostly studied and developed cutting planes techniques for MINLP. The main contributions of the thesis, grouped by chapters, are the following.

Chapter 1 The exposition of monoidal strengthening.

Chapter 2 The interpretation of Veinott's Supporting Hyperplane algorithm as a particular case of Kelley's Cutting Plane algorithm. The extension of Veinott's algorithm to the case where the feasible region is represented by, possibly, nonconvex and non-differentiable functions.

Chapter 3 The observation that the point we want to separate allows to reduce the feasible region while still ensuring that every separating hyperplane is valid. The formalization of the above observation using the reverse polar and visible points. The characterization of visible points for quadratic constraints.

Chapter 4 The framework for generating intersection cuts for factorable MINLPs. The construction of a concave underestimator of a general factorable function.

Chapter 5 The definition of a point exposing an inequality at infinity. The construction of maximal quadratic-free sets.

We offer a final summary of the main ideas of the different chapters. For Chapter 2, while trying to understand when gradient cuts of functions yield supporting hyperplanes, we observed that Veinott's algorithm naturally appeared. This led to the observation that Veinott's algorithm is just Kelley's algorithm in disguise. The disguise is changing the constraints representing the feasible region, C, by the gauge of C. With this insight it is natural to consider extensions of these algorithm to cases where the constraint function are not convex nor differentiable.

Chapter 3 is based on a simple observation: if a cut separating x is *invalid*, then it must be separate a feasible near x. In other words, the point x defines a subset of the feasible region, V, such that if a cut separates x from V, then it separates x from the feasible region. The way we capitalized on this observation was to find better bounds for the variables based on V. Then, cutting plane methods that exploit bounds can produce stronger cuts.

The main motivation for Chapter 4 came from the parallel between solving MILPs and MINLPs. Currently, a solver for MINLPs such as SCIP would not generate cuts for a violated constraint of the form $x(1-x) \leq 0$. However, if the constraint is written as $x \in 0, 1$, then SCIP would try to generate, for example, Gomory cuts. Applying the same deduction of Gomory cuts to a general nonlinear constraint, leads to a reformulation of the nonlinear constraint. The advantage of this reformulation is that the point to separate is now the vertex. This alone allows to recover Gomory cuts for $x(1-x) \leq 0$, but only because $x(1-x)$ is a concave function. Thus, it is natural to seek for a concave underestimator in order to be able to deduce an intersection cut.

The question motivating Chapter 5 is natural, as maximal S-free sets yield the strongest intersection cuts.

Bibliography

The page numbers in brackets at the end of each citation refer to the text.

T. Achterberg. *Constraint Integer Programming.* PhD thesis, 2009. [134]

T. Achterberg and R. Wunderling. Mixed integer programming: Analyzing 12 years of progress. In *Facets of Combinatorial Optimization*, pages 449–481. Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-38189-8_18. [75]

K. Andersen, Q. Louveaux, R. Weismantel, and L. A. Wolsey. Inequalities from two rows of a simplex tableau. In *Integer Programming and Combinatorial Optimization*, pages 1–15. Springer Berlin Heidelberg, 2007. doi: 10.1007/978-3-540-72792-7_1. [28]

T. Arnold, R. Henrion, Möller, A., and S. Vigerske. A mixed-integer stochastic nonlinear optimization problem with joint probabilistic constraints, 2013. URL `https://edoc.hu-berlin.de/handle/18452/9087`. [39]

A. Bagirov, N. Karmitsa, and M. M. Mäkelä. *Introduction to Nonsmooth Optimization.* Springer International Publishing, 2014. doi: 10.1007/978-3-319-08114-4. [49]

E. Balas. Intersection cuts—a new type of cutting planes for integer programming. *Operations Research*, 19(1):19–39, feb 1971. doi: 10.1287/opre.19.1.19. [10, 75]

E. Balas. Disjunctive programming. In *Discrete Optimization II, Proceedings of the Advanced Research Institute on Discrete Optimization and Systems Applications of the Systems Science Panel of NATO and of the Discrete Optimization Symposium co-sponsored by IBM Canada and SIAM Banff, Aha. and Vancouver*, pages 3–51. Elsevier BV, 1979. doi: 10.1016/s0167-5060(08)70342-x. [25, 76]

E. Balas. Disjunctive programming: Properties of the convex hull of feasible points. *Discrete Applied Mathematics*, 89(1-3):3–44, dec 1998. doi: 10.1016/s0166-218x(98)00136-x. [54]

