Modeling Cookbook

revision 2.0.1

MOSEK ApS

March 17, 2016

CONTENTS

1 **Preface** **1**

2 **Linear optimization** **3**
- 2.1 Introduction . . . 3
- 2.2 Linear modeling . . . 5
- 2.3 Duality in linear optimization . . . 11
- 2.4 Infeasibility in linear optimization . . . 13
- 2.5 Bibliography . . . 15

3 **Conic quadratic optimization** **17**
- 3.1 Introduction . . . 17
- 3.2 Conic quadratic modeling . . . 18
- 3.3 Conic quadratic optimization examples . . . 28
- 3.4 Duality in conic quadratic optimization . . . 30
- 3.5 Infeasibility in conic quadratic optimization . . . 33
- 3.6 Bibliography . . . 34

4 **Semidefinite optimization** **35**
- 4.1 Introduction . . . 35
- 4.2 Semidefinite modeling . . . 41
- 4.3 Semidefinite optimization examples . . . 49
- 4.4 Bibliography . . . 60

5 **Quadratic optimization** **61**
- 5.1 Introduction . . . 61
- 5.2 Convex quadratic optimization . . . 61
- 5.3 Quadratically constrained optimization . . . 64
- 5.4 Separable quadratic optimization . . . 66

6 **Mixed integer optimization** **69**
- 6.1 Introduction . . . 69
- 6.2 Integer modeling . . . 69
- 6.3 Boolean primitives . . . 75
- 6.4 Integer optimization . . . 77
- 6.5 Bibliography . . . 77

7 **Practical optimization** **79**
- 7.1 Introduction . . . 79
- 7.2 The quality of a solution . . . 79

7.3 Distance to a cone 80

8 Case studies **83**
8.1 Introduction 83
8.2 A resource constrained production and inventory problem 83
8.3 Markowitz portfolio optimization 84

9 Duality in conic optimization **89**
9.1 Introduction 89
9.2 The Lagrangian 89
9.3 The dual function and problem 90
9.4 Weak duality 90
9.5 Strong duality 90

10 Notation and definitions **95**

Bibliography **97**

CHAPTER ONE

PREFACE

This cookbook is about model building using convex optimization. It is intended as a modeling guide for the MOSEK optimization package. However, the style is intentionally quite generic without specific MOSEK commands or API descriptions.

There are several excellent books available on this topic, for example the recent books by Ben-Tal and Nemirovski *[BenTalN01]* and Boyd and Vandenberghe *[BV04]*, which have both been a great source of inspiration for this manual. The purpose of this manual is to collect the material which we consider most relevant to our customers and to present it in a practical self-contained manner; however, we highly recommend the books as a supplement to this manual.

Some textbooks on building models using optimization (or mathematical programming) introduce different concept through practical examples. In this manual we have chosen a different route, where we instead show the different sets and functions that can be modeled using convex optimization, which can subsequently be combined into realistic examples and applications. In other words, we present simple *convex building blocks*, which can then be combined into more elaborate convex models. With the advent of more expressive and sophisticated tools like conic optimization, we feel that this approach is better suited.

The first three chapters discuss self-dual conic optimization, namely linear optimization (Chap. *2*), conic quadratic optimization (Chap. *3*) and semidefinite optimization (Chap. *4*), which should be read in succession. Chap. *5* discusses quadratic optimization, and has Chap. *3* as a prerequisite. Finally, the last chapter (Chap. *6*) diverges from the path of convex optimization and discusses mixed integer conic optimization. App. *10* contains details on notation used in the manual.

CHAPTER

TWO

LINEAR OPTIMIZATION

2.1 Introduction

In this chapter we discuss different aspects of linear optimization. We first introduce the basic concepts of a linear optimization and discuss the underlying geometric interpretations. We then give examples of the most frequently used reformulations or modeling tricks used in linear optimization, and we finally discuss duality and infeasibility theory in some detail.

2.1.1 Basic notions

The most basic class of optimization is *linear optimization*. In linear optimization we minimize a linear function given a set of linear constraints. For example, we may wish to minimize a linear function

$$x_1 + 2x_2 - x_3$$

under the constraints that

$$x_1 + x_2 + x_3 = 1, \quad x_1, x_2, x_3 \geq 0.$$

The function we minimize is often called the *objective function*; in this case we have a linear objective function. The constraints are also linear and consists of both linear *equality* constraints and linear *inequality* constraints. We typically use a more compact notation

$$\begin{array}{ll} \mbox{minimize} & x_1 + 2x_2 - x_3 \\ \mbox{subject to} & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0, \end{array} \tag{2.1}$$

and we call (2.1) an *linear optimization problem* The domain where all constraints are satisfied is called the *feasible set*; the feasible set for (2.1) is shown in Fig. 2.1.

For this simple problem we see by inspection that the *optimal value* of the problem is -1 obtained by the *optimal solution*

$$(x_1^\star, x_2^\star, x_3^\star) = (0, 0, 1).$$

Linear optimization problems are typically formulated with a notation similar to

$$\begin{array}{ll} \mbox{minimize} & c^T x \\ \mbox{subject to} & Ax = b \\ & x \geq 0. \end{array}$$

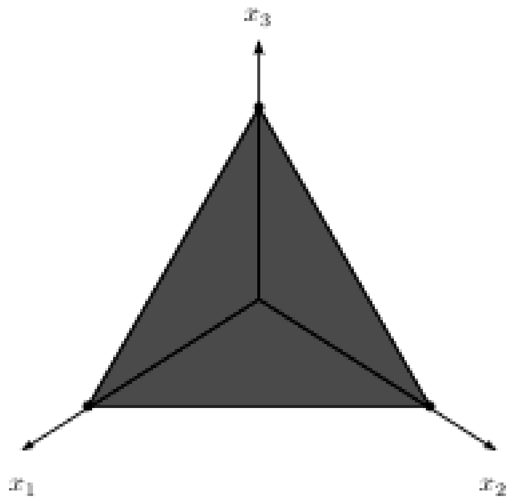


Fig. 2.1: Feasible set for $x_1 + x_2 + x_3 = 1$ and $x_1, x_2, x_3 \geq 0$.

For example, we can pose (2.1) in this form with

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

There are many other formulations for linear optimization problems; we can have different types of constraints,

$$Ax = b, \quad Ax \geq b, \quad Ax \leq b, \quad l^c \leq Ax \leq u^c,$$

and different bounds on the variables

$$l^x \leq x \leq u^x$$

or we may have no bounds on some x_i, in which case we say that x_i is a *free variable*.

All these formulations are equivalent in the sense that by simple transformations and introduction of auxiliary variables they can represent the same problem with the same optimal solution. The important feature is that the objective function and the constraints are all *linear* in x.

2.1.2 Geometry of linear optimization

A *hyperplane* is a set $\{x \mid a^T(x - x_0) = 0\}$ or equivalenty $\{x \mid a^T x = \gamma\}$ with $a^T x_0 = \gamma$, see Fig. 2.2.

Thus linear constraints

$$Ax = b$$

with $A \in \mathbb{R}^{m \times n}$ represents an intersection of m hyperplanes.

Next consider a point x above the hyperplane in Fig. 2.2. Since $x - x_0$ forms an acute angle with a we have that $a^T(x - x_0) \geq 0$, or $a^T x \geq \gamma$. The set $\{x \mid a^T x \geq \gamma\}$ is called a halfspace, see Fig. 2.3. Similarly the set $\{x \mid a^T x \leq 0\}$ forms another halfspace; in Fig. 2.3 it corresponds to the area below the dashed line.

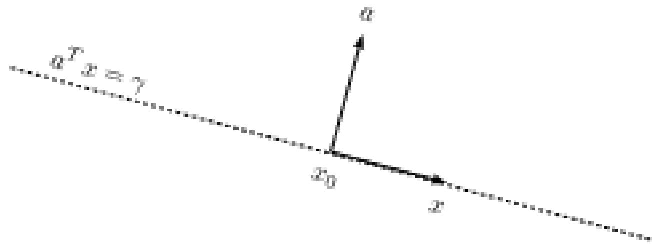


Fig. 2.2: The dashed line illustrates a hyperplane $\{x \mid a^T x = \gamma\}$

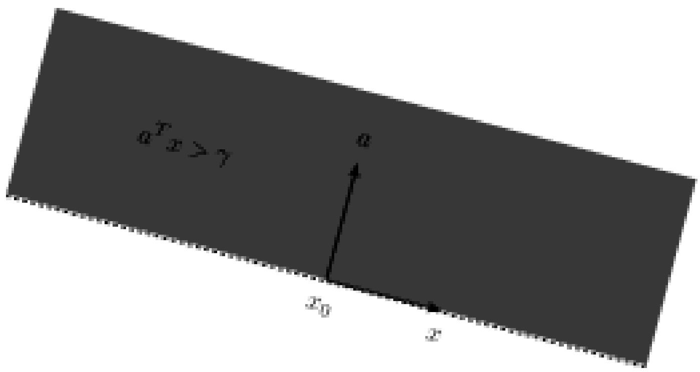


Fig. 2.3: The gray area is the halfspace $\{x \mid a^T x \geq \gamma\}$

A set of linear inequalities

$$Ax \leq b$$

corresponds to an intersection of halfspaces and forms a *polyhedron*, see Fig. 2.4.

The polyhedral description of the feasible set gives us a very intuitive intepretation of linear optimization, which is illustrated in Fig. 2.5.

The dashed lines are normal to the objective $c = (-1, 1)$, and to minimize $c^T x$ we move as far as possible in the opposite direction of c, to a point where one the normals intersect the polyhedron; an optimal solution is therefore always either a vertex of the polyhedron, or an entire facet of the polyhedron may be optimal.

The polyhedron shown in the figure in bounded, but this is not always the case for polyhedra coming from linear inequalities in optimization problems. In such cases the optimization problem may be unbounded, which we will discuss in detail in Section *2.4*.

2.2 Linear modeling

In this section we discuss both useful reformulation techniques to model different functions using linear programming, as well as common practices that are best avoided. By modeling we mean equivalent reformulations that lead to the same optimal solution; there is no approximation or modeling error involved.

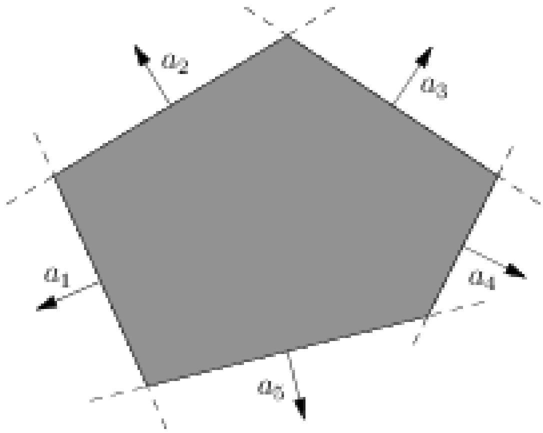


Fig. 2.4: A polyhedron formed as an intersection of halfspaces.

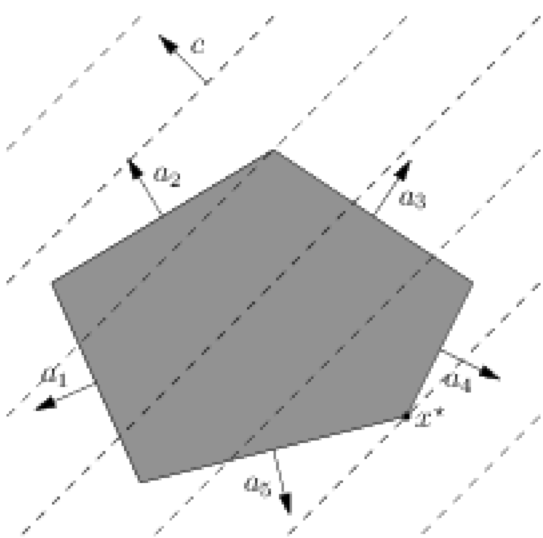


Fig. 2.5: Geometric interpretation of linear optimization. The optimal solution $x^\star$ is at point where the normals to c (the dashed lines) intersect the polyhedron.

2.2.1 Convex piecewise-linear functions

Perhaps the most basic and frequently used reformulation for linear optimization involves modeling a convex piecewise-linear function by introducing a number of linear inequalities. Consider the convex piecewise-linear (or rather piecewise-affine) function illustrated in Fig. 2.6, where the function can be described as $\max\{a_1x+b_1, a_2x+b_2, a_3x+b_3\}$.

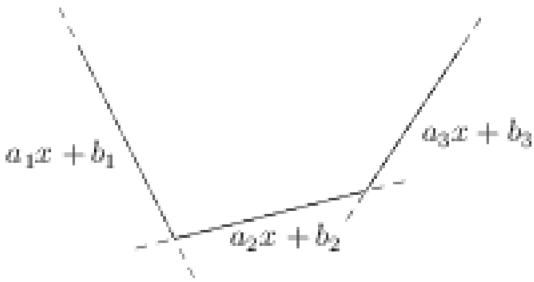


Fig. 2.6: A convex piecewise-linear function (solid lines) of a single variable x. The function is defined as the maximum of 3 affine functions.

More generally, we consider convex piecewise-linear functions $f : \mathbb{R}^n \mapsto \mathbb{R}$ defined as

$$f(x) := \max_{i=1,\dots,m} \{a_i^T x + b_i\},$$

where the important notion is that f is defined as the maximum of a set of affine functions. To model the *epigraph* $f(x) \leq t$ (see App. 10), we note that $t \geq \max_i\{a_i^T x + b_i\}$ if and only if t is larger than all the affine functions, i.e., we have an equivalent formulation with m inequalities,

$$a_i^T x + b_i \leq t, \quad i = 1, \dots, m.$$

Piecewise-linear functions have many uses linear in optimization; either we have a convex piecewise-linear formulation from the onset, or we may approximate a more complicated (nonlinear) problem using piecewise-linear approximations.

A word of caution is in order at this point; in the past the use of linear optimization was significantly more widespread than the use of nonlinear optimation, and piecewise-linear modeling was used to approximately solve nonlinear optimization problems. The field of nonlinear optimization (especially conic optimization) has since matured, and with modern optimization software it is both easier and more efficient to directly formulate and solve conic problems without piecewise-linear approximations.

2.2.2 Absolute values

The absolute value of a scalar variable is an example of a simple convex piecewise-linear function,

$$|x| := \max\{\ x,\ -x\ \},$$

so we model $|x| \leq t$ using two inequalities

$$-t \leq x \leq t.$$

2.2.3 The ℓ_1 norm

All norms are convex functions, but the ℓ_1 and ℓ_∞ norms are of particular interest for linear optimization. The ℓ_1 norm of vector $x \in \mathbb{R}^n$ is defined as

$$\|x\|_1 := |x_1| + |x_2| + \cdots + |x_n|.$$

To model the epigraph

$$\|x\|_1 \leq t, \tag{2.2}$$

we first introduce the following system

$$|x_i| \leq z_i, \; i = 1, \ldots, n, \quad \sum_{i=1}^n z_i = t, \tag{2.3}$$

with additional (auxiliary) variable $z \in \mathbb{R}^n$, and we claim that (2.2) and (2.3) are equivalent. They are equivalent if all $z_i = |x_i|$ in (2.3). Suppose that t is optimal (as small as possible) for (2.3), but that some $z_i > |x_i|$. But then we could reduce t further by reducing z_i contradicting the assumption that t optimal, so the two formulations are equivalent. Therefore, we can model (2.2) using linear (in)equalities

$$-z_i \leq x_i \leq z_i, \quad \sum_{i=1}^n z_i = t,$$

with auxiliary variables z. Similarly, we can describe the epigraph of the norm of an affine function of x,

$$\|Ax - b\|_1 \leq t$$

as

$$-z_i \leq a_i^T x - b_i \leq z_i, \quad \sum_{i=1}^n z_i = t,$$

where a_i is the $i-$th row of A (taken as a column-vector).

The ℓ_1 norm is overwhelmingly popular as a convex approximation of the cardinality (i.e., number on nonzero elements) of a vector x. For example, suppose we are given an overdetermined linear system

$$Ax = b$$

where $A \in \mathbb{R}^{m \times n}$ and $m << n$. The basis pursuit problem

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax = b, \end{array} \tag{2.4}$$

uses the ℓ_1 norm of x as a heuristic of finding a sparse solution (one with many zero elements) to $Ax = b$, i.e., it aims to represent b using few columns of A. Using the reformulation above we can pose the problem as a linear optimization problem,

$$\begin{array}{llcccc} \text{minimize} & & & e^T z & & \\ \text{subject to} & -z & \leq & x & \leq & z \\ & & & Ax & = & b, \end{array} \tag{2.5}$$

where $e = (1, \ldots, 1)^T$.

2.2.4 The ℓ_∞ norm

The ℓ_∞ norm of a vector $x \in \mathbb{R}^n$ is defined as

$$\|x\|_\infty := \max_{i=1,\dots,n} |x_i|,$$

which is another example of simple piecewise-linear functions. To model

$$\|x\|_\infty \leq t \tag{2.6}$$

we use that $t \geq \max_{i=1,\dots,n} |x_i|$ if and only if t is greater than each term, i.e., we can model (2.6) as

$$-t \leq x_i \leq t, \quad i = 1, \dots, n.$$

Again, we can also consider an affine function of x, i.e.,

$$\|Ax - b\|_\infty \leq t,$$

which can be described as

$$-t \leq a_i^T x - b \leq t, \quad i = 1, \dots, n.$$

It is interesting to note that the ℓ_1 and ℓ_∞ norms are dual norms. For any norm $\|\cdot\|$ on $\mathbb{R}^n$, the dual norm $\|\cdot\|_*$ is defined as

$$\|x\|_* = \max\{x^T v \mid \|v\| \leq 1\}.$$

Let us verify that the dual of the ℓ_∞ norm is the ℓ_1 norm. Consider

$$\|x\|_{*,\infty} = \max\{x^T v \mid \|v\|_\infty \leq 1\}.$$

Obviously the maximum is attained for

$$v_i = \begin{cases} +1, & x_i \geq 0, \\ -1, & x_i < 0, \end{cases}$$

i.e., $\|x\|_{*,\infty} = \|x\|_1 = \sum_i |x_i|$. Similarly, consider the dual of the ℓ_1 norm,

$$\|x\|_{\star,1} = \max\{x^T v \mid \|v\|_1 \leq 1\}.$$

To maximize $x^T v$ subject to $|v_1| + \cdots + |v_n| \leq 1$ we identify the largest element of x, say $|x_k|$, The optimizer v is then given by $v_k = \pm 1$ and $v_i = 0$, $i \neq k$, i.e., $\|x\|_{*,1} = \|x\|_\infty$. This illustrates a more general property of dual norms, namely that $\|x\|_{**} = \|x\|$.

2.2.5 Avoid ill-posed problems

A problem is *ill posed* if small perturbations of the problem data result in arbitrarily large perturbations of the solution, or change feasibility of the problem. Such problem formulations should always be avoided as even the smallest numerical perturbations (for example rounding errors, or solving the problem on a different computer) can result in different or wrong solutions. Additionally, from an algorithmic point of view, even computing a *wrong* solution is very difficult for ill-posed problems.

A rigorous definition of the degree of *ill-posedness* is possible by defining a condition number for a linear optimization, but unfortunately this is not a very practical metric, as evaluating such a condition number requires solving several optimization problems. Therefore even though being able to quatify the difficulty of an optimization problem from a condition number is very attractive, we only make the modest recommendations to avoid problems

- that are nearly infeasible,
- with constraints that are linearly dependent, or nearly linearly dependent.

Typical examples of problems with nearly linearly dependent constraints are discretizations of continuous processes, where the constraints invariably become more correlated as we make the discretization finer; as such there may be nothing wrong with the discretization or problem formulation, but we should expect numerical difficulties for sufficiently fine discretizations.

2.2.6 Avoid badly scaled problems

Another difficulty encountered in practice is with model that are badly scaled. Loosely defined, we consider a model to be badly scaled if

- variables are measured on very different scales,
- constraints or bounds are measured on very different scales.

For example if one variable x_1 has units of molecules and another variable x_2 measures temperature, we expect both the coefficients c_1 and c_2 to be of a very different scale. In that case we might have an objective

$$10^{12}x_1 + x_2$$

and if we normalize the objective by 10^{12} and round coefficients to 8 digits, we see that x_2 effectively disappears from the objective. In practice we do not round coefficients, but in finite precision arithmetic any contribution of x_2 to the objective will be insignificant (and unreliable).

A similar situation (from a numerical point of view) is encountered with the method of *penalization* or $big - M$ strategies. For sake of argument, assume we have a standard linear optimization problem (2.7) with the additional constraint that $x_1 = 0$. We may eliminate x_1 completely from the model, or we might add an additional constraint, but suppose we choose to formulate a penalized problem instead

$$\begin{array}{ll} \text{minimize} & c^T x + 10^{12} x_1 \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array}$$

reasoning that the large penalty term will force $x_1 = 0$. However, if $\|c\| < 1$ we have the exact same problem, namely that in finite precision the penalty term will complety dominate the objective and render the contribution $c^T x$ insignificant or unreliable. Therefore, penality or big-M terms should never be used explicitly as part of an optimization model.

As a different example, consider again a problem (2.7), but with additional (redundant) constraints $x_i \leq \gamma$. For optimization practioners this is a common approach for stabilizing the optimization algorithm since it improves conditioning of the linear systems solved by an interior-point algorithm. The problem we solve is then

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \\ & x \leq \gamma e, \end{array}$$

with a dual problem

$$\begin{array}{ll} \text{maximize} & b^T y - \gamma e^T z \\ \text{subject to} & A^T y + s - z = c \\ & s, z \geq 0. \end{array}$$

Suppose we do not know *a-priori* an upper bound on $\|x\|_\infty$, so we choose $\gamma = 10^{12}$ reasoning that this will not change the optimal solution. Note that the large variable bound becomes a penalty term in the dual problem; in finite precision such a large bound will effectively destroy accuracy of the solution.

2.3 Duality in linear optimization

Duality theory is a rich and powerful area of convex optimization, and central to understanding sensitivity analysis and infeasibility issues in linear (and convex) optimization. Furthermore,it provides a simple and systematic way of obtaining non-trivial lower bounds on the optimal value for many difficult non-convex problem. In this section we only discuss duality theory at a descriptive level suited for practioners; we refer to Appendix *9* for a more advanced treatment.

2.3.1 The dual problem

Initially, consider the standard linear optimization problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \tag{2.7}$$

Associated with (2.7) is a so-called *Lagrangian* function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}$ that augments the objective with a weighted combination of all the constraints,

$$L(x, y, s) = c^T x + y^T (b - Ax) - s^T x,$$

The variables $y \in \mathbb{R}^p$ and $s \in \mathbb{R}^n_+$ are called *Lagrange multipliers* or *dual variables*. It is easy to verify that

$$L(x, y, s) \leq c^T x$$

for any feasible x. Indeed, we have $b - Ax = 0$ and $x^T s \geq 0$ since $x, s \geq 0$, i.e., $L(x, y, s) \leq c^T x$. Note the importance of nonnegativity of s; more generally of all Langrange multipliers associated with inequality constraints. Without the nonnegativity constraint the Lagrangian function is not a lower bound.

The *dual function* is defined as the minimum of $L(x, y, s)$ over x. Thus the dual function of (2.7) is

$$g(y, s) = \min_x L(x, y, s) = \min_x x^T (c - A^T y - s) + b^T y.$$

We see that the Langrangian function is linear in x, so it is unbounded below unless when $c - A^T y - s = 0$, i.e.,

$$g(y,s) = \begin{cases} b^T y, & c - A^T y - s = 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

Finally, we get a *dual problem* is by maximizing $g(y,s)$. The dual problem of (2.7) is

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \geq 0. \end{array} \tag{2.8}$$

Example 2.1 (Dual of basis pursuit). As another example, let us derive the dual of the basis pursuit formulation (2.5). The Lagrangian function is

$$L(x,z,y,u,v) = e^T z + u^T(x-z) - v^T(x+z) + y^T(b - Ax)$$

with Lagrange multipliers $y \in \mathbb{R}^m$ and $u, v \in \mathbb{R}^n_+$. The dual function

$$g(y,u,v) = \min_{x,z} L(x,z,y,u,v) = \min_{x,z} z^T(e-u-v) + x^T(u - v - A^T y) + y^T b$$

is linear in z and x so unbounded below unless $e = u + v$ and $A^T y = u - v$, i.e., the dual problem is

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & e = u + v, \\ & A^T y = u - v \\ & u, v \geq 0. \end{array} \tag{2.9}$$

Example 2.2 (Dual of basis pursuit revisited). We can also derive the dual of the basis pursuit formulation (2.4) directly. The Lagrangian is

$$L(x,y) = \|x\|_1 + y^T(Ax - b) = \|x\|_1 + x^T A^T y - b^T y$$

with a dual function

$$g(y,s) = -b^T y + \min_x (\|x\|_1 + x^T A^T y).$$

The term $\min_x(\|x\|_1 + x^T A^T y)$ can be simplified as

$$\begin{aligned} \min_x (\|x\|_1 + x^T A^T y) &= \min_{t \geq 0} \min_{\|z\|_1 = 1} (t\|z\|_1 + t z^T A^T y) \\ &= \min_{t \geq 0} t(1 - \max_{\|z\|_1 = 1} z^T A^T y) \\ &= \min_{t \geq 0} t(1 - \|A^T y\|_\infty), \end{aligned} \tag{2.10}$$

where we used the definition of the dual norm in the last line. Finally $\min_{t \geq 0} t(1 - \|A^T y\|_\infty)$ is 0 if $\|A^T y\|_\infty \leq 1$ and unbounded below otherwise. In other words, we get a dual function

$$g(y) = \begin{cases} -b^T y, & \|A^T y\|_\infty \leq 1, \\ -\infty, & \text{otherwise,} \end{cases}$$

and a dual problem

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \|A^T y\|_\infty \leq 1. \end{array} \tag{2.11}$$

Not surprisingly, (2.9) and (2.11) are equivalent in the sense that

$$e = u + v, \quad A^T y = u - v, \quad u, v \geq 0 \qquad \Longleftrightarrow \qquad \|A^T y\|_\infty \leq 1.$$

2.3.2 Duality properties

Many important observations can be made for the dual problem (2.8), which we discuss in details in Appendix 9. We summarize the most important properties below; to that end let $x^\star$ and $(y^\star, s^\star)$ denote the optimal primal and dual solutions with optimal values $p^\star := c^T x^\star$ and $d^\star := b^T y^\star$, respectively.

- In both the primal and dual problem we optimize over the nonnegative orthant, i.e., $x, s \geq 0$. The number of constraints in the primal problem becomes the number of variables in the dual problem, and vice versa.
- The dual problem gives a lower bound on the primal problem for all (x, y, s). From the dual problem we have $c - A^T y = s$, so

  $$x^T (c - A^T y) = x^T s \geq 0,$$

  and since $Ax = b$ we have

  $$c^T x - b^T y = x^T s \geq 0$$

  i.e., $b^T y \leq c^T x$. This is called the *weak duality* property, and the nonnegative entity $x^T s$ is called the *complementarity gap*.
- If the primal feasible set is nonempty

  $$\mathcal{F}_p = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\} \neq \emptyset$$

  or the dual feasible set is nonempty

  $$\mathcal{F}_d = \{y \in \mathbb{R}^m \mid c - A^T y \geq 0\} \neq \emptyset$$

  then the optimal primal and dual values coincide, i.e.,

  $$d^\star = c^T x^\star = b^T y^\star = p^\star.$$

  This remarkable fact (see Appendix 9.5 for a proof) is called the *strong duality* property. If strong duality holds then $(x^\star)^T s^\star = 0$ so $x^\star$ and $s^\star$ are complementary; since $x^\star, s^\star \geq 0$ we have that $x_i^\star > 0 \Leftrightarrow s_i^\star = 0$ and vice versa.

