SHAPE CONSTRAINED OPTIMIZATION

WITH APPLICATIONS IN FINANCE AND ENGINEERING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MANAGEMENT SCIENCE & ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Alexandre d'Aspremont
March 2004

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Stephen Boyd (EE)
(Principal Adviser)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

James Primbs

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Laurent El Ghaoui
(EECS, UC Berkeley)

Approved for the University Committee on Graduate Studies:

Abstract

In the first part of this work, we study a particular class of infinite dimensional linear programs on the value of a function at a given number of points, with the additional constraint that this function be convex. Convexity is shown to be the key ingredient making these problems tractable. We detail some applications, with a particular focus on arbitrage constraints between call options.

In the second part, we use the results of chapter one to compute tractable relaxations to some multivariate or basket option pricing problems. We then derive tight price bounds on basket options in some particular cases.

Finally, part three uses some recent results in moment theory and semidefinite programming to refine the convex relaxation techniques of part two and compute tighter constraints linking the prices of basket options.

Contents

List of Tables

List of Figures

Chapter 1

Convexity constraints

In this chapter, we start by presenting an efficient technique to solve the problem of finding optimal bounds on a function at a given point under the following conditions: it must satisfy a set of polyhedral constraints on its value at some given points and must be convex and/or monotone. We observe that this infinite dimensional problem can be cast as a linear program in $O(nm)$ variables and $O(m^2)$ constraints, where n is the dimension and m the number of data points. Finally, we detail various applications in consumer preference assessment, nonlinear pricing, imaging, norm problems, finance, etc.

1.1 Introduction

Infinite dimensional linear programs show up in a wide variety of applications under various names: semi infinite programs, infinite linear programs, continuous linear programs, chance constrained programs, etc... In stark contrast with their finite dimensional counterparts, infinite dimensional linear programs are intractable (NP-Hard) in general as they include for example all polynomial optimization problems. However, a number of good heuristic methods exist to solve some particular cases (see Hettich & Kortanek (1993) for a survey). In this chapter, we observe that a particular class of infinite dimensional linear programs formulated on the values and variations of a convex function at a finite set of points can be reduced to a polynomial size linear program. This simple result turns out to have a broad range of applications and we provide various examples in finance, economics, imaging, etc... Moreover, we will observe in some of these examples that the convexity constraint provides tight relaxations to otherwise intractable problems. Finally, we explore in detail an application of this result to multivariate option pricing.

1.1.1 Problem statement

We shall denote as $z = [f(x_1), \ldots, f(x_m), g_1^T, \ldots, g_m^T]$, the vector composed of the values of a convex continuous function $f : \mathbf{R}^n \to \mathbf{R}$ at a finite set of points $x_1, \ldots, x_m \in \mathbf{R}^n$ and its subgradients at those points $g_1, \ldots, g_m \in \mathbf{R}^n$. We consider the following infinite linear program:

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & Az \preceq b,\ Cz = d \\ & z = \left[f(x_1), \ldots, f(x_m), g_1^T, \ldots, g_m^T\right]^T \\ & f \text{ monotone, convex} \\ & g_i \text{ subgradient of } f \text{ at } x_i, \quad i = 1, \ldots, m, \end{array} \tag{1.1}$$

in the variables $f \in \mathcal{C}(\mathbf{R}^n)$, $z \in \mathbf{R}^{(n+1)m}$ and $g_1, \ldots, g_m \in \mathbf{R}^n$, with parameters $A \in \mathbf{R}^{p\times(n+1)m}$, $C \in \mathbf{R}^{q\times(n+1)m}$, $c \in \mathbf{R}^{(n+1)m}$, $b \in \mathbf{R}^p$, $d \in \mathbf{R}^q$. The feasible set of this problem is convex as the intersection of the set of convex functions with a polyhedron and the objective is linear, which means that problems of the form (1.1) are convex. However, problem (1.1) is infinite dimensional since the main variable f lies in the space $\mathcal{C}(\mathbf{R}^n)$ of continuous functions. Thus, despite its convexity, problem (1.1) is potentially intractable. In what follows, we will observe that a solution can be found efficiently by solving a finite dimensional linear program. We will also detail the solution to a variant of the generic problem (1.1) where the convexity constraint is dropped to keep only a monotonicity requirement.

1.1.2 Existing results

Hereafter, we call infinite linear program (ILP) the infinite dimensional linear programs such as (1.1). These were first formally introduced in Bellman (1953) and later used in, for example, the "bottleneck problem" in Bellman (1957). The entire class of linear programming problems in infinite dimensional spaces is explicitly treated in the book by Anderson & Nash (1987). Some particular instances where it is possible to get strong duality results very similar to those derived in the finite case have received special attention in the literature, these include semi-infinite programs where either the number of variables or the number of constraints is finite (see Hettich & Kortanek (1993) for a recent survey) and separable continuous linear programs (see Bellman (1953) and Pullan (2000)) where the convergence of a general class of algorithms that are not based on discretization can be proved. Here, we will focus on a class of problems where the solution can be obtained with *no discretization error*.

1.2 Linear programming solution

1.2.1 Generic problem

Let us first reformulate (1.1) as a standard form infinite linear program. We can write it as:

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & Az \preceq b,\ Cz = d \\ & z = [f(x_1), \dots, f(x_m), g_1, \dots, g_m]^T \\ & f(y) \leq f(z) \quad \text{if } y \leq z, \quad z, y \in \mathbf{R}^n \\ & f(\theta z + (1-\theta)y) \leq \theta f(z) + (1-\theta) f(y), \quad \theta \in [0,1] \\ & g_i^T (y - x_i) \leq f(y) - f(x_i), \quad i = 1, \dots, m, \end{array} \tag{1.2}$$

and because all the constraints on the variables f and g are linear, the problem is an infinite linear program on the product space $\mathcal{C}(\mathbf{R}^n) \times \mathbf{R}^{(n+1)m}$, with parameters $A \in \mathbf{R}^{p \times (n+1)m}$, $C \in \mathbf{R}^{q \times (n+1)m}$, $c \in \mathbf{R}^{(n+1)m}$, $b \in \mathbf{R}^p$, $d \in \mathbf{R}^q$.

We first look for a lower bound on the solution of that program. We can form a finite-dimensional problem by only enforcing the convexity and subgradient constraints at the points x_i:

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & Az \preceq b,\ Cz = d \\ & z = \left[f(x_1), \dots, f(x_m), g_1^T, \dots, g_m^T\right]^T \\ & g_i^T (x_j - x_i) \leq f(x_j) - f(x_i), \quad i, j = 1, \dots, m, \end{array} \tag{1.3}$$

which is a linear program in the variables $f(x_i)$ and g in $\mathbf{R}^n \times \mathbf{R}^{n \times m}$. Note that the monotonicity constraints are simple positivity constraints on the subgradients, and have been implicitly included in the $Az \preceq b$ inequalities for simplicity.

This linear program produces a lower bound on the optimal value of the original ILP (1.2) because the LP problem (1.3) is formed using a finite subset of the constraints in problem (1.2). We will now show that this bound is actually equal to the optimal value of (1.2).

Proposition 1 *The infinite linear program (1.2) and the linear program (1.3) have the same optimal value and problem (1.2) has a piecewise affine solution that can be constructed explicitly from the solution to (1.3).*

Proof. We shall denote as

$$z^{\text{opt}} = \left[f^{\text{opt}}(x_1), \ldots, f^{\text{opt}}(x_m), g_1^{\text{opt}T}, \ldots, g_m^{\text{opt}T} \right]^T$$

the optimal solution to the linear programming problem (1.3) above and define:

$$s(x) = \max_{i=1,\ldots,m} \left\{ f^{\text{opt}}(x_i) + g_i^{\text{opt}T}(x - x_i) \right\}.$$

$s(x)$ satisfies:

$$s(x_i) = f^{\text{opt}}(x_i), \quad i = 1, \ldots, m,$$

and, by construction, $s(x)$ attains the lower bound $c^T z^{\text{opt}}$ computed by solving problem (1.3). Because $s(x)$ is monotone convex as the pointwise maximum of monotone affine functions, it is also a feasible point of problem (1.2), hence problem (1.2) and (1.3) must have the same optimal value and $s(x)$ is an optimal solution to the infinite linear program in (1.2). ■

The result above shows that, because of the convex shape constraints, a global optimum to the infinite program can be found by sampling the constraints only at the data points.

1.2.2 Monotone variation

Let us introduce a small variation to the generic problem in (1.2) where we drop the convex constraint. The only global requirement that remains is that the function be monotone. The problem now becomes:

$$
\begin{array}{ll}
\text{minimize} & c^T z \\
\text{subject to} & Az \preceq b, \; Cz = d \\
& z = [f(x_1), \ldots, f(x_m)]^T \\
& f \text{ monotone.}
\end{array}
\tag{1.4}
$$

The problem is again an infinite linear program in the variables (z, f), on the product of the space $BV(\mathbf{R}^n)$ of functions on $\mathbf{R}^n$ with bounded variations and $\mathbf{R}^m$, with parameters $A \in \mathbf{R}^{p\times m}$, $C \in \mathbf{R}^{q\times m}$, $c \in \mathbf{R}^m$, $b \in \mathbf{R}^p$, $d \in \mathbf{R}^m$. Without loss of generality, we can focus on the case where f has to be nondecreasing. Again, we first look for a lower bound and form a finite dimensional problem by only enforcing the monotonicity constraints at the points x_i:

$$
\begin{array}{ll}
\text{minimize} & c^T z \\
\text{subject to} & Az \preceq b, \; Cz = d \\
& z = [f(x_1), \ldots, f(x_m)]^T \\
& f(x_i) \leq f(x_j) \text{ if } x_i \preceq x_j, \quad (i,j) = 1, \ldots, m,
\end{array}
\tag{1.5}
$$

this is a linear program in the variables $(f(x_1), \ldots, f(x_m), z)$ in $\mathbf{R}^m \times \mathbf{R}^m$, with the same set of parameters as above. Because this linear program forms a set of necessary conditions on the values of the feasible points of (1.4), the optimum of (1.5) is a lower bound on that of (1.4). Again, let us show that this bound is actually equal to the optimal value of the infinite program (1.4).

Proposition 2 *The infinite linear program (1.4) and the linear program (1.5) have the same optimal value and problem (1.4) has a piecewise constant solution that can be constructed explicitly from the solution to (1.5).*

Proof. We shall denote as $z^{\text{opt}} = [f^{\text{opt}}(x_1), \ldots, f^{\text{opt}}(x_m)]^T$ the optimal solution to the linear programming problem (1.5) above. Let $\{P_i\}_{i=1,\ldots,l}$ be a partition of $\mathbf{R}^n$ in rectangles such that $x_i \in P_i$ for $i = 1, \ldots, m$ and $x_j \notin P_i$ for $i \neq j$ with $i, j = 1, \ldots, m$. Let us define a piecewise constant function f such that $f(x) = f^{\text{opt}}(x_i)$ for $x \in P_i$. To complete the definition of f as a monotone function, we pick some additional points $x_i \in P_i$, for $i = (m+1), \ldots, l$ and set:

$$f(x_i) = \max \{f(x_j) \mid i \neq j,\ x_j \preceq x_i\}$$

and we take $f(x_i) = \min_{j=1,\ldots,m} f(x_j)$ if the set $\{x_j \mid i \neq j,\ x_j \preceq x_i\}$ is empty. By construction, f attains the lower bound found in (1.5) and is monotone, hence it is an optimal solution of (1.4). ■

1.3 Applications

In this section, we detail some applications of the result in proposition 1.