E. Balas and R. G. Jeroslow. Strengthening cuts for mixed integer programs. *European Journal of Operational Research*, 4(4):224–234, apr 1980. doi: 10.1016/0377-2217(80)90106-x. [23, 26, 28, 29, 76, 85, and 87]

E. Balas and F. Margot. Generalized intersection cuts and a new cut generating paradigm. *Mathematical Programming*, 137(1-2):19–35, aug 2011. doi: 10.1007/s10107-011-0483-x. [76]

E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0–1 programs. *Mathematical Programming*, 58(1-3):295–324, jan 1993. doi: 10.1007/bf01581273. [77]

A. Basu, M. Conforti, G. Cornuéjols, and G. Zambelli. Maximal lattice-free convex sets in linear subspaces. *Mathematics of Operations Research*, 35(3):704–720, Aug. 2010a. doi: 10.1287/moor.1100.0461. URL `https://doi.org/10.1287/moor.1100.0461`. [92]

A. Basu, M. Conforti, G. Cornuéjols, and G. Zambelli. Minimal inequalities for an infinite relaxation of integer programs. *SIAM Journal on Discrete Mathematics*, 24(1):158–168, 2010b. doi: 10.1137/090756375. [28]

A. Basu, G. Cornuéjols, and G. Zambelli. Convex sets and minimal sublinear functions. *Journal of Convex Analysis*, 18(2):427–432, 2011. [76]

A. Basu, M. Campelo, M. Conforti, G. Cornuéjols, and G. Zambelli. Unique lifting of integer variables in minimal inequalities. *Mathematical Programming*, 141(1-2):561–576, jun 2012. doi: 10.1007/s10107-012-0560-9. [28]

A. Basu, S. S. Dey, and J. Paat. Nonunique lifting of integer variables in minimal inequalities. *SIAM Journal on Discrete Mathematics*, 33(2):755–783, jan 2019. doi: 10.1137/17m1117070. [94]

P. Belotti. Disjunctive cuts for nonconvex MINLP. In *Mixed Integer Nonlinear Programming*, pages 117–144. Springer New York, nov 2011. doi: 10.1007/978-1-4614-1927-3_5. [77]

P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wächter. Branching and bounds tightening techniques for non-convex MINLP. *Optimization Methods & Software*, 24(4-5):597–634, 2009. [38, 41, and 74]

D. Bienstock, C. Chen, and G. Munoz. Outer-product-free sets for polynomial optimization and oracle-based cuts. *arXiv preprint arXiv:1610.04604*, 2016. [vii, 77, 90, 99, 128, 132, and 136]

D. Bienstock, C. Chen, and G. Muñoz. Intersection cuts for polynomial optimization. In A. Lodi and V. Nagarajan, editors, *Integer Programming and Combinatorial Optimization*, pages 72–87, Cham, 2019. Springer International Publishing. ISBN 978-3-030-17953-3. doi: 10.1007/978-3-030-17953-3_6. [77, 89, 90, 91, 132, 133, and 136]

M. Bodur, S. Dash, and O. Günlük. Cutting planes from extended lp formulations. *Mathematical Programming*, 161(1-2):159–192, 2017. [91]

P. Bonami, J. Linderoth, and A. Lodi. Disjunctive cuts for mixed integer nonlinear programming problems. *Progress in Combinatorial Optimization*, pages 521–541, 2011. [77]

F. Boukouvala, R. Misener, and C. A. Floudas. Global optimization advances in mixed-integer nonlinear programming, MINLP, and constrained derivative-free optimization, CDFO. *European Journal of Operational Research*, 252(3):701–727, aug 2016. doi: 10.1016/j.ejor.2015.12.018. [1]

S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004. [6]

A. Brondsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Proceedings of the American Mathematical Society*, 16(4):605, aug 1965. doi: 10.2307/2033889. [75]

C. Buchheim and C. D'Ambrosio. Monomial-wise optimal separable underestimators for mixed-integer polynomial optimization. *Journal of Global Optimization*, 67(4): 759–786, may 2016. doi: 10.1007/s10898-016-0443-3. [77]

C. Buchheim and E. Traversi. Separable non-convex underestimators for binary quadratic programming. In *Experimental Algorithms*, pages 236–247. Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-38527-8_22. [77]