  For linear (and conic) optimization strong duality means that we have the choice of solving either the primal or the dual problem (or both).

2.4 Infeasibility in linear optimization

2.4.1 Basic concepts

In Sec. 2.3.2 we summarized the main duality properties, namely weak and strong duality properties. In this section we discuss situations where strong duality does not hold. Those situations are captured by the following two results known as (variations of) Farkas' lemma; for proofs see Appendix 9.

Lemma 2.1 (Farkas' lemma)**.** *Given A and b, exactly one of the two statements are true:*

1. *There exists an $x \geq 0$ such that $Ax = b$.*
2. *There exists a y such that $A^T y \leq 0$ and $b^T y > 0$.*

The Farkas lemma tells us that either the primal problem (2.7) is feasible ($\mathcal{F}_p \neq \emptyset$) or there exists a y such that $A^T y \leq 0$ and $b^T y > 0$. In other words, any y satisfying

$$A^T y \leq 0, \quad b^T y > 0$$

is a *certificate of primal infeasibility*. We can also think of a Farkas certificate as an unbounded direction for the dual problem; to that end assume that

$$\nexists x \geq 0 \ : \ Ax = b,$$

so we have a y satisfying $A^T y \leq 0$ and $b^T y > 0$. If we further assume existence of point y_0 satifying

$$c - A^T y_0 \geq 0$$

then the dual remains feasible in the direction of y,

$$c - A^T (ty + y_0) \geq 0, \quad \forall t \geq 0$$

with an unbounded objective $b^T (ty + y_0) \to \infty$ for $t \to \infty$, i.e., $d^\star = \infty$.

Similarly, the dual variant of Farkas' lemma states that either the dual problem is feasible ($\mathcal{F}_d \neq \emptyset$) or there exists an $x \geq 0$ such that $Ax = 0$ and $c^T x < 0$. More precisely

Lemma 2.2 (Farkas' lemma dual variant)**.** *Given A and c, exactly one of the two statements are true:*

1. *There exists an $x \geq 0$ such that $Ax = 0$ and $c^T x < 0$.*
2. *There exists a y such $c - A^T y \geq 0$.*

In other words, any $x \geq 0$ satisfying $Ax = 0$ and $c^T x < 0$ is a *certificate of dual infeasibility*. If the primal problem is feasible, then the certificate is a feasible unbounded direction for the primal objection, i.e., $p^\star = -\infty$.

Below we summarize the different cases that can occur in linear optimization:

- If the either the primal or dual problems are feasible, we have strong duality, i.e., $p^\star = d^\star$.

- If the primal problem is infeasible ($p^\star = \infty$), then from Farkas' lemma the dual problem is unbounded ($d^\star = \infty$) or infeasible ($d^\star = -\infty$).
- If the primal problem is unbounded ($p^\star = -\infty$), then from weak duality the dual problem is infeasible ($d^\star = -\infty$).
- If the dual problem is infeasible ($d^\star = -\infty$), then from Farkas' dual lemma then the primal problem is unbounded ($p^\star = -\infty$) or infeasible ($p^\star = \infty$).

Example 2.3 (Primal and dual infeasibility). As an example exhibiting both primal and dual infeasibility consider the problem

$$\begin{array}{ll} \text{minimize} & -x_1 - x_2 \\ \text{subject to} & x_1 = -1 \\ & x_1,\, x_2 \geq 0 \end{array}$$

with a dual problem

$$\begin{array}{ll} \text{maximize} & -y \\ \text{subject to} & \left[\begin{array}{c} 1 \\ 0 \end{array}\right] y \leq \left[\begin{array}{c} -1 \\ -1 \end{array}\right]. \end{array}$$

Both the primal and dual problems are trivially infeasible; $y = -1$ serves as a certificate of primal infeasibility, and $x = (0, 1)$ is a certificate of dual infeasiblity.

2.4.2 Locating infeasibility

In some cases we are interested in locating the cause of infeasibility in a model, for example if expect the infeasibility to be caused by an error in the problem formulation. This can be difficult in practice, but a Farkas certificate lets us reduce the dimension of the infeasible problem, which in some cases pinpoints the cause of infeasibility.

To that end, suppose we are given a certificate of primal infeasibility,

$$A^T y \leq 0, \quad b^T y > 0,$$

and define the index-set

$$\mathcal{I} = \{i \mid y_i \neq 0\}.$$

Consider the reduced set of constraints

$$A_{\mathcal{I},:} x = b_{\mathcal{I}}, \quad x \geq 0.$$

It is then easy to verify that

$$A_{\mathcal{I},:}^T y_{\mathcal{I}} \leq 0, \quad b_{\mathcal{I}}^T y_{\mathcal{I}} > 0$$

is an infeasibility certificate for the reduced problem with fewer constraints. If the reduced system is sufficiently small, it may be possible to locate the cause of infeasibility by manual inspection.

2.5 Bibliography

The material in this chapter is very basic, and can be found in any textbook on linear optimization. For further details, we suggest a few standard references *[Chv83]*, *[BT97]* and *[PS98]*, which all cover much more that discussed here. Although the first edition of is not very recent, it is still considered one of the best textbooks on graph and network flow algorithms. *[NW06]* gives a more modern treatment of both theory and algorithmic aspects of linear optimization.

CHAPTER THREE

CONIC QUADRATIC OPTIMIZATION

3.1 Introduction

This chapter extends the notion of linear optimization with *quadratic cones*; conic quadratic optimization is a straightforward generalization of linear optimization, in the sense that we optimize a linear function with linear (in)equalities with variables belonging to one or more (rotated) quadratic cones. In general we also allow some of the variables to be linear variables as long as some of the variables belong to a quadratic cone.

We discuss the basic concept of quadratic cones, and demonstrate the surprisingly large flexibility of conic quadratic modeling with several examples of (non-trivial) convex functions or sets that be represented using quadratic cones. These convex sets can then be combined arbitrarily to form different conic quadratic optimization problems.

We finally extend the duality theory and infeasibility analysis from linear to conic optimization, and discuss infeasibility of conic quadratic optimization problems.

3.1.1 Quadratic cones

We define an n-dimensional quadratic cone as

$$\mathcal{Q}^n = \left\{ x \in \mathbb{R}^n \mid x_1 \geq \sqrt{x_2^2 + x_3^2 + \cdots + x_n^2} \right\}. \tag{3.1}$$

The geometric interpretation of a quadratic (or second-order) cone is shown in Fig. 3.1 for a cone with three variables, and illustrates how the exterior of the cone resembles an ice-cream cone.

The 1-dimension quadratic cone simply implies standard nonnegativity $x_1 \geq 0$.

A set S is called a *convex cone* if for any $x \in S$ we have $\alpha x \in S$, $\forall \alpha \geq 0$. From the definition (3.1) it is clear that if $x \in \mathcal{Q}^n$ then obviously $\alpha x \in \mathcal{Q}^n$, $\forall \alpha \geq 0$, which justifies the notion *quadratic cone*.

3.1.2 Rotated quadratic cones

An $n-$dimensional *rotated quadratic cone* is defined as

$$\mathcal{Q}_r^n = \left\{ x \in \mathbb{R}^n \mid 2x_1x_2 \geq x_3^2 + \cdots + x_n^2,\ x_1, x_2 \geq 0 \right\}. \tag{3.2}$$

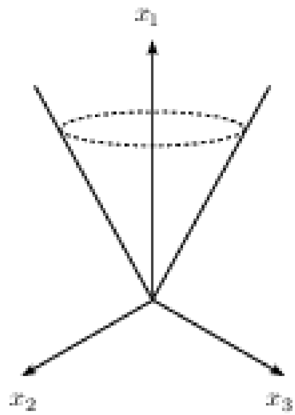


Fig. 3.1: A quadratic or second-order cone satisfying $x_1 \geq \sqrt{x_2^2 + x_3^2}$.

As the name indicates, there is simple relationship between quadratic and rotated quadratic cones. Define an orthogonal transformation

$$T_n := \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & I_{n-2} \end{bmatrix}. \tag{3.3}$$

Then it is easy to verify that

$$x \in \mathcal{Q}^n \quad \Longleftrightarrow \quad (T_n x) \in \mathcal{Q}_r^n,$$

and since T is orthogonal we call $\mathcal{Q}_r$ a rotated cone; the transformation corresponds to a rotation of $\pi/4$ of the (x_1, x_2) axis. For example if $x \in \mathcal{Q}^3$ and

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}(x_1 + x_2) \\ \frac{1}{\sqrt{2}}(x_1 - x_2) \\ x_3 \end{bmatrix}$$

then

$$2z_1 z_2 \geq z_3^2, \; z_1, z_2 \geq 0 \quad \Longrightarrow \quad (x_1^2 - x_2^2) \geq x_3^2, \; x_1 \geq 0,$$

and similarly (by interchanging roles of x and z) we see that

$$x_1^2 \geq x_2^2 + x_3^2, \; x_1 \geq 0 \quad \Longrightarrow \quad 2x_1 x_2 \geq x_3, \; x_1, x_2 \geq 0.$$

Thus, one could argue that we only need quadratic cones, but there are many examples of functions where using an explicit rotated quadratic conic formulation is more natural; in Sec. 3.2 we discuss many examples involving both quadratic cones and rotated quadratic cones.

3.2 Conic quadratic modeling

In the following we describe several convex sets that can be modeled using conic quadratic formulations. We describe the convex sets in their simplest embodiment, which we interpret as *conic* *quadratic building blocks*; it is then straightforward to combine those in arbitrary intersections and affine mappings to model complex sets. For example we will show that the set

$$\frac{1}{x} \leq t, \quad x \geq 0$$

can be represented using quadratic cones. Sets that can be represented using quadratic cones are called *conic quadratic representable* sets.

3.2.1 Absolute values

In Sec. 2.2.2 we saw how to model $|x| \leq t$ using two linear inequalities, but in fact the epigraph of the absolute value is just the definition of a two-dimensional quadratic cone, i.e.,

$$|x| \leq t \quad \iff \quad (t, x) \in \mathcal{Q}^2.$$

3.2.2 Euclidean norms

The Euclidean norm of $x \in \mathbb{R}^n$,

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

essentially defines the quadratic cone, i.e.,

$$\|x\|_2 \leq t \quad \iff \quad (t, x) \in \mathcal{Q}^{n+1}.$$

3.2.3 Squared Euclidean norms

The epigraph of the squared Euclidean norm can be described as the intersection of a rotated quadratic cone with an affine hyperplane,

$$\|x\|_2^2 \leq t \quad \iff \quad (1/2, t, x) \in \mathcal{Q}_r^{n+2}.$$

3.2.4 Convex quadratic sets

Assume $Q \in \mathbb{R}^{n \times n}$ is a symmetric positive semidefinite matrix. The convex inequality

$$(1/2)x^T Q x + c^T x + r \leq 0$$

may be rewritten as

$$\begin{array}{rcl} t + c^T x + r & = & 0, \\ x^T Q x & \leq & 2t. \end{array} \tag{3.4}$$

Since Q is symmetric positive semidefinite the epigraph

$$x^T Q x \leq 2t \tag{3.5}$$

is a convex set and there exists a matrix $F \in \mathbb{R}^{k\times n}$ such that

$$Q = F^T F, \tag{3.6}$$

(see Chap. *4* for properties of semidefinite matrices); for instance F could be the Cholesky factorization of Q. We then have an equivalent characterization of (3.5) as

$$\begin{array}{rcl} Fx - y & = & 0, \\ s & = & 1, \\ \|y\|^2 & \leq & 2st, \end{array}$$

or more succinctly; if $Q = F^T F$ then

$$(1/2)x^T Q x \leq t \quad \Longleftrightarrow \quad (t, 1, Fx) \in \mathcal{Q}_r^{2+k}.$$

Frequently Q has the structure

$$Q = I + F^T F$$

where I is the identity matrix, so

$$x^T Q x = x^T x + x^T F^T F x$$

and hence

$$\begin{array}{rcl} Fx - y & = & 0, \\ f + h & = & t \\ \|x\|^2 & \leq & 2ef, \quad e = 1, \\ \|y\|^2 & \leq & 2gh, \quad g = 1. \end{array}$$

In other words,

$$(f, 1, x) \in \mathcal{Q}_r^{2+n}, \qquad (h, 1, Fx) \in \mathcal{Q}_r^{2+k}, \qquad f + h = t$$

is a conic quadratic representation of (3.5).

3.2.5 Second-order cones

A second-order cone is occasionally specified as

$$\|Ax + b\|_2 \leq c^T x + d \tag{3.7}$$

where $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^n$. The formulation (3.7)) is equivalent to

$$(c^T x + d, Ax + b) \in \mathcal{Q}^{m+1}$$

or equivalently

$$\begin{array}{rcl} s & = & Ax + b, \\ t & = & c^T x + d, \\ (t, s) & \in & \mathcal{Q}^{m+1}. \end{array} \tag{3.8}$$

As will be explained later, we refer to (3.7) as the dual form and (3.8) as the primal form, respectively. An alternative characterization of (3.7) is

$$\|Ax + b\|_2^2 - (c^T x + d)^2 \leq 0, \qquad c^T x + d \geq 0 \tag{3.9}$$

which shows that certain quadratic inequalities are conic quadratic representable. The next section shows such an example.

3.2.6 Simple polynomial sets

We next show how to represent some frequently occuring convex sets represented by polynomials. The result is stated in the following lemma:

Lemma 3.1. *The following five propositions are true.*

1. $\{(t,x) \mid t \leq \sqrt{x},\ x \geq 0\} = \left\{(t,x) \mid (x, 1/2, t) \in \mathcal{Q}_r^3\right\}$.
2. $\left\{(t,x) \mid t \geq \frac{1}{x},\ x \geq 0\right\} = \left\{(t,x) \mid (x, t, \sqrt{2}) \in \mathcal{Q}_r^3\right\}$.
3. $\left\{(t,x) \mid t \geq x^{3/2},\ x \geq 0\right\} = \left\{(t,x) \mid (s,t,x), (x, 1/8, s) \in \mathcal{Q}_r^3\right\}$.
4. $\left\{(t,x) \mid t \geq x^{5/3},\ x \geq 0\right\} = \left\{(t,x) \mid (s,t,x), (1/8, z, s), (s, x, z) \in \mathcal{Q}_r^3\right\}$.
5. $\left\{(t,x) \mid t \geq \frac{1}{x^2},\ x \geq 0\right\} = \left\{(t,x) \mid (t, 1/2, s), (x, s, \sqrt{2}) \in \mathcal{Q}_r^3\right\}$.
6. $\left\{(t,x,y) \mid t \geq \frac{|x|^3}{y^2}, y \geq 0\right\} = \left\{(t,x,y) \mid (z,x) \in \mathcal{Q}^2, (\frac{y}{2}, s, z), (\frac{t}{2}, z, s) \in \mathcal{Q}_r^3\right\}$.

Proof. Proposition *(i)* follows from

$$\begin{array}{ll} & (x, \frac{1}{2}, t) \in \mathcal{Q}_r^3 \\ \Leftrightarrow & x \geq t^2,\ x \geq 0 \\ \Leftrightarrow & \sqrt{x} \geq t,\ x \geq 0. \end{array}$$

Proposition *(ii)* follows from

$$\begin{array}{ll} & (x, t, \sqrt{2}) \in \mathcal{Q}_r^3 \\ \Leftrightarrow & 2xt \geq 2,\ x,t \geq 0 \\ \Leftrightarrow & t \geq \frac{1}{x},\ x \geq 0. \end{array}$$

Proposition *(iii)* follows from

$$\begin{array}{ll} & (s,t,x), (x, \frac{1}{8}, s) \in \mathcal{Q}_r^3 \\ \Leftrightarrow & 2st \geq x^2,\ \frac{1}{4}x \geq s^2,\ s,t,x \geq 0 \\ \Leftrightarrow & \sqrt{x}t \geq x^2,\ t,x \geq 0 \\ \Leftrightarrow & t \geq x^{3/2},\ x \geq 0. \end{array}$$

Proposition *(iv)* follows from

$$\begin{array}{ll} & (s,t,x), (1/8, z, s), (s, x, z) \in \mathcal{Q}_r^3 \\ \Leftrightarrow & \frac{1}{4}z \geq s^2,\ 2sx \geq z^2,\ 2st \geq x^2,\ s,t,x \geq 0 \\ \Leftrightarrow & 2sx \geq (4s^2)^2,\ 2st \geq x^2,\ s,t,x \geq 0 \\ \Leftrightarrow & x \geq 8s^3,\ 2st \geq x^2,\ x,s \geq 0 \\ \Leftrightarrow & x^{1/3}t \geq x^2,\ x \geq 0 \\ \Leftrightarrow & t \geq x^{5/3},\ x \geq 0. \end{array}$$

Proposition *(v)* follows from

$$\begin{array}{ll} & (t, \frac{1}{2}, s), (x, s, \sqrt{2}) \in \mathcal{Q}_r^3 \\ \Leftrightarrow & t \geq s^2,\ 2xs \geq 2,\ t,x \geq 0 \\ \Leftrightarrow & t \geq \frac{1}{x^2}. \end{array}$$

Finally, proposition *(vi)* follows from

$$\begin{array}{ll} & (z,x) \in \mathcal{Q}^2, (\frac{y}{2}, s, z), (\frac{t}{2}, z, s) \in \mathcal{Q}_r^3 \\ \Leftrightarrow & z \geq |x|,\ ys \geq z^2,\ zt \geq s^2,\ z,y,s,t \geq 0 \\ \Leftrightarrow & z \geq |x|,\ zt \geq \frac{z^4}{y^2},\ z,y,t \geq 0 \\ \Leftrightarrow & t \geq \frac{|x|^3}{y^2},\ y \geq 0. \end{array}$$

In the above proposition the terms $x^{3/2}$ and $x^{5/3}$ appeared. Those terms are special cases of

$$t \geq x^{\frac{2k-1}{k}},\ x \geq 0 \tag{3.10}$$

defined for $k = 2, 3, 4, 5, ...$ (3.10) is identical to

$$tx^{1/k} \geq x^2,\ x \geq 0$$

which implies

$$\begin{array}{rcll} 2tz & \geq & x^2, & x, z \geq 0, \\ x^{1/k} & \geq & 2z. & \end{array}$$

The latter can be rewritten as

$$\begin{array}{rcll} 2tz & \geq & x' & x, s \geq 0, \\ x & \geq & 2^k z^k & \end{array}$$

for which it is easy to build a conic quadratic representation. Therefore we will leave it as an exercise for reader.

3.2.7 Geometric mean

Closely related is the hypograph of the geometric mean of nonnegative variables x_1 and x_2,

$$K_{\text{gm}}^2 = \{(x_1, x_2, t) \mid \sqrt{x_1 x_2} \geq t,\ x_1, x_2 \geq 0\}.$$

This corresponds to a scaled rotated quadratic cone

$$K_{\text{gm}}^2 = \{(x_1, x_2, t) \mid (x_1, x_2, \sqrt{2}t) \in \mathcal{Q}_r^3\}.$$

More generally, we can obtain a conic quadratic representation of the hypograph

$$K_{\text{gm}}^n = \left\{(x, t) \in \mathbb{R}^{n+1} \mid (x_1 x_2 \cdots x_n)^{1/n} \geq t, x \geq 0\right\}.$$

Initially let us assume that $n = 2^l$ for an integer $l \geq 1$. In particular, let us assume $l = 3$, i.e., we seek a conic quadratic representation of

$$t \leq (x_1 x_2 \cdots x_8)^{1/8}, \quad x \geq 0. \tag{3.11}$$

If we introduce the cone relationships

$$(x_1, x_2, y_1), (x_3, x_4, y_2), (x_5, x_6, y_3), (x_7, x_8, y_4) \in \mathcal{Q}_r^3, \tag{3.12}$$

then

$$2x_1 x_2 \geq y_1^2$$

and so forth, which is obviously equivalent to

$$x_1 x_2 \geq (1/2) y_1^2.$$

This leads to a characterization

$$((1/2)y_1^2 (1/2)y_2^2 (1/2)y_3^2 (1/2)y_4^2)^{1/8} \leq (x_1 x_2 \cdots x_8)^{1/8},$$

or equivalently

$$\frac{1}{\sqrt{2}}(y_1y_2y_3y_4)^{1/4} \leq (x_1x_2\cdots x_8)^{1/8}.$$

Hence, by introducing 4 3-dimensional rotated quadratic cones, we have obtained a simpler problem with 4 y variables instead of 8 x variables. Clearly, we can apply the same idea to the reduced problem. To that end, introduce the inequalities

$$(y_1, y_2, z_1), (y_3, y_4, z_2) \in \mathcal{Q}_r^3,$$

implying that

$$\frac{1}{\sqrt{2}}(z_1z_2)^{1/2} \leq (y_1y_2y_3y_4)^{1/4}.$$

Finally, if we introduce the conic relationship

$$(z_1, z_2, w_1) \in \mathcal{Q}_r^3$$

we get

$$w_1 \leq \sqrt{2}(z_1z_2)^{1/2} \leq \sqrt{4}(y_1y_2y_3y_4)^{1/4} \leq \sqrt{8}(x_1x_2\cdots x_8)^{1/8}.$$

Therefore

$$\begin{array}{rcl} (x_1,x_2,y_1),(x_3,x_4,y_2),(x_5,x_6,y_3),(x_7,x_8,y_4) & \in & \mathcal{Q}_r^3,\\ (y_1,y_2,z_1),(y_3,y_4,z_2) & \in & \mathcal{Q}_r^3,\\ (z_1,z_2,w_1) & \in & \mathcal{Q}_r^3,\\ w_1 & = & \sqrt{8}t \end{array} \tag{3.13}$$

is a representation of the set (3.11), which is by construction conic quadratic representable. The relaxation using l levels of auxiliary variables and rotated quadratic cones can represented using a tree-structure as shown shown in Fig. 3.2.

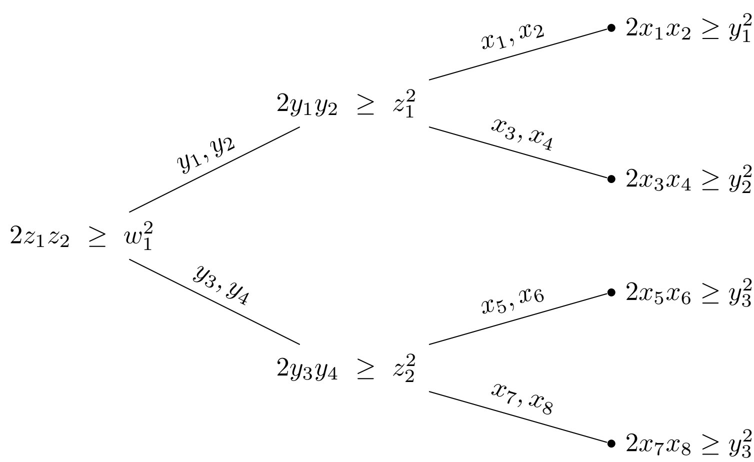


Fig. 3.2: Tree-structure of the relaxations as in (3.13): on the edges we report the additional variables, on the nodes the additional contraints.

Next let us assume that n is not a power of two let, for example $n = 6$. We then wish to characterize the set

$$t \leq (x_1x_2x_3x_4x_5x_6)^{1/6}, \quad x \geq 0.$$

which is obviously equivalent to

$$t \le (x_1 x_2 x_3 x_4 x_5 x_6)^{1/8}, \quad x_7 = x_8 = t, \quad x \ge 0.$$

In other words, we can reuse the result (3.11) if we add simple equality constraints

$$x_7 = x_8 = t.$$

Thus, if n is not a power of two, we take $l = \lceil \log_2 n \rceil$ and build the conic quadratic quadratic representation for that set, and we add $2^l - n$ simple equality constraints.

3.2.8 Harmonic mean

We can also consider the hypograph of the harmonic mean,

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i^{-1} \right)^{-1} \ge t, \quad x \ge 0.$$

This is a convex inequality, but not conic quadratic representable. However, the reciprocal inequality

$$\frac{1}{n} \sum_{i=1}^{n} x_i^{-1} \le y, \quad x \ge 0. \tag{3.14}$$

with $y = 1/t$ can be characterized as

$$x_i z_i \ge 1, \quad \sum_{i=1}^{n} z_i = ny.$$

In other words, the set (3.14) corresponding to the epigraph of the reciprocal harmonic mean of x can be described using an intersection of rotated quadratic cones an affine hyperplanes

$$2 x_i z_i \ge w_i^2, \quad x_i, z_i \ge 0, \quad w_i = \sqrt{2}, \quad \sum_{i=1}^{n} z_i = ny,$$

or $(x_i, z_i, \sqrt{2}) \in \mathcal{Q}_r^3$, $e^T z = ny$.

3.2.9 Convex increasing rational powers

We can extend the formulations in Sec. *3.2.7* to describe the epigraph

$$K^{p/q} = \{(x,t) \mid x^{p/q} \le t, \ x \ge 0\}$$

for any rational convex power $p/q \ge 1$, $p, q \in \mathbb{Z}_+$. For example, consider

$$x^{5/3} \le t, \quad x \ge 0.$$

We rewrite this as

$$x^5 \le t^3, \quad x \ge 0,$$

which can be described (using the approach in Sec. 3.2.7) as

$$0 \le x \le (y_1 y_2 y_3 y_4 y_5)^{1/5}, \quad y_1 = y_2 = y_3 = t, \quad y_4 = y_5 = 1,$$

where we note that

$$y_1 y_2 y_3 y_4 y_5 = t^3.$$

The inequality

$$0 \le x \le (y_1 y_2 y_3 y_4 y_5)^{1/5}$$

is a geometric mean inequality discussed in Sec. 3.2.7. If we let $e_k \in \mathbb{R}^k$ denote the vector of all ones, then for general $p/q \ge 1$, $p, q \in \mathbb{Z}_+$ we have

$$K^{p/q} = \{(x,t) \mid (e_q t, e_{p-q}, x) \in K_{\mathrm{gm}}^p\}.$$

3.2.10 Convex decreasing rational powers

In a similar way we can describe the epigraph

$$K^{-p/q} = \{(x,t) \mid x^{-p/q} \le t,\ x \ge 0\}$$

for any $p, q \in \mathbb{Z}_+$. For example, consider

$$x^{-5/2} \le t, \quad x \ge 0,$$

which we rewrite as

$$1 \le x^5 t^2, \quad x \ge 0,$$

or equivalently as

$$1 \le (y_1 y_2 \dots y_7)^{1/7}, \quad y_1 = \cdots = y_5 = x, \quad y_6 = y_7 = t, \quad y, t \ge 0.$$

Let $e_k \in \mathbb{R}^k$ denote the vector of all ones. For general $p, q \in \mathbb{Z}_+$ we then have

$$K^{-p/q} = \{(x,t) \mid (e_p x, e_q t, 1) \in K_{\mathrm{gm}}^{p+q}\}.$$

3.2.11 Power cones

The $(n+1)$-dimensional power cone is defined by

$$K_\alpha^{n+1} = \left\{ (x,y) \in \mathbb{R}_+^n \times \mathbb{R} \mid |y| \le \prod_{j=1}^n x_j^{\alpha_j} \right\} \tag{3.15}$$

where $\alpha > 0$ and

$$\sum_{j=1}^n \alpha_j = 1.$$

Of particular interest is the three-dimensional power cone given by

$$K_\alpha^3 = \left\{ (x,y) \in \mathbb{R}_+^2 \times \mathbb{R} \mid |y| \le x_1^{\alpha_1} x_2^{1-\alpha_1} \right\}. \tag{3.16}$$

For example, the intersection of the three-dimensional power cone with an affine hyperplane

$$|y| \leq x_1^{\alpha_1} x_2^{1-\alpha_1}, \qquad x_2 = 1$$

is equivalent to

$$|y|^{1/\alpha_1} \leq x_1,$$

i.e., we can model the epigraph of convex increasing terms ($1/\alpha_1 \geq 1$) using a three-dimensional power cone. As another example, the cone

$$|y| \leq x_1^{\frac{2}{3}} x_2^{1-\frac{2}{3}}$$

is equivalent to

$$\frac{|y|^3}{x_1^2} \leq x_2.$$

Assume next that $\alpha_1 = p/q$ where p and q are positive integers and $p \leq q$. Then

$$K^3_{(\frac{p}{q},1-\frac{p}{q})} = \left\{(x,y) \in \mathbb{R}^2_+ \times \mathbb{R} \mid |y| \leq x_1^{\frac{p}{q}} x_2^{(1-\frac{p}{q})}\right\}. \tag{3.17}$$

Now

$$|y| \leq x_1^{\frac{p}{q}} x_2^{(1-\frac{p}{q})}$$

is equivalent to

$$|y| \leq (x_1^p x_2^{q-p})^{\frac{1}{q}}, \tag{3.18}$$

so by introducing additional z variables and the constraints

$$z_1 = z_2 = \ldots = z_{q-p} = x_1, \qquad z_{q-p+1} = z_{q-p+2} = \ldots = z_q = x_2$$

we can rewrite (3.18) as

$$|y| \leq (z_1 z_2 \ldots z_q)^{\frac{1}{p}} \tag{3.19}$$

which essentially is the geometric mean inequality discussed above.