1.3.1 Preference assessment

Here, we show that the concave preference assessment problem as posed by Meyer & Pratt (1968) (see also Pratt, Raiffa & Schlaifer (1964), Keeney & Raiffa (1993) or Keeney (1977)) can be solved exactly and we extend this result to multidimensional utilities.

We consider different *baskets of goods*, consisting of different amounts of n consumer goods. A goods basket is specified by a vector $x \in \mathbf{R}^n$ where x_i denotes the amount of consumer good i. We assume the amounts are normalized so that $0 \leq x_i \leq 1$, *i.e.*, $x_i = 0$ is the minimum and $x_i = 1$ is the maximum possible amount of good i. Given two baskets of goods x and $\tilde{x}$, a consumer can either prefer x to $\tilde{x}$, or prefer $\tilde{x}$ to x, or consider x and $\tilde{x}$ equally attractive. We consider one model consumer, whose choices are repeatable.

We model consumer preference in the following way. We assume there is an underlying *utility function* $u : \mathbf{R}^n \to \mathbf{R}$, with domain $[0,1]^n$; $u(x)$ gives a measure of the utility derived by the consumer from the goods basket x. Given a choice between two baskets of goods, the consumer chooses the one that has larger utility, and will be ambivalent when the two baskets have equal utility. It is reasonable to assume that u is monotone nondecreasing. This means that the consumer always prefers to have more of any good, with the amounts of all other goods the same. It is also reasonable to assume that u is concave. This models *satiation*, or decreasing marginal utility as we increase the amount of goods.

Now suppose we are given some consumer preference data, but we do not know the underlying utility function u. Specifically, we have a set of goods baskets $a_1, \ldots, a_m \in$

$[0,1]^n$, and some information about preferences among them:

$$u(a_i) > u(a_j) \text{ for } (i,j) \in \mathcal{P}, \qquad u(a_i) \geq u(a_j) \text{ for } (i,j) \in \mathcal{P}_{\text{weak}}, \tag{1.6}$$

where $\mathcal{P},\ \mathcal{P}_{\text{weak}} \subseteq \{1,\ldots,m\} \times \{1,\ldots,m\}$ are given. Here $\mathcal{P}$ gives the set of known preferences: $(i,j) \in \mathcal{P}$ means that basket a_i is known to be preferred to basket a_j. The set $\mathcal{P}_{\text{weak}}$ gives the set of known weak preferences: $(i,j) \in \mathcal{P}_{\text{weak}}$ means that basket a_i is preferred to basket a_j, or that the two baskets are equally attractive.

We first consider the following question: How can we determine if the given data are consistent, *i.e.*, whether or not there exists a concave nondecreasing utility function u for which (1.6) holds? This is equivalent to solving the feasibility problem

$$\begin{array}{ll} \text{find} & u \\ \text{subject to} & u : \mathbf{R}^n \to \mathbf{R} \text{ concave and nondecreasing} \\ & u(a_i) > u(a_j), \quad (i,j) \in \mathcal{P} \\ & u(a_i) \geq u(a_j), \quad (i,j) \in \mathcal{P}_{\text{weak}}, \end{array} \tag{1.7}$$

with the function u as the (infinite-dimensional) optimization variable. Since the constraints in (1.7) are all homogeneous, we can express the problem in the equivalent form

$$\begin{array}{ll} \text{find} & u \\ \text{subject to} & u : \mathbf{R}^n \to \mathbf{R} \text{ concave and nondecreasing} \\ & u(a_i) \geq u(a_j) + 1, \quad (i,j) \in \mathcal{P} \\ & u(a_i) \geq u(a_j), \quad (i,j) \in \mathcal{P}_{\text{weak}}, \end{array} \tag{1.8}$$

which uses only nonstrict inequalities. (It is clear that if u satisfies (1.8), then it must satisfy (1.7); conversely, if u satisfies (1.7), then it can be scaled to satisfy (1.8).) This problem, in turn, can be cast as a (finite-dimensional) linear programming feasibility problem, using the interpolation result in prop. 1:

$$\begin{array}{ll} \text{find} & u_1, \dots, u_m,\ g_1, \dots, g_m \\ \text{subject to} & g_i \succeq 0, \quad i = 1, \dots, m \\ & u_j \le u_i + g_i^T (a_j - a_i), \quad i, j = 1, \dots, m \\ & u_i \ge u_j + 1, \quad (i, j) \in \mathcal{P} \\ & u_i \ge u_j, \quad (i, j) \in \mathcal{P}_{\text{weak}}. \end{array} \tag{1.9}$$

By solving this linear programming feasibility problem, we can determine whether there exists a concave, nondecreasing utility function that is consistent with the given sets of strict and nonstrict preferences. If (1.9) is feasible, there is at least one such utility function (and indeed, we can construct one that is piecewise-linear, from a feasible $u_1, \dots, u_m$, $g_1, \dots, g_m$). If (1.9) is not feasible, we can conclude that there is no concave increasing utility function that is consistent with the given sets of strict and nonstrict preferences.

As an example, suppose that $\mathcal{P}$ and $\mathcal{P}_{\text{weak}}$ are consumer preferences that are known to be consistent with at least one concave increasing utility function. Consider a pair (k, l) that is not in $\mathcal{P}$ or $\mathcal{P}_{\text{weak}}$, *i.e.*, consumer preference between baskets k and l is not known. In some cases we can conclude that a preference holds between basket k and l, even without knowing the underlying preference function. To do this we augment the known preferences (1.6) with the inequality $u(a_k) \le u(a_l)$, which means that basket l is preferred to basket k, or they are equally attractive. We then solve the feasibility linear program (1.9), including the extra weak preference $u(a_k) \le u(a_l)$. If the augmented set of preferences is infeasible, it means that any concave nondecreasing utility function that is consistent with the original given consumer preference data must also satisfy $u(a_k) > u(a_l)$. In other words, we can conclude that basket k is preferred to basket l, without knowing the underlying utility function.

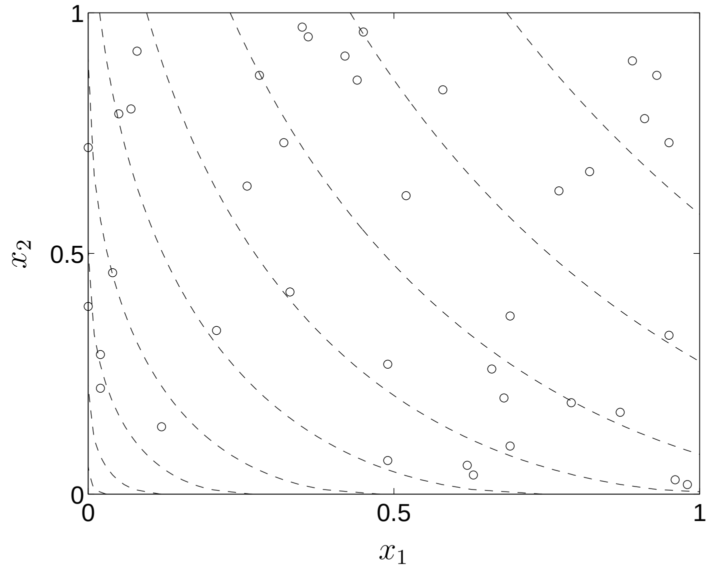

Figure 1.1: Forty goods baskets $a_1, \ldots, a_{40}$, shown as circles. The 0.1, 0.2, ..., 0.9 level curves of the true utility function u are shown as dashed lines. This utility function is used to find the consumer preference data $\mathcal{P}$ among the 40 baskets.

Example

Here we give a simple numerical example that illustrates the discussion above. We consider baskets of two goods (so we can easily plot the goods baskets). To generate the consumer preference data $\mathcal{P}$, we compute 40 random points in $[0,1]^2$, and then compare them using the utility function

$$u(x_1, x_2) = (1.1x_1^{1/2} + 0.8x_2^{1/2})/1.9.$$

These goods baskets, and a few level curves of the utility function u, are shown in figure 1.1.

We now use the consumer preference data (but not, of course, the true utility function u) to compare each of these 40 goods baskets to the basket $a_0 = (0.5, 0.5)$. For each original basket a_i, we solve the linear programming feasibility problem described above, to see if we can conclude that basket a_0 is preferred to basket a_i. Similarly, we check whether we can conclude that basket a_i is preferred to basket a_0. For each basket a_i, there are three possible outcomes: we can conclude that a_0 is definitely preferred to a_i, that a_i is definitely preferred to a_0, or (if both LP feasibility problems are feasible) that no conclusion is possible. (Here, *definitely preferred* means that the preference holds for any concave nondecreasing utility function that is consistent with the original given data.)

We find that 21 of the baskets are definitely rejected in favor of $(0.5, 0.5)$, and 14 of the baskets are definitely preferred. We cannot make any conclusion, from the consumer preference data, about the remaining 5 baskets. These results are shown in figure 1.2. The vertical and horizontal lines passing through $(0.5, 0.5)$ divide $[0, 1]^2$ into four quadrants. Points in the upper right quadrant must be preferred to $(0.5, 0.5)$, by the monotonicity assumption on u. Similarly, $(0.5, 0.5)$ must be preferred to the points in the lower left quadrant. For the points in the other two quadrants, the results are not obvious. So for these 17 points, there is no need to solve the feasibility LP (1.9). Classifying the 23 points in the other two quadrants, however, requires the concavity assumption, and solving the feasibility LP (1.9).

1.3.2 Multidimensional screening

We describe here the multidimensional screening problem formulated by Rochet & Chone (1998)[1], in a setting derived from the general nonlinear pricing and optimal taxation problems detailed in Wilson (1993) or Mirrlees (1971) respectively.

A monopolist has to design a product line $Q \subseteq \mathbf{R}^n$ and a price schedule $p : Q \to \mathbf{R}$, to jointly maximize her profit based on the knowledge of the distribution of consumer types. We shall denote this distribution's density $f(t) \in \mathcal{C}^1\left(\mathbf{R}^n_+\right)$. As in Rochet & Chone (1998), without loss of generality, we can assume that a consumer of type t has a bilinear individual utility function given by $u(t, q) = t^T q$ for $(t, q) \in \mathbf{R}^n_+ \times Q$.

[1]We are very grateful to Bob Wilson for suggesting this example.