A. Chmiela. Intersection cuts for non-convex MINLP. Master's thesis, Technische Universitët Berlin, 2020. [89, 134, and 135]

F. H. Clarke. *Optimization and Nonsmooth Analysis.* Society for Industrial and Applied Mathematics, Jan. 1990. doi: 10.1137/1.9781611971309. [36, 47]

F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory.* Springer New York, 1998. doi: 10.1007/b97650. [47, 48, 49, and 51]

M. Conforti and L. A. Wolsey. "Facet" separation with one linear program. *Mathematical Programming*, may 2018. doi: 10.1007/s10107-018-1299-8. [57]

M. Conforti, G. Cornuéjols, and G. Zambelli. A geometric perspective on lifting. *Operations Research*, 59(3):569–577, jun 2011a. doi: 10.1287/opre.1110.0916. [28, 76, and 94]

M. Conforti, G. Cornuéjols, and G. Zambelli. Corner polyhedron and intersection cuts. *Surveys in Operations Research and Management Science*, 16(2):105–120, jul 2011b. doi: 10.1016/j.sorms.2011.03.001. [17]

M. Conforti, G. Cornuéjols, and G. Zambelli. *Integer Programming.* Springer International Publishing, 2014. ISBN 978-3-319-11008-0. doi: 10.1007/978-3-319-11008-0. [28, 54, 92, and 97]

M. Conforti, G. Cornuéjols, A. Daniilidis, C. Lemaréchal, and J. Malick. Cut-generating functions and S-free sets. *Mathematics of Operations Research*, 40 (2):276–391, may 2015. doi: 10.1287/moor.2014.0670. [17, 75, 76, and 90]

W. de Oliveira. Regularized optimization methods for convex MINLP problems. *TOP*, 24(3):665–692, mar 2016. doi: 10.1007/s11750-016-0413-4. [39]

F. Deutsch, H. Hundal, and L. Zikatanov. Visible points in convex sets and best approximation. In *Computational and Analytical Mathematics*, pages 349–364. Springer New York, 2013. doi: 10.1007/978-1-4614-7621-4_15. [54, 61, and 68]

S. S. Dey and L. A. Wolsey. Two row mixed-integer cuts via lifting. *Mathematical Programming*, 124(1-2):143–174, may 2010. doi: 10.1007/s10107-010-0362-x. [10, 28, 76, and 91]

M. A. Duran and I. E. Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming*, 36(3):307–339, oct 1986. doi: 10.1007/bf02592064. [37]

J. Dutta and C. S. Lalitha. Optimality conditions in convex optimization revisited. *Optimization Letters*, 7(2):221–229, Oct. 2011. doi: 10.1007/s11590-011-0410-3. [36, 47, 50, and 51]

V.-P. Eronen, M. M. Mäkelä, and T. Westerlund. On the generalization of ECP and OA methods to nonsmooth convex MINLP problems. *Optimization*, 63(7):1057–1073, aug 2012. doi: 10.1080/02331934.2012.712118. URL `https://doi.org/10.1080%2F02331934.2012.712118`. [37, 38]

V.-P. Eronen, M. M. Mäkelä, and T. Westerlund. Extended cutting plane method for a class of nonsmooth nonconvex MINLP problems. *Optimization*, pages 1–21, jun 2013. doi: 10.1080/02331934.2013.796473. [38]

V.-P. Eronen, J. Kronqvist, T. Westerlund, M. M. Mäkelä, and N. Karmitsa. Method for solving generalized convex nonsmooth mixed-integer nonlinear programming problems. *Journal of Global Optimization*, 69(2):443–459, May 2017. doi: 10.1007/s10898-017-0528-7. [36, 50]

G. Fasano and R. Pesenti. Conjugate direction methods and polarity for quadratic hypersurfaces. *Journal of Optimization Theory and Applications*, 175(3):764–794, Oct. 2017. doi: 10.1007/s10957-017-1180-6. [68]

M. Fischetti and M. Monaci. A branch-and-cut algorithm for mixed-integer bilinear programming. *European Journal of Operational Research*, sep 2019. doi: 10.1016/j.ejor.2019.09.043. [77]

M. Fischetti, I. Ljubić, M. Monaci, and M. Sinnl. Intersection cuts for bilevel optimization. In *Integer Programming and Combinatorial Optimization*, pages 77–88. Springer International Publishing, 2016. doi: 10.1007/978-3-319-33461-5_7. [77]