We next consider the $n+1$ dimensional power cone with rational powers,

$$K^{n+1}_\alpha = \left\{(x,y) \in \mathbb{R}^n_+ \times \mathbb{R} \mid |y| \leq \prod_{j=1}^n x_j^{p_j/q_j}\right\} \tag{3.20}$$

where p_j, q_j are integers satisfying $0 < p_j \leq q_j$ and $\sum_{j=1}^n p_j/q_j = 1$. To motivate the general procedure, we consider a simple example.

Consider the 4 dimensional power cone

$$K^4_{(\frac{1}{4},\frac{3}{8},\frac{3}{8})} = \left\{(x,y) \in \mathbb{R}^3_+ \times \mathbb{R} \mid |y| \leq (x_1^{\frac{1}{4}} x_2^{\frac{3}{8}} x_3^{\frac{3}{8}})\right\}.$$

An equivalent representation is given by the geometric mean inequality

$$|y| \leq (z_1 z_2 \ldots z_8)^{1/8}$$

with

$$z_1 = z_2 = x_1, \qquad z_3 = z_4 = z_5 = x_2, \qquad z_6 = z_7 = z_8 = x_3.$$

In the general case, let

$$\beta = \mathrm{lcm}(q_1, q_2, \ldots, q_n)$$

denote the least common multiple of $\{q_i\}$. Then

$$K_\alpha^{n+1} = \left\{ (x,y) \in \mathbb{R}_+^n \times \mathbb{R} \mid |y| \leq \left(\prod_{j=1}^n x_j^{\frac{\beta p_j}{q_j}} \right)^{\frac{1}{\beta}} \right\},$$

where we note that $\frac{\beta p_j}{q_j}$ are integers and $\sum_{j=1}^n \frac{\beta p_j}{q_j} = \beta$. If we define

$$s_j = \sum_{k=1}^j \frac{\beta p_k}{q_k}, \quad j = 1, \ldots, n-1,$$

we can then reformulate (3.20) as

$$|y| \leq (z_1 z_2 \ldots z_\beta)^{\frac{1}{q}}$$
$$z_1 = \ldots = z_{s_1} = x_1, \quad z_{s_1+1} = \ldots = z_{s_2} = x_2, \quad \ldots \quad z_{s_{n-1}+1} = \ldots = z_\beta = x_n.$$

3.2.12 Quadratic forms with one negative eigenvalue

Assume that $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix with exactly one negative eigenvalue, i.e., A has a spectral factorization (i.e., eigenvalue decomposition)

$$A = Q \Lambda Q^T = -\alpha_1 q_1 q_1^T + \sum_{i=2}^n \alpha_i q_i q_i^T$$

where $Q^T Q = I$, $\Lambda = \mathbf{diag}(-\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\alpha_i \geq 0$, $\forall i$. Then

$$x^T A x \leq 0$$

is equivalent to

$$\sum_{j=2}^n \alpha_j (q_j^T x)^2 \leq \alpha_1 (q_1^T x)^2. \tag{3.21}$$

Suppose $q_1^T x \geq 0$. We can characterize (3.21) in different ways, e.g., as

$$(\sqrt{\alpha_1/2} q_1^T x, \sqrt{\alpha_1/2} q_1^T x, \sqrt{\alpha_2} q_2^T x, \ldots, \sqrt{\alpha_n} q_n^T x) \in \mathcal{Q}_r^{n+1} \tag{3.22}$$

or as

$$(\sqrt{\alpha_1} q_1^T x, \sqrt{\alpha_2} q_2^T x, \ldots, \sqrt{\alpha_n} q_n^T x) \in \mathcal{Q}^n. \tag{3.23}$$

The latter equivalence follows directly by using the orthogonal transformation T_{n+1} on (3.22), i.e.,

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & & & \\ 1/\sqrt{2} & -1/\sqrt{2} & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\alpha_1/2} q_1^T x \\ \sqrt{\alpha_1/2} q_1^T x \\ \sqrt{\alpha_2} q_2^T x \\ \vdots \\ \sqrt{\alpha_n} q_n^T x \end{bmatrix} = \begin{bmatrix} \sqrt{\alpha_1} q_1^T x \\ 0 \\ \sqrt{\alpha_2} q_2^T x \\ \vdots \\ \sqrt{\alpha_n} q_n^T x \end{bmatrix}.$$

3.2.13 Ellipsoidal sets

The set

$$\mathcal{E} = \{x \in \mathbb{R}^n \mid \|P(x-c)\|_2 \leq 1\}$$

describes an ellipsoid centered at c. It has a natural conic quadratic representation, i.e., $x \in \mathcal{E}$ if and only if

$$y = P(x-c), \quad (t, y) \in \mathcal{Q}^{n+1}, \quad t = 1.$$

Suppose P is nonsingular. We then get an alternatively characterization

$$\mathcal{E} = \{x \in \mathbb{R}^n \mid x = P^{-1}y + c, \ \|y\|_2 \leq 1\}.$$

Depending on the context one characterization may be more useful than the other.

3.3 Conic quadratic optimization examples

In this section we give different instructive examples of conic quadratic optimization problems formed using the formulations from Sec. 3.2.

3.3.1 Quadratically constrained quadratic optimization

A general convex quadratically constrained quadratic optimization problem can be written as

$$\begin{array}{ll} \text{minimize} & (1/2)x^T Q_0 x + c_0^T x + r_0 \\ \text{subject to} & (1/2)x^T Q_i x + c_i^T x + r_i \leq 0, \quad i = 1, \ldots, p, \end{array} \tag{3.24}$$

where all Q_i are symmetric positive. Let

$$Q_i = F_i^T F_i, \quad i = 0, \ldots, p,$$

where $F_i \in \mathbb{R}^{k_i \times n}$. Using the formulations in Sec. 3.2.4 we then get an equivalent conic quadratic problem

$$\begin{array}{lll} \text{minimize} & t_0 + c_0^T x + r_0 & \\ \text{subject to} & t_i + c_i^T x + r_i = 0, & i = 1, \ldots, p, \\ & (t_i, 1, F_i x) \in \mathcal{Q}_r^{k_i+2}, & i = 0, \ldots, p. \end{array} \tag{3.25}$$

Assume next that Q_i is a rank 1 matrix, i.e., F_i has 1 row and n columns. Storing Q_i requires about $n^2/2$ space whereas storing F_i then only requires n space. Moreover, the amount of work required to evaluate

$$x^T Q_i x$$

is proportional to n^2 whereas the work required to evaluate

$$x^T F_i^T F_i x = \|F_i x\|^2$$

is proportional to n only. In other words, if Q_i have low rank, then (3.25) will require much less space to solve than (3.24). This fact usually also translates into much faster solution times.

3.3.2 Robust optimization with ellipsoidal uncertainties

Often in robust optimization some of the parameters in the model are assumed to be unknown, but we assume that the unknown parameters belong to a simple set describing the uncertainty. For example, for a standard linear optimization problem (2.3.1) we may wish to find a robust solution for all vectors c in an ellipsoid

$$\mathcal{E} = \{c \in \mathbb{R}^n \mid c = Fy + g,\ \|y\|_2 \leq 1\}.$$

A common approach is then to optimize for the *worst-case* realization of c, i.e., we get a robust version

$$\begin{array}{ll} \text{minimize} & \sup_{c\in\mathcal{E}} c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \tag{3.26}$$

The worst-case objective can be evaluated as

$$\sup_{c\in\mathcal{E}} c^T x = g^T x + \sup_{\|y\|_2\leq 1} y^T F^T x = g^T x + \|F^T x\|_2$$

where we used that $\sup_{\|u\|_2\leq 1} v^T u = (v^T v)/\|v\|_2 = \|v\|_2$. Thus the robust problem (3.26) is equivalent to

$$\begin{array}{ll} \text{minimize} & g^T x + \|F^T x\|_2 \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array}$$

which can be posed as a conic quadratic problem

$$\begin{array}{ll} \text{minimize} & g^T x + t \\ \text{subject to} & Ax = b \\ & (t, F^T x) \in \mathcal{Q}^{n+1} \\ & x \geq 0. \end{array} \tag{3.27}$$

3.3.3 Markowitz portfolio optimization

In classical Markowitz portfolio optimization we consider investment in n stocks or assets held over a period of time. Let x_i denote the amount we invest in asset i, and assume a stochastic model where the return of the assets is a random variable r with known mean

$$\mu = \mathbf{E}r$$

and covariance

$$\Sigma = \mathbf{E}(r-\mu)(r-\mu)^T.$$

The return of our investment is also a random variable $y = r^T x$ with mean (or expected return)

$$\mathbf{E}y = \mu^T x$$

and variance (or risk)

$$(y - \mathbf{E}y)^2 = x^T \Sigma x.$$

We then wish to rebalance our portfolio to achieve a compromise between risk and expected return, e.g., we can maximize the expected return given an upper bound γ on the tolerable risk and a constraint that our total investment is fixed,

$$\begin{array}{ll} \text{maximize} & \mu^T x \\ \text{subject to} & x^T \Sigma x \leq \gamma \\ & e^T x = 1 \\ & x \geq 0. \end{array}$$

Suppose we factor $\Sigma = GG^T$ (e.g., using a Cholesky or a eigenvalue decomposition). We then get a conic formulation

$$\begin{array}{ll} \text{maximize} & \mu^T x \\ \text{subject to} & (\sqrt{\gamma}, G^T x) \in \mathcal{Q}^{n+1} \\ & e^T x = 1 \\ & x \geq 0. \end{array}$$

In practice both the average return and covariance are estimated using historical data. A recent trend is then to formulate a robust version of the portfolio optimization problem to combat the inherent uncertainty in those estimates, e.g., we can constrain μ to an ellipsoidal uncertainty set as in Sec. *3.3.2*.

3.4 Duality in conic quadratic optimization

To discuss conic quadratic duality, we consider a primal problem

$$\begin{array}{ll} \text{minimize} & \langle c^l, x^l \rangle + \sum_{j=1}^{n_q} \langle c_j^q, x_j^q \rangle \\ \text{subject to} & A^l x^l + \sum_{j=1}^{n_q} A_j^q x_j^q = b \\ & x^l \in \mathbb{R}_+^{n_l}, \ x_j^q \in \mathcal{Q}^{n_j}, \end{array} \tag{3.28}$$

where $c^l \in \mathbb{R}^{n_l}$, $c_j^q \in \mathbb{R}^{n_j}$, $A^l \in \mathbb{R}^{m \times n^l}$, $A_j^q \in \mathbb{R}^{m \times n_j}$ and $b \in \mathbb{R}^m$, i.e., a problem with both a linear cone and n_q quadratic cones. We can also write (3.28) more compactly as

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{C}, \end{array} \tag{3.29}$$

with

$$x = \begin{pmatrix} x^l \\ x_1^q \\ \vdots \\ x_{n_q}^q \end{pmatrix}, \quad c = \begin{pmatrix} c^l \\ c_1^q \\ \vdots \\ c_{n_q}^q \end{pmatrix}, \quad A = \begin{pmatrix} A^l & A_1^q & \dots & A_{n_q}^q \end{pmatrix},$$

for the cone $\mathcal{C} = \mathbb{R}_+^l \times \mathcal{Q}^{n_1} \times \cdots \times \mathcal{Q}^{n_q}$. We note that problem (3.29) resembles a standard linear optimization, except for a more general cone. The Lagrangian function is

$$L(x, y, s) = \sum c^T x - y^T (Ax - b) - x^T s.$$

For linear optimization the choice of nonnegative Lagrange multipliers $s \geq 0$ ensures that $x^T s \geq 0$, $\forall x \geq 0$ so the Lagrangian function provides a lower bound on the primal problem. For conic quadratic optimization the choice of Lagrange multipliers is more involved, and we need the notion of a *dual cone*

$$(\mathcal{Q}^n)^* = \{v \in \mathbb{R}^n \mid u^T v \geq 0,\ \forall u \in \mathcal{Q}^n\}.$$

Then for $s \in (\mathcal{Q}^n)^*$ we similarly have $x^T s \geq 0, \forall x \in \mathcal{Q}^n$, so the Lagrange function becomes a global lower bound.

The quadractic cone is self-dual, $(\mathcal{Q}^n)^* = \mathcal{Q}^n$.

Assume first that $u_1 \geq \|u_{2:n}\|$ and $v_1 \geq \|v_{2:n}\|$. Then

$$u^T v = u_1 v_1 + u_{2:n}^T v_{2:n} \geq u_1 v_1 - \|u_{2:n}\| \cdot \|v_{2:n}\| \geq 0.$$

Conversely, assume that $v_1 < \|v_{2:n}\|$. Then $u := (1, -v_{2:n}/\|v_{2:n}\|) \in Q^n$ satisfies

$$u^T v = v_1 - \|v_{2:n}\| < 0,$$

i.e., $v \notin (\mathcal{Q}^n)^*$.

The notion of a dual cone allows us to treat (3.29) as a linear optimization problem, where the dual variables belong to the dual cone, i.e., we have a dual problem

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \in \mathcal{C}^*, \end{array} \tag{3.30}$$

with a dual cone

$$\mathcal{C}^* = \mathcal{C} = \mathbb{R}^l_+ \times \mathcal{Q}^{n_1} \times \cdots \times \mathcal{Q}^{n_q}.$$

Weak duality follows directly by taking inner product of x and the dual constraint,

$$c^T x - x^T A^T y = c^T x - b^T y = x^T s \geq 0.$$

3.4.1 Primal versus dual form

Let us consider the dual formulation (3.30) without linear terms. If we eliminate the dual variable s, we can rewrite (3.30) as

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & (c_j^q - (A_j^q)^T y) \in \mathcal{Q}^{n_j}, \quad j = 1, \ldots, n_q, \end{array}$$

or equivalently

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & \|(c_j^q)_{2:n_j} - ((A_j^q)_{2:n_j,:})^T y\|_2 \leq (c_j^q)_1 - ((A_j^q)_{1,:})^T y, \quad j = 1, \ldots, n_j, \end{array}$$

where the dual constraints

$$\|(c_j^q)_{2:n_j} - ((A_j^q)_{2:n_j,:})^T y\|_2 \leq (c_j^q)_1 - ((A_j^q)_{1,:})^T y$$

correspond to the so-called *dual form* conic constraints in Sec. *3.2.5*. If the matrix

$$A = \left(\begin{array}{ccc} A_1^q & \ldots & A_{n_q}^q \end{array} \right),$$

has more columns than rows then usually it is more efficient to solve the primal problem (3.29), and if the opposite is the case, then it is usually more efficient to solve the dual problem (3.30).

3.4.2 Examples of bad duality states

The issue of strong duality is more complicated than for linear programming, for example the primal (or dual) optimal values may be finite but unattained, as illustrated in the following example.

Example 3.1 (Positive duality gap). The following example is from *[ART02]*: consider the problem

$$\begin{array}{ll} \text{minimize} & x_1 - x_2 \\ \text{subject to} & x_3 = 1 \\ & x_1 \geq \sqrt{x_2^2 + x_3^2}, \end{array} \tag{3.31}$$

with feasible set $\{x_1 \geq \sqrt{1+x_2^2},\ x_3 = 1\}$. The optimal objective value is $p^\star = 0$ for $(x_1, x_2) \to \infty$, and we say that the objective is unattained. The dual problem is

$$\begin{array}{ll} \text{maximize} & y \\ \text{subject to} & 1 \geq \sqrt{1+y^2}, \end{array} \tag{3.32}$$

with feasible set $\{y = 0\}$ and optimal value $d^\star = 0$, so a positive duality gap exists for all bounded (x, y).

It is also possible to have a positive duality gap even when both the primal and dual optimal objectives are attained.

Example 3.2 (Duality gap). We consider the problem

$$\begin{array}{lrcl} \text{minimize} & & x_3 & \\ \text{subject to} & x_1 + x_2 + u_1 + u_2 & = & 0 \\ & x_3 + u_2 & = & 1 \\ & \sqrt{x_2^2 + x_3^2} & \leq & x_1 \\ & \sqrt{u_2^2} & \leq & u_1, \end{array} \tag{3.33}$$

with two quadratic cones

$x \in \mathcal{Q}^3,\ u \in \mathcal{Q}^2$. It follows from the

inequality constraints that

$$x_1 \geq |x_2|, \qquad u_1 \geq |u_2|$$

and therefore $x_1 + x_2 \geq 0$ and $u_1 + u_2 \geq 0$. But then $x_1 + x_2 = 0$ and $u_1 + u_2 = 0$, which combined with $x_3 = 1 - u_2$ gives a simplified problem

$$\begin{array}{lrcl} \text{minimize} & & 1 - u_2 & \\ & \sqrt{x_1^2 + (1-u_2)^2} & \leq & x_1, \end{array}$$

with feasible set $\{u_2 = 1,\ x_1 \geq 0\}$ and optimal value $p^\star = 0$. The dual problem of (3.33) is

$$\begin{array}{ll} \text{maximize} & y_2 \\ \text{subject to} & y_1 \geq \sqrt{y_1^2 + (1-y_2)^2} \\ & y_1 \geq \sqrt{(y_1 - y_2)^2}, \end{array} \tag{3.34}$$

with feasible set $\{y_2 = 1,\ y_1 \geq \frac{1}{2}\}$ and the optimal value is $d^\star = -1$. For this example, both the primal and dual optimal values are attained, but we have positive duality gap $p^\star - d^\star = 1$.

To ensure strong duality for conic quadratic optimization, we need an additional regularity assumption. Consider the primal feasible set for (3.29),

$$\mathcal{F}_p = \{x = (x^l, x_1^q, \dots, x_{n_q}^q) \mid Ax = b,\ x^l \geq 0,\ x_j^q \in \mathcal{Q}^{n_j}\}.$$

and dual feasible set for (3.30),

$$\mathcal{F}_d = \{y \in \mathbb{R}^m \mid c - A^T y = (s^l, s_1^q, \dots, s_{n_q}^q),\ s^l \geq 0,\ s_j^q \in \mathcal{Q}^{n_j}\},$$

respectively. If

$$\exists x \in \mathcal{F}_p \ :\ (x_j^q)_1 > \|(x_j^q)_{2:n_j}\|,\ j = 1, \dots, n_q,$$

or

$$\exists y \in \mathcal{F}_p \ :\ c - A^T y = (s^l, s_1^q, \dots, s_{n_q}^q),\ (s_j^q)_1 > \|(s_j^q)_{2:n_j}\|,\ j = 1, \dots, n_q,$$

then strong duality between the problems (3.29) and (3.30) holds. The additional regularity assumptions are called a *Slater constraint qualification*. In other words, strong duality holds if the *conic inequalities* in the primal or dual problem are *strictly feasible* (the proofs in App. *9.5* can be adapted to handle this case).

Thus the deficiency of positive duality gap in Example 3.4.2 is caused by the fact that neither the primal or dual problem is strictly feasible; the inequality in the first problem

$$u_1 \geq \sqrt{u_2^2}$$

is not strict for any u in the primal feasible set, and similarly the inequality

$$y_1 \geq \sqrt{y_1^2 + (1 + y_2)^2}$$

is not strict for any y in the dual feasible set.

3.5 Infeasibility in conic quadratic optimization

Since duality theory for linear and conic quadratic optimization is almost identical, the same is true for infeasibility concepts. In particular, consider a pair of primal and dual conic optimization problems (3.29) and (3.30). We then have generalized versions of Farkas Lemma.

Lemma 3.2 (Generalized Farkas)**.** *Given A and b, exactly one of the two statements are true:*

1. *There exists an $x \in \mathcal{C}$ such that $Ax = b$.*
2. *There exists a y such that $-A^T y \in \mathcal{C}$ and $b^T y > 0$.*

Farkas' lemma tells us that either the primal problem (3.29) is feasible or there exists a y such that $-A^T y \in \mathcal{C}$ and $b^T y > 0$. In other words, any y satisfying

$$-A^T y \in \mathcal{C}, \quad b^T y > 0$$

is a *certificate of primal infeasibility*. Similarly, the dual variant of Farkas' lemma states that either the dual problem (3.30) is feasible or there exists an $x \in \mathcal{C}$ such that $Ax = 0$ and $c^T x < 0$. More precisely

Lemma 3.3 (Generalized Farkas dual variant.)**.** *Given A and c, exactly one of the two statements are true:*

1. *There exists an* $x \in \mathcal{C}$ *such that* $Ax = 0$ *and* $c^T x < 0$.
2. *There exists a* y *such* $(c - A^T y) \in \mathcal{C}$.

In other words, any $x \in \mathcal{C}$ satisfying $Ax = 0$ and $c^T x < 0$ is a *certificate of dual infeasibility*.

Consider the conic quadratic problem

$$\begin{array}{lrcl} \text{minimize} & x_1 & & \\ \text{subject to} & 2x_1 - x_2 & = & 0 \\ & x_1 - s & = & 1 \\ & \sqrt{x_2^2} & \leq & x_1 \\ & s & \geq & 0, \end{array}$$

which is infeasible since $x_1 \geq 2|x_1|$ and $x_1 \geq 1$. The dual problem is

$$\begin{array}{ll} \text{maximize} & y_2 \\ \text{subject to} & \left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right) + \left(\begin{array}{c} 2y_1 + y_2 \\ -y_1 \\ -y_2 \end{array}\right) \in \mathcal{Q}^3. \end{array}$$

Any point $y = (0, t)$, $t \geq 0$ is a certificate of infeasibility since $(t, 0, -t) \in \mathcal{Q}^3$; in fact $(0, t)$ is a dual feasible direction with unbounded objective.

3.6 Bibliography

The material in this chapter is based on the paper *[LVBL98]* and the books *[BenTalN01]*, *[BV04]*. The papers *[AG03]*, *[ART03]* contain additional theoretical and algorithmic aspects.

CHAPTER

FOUR

SEMIDEFINITE OPTIMIZATION

4.1 Introduction

In this chapter we extend the conic optimization framework from Chap. *2* and *3* with symmetric positive semidefinite matrix variables.

4.1.1 Semidefinite matrices and cones

A symmetric matrix $X \in \mathcal{S}^n$ is called *symmetric positive semidefinite* if

$$z^T X z \geq 0, \quad \forall z \in \mathbb{R}^n.$$

We then define the cone of symmetric positive semidefinite matrices as

$$\mathcal{S}_+^n = \{X \in \mathcal{S}^n \mid z^T X z \geq 0, \, \forall z \in \mathbb{R}^n\}. \tag{4.1}$$

For brevity we will often use the shorter notion *semidefinite* instead of *symmetric positive semidefinite*, and we will write $X \succeq Y$ ($X \preceq Y$) as shorthand notation for $(X - Y) \in \mathcal{S}_+^n$ ($(Y - X) \in \mathcal{S}_+^n$). As inner product for semidefinite matrices, we use the standard trace inner product for general matrices, i.e.,

$$\langle A, B \rangle := \mathrm{tr}(A^T B) = \sum_{ij} a_{ij} b_{ij}.$$

It is easy to see that (4.1) indeed specifies a convex cone; it is pointed (with origin $X = 0$), and $X, Y \in \mathcal{S}_+^n$ implies that $(\alpha X + \beta Y) \in \mathcal{S}_+^n$, $\alpha, \beta \geq 0$. Let us a review a few equivalent definitions of $\mathcal{S}_+^n$. It is well-known that every symmetric matrix A has a spectral factorization

$$A = \sum_{i=1}^n \lambda_i q_i q_i^T.$$

where $q_i \in \mathbb{R}^n$ are the (orthogonal) eigenvectors and λ_i are eigenvalues of A. Using the spectral factorization of A we have

$$x^T A x = \sum_{i=1}^n \lambda_i (x^T q_i)^2,$$

which shows that $x^T A x \geq 0 \;\Leftrightarrow\; \lambda_i \geq 0,\; i = 1,\dots,n$. In other words,

$$\mathcal{S}_+^n = \{X \in \mathcal{S}^n \mid \lambda_i(X) \geq 0,\; i = 1,\dots,n\}. \tag{4.2}$$

Another useful characterization is that $A \in \mathcal{S}_+^n$ if and only if it is a *Grammian matrix* $A = V^T V$. Using the Grammian representation we have

$$x^T A x = x^T V^T V x = \|Vx\|_2^2,$$

i.e., if $A = V^T V$ then $x^T A x \geq 0, \forall x$. On the other hand, from the positive spectral factorization $A = Q\Lambda Q^T$ we have $A = V^T V$ with $V = \Lambda^{1/2} Q^T$, where $\Lambda^{1/2} = \mathbf{diag}(\sqrt{\lambda_1},\dots,\sqrt{\lambda_n})$. We thus have the equivalent characterization

$$\mathcal{S}_+^n = \{X \in \mathcal{S}_+^n \mid X = V^T V \text{ for some } V \in \mathbb{R}^{k\times n}\}. \tag{4.3}$$

In a completely analogous way we define the cone of *symmetric positive definite matrices* as

$$\begin{aligned}\mathcal{S}_{++}^n &= \{X \in \mathcal{S}^n \mid z^T X z > 0,\; \forall z \in \mathbb{R}^n\}\\ &= \{X \in \mathcal{S}^n \mid \lambda_i(X) > 0,\; i = 1,\dots,n\}\\ &= \{X \in \mathcal{S}_+^n \mid X = V^T V \text{ for some } V \in \mathbb{R}^{k\times n},\; \mathbf{rank}(V) = n\},\end{aligned}$$

and we write $X \succ Y$ ($X \prec Y$) as shorthand notation for $(X - Y) \in \mathcal{S}_{++}^n$ ($(Y - X) \in \mathcal{S}_{++}^n$).