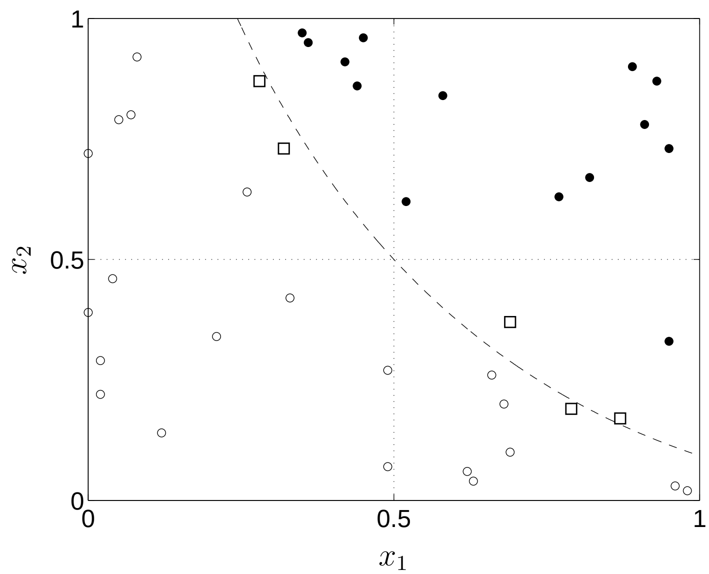

Figure 1.2: Results of consumer preference analysis using the LP (1.9), for a new goods basket $a_0 = (0.5, 0.5)$. The original baskets are displayed as open circles if they are definitely rejected ($u(a_k) < u(a_0)$), as solid black circles if they are definitely preferred ($u(a_k) > u(a_0)$), and as squares when no conclusion can be made. The level curve of the underlying utility function, that passes through $(0.5, 0.5)$, is shown as a dashed curve.

We suppose that there is an outside good with quality $q_0 \in Q$ and price p_0 so that the investor will choose the product q solution to:

$$U(t) = \max_{q \in Q} tq - p(q)$$

if and only if $U(t) \geq q_0 t - p_0$ and the outside product otherwise. Finally, we make the simplifying assumption that the monopolist's cost function is linear:

$$C(q) = c^T q, \quad q \in Q.$$

Following Rochet & Chone (1998) we know that the monopolist's problem is then given by:

$$\begin{array}{ll} \text{maximize} & \int_{\mathbf{R}_+^n} \left((t-c)\nabla U(t) - U(t)\right) f(t)dt \\ \text{subject to} & U \text{ convex continuous} \\ & U(t) \geq tq_0 - p_0 \end{array} \tag{1.10}$$

We can approximate the expected profit $\mathbf{E}\left((t-c)^T \nabla U(t) - U(t)\right)$ using Monte-Carlo as

$$\mathbf{E}\left((t-c)\nabla U(t) - U(t)\right) = 1/N \sum_{i=1}^{N} (t_i - c)\nabla U(t_i) - U(t_i),$$

where the t_i are N random sample points distributed with density $f(t)$. The monopolist's problem then becomes:

$$\begin{array}{ll} \text{maximize} & 1/N \sum_{i=1}^{N} (t_i - c)\nabla U(t_i) - U(t_i) \\ \text{subject to} & U \text{ convex continuous} \\ & U(t) \geq tq_0 - p_0 \end{array}$$

which can be solved efficiently using proposition 1. Hence, in a particular case, proposition 1 has allowed us to reduce the complex variational problem in (1.10) into a linear program.

1.3.3 Imaging

In a first simple example, suppose that a laser is scanning a convex object and allows the detection of both range and gradient at a set of points x_i on the object's surface. The problem of reconstructing an approximate image of the object using this

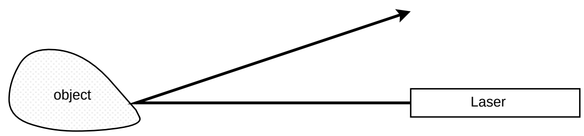

Figure 1.3: 3D laser scanning.

information can be formulated as:

$$\begin{array}{ll} \text{find} & g_i \\ \text{subject to} & g_i^T(x_j - x_i) \leq f(x_j) - f(x_i) \quad i, j = 1, \ldots, m, \end{array}$$

and proposition 1 allows a direct reconstruction of an estimated shape by solving for the subgradients at the points x_i.

Hansen & Lauritzen (1998) or Hansen, Gill & Baddeley (1996) show a similar result on the linear contact distribution. If we let X be a stationary random set in $\mathbf{R}^n$. The problem is, based on the observable data on X, to estimate the function F defined as follows. Let

$$\rho_B(A) = \inf\{r \geq 0 : \; rB \cap A \neq \emptyset\}$$

the function F is then defined as

$$F(r) = P(\rho_B(X) \leq r).$$

Hansen et al. (1996) show that if X is a random closed set in $\mathbf{R}^n$, then F must be concave. This is used to discriminate between different kinds of heat treatments for milk.

1.3.4 Statistics

We know from Grenander (1956) or, more recently, Groeneboom, Jongbloed & Wellner (2001) that the maximum likelihood estimator for a distribution with concave CDF, is precisely the least concave majorant of the sample CDF. Thus, the estimation problem can be reduced to a concave interpolation problem as in (1.1). Further applications of the linear contact distribution to the study of spatial patterns can be found in Hansen et al. (1996), Hansen & Lauritzen (1998) or Serra (1982), while applications of nonparametric estimation with monotonicity and/or convexity constraints are discussed in Groeneboom et al. (2001).

1.3.5 Norm problems

Here, we are given a set of points x_i in $\mathbf{R}^n$ and a corresponding set of intervals $[\alpha_i, \beta_i] \subset \mathbf{R}_+$ and we want to determine if there is a norm f such that $f(x_i) \in [\alpha_i, \beta_i]$ for $i = 1, \ldots, m$. We also want to compute bounds on the norm of another element x_0. In general, we can write this type of problems as the following infinite dimensional program:

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & Az \preceq b,\ Cz = d \\ & z = [f(x_0), \ldots, f(x_m)]^T \\ & f \text{ is a norm,} \end{array} \tag{1.11}$$

in the variables $(f, z) \in \mathcal{C}(\mathbf{R}^n) \times \mathbf{R}^{m+1}$, with parameters $A \in \mathbf{R}^{p\times(m+1)}$, $C \in \mathbf{R}^{q\times(m+1)}$, $c \in \mathbf{R}^{m+1}$, $b \in \mathbf{R}^p$, $d \in \mathbf{R}^q$.

Proposition 3 *We can find the solution to the problem in (1.11) by solving:*

$$\begin{array}{ll} \textit{minimize} & c^T z \\ \textit{subject to} & [A,0]z \preceq b,\ [C,0]z = d \\ & z = \left[f(x_0), \ldots, f(x_m), g_1^T, \ldots, g_m^T\right]^T \\ & g_i^T(x_j - x_i) \leq f(x_j) - f(x_i), \quad i,j = 0,\ldots,m \\ & g_i^T x_i = f(x_i), \quad i = 0,\ldots,m, \end{array} \tag{1.12}$$

which is a linear program in the variables $f(x_i) \in \mathbf{R}$ *and* $g_i \in \mathbf{R}^n$*.*

Proof. Let us first show that the above set of conditions is necessary. To be a norm, the real-valued function f must satisfy the following properties:

- $f(x) > 0$ for all $x \neq 0$,
- $f(x_1 + x_2) \leq f(x_1) + f(x_2)$ for all $x_1,\ x_2 \in \mathbf{R}^n$,
- $f(\lambda x) = |\lambda|\, f(x)$ for all $x \in \mathbf{R}^n,\ \lambda \in \mathbf{R}$.

Using Euler's theorem for homogeneous functions we also know that if f is homogeneous of degree r and differentiable, then for any $x \in \mathbf{R}^n$, $\nabla f(x)^T x = r f(x)$. If f is convex, we can replace ∇f by a subgradient g. This proves that the constraints in program (1.12) are necessary conditions on f, hence problem (1.12) provides a lower bound on the optimum of problem (1.11). Let

$$z^{\text{opt}} = \left[f^{\text{opt}}(x_1), \ldots, f^{\text{opt}}(x_m), g_1^{\text{opt}T}, \ldots, g_m^{\text{opt}T}\right]$$

be the optimal solution to the linear programming problem (1.12) above. Let h_i for $i = 1, \ldots, l$, be a set of vectors such that

$$\text{span}(h_i)_{i=1,\ldots,l} \oplus \text{span}(g_i^{\text{opt}})_{i=1,\ldots,m} = \mathbf{R}^n$$

where we choose h such that $\max_{i,j}\{\left|h_i^T x_j\right|\} \leq \min_{i,j}\{\left|g_i^{\mathrm{opt}T} x_j\right|\}$. We then define:

$$n(x) = \max\left\{ \max_{i=1,\ldots,l} \left\{\left|h_i^T x\right|\right\}, \max_{i=1,\ldots,m} \left\{\left|g_i^{\mathrm{opt}T} x\right|\right\} \right\}.$$

Let us show that $n(x)$ is in fact a norm:

- $n(x) > 0$, for all $x \neq 0$. Suppose $n(x) = 0$, then $h_i^T x = 0$ for $i = 1, \ldots, l$ and $g_i^T x = 0$ for $i = 1, \ldots, m$, which implies $x = 0$ because $\mathrm{span}(h_i)_{i=1,\ldots,l} \oplus \mathrm{span}(g_i)_{i=1,\ldots,m} = \mathbf{R}^n$.
- $n(\lambda x) = |\lambda|\, n(x)$, for all $x \in \mathbf{R}^n$, $\lambda \in \mathbf{R}$. This holds true because n is the maximum of homogeneous functions of degree one and $n(x) = n(-x)$.
- $n(x_1 + x_2) \leq n(x_1) + n(x_2)$, for all x_1, $x_2 \in \mathbf{R}^n$. Because n is homogeneous of degree one and convex as the pointwise maximum of affine functions, it satisfies the triangle inequality.

Also, by construction, we have $n(x_i) = f^{\mathrm{opt}}(x_i)$, so $n(x)$ is a norm that attains the lower bound computed in (1.12), hence it is an optimal solution of the original problem (1.11). ■

1.3.6 Interpolation with convex functions

Perhaps the simplest application example is to compute the least-squares fit of a convex function to a given data set (u_i, y_i), $i = 1, \ldots, m$:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m (y_i - f(u_i))^2 \\ \text{subject to} & f : \mathbf{R}^n \to \mathbf{R} \text{ is convex}, \quad \mathbf{dom}\, f = \mathbf{R}^n. \end{array}$$

This is an infinite-dimensional problem, since the variable is f, which is in the space of continuous real-valued functions on $\mathbf{R}^n$. Using the result in proposition 1, we can

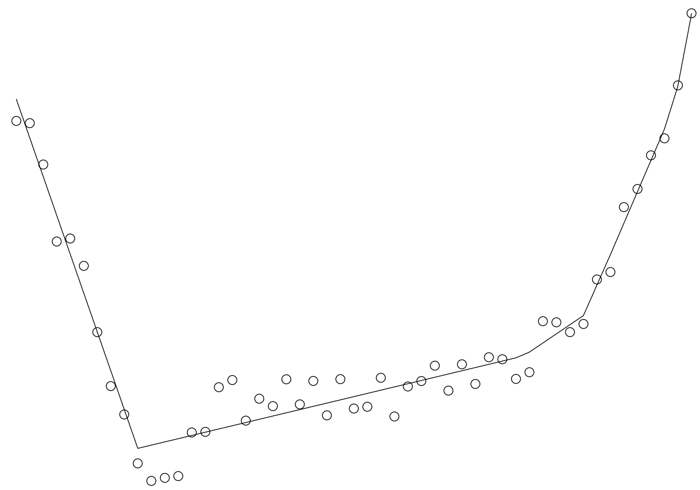

Figure 1.4: Least-squares fit of a convex function to data, shown as circles. The (piecewise-linear) function shown minimizes the sum of squared fitting error, over all convex functions.

formulate this problem as

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m (y_i - \hat{y}_i)^2 \\ \text{subject to} & \hat{y}_j \geq \hat{y}_i + g_i^T(u_j - u_i), \quad i,\ j = 1, \ldots, m, \end{array}$$

which is a quadratic program with variables $\hat{y} \in \mathbf{R}^m$ and $g_1, \ldots, g_m \in \mathbf{R}^n$. The optimal value of this problem is zero if and only if the given data can be interpolated by a convex function, *i.e.*, if there is a convex function that satisfies $f(u_i) = y_i$. An example is shown in figure (1.4).