M. Fischetti, I. Ljubić, M. Monaci, and M. Sinnl. A new general-purpose algorithm for mixed-integer bilevel linear programs. *Operations Research*, 65(6):1615–1637, dec 2017. doi: 10.1287/opre.2017.1650. [77]

R. Fletcher and S. Leyffer. Solving mixed integer nonlinear programs by outer approximation. *Mathematical Programming*, 66(1):327–349, 1994. ISSN 1436-4646. doi: 10.1007/BF01581153. [37]

G. Gamrath, D. Anderson, K. Bestuzheva, W.-K. Chen, L. Eifler, M. Gasse, P. Gemander, A. Gleixner, L. Gottwald, K. Halbig, G. Hendel, C. Hojny, T. Koch, P. L. Bodic, S. J. Maher, F. Matter, M. Miltenberger, E. Mühmer, B. Müller, M. Pfetsch, F. Schlösser, F. Serrano, Y. Shinano, C. Tawfik, S. Vigerske, F. Wegscheider, D. Weninger, and J. Witzig. The SCIP Optimization Suite 7.0. ZIB-Report 20-10, Zuse Institute Berlin, March 2020. URL `http://nbn-resolving.de/urn:nbn:de:0297-zib-78023`. [134]

A. M. Geoffrion. Generalized benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, oct 1972. doi: 10.1007/bf00934810. [37]

F. Glover. Convexity cuts and cut search. *Operations Research*, 21(1):123–134, feb 1973. doi: 10.1287/opre.21.1.123. [10, 75]

F. Glover. Polyhedral convexity cuts and negative edge extensions. *Zeitschrift für Operations Research*, 18(5):181–186, oct 1974. doi: 10.1007/bf02026599. [14, 76]

M. Goberna, E. González, J. Martínez-Legaz, and M. Todorov. Motzkin decomposition of closed convex sets. *Journal of Mathematical Analysis and Applications*, 364(1):209–221, apr 2010. doi: 10.1016/j.jmaa.2009.10.015. [94]

R. Gomory. An algorithm for the mixed integer problem. Technical report, RAND CORP SANTA MONICA CA, 1960. [24]

R. E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64(5):275–279, sep 1958. doi: 10.1090/s0002-9904-1958-10224-4. [33]

M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer Berlin Heidelberg, 1993. doi: 10.1007/978-3-642-78240-4. [54]

A. S. E. D. Hamed and G. P. McCormick. Calculation of bounds on variables satisfying nonlinear inequality constraints. *Journal of Global Optimization*, 3(1): 25–47, 1993. doi: 10.1007/bf01100238. [55]

M. M. F. Hasan. An edge-concave underestimator for the global optimization of twice-differentiable nonconvex problems. *Journal of Global Optimization*, 71(4): 735–752, mar 2018. doi: 10.1007/s10898-018-0646-x. [77]

J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms II*. Springer Berlin Heidelberg, 1993. doi: 10.1007/978-3-662-06409-2. [39]

R. Horst and H. Tuy. *Global Optimization*. Springer Nature, 1990. doi: 10.1007/978-3-662-02598-7. [5, 35, 37, 38, 45, and 46]

J. J. E. Kelley. The cutting-plane method for solving convex programs. *Journal of the Society for Industrial and Applied Mathematics*, 8(4):703–712, dec 1960. doi: 10.1137/0108053. [32, 35, and 37]

V. Jeyakumar and D. T. Luc. Nonsmooth calculus, minimality, and monotonicity of convexificators. *Journal of Optimization Theory and Applications*, 101(3):599–621, Jun 1999. ISSN 1573-2878. doi: 10.1023/a:1021790120780. [36, 48]

A. Kabgani, M. Soleimani-damaneh, and M. Zamani. Optimality conditions in optimization problems with convex feasible set using convexificators. *Mathematical Methods of Operations Research*, 86(1):103–121, Apr 2017. ISSN 1432-5217. doi: 10.1007/s00186-017-0584-2. [36, 47, and 50]

O. Khamisov. On optimization properties of functions, with a concave minorant. *Journal of Global Optimization*, 14(1):79–101, 1999. doi: 10.1023/a:1008321729949. [75, 77]

M. R. Kılınç and N. V. Sahinidis. Exploiting integrality in the global optimization of mixed-integer nonlinear programming problems with BARON. *Optimization Methods and Software*, 33(3):540–562, jul 2017. doi: 10.1080/10556788.2017.1350178. [74]