The one dimensional cone $\mathcal{S}_+^1$ simply corresponds to $\mathbb{R}_+$. Similarly consider

$$X = \left[\begin{array}{cc} x_1 & x_3\\ x_3 & x_2\end{array}\right]$$

with determinant $\det(X) = x_1 x_2 - x_3^2 = \lambda_1\lambda_2$ and trace $\mathrm{tr}(X) = x_1 + x_2 = \lambda_1 + \lambda_2$. Therefore X has positive eigenvalues if and only if

$$x_1 x_2 \geq x_3^2, \quad x_1, x_2 \geq 0,$$

which characterizes a three dimensional scaled rotated cone, i.e.,

$$\left[\begin{array}{cc} x_1 & x_3\\ x_3 & x_2\end{array}\right] \in \mathcal{S}_+^2 \quad\Longleftrightarrow\quad (x_1, x_2, x_3/\sqrt{2}) \in \mathcal{Q}_r^2.$$

More generally we have

$$x \in \mathbb{R}_+^n \quad\Longleftrightarrow\quad \mathbf{Diag}(X) \in \mathcal{S}_+^n$$

and

$$(t, x) \in \mathcal{Q}^{n+1} \quad\Longleftrightarrow\quad \left[\begin{array}{cc} t & x^T\\ x & tI\end{array}\right] \in \mathcal{S}_+^{n+1},$$

where the latter equivalence follows immediate from Lemma 4.1.2. Thus both the linear and conic quadratic cone are embedded in the semidefinite cone. In practice, however, linear and conic quadratic should never be described using semidefinite cones, which would result in a large performance penalty by squaring the number of variables.

Example 4.1. As a more interesting example, consider the symmetric matrix

$$A(x,y,z) = \left[\begin{array}{ccc} 1 & x & y\\ x & 1 & z\\ y & z & 1\end{array}\right] \tag{4.4}$$

parametrized by (x, y, z). The set

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid A(x, y, z) \in \mathcal{S}_+^3\},$$

(shown in Fig. 4.1) is called a *spectrahedron* and is perhaps the simplest bounded semidefinite representable set, which cannot be represented using (finitely many) linear or conic quadratic cones. To gain a geometric intuition of S, we note that

$$|A(x, y, z)| = -(x^2 + y^2 + z^2 - 2xyz - 1),$$

so the boundary of S can be characterized as

$$x^2 + y^2 + z^2 - 2xyz = 1,$$

or equivalently as

$$\left[\begin{array}{c} x \\ y \end{array}\right]^T \left[\begin{array}{cc} 1 & -z \\ -z & 1 \end{array}\right] \left[\begin{array}{c} x \\ y \end{array}\right] = 1 - z^2.$$

For $z = 0$ this describes a circle in the (x, y)-plane, and for $-1 \leq z \leq 1$ it characterizes an ellipse (for a fixed z).

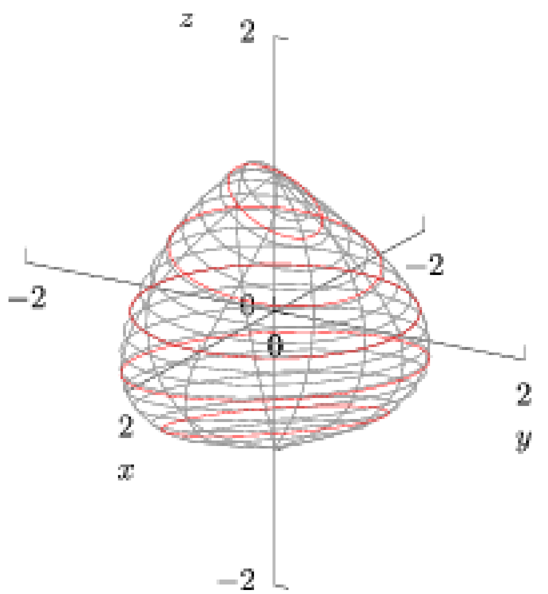


Fig. 4.1: Plot of spectrahedron $S = \{(x, y, z) \in \mathbb{R}^3 \mid A(x, y, z) \succeq 0\}$.

4.1.2 Properties of semidefinite matrices

Many useful properties of (semi)definite matrices follow directly from the definitions (4.1)-(4.3) and the definite counterparts.

- The diagonal elements of $A \in \mathcal{S}_+^n$ are nonnegative. Let e_i denote the ith standard basis vector (i.e., $[e_i]_j = 0,\ j \neq i$, $[e_i]_i = 1$). Then $A_{ii} = e_i^T A e_i$, so (4.1) implies that $A_{ii} \geq 0$.
- A block-diagonal matrix $A = \mathbf{diag}(A_1, \ldots A_p)$ is (semi)definite if and only if each diag-

onal block A_i is (semi)definite.

- Given a quadratic transformation $M := B^T AB$. Then $M \succ 0$ if and only if $A \succ 0$ and B has full rank. This follows directly from the Grammian characterization $M = (VB)^T(VB)$. For $M \succeq 0$ we only require that $A \succeq 0$. As an example, if A is (semi)definite then so is any permutation $P^T AP$.
- Any submatrix of $A \in \mathcal{S}^n_{++}$ ($A \in \mathcal{S}^n_+$) is positive (semi)definite; this follows from the Grammian characterization $A = V^T V$ since any subset of the columns of V is linearly independent.
- The inner product of positive (semi)definite matrices is positive (nonnegative). For any $A, B \in \mathcal{S}^n_{++}$ let $A = U^T U$ and $B = V^T V$ where U and V have full rank. Then

$$\langle A, B\rangle = \mathrm{tr}(U^T UV^T V) = \|UV^T\|_F^2 > 0,$$

where strict positivity follows from the assumption that U has full column-rank, i.e., $UV^T \neq 0$.

- The inverse of a positive definite matrix is positive definite. This follows from the positive spectral factorization $A = Q\Lambda Q^T$, which gives us

$$A^{-1} = Q^T \Lambda^{-1} Q$$

where $\Lambda_{ii} > 0$. If A is semidefinite then the *pseudo-inverse* of A is semidefinite.

- Consider a matrix $X \in \mathcal{S}^n$ partitioned as

$$X = \left[\begin{array}{cc} A & B^T \\ B & C \end{array}\right].$$

Let us find necessary and sufficient conditions for $X \succ 0$. We know that $A \succ 0$ and $C \succ 0$ (since any submatrix must be positive definite). Furthermore, we can simplify the analysis using a nonsingular transformation

$$L = \left[\begin{array}{cc} I & 0 \\ F & I \end{array}\right]$$

to diagonalize X as $LXL^T = D$, where D is block-diagonal. Note that $\det(L) = 1$, so L is indeed nonsingular. Then $X \succ 0$ if and only if $D \succ 0$. Expanding $LXL^T = D$, we get

$$\left[\begin{array}{cc} A & AF^T + B^T \\ FA + B & FAF^T + FB^T + BF^T + C \end{array}\right] = \left[\begin{array}{cc} D_1 & 0 \\ 0 & D_2 \end{array}\right].$$

Since $\det(A) \neq 0$ (by assuming that $A \succ 0$) we see that $F = -BA^{-1}$ and direct substition gives us

$$\left[\begin{array}{cc} A & 0 \\ 0 & C - BA^{-1}B^T \end{array}\right] = \left[\begin{array}{cc} D_1 & 0 \\ 0 & D_2 \end{array}\right].$$

We have thus established the following useful result.
Lemma 4.1 (Schur complement lemma)**.** *A symmetric matrix*

$$X = \left[\begin{array}{cc} A & B^T \\ B & C \end{array}\right].$$

is positive definite if and only if

$$A \succ 0, \quad C - BA^{-1}B^T \succ 0.$$

4.1.3 Semidefinite duality

Semidefinite duality is largely identical to conic quadratic duality. We define a primal semidefinite optimization problem as

$$\begin{array}{ll} \text{minimize} & \langle C, X\rangle \\ \text{subject to} & \langle A_i, X\rangle = b_i, \quad i=1,\dots,m \\ & X \in \mathcal{S}_+^n \end{array} \tag{4.5}$$

with a dual cone

$$(\mathcal{S}_+^n)^* = \{Z \in \mathbb{R}^{n\times n} \mid \langle X, Z\rangle \geq 0,\ \forall X \in \mathcal{S}_+^n\}.$$

Lemma 4.2. *The semidefinite cone is self-dual,* $(\mathcal{S}_+^n)^* = \mathcal{S}_+^n$.

Proof. We have already seen that $X, Z \in \mathcal{S}_+^n \Rightarrow \langle X, Z\rangle \geq 0$. Conversely, assume that $Z \not\in \mathcal{S}_+^n$. Then there exists a w satisfying $w^T Z w < 0$, i.e., $X = ww^T \in \mathcal{S}_+^n$ satisfies $\langle X, Z\rangle < 0$, so $Z \not\in (\mathcal{S}_+^n)^*$.

We can now define a Lagrange function

$$\begin{array}{ll} L(X, y, S) & = \langle C, X\rangle - \sum_{i=1}^m y_i(\langle A_i, X\rangle - b_i) - \langle X, S\rangle \\ & = \langle X, C - \sum_{i=1}^m y_i A_i - S\rangle + b^T y \end{array}$$

with a dual variable $S \in \mathcal{S}_+^n$. The Lagrange function is unbounded below unless $C - \sum_i y_i A_i = S$, so we get a dual problem

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & C - \sum_{i=1}^m y_i A_i = S \\ & S \in \mathcal{S}_+^n. \end{array} \tag{4.6}$$

The primal constraints are a set of equality constraints involving inner products $\langle A_i, X\rangle = b_i$, i.e., an intersection of $n(n+1)/2$-dimensional affine hyperplanes. If we eliminate S from the dual constraints,

$$C - \sum_{i=1}^m y_i A_i \succeq 0$$

we get an affine matrix-valued function with coefficients $(C, A_i) \in \mathcal{S}^n$ and variable $y \in \mathbb{R}^m$. Such a matrix-valued affine inequality is called a *linear matrix inequality*, which we discuss in greater detail in the next section.

Weak duality follows immediately by taking the inner product between X and the dual constraint,

$$\langle X, C - \sum_{i=1}^m y_i A_i\rangle = \langle C, X\rangle - b^T y = \langle X, S\rangle \geq 0,$$

but (similar to conic quadratic optimization) the issue of strong duality is more complicated than for linear optimization. In particular, it it possible that the primal or dual optimal values are unattained, and it is also possible for a problem to have a positive duality gap, even though both the primal and dual optimal values are attained. This is illustrated in the following example.

Example 4.2 (Duality). We consider a problem

$$
\begin{array}{ll}
\text{minimize} & x_1 \\
\text{subject to} & \left[\begin{array}{ccc} 0 & x_1 & 0 \\ x_1 & x_2 & 0 \\ 0 & 0 & 1+x_1 \end{array}\right] \in \mathcal{S}_+^3,
\end{array}
$$

with feasible set $\{x_1 = 0, x_2 \geq 0\}$ and optimal value $p^\star = 0$. The dual problem can be formulated as

$$
\begin{array}{ll}
\text{maximize} & -z_2 \\
\text{subject to} & \left[\begin{array}{ccc} z_1 & (1-z_2)/2 & 0 \\ (1-z_2)/2 & 0 & 0 \\ 0 & 0 & z_2 \end{array}\right] \in \mathcal{S}_+^3,
\end{array}
$$

which has a feasible set $\{z_1 \geq 0, z_2 = 1\}$ and dual optimal value $d^\star = -1$. Both problems are feasible, but not strictly feasible.

Just as for conic quadratic optimization, existence of either a strictly feasible primal or dual point (i.e., a Slater constraint qualification) ensures strong duality for the pair of primal and dual semidefinite programs.

As discussed in the introduction, we can embed the nonnegative orthant and quadratic cone within the semidefinite cone, and the Carthesian product of several semidefinite cones can be modeled as a larger block-diagonal semidefinite matrix. For an optimization solver, however, it is more efficient to explicitly include multiple cones in the problem formulation. A general linear, conic quadratic and semidefinite problem format often used by optimization solvers is

$$
\begin{array}{ll}
\text{minimize} & \langle c^l, x^l \rangle + \sum_{j=1}^{n_q} \langle c_j^q, x_j^q \rangle + \sum_{j=1}^{n_s} \langle C_j, X_j \rangle \\
\text{subject to} & A^l x^l + \sum_{j=1}^{n_q} A_j^q x_j^q + \sum_{j=1}^{n_s} \mathcal{A}_j(X_j) = b \\
& x^l \in \mathbb{R}_+^{n_l},\ x_j^q \in \mathcal{Q}^{q_j},\ X_j \in \mathcal{S}_+^{s_j}\ \forall j
\end{array}
\tag{4.7}
$$

with a linear variable x^l, conic quadratic variables x_j^q, semidefinite variables X_j and a linear mapping $\mathcal{A}_j : \mathcal{S}^{s_j} \mapsto \mathbb{R}^m$,

$$
\mathcal{A}_j(X_j^s) := \left[\begin{array}{c} \langle A_{1j}^s, X_j^s \rangle \\ \vdots \\ \langle A_{mj}^s, X_j^s \rangle \end{array}\right].
$$

All the duality results we have presented for the simpler linear, conic quadratic and semidefinite problems translate directly to (4.7); the dual problem is

$$
\begin{array}{lrcll}
\text{maximize} & & b^T y & & \\
\text{subject to} & c^l - (A^l)^T y & = & s^l & \\
& c_j^q - (A_j^q)^T y & = & s_j^q, & j = 1, \dots, n_q \\
& C_j - \mathcal{A}_j^*(y) & = & S_j, & j = 1, \dots, n_s,
\end{array}
\tag{4.8}
$$

where

$$
\mathcal{A}_j^*(y) := \sum_{i=1}^m y_i A_{ij}^s,
$$

and strong duality holds if a strictly feasible point exists for either (4.7) or (4.8). Similarly, if there exist

$$
x^l \in \mathbb{R}_+^{n_l}, \qquad x_j^q \in \mathcal{Q}^{q_j}\ \forall j, \qquad X_j \in \mathcal{S}_+^{s_j}\ \forall j
$$

satisfying

$$A^l x^l + \sum_{j=1}^{n_q} A_j^q x_j^q + \sum_{j=1}^{n_s} \mathcal{A}_j(X_j) = 0, \quad \langle c^l, x^l \rangle + \sum_{j=1}^{n_q} \langle c_j^q, x_j^q \rangle + \sum_{j=1}^{n_s} \langle C_j, X_j \rangle < 0,$$

then the dual problem is infeasible.

4.2 Semidefinite modeling

Having discussed different characterizations and properties of semidefinite matrices, we next turn to different functions and sets that can be modeled using semidefinite cones and variables. Most of those representations involve semidefinite matrix-valued affine functions, which we discuss next.

4.2.1 Linear matrix inequalities

A linear matrix inequality is an affine matrix-valued mapping $A : \mathbb{R}^n \mapsto \mathcal{S}^m$,

$$A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0, \tag{4.9}$$

in the variable $x \in \mathbb{R}^n$ with symmetric coefficients $A_i \in \mathcal{S}^m$, $i = 0, \ldots, n$. As a simple example consider the matrix in (4.4),

$$A(x, y, z) = A_0 + xA_1 + yA_2 + zA_3 \succeq 0$$

with

$$A_0 = I, \quad A_1 = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right], \quad A_2 = \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right], \quad A_3 = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right].$$

Alternatively, we can describe the linear matrix inequality $A(x, y, z) \succeq 0$ as

$$X \in \mathcal{S}_+^3, \qquad x_{11} = x_{22} = x_{33} = 1,$$

i.e., as a semidefinite variable with fixed variables; these two alternative formulations illustrate the difference between primal and dual semidefinite formulations, discussed in Section *4.1.3*.

4.2.2 Eigenvalue optimization

Consider a symmetric matrix valued function $A : \mathbb{R}^n \mapsto \mathcal{S}^m$,

$$A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n,$$

and let the eigenvalues of $A(x)$ be denoted by

$$\lambda_1(A(x)) \geq \lambda_2(A(x)) \geq \cdots \geq \lambda_m(A(x)).$$

A number of different functions of $\lambda_i(A(x))$ can then be described using simple linear matrix inequalites.

Sum of eigenvalues

The sum of the eigenvalues corresponds to

$$\sum_{i=1}^{m} \lambda_i(A(x)) = \mathrm{tr}(A(x)) = \mathrm{tr}(A_0) + \sum_{i=1}^{n} x_i \mathrm{tr}(A_i).$$

Largest eigenvalue

The largest eigenvalue can be characterized in epigraph form $\lambda_1(A(x)) \leq t$ as

$$K = \{(x,t) \in \mathbb{R}^{n+1} \mid tI - A(x) \succeq 0\}. \tag{4.10}$$

To verify this, suppose we have a spectral factorization $A(x) = Q(x)\Lambda(x)Q(x)^T$ where $Q(x)$ is orthogonal and $\Lambda(x)$ is diagonal. Then t is an upper bound on the largest eigenvalue if and only if

$$Q(x)^T(tI - A(x))Q(x) = tI - \Lambda(x) \succeq 0.$$

Thus we can minimize the largest eigenvalue of $A(x)$.

Smallest eigenvalue

The smallest eigenvalue can be described in hypograph form $\lambda_m(A(x)) \geq t$ as

$$K = \{(x,t) \in \mathbb{R}^{n+1} \mid A(x) \succeq tI\}, \tag{4.11}$$

i.e., we can maximize the smallest eigenvalue of $A(x)$.

Eigenvalue spread

The eigenvalue spread can be modeled in epigraph form

$$\lambda_1(A(x)) - \lambda_m(A(x)) \leq t$$

by combining the two linear matrix inequalities in (4.10) and (4.11), i.e.,

$$K = \{(x,t) \in \mathbb{R}^{n+1} \mid zI \preceq A(x) \preceq sI,\ s - z \leq t\}. \tag{4.12}$$

Spectral radius

The spectral radius $\rho(A(x)) := \max_i |\lambda_i(A(x))|$ can be modeled in epigraph form $\rho(A(x)) \leq t$ using two linear matrix inequalities

$$tI \pm A(x) \succeq 0,$$

i.e., the epigraph is

$$K = \{(x,t) \in \mathbb{R}^{n+1} \mid tI \succeq A(x) \succeq -tI\}. \tag{4.13}$$

Condition number of a positive definite matrix

The condition number of a positive definite matrix can be minimized by noting that $\lambda_1(A(x))/\lambda_m(A(x)) \leq t$ if and only if there exists a $\mu > 0$ such that

$$\mu I \preceq A(x) \preceq \mu t I,$$

or equivalently if and only if $I \preceq \mu^{-1}A(x) \preceq tI$. In other words, if we make a change of variables $z := x/\mu$, $\nu := 1/\mu$ we can minimize the condition number as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & I \preceq \nu A_0 + \sum_{i=1}^m z_i A_i \preceq tI, \end{array} \tag{4.14}$$

and subsequently recover the solution $x = \nu z$. In essence, we first normalize the spectrum by the smallest eigenvalue, and then minimize the largest eigenvalue of the normalized linear matrix inequality.

Roots of the determinant

The determinant of a semidefinite matrix

$$\det(A(x)) = \prod_{i=1}^m \lambda_i(A(x))$$

is neither convex or concave, but rational powers of the determinant can be modeled using linear matrix inequalities. For a rational power $0 \leq q \leq 1/m$ we have that

$$t \leq \det(X)^q$$

if and only if

$$\left[\begin{array}{cc} A(x) & Z \\ Z^T & \mathbf{Diag}(Z) \end{array}\right]$$

$$(Z_{11}Z_{22}\cdots Z_{mm})^q \geq t, \tag{4.15}$$

where $Z \in \mathbb{R}^{m\times m}$ is a lower-triangular matrix, and the inequality (4.15) can be modeled using the formulations in Section *3.2.7*; see for a proof. Similarly we can model negative powers of the determinant, i.e., for any rational $q > 0$ we have

$$t \geq \det(X)^{-q}$$

if and only if

$$\begin{array}{c} \left[\begin{array}{cc} A(x) & Z \\ Z^T & \mathbf{Diag}(Z) \end{array}\right] \succeq 0 \\ (Z_{11}Z_{22}\cdots Z_{mm})^{-q} \geq t \end{array} \tag{4.16}$$

for a lower triangular Z.

4.2.3 Singular value optimization

We next consider a nonsymmetric matrix valued function $A : \mathbb{R}^n \mapsto \mathbb{R}^{m\times p}$,

$$A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n,$$

where $A_i \in \mathbb{R}^{m\times p}$. Assume $p \leq m$ and denote the singular values of $A(x)$ by

$$\sigma_1(A(x)) \geq \sigma_2(A(x)) \geq \cdots \geq \sigma_p(A(x)) \geq 0.$$

The singular values are simply connected to the eigenvalues of $A(x)^T A(x)$,

$$\sigma_i(A(x)) = \sqrt{\lambda_i(A(x)^T A(x))}, \tag{4.17}$$

and if $A(x)$ is symmetric the singular values corresponds to the absolute of the eigenvalues. We can then optimize several functions of the singular values.

Largest singular value

The epigraph $\sigma_1(A(x)) \leq t$ can be characterized using (4.17) as

$$A(x)^T A(x) \preceq t^2 I$$

which from Schur's lemma is equivalent to the linear matrix inequality

$$\left[\begin{array}{cc} tI & A(x) \\ A(x)^T & tI \end{array}\right] \succeq 0. \tag{4.18}$$

The largest singular value $\sigma_1(A(x))$ is also called the *spectral norm* or the ℓ_2-norm of $A(x)$, $\|A(x)\|_2 := \sigma_1(A(x))$.

Sum of singular values

The *trace* norm or the *nuclear* norm of $A(x)$ is the dual of the ℓ_2-norm. We can characterize it as

$$\|X\|_* = \sup_{\|Z\|_2\leq 1} \mathrm{tr}(X^T Z). \tag{4.19}$$

It turns out that the nuclear norm corresponds to the sum of the singular values,

$$\|X\|_* = \sum_{i=1}^m \sigma_i(X), \tag{4.20}$$

which is easy to verify using singular value decomposition $X = U\Sigma V^T$. We then have

$$\begin{aligned} \sup_{\|Z\|_2\leq 1} \mathrm{tr}(X^T Z) &= \sup_{\|X\|_2\leq 1} \mathrm{tr}(\Sigma^T U^T Z V) \\ &= \sup_{\|U^T Z V\|_2\leq 1} \mathrm{tr}(\Sigma^T U^T Z V) \\ &= \sup_{\|Y\|_2\leq 1} \mathrm{tr}(\Sigma^T Y), \end{aligned}$$

where we used the unitary invariance of the norm $\|\cdot\|_2$. We consider $Y = \mathbf{diag}(y_1, \dots, y_p)$, so that

$$\sup_{\|Z\|_2\le 1} \mathrm{tr}(X^T Z) = \sup_{|y_i|\le 1} \sum_{i=1}^p \sigma_i y_i = \sum_{i=1}^p \sigma_i,$$

which shows (4.20). Alternatively, we can express (4.19) as the solution to

$$\begin{array}{ll} \mbox{maximize} & \mathrm{tr}(X^T Z)\\ \mbox{subject to} & \left[\begin{array}{cc} I & Z^T\\ Z & I \end{array}\right] \succeq 0, \end{array}$$

with the dual problem

$$\begin{array}{ll} \mbox{minimize} & \mathrm{tr}(U+V)/2\\ \mbox{subject to} & \left[\begin{array}{cc} U & X^T\\ X & V \end{array}\right] \succeq 0. \end{array}$$

In other words, using strong duality we can characterize the epigraph $\|A(x)\|_* \le t$ using a linear matrix inequality

$$\left[\begin{array}{cc} U & A(x)^T\\ A(x) & V \end{array}\right] \succeq 0, \quad \mathrm{tr}(U+V)/2 \le t. \tag{4.21}$$

For a symmetric matrix the nuclear norm corresponds to the sum of the absolute eigenvalues, and for a semidefinite matrix it simply corresponds to the trace of the matrix.

4.2.4 Matrix inequalities from Schur's Lemma

Several quadratic or quadratic-over-linear matrix inequalities follow immediately from Schur's lemma. Suppose $A : \mathbb{R}^n \mapsto \mathbb{R}^{m\times p}$ and $B : \mathbb{R}^n \mapsto \mathbb{R}^{p\times p}$ are matrix-valued affine functions

$$\begin{array}{ll} A(x) & = A_0 + x_1 A_1 + \cdots + x_n A_n\\ B(y) & = B_0 + y_1 B_1 + \cdots + y_r B_r \end{array}$$

where $A_i \in \mathbb{R}^{m\times p}$, $B_i \in \mathcal{S}^p$, $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^r$. Then

$$A(x)^T B(y)^{-1} A(x) \preceq C$$

if and only if

$$\left[\begin{array}{cc} C & A(x)^T\\ A(x) & B(y) \end{array}\right] \succeq 0.$$

4.2.5 Nonnegative polynomials

We next consider characterizations of polynomials constrained to be nonnegative on the real axis. To that end, consider a polynomial *basis function*

$$v(t) = \left(1, t, \dots, t^{2n}\right).$$

It is then well-known that a polynomial $f : \mathbb{R} \mapsto \mathbb{R}$ of even degree $2n$ is nonnegative on the entire real axis

$$f(t) := x^T v(t) = x_0 + x_1 t + \cdots + x_{2n} t^{2n} \ge 0, \ \forall t \tag{4.22}$$

if and only if it can be written as a sum of squared polynomials of degree n (or less), i.e., for some $q_1, q_2 \in \mathbb{R}^{n+1}$

$$f(t) = (q_1^T u(t))^2 + (q_2^T u(t))^2, \qquad u(t) := (1, t, \ldots, t^n). \tag{4.23}$$

It turns out that an equivalent characterization of $\{x \mid f(t) \geq 0,\ \forall t\}$ can be given in terms of a semidefinite variable X,

$$x_i = \langle X, H_i \rangle, \quad i = 0, \ldots, 2n, \qquad X \in \mathcal{S}_+^{n+1}. \tag{4.24}$$

where $H_i^{n+1} \in \mathbb{R}^{(n+1)\times(n+1)}$ are Hankel matrices

$$[H_i]_{kl} = \begin{cases} 1, & k + l = i \\ 0, & \text{otherwise.} \end{cases}$$

When there is no ambiguity, we drop the superscript on H_i. For example, for $n = 2$ we have

$$H_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad H_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \ldots \quad H_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

To verify that (4.22) and (4.24) are equivalent, we first note that

$$u(t)u(t)^T = \sum_{i=0}^{2n} H_i v_i(t),$$

i.e.,

$$\begin{bmatrix} 1 \\ t \\ \vdots \\ t^n \end{bmatrix} \begin{bmatrix} 1 \\ t \\ \vdots \\ t^n \end{bmatrix}^T = \begin{bmatrix} 1 & t & \ldots & t^n \\ t & t^2 & \ldots & t^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ t^n & t^{n+1} & \ldots & t^{2n} \end{bmatrix}.$$

Assume next that $f(t) \geq 0$. Then from (4.23) we have

$$\begin{aligned} f(t) &= (q_1^T u(t))^2 + (q_2^T u(t))^2 \\ &= \langle q_1 q_1^T + q_2 q_2^T, u(t)u(t)^T \rangle \\ &= \textstyle\sum_{i=0}^{2n} \langle q_1 q_1^T + q_2 q_2^T, H_i \rangle v_i(t), \end{aligned}$$

i.e., we have $f(t) = x^T v(t)$ with $x_i = \langle X, H_i \rangle$, $X = (q_1 q_1^T + q_2 q_2^T) \succeq 0$. Conversely, assume that (4.24) holds. Then

$$f(t) = \sum_{i=0}^{2n} \langle H_i, X \rangle v_i(t) = \langle X, \sum_{i=0}^{2n} H_i v_i(t) \rangle = \langle X, u(t)u(t)^T \rangle \geq 0$$

since $X \succeq 0$. In summary, we can characterize the cone of nonnegative polynomials over the real axis as

$$K_\infty^n = \{x \in \mathbb{R}^{n+1} \mid x_i = \langle X, H_i \rangle,\ i = 0, \ldots, 2n,\ X \in \mathcal{S}_+^{n+1}\}. \tag{4.25}$$

Checking nonnegativity of a univariate polynomial thus corresponds to a semidefinite feasibility problem.