1.4 Option pricing

The now classic Black & Scholes (1973) option pricing model is based on two fundamental assumptions. The first supposes that agents can trade dynamically in stocks and cash to hedge their position in the option, this is called *dynamic hedging*. The second models the asset dynamics as following *lognormal processes*. It is then quite natural that the option prices obtained via these modelling assumptions be very sensitive to illiquidity, transaction costs and model risk (stocks not being lognormal). Here we take the complementary approach, asking what can be said about option prices with a very minimal set of assumptions, *i.e.* with little trading involved and no assumption on the assets' distribution.

The valuation of derivative securities (see Duffie (1996) for example) is based on the conic duality between portfolios and probability measures in the *fundamental theorem of asset pricing*, which (roughly) states that:

$$\text{Absence of Arbitrage} \quad \Leftrightarrow \quad \text{Price} = \mathbf{E}_\pi[\text{Payoff}]$$

where π is a probability measure. The exact meaning of arbitrage depends on what set of trading strategies one is willing to allow, but this result says that there is no arbitrage if and only if we can evaluate securities in a very natural way: by computing their expected payoff with respect to a certain probability measure.

In the classic Black & Scholes (1973) model case, these strategies are all continuous self-financing portfolios (barring some extreme doubling strategies, etc...). Here however, we will only allow trading today and at the option's maturity and we will make no assumption on the assets' distribution. The corresponding notion of *arbitrage* is then very much immune to liquidity issues and independent of any assumption on the dynamics of the assets. This means that we can expect this method to give much more *robust* information on prices, however we also have to expect this information to be much *coarser*.

1.4.1 Some definitions

In our context, we define an *arbitrage* as the following buy-and-hold strategy:

- form a portfolio at no cost today with a strictly positive payoff,
- carry it without further trading until maturity,
- liquidate the entire position at maturity.

When such a strategy exists, it allows for unlimited profits with zero initial investment, hence prices allowing for such arbitrage opportunities are not viable. We will sometimes refer to such a strategy as a *static arbitrage* to clearly distinguish it from the dynamic arbitrage strategies defining prices in the Black & Scholes (1973).

A *call* with maturity T and *strike* K on the stock S is a security that pays $(S-K)_+$ at a certain maturity date T, where S is the value of the stock at time T. A *put* is defined in a similar way, except that the payoff is given instead by $(K-S)_+$. We shall denote as $C(K)$ the price today of the call with terminal payoff $(S-K)_+$. As an example, in figure (1.5), we plot a set of market prices for IBM calls with maturity one week, arranged by strike price. Finally, the *forward price* of the stock S is the price today of getting the asset S at maturity T. We now look at what constraints link option prices in the absence of static arbitrage in this setting.

Call put parity

If we know the forward prices, then we can deduce call prices from puts using the classical "call put parity" relationship. It states that buying a put option and selling a call with identical strike K produces the same payoff at maturity as selling (short) the stock and holding K in cash. Then, absence of arbitrage states that we must have:

$$\text{put} - \text{call} = K - S.$$

This is illustrated in figure (1.6).

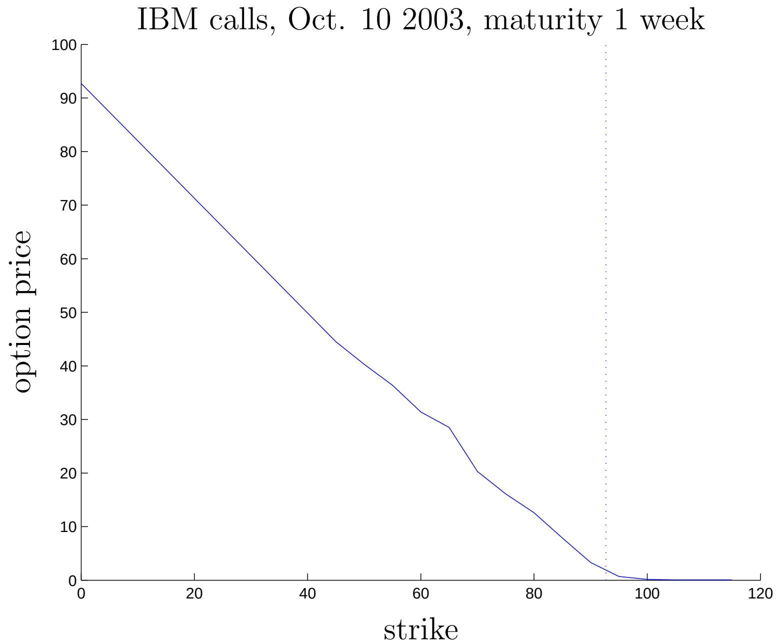

Figure 1.5: IBM call prices vs. strike K, vertical dashed line is the forward price (source: REUTERS).

Call spread

A *call spread* is formed by buying a call with strike price K, while selling a call with a higher strike price $K+\epsilon$, see figure (1.7). Since the payoff of such a strategy is strictly positive, absence of arbitrage implies that its price today be positive. So call prices must be *decreasing* with strike to preclude arbitrage:

$$C(K+\epsilon) - C(K) \leq 0,$$

for $\epsilon > 0$.

Butterfly spread

A *butterfly spread* is formed by buying a call with strike price $K-\epsilon$, while selling two calls with a higher strike price K and buying another call with strike price $K+\epsilon$,

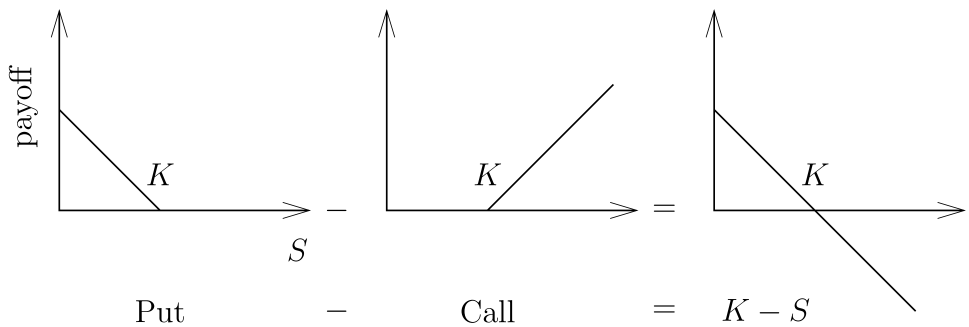

Figure 1.6: Call put parity.

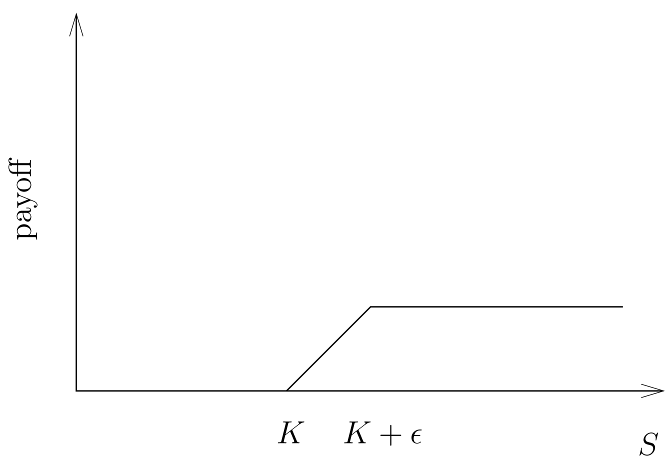

Figure 1.7: Call spread.

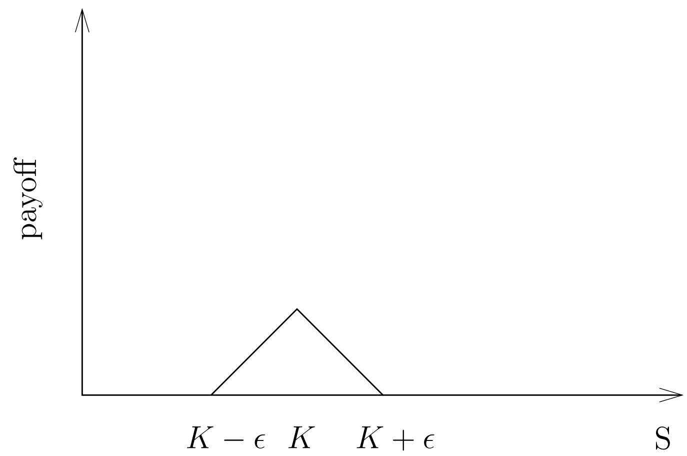

Figure 1.8: Butterfly spread.

see figure (1.8). Again, since the payoff of such a strategy is strictly positive, absence of arbitrage implies that the price of a butterfly spread be positive, hence call prices must be *convex* with strike:

$$C(K+\epsilon) - 2C(K) + C(K-\epsilon) \geq 0,$$

for $\epsilon > 0$.

1.4.2 Sufficient conditions

To summarize, the *absence of arbitrage* implies that if $C(K)$ is a function giving the price of an option of strike K, then $C(K)$ must be positive, nonincreasing and convex. With $C(0) = S$, we have a set of *necessary* conditions on call prices for the absence of arbitrage

In fact, Breeden & Litzenberger (1978), Laurent & Leisen (2000) or Bertsimas & Popescu (2002) among others, show that these conditions are also *sufficient*:

Proposition 4 *Suppose we have a set of market prices for calls $C(K_i) = p_i$, then there is no arbitrage iff there is a function C(K) such that:*

1. $C(K)$ *positive*
2. $C(K)$ *decreasing*
3. $C(K)$ *convex*
4. $C(K_i) = p_i$ *and* $C(0) = S, \quad i = 1, \ldots, m.$

We remark that this is a shape constrained feasibility problem of the form (1.1). If we look back at the market prices in figure (1.5), we notice that the prices are far from satisfying the arbitrage conditions above. In particular, the convexity condition is violated around $K = 65$. How can we explain that such an obvious violation of basic arbitrage relationships is allowed to persist? The first possible explanation is that all the call prices in the data set are *last quotes*, hence do not reflect simultaneous option transactions. In particular, some of these options might be illiquid enough to prevent any arbitrage strategy (a butterfly spread in this case) to be formed. Finally, we haven't accounted for *transaction costs* which might also hamper arbitrage trading.