J. Kronqvist, A. Lundell, and T. Westerlund. The extended supporting hyperplane algorithm for convex mixed-integer nonlinear programming. *Journal of Global Optimization*, 64(2):249–272, 2016. ISSN 1573-2916. doi: 10.1007/s10898-015-0322-3. [33, 35, 44, 45, and 51]

J. Kronqvist, D. E. Bernal, A. Lundell, and I. E. Grossmann. A review and comparison of solvers for convex MINLP. *Optimization and Engineering*, 20(2):397–455, dec 2018. doi: 10.1007/s11081-018-9411-8. [34, 52]

J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. [91]

J. B. Lasserre. On representations of the feasible set in convex optimization. *Optimization Letters*, 4(1):1–5, oct 2009. doi: 10.1007/s11590-009-0153-6. [35, 47, and 50]

J. B. Lasserre. On convex optimization without convex representation. *Optimization Letters*, 5(4):549–556, apr 2011. doi: 10.1007/s11590-011-0323-1. [36, 47, and 50]

J. B. Lasserre. Erratum to: On convex optimization without convex representation. *Optimization Letters*, 8(5):1795–1796, Apr. 2014. doi: 10.1007/s11590-014-0735-9. [36, 47]

M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, pages 157–270. Springer, 2009. [91]

C. Lemaréchal. An introduction to the theory of nonsmooth optimization. *Optimization*, 17(6):827–858, 1986. doi: 10.1080/02331938608843204. [36]

Y. Lin and L. Schrage. The global solver in the LINDO API. *Optimization Methods and Software*, 24(4-5):657–668, oct 2009. doi: 10.1080/10556780902753221. [74]

M. Lubin, D. Bienstock, and J. P. Vielma. Two-sided linear chance constraints and extensions. *arXiv preprint arXiv:1507.01995*, 2015. [38]

J. E. Martínez-Legaz. Optimality conditions for pseudoconvex minimization over convex sets defined by tangentially convex constraints. *Optimization Letters*, 9(5): 1017–1023, Oct. 2014. doi: 10.1007/s11590-014-0822-y. [36, 47, and 50]

G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part i — convex underestimating problems. *Mathematical Programming*, 10(1):147–175, dec 1976. doi: 10.1007/bf01580665. [4, 6, 69, 77, and 78]

MINLPLIB. MINLP library. `http://www.minlplib.org`. [135]

R. Misener and C. A. Floudas. ANTIGONE: Algorithms for coNTinuous / Integer Global Optimization of Nonlinear Equations. *Journal of Global Optimization*, 59 (2-3):503–526, mar 2014. doi: 10.1007/s10898-014-0166-2. [74]

S. Modaresi, M. R. Kılınç, and J. P. Vielma. Intersection cuts for nonlinear integer programming: convexification techniques for structured sets. *Mathematical Programming*, 155(1-2):575–611, feb 2015. doi: 10.1007/s10107-015-0866-5. [76]

D. Morán and S. S. Dey. On maximal s-free convex sets. *SIAM Journal on Discrete Mathematics*, 25(1):379–393, jan 2011. doi: 10.1137/100796947. [28, 95]

G. Muñoz and F. Serrano. Maximal quadratic-free sets. In *Integer Programming and Combinatorial Optimization*, pages 307–321. Springer International Publishing, 2020. doi: 10.1007/978-3-030-45771-6_24. URL `https://doi.org/10.1007/978-3-030-45771-6_24`. [90]

F. Plastria. Lower subdifferentiable functions and their minimization by cutting planes. *Journal of Optimization Theory and Applications*, 46(1):37–53, may 1985. doi: 10.1007/bf00938758. [38]

M. Porembski. How to extend the concept of convexity cuts to derive deeper cutting planes. *Journal of Global Optimization*, 15(4):371–404, 1999. doi: 10.1023/a:1008315229750. [76]

M. Porembski. Finitely convergent cutting planes for concave minimization. *Journal of Global Optimization*, 20(2):109–132, 2001. doi: 10.1023/a:1011240309783. [76]

B. H. Pourciau. Modern multiplier rules. *The American Mathematical Monthly*, 87 (6):433–452, jun 1980. doi: 10.1080/00029890.1980.11995060. [23]