Nonnegativity on a finite interval

As an extension we consider a basis function of degree n,

$$v(t) = (1, t, \ldots, t^n).$$

A polynomial $f(t) := x^T v(t)$ is then nonnegative on a subinterval $I = [a, b] \subset \mathbb{R}$ if and only $f(t)$ can be written as a *sum of weighted squares*,

$$f(t) = w_1(t)(q_1^T u_1(t))^2 + w_2(t)(q_2^T u_2(t))^2$$

where $w_i(t)$ are nonnegative polynomial $w_i(t) \geq 0$, $i = 1, 2$, $\forall t \in I$. To describe the cone

$$K_{a,b}^n = \{x \in \mathbb{R}^{n+1} \mid f(t) = x^T v(t) \geq 0,\ \forall t \in [a, b]\}$$

we distinguish between polynomials of odd and even degree.

- *Even degree*. Let $n = 2m$ and denote

  $$u_1(t) = (1, t, \ldots, t^m), \qquad u_2(t) = (1, t, \ldots, t^{m-1}).$$

  We choose $w_1(t) = 1$ and $w_2(t) = (t-a)(b-t)$ and note that $w_2(t) \geq 0$ on $[a, b]$. Then $f(t) \geq 0, \forall t \in [a, b]$ if and only if

  $$f(t) = (q_1^T u_1(t))^2 + w_2(t)(q_2^T u_2(t))^2$$

  for some q_1, q_2, and an equivalent semidefinite characterization can be found as

  $$\begin{aligned} K_{a,b}^n = \{x \in \mathbb{R}^{n+1} \mid x_i = \langle X_1, H_i^m \rangle + \langle X_2, (a+b)H_{i-1}^{m-1} - abH_i^{m-1} - H_{i-2}^{m-1} \rangle, \\ i = 0, \ldots, n,\ X_1 \in \mathcal{S}_+^m,\ X_2 \in \mathcal{S}_+^{m-1}\}. \end{aligned} \tag{4.26}$$

- *Odd degree*. Let $n = 2m+1$ and denote $u(t) = (1, t, \ldots, t^m)$. We choose $w_1(t) = (t-a)$ and $w_2(t) = (b-t)$. We then have that $f(t) = x^T v(t) \geq 0, \forall t \in [a, b]$ if and only if

  $$f(t) = (t-a)(q_1^T u(t))^2 + (b-t)(q_2^T u(t))^2$$

  for some q_1, q_2, and an equivalent semidefinite characterization can be found as

  $$\begin{aligned} K_{a,b}^n = \{x \in \mathbb{R}^{n+1} \mid x_i = \langle X_1, H_{i-1}^m - aH_i^m \rangle + \langle X_2, bH_i^m - H_{i-1}^m \rangle, \\ i = 0, \ldots, n,\ X_1, X_2 \in \mathcal{S}_+^m\}. \end{aligned} \tag{4.27}$$

4.2.6 Hermitian matrices

Semidefinite optimization can be extended to complex-valued matrices. To that end, let $\mathcal{H}^n$ denote the cone of Hermitian matrices of order n, i.e.,

$$X \in \mathcal{H}^n \qquad \Longleftrightarrow \qquad X \in \mathbb{C}^{n \times n}, \quad X^H = X \tag{4.28}$$

where superscript 'H' denotes Hermitian (or complex) transposition. Then $X \in \mathcal{H}_+^n$ if and only if

$$\begin{aligned} z^H X z &= (\Re z - i\Im z)^T (\Re X + i\Im X)(\Re z + i\Im z) \\ &= \left[\begin{array}{c} \Re z \\ \Im z \end{array}\right]^T \left[\begin{array}{cc} \Re X & -\Im X \\ \Im X & \Re X \end{array}\right] \left[\begin{array}{c} \Re z \\ \Im z \end{array}\right] \geq 0, \quad \forall z \in \mathbb{C}^n. \end{aligned}$$

In other words,

$$X \in \mathcal{H}_+^n \qquad \Longleftrightarrow \qquad \left[\begin{array}{cc} \Re X & -\Im X \\ \Im X & \Re X \end{array}\right] \in \mathcal{S}_+^{2n}. \tag{4.29}$$

Note that (4.29) implies skew-symmetry of $\Im X$, i.e., $\Im X = -\Im X^T$.

4.2.7 Nonnegative trigonometric polynomials

As a complex-valued variation of the sum-of-squares representation we consider trigonometric polynomials; optimization over cones of nonnegative trigonometric polynomials has several important engineering applications. Consider a trigonometric polynomial evaluated on the complex unit-circle

$$f(z) = x_0 + 2\Re(\sum_{i=1}^{n} x_i z^{-i}), \qquad |z| = 1 \tag{4.30}$$

parametrized by $x \in \mathbb{R} \times \mathbb{C}^n$. We are interested in characterizing the cone of trigonometric polynomials that are nonnegative on the angular interval $[0, \pi]$,

$$K^n_{0,\pi} = \{x \in \mathbb{R} \times \mathbb{C}^n \mid x_0 + 2\Re(\sum_{i=1}^{n} x_i z^{-i}) \geq 0,\ \forall z = e^{jt}, t \in [0, \pi].\}.$$

Consider a complex-valued basis function

$$v(z) = (1, z, \dots, z^n).$$

The Riesz-Fejer Theorem states that a trigonometric polynomial $f(z)$ in (4.30) is nonnegative (i.e., $x \in K_{0,\pi}$) if and only if for some $q \in \mathbb{C}^{n+1}$

$$f(z) = |q^H v(z)|^2. \tag{4.31}$$

Analogously to Section *4.2.5* we have a semidefinite characterization of the sum-of-squares representation, i.e., $f(z) \geq 0, \forall z = e^{jt}$ if and only if

$$x_i = \langle X, T_i \rangle, \quad i = 0, \dots, n, \qquad X \in \mathcal{H}^{n+1}_+ \tag{4.32}$$

where $T_i^{n+1} \in \mathbb{R}^{(n+1)\times(n+1)}$ are Toeplitz matrices

$$[T_i]_{kl} = \left\{ \begin{array}{ll} 1, & k - l = i \\ 0, & \text{otherwise} \end{array} \right., \quad i = 0, \dots, n.$$

When there is no ambiguity, we drop the superscript on T_i. For example, for $n = 2$ we have

$$T_0 = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right], \quad T_1 = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right], \quad T_2 = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{array} \right].$$

To prove correctness of the semidefinite characterization, we first note that

$$v(z)v(z)^H = I + \sum_{i=1}^{n} (T_i v_i(z)) + \sum_{i=1}^{n} (T_i v_i(z))^H$$

i.e.,

$$\left[\begin{array}{c} 1 \\ z \\ \vdots \\ z^n \end{array} \right] \left[\begin{array}{c} 1 \\ z \\ \vdots \\ z^n \end{array} \right]^H = \left[\begin{array}{cccc} 1 & z^{-1} & \dots & z^{-n} \\ z & 1 & \dots & z^{1-n} \\ \vdots & \vdots & \ddots & \vdots \\ z^n & z^{n-1} & \dots & 1 \end{array} \right].$$

Next assume that (4.31) is satisfied. Then

$$\begin{array}{ll} f(z) & = \langle qq^H, v(z)v(z)^H \rangle \\ & = \langle qq^H, I\rangle + \langle qq^H, \sum_{i=1}^n T_i v_i(z)\rangle + \langle qq^H, \sum_{i=1}^n T_i^T \overline{v_i(z)}\rangle \\ & = \langle qq^H, I\rangle + \sum_{i=1}^n \langle qq^H, T_i\rangle v_i(z) + \sum_{i=1}^n \langle qq^H, T_i^T\rangle \overline{v_i(z)} \\ & = x_0 + 2\Re(\sum_{i=1}^n x_i v_i(z)) \end{array}$$

with $x_i = \langle qq^H, T_i\rangle$. Conversely, assume that (4.32) holds. Then

$$f(z) = \langle X, I\rangle + \sum_{i=1}^n \langle X, T_i\rangle v_i(z) + \sum_{i=1}^n \langle X, T_i^T\rangle \overline{v_i(z)} = \langle X, v(z)v(z)^H\rangle \geq 0.$$

In other words, we have shown that

$$K_{0,\pi}^n = \{x \in \mathbb{R}\times\mathbb{C}^n \mid x_i = \langle X, T_i\rangle,\ i=0,\dots,n,\ X \in \mathcal{H}_+^{n+1}\}. \tag{4.33}$$

Nonnegativity on a subinterval

We next sketch a few useful extensions. An extension of the Riesz-Fejer Theorem states that a trigonometric polynomial $f(z)$ of degree n is nonnegative on $I(a,b) = \{z \mid z = e^{jt},\ t \in [a,b] \subseteq [0,\pi]\}$ if and only if it can be written as a weighted sum of squared trigonometric polynomials

$$f(z) = |f_1(z)|^2 + g(z)|f_2(z)|^2$$

where f_1, f_2, g are trigonemetric polynomials of degree $n, n-d$ and d, respectively, and $g(z) \geq 0$ on $I(a,b)$. For example $g(z) = z + z^{-1} - 2\cos\alpha$ is nonnegative on $I(0,\alpha)$, and it can be verified that $f(z) \geq 0,\ \forall z \in I(0,\alpha)$ if and only if

$$x_i = \langle X_1, T_i^{n+1}\rangle + \langle X_2, T_{i+1}^n\rangle + \langle X_2, T_{i-1}^n\rangle - 2\cos(\alpha)\langle X_2, T_i^n\rangle,$$

for $X_1 \in \mathcal{H}_+^{n+1}$, $X_2 \in \mathcal{H}_+^n$, i.e.,

$$\begin{array}{rl} K_{0,\alpha}^n = \{x \in \mathbb{R}\times\mathbb{C}^n \mid x_i = & \langle X_1, T_i^{n+1}\rangle + \langle X_2, T_{i+1}^n\rangle + \langle X_2, T_{i-1}^n\rangle \\ & -2\cos(\alpha)\langle X_2, T_i^n\rangle,\ X_1 \in \mathcal{H}_+^{n+1}, X_2 \in \mathcal{H}_+^n\}. \end{array} \tag{4.34}$$

Similarly $f(z) \geq 0,\ \forall z \in I(\alpha,\pi)$ if and only if

$$x_i = \langle X_1, T_i^{n+1}\rangle + \langle X_2, T_{i+1}^n\rangle + \langle X_2, T_{i-1}^n\rangle - 2\cos(\alpha)\langle X_2, T_i^n\rangle$$

i.e.,

$$\begin{array}{rl} K_{\alpha,\pi}^n = \{x \in \mathbb{R}\times\mathbb{C}^n \mid x_i = & \langle X_1, T_i^{n+1}\rangle + \langle X_2, T_{i+1}^n\rangle + \langle X_2, T_{i-1}^n\rangle \\ & +2\cos(\alpha)\langle X_2, T_i^n\rangle,\ X_1 \in \mathcal{H}_+^{n+1}, X_2 \in \mathcal{H}_+^n\}. \end{array} \tag{4.35}$$

4.3 Semidefinite optimization examples

In this section we give examples of semidefinite optimization problems using some of the results from Section *4.2*.

4.3.1 Nearest correlation matrix

We consider the set

$$S = \{X \in \mathcal{S}_+^n \mid X_{ii} = 1,\ i = 1, \ldots, n\}$$

(shown in Fig. 4.1 for $n = 3$). For $A \in \mathcal{S}^n$ the nearest correlation matrix is

$$X^\star = \arg\min_{X \in S} \|A - X\|_F,$$

i.e., the projection of A onto the set S. To pose this as a conic optimization we define the linear operator

$$\mathbf{svec}(U) = (U_{11}, \sqrt{2}U_{21}, \ldots, \sqrt{2}U_{n1}, U_{22}, \sqrt{2}U_{32}, \ldots, \sqrt{2}U_{n2}, \ldots, U_{nn}),$$

which extracts and scales the lower-triangular part of U. We then get a conic formulation of the nearest correlation problem exploiting symmetry of $A - X$,

$$\begin{array}{lcl} \text{minimize} & t & \\ \text{subject to} & \|z\|_2 & \leq t \\ & \mathbf{svec}(A - X) & = z \\ & \mathbf{diag}(X) & = e \\ & X \succeq 0. & \end{array} \tag{4.36}$$

This is an example of a problem with both conic quadratic and semidefinite constraints in *primal form*, i.e., it matches the generic format (4.7).

We can add different constraints to the problem, for example a bound γ on the smallest eigenvalue

$$X - \gamma I \succeq 0.$$

A naive way to add the eigenvalue bound would be to introduce a new variable U,

$$U = X - \gamma I, \qquad U \in \mathcal{S}_+^n,$$

which would approximately double the number of variables and constraints in the model. Instead we should just interpret U as a change of variables leading to a problem

$$\begin{array}{lcl} \text{minimize} & t & \\ \text{subject to} & \|z\|_2 & \leq t \\ & \mathbf{svec}(A - U - \lambda I) & = z \\ & \mathbf{diag}(U + \lambda I) & = e \\ & U \succeq 0. & \end{array}$$

4.3.2 Extremal ellipsoids

Given a polytope we can find the largest ellipsoid contained in the polytope, or the smallest ellipsoid containing the polytope (for certain representations of the polytope).

Maximal inscribed ellipsoid

Consider a polytype

$$S = \{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \ i = 1, \ldots, m\}.$$

The ellipsoid

$$\mathcal{E} := \{x \mid x = Cu + d, \ \|u\| \leq 1\}$$

is contained in S if and only if

$$\max_{\|u\|_2 \leq 1} a_i^T(Cu + d) \leq b_i, \quad i = 1, \ldots, m$$

or equivalently, if and only if

$$\|Ca_i\|_2 + a_i^T d \leq b_i, \quad i = 1, \ldots, m.$$

Since $\mathbf{Vol}(\mathcal{E}) \approx \det(C)^{1/n}$ the maximum-volume inscribed ellipsoid is the solution to

$$\begin{array}{ll} \text{maximize} & \det(C)^{1/n} \\ \text{subject to} & \|Ca_i\|_2 + a_i^T d \quad \leq \quad b_i \quad i = 1, \ldots, m \\ & C \succeq 0. \end{array}$$

If we use the semidefinite characterization of a fractional power of the determinant of a positive definite matrix, we get an equivalent conic formulation,

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & \|Ca_i\|_2 + a_i^T d \leq b_i, \quad i = 1, \ldots, m \\ & \left[\begin{array}{cc} C & Z \\ Z^T & \mathbf{Diag}(Z) \end{array}\right] \succeq 0 \\ & t \leq (Z_{11} Z_{22} \cdots Z_{nn})^{1/n}, \end{array}$$

where $t \leq (Z_{11} Z_{22} \cdots Z_{nn})^{1/n}$ can be modeled as the intersection of rotated quadratic cones, see Section *3.2.7*.

Minimal enclosing ellipsoid

Next consider a polytope given as the convex hull of a set of points,

$$S' = \mathbf{conv}\{x_1, x_2, \ldots, x_m\}, \quad x_i \in \mathbb{R}^n.$$

The ellipsoid

$$\mathcal{E}' := \{x \mid \|P(x - c)\|_2 \leq 1\}$$

has $\mathbf{Vol}(\mathcal{E}') \approx \det(P)^{-1/n}$, so the minimum-volume enclosing ellipsoid is the solution to

$$\begin{array}{ll} \text{minimize} & \det(P)^{-1/n} \\ \text{subject to} & \|P(x_i - c)\|_2 \quad \leq 1, \quad i = 1, \ldots, m \\ & P \succeq 0. \end{array}$$

Alternatively, we can maximize $\det(P)^{1/n}$ to get an equivalent formulation

$$\begin{array}{ll}\mbox{maximize} & t\\ \mbox{subject to} & \|P(x_i-c)\|_2\leq 1,\quad i=1,\ldots,m\\ & \left[\begin{array}{cc}P & Z\\ Z^T & \mathbf{Diag}(Z)\end{array}\right]\succeq 0\\ & t\leq (Z_{11}Z_{22}\cdots Z_{nn})^{1/n}.\end{array}$$

In Fig. 4.2 we illustrate the inner and outer ellipsoidal approximations of a polytope characterized by 5 points in $\mathbb{R}^2$.

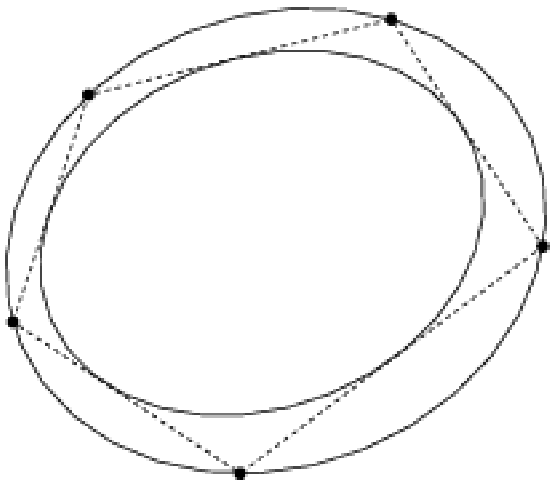


Fig. 4.2: Example of inner and outer ellipsoidal approximations.

4.3.3 Polynomial curve-fitting

Consider a univariate polynomial of degree n,

$$f(t)=x_0+x_1t+x_2t^2+\cdots+x_nt^n.$$

Often we wish to fit such a polynomial to a given set of measurements or control points

$$\{(t_1,y_1),(t_2,y_2),\ldots,(t_m,y_m)\},$$

i.e., we wish to determine coefficients x_i, $i=0,\ldots,n$ such that

$$f(t_j)\approx y_j,\quad j=1,\ldots,m.$$

To that end, define the *Vandermonde* matrix

$$A=\left[\begin{array}{ccccc}1 & t_1 & t_1^2 & \ldots & t_1^n\\ 1 & t_2 & t_2^2 & \ldots & t_2^n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & t_m & t_m^2 & \ldots & t_m^n\end{array}\right].$$

We can then express the desired curve-fit compactly as

$$Ax\approx y,$$

i.e., as a linear expression in the coefficients x. When the degree of the polynomial equals the number measurements, $n = m$, the matrix A is square and non-singular (provided there are no duplicate rows), so we can can solve

$$Ax = y$$

to find a polynomial that passes through all the control points (t_i, y_i). Similarly, if $n > m$ there are infinitely many solutions satisfying the *underdetermined* system $Ax = y$. A typical choice in that case is the *least-norm* solution

$$x_{\mathrm{ln}} = \arg\min_{Ax=y} \|x\|_2$$

which (assuming again there are no duplicate rows) has a simply solution

$$x_{\mathrm{ln}} = A^T(AA^T)^{-1}y.$$

On the other hand, if $n < m$ we generally cannot find a solution to the *overdetermined* system $Ax = y$, and we typically resort to a *least-squares* solution

$$x_{\mathrm{ls}} = \arg\min \|Ax - y\|_2$$

which has a simple solution

$$x_{\mathrm{ls}} = (A^TA)^{-1}A^Ty.$$

In the following we discuss how the semidefinite characterizations of nonnegative polynomials (see Section *4.2.5*) lead to more advanced and useful polynomial curve-fitting constraints.

- *Nonnegativity*. One possible constraint is nonnegativity on an interval,

  $$f(t) := x_0 + x_1 t + \cdots + x_n t^n \geq 0, \, \forall t \in [a, b]$$

  with a semidefinite characterization embedded in $x \in K^n_{a,b}$, see (4.26).

- *Monotonicity*. We can ensure monotonicity of $f(t)$ by requiring that $f'(t) \geq 0$ (or $f'(t) \leq 0$), i.e.,

  $$(x_1, 2x_2, \ldots, nx_n) \in K^{n-1}_{a,b},$$

  or

  $$-(x_1, 2x_2, \ldots, nx_n) \in K^{n-1}_{a,b},$$

  respectively.

- *Convexity or concavity*. Convexity (or concavity) of $f(t)$ corresponds to $f''(t) \geq 0$ (or $f''(t) \leq 0$), i.e.,

  $$(2x_2, 6x_3, \ldots, (n-1)nx_n) \in K^{n-2}_{a,b},$$

  or

  $$-(2x_2, 6x_3, \ldots, (n-1)nx_n) \in K^{n-2}_{a,b},$$

  respectively.

As an example, we consider fitting a smooth polynomial

$$f_n(t) = x_0 + x_1 t + \cdots + x_n t^n$$

to the points $\{(-1,1),(0,0),(1,1)\}$, where smoothness is implied by bounding $|f_n'(t)|$. More specifically, we wish to solve the problem

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & |f_n'(t)| \leq z, \quad \forall t \in [-1,1] \\ & f_n(-1) = 1, \quad f_n(0) = 0, \quad f_n(1) = 1, \end{array}$$

or equivalently

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & z - f_n'(t) \geq 0, \quad \forall t \in [-1,1] \\ & f_n'(t) - z \geq 0, \quad \forall t \in [-1,1] \\ & f_n(-1) = 1, \quad f_n(0) = 0, \quad f_n(1) = 1. \end{array}$$

Finally, we use the characterizations $K^n_{a,b}$ to get a conic problem

$$\begin{array}{ll} \text{minimize} & z \\ \text{subject to} & (z - x_1, -2x_2, \ldots, -nx_n) \in K^{n-1}_{-1,1} \\ & (x_1 - z, 2x_2, \ldots, nx_n) \in K^{n-1}_{-1,1} \\ & f_n(-1) = 1, \quad f_n(0) = 0, \quad f_n(1) = 1. \end{array}$$

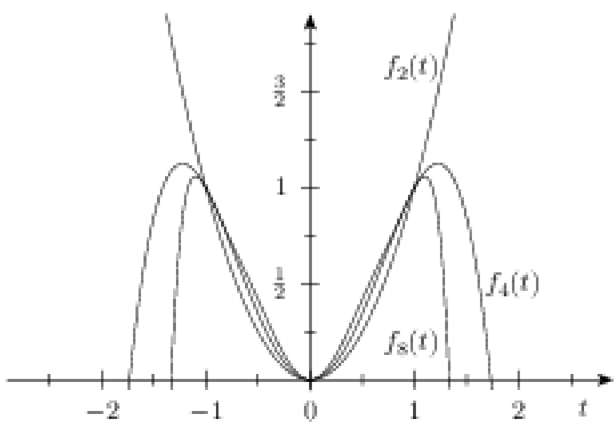


Fig. 4.3: Graph of univariate polynomials of degree 2, 4, and 8, respectively, passing through $\{(-1,1),(0,0),(1,1)\}$. The higher-degree polynomials are increasingly smoother on $[-1,1]$.

In Fig. 4.3 we show the graphs for the resulting polynomails of degree 2, 4 and 8, respectively. The second degree polynomial is uniquely determined by the three constraints $f_2(-1) = 1$, $f_2(0) = 0$, $f_2(1) = 1$, i.e., $f_2(t) = t^2$. Also, we obviously have a lower bound on the largest derivative $\max_{t\in[-1,1]} |f_n'(t)| \geq 1$. The computed fourth degree polynomial is given by

$$f_4(t) = \frac{3}{2}t^2 - \frac{1}{2}t^4$$

after rounding coefficients to rational numbers. Furthermore, the largest derivative is given by

$$f_4'(1/\sqrt{2}) = \sqrt{2},$$

and $f_4''(t) < 0$ on $(1/\sqrt{2}, 1]$ so, although not visibly clear, the polynomial is nonconvex on $[-1, 1]$. In Fig. 4.4 we show the graphs of the corresponding polynomials where we added a convexity constraint $f_n''(t) \geq 0$, i.e.,

$$(2x_2, 6x_3, \ldots, (n-1)nx_n) \in K^{n-2}_{-1,1}.$$

In this case, we get

$$f_4(t) = \frac{6}{5}t^2 - \frac{1}{5}t^4$$

and the largest derivative increases to $\frac{8}{5}$.

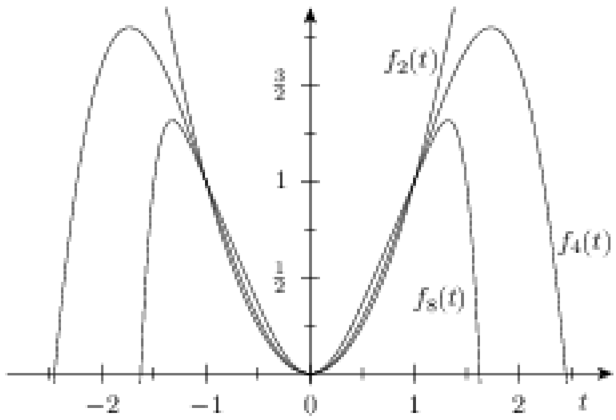


Fig. 4.4: Graph of univariate polynomials of degree $2, 4$, and 8, respectively, passing through $\{(-1, 1), (0, 0), (1, 1)\}$. The polynomials all have positive second derivative (i.e., they are convex) on $[-1, 1]$ and the higher-degree polynomials are increasingly smoother on that interval.

4.3.4 Filter design problems

An important signal processing application is filter design problems over trigonometric polynomials

$$\begin{aligned} H(\omega) &= x_0 + 2\Re(\textstyle\sum_{k=1}^n x_k e^{-j\omega k}) \\ &= a_0 + 2\textstyle\sum_{k=1}^n (a_k \cos(\omega k) + b_k \sin(\omega k)) \end{aligned}$$

where $a_k := \Re(x_k)$, $b_k := \Im(x_k)$. If the function $H(\omega)$ is nonnegative we call it a transfer function, and it describes how different harmonic components of a periodic discrete signal are attenuated when a filter with transfer function $H(\omega)$ is applied to the signal.

We often wish a transfer function where $H(\omega) \approx 1$ for $0 \leq \omega \leq \omega_p$ and $H(\omega) \approx 0$ for $\omega_s \leq \omega \leq \pi$ for given constants ω_p, ω_s. One possible formulation for achieving this is

$$\begin{array}{lccccll} \text{minimize} & & & t & & & \\ \text{subject to} & 0 & \leq & H(\omega) & & & \forall \omega \in [0, \pi] \\ & 1-\delta & \leq & H(\omega) & \leq & 1+\delta & \forall \omega \in [0, \omega_p] \\ & & & H(\omega) & \leq & t & \forall \omega \in [\omega_s, \pi], \end{array}$$

which corresponds to minimizing $H(w)$ on the interval $[\omega_s, \pi]$ while allowing $H(w)$ to depart from unity by a small amount δ on the interval $[0, \omega_p]$. Using the results from Section 4.2.7 (in particular (4.33), (4.34) and (4.35)), we can pose this as a conic optimization problem

$$\begin{array}{lrcl} \text{minimize} & t & & \\ \text{subject to} & x & \in & K^n_{0,\pi} \\ & (x_0-(1-\delta), x_{1:n}) & \in & K^n_{0,\omega_p} \\ & -(x_0-(1+\delta), x_{1:n}) & \in & K^n_{0,\omega_p} \\ & -(x_0-t, x_{1:n}) & \in & K^n_{\omega_s,\pi}, \end{array} \tag{4.37}$$

which is a semidefinite optimization problem. In Fig. 4.5 we show $H(\omega)$ obtained by solving (4.37) for $n=10$, $\delta=0.05$, $\omega_p=\pi/4$ and $\omega_s=\omega_p+\pi/8$.