However, the key problem here is not that the option prices published are not viable. The danger is that this data is used by derivatives trading desks to calibrate more complex models and price other derivative products. If the call pricing data is not viable as is in (1.5), then calibration of any model on these prices is *impossible*. Worst, as algorithms used in practice to calibrate on this data often use a "best-fit" approach, this problem in the market data could go unnoticed and get incorporated into the model, with unpredictable consequences.

1.4.3 Basket options

In the case of simple call and put options, we have seen that the viability of prices is equivalent to a set of shape constraints on prices as functions of the option's strike. In particular, these conditions can be checked by inspection. For multivariate options, or *basket options*, the situation is entirely different.

A *basket call* payoff is given by

$$\left(\sum_{i=1}^{n} w_i S_i - K\right)_+$$

where $w_1, \ldots, w_n$ are the basket's weights and K is the option's strike price. These options appear in many markets, examples including: index options, spread options (options on the difference of two assets), swaptions (options on a swap rate)..., where basket option prices are used to gather information on the *correlation* among assets. We shall denote here as $C(w, K)$ the price of such an option. The question is now: Can we get tractable conditions to test basket price data for arbitrage opportunities? Arguments similar to the ones used above on simple calls an puts allow us to show the following result.

Proposition 5 *Given a set of market prices for basket calls $C(w_i, K_i) = p_i$, suppose there is no arbitrage, then the function $C(w, K)$ must satisfy:*

1. *$C(w, K)$ positive*
2. *$C(w, K)$ decreasing in K, increasing in w*
3. *$C(w, K)$ jointly convex in (w, K)*
4. *$C(w_i, K_i) = p_i$ and $C(0) = S, \quad i = 1, \ldots, m.$*

As in the unidimensional case, this is a shape constrained feasibility problem of the form (1.1), hence is tractable as a polynomial size linear program. However, in a key difference with dimension one, the conditions above are necessary *but not sufficient.* In particular Bertsimas & Popescu (2002) show that the exact problem (showing the absence of arbitrage between basket options) is NP-Hard.

The exact conditions for a function to represent the price of basket calls with weights w and strike K were derived by Henkin & Shananin (1990) in their investigation of production functions:

Proposition 6 *A function can be written*

$$C(w,K)=\int_{\mathbf{R}^n_+}(w^Tx-K)_+d\pi(x)$$

with $w\in\mathbf{R}^n_+$ *and* $K>0$*, if and only if:*

- $C(w,K)$ *is convex and homogenous of degree one*
- *for every* $w\in\mathbf{R}^n_{++}$*, we have* $\lim_{K\to\infty}C(w,K)=0$ *and* $\lim_{K\to 0^+}\frac{\partial C(w,K)}{\partial K}=-1$
- $F(w)=\displaystyle\int_0^\infty e^{-K}d\left(\frac{\partial C(w,K)}{\partial K}\right)$ *belongs to* $C_0^\infty(\mathbf{R}^n_+)$
- *For some* $\tilde{w}\in\mathbf{R}^n_+$, $(-1)^{k+1}D_{\xi_1}...D_{\xi_k}F(\lambda\tilde{w})\geq 0$, *for all positive integers* k *and* $\lambda\in\mathbf{R}_{++}$ *and all* $\xi_1,\ldots,\xi_k$ *in* $\mathbf{R}^n_+$.

If we want to keep the problem tractable as in (1.1), we can use only a subset of the conditions above. Given a set of market prices for basket calls $C(\omega_i,K_i)$, testing for the absence of arbitrage supposes that we solve the following shape constrained feasibility problem:

$$\begin{array}{ll}\text{find} & C(\omega,K)\\ \text{subject to} & C(\omega_i,K_i)=p_i,\quad i=1,\ldots,m\\ & C(\omega,K)\text{ convex in }(K,\omega)\\ & C(\omega,K)\text{ nonincreasing in }K\text{, nondecreasing in }\omega\\ & \partial C(\omega,K,0)/\partial K\geq -1\text{ and }\lim_{K\to\infty}C(\omega,K,0)=0\\ & C(\omega,K,T)\text{ homogeneous of degree 1 in }(K,\omega)\,,\end{array}\tag{1.13}$$

which is formed using a subset of the exact conditions in the result by Henkin & Shananin (1990). Using the conditions above and given prices p_i for basket calls $C(\omega_i,K_i)$, we can also compute upper and lower bounds on the price of another basket $C(\omega_0, K_0)$ by solving:

$$\begin{array}{ll} \text{max./min.} & C(\omega_0, K_0) \\ \text{subject to} & C(\omega_i, K_i) = p_i, \quad i = 1, \ldots, m \\ & C(\omega, K) \text{ convex in } (K, \omega) \\ & C(\omega, K) \text{ nonincreasing in } K, \text{ nondecreasing in } \omega \\ & \partial C(\omega, K, 0)/\partial K \geq -1 \text{ and } \lim_{K\to\infty} C(\omega, K, 0) = 0 \\ & C(\omega, K, T) \text{ homogeneous of degree 1 in } (K, \omega)\,. \end{array} \tag{1.14}$$

We notice that this is an infinite linear program of the form (1.1) and can be solved using the result in proposition 1.

Proposition 7 *If the following finite LP:*

$$\begin{array}{ll} \textit{maximize/minimize} & p_0 \\ \textit{subject to} & g_i^T((w_j, K_j) - (w_i, K_i)) \leq p_j - p_i, \quad i, j = 0, \ldots, m \\ & g_{i,j} \geq 0, -1 \leq g_{i,n+1} \leq 0, \quad i = 0, ..., m, \quad j = 1, \ldots, n \\ & g_i^T((w_i, K_i)) = p_i, \quad i = 0, \ldots, m, \end{array} \tag{1.15}$$

in the variables $p_0 \in \mathbf{R}_+$ and $g_i \in \mathbf{R}^{n+1}$ for $i = 0, ..., m$, is strictly feasible and its optimal value is finite (hence it is attained), the infinite program (1.14) and its discretization (1.15) have the same optimal value. Furthermore, an optimal point of (1.14) can be constructed from the optimal solution to (1.15).

Solving this linear program will produce *outer* price bounds, *i.e.* a lower bound on the lower bound and an upper bound on the upper bound. The numerical cost involved, in $O(mn + m^2)$, is minimal, however these conditions are of course not sufficient for the absence of arbitrage between basket call options.

1.4.4 Numerical example

Discrete model

Here, we try to quantify on simple examples the magnitude of the gap between the relaxed conditions in (1.13) and the exact ones in (6). To do this, we simulate a set of arbitrage free basket call prices using a simple model. Given these prices and the absence of arbitrage between basket calls, we study the price bounds induced on another basket call. We will discuss the exact interpretation of these bounds in much more detail in the next chapter, here we just want to get a rough idea of how good the convex relaxation in (1.13) is.

We can't compare these bounds with the exact solution since the exact problem is intractable in the general case, however we can try to compare these bounds with *inner* bounds obtained by maximizing and minimizing the price $C(\omega_0, K_0)$ over a set of probability measures satisfying the price constraints $C(\omega_i, K_i) = p_i$. If we consider only discrete measures, this becomes a (exponentially large) linear program.

We suppose here that the asset price at maturity T lies within the unit box $[0,1]^n$. We then discretize the probability density using a grid with N bins per asset. The problem of finding (inner) upper and lower bounds on a basket (ω_0, K_0) can then be written as:

$$\begin{array}{ll} \text{max./min.} & \mathbf{E}_\nu \left(\omega_0^T x - K_0\right)_+ \\ \text{subject to} & \mathbf{E}_\nu \left(\omega_i^T x - K_i\right)_+ = p_i, \quad i = 1, \ldots, m, \end{array} \tag{1.16}$$

which is a linear program of (exponential) size N^n in the (discrete) measure ν.

We test these two sets of upper and lower bounds in dimension two. Program (1.16) can be written:

$$\begin{array}{ll} \text{max./min.} & \sum_{k,l=0,\ldots,N} \nu_{kl} \left(\omega_0^T [k/N, l/N]^T - K_0\right)_+ \\ \text{subject to} & \sum_{k,l=0,\ldots,N} \nu_{kl} \left(\omega_i^T [k/N, l/N]^T - K_i\right)_+ = p_i, \quad i = 1, \ldots, m. \end{array}$$

which is a linear program in the variable $\nu \in \mathbf{R}^{N \times N}$. The assets are noted x_1, x_2 and we look for bounds on the price of an index option with payoff $(x_1 + x_2 - K)_+$. To

x	$(0,0)$	$(0,.8)$	$(.8,.3)$	$(.6,.6)$	$(.1,.4)$	$(1,1)$
probability	.2	.2	.2	.1	.1	.2

Table 1.1: Asset distribution.

produce price data, we use a simple discrete model for the assets, their distribution has finite support and is given by table (1.1). The input data set is composed of the forward prices together with the following call prices:

$$\begin{aligned}&(.2x_1 + x_2 - .1)_+,\ (.5x_1 + .8x_2 - .8)_+,\ (.5x_1 + .3x_2 - .4)_+,\\&(x_1 + .3x_2 - .5)_+,\ (x_1 + .5x_2 - .5)_+,\ (x_1 + .4x_2 - 1)_+,\ (x_1 + .6x_2 - 1.2)_+.\end{aligned}$$

We plot the inner and outer bounds obtained using this data in figure (1.9). We observe that sometimes the bounds match, *i.e.* the price bounds given by the relaxation are tight, while sometime there is a gap and not much can be said about the relaxation's suboptimality. In the next chapter, we will detail some cases where the exact problem is tractable and test the relaxation's performance in those cases.

We also examine how these bounds evolve as more and more instruments are incorporated into the data set. For a particular choice of strike price (here $K = 1$), we compute the outer bounds (1.15) and inner bounds (1.16) obtained when using only the k first instruments in the data set, for $k = 2, \ldots, 7$. The result is plotted in figure (1.10). Finally, we plot in figure (1.11) the distribution ν obtained while solving for the last lower inner bound in the previous example.

Black-Scholes model

Here again, we simulate a set of arbitrage free basket call prices using this time a multivariate Black & Scholes (1973) model. Given these prices and the absence of arbitrage between basket calls, we study the price bounds induced on another basket call.

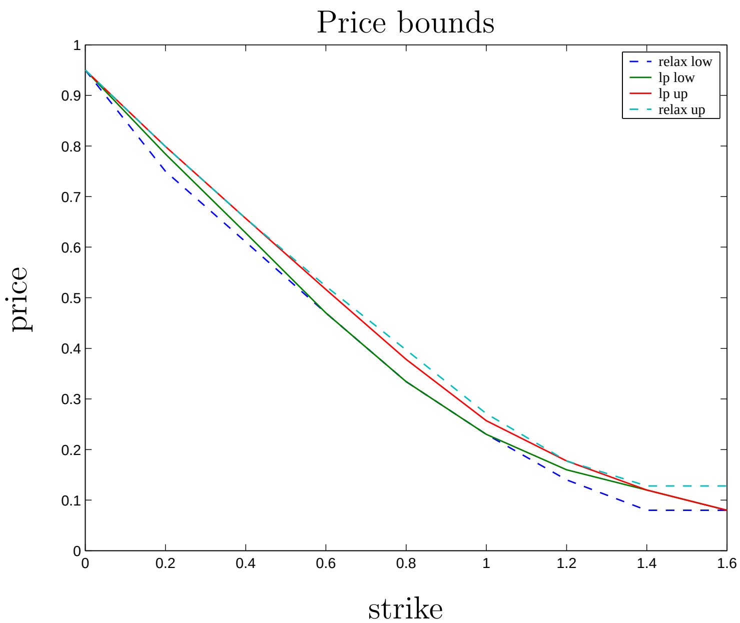

Figure 1.9: Comparison of the inner bounds computed by discretization (solid lines, computed using (1.16)) and the outer bounds obtained by relaxation (dashed lines, computed using (1.14)).