V. Powers and B. Reznick. Polynomials that are positive on an interval. *Transactions of the American Mathematical Society*, 352(10):4677–4693, oct 2000. doi: 10.1090/s0002-9947-00-02595-2. [69]

A. Prékopa. *Stochastic Programming.* Springer Netherlands, 1995. doi: 10.1007/978-94-017-3087-7. [39]

A. Prékopa and T. Szántai. Flood control reservoir system design using stochastic programming. In *Mathematical Programming in Use*, pages 138–151. Springer Berlin Heidelberg, 1978. doi: 10.1007/bfb0120831. [39]

B. N. Pshenichnyi. *Necessary Conditions for an Extremum.* Marcel Dekker Inc, New York, 1971. [36]

Y. Puranik and N. V. Sahinidis. Domain reduction techniques for global NLP and MINLP optimization. *Constraints*, 22(3):338–376, jan 2017. doi: 10.1007/s10601-016-9267-5. [55]

I. Quesada and I. E. Grossmann. An lp/nlp based branch and bound algorithm for convex minlp optimization problems. *Computers & chemical engineering*, 16 (10-11):937–947, 1992. [37]

R. T. Rockafellar. *Convex analysis.* Princeton University Press, 1970. [6, 22, 23, 34, 40, 42, 57, 62, 63, 64, 65, 97, and 138]

P. Ruys. *Public goods and decentralization: the duality approach in the theory of value.* PhD thesis, Tilburg University, 1974. [57]

A. Saxena, P. Bonami, and J. Lee. Convex relaxations of non-convex mixed integer quadratically constrained programs: extended formulations. *Mathematical Programming*, 124(1-2):383–411, may 2010a. doi: 10.1007/s10107-010-0371-9. [77]

A. Saxena, P. Bonami, and J. Lee. Convex relaxations of non-convex mixed integer quadratically constrained programs: projected formulations. *Mathematical Programming*, 130(2):359–413, mar 2010b. doi: 10.1007/s10107-010-0340-3. [77]

S. Scholtes. *Introduction to Piecewise Differentiable Equations.* Springer New York, 2012. doi: 10.1007/978-1-4614-4340-7. [51]

A. Schrijver. *Theory of Linear and Integer Programming.* Wiley, 1998. [6]

S. Sen and H. D. Sherali. Facet inequalities from simple disjunctions in cutting plane theory. *Mathematical Programming*, 34(1):72–83, jan 1986. doi: 10.1007/bf01582164. [76]

S. Sen and H. D. Sherali. Nondifferentiable reverse convex programs and facetial convexity cuts via a disjunctive characterization. *Mathematical Programming*, 37 (2):169–183, jun 1987. doi: 10.1007/bf02591693. [76]

F. Serrano. Intersection cuts for factorable MINLP. In *Integer Programming and Combinatorial Optimization*, pages 385–398. Springer International Publishing, 2019. doi: 10.1007/978-3-030-17953-3_29. [73]

N. Z. Shor. Quadratic optimization problems. *Soviet Journal of Computer and Systems Sciences*, 25:1–11, 1987. [91]

V. N. Solovev. On a criterion for convexity of a positive-homogeneous function. *Mathematics of the USSR-Sbornik*, 46(2):285–290, Feb. 1983. doi: 10.1070/sm1983v046n02abeh002787. URL `https://doi.org/10.1070/sm1983v046n02abeh002787`. [108]

Szántai. A computer code for solution of probabilistic-constrained stochastic programming problems. In Y. Ermoliev and R.-B. Wets, editors, *Numerical Techniques for Stochastic Optimization*, pages 229–235. Springer Verlag, 1988. [39]

M. Tawarmalani and N. V. Sahinidis. Convex extensions and envelopes of lower semi-continuous functions. *Mathematical Programming*, 93(2):247–263, dec 2002. doi: 10.1007/s10107-002-0308-z. [75]

M. Tawarmalani and N. V. Sahinidis. A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103(2):225–249, may 2005. doi: 10.1007/s10107-005-0581-8. [74]

J. Tind and L. A. Wolsey. On the use of penumbras in blocking and antiblocking theory. *Mathematical Programming*, 22(1):71–81, dec 1982. doi: 10.1007/bf01581026. [57]

E. Towle and J. Luedtke. Intersection disjunctions for reverse convex sets. *arXiv preprint arXiv:1901.02112*, 2019. [77]