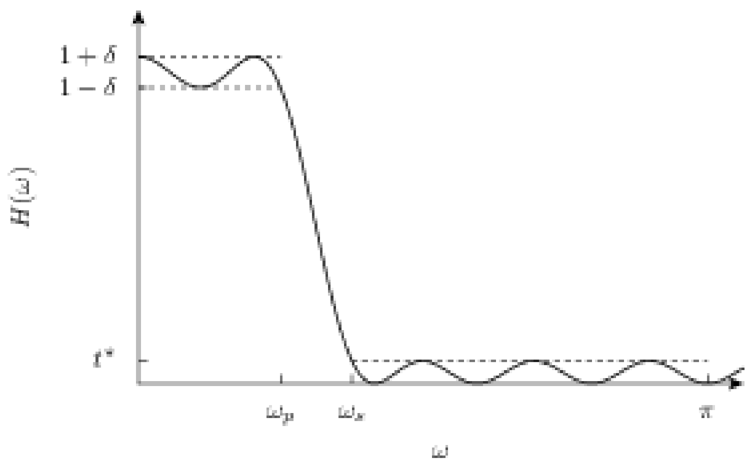


Fig. 4.5: Plot of lowpass filter transfer function.

4.3.5 Relaxations of binary optimization

Semidefinite optimization is also useful for computing bounds on difficult non-convex or combinatorial optimization problems. For example consider the binary linear optimization problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \{0,1\}^n. \end{array} \tag{4.38}$$

In general, problem (4.38) is a very difficult non-convex problem where we have to explore 2^n different objectives. Alternatively, we can use semidefinite optimization to get a lower bound on the optimal solution with polynomial complexity. We first note that

$$x_i \in \{0,1\} \quad \Longleftrightarrow \quad x_i^2 = x_i,$$

which is, in fact, equivalent to a rank constraint on a semidefinite variable,

$$X = xx^T, \qquad \mathbf{diag}(X) = x.$$

By relaxing the rank 1 constraint on X we get a semidefinite relaxation of (4.38),

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & \mathbf{diag}(X) = x \\ & X \succeq xx^T, \end{array} \tag{4.39}$$

where we note that

$$X \succeq xx^T \quad \Longleftrightarrow \quad \left[\begin{array}{cc} 1 & x^T \\ x & X \end{array}\right] \succeq 0.$$

Since (4.39) is a semidefinite optimization problem, it can be solved very efficiently. Suppose $x^\star$ is an optimal solution for (4.38); then $(x^\star, x^\star(x^\star)^T)$ is also feasible for (4.39), but the feasible set for (4.39) is larger than the feasible set for (4.38), so in general the optimal solution of (4.39) serves as a lower bound. However, if the optimal solution $X^\star$ of (4.39) has rank 1 we have found a solution to (4.38) also.

The semidefinite relaxation (4.39) can be derived more systematically. The Lagrangian function of (4.38) is

$$\begin{array}{ll} L(x,y) & = c^T x + y^T(Ax-b) + \sum_{i=1}^n z_i(x_i^2 - x_i) \\ & = x^T(c + A^T y - z) + x^T \mathbf{diag}(z) x - b^T y, \end{array}$$

which is bounded below only if $\mathbf{diag}(z) \succeq 0$ and $(c + A^T y - z) \in \mathcal{R}(\mathbf{diag}(z))$. Assume for simplicity that $\mathbf{diag}(z) \succ 0$. Then the minimizer of $L(x,y)$ is

$$\arg\min_x L(x,y) = -(1/2)\mathbf{diag}(z)^{-1}(c + A^T y - z)$$

and we get a dual problem

$$\begin{array}{ll} \mbox{maximize} & -(1/4)(c + A^T y - z)^T \mathbf{diag}(z)^{-1}(c + A^T y - z) - b^T y \\ \mbox{subject to} & z > 0. \end{array}$$

From Schur's Lemma Lemma 4.1.2 this is equivalent to a standard semidefinite optimization problem

$$\begin{array}{ll} \mbox{maximize} & -y^T b - t \\ \mbox{subject to} & \left[\begin{array}{cc} t & \frac{1}{2}(c + A^T y - z) \\ \frac{1}{2}(c + A^T y - z)^T & \mathbf{diag}(z) \end{array}\right] \succeq 0. \end{array} \tag{4.40}$$

This dual problem gives us a lower bound. To obtain an explicit relaxation of (4.38) we derive the dual once more, but this time of problem (4.40). The Lagrangian function of (4.40) is

$$\begin{array}{ll} L(t,y,X) & = -y^T b - t + \langle \left[\begin{array}{cc} \nu & x^T \\ x & X \end{array}\right], \left[\begin{array}{cc} t & \frac{1}{2}(c + A^T y - z) \\ \frac{1}{2}(c + A^T y - z)^T & \mathbf{diag}(z) \end{array}\right] \rangle \\ & = t(\nu - 1) + y^T(Ax - b) + z^T(\mathbf{Diag}(X) - x) + c^T x \end{array}$$

which is only bounded above if $\nu = 1$, $Ax = b$, $\mathbf{Diag}(X) = x$, i.e., we get the dual problem (4.39), which is called a *Lagrangian relaxation*. From strong duality, problem (4.39) has the same optimal value as problem (4.40).

We can tighten (or improve) the relaxation (4.39) by adding other constraints the cuts away part of the feasible set, without excluding rank 1 solutions. By tightening the relaxation, we reduce the duality gap between the optimal values of the original problem and the relxation. For example, we can add the constraints

$$0 \leq x_i \leq 1, \quad i = 1, \dots, n$$

and

$$0 \leq X_{ij} \leq 1, \quad i = 1, \dots, n, \; j = 1, \dots, n.$$

A semidefinite matrix X with $X_{ij} \geq 0$ is called a *doubly nonnegative* matrix. In practice, constraining a semidefinite matrix to be doubly nonnegative has a dramatic impact on the solution time and memory requirements of an optimization problem, since we add n^2 linear inequality constraints. We refer to for a recent survey on semidefinite relaxations for integer programming.

4.3.6 Relaxations of boolean optimization

Similarly to Section *4.3.5* we can use semidefinite relaxations of boolean constraints $x \in \{-1,+1\}^n$. To that end, we note that

$$x \in \{-1,+1\}^n \quad \Longleftrightarrow \quad X = xx^T, \quad \mathbf{diag}(X) = e, \tag{4.41}$$

with a semidefinite relaxation $X \succeq xx^T$ of the rank-1constraint.

As a (standard) example of a combinatorial problem with boolean constraints, let us consider an undirected graph G described by a set of vertices $V = \{v_1, \dots, v_n\}$ and a set of edges $E = \{(i,j) \mid i,j \in V,\ i \neq j\}$, and we wish to find the cut of maximum capacity. A cut C partitions the nodes V into two disjoint sets S and $T = V \setminus S$, and the cut-set I is the set of edges with one node in S and another in T, i.e.,

$$I = \{(u,v) \in E \mid u \in S,\ v \in T\}.$$

The capacity of a cut is then defined as the number of edges in the cut-set $|I|$; for example Fig. 4.6 illustrates a cut $\{v_2, v_4, v_5\}$ of capacity 9.

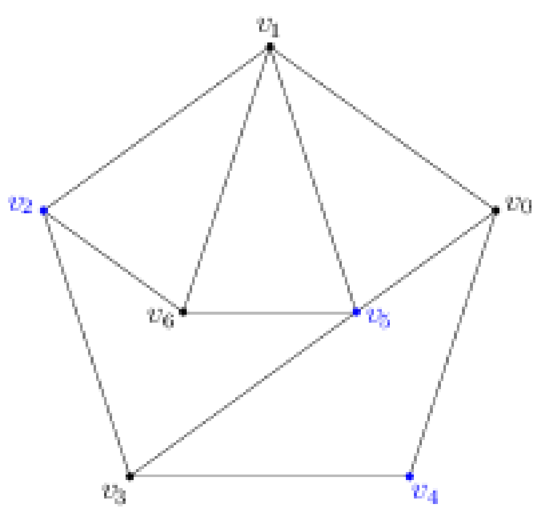


Fig. 4.6: Undirected graph. The cut $\{v_2, v_4, v_5\}$ has capacity 9.

To maximize the capacity of a cut we define the symmetric *adjacency* matrix $A \in \mathcal{S}^n$,

$$[A]_{ij} = \begin{cases} 1, & (v_i, v_j) \in E \\ 0, & \text{otherwize} \end{cases}$$

where $n = |V|$, and let

$$x_i = \begin{cases} +1, & v_i \in S \\ -1, & v_i \notin S. \end{cases}$$

Suppose $v_i \in S$. Then $1 - x_i x_j = 0$ if $v_j \in S$ and $1 - x_i x_j = 2$ if $v_j \notin S$, so we get an expression of the capacity as

$$c(x) = \frac{1}{4}\sum_{ij}(1 - x_i x_j)A_{ij} = \frac{1}{4}e^T A e - \frac{1}{4}x^T A x,$$

and discarding the constant term $e^T Ae$ gives us the MAX-CUT problem

$$\begin{array}{ll} \text{minimize} & x^T Ax \\ \text{subject to} & x \in \{-1, +1\}^n, \end{array} \tag{4.42}$$

with a semidefinite relaxation

$$\begin{array}{ll} \text{minimize} & \langle A, X \rangle \\ \text{subject to} & \mathbf{diag}(X) = e \\ & X \succeq 0. \end{array} \tag{4.43}$$

The MAX-CUT relaxation can also be derived using Lagrangian duality, analogous to Section *4.3.5*.

4.3.7 Converting a problem to standard form

From a modeling perspective it does not matter whether constraints are given as linear matrix inequalities or as an intersection of affine hyperplanes; one formulation is easily converted to other using extraneous variables and constraints, and this transformation is often done transparently by optimization software. Nevertheless, it is instructive to study an explicit example of how to carry out this transformation.

Optimization solvers require a problem to be posed either in primal form (4.7) or dual form (4.8), but problems often have both linear equality constraints and linear matrix inequalities, so one must be converted to the other. A linear matrix inequality

$$A_0 + x_1 A_1 + \cdots + x_n A_n \succeq 0$$

where $A_i \in \mathcal{S}^m$ is converted to a set of linear equality constraints using a slack variable

$$A_0 + x_1 A_1 + \cdots + x_n A_n = S, \qquad S \succeq 0. \tag{4.44}$$

Apart from introducing an additional semidefinite variable $S \in \mathcal{S}_+^m$, we also add $m(m+1)/2$ equality constraints. On the other hand, a semidefinite variable $X \in \mathcal{S}_+^n$ can be rewritten as a linear matrix inequality with $n(n+1)/2$ scalar variables

$$X = \sum_{i=1}^n e_i e_i^T x_{ii} + \sum_{i=1}^n \sum_{j=i+1}^n (e_i e_j^T + e_j e_i^T) x_{ij} \succeq 0. \tag{4.45}$$

Such convertions are often necessary, but they can make the computational cost of the resulting problem prohibitively high.

Obviously we should only use the the transformations (4.44) and (4.45) when necessary; if we have a problem that is more naturally interpreted in either primal or dual form, we should be careful to recognize that structure. For example, consider a problem

$$\begin{array}{ll} \text{minimize} & e^T z \\ \text{subject to} & A + \mathbf{Diag}(z) = X \\ & X \succeq 0. \end{array}$$

with the variables $X \in \mathcal{S}_+^n$ and $z \in \mathbb{R}^n$. This is a problem in primal form with $n(n+1)/2$ equality constraints, but they are more naturally interpreted as a linear matrix inequality

$$A + \sum_i e_i e_i^T z_i \succeq 0.$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -\langle A, Z\rangle \\ \text{subject to} & \mathbf{diag}(Z) = e \\ & Z \succeq 0, \end{array}$$

in the variable $Z \in \mathcal{S}_+^n$, which corresponds to the MAX-CUT relaxation. The dual problem has only n equality constraints, which is a vast improvement over the $n(n+1)/2$ constraints in the primal problem.

4.4 Bibliography

Much of the material in this chapter is based on the paper *[VB96]* and the books *[BenTalN01]*, *[BKVH04]*. The section on optimization over nonnegative polynomials is based on *[Nes99]*, *[Hac03]*. We refer to *[LR05]* for a comprehensive survey on semidefinite optimization and relaxations in combinatorial optimization.

CHAPTER

FIVE

QUADRATIC OPTIMIZATION

5.1 Introduction

In this chapter we discuss convex quadratic optimization. Our discussion is fairly brief compared to the previous chapters for three reasons; (*i*) quadratic optimization is a special case of conic quadratic optimization, (*ii*) for most problems it is actually more efficient for the solver to pose the problem in conic form, and (*iii*) duality theory (including infeasibility certificates) are much simpler for conic quadratic optimization. Therefore, we generally recommend a conic quadratic formulation.

5.2 Convex quadratic optimization

A standard (convex) quadratic optimization problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^TQx + c^Tx \\ \text{subject to} & Ax = b \\ & x \geq 0, \end{array} \tag{5.1}$$

with $Q \in \mathcal{S}_+^n$ is conceptually a simple extension of a linear optimization problem (2.3.1) with a quadratic term x^TQx. Note the important requirement that Q is symmetric positive semidefinite; otherwise the objective function is not convex.

5.2.1 Geometry of quadratic optimization

Quadratic optimization has a simple geometric interpretation; we minimize a convex quadratic function over a polyhedron. In Fig. 5.1 we illustrate this interpretation for the same example as in Fig. 2.5 extended with an additional quadratic term with

$$Q = \left[\begin{array}{cc} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{array} \right].$$

It is intuitively clear that the following different cases can occur:

- The optimal solution $x^\star$ is at the boundary of the polyhedron (as shown in Fig. 5.1). At $x^\star$ one of the hyperplanes is tangential to an ellipsoidal level curve.
- The optimal solution is inside the polyhedron; this occurs if the unconstrained minimizer $\arg\min_x \frac{1}{2}x^TQx + c^Tx = -Q^\dagger c$ (i.e., the center of the ellipsoidal level curves) is inside the polyhedron.

- If the polyhedron is unbounded in the opposite direction of c, and if the ellipsoid level curves are degenerate in that direction (i.e., $Qc = 0$), then the problem is unbounded. If $Q \in \mathcal{S}^n_{++}$, then the problem cannot be unbounded.
- The problem is infeasible, i.e., $\{x \mid Ax = b,\ x \geq 0\} = \emptyset$.

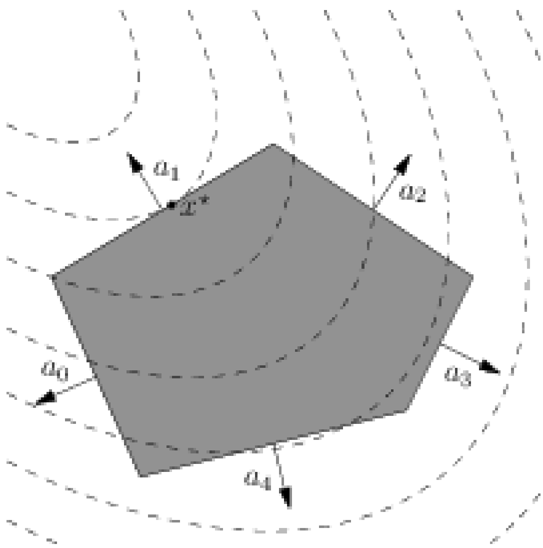


Fig. 5.1: Geometric interpretation of quadratic optimization. At the optimal point $x^\star$ the hyperplane $\{x \mid a_1^T x = b\}$ is tangential to an ellipsoidal level curve.

Possibly because of its simple geometric interpretation and similarity to linear optimization, quadratic optimization has been more widely adopted by optimization practioners than conic quadratic optimization.

5.2.2 Duality in quadratic optimization

The Lagrangian function for (5.1) is

$$L(x, y, s) = \frac{1}{2}x^T Qx + x^T (c - A^T y - s) + b^T y \tag{5.2}$$

with Lagrange multipliers $s \in \mathbb{R}^n_+$, and from $\nabla_x L(x,y,s) = 0$ we get the necessary first-order optimality condition

$$Qx = A^T y + s - c,$$

i.e., $(A^T y + s - c) \in \mathcal{R}(Q)$. Then

$$\arg\min_x L(x,y,s) = Q^\dagger (A^T y + s - c),$$

which can be substituted into (5.2) to give a dual function

$$g(y,s) = \begin{cases} b^T y - \frac{1}{2}(A^T y + s - c)Q^\dagger (A^T y + s - c), & (A^T y + s - c) \in \mathcal{R}(Q) \\ -\infty & \text{otherwise,} \end{cases}$$

Thus we get a dual problem

$$\begin{array}{ll} \text{maximize} & b^T y - \frac{1}{2}(A^T y + s - c)Q^\dagger (A^T y + s - c) \\ \text{subject to} & (A^T y + s - c) \in \mathcal{R}(Q) \\ & s \geq 0. \end{array} \tag{5.3}$$

Alternatively, we can characterize the constraint $(A^T y + s - c) \in \mathcal{R}(Q)$ as $Qw = A^T y + s - c$ with an extraneous variable w to get an equivalent dual problem

$$\begin{array}{ll} \text{maximize} & b^T y - \frac{1}{2} w^T Q w \\ \text{subject to} & Qw = A^T y - c + s \\ & s \geq 0. \end{array} \tag{5.4}$$

Note from the optimality conditions that $w = x$, so (5.4) is an unusual dual problem in the sense that it involves both primal and dual variables.

Weak duality follows from the inner product between x and the dual equality constraint,

$$x^T Q x = b^T y - c^T x + x^T s,$$

which shows that

$$\frac{1}{2} x^T Q x + c^T x - \left(b^T y - \frac{1}{2} x^T Q x \right) = x^T s \geq 0.$$

Furthermore, strong duality holds if a Slater constraint qualification is satisfied, i.e., if either the primal or dual problem is strictly feasible (as in the previous chapters).

5.2.3 Infeasibility in quadratic optimization

Feasibility of the primal problem (5.1) is described by Lemma 9.5; either problem (5.1) is feasible or there exists a certificate y of primal infeasibility satisfying

$$A^T y \leq 0, \qquad b^T y > 0.$$

In order to characterize feasibility of the dual problem, we need another pair of alternatives, which follows directly from Lemma 9.5 by considering

$$A' := \left(\begin{array}{c} A \\ -Q \end{array} \right).$$

Either there exists an $x \geq 0$ satisfying

$$Ax = 0, \qquad Qx = 0, \qquad c^T x < 0$$

or there exists (y, w) satisfying

$$Qw - A^T y + c \geq 0.$$

5.3 Quadratically constrained optimization

A general convex quadratically constrained quadratic optimization problem is

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Q_0 x + c_0^T x + r_0 \\ \text{subject to} & \frac{1}{2}x^T Q_i x + c_i^T x + r_i \leq 0, \quad i = 1, \dots, m, \end{array} \tag{5.5}$$

where $Q_i \in \mathcal{S}_+^n, \ \forall i$. This corresponds to minimizing a convex quadratic function over an intersection of ellipsoids (or affine halfspaces if some of Q_i are zero). Note the important requirement $Q_i \succeq 0, \ \forall i$, so that the objective function is convex and the constraints

$$\frac{1}{2}x^T Q_i x + c_i^T x + r_i \leq 0$$

characterize convex sets. It is important to note that neither quadratic equalities

$$\frac{1}{2}x^T Q_i x + c_i^T x + r_i = 0$$

or inequalities of the form

$$\frac{1}{2}x^T Q_i x + c_i^T x + r_i \geq 0$$

characterize convex sets, and therefore cannot be included.

5.3.1 Duality in quadratically constrained optimization

The Lagrangian function for (5.5) is

$$L(x, \lambda) = \frac{1}{2}x^T Q_0 x + c_0^T x + r_0 + \sum_{i=1}^m \lambda_i \left(\frac{1}{2}x^T Q_i x + c_i^T x + r_i \right) = \frac{1}{2}x^T Q(\lambda) x + c(\lambda)^T x + r(\lambda),$$

where

$$Q(\lambda) = Q_0 + \sum_{i=1}^m \lambda_i Q_i, \quad c(\lambda) = c_0 + \sum_{i=1}^m \lambda_i c_i, \quad r(\lambda) = r_0 + \sum_{i=1}^m \lambda_i r_i.$$

From the Lagrangian we get the first-order optimality conditions

$$Q(\lambda) x = -c(\lambda), \tag{5.6}$$

and similar to the case of quadratic optimization we get a dual problem

$$\begin{array}{ll} \text{maximize} & -\frac{1}{2}c(\lambda)^T Q(\lambda)^\dagger c(\lambda) + r(\lambda) \\ \text{subject to} & c(\lambda) \in \mathcal{R}\left(Q(\lambda)\right) \\ & \lambda \geq 0, \end{array} \tag{5.7}$$

or equivalently

$$\begin{array}{ll} \text{maximize} & -\frac{1}{2}w^T Q(\lambda) w + r(\lambda) \\ \text{subject to} & Q(\lambda) w = -c(\lambda) \\ & \lambda \geq 0, \end{array} \tag{5.8}$$

Weak duality is easily verified; from (5.6) we have

$$x^T Q(\lambda)x + c(\lambda)^T x = 0$$

which implies

$$\frac{1}{2}x^T Q_0 x + c_0^T x + r_0 - \left(-\frac{1}{2}x^T Q(\lambda)x + r(\lambda)\right) = -\sum_{i=1}^m \lambda_i \left(\frac{1}{2}x^T Q_i x + c_i^T x + r_i\right) \geq 0,$$

and strong duality holds provided the quadratic inequalities are strictly feasible,

$$\frac{1}{2}x^T Q_i x + c_i^T x + r_i < 0, \quad i = 1, \dots, m$$

for some $x \in \mathbb{R}^n$. Using a general version of Lemma [lm:schur] for singular matrices (see, e.g.,), we can also write (5.7) as an equivalent semidefinite optimization problem,

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & \left[\begin{array}{cc} 2(r(\lambda)-t) & c(\lambda)^T \\ c(\lambda) & Q(\lambda) \end{array}\right] \succeq 0 \\ & \lambda \geq 0. \end{array} \tag{5.9}$$

5.3.2 Infeasibility in quadratically constrained optimization

Feasibility of (5.5) is characterized by the following lemma.

Given $Q_i \succeq 0$ $i = 1, \dots, m$, either there exists an $x \in \mathbb{R}^n$ satisfying

$$\frac{1}{2}x^T Q_i x + c_i^T x + r_i < 0, \quad i = 1, \dots, m \tag{5.10}$$

or there exists $\lambda \geq 0,\ \lambda \neq 0$ satisfying

$$\sum_{i=1}^m \lambda_i \left[\begin{array}{cc} 2r_i & c_i^T \\ c_i & Q_i \end{array}\right] \succeq 0. \tag{5.11}$$

Suppose (5.10) and (5.11) are both satisfied. Then

$$\sum_{i=1}^m \lambda_i \left[\begin{array}{c} 1 \\ x \end{array}\right]^T \left[\begin{array}{cc} 2r_i & c_i^T \\ c_i & Q_i \end{array}\right] \left[\begin{array}{c} 1 \\ x \end{array}\right] = 2\sum_{i=1}^m \lambda_i \left(\frac{1}{2}x^T Q_i x + c_i^T x + r_i\right) \geq 0,$$

which is a contradiction. Conversely, assume (5.11) is not satisfied, i.e.,

$$\sum_{i=1}^m \lambda_i \left[\begin{array}{cc} 2r_i & c_i^T \\ c_i & Q_i \end{array}\right] \prec 0, \quad \forall \lambda \geq 0,\ \lambda \neq 0.$$

But then $X = \left(\begin{smallmatrix} 1 & x^T \\ x & xx^T \end{smallmatrix}\right)$ satisfies

$$\left\langle \sum_{i=1}^m \lambda_i \left[\begin{array}{cc} 2r_i & c_i^T \\ c_i & Q_i \end{array}\right], \left[\begin{array}{c} 1 \\ x \end{array}\right]\left[\begin{array}{c} 1 \\ x \end{array}\right]^T \right\rangle < 0, \quad \forall \lambda > 0,\ \lambda \neq 0$$

or equivalently

$$2\sum_{i=1}^m \lambda_i \left(\frac{1}{2}x^T Q_i x + c_i^T x + r_i\right) < 0, \quad \forall \lambda > 0,\ \lambda \neq 0$$

which in turn implies (5.10).

Dual infeasibility is characterized by the following set of alternatives; either problem (5.8) is feasible or there exists a certificate x of dual infeasibility satisfying (5.13).

Given $Q_i \succeq 0$, $i = 0, \ldots, m$, either there exists (λ, w) satisfying

$$Q(\lambda)w + c(\lambda) = 0, \qquad \lambda \geq 0, \tag{5.12}$$

or there exists $x \in \mathbb{R}^n$ satisfying

$$Q_0 x = 0, \qquad c_0^T x < 0, \qquad Q_i x = 0, \qquad c_i^T x = 0, \quad i = 1, \ldots, m. \tag{5.13}$$

Both cannot be true, because then

$$c_0^T x = -x^T(Q_0 + \sum_{i=1}^m \lambda_i Q_i)w - \sum_{i=1}^m \lambda_i c_i^T x = 0,$$

which is a contradiction. On the other hand, suppose

$$Q(\lambda)w + c(\lambda) \neq 0, \quad \forall w, \forall \lambda \geq 0,$$

i.e., $Q(\lambda)$ is singular $\forall \lambda \geq 0$, so there exists $x \in \mathbb{R}^n$ satisfying $Q_i x = 0,\ i = 0, \ldots, m$.

5.4 Separable quadratic optimization

Often a factorization of $Q = F^T F$ is either known or readily available, in which case we get an alternative formulation of (5.1) as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2} z^T z + c^T x \\ \text{subject to} & Ax = b \\ & Fx = z \\ & x \geq 0. \end{array} \tag{5.14}$$

Formulation (5.14) has several advantages; convexity of the problem is obvious (which can occasionally be difficult to detect in finite precision arithmetic), and the structure and sparsity of the separable reformulation is typically more efficient for an optimization solver. In the following example we consider a related reformulation, which can result in a significant reduction of complexity and solution time (see also Sec. *3.3.3*).

Consider a standard quadratic optimization problem (5.14) where

$$Q = \mathbf{diag}(w) + F^T F, \quad w > 0.$$

We then get a separable problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2} x^T \mathbf{diag}(w) x + z^T z + c^T x \\ \text{subject to} & Ax = b \\ & Fx = z \\ & x \geq 0. \end{array}$$

In general we can reformulate (5.5) as

$$\begin{array}{lll} \text{minimize} & \frac{1}{2} z_0^T z_0 + c_0^T x + r_0 & \\ \text{subject to} & \frac{1}{2} z_i^T z_i + c_i^T x + r_i \leq 0, & i = 1, \ldots, m \\ & F_i x = z_i, & i = 0, \ldots, m, \end{array} \tag{5.15}$$

which is very similar to the conic formulation (3.3.1). If the formulation (5.15) is sparse compared to (5.5) (i.e., if it is described using fewer non-zeros), then it generally results in a reduction of computation time. On the other hand, if we form (5.15), we might as well use the conic formulation (3.3.1) instead.