There is no closed form formula to compute the price of a basket call option in a multivariate model, and to get inner bounds on the basket price, we use the approximation technique in d'Aspremont (2003). In this market, the dynamics of the assets F_t^i are given by:

$$dF_s^i = F_s^i \sigma^i dW_s,$$

where W_s is a d-dimensional Brownian motion and $\sigma = (\sigma^i)_{i=1,\dots,n} \in \mathbf{R}^{n \times d}$ is the volatility matrix. We shall denote as $\Gamma \in \mathbf{S}^n$ the corresponding covariance matrix, with $\Gamma_{ij} = \sigma^{iT}\sigma^j$. The sum of lognormally distributed assets is not lognormal, but we can approximate it by a lognormal distribution to price a basket call. We use the following result from d'Aspremont (2003):

Proposition 8 *The price of a basket call with payoff:*

$$\left(\sum_{i=1}^{n} w_i F_T^i - K\right)_+$$

at time T, can be approximated by a Black & Scholes (1973) call price using an appropriate variance V_T such that:

$$C = BS(T, w^T F_t, V_T) = (w^T F_t)\mathcal{N}(h(V_T)) - \kappa \mathcal{N}\left(h(V_T) - \sqrt{V_T}\right), \tag{1.17}$$

where $\mathcal{N}(x)$ is the CDF of the normal distribution,

$$h\left(V_T\right) = \frac{\left(\ln\left(\frac{w^T F_t}{\kappa}\right) + \frac{1}{2}V_T\right)}{\sqrt{V_T}},$$

and with

$$V_T = \mathbf{Tr}\left(\Omega_t \Gamma\right) T,$$

where

$$\Omega_t = \hat{w}_t \hat{w}_t^T \quad \textit{and} \quad \hat{w}_{i,t} = \frac{w_i F_t^i}{w^T F_t}.$$

Since the Black & Scholes (1973) formula is increasing with its variance term V_T, computing bounds on the model price of a basket call given market price data on other baskets is equivalent to solving the following semidefinite program:

$$\begin{array}{ll} \text{max./min.} & \mathbf{Tr}(\Omega_{0,t}X) \\ \text{subject to} & \mathbf{Tr}(\Omega_{i,t}X) = V_{T,i}, \quad i = 1, \ldots, m \\ & X \succeq 0, \end{array} \tag{1.18}$$

with $V_{T,i}$ such that:

$$BS(T, w_i^T F_t, V_{i,T}) = p_i, \quad i = 1, \ldots, m,$$

where $p_i \in \mathbf{R}^n$ are the market prices of basket call options with weights w_i.

We look for bounds on the price of an index option: $(.2 \sum_{i=1}^{5} x_i - K)_+$, given the price of at-the-money options with the following weights:

$$\begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \\ 0.33 & 0.33 & 0.33 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.33 & 0.33 & 0.33 \\ 0.40 & 0.20 & 0.20 & 0.20 & 0.00 \end{bmatrix}$$

The assets initial values F_t^i are $(0.03, 0.03, .05, .07, .07)$ and the model covariance matrix Γ is given by:

$$\begin{pmatrix} 0.06 & 0.04 & 0.04 & 0.04 & 0.04 \\ 0.04 & 0.06 & 0.04 & 0.04 & 0.04 \\ 0.04 & 0.04 & 0.06 & 0.04 & 0.04 \\ 0.04 & 0.04 & 0.04 & 0.06 & 0.04 \\ 0.04 & 0.04 & 0.04 & 0.04 & 0.06 \end{pmatrix}$$

We plot the inner and outer bounds obtained using this data in figure (1.12).

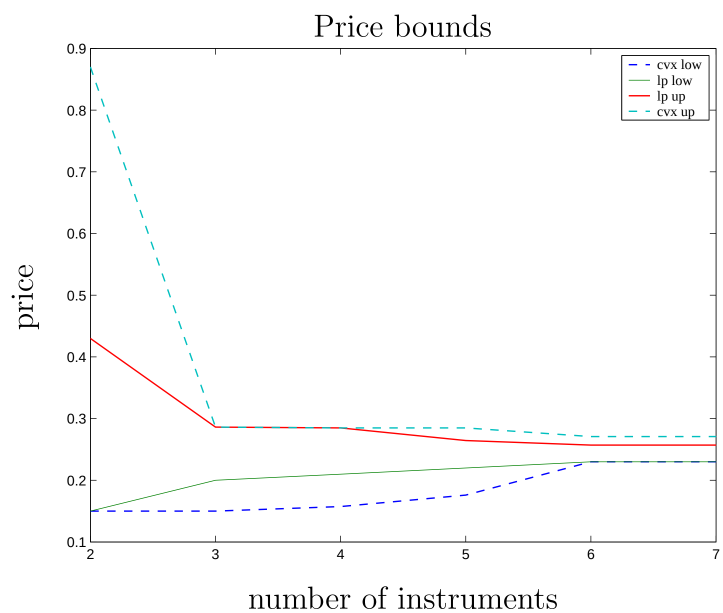

Figure 1.10: Inner bounds (solid lines, computed using (1.16)) and outer bounds (dotted lines, computed using (1.14)) versus number of instruments in the data set.

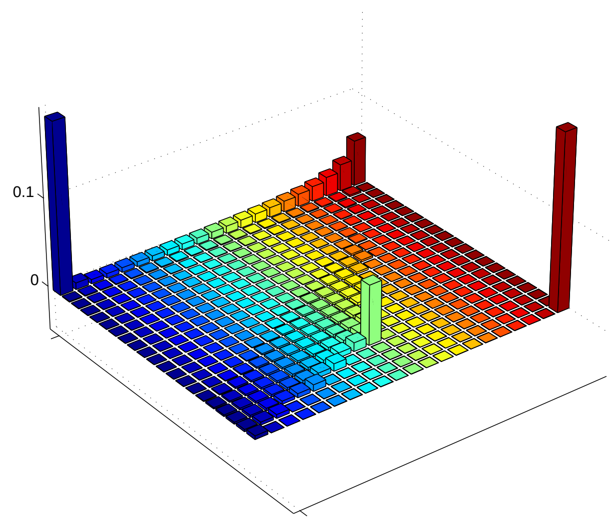

Figure 1.11: An example of discrete distribution minimizing the price of the basket $(x_1 + x_2 - K)_+$.

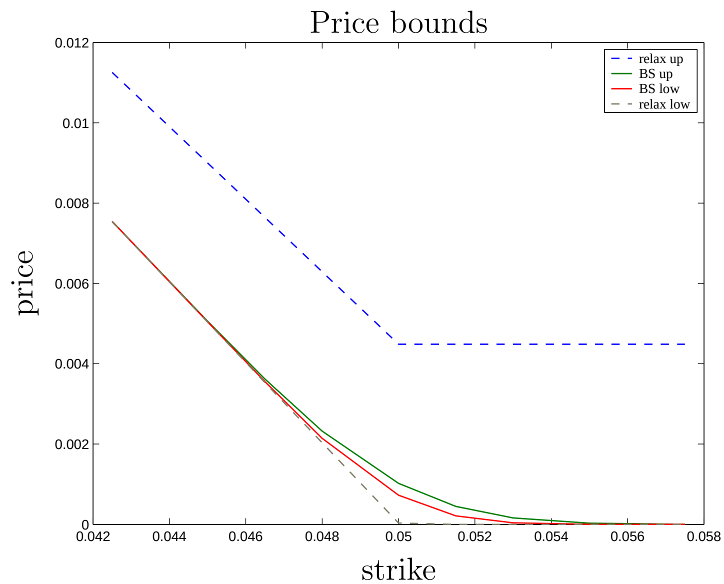

Figure 1.12: Comparison of the inner bounds computed using the Black & Scholes model (solid lines, computed using (1.18)) and the outer bounds obtained by relaxation (dashed lines, computed using (1.14)).

Chapter 2

Tightness results

In the previous chapter, we obtained bounds on the price of basket options using a linear programming relaxation. In this chapter, we consider again the problem of computing upper and lower bounds on the price of a European basket call option, given prices on other similar baskets. Although we have seen that this problem is very hard to solve exactly in the general case, we show that in some instances the *exact* upper and lower bounds can be computed via simple closed-form expressions, or linear programs. We also show that the relaxation (1.15) discussed in the previous chapter is tight in some of the special cases examined.

Notation

Hereafter, x_+ will denote the positive part of x, which is the vector with components $\max(x_i, 0)$. e is the n-vector with all components equal to one, and e_i is the i-th unit vector of $\mathbf{R}^n$. The set $\mathbf{R}^n_+$ denotes the set of n-vectors with non-negative components, and $\mathbf{R}^n_{++}$ its interior. The cone of nonnegative measures with support included in $\mathbf{R}^n_+$ is denoted by $\mathcal{K}$. For $w \in \mathbf{R}^s$, $K \in \mathbf{R}$ and $g \in \mathbf{R}^{s+1}$, the notation $\langle g, (w, K) \rangle$ denotes the scalar product $\tilde{g}^T w + g_{m+1} K$, where $\tilde{g}$ contains the first s elements of g.

2.1 Introduction

2.1.1 Problem setup

Let $p \in \mathbf{R}^m_+$, $K \in \mathbf{R}^m_+$, $w \in \mathbf{R}^n$, $w_i \in \mathbf{R}^n$, $i = 1, \ldots, m$ and $K_0 \geq 0$. We consider the problem of computing upper and lower bounds on the price of an European basket call option with strike K_0 and weight vector w_0:

$$\mathbf{E}_\pi (w_0^T x - K_0)_+, \tag{2.1}$$

with respect to all probability distributions $\pi \in \mathcal{K}$ on the asset price vector x, consistent with a given set of observed prices p_i of options on other baskets, that is, given

$$\mathbf{E}_\pi (w_i^T x - K_i)_+ = p_i, \quad i = 1, \ldots, m. \tag{2.2}$$

Note that we implicitly assume that all the options have the same maturity, and that, without loss of generality, the risk-free interest rate is zero (we compare prices in the forward market).

As in the previous chapter, we seek non-parametric bounds, that is, we do not assume any specific model for the underlying asset prices. Our sole assumption is the absence of a static arbitrage today (*i.e.* the absence of an arbitrage that only requires trading today and at maturity, see §1.4 for a complete definition and simple examples. The primary objective of these bounds is not to detect and exploit arbitrage opportunities in the basket vs. vanilla market near the money, the amplitude of the Bid-Ask spreads being likely to make those opportunities very rare. However, the data on basket prices (index options in equity markets or swaptions in fixed income) is usually *very sparse* and traders often rely on intuitive guesses to extrapolate the remaining points. Our results provide a simple method to check the validity of these extrapolated prices where they are the most likely to create static arbitrage opportunities, *i.e.* very far in or out of the money (cf. figure (1.5)).