A. Tsoukalas and A. Mitsos. Multivariate McCormick relaxations. *Journal of Global Optimization*, 59(2-3):633–662, apr 2014. doi: 10.1007/s10898-014-0176-0. [78, 80]

H. Tuy. Concave programming with linear constraints. *Doklady Akademii Nauk*, 159 (1):32–35, 1964. [10, 75, 76, and 83]

H. Tuy. *Convex Analysis and Global Optimization*. Springer International Publishing, 2016. doi: 10.1007/978-3-319-31484-6. [42]

W. van Ackooij and W. de Oliveira. Convexity and optimization with copulæ structured probabilistic constraints. *Optimization*, 65(7):1349–1376, May 2016. doi: 10.1080/02331934.2016.1179302. [39]

W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. Joint chance constrained programming for hydro reservoir management. *Optimization and Engineering*, oct 2013. doi: 10.1007/s11081-013-9236-4. [39]

W. van Ackooij, E. C. Finardi, and G. M. Ramalho. An exact solution method for the hydrothermal unit commitment under wind power uncertainty with joint probability constraints. *IEEE Transactions on Power Systems*, 33(6):6487–6500, nov 2018. doi: 10.1109/tpwrs.2018.2848594. [39]

A. F. Veinott. The supporting hyperplane method for unimodal programming. *Operations Research*, 15(1):147–152, feb 1967. doi: 10.1287/opre.15.1.147. [33, 35, 37, 45, 50, and 51]

S. Venkatachalam and L. Ntaimo. Integer Set Reduction for Stochastic Mixed-Integer Programming. *arXiv e-prints*, art. arXiv:1605.05194, Apr 2016. [56, 67]

S. Vigerske. *Decomposition in multistage stochastic programming and a constraint integer programming approach to mixed-integer nonlinear programming*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2013. [55]

S. Vigerske and A. Gleixner. SCIP: global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. *Optimization Methods and Software*, 33 (3):563–593, jun 2017. doi: 10.1080/10556788.2017.1335312. [74]

S. Vigerske and A. Gleixner. Scip: global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. *Optimization Methods and Software*, 33 (3):563–593, 2018. doi: 10.1080/10556788.2017.1335312. URL https://doi.org/10.1080/10556788.2017.1335312. [134]

Z. Wei and M. M. Ali. Outer approximation algorithm for one class of convex mixed-integer nonlinear programming problems with partial differentiability. *Journal of Optimization Theory and Applications*, 167(2):644–652, mar 2015a. doi: 10.1007/s10957-015-0715-y. [37]

Z. Wei and M. M. Ali. Convex mixed integer nonlinear programming problems and an outer approximation algorithm. *Journal of Global Optimization*, 63(2):213–227, feb 2015b. doi: 10.1007/s10898-015-0284-5. [37]

Z. Wei and M. M. Ali. Generalized benders decomposition for one class of MINLPs with vector conic constraint. *SIAM Journal on Optimization*, 25(3):1809–1825, jan 2015c. doi: 10.1137/140967519. [38]

T. Westerlund and F. Pettersson. An extended cutting plane method for solving convex MINLP problems. *Computers & Chemical Engineering*, 19:131–136, jun 1995. doi: 10.1016/0098-1354(95)87027-x. [38]

T. Westerlund, H. Skrifvars, I. Harjunkoski, and R. Pörn. An extended cutting plane method for a class of non-convex MINLP problems. *Computers & Chemical Engineering*, 22(3):357–365, feb 1998. doi: 10.1016/s0098-1354(97)00000-8. [38]

T. Westerlund, V.-P. Eronen, and M. M. Mäkelä. On solving generalized convex MINLP problems using supporting hyperplane techniques. *Journal of Global Optimization*, 71(4):987–1011, mar 2018. doi: 10.1007/s10898-018-0644-z. [38]

S. Wiese. *On the interplay of Mixed Integer Linear, Mixed Integer Nonlinear and Constraint Programming.* PhD thesis, Alma Mater Studiorum - Università di Bologna, 2016. [23]

A. Zaffaroni. Convex coradiant sets with a continuous concave cogauge. *Journal of Convex Analysis*, 15(2):325–343, 2008. [54]

G. M. Ziegler. *Lectures on Polytopes.* Springer New York, 1995. doi: 10.1007/978-1-4613-8431-1. URL https://doi.org/10.1007/978-1-4613-8431-1. [18]