CHAPTER

SIX

MIXED INTEGER OPTIMIZATION

6.1 Introduction

In the previous chapters we have considered different classes of convex problems with continuous variables. In this chapter we consider a much wider range of problems by allowing integer variables. In particular, we consider conic and convex quadratic optimization problems where some of the variables are constrained to be integer valued. In principle, this allows us to model a vastly larger range of problems, but in practice we may not able to solve the resulting integer optimization problems within a reasonable amount of time.

6.2 Integer modeling

In this section we consider different building blocks for integer optimization, which we think of as building blocks for general integer problems.

6.2.1 Implication of positivity

Often we have a real-valued variable $x \in \mathbb{R}$ satisfying $0 \leq x < M$ for a known upper bound M, and we wish to model the implication

$$(x > 0) \rightarrow (z = 1) \tag{6.1}$$

where $z \in \{0, 1\}$ is an *indicator variable*. Since $0 \leq x < M$ we can rewrite (6.1) as a linear inequality

$$x \leq Mz, \tag{6.2}$$

where $z \in \{0, 1\}$. From the characterization (6.2) we see that that (6.1) is equivalent to

$$(z = 0) \rightarrow (x = 0),$$

which is the well-known property of *contraposition*. In Sec. *6.3* we discuss such standard boolean expressions in more details.

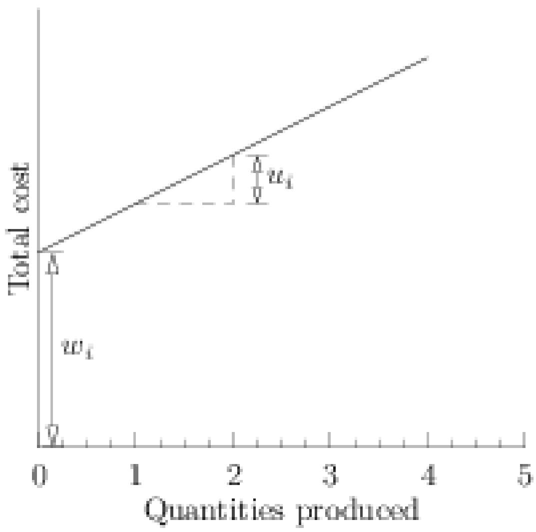


Fig. 6.1: Production cost with fixed charge w_i.

Example 6.1 (Fixed charge cost). Assume that production of a specific item i has a production cost of u_i per unit produced with an additional fixed charge of w_i if we produce item i, i.e., a discontinuous production cost

$$c_i(x_i) = \left\{ \begin{array}{ll} w_i + u_i x_i, & x_i > 0 \\ 0, & x_i = 0. \end{array} \right.$$

shown in Fig. 6.1. If we let M denote an upper bound on the quantities we can produce, we get an alternative expression for the production cost,

$$c_i(x_i) = u_i x_i + w_i z_i, \qquad x_i \leq M z_i, \quad z_i \in \{0, 1\},$$

where z_i is a binary variable indicating whether item i is produced or not. We can then minimize the total production cost under an affine constraint $Ax = b$ as

$$\begin{array}{ll} \mbox{minimize} & u^T x + w^T z \\ \mbox{subject to} & Ax = b \\ & x_i \leq M z_i, \quad i = 1, \dots, n \\ & x \geq 0 \\ & z \in \{0, 1\}^n, \end{array}$$

which is a linear mixed-integer optimization problem. Note that by minimizing the production cost, we rule out the possibility that $z_i = 1$ and $x_i = 0$.

On the other hand, suppose x satisfies $0 < m \leq x$ for a known lower bound m. We can then model the implication

$$(z = 1) \rightarrow (x > 0), \tag{6.3}$$

(which is the reverse implication of (6.1)) as a linear inequality

$$x \geq mz, \tag{6.4}$$

where $z \in \{0, 1\}$.

6.2.2 Semi-continuous variables

We can also model *semi-continuity* of a variable $x \in \mathbb{R}^n$,

$$x \in 0 \cup [a, b] \tag{6.5}$$

where $0 < a \leq b$ using a double inequality

$$az \leq x \leq bz \tag{6.6}$$

where $z \in \{0, 1\}$.

6.2.3 Constraint satisfaction

Suppose we have an upper bound M on the affine expression $a^T x - b$ where $x \in \mathbb{R}^n$. We can then model the implication $(z = 1) \rightarrow (a^T x \leq b)$ as

$$a^T x \leq b + M(1 - z) \tag{6.7}$$

where $z \in \{0, 1\}$, i.e., z indicates whether or not the constraint $a^T x \leq b$ is satisfied. To model the reverse implication $(a^T x \leq b) \rightarrow (z = 1)$ we equivalently consider the contraposition $(z = 0) \rightarrow (a^T x > b)$, which we can rewrite as

$$a^T x > b + mz$$

where $m < a^T x - b$ is a lower bound. In practice, we relax the strict inequality using a small amount of slack, i.e.,

$$a^T x \geq b + (m - \epsilon)z + \epsilon. \tag{6.8}$$

Collectively, we thus model $(a^T x \leq b) \leftrightarrow (z = 1)$ as

$$a^T x \leq b + M(1 - z), \qquad a^T x \geq b + (m - \epsilon)z + \epsilon. \tag{6.9}$$

In a similar fashion, we can model $(z = 1) \rightarrow (a^T x \geq b)$ as

$$a^T x \geq b + m(1 - z) \tag{6.10}$$

and $(z = 0) \rightarrow (a^T x < b)$ as

$$a^T x \leq b + (M + \epsilon)z - \epsilon, \tag{6.11}$$

so we model $(z = 1) \leftrightarrow (a^T x \geq b)$ as

$$a^T x \geq b + m(1 - z), \qquad a^T x \leq b + (M + \epsilon)z - \epsilon, \tag{6.12}$$

using the alower bound $m < a^T x - b$ and upper bound $M > a^T x - b$. We can combine (6.9) and (6.9) to model indication of equality constraints $a^T x = b$. To that end, we note that $(z = 1) \rightarrow (a^T x = b)$ is equivalent to both (6.7) and (6.10), i.e.,

$$a^T x \leq b + M(1 - z), \qquad a^T x \geq b + m(1 - z).$$

On the other hand, $(z = 0) \rightarrow (a^T x \neq b)$ (or equivalently $(z = 0) \rightarrow (a^T x > b) \vee (a^T x < b)$) can be written using (6.8) and (6.11) as

$$a^T x \geq b + (m - \epsilon)z_1 + \epsilon, \quad a^T x \leq b + (M + \epsilon)z_2 - \epsilon, \quad z_1 + z_2 - z \leq 1,$$

where $z_1 + z_2 - z \leq 1$ is equivalent to $(z = 0) \rightarrow (z_1 = 0) \vee (z_2 = 0)$.

6.2.4 Disjunctive constraints

With disjunctive constraints we require that at least one out of a set of constraints is satisfied. For example, we may require that

$$(a_1^T x \leq b_1) \vee (a_2^T x \leq b_2) \vee \cdots \vee (a_k^T x \leq b_k).$$

To that end, let $M > a_j^T x - b_j,\ \forall j$ be a collective upper bound. Then we can model

$$(z = 1) \rightarrow (a_1^T x \leq b_1) \vee (a_2^T x \leq b_2) \vee \cdots \vee (a_k^T x \leq b_k)$$

as

$$z = z_1 + \cdots + z_k \geq 1, \qquad a_j^T x \leq b_j + M(1 - z_j)\ \forall j, \tag{6.13}$$

where $z_j \in \{0, 1\}$ are binary variables. To characterize the reverse implication $(a_1^T x \leq b_1) \vee (a_2^T x \leq b_2) \vee \cdots \vee (a_k^T x \leq b_k) \rightarrow (z = 1)$ we consider the contrapositive

$$(z = 0) \rightarrow (a_1^T x > b_1) \wedge (a_2^T x > b_2) \wedge \cdots \wedge (a_k^T x > b_k),$$

which we can write as

$$a_j^T x \geq b + (m - \epsilon) z + \epsilon, \quad j = 1, \ldots, k,$$

for a lower bound $m < a_j^T x - b_j,\ \forall j$.

6.2.5 Pack constraints

For a *pack constraint* we require that at most one of the constraints are satisfied. Considering (6.13) we can formulate this as

$$z_1 + \cdots + z_k \leq 1, \qquad a_j^T x \leq b_j + M(1 - z_j)\ \forall j.$$

6.2.6 Partition constraints

For a *partition constraint* we require that exactly one of the constraints are satisfied. Considering (6.13) we can formulate this as

$$z_1 + \cdots + z_k = 1, \qquad a_j^T x \leq b_j + M(1 - z_j)\ \forall j.$$

6.2.7 Continuous piecewise-linear functions

Consider the continuous univariate piecewise-linear *non-convex* function $f : [\alpha_1, \alpha_5] \mapsto \mathbb{R}$ shown in Fig. 6.2. At the interval $[\alpha_j, \alpha_{j+1}]$, $j = 1, 2, 3, 4$ we can describe the function as

$$f(x) = \lambda_j f(\alpha_j) + \lambda_{j+1} f(\alpha_{j+1})$$

for some $\lambda_j, \lambda_{j+1} \geq 0$ with $\lambda_j + \lambda_{j+1} = 1$. If we add a constraint that only two (adjacent) variables λ_j, λ_{j+1} can be nonzero we can characterize $f(x)$ over the entire interval $[\alpha_1, \alpha_5]$ as a convex combination,

$$f(x) = \sum_{j=1}^{4} \lambda_j f(\alpha_j).$$

The condition that only two adjacant variables can be nonzero is sometimes called an SOS2 constraint. If we introduce indicator variables z_i for each pair of adjacent variables $(\lambda_i, \lambda_{i+1})$, we can model it as

$$\begin{array}{c}\lambda_1 \leq z_1, \quad \lambda_2 \leq z_1 + z_2, \quad \lambda_3 \leq z_2 + z_3, \quad \lambda_4 \leq z_4 + z_3, \quad \lambda_5 \leq z_4\\ z_1 + z_2 + z_3 + z_4 = 1, \quad z \in \{0,1\}^4,\end{array}$$

which satisfies $(z_j = 1) \rightarrow \lambda_i = 0, i \neq \{j, j+1\}$. Collectively, we can then model the epigraph $f(x) \leq t$ as

$$\begin{array}{cccc} x = \sum_{j=1}^{n} \lambda_j \alpha_j, & \sum_{j=1}^{n} \lambda_j f(\alpha_j) \leq t & & \\ \lambda_1 \leq z_1, & \lambda_j \leq z_j + z_{j-1},\ j = 2, \ldots, n-1, & & \lambda_n \leq z_{n-1}, \\ \lambda \geq 0, & \sum_{j=1}^{n} \lambda_j = 1, & \sum_{j=1}^{n-1} z_j = 1, & z \in \{0,1\}^{n-1}, \end{array} \tag{6.14}$$

for a piecewise-linear function $f(x)$ with n terms. This approach is often called the *lambda-method*.

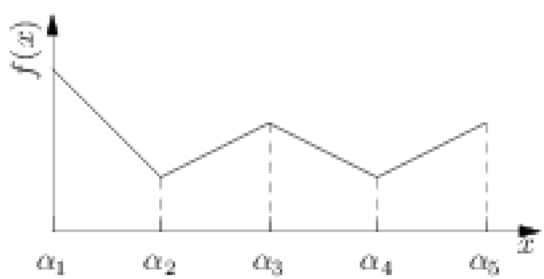
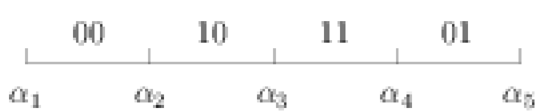


Fig. 6.2: A univariate piecewise-linear non-convex function.

For the function in Fig. 6.2 we can reduce the number of integer variables by using a *Gray encoding*



of the intervals $[\alpha_j, \alpha_{j+1}]$ and an indicator variable $y \in \{0,1\}^2$ to represent the four different values of Gray code. We can then describe the constraints on λ using only two indicator variables,

$$\begin{array}{ccc} (y_1 = 0) & \rightarrow & \lambda_3 = 0 \\ (y_1 = 1) & \rightarrow & \lambda_1 = \lambda_5 = 0 \\ (y_2 = 0) & \rightarrow & \lambda_4 = \lambda_5 = 0 \\ (y_2 = 1) & \rightarrow & \lambda_1 = \lambda_2 = 0, \end{array}$$

which leads to a more efficient characterization of the epigraph $f(x) \leq t$,

$$\begin{array}{c} x = \sum_{j=1}^{5} \lambda_j \alpha_j, \quad \sum_{j=1}^{5} \lambda_j f(\alpha_j) \leq t, \\ \lambda_3 \leq y_1, \quad \lambda_1 + \lambda_5 \leq (1 - y_1), \quad \lambda_4 + \lambda_5 \leq y_2, \quad \lambda_1 + \lambda_2 \leq (1 - y_2), \\ \lambda \geq 0, \quad \sum_{j=1}^{5} \lambda_j = 1, \quad y_1 + y_2 = 1, \quad y \in \{0,1\}^2, \end{array}$$

The lambda-method can also be used to model multivariate continuous piecewise-linear non-convex functions, specified on a set of polyhedra P_k. For example, for the function shown in Fig. 6.3 we can model the epigraph $f(x) \le t$ as

$$
\begin{array}{c}
x = \sum_{i=1}^6 \lambda_i v_i, \qquad \sum_{i=1}^6 \lambda_i f(v_i) \le t,\\
\lambda_1 \le z_1 + z_2, \qquad \lambda_2 \le z_1, \qquad \lambda_3 \le z_2 + z_3,\\
\lambda_4 \le z_1 + z_2 + z_3 + z_4, \qquad \lambda_5 \le z_3 + z_4, \qquad \lambda_6 \le z_4,\\
\lambda \ge 0, \qquad \sum_{i=1}^6 \lambda_i = 1, \qquad \sum_{i=1}^4 z_i = 1,\\
z \in \{0,1\}^4.
\end{array}
\tag{6.15}
$$

Note, for example, that $z_2 = 1$ implies that $\lambda_2 = \lambda_5 = \lambda_6 = 0$ and $x = \lambda_1 v_1 + \lambda_3 v_3 + \lambda_4 v_4$.

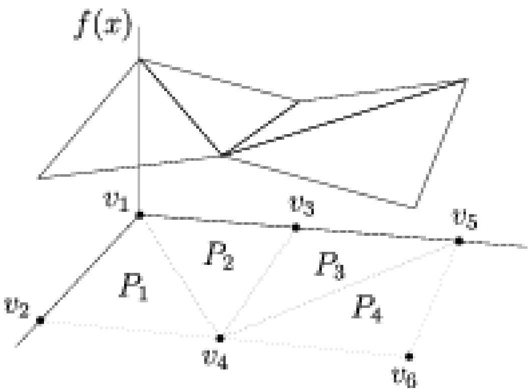


Fig. 6.3: A multivariate continuous piecewise-linear non-convex function.

For this function we can also reduce the number of indicator variables using a Gray encoding of decision variables for the different polyhedra. If we use a decision variable $y \in \{0,1\}^2$ with a Gray encoding

$$
(P_1, P_2, P_3, P_4) \rightarrow (00, 10, 11, 01),
$$

we have

$$
\begin{array}{rcl}
(y_1 = 0) & \rightarrow & \lambda_3 = 0\\
(y_1 = 1) & \rightarrow & \lambda_2 = \lambda_6 = 0\\
(y_2 = 0) & \rightarrow & \lambda_5 = \lambda_6 = 0\\
(y_2 = 1) & \rightarrow & \lambda_1 = \lambda_2 = 0,
\end{array}
$$

which gives the characterization of the epigraph

$$
\begin{array}{c}
x = \sum_{i=1}^6 \lambda_i v_i, \qquad \sum_{i=1}^6 \lambda_i f(v_i) \le t,\\
\lambda_1 \le y_1, \quad \lambda_1 + \lambda_5 \le y_2, \quad \lambda_4 + \lambda_5 \le y_3, \quad \lambda_1 + \lambda_2 \le (1 - y_2),\\
\lambda \ge 0, \qquad \sum_{i=1}^6 \lambda_i = 1, \qquad \sum_{i=1}^2 y_i = 1, \qquad y \in \{0,1\}^2.
\end{array}
\tag{6.16}
$$

For multidimensional piecewise linear functions, a reduction of the number of binary decision variables is only possible when the polyhedral regions have a special topoly, for example a "Union Jack" arrangement as shown in Fig. 6.3. We refer to the recent survey *[VAN10]* for further references and comparisons between different mixed-integer formulations for piecewise linear functions.

6.2.8 Lower semicontinuous piecewise-linear functions

The ideas in Sec. *6.2.7* can be applied to *lower semicontinuous* piecewise-linear functions as well. For example, consider the function shown in Fig. 6.4. If we denote the one-sided limits by $f_-(c) := \lim_{x\uparrow c} f(x)$ and $f_+(c) := \lim_{x\downarrow c} f(x)$, respectively, the one-sided limits, then we can describe the epigraph $f(x) \leq t$ for the function in Fig. 6.4 as

$$
\begin{array}{c}
x = \lambda_1\alpha_1 + (\lambda_2+\lambda_3+\lambda_4)\alpha_2 + \lambda_5\alpha_5,\\
\lambda_1 f(\alpha_1) + \lambda_2 f_-(\alpha_2) + \lambda_3 f(\alpha_2) + \lambda_4 f_+(\alpha_2) + \lambda_5 f(\alpha_3) \leq t,\\
\lambda_1+\lambda_2 \leq z_1, \quad \lambda_3 \leq z_2, \quad \lambda_4+\lambda_5 \leq z_3,\\
\lambda \geq 0, \qquad \sum_{i=5}^{6}\lambda_i = 1, \qquad \sum_{i=1}^{3} z_i = 1, \qquad z \in \{0,1\}^3,
\end{array}
\tag{6.17}
$$

where we a different decision variable for the intervals $[\alpha_1, \alpha_2)$, $[\alpha_2, \alpha_2]$, and $(\alpha_2, \alpha_3]$. As a special case this gives us an alternative characterization of fixed charge models considered in Sec. *6.2.1*.

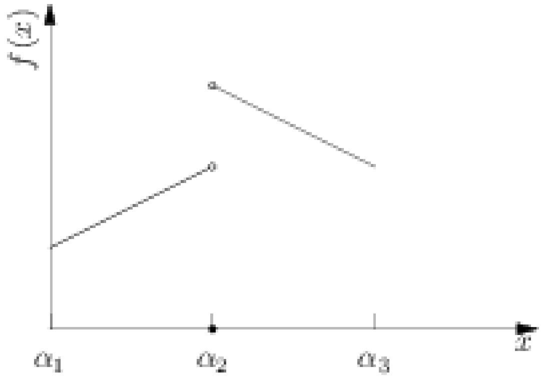


Fig. 6.4: A univariate lower continuous piecewise-linear function.

6.3 Boolean primitives

In the previous sections we used a binary decision variables to indicate whether or not a real valued affine expression was positive. In general, we can express a multitude of boolean expressions as linear inequality constraints involving integer variables. The basic operators from boolean algebra are *complement*, *conjuction* and *disjunction*, which are used to construct more complicated boolean expressions. In the following we let x and y denote binary variables with values $\{0, 1\}$.

6.3.1 Complement

Logical complement $\neg x$ is defined by the following truth table.

x	$\neg x$
0	1
1	0

Thus in terms of an affine relationship,

$$\neg x = 1 - x. \tag{6.18}$$

6.3.2 Conjunction

Logical conjuction $(x \wedge y)$ is defined by the following truth table.

x	y	$x \wedge y$
0	0	0
1	0	0
0	1	0
1	1	1

Thus $z = (x \wedge y)$ is equivalent to

$$z + 1 \geq x + y, \qquad x \geq z, \qquad y \geq z. \tag{6.19}$$

6.3.3 Disjunction

Logical disjuction $(x \vee y)$ is defined by the following truth table.

x	y	$x \vee y$
0	0	0
1	0	1
0	1	1
1	1	1

Thus $(x \vee y)$ is equivalent to

$$x + y \geq 1. \tag{6.20}$$

The following table shows basic properties that are easily verified using the truth tables for complement, conjunction and disjunction.

Associativity	$x \vee (y \vee z) = (x \vee y) \vee z$	$x \wedge (y \wedge z) = (x \wedge y) \wedge z$
Commutativity	$x \vee y = y \vee x$	$x \wedge y = y \wedge x$
Distributivity	$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$	$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$
De Morgan	$\neg(x \wedge y) = \neg x \vee \neg y$	$\neg(x \vee y) = \neg x \wedge \neg y$

The last row in the table shows the properties called *De Morgan's laws*, which shows that conjugation and disjunction in a sense are dual operators.

6.3.4 Implication

Logical implication $(x \rightarrow y)$ is defined by the following truth table.

x	y	$x \rightarrow y$
0	0	1
1	0	0
0	1	1
1	1	1

It is easy verify that that $(x \rightarrow y)$ is equivalent to both $(\neg x \vee y)$ and

$$\neg y \rightarrow \neg x. \tag{6.21}$$

In terms of indicator variables we have that $x \rightarrow y$ is equivalent to

$$x \leq y. \tag{6.22}$$

Example 6.2 (Implication). Consider two real-valued variables $0 \leq u, v \leq M$. We can model complementary $uv = 0$ by introducing two indicator variables x and y,

$$(u > 0) \rightarrow (x = 1), \qquad (v > 0) \rightarrow (y = 1),$$

or equivalently

$$\neg x \rightarrow (u = 0), \qquad \neg y \rightarrow (v = 0).$$

We then model $(u = 0) \vee (v = 0)$ as $(\neg x \vee \neg y)$, which can be written as

$$u \leq Mx, \qquad v \leq My, \qquad (1 - x) + (1 - y) \geq 1.$$

6.3.5 Exclusive disjunction

Logical exclusive disjunction $(x \oplus y)$ is defined by the following truth table.

x	y	$x \oplus y$
0	0	0
1	0	1
0	1	1
1	1	0

Thus $z = x \oplus y$ is equivalent to

$$z \leq x + y, \qquad z \geq x - y, \qquad z \geq -x + y, \qquad z \leq 2 - x - y. \tag{6.23}$$

It is easy to verify from the truth table that $(x \oplus y)$ is equivalent to $(x \vee y) \wedge \neg(x \wedge y)$.

6.4 Integer optimization

A general mixed integer conic optimization problem has the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in \mathcal{C} \\ & x_i \in \mathbb{Z}, \quad \forall i \in \mathcal{I}, \end{array}$$

where $\mathcal{C}$ is direct product of self-dual cones (discussed in Chaps. 2-4) and $\mathcal{I} \subseteq \{1, \ldots, n\}$ denotes the set of variables that are constrained to be integers.

6.5 Bibliography

This chapter is based on the books *[NW88]*, *[Wil93]*. Modeling of piecewise linear functions is described in the survey paper *[VAN10]*.

CHAPTER

SEVEN

PRACTICAL OPTIMIZATION

7.1 Introduction

In this chapter we discuss different practical aspects of creating optimization models.

7.2 The quality of a solution

In this section we will discuss how to validate an obtained solution. Assume we have a conic model with continuous variables only and that the optimization software has reported an optimal primal and dual solution. Given such a solution, we might ask how can we verify that the reported solution is indeed optimal?

To that end, consider a simple the model

$$\begin{array}{lrcl} \text{minimize} & -x_2 & & \\ \text{subject to} & x_1 + x_2 & \leq & 2, \\ & x_2 & \leq & \sqrt{2}, \\ & x_1, x_2 & \geq & 0, \end{array} \tag{7.1}$$

where MOSEK might approximate the solution as

$$x_1 = 0.0000000000000000 \text{ and } x_2 = 1.4142135623730951.$$

and therefore the approximate optimal objective value is

$$1.4142135623730951.$$

The true objective value is $-\sqrt{2}$, so the approximate objective value is wrong by the amount

$$1.4142135623730951 - \sqrt{2}.$$

Most likely this difference is irrelevant for all practical purposes. Nevertheless, in general a solution obtained using floating point arithmetic is only an approximation. Most (if not all) commercial optimization software uses double precision floating point arithmetic, implying that about 16 digits are correct in the computations performed internally by the software. This also means that irrational numbers such as $\sqrt{2}$ and π can only be stored accurately within 16 digits.

A good practice after solving an optimization problem is to evaluate whether the reported solution is indeed an optimal solution; at the very least this process should be caried out during the initial phase of building a model, or if the reported solution is unexpected in some way. The first step in that process is to check that is it feasible; in case of the small example (7.1) this amounts to verifying that

$$\begin{array}{rcrcl} x_1 + x_2 & = & 0.0000000000000000 + 1.4142135623730951 & \leq & 2, \\ x_2 & = & 1.4142135623730951 & \leq & \sqrt{2}, \\ x_1 & = & 0.0000000000000000 & \geq & 0, \\ x_2 & = & 1.4142135623730951 & \geq & 0, \end{array}$$

which demonstrates that the solution is feasible. However, in general the constraints might be violated due to computations in finite precision. It can be difficult to access the significance of a specific violation; for example a violation of one unit in

$$x_1 + x_2 \leq 1$$

may or may not be more significant than a violation of one unit in

$$x_1 + x_2 \leq 10^9.$$

The right-hand side of 10^9 may itself be the result of a computation in finite precision, and my only be known with, say 3 digits of accuracy. Therefore, a violation of 1 unit is not significant since the true right-hand side could just as well be $1.001 \cdot 10^9$.

Another question is how to verify that a feasible approximate solutionis actually optimal, which is answered by duality theory. From the discussion of duality in Sec. *2.3*, the dual problem to (7.1) is

$$\begin{array}{ll} \text{maximize} & 2y_1 + \sqrt{2}y_2 \\ \text{subject to} & \begin{array}{rcl} y_1 & \leq & 0, \\ y_1 + y_2 & \leq & -1. \end{array} \end{array}$$

and weak duality mandates thats the primal objective value is greater than dual objective value for any dual feasible solution, i.e.,

$$-x_2 \geq 2y_1 + \sqrt{2}y_2.$$

Furthermore, if the bound is tight (i.e., if the inequality holds as an equality), then the primal solution must be optimal. In other words, if there exists a dual feasible solution such that the dual objective value is identical to $-\sqrt{2}$, then we know (from duality) that the primal solution is optimal. For example, the optimization software may report

$$y_1 = 0 \text{ and } y_2 = -1,$$

which is easily verified to be a feasible solution with dual objective value $-\sqrt{2}$, proving optimality of both the primal and solution.

7.3 Distance to a cone

Assume we want to solve the conic optimization problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{array}{rcl} Ax & = & b, \\ x \in \mathcal{K} & & \end{array} \end{array}$$

where $\mathcal{K}$ is a convex cone. Let $P(x,\mathcal{K})$ denotes the projection of x onto $\mathcal{K}$. The distance measure

$$d(x,\mathcal{K}) := \min_{y\in\mathcal{K}} \|x-y\|_2 = \|x - P(x,\mathcal{K})\|_2 \tag{7.2}$$

then gives the distance from x to the nearest feasible point in $\mathcal{K}$, and obviously

$$x \in \mathcal{K} \quad \Leftrightarrow \quad d(x,\mathcal{K}) = 0.$$

Thus $d(x,\mathcal{K})$ is natural measure of the cone-feasibility of a point x; if $d(x,\mathcal{K}) > 0$ then $x \notin \mathcal{K}$ (i.e., x is infeasible) and the smaller the value, the closer x is to $\mathcal{K}$. For $x \in \mathbb{R}$, let

$$[x]_+ := \max\{x, 0\}.$$

The following lemmas then give the projections and distance measure for the different cones.