From a financial point of view, our approach can be seen as a one-period, non-parametric computation of the upper and lower hedging prices defined in El Karoui & Quenez (1991), El Karoui & Quenez (1995) or Karatzas & Shreve (1998)). The necessary conditions we detail in §2.3 in a multidimensional setup have been extensively used in the unidimensional case to infer information on the state-price density given option prices (see Breeden & Litzenberger (1978), Buchen & Kelly (1996) or Laurent & Leisen (2000) among others).

From an optimization point of view, problems such as the one above have received a significant amount of attention in various forms. First, we can think of (2.2) as a *linear semi-infinite program*, *i.e.* a linear program with a finite number of linear constraints on an infinite dimensional variable. We use this interpretation and the related duality results to compute closed-form solutions, for a particular subclass of problems. Secondly, we can see (2.2) as *generalized moment constraints*. This approach was successfully used in dimension one by Bertsimas & Popescu (2002) and we will come back to this in the last chapter. In higher dimensions however, the proposed relaxation algorithm in Bertsimas & Popescu (2002) requires the solution of a number of linear programs that is potentially exponential in n, the number of underlying assets. This makes the method prohibitive for large-scale problems. Finally, as Henkin & Shananin (1990), one can think of (2.2) as an *integral transform inversion problem*. This is the approach we adopt to design an efficient relaxation in the general case, based on shape constraints on the call price as a function of the weight vector w and strike price K.

We examine in detail a special case of the problem, in which prices on options of individual assets, as well as forward prices, are given, and the option to be priced involves a non-negative weight vector w. Our contribution there is to provide a solution that is polynomial-time in the number of assets, involving a linear program with $O(n)$ variables and constraints, where n is the number of assets. We prove that our bounds are exact in most of the cases, that is, they are attained (possibly in the limit) by some distribution π consistent with observed option prices. We obtain expressions for these optimal measures, and use them to prove tightness of the linear programming relaxations (1.15) in some special cases.

This chapter is organized as follows. We first detail a dual formulation of the general problem in §2.1.2. We obtain in §2.2 upper and lower bounds, in the special cases referred to above, and then for the general problem through a relaxation of the type (1.15). We discuss the tightness of the bounds in §2.4 and summarize these results in §2.5. Finally, §2.6 provides some numerical examples.

2.1.2 Dual program: a portfolio problem

In the general case, we can write the upper bound problem as a semi-infinite program:

$$p^{\text{sup}} := \sup_{\pi \in \mathcal{K}} \int_{\mathbf{R}^n_+} \psi(x)\pi(x)dx \text{ subject to} \int_{\mathbf{R}^n_+} \phi(x)\pi(x)dx = p, \quad \int_{\mathbf{R}^n_+} \pi(x)dx = 1, \quad (2.3)$$

where

$$\psi(x) := (w^T x - K_0)_+, \quad \phi_i(x) := (w_i^T x - K_i)_+, \quad i = 1, \ldots, m.$$

We define the Lagrangian (on $\mathcal{K} \times \mathbf{R}^{m+1}$):

$$L(\pi, \lambda, \lambda_0) = \int_{\mathbf{R}^n_+} \psi(x)\pi(x)dx + \lambda^T \left(p - \int_{\mathbf{R}^n_+} \phi(x)\pi(x)dx \right) + \lambda_0 (1 - \int_{\mathbf{R}^n_+} \pi(x)dx),$$

and, as in Hettich & Kortanek (1993), we can explicit the dual of (2.3):

$$\begin{array}{rcl} d^{\text{sup}} & := & \inf\limits_{\lambda_0, \lambda} : \ \lambda^T p + \lambda_0 \ : \ \lambda^T \phi(x) + \lambda_0 \geq \psi(x) \text{ for every } x \in \mathbf{R}^n_+ \\ & = & \inf\limits_{\lambda} : \ \sup\limits_{x \geq 0} : \lambda^T p + \psi(x) - \lambda^T \phi(x). \end{array} \quad (2.4)$$

Both primal and dual problems have very intuitive financial interpretations. The primal problem looks for a state price density (see for example Duffie (1996)) that maximizes the target option while satisfying the pricing constraints imposed by the current market conditions. The dual problem looks for the least expensive portfolio of options plus cash, $\lambda^T\phi(x)+\lambda_0$, that dominates the option payoff $\psi(x)$. Of course, the dual problem above yields an upper bound on the upper bound.

Similarly, the computation of the lower bound involves

$$p^{\inf} := \inf_{\pi\in\mathcal{K}} \int_{\mathbf{R}_+^n} \psi(x)\pi(x)dx \text{ subject to } \int_{\mathbf{R}_+^n} \phi(x)\pi(x)dx = p, \ \int_{\mathbf{R}_+^n} \pi(x)dx = 1, \quad (2.5)$$

whose dual is

$$\begin{aligned} d^{\inf} &:= \sup_{\lambda_0,\lambda} : \lambda^T p + \lambda_0 \ : \ \lambda^T\phi(x)+\lambda_0 \le \psi(x) \text{ for every } x \in \mathbf{R}_+^n \\ &= \sup_{\lambda} : \inf_{x\ge 0} : \lambda^T p + \psi(x) - \lambda^T\phi(x). \end{aligned} \quad (2.6)$$

Here, the dual problem provides a lower bound on the lower bound.

General results on semi-infinite linear programs establish the equivalence between the primal and dual formulations. We cite here a sufficient constraint qualification condition for perfect duality from Hettich & Kortanek (1993), which makes an assumption about the support of optimal distributions. (We focus now on the lower bound; a similar result holds for the upper bound problem.)

Proposition 9 *Assume that for problem (2.6), without loss of generality, the support of the asset price distribution can be restricted to a given compact set $B \subset \mathbf{R}_+^n$. Assume further that there exist a pair $(\lambda_0,\lambda)\in\mathbf{R}^{n+1}$ such that:*

$$\lambda^T\phi(x)+\lambda_0 < \psi(x) \textit{ for all } x \in B.$$

Then if $d^{\inf}$ is finite, perfect duality holds, namely $d^{\inf}=p^{\inf}$.

This constraint qualification condition trivially holds when $\phi(x)$ and $\psi(x)$ are call option payoffs hence we have $d^{\inf}=p^{\inf}$, provided that the support of distributions feasible for our problem can be restricted to some compact $B\subset\mathbf{R}_+^n$. However, this may not be the case for the bounds detailed below and we will prove perfect duality directly whenever possible.

2.2 Upper and lower bounds

In this section, we address the problem of computing the bounds. We first consider the case when the observed prices correspond to options on each individual assets. In practice, these observations always include the forward contract prices $\mathbf{E}_\pi x_i = q_i$, $i = 1, \ldots, n$, which are quoted by the market, and we seek to exploit the forward price information. Then we specialize in §2.2.2 these results to the case when the forward prices are ignored; we examine this case because it is useful in the proofs of perfect duality in §2.4. Finally we address the general case in §2.3.

2.2.1 Option and forward price constraints

We examine the problem of computing upper and lower bounds on

$$\mathbf{E}_\pi (w^T x - K_0)_+,$$

given the $2n$ constraints

$$\mathbf{E}_\pi (x_i - K_i)_+ = p_i, \quad \mathbf{E}_\pi x_i = q_i, \quad i = 1, \ldots, n, \tag{2.7}$$

where $K_0 > 0$ and w, K, p, q are given vectors of $\mathbf{R}^n_{++}$.

We will assume that $0 \leq p \leq q \leq p + K$, which is a necessary and sufficient condition for the problem above to be feasible. Sufficiency is obtained with the discrete distribution defined by

$$x = \begin{cases} 2p + K & \text{with probability } 1/2, \\ 2(q-p) - K & \text{with probability } 1/2. \end{cases} \tag{2.8}$$

From the form of the constraints, we also observe that the constraints $0 \leq p \leq q \leq p + K$ are necessary.

Upper bound

In this section, we apply the duality formalism to the upper bound problem with constraints described in (2.7).

In view of the general result (2.4), the dual problem can be expressed as

$$d^{\sup} \;=\; \inf_{\lambda+\mu\geq w}\ \sup_{x\geq 0}\ \lambda^T p + \mu^T q + (w^T x - K_0)_+ - \lambda^T (x-K)_+ - \mu^T x, \tag{2.9}$$

where, without loss of generality, we have included the constraint $\lambda+\mu \geq w$, in order to ensure that the inner supremum is finite. We introduce a partition of $\mathbf{R}^n_+$ as follows. To a given subset I of $\{1,\ldots,n\}$, we associate a subset D_I of $\mathbf{R}^n_+$, defined by

$$D_I = \{x \ : \ x_i > K_i, \quad i\in I, \quad 0\leq x_i \leq K_i, \quad i \in J\},$$

where J denotes the complement of I in $\{1,\ldots,n\}$. For $z\in\mathbf{R}^n$, let z_I be the vector formed with the elements $(z_i)_{i\in I}$, in the ascending order of indices in I.

We have

$$\begin{aligned} d^{\sup} &= \inf_{\lambda+\mu\geq w}\ \max_{t\in\{0,1\}}\ \max_{I\subseteq\{1,\ldots,n\}}\ \sup_{x\in D_I}\ \lambda^T p + \mu^T q + t(w^T x - K_0) \\ &\quad -\lambda_I^T (x_I - K_I) - \mu^T x \\ &= \inf_{\lambda+\mu\geq w}\ \max_{t\in\{0,1\}}\ \max_{I\subseteq\{1,\ldots,n\}}\ \lambda^T p + \mu^T q + h(\lambda,\mu,I,t), \end{aligned}$$

where $h(\lambda,\mu,I,t)$ is given by

$$\begin{aligned} & h(\lambda,\mu,I,t) \\ := \;& \sup_{x\in D_I}\ t(w^T x - K_0) - \lambda_I^T (x_I - K_I) - \mu^T x \\ = \;& \sup_{0\leq x_J\leq K_J}\ (tw_J - \mu_J)^T x_J - tK_0 + \lambda_I^T K_I \\ & + \sup_{x_I > K_I}\ (tw_I - \mu_I - \lambda_I)^T x_I \\ = \;& \begin{cases} (tw_J-\mu_J)_+^T K_J - tK_0 + (tw_I - \mu_I)^T K_I & \text{if } \lambda_I + \mu_I \geq tw_I, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

We note that finiteness of $h(\lambda, \mu, I, t)$ is guaranteed by $\lambda + \mu \geq w$ and $t \geq 0$. When these conditions hold, the maximum value of $h(\lambda, \mu, I, t)$ over $I \subseteq \{1, \ldots, n\}$ is obtained when the complement J is the full set, that is, when I is empty. We obtain

$$\max_{I \subseteq \{1,\ldots,n\}} h(\lambda, \mu, I, t) \quad = \quad (tw - \mu)_+^T K - tK_0.$$

Optimizing over t, we obtain

$$\max_{t \in \{0,1\}} \max_{I \subseteq \{1,\ldots,n\}} h(\lambda, \mu, I, t) = \max \left((-\mu)_+^T K, (w - \mu)_+^T K - K_0 \right).$$