Lemma 7.1 (Distance to linear cones)**.** *The projection is*

$$P(x,\mathbb{R}_+) = [x]_+$$

and the distance is

$$d(x,\mathbb{R}_+) = [-x]_+.$$

Lemma 7.2 (Distance to quadratic cones)**.** *Let $x := (x_1, x_2) \in \mathbb{R}\times\mathbb{R}^{n-1}$. Then*

$$P(x,\mathcal{Q}^n) = \begin{cases} x, & x_1 \geq \|x_2\|, \\ \frac{[x_1+\|x_2\|]_+}{2\|x_2\|}(\|x_2\|, x_2), & x_1 < \|x_2\|, \end{cases}$$

see [FLT02] *prop 3.3, and*

$$d(x,\mathcal{Q}^n) = \begin{cases} \frac{[\|x_2\|-x_1]_+}{\sqrt{2}}, & x_1 \geq -\|x_{2:n}\|, \\ \|x\|, & x_1 < -\|x_{2:n}\|. \end{cases}$$

Lemma 7.3 (Distance to rotated quadratic cones)**.** *Let $x := (x_1, x_2, x_3) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^{n-2}$. Then*

$$P(x,\mathcal{Q}_r^n) = T_n P(T_n x, \mathcal{Q}^n)$$

and

$$d(x,\mathcal{Q}_r^n) = d(T_n x, \mathcal{Q}^n),$$

where T_n is the orthogonal transformation defined in (3.3)*.*

Lemma 7.4 (Distance to semidefinite cones)**.** *Let $X = \sum_{i=1}^n \lambda_i q_i q_i^T \in \mathcal{S}^n$ be an eigendecomposition. Then*

$$P(X,\mathcal{S}_+^n) = \sum_{i=1}^n [\lambda_i]_+ q_i q_i^T,$$

and

$$d(X,\mathcal{S}_+^n) = \max_i [-\lambda_i]_+.$$

CHAPTER

EIGHT

CASE STUDIES

8.1 Introduction

This chapter discusses a number of case studies and examples with details that are beyound the general character of the previous chapters.

8.2 A resource constrained production and inventory problem

The resource constrained production and inventory problem [Zie82] can be formulated as follows:

$$\begin{array}{lll} \mbox{minimize} & \sum_{j=1}^n (d_j x_j + e_j/x_j) & \\ \mbox{subject to} & \sum_{j=1}^n r_j x_j \leq b, & \\ & x_j \geq 0, & j=1,\ldots,n, \end{array} \tag{8.1}$$

where n denotes the number of items to be produced, b denotes the amount of common resource, and r_j is the consumption of the limited resource to produce one unit of item j. The objective function represents production and inventory costs. Let c_j^p denote the cost per unit of product j and c_j^i denote the rate of of holding costs, respectively. Further, let

$$d_j = \frac{c_j^p c_j^r}{2}$$

so that

$$d_j x_j$$

is the average holding costs for product j. If D_j denotes the total demand or product j and c_j^o the ordering cost per order of product j then let

$$e_j = c_j^o D_j$$

and hence

$$\frac{e_j}{x_j} = \frac{c_j^o D_j}{x_j}$$

is the average ordering costs for product j. It is not always possible to produce a fractional number of items. In such cases a constraint saying x_j has to be integer valued should be added.

In summary, the problem finds the optimal batch size such that the inventory and ordering cost are minimized while satisfying the constraints on the common resource. Given $d_j, e_j \geq 0$ problem (8.1) is equivalent to the conic quadratic problem

$$
\begin{array}{lll}
\text{minimize} & \sum_{j=1}^n (d_j x_j + e_j t_j) & \\
\text{subject to} & \sum_{j=1}^n r_j x_j \leq b, & \\
& (t_j, x_j, \sqrt{2}) \in \mathcal{Q}_r^3, & j = 1, \ldots, n, \\
& x_j \geq 0, & j = 1, \ldots, n.
\end{array}
$$

8.3 Markowitz portfolio optimization

8.3.1 Basic notions

Recall from Sec. *3.3.3* that the standard Markowitz portfolio optimization problem is

$$
\begin{array}{ll}
\text{maximize} & \mu^T x \\
\text{subject to} & x^T \Sigma x \leq \gamma \\
& e^T x = 1 \\
& x \geq 0,
\end{array} \tag{8.2}
$$

where $\mu \in \mathbb{R}^n$ is a vector of expected returns for n different assets and $\Sigma \in \mathcal{S}_+^n$ denotes the corresponding covariance matrix.

Problem (8.2) maximizes the expected return of investment given a budget constraint and an upper bound γ on the allowed risk. Alternatively, we can minimize the risk given a lower bound δ on the expected return of investment, i.e., we can solve a problem

$$
\begin{array}{ll}
\text{minimize} & x^T \Sigma x \\
\text{subject to} & \mu^T x \geq \delta \\
& e^T x = 1 \\
& x \geq 0.
\end{array} \tag{8.3}
$$

Both problems (8.2) and (8.3) are equivalent in the sense that they describe the same Pareto-optimal trade-off curve by varying δ and γ. In fact, given a solution to either problem, we can easily find the parameters for the other problem, which results in the same solution. The optimality conditions for (8.3) are

$$
2\Sigma x = \nu\mu + \lambda e + z, \quad \nu(\mu^T x - \delta) = 0, \quad e^T x = 1, \quad x, z, \nu \geq 0, \quad x^T z = 0, \tag{8.4}
$$

where $\nu \in \mathbb{R}_+$, $\lambda \in \mathbb{R}$, $z \in \mathbb{R}_+^n$ are dual variables, and the optimality conditions for (8.2) are

$$
2\alpha\Sigma x = \mu + \zeta e + v, \quad \alpha(x^T \Sigma x - \gamma) = 0, \quad e^T x = 1, \quad \alpha, v \geq 0, \quad x^T v = 0, \tag{8.5}
$$

where $\alpha \in \mathbb{R}_+$, $\zeta \in \mathbb{R}$, $v \in \mathbb{R}_+^n$ are dual variables. Furthermore, from (8.4) we see that

$$
x^T \Sigma x = (1/2)(\nu \mu^T x + \lambda e^T x) = (1/2)(\nu\delta + \lambda).
$$

Thus, given a solution $(x^\star, \nu^\star, \lambda^\star, z^\star)$ to (8.4), we see that for $\gamma = (x^\star)^T \Sigma x^\star = (1/2)(\nu^\star \delta + \lambda^\star)$,

$$
x = x^\star, \quad \alpha = 1, \quad \zeta = \lambda^\star, \quad v = z^\star
$$

is a solution to (8.5).

Next consider a factorization

$$\Sigma = V^T V \tag{8.6}$$

for some $V \in \mathbb{R}^{k\times n}$. We can then rewrite both problems (8.2) and (8.3) in conic quadratic form as

$$\begin{array}{ll} \text{maximize} & \mu^T x \\ \text{subject to} & (1/2, \gamma, Vx) \in \mathcal{Q}_r^{k+2} \\ & e^T x = 1 \\ & x \geq 0, \end{array} \tag{8.7}$$

and

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & (1/2, t, Vx) \in \mathcal{Q}_r^{k+2} \\ & \mu^T x \geq \delta \\ & e^T x = 1 \\ & x \geq 0, \end{array} \tag{8.8}$$

respectively. In practice, a suitable factorization (8.6) is either readily available, or it easily obtained. We mention typical choices below.

- *Data-matrix*. Σ might be specified directly in the form (8.6), where V is a normalized data-matrix with k observations of market data (for example daily returns) of the n assets. When the observation horizon k is shorter than n, which is typically the case, the conic representation is both more parsimonious and has better numerical properties.
- *Cholesky*. Given $\Sigma \succ 0$, we may factor it as (8.6) where V is an upper-triangular Cholesky factor. In this case $k = n$, i.e., $V \in \mathbb{R}^{n\times n}$, so there is little difference in complexity between the conic and quadratic formulations.
- *Factor model*. For a factor model we have

  $$\Sigma = D + U^T U$$

  where $D \succ 0$ is a diagonal matrix, and $U \in \mathbb{R}^{k\times n}$ is a low-rank matrix ($k \ll n$). The factor model directly gives us a factor

  $$V = \left[\begin{array}{c} D^{1/2} \\ U \end{array} \right] \tag{8.9}$$

  of dimensions $(n + k) \times n$. The dimensions of V in (8.9) are larger than the dimensions of the Cholesky factors of Σ, but V in (8.9) is very sparse, which usually results in a significant reduction of solution time.

We can also combine the formulations (8.2) and (8.3) into an alternative problem that more explicitly computes a trade-off between risk and return,

$$\begin{array}{ll} \text{maximize} & \mu^T x - \lambda x^T \Sigma x \\ \text{subject to} & e^T x = 1 \\ & x \geq 0, \end{array} \tag{8.10}$$

with a trade-off parameter $\lambda \geq 0$. One advantage of (8.10) is that the problem is always feasible. This formulation also produces the same optimal trade-off curve, i.e., for a particular (feasible) choice of δ (or γ), there exists a value of λ for which the different problems have the same solution x. The conic formulation of (8.10) is

$$\begin{array}{ll} \text{maximize} & \mu^T x - \lambda t \\ \text{subject to} & (1/2, t, Vx) \in \mathcal{Q}_r^{k+2} \\ & e^T x = 1 \\ & x \geq 0. \end{array} \tag{8.11}$$

8.3.2 Slippage costs

In a more realistic model we correct the expected return of investment by *slippage* terms $T_j(x_j)$,

$$\begin{array}{ll} \text{maximize} & \mu^T x - \sum_j T_j(x_j) \\ \text{subject to} & x^T \Sigma x \leq \gamma \\ & e^T x = 1 \\ & x \geq 0, \end{array} \tag{8.12}$$

which accounts for several different factors.

Market impact

The phenomenon that the price of an asset changes when large volumes are traded is called *market impact*, and for large-volume trading it is the most important slippage contribution. Empirically it has been demonstrated that market impact, i.e., the additional cost of buying or selling an asset is well-approximated by the conic representable function

$$T_j(x_j) = \alpha_j x_j^{3/2}$$

with coefficients $\alpha_j \geq 0$ estimated from historical data. Using Lemma XXX Prop. (iv), we can write the epigraph

$$x_j^{3/2} \leq t_j$$

as

$$(s_j, t_j, x_j), (1/8, x_j, s_j) \in \mathcal{Q}_r^3,$$

i.e, we get a conic formulation of (8.12)

$$\begin{array}{ll} \text{maximize} & \mu^T x - \alpha^T t \\ \text{subject to} & (1/2, \gamma, Vx) \in \mathcal{Q}_r^{k+2} \\ & (s_j, t_j, x_j) \in \mathcal{Q}_r^3 \\ & (1/8, x_j, s_j) \in \mathcal{Q}_r^3 \\ & e^T x = 1 \\ & x \geq 0. \end{array} \tag{8.13}$$

Transaction costs

For smaller volume trading the main slippage contribution is transaction costs, which can modeled as a linear function

$$T_j(x_j) = \alpha_k x_j$$

or as a fixed-charge plus linear function

$$T_j(x_j) = \begin{cases} 0, & x = 0, \\ \alpha_j x_j + \beta_k, & 0 < x_j \le \rho_j, \end{cases}$$

which can be modeled as using indicator variables (see Section *6.2.1*). To that end, let z_j be an indicator variable,

$$(z_j = 0) \rightarrow (x_j = 0)$$

which we can express as

$$x_j \le \rho_j z_j, \quad z_j \in \{0, 1\}. \tag{8.14}$$

We thus get a *conic integer problem*

$$\begin{array}{ll} \text{maximize} & \mu^T x - e^T t \\ \text{subject to} & (1/2, \gamma, Vx) \in \mathcal{Q}_r^{k+2} \\ & \alpha_j x_j + \beta_j z_j \le t_j \\ & x_j \le \rho_j z_j \\ & e^T x = 1 \\ & x \ge 0 \\ & z \in \{0, 1\}^n. \end{array} \tag{8.15}$$

Since the objective function minimizes t_j we exclude the possibility that $x_j = 0$ and $z_j = 1$.

8.3.3 Maximizing the Sharpe ratio

The Sharpe ratio defines an efficiency metric of a portfolio as the expected return per unit risk, i.e.,

$$S(x) = \frac{\mu^T x - r_f}{(x^T \Sigma x)^{1/2}},$$

where r_f denotes the return of a *risk-free* asset. By assumption, we have that $\mu^T x > r_f$ (i.e., there exists a risk-associated portfolio with a greater yield than the risk-free asset), so maximizing the Sharpe ratio is equivalent to minimizing $1/S(x)$. In other words, we have the following problem

$$\begin{array}{lccc} \text{minimize} & \frac{\|V^T x\|}{\mu^T x - r_f} & & \\ \text{subject to} & e^T x & = & 1 \\ & x & \ge & 0. \end{array}$$

We next introduce a variable transformation,

$$y = \gamma \mu^T x, \qquad \gamma \ge 0.$$

Since a positive γ can be chosen arbitrarily and $(\mu - r_f e)^T x > 0$, we have without loss of generality that

$$(\mu - r_f e)^T y = 1.$$

Thus, we obtain the following conic problem for maximizing the Sharpe ratio,

$$\begin{array}{lccc} \text{minimize} & t & & \\ \text{subject to} & (t, Vy) & \in & \mathcal{Q}^{k+1} \\ & e^T y & = & \gamma \\ & (\mu - r_f e)^T y & = & 1 \\ & \gamma & \ge & 0 \\ & y & \ge & 0, \end{array}$$

and we recover $x := y/\gamma$.

CHAPTER

NINE

DUALITY IN CONIC OPTIMIZATION

9.1 Introduction

Duality theory is a rich and powerful area of convex optimization, and central to understanding sensitivity analysis and infeasibility issues in linear (and convex) optimization. Furthermore, it provides a simple and systematic way of obtaining non-trivial lower bounds on the optimal value for many difficult non-convex problem.

9.2 The Lagrangian

Duality theory is inherently connected with a (primal) optimization problem. Initially we consider an optimization problem in a general form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f(x) \leq 0 \\ & h(x) = 0, \end{array} \tag{9.1}$$

where $f_0 : \mathbb{R}^n \mapsto \mathbb{R}$ is the objective funtion, $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ encodes inequality constraints, and $h : \mathbb{R}^n \mapsto \mathbb{R}^p$ encodes equality constraints. We will denote (9.1) the *primal problem*, and its optimal value by $p^\star$. A general problem such as (9.1) is well-suited for our initial discussion of duality (especially weak duality); for our discussion of strong duality in Section *9.5* we only consider a particular class of linear optimization problems.

The *Lagrangian* for (9.1) is a function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \mapsto \mathbb{R}$ that augments the objective with a weigthed combination of all the constraints,

$$L(x, y, s) = f_0(x) + y^T h(x) + s^T f(x). \tag{9.2}$$

The variables $y \in \mathbb{R}^p$ and $s \in \mathbb{R}^m_+$ are called *Lagrange multipliers* or *dual variables*. Readers familiar with the method of penalization for enforcing constraints may recognize the Lagrangian as something similar; if we fix y and s large enough, then the minimizer of $L(x)$ may approximately satisfy $h(x) = 0$ and $f(x) \leq 0$. That is just a (very) simplifying interpretation, however. The main property of the Lagrangian is that it provides a lower bound on feasible solutions to (9.1); for a feasible x we have $y^T h(x) = 0$ and $s^T f(x) \leq 0$ (since $s \geq 0$), so obviously $L(x, y, s) \leq f_0(x)$.

9.3 The dual function and problem

The Lagrange dual function is defined as the minimum of the Lagrangian over x,

$$g(y, s) = \inf_x L(x, y, s),$$

and where the Lagrangian is unbounded below in x, we assign the dual function the value $-\infty$. For fixed (y, s) the dual function is an affine function, so we can think of $g(y, s)$ as the pointwise infimum of infinitely many such affine functions (recall our discussion in Section *2.2.1* of convex piecewise-linear functions defined as the maximum of a set of affine function); thus $g(y, s)$ is a concave function, even when the problem (9.1) is not convex.

The Lagrange dual problem is then obtained by maximixing $g(x, y)$, i.e., is it defined as

$$\begin{array}{ll} \text{maximize} & g(y, s) \\ \text{subject to} & s \geq 0, \end{array} \tag{9.3}$$

which is then always a concave optimization problem.

9.4 Weak duality

We already established that

$$L(x, y, s) \leq f_0(x),$$

which is a global (and weak) inequality, i.e., it holds for all (x, y, s). The Lagrange dual

$$g(y, s) = \inf_x L(x, y, s)$$

provides a bound on the optimal value $p^\star = f_0(x^\star)$ by minimizing over x. The best such lower bound is given by the dual problem (i.e., $d^\star = \sup_{y,s} g(y, s)$), and the relationship

$$p^\star \geq d^\star$$

is called the *weak duality* property, and holds for any kind of optimization problem with a well-defined Lagrangian function (including, e.g., nonconvex problems, but excluding general integer problems).

9.5 Strong duality

Strong duality is one the most important concepts in convex optimization. The problem (9.3) is generally not convex; for that we require that the objective function $f_0(x)$ and the inequalities $f_i(x),\ i = 1, \ldots, m$ are convex functions, and that the equality constraints are affine functions. Then a sufficient condition for strong duality is known as *Slater's constraint qualifications*, which basically assumes that the existence of a strictly feasible point exists for either the primal or dual problem.

In this section we only show strong duality for the linear optimization problem (2.3.1). Since all linear optimization can be rewritten into this form, this incurs no loss of generality. For linear optimization strong duality is traditionally proved using ideas from the simplex algorithm; our proof is based on a more recent approach using separating hyperplanes and Farkas' lemma, which are also important for the discussion of infeasibility in Section *2.4*.

To prove strong duality we need two intermediate results. The first (which we state without proof) is a special case of a more general separating hyperplane theorem.

Theorem 9.1 (Separating hyperplane theorem)**.** *Let S by a closed convex set, and $b \notin S$. Then there exists a strictly separating hyperplane a such that*

$$a^T b > a^T x, \quad \forall x \in S.$$

The geometric interpretating of the (point) *Separating hyperplane theorem* is given in Fig. 9.1, and indicates that a constructive proof follows from projecting b onto S.

We next show an important result, which makes use of the separating hyperplane theorem.

Lemma 9.1 (Farkas lemma)**.** *Given A and b, exactly one of the two propositions is true:*

1. $\exists x \ : \ Ax = b, \quad x \geq 0,$
2. $\exists y \ : \ A^T y \leq 0, \quad b^T y > 0.$

Proof. Both cannot be true, because then

$$b^T y = x^T A^T y \leq 0,$$

which is a contradiction. Next assume that 1. is not true, i.e., $b \notin S = \{Ax \mid x \geq 0\}$. The set S is a closed and convex, so there exists a separating hyperplane y (separating b and S) such that

$$y^T b > y^T Ax, \quad \forall x \geq 0$$

But this implies that $b^T y > 0$ and $A^T y \leq 0$.

We can now prove strong duality.

Proof Assume that the dual optimal value problem (2.3.1) is attained (i.e., $d^\star := b^T y^\star < \infty$). Since $s^\star = c - A^T y^\star \geq 0$ there can then be no z satisfying

$$b^T z > 0, \qquad A^T z \leq 0, \tag{9.4}$$

because otherwise

$$b^T (y^\star + z) > d^\star, \qquad c - A^T (y^\star + z) \geq 0$$

contradicting the optimality assumption on $y^\star$. From Farkas' lemma (9.4) is infeasible if and only if there exists x such that

$$Ax = b, \qquad x \geq 0.$$

For that particular choice of x we have

$$c^T x = x^T A^T y = b^T y^\star,$$

i.e., $p^\star = d^\star$. On the other hand, assume that $-\infty < p^\star < \infty$. From weak duality we then have $d^\star < \infty$, and from Farkas' lemma $d^\star > -\infty$, i.e., the dual optimal value is attained. In other words, if either the primal or dual optimal values are attained, we have $p^\star = d^\star$.

Farkas' lemma is an example of theorems of alternatives, where exactly one of two systems have a solution, and we can also formulate a dual variant.

Lemma 9.2 (Dual variant of Farkas' lemma)**.** *Given A and c, exactly one of the two propositions is true:*

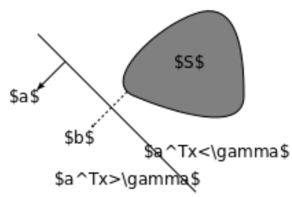


Fig. 9.1: Separating Hyperplane for a closed conves set S and a point $b \in S$.

1. $\exists x \ : \ Ax = 0, \quad x \geq 0, \quad c^T x < 0$,
2. $\exists y \ : \ c - A^T y \geq 0$.

Proof. Both cannot be true, because then

$$c^T x = x^T (c - A^T y) \geq 0.$$

Next assume that 2. is not true, i.e., $\forall y \ : \ c < A^T y$. But then for $x \geq 0$ we have

$$c^T x < x^T A^T y, \quad \forall y$$

implying that $c^T x < 0$ and $Ax = 0$.

CHAPTER TEN

NOTATION AND DEFINITIONS

Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{Z}$ the set of integers, and $\{0,1\}$ the boolean set $\{0,1\}$, respectively. $\mathbb{R}^n$ denotes the set n dimensional vectors of real numbers (and similary for $\mathbb{Z}^n$ and $\{0,1\}^n$); in most cases we denote such vectors by lower case letters, e.g., $a \in \mathbb{R}^n$. A subscripted value a_i then refers to the ith entry in a, i.e.,

$$a = (a_1, a_2, \dots, a_n).$$

All vectors considered in this manual are interpreted as *column-vectors*. For $a, b \in \mathbb{R}^n$ we use the standard inner product,

$$\langle a, b \rangle := a_1 b_1 + a_2 b_2 + \cdots + a_n b_n,$$

which we also write as $a^T b := \langle a, b \rangle$. We let $\mathbb{R}^{m \times n}$ denote the set of $m \times n$ matrices, and we use upper case letters to represent them, e.g., $B \in \mathbb{R}^{m \times n}$ organized as

$$B = \left[\begin{array}{cccc} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{array} \right],$$

i.e, $B_{ij} = b_{ij}$, and for matrices $A, B \in \mathbb{R}^{m \times n}$ we use the inner product

$$\langle A, B \rangle := \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij}.$$

For a vector $x \in \mathbb{R}^n$ we have

$$\mathbf{Diag}(x) := \left[\begin{array}{cccc} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n \end{array} \right],$$

i.e., a square matrix with x on the diagonal and zero elsewhere. Similarly, for a square matrix $X \in \mathbb{R}^{n \times n}$ we have

$$\mathbf{diag}(X) := (x_{11}, x_{22}, \dots, x_{nn}).$$

A set $S \subseteq \mathbb{R}^n$ is *convex* if and only if for any $x, y \in S$ we have

$$\theta x + (1-\theta)y \in S$$

for all $\theta \in [0,1]$. A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex if and only if $\mathbf{dom}(f)$ is convex and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y).$$

for all $\theta \in [0,1]$, where $\mathbf{dom}(f)$ is the *domain* of the function f. A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is *concave* if and only if $-f$ is convex. For example, the function $f : \mathbb{R} \mapsto \mathbb{R}$ given by

$$f(x) = -\log(x)$$

has $\mathbf{dom}(f) = \{x \mid x > 0\}$ and is convex. The *epigraph* of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the set

$$\mathbf{epi}(f) := \{(x,t) \mid x \in \mathbf{dom}(f),\ f(x) \leq t\},$$

shown in Fig. 10.1.

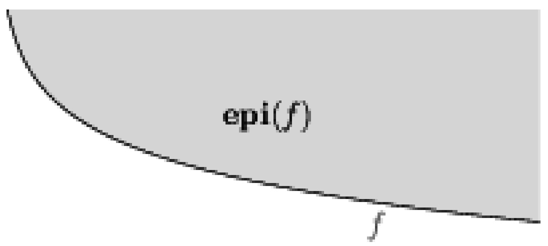


Fig. 10.1: The shaded region is the epigraph of the function $f(x) = -\log(x)$.

Thus, minimizing over the epigraph

$$\begin{array}{ll} \mbox{minimize} & t \\ \mbox{subject to} & f(x) \leq t \end{array}$$

is equivalent to minimizing $f(x)$. Furthermore, f is convex if and only if $\mathbf{epi}(f)$ is a convex set. Similarly, the *hypograph* of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the set

$$\mathbf{hypo}(f) := \{(x,t) \mid x \in \mathbf{dom}(f),\ f(x) \geq t\}.$$

Maximizing f is equivalent to maximizing over the hypograph

$$\begin{array}{ll} \mbox{maximize} & t \\ \mbox{subject to} & f(x) \geq t, \end{array}$$

and f is concave if and only if $\mathbf{hypo}(f)$ is a convex set.

BIBLIOGRAPHY

[AG03] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Math. Programming*, 95(1):3–51, 2003.

[ART02] E. D. Andersen, C. Roos, and T. Terlaky. Notes on duality in second order and p-order cone optimization. *Optimization*, 51(4):627–643, 2002. URL: http://dx.doi.org/10.1080/0233193021000030751, doi:10.1080/0233193021000030751.

[ART03] E. D. Andersen, C. Roos, and T. Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. *Math. Programming*, February 2003.

[BT97] D. Bertsimas and J. N. Tsitsiklis. *Introduction to linear optimization*. Athena Scientific, 1997.

[BKVH04] S.P. Boyd, S.J. Kim, L. Vandenberghe, and A. Hassibi. A Tutorial on Geometric Programming. Technical Report, ISL, Electrical Engineering Department, Stanford University, Stanford, CA, 2004. Available at http://www.stanford.edu/ boyd/gp_tutorial.html.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004. http://www.stanford.edu/ boyd/cvxbook/.

[Chv83] V. Chvátal. *Linear programming*. W.H. Freeman and Company, 1983.

[FLT02] M. Fukushima, Z.Q. Luo, and P. Tseng. Smoothing functions for second-order-cone complementarity problems. *SIAJO*, 12(2):436–460, 2002.

[Hac03] Y. Hachez. *Convex optimization over non-negative polynomials: structured algorithms and applications*. PhD thesis, Université Catholique De Lovain, 2003.

[LR05] M. Laurent and F. Rendl. Semidefinite programming and integer programming. *Handbooks in Operations Research and Management Science*, 12:393–514, 2005.

[LVBL98] M. S. Lobo, L. Vanderberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra Appl.*, 284:193–228, November 1998.

[NW88] G. L. Nemhauser and L. A. Wolsey. *Integer programming and combinatorial optimization*. John Wiley and Sons, New York, 1988.

[Nes99] Yu. Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*. 1999.

[NW06] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science, 2nd edition, 2006.

[PS98] C. H. Papadimitriou and K. Steiglitz. *Combinatorial optimization: algorithms and complexity*. Dover publications, 1998.

[VB96] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, March 1996.

[VAN10] J. P. Vielma, S. Ahmed, and G. Nemhauser. Mixed-integer models for nonseparable piecewise-linear optimization: unifying framework and extensions. *Operations research*, 8():03–315, 2010.

[Wil93] H. P. Williams. *Model building in mathematical programming*. John Wiley and Sons, 3 edition, 1993.

[Zie82] H Ziegler. Solving certain singly constrained convex optimization problems in production planning. *Operations Research Letters*, 1982. URL: http://www.sciencedirect.com/science/article/pii/016763778290030X.

[BenTalN01] A. Ben-Tal and A Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS/SIAM Series on Optimization. SIAM, 2001.