This results in the following expression for $d^{\sup}$:

$$\begin{aligned} d^{\sup} \quad &= \quad \inf_{\lambda + \mu \geq w} \lambda^T p + \mu^T q + \max \left((-\mu)_+^T K, (w - \mu)_+^T K - K_0 \right) \\ &= \quad \inf_{\mu} w^T p + \mu^T (q - p) + \max \left((-\mu)_+^T K, (w - \mu)_+^T K - K_0 \right), \end{aligned} \tag{2.10}$$

which admits the following linear programming representation:

$$\begin{aligned} d^{\sup} = \inf_{\mu, t, v, z} : w^T p + \mu^T (q - p) + t \qquad & t \geq v^T K, \quad v \geq 0, \quad v + \mu \geq 0 \\ & t \geq z^T K - K_0, \quad z \geq 0, \quad z + \mu \geq w. \end{aligned}$$

The problem is feasible, and is thus equivalent to its dual. After some elimination of dual variables, the dual writes

$$\begin{aligned} d^{\sup} = \max_{y, \beta} w^T p + w^T y - \beta K_0 \ : \quad & (1 - \beta) K \geq q - p - y \geq 0 \\ & \beta K \geq y \geq 0. \end{aligned}$$

Note that the above problem is feasible if and only if $p \leq q \leq p + K$. We thus recover the primal feasibility condition mentioned before. This condition ensures that the dual bound $d^{\sup}$ is finite. The above further reduces to the one-dimensional problem:

$$d^{\sup} = \max_{0 \le \beta \le 1} : w^T p + \sum_i w_i \min(q_i - p_i, \beta K_i) - \beta K_0. \tag{2.11}$$

The above problem is the maximization of a piecewise linear concave function of one variable, thus the maximum is attained at one of the break points $\beta_j := (q_j - p_j)/K_j \in [0 : 1]$, $j = 1, \ldots, n$, or for $\beta = 0, 1$. This way, we can obtain a closed-form expression for the upper bound, namely

$$d^{\sup} = \max_{0 \le j \le n+1} w^T p + \sum_i w_i \min(q_i - p_i, \beta_j K_i) - \beta_j K_0,$$

with the convention $\beta_0 = 0$, $\beta_{n+1} = 1$.

We can check that the above bound satisfies some basic properties: it is convex in w and concave in p, q. Also, when $w = e_i$ (the i-th unit vector), and $K_0 = K_i$, we obtain $d^{\sup} = p_i$, while for $K_i = 0$, we obtain $d^{\sup} = q_i$.

Lower bound

In the lower bound case, the dual problem is

$$d^{\inf} = \sup_{\lambda + \mu \le w} \inf_{x \ge 0} \lambda^T p + \mu^T q + (w^T x - K_0)_+ - \lambda^T (x - K)_+ - \mu^T x,$$

where we exploited the fact that the inner infimum is $-\infty$ unless $\lambda + \mu \le w$.

Let us use the same notation as before. We have

$$\begin{aligned} d^{\inf} &= \sup_{\lambda + \mu \le w} \min_{I \subseteq \{1, \ldots, n\}} \inf_{x \in D_I} \lambda^T p + \mu^T q + (w^T x - K_0)_+ - \lambda_I^T (x_I - K_I) - \mu^T x \\ &= \sup_{\lambda + \mu \le w} \min_{I \subseteq \{1, \ldots, n\}} \lambda^T p + \mu^T q + h(\lambda, \mu, I), \end{aligned}$$

where

$$h(\lambda,\mu,I) \quad = \quad \inf_{x,y_0} y_0 - \lambda_I^T(x_I - K_I) - \mu^T x \ : \ x \in D_I,\ y_0 \geq w^T x - K_0,\ y_0 \geq 0.$$

We have by linear programming duality

$$h(\lambda,\mu,I) = \sup\,(\alpha w - \mu)^T K - \alpha K_0 - (\alpha w_J - \mu_J)_+^T K_J \ : \ \begin{array}{l} \alpha w_I - \lambda_I - \mu_I \geq 0 \\ 0 \leq \alpha \leq 1 \end{array}$$

Thus

$$d^{\inf} \quad = \quad \sup_{\lambda+\mu\leq w} \lambda^T p + \mu^T(q-K) + \min_{I\subseteq\{1,\ldots,n\}} f(\lambda,\mu,I),$$

where

$$f(\lambda,\mu,I) := \sup_{\underline{\alpha}(\lambda,\mu,I)\leq\alpha\leq 1} \alpha(w^T K - K_0) - (\alpha w_J - \mu_J)_+^T K_J,$$

and

$$\underline{\alpha}(\lambda,\mu,I) := \max_{i\in I} \frac{(\lambda_i+\mu_i)_+}{w_i},$$

with the convention that $\underline{\alpha}(\lambda,\mu,I) = 0$ when I is empty.

Let I be a non-empty subset of $\{1,\ldots,n\}$. Let $i \in \arg\max_{i\in I}(\lambda_i+\mu_i)_+/w_i$. We observe that

$$\underline{\alpha}(\lambda,\mu,I) = \underline{\alpha}(\lambda,\mu,\{i\}),$$

and

$$f(\lambda,\mu,I) \geq f(\lambda,\mu,\{i\}),$$

which dramatically reduces the complexity of the minimization subproblem: instead of computing the minimum over all 2^n sets $I \subseteq \{1,\ldots,n\}$ it is sufficient to pick I in the set of *singletons* of $\{1,\ldots,n\}$, or $I = \emptyset$. Therefore, the problem reads as a linear program

$$
\begin{array}{lll}
d^{\inf} = & \sup_{\lambda,\mu,\alpha_0,\ldots,\alpha_n} & \lambda^T p + \mu^T (q - K) + h \\
\text{subject to} & \lambda + \mu \leq w & \\
& h \leq \alpha_0 (w^T K - K_0) - (\alpha_0 w - \mu)_+^T K & \\
& 0 \leq \alpha_0 \leq 1 & \\
& \forall i \ : \ h \leq \alpha_i (w^T K - K_0) - \sum_{j \neq i} (\alpha_i w_j - \mu_j)_+ K_j & \\
& (\lambda_i + \mu_i)_+ / w_i \leq \alpha_i \leq 1, &
\end{array}
\tag{2.12}
$$

and can be solved efficiently, since it has $O(n)$ constraints and variables.

2.2.2 Ignoring forward price constraints

In this section, we examine the problem in the case when the forward price constraints $\mathbf{E}_\pi x = q$ are ignored. The simple bounds we obtain in this setting will prove useful for obtaining perfect duality results later.

Upper bound

The new upper bound is readily obtained by setting the variable μ, which is the variable dual to the constraint $\mathbf{E}_\pi x = q$, to zero in the expression (2.10). We get the simple closed-form expression

$$
d^{\sup} = w^T p + (w^T K - K_0)_+, \tag{2.13}
$$

which can be obtained as a direct consequence of Jensen's inequality applied to the function $x \to x_+$.

Lower bound

A closed-form expression. For the lower bound, we again set the dual variable μ to zero in the expression (2.12). We obtain

$$d^{\inf} \quad = \quad \sup_{0 \le \xi \le e} \; p(w)^T \xi + h \; : \; h \le 0, \; h \le \xi_i (w_i K_i - K_0), \; 1 \le i \le n. \tag{2.14}$$

We note that $d^{\inf}$ can be expressed as the solution of a non-linear, convex optimization problem:

$$d^{\inf} = \sup_{\xi} \; p(w)^T \xi - \max_{1 \le i \le n} \; \xi_i (K_0 - w_i K_i)_+ \; : \; 0 \le \xi \le e, \tag{2.15}$$

or its dual:

$$d^{\inf} = \inf_{\nu} \; \sum_{i=1}^{n} \left(p_i w_i - \nu_i (K_0 - w_i K_i)_+ \right)_+ \; : \; \nu^T e = 1, \; \nu \ge 0. \tag{2.16}$$

We can reduce the optimization problem to a line search over a scalar parameter, by elimination of the variable ξ. We obtain

$$d^{\inf} = \sum_{i \; : \; K_i w(i) \ge K_0} p_i w(i) + \sup_{v \ge 0} \sum_{i \; : \; K_i w(i) < K_0} p_i w(i) \min(1, \frac{v}{K_0 - K_i w(i)}) - v.$$

The minimization above can be further reduced to a closed-form expression by noting that the piecewise-linear function (of v) involved has break points at $\gamma_i = K_0 - K_i w_i$ (for i such that $\gamma_i > 0$) and 0. Thus:

$$\begin{aligned} d^{\inf} = \; & \textstyle\sum_{i \; : \; K_i w_i \ge K_0} p_i w_i \\ & + \max_{j \; : \; K_j w_j < K_0} \left(\sum_{i \; : \; K_i w_i < K_0} p_i w_i \min(1, \frac{K_0 - K_j w_j}{K_0 - K_i w_i}) - K_0 + w_j K_j \right)_+ . \end{aligned} \tag{2.17}$$

Interpretation in terms of portfolios. Although the development above has the definite advantage of being completely constructive, we can get a more direct and perhaps more intuitive proof of (2.17) by interpreting its equivalent form (2.15) in terms of portfolio inequalities. Without loss of generality, we can assume that $w_0 = e$, where e is the n-vector of ones. To show

$$d^{\inf} = \sup_{0 \le \xi \le e} p^T \xi - \max_{1 \le i \le n} \xi_i (K_0 - K_i)_+,$$

it suffices to show that

$$\xi^T (x - K)_+ - \max_{1 \le i \le n} \xi_i (K_0 - K_i)_+ \le \left(e^T x - K_0\right)_+ \text{ for all } x \in \mathbf{R}^n_+ \tag{2.18}$$

holds for every ξ such that $0 \le \xi \le e$. The above can be interpreted as a portfolio inequality: the price of the options portfolio $\xi^T (x - K)_+$, together with a certain amount of cash (negative values meaning borrowing), is dominated by the payoff.

Let us prove the portfolio inequality above. Let ξ be such that $0 \le \xi \le e$. Condition (2.18) trivially holds when $x = 0$. Let us now consider $x \in \mathbf{R}^n_+$, $x \ne 0$. Then, $e^T x > 0$. First, assume $(0 <) e^T x \le K_0$, then:

$$\xi^T (x - K)_+ \le \xi^T \left(\frac{x}{e^T x} K_0 - K \right)_+ ,$$

and by convexity of the function $x \to x_+$, we have:

$$\begin{aligned} \xi^T \left(\frac{x}{e^T x} K_0 - K \right)_+ &\le \sum_{i=1}^n \xi_i \frac{x_i}{e^T x} (K_0 - K_i)_+ \\ &\le \max_{i=1,\ldots,n} \xi_i (K_0 - K_i)_+ . \end{aligned} \tag{2.19}$$

Assume now that $e^T x \ge K_0$, and let $i_0 = \arg\max_{i=1,\ldots,n} \xi_i (K_0 - K_i)_+$, we can write

$$\sum_{i=1, i\neq i_0}^{n} \xi_i x_i + \xi_{i_0} \left(x_{i_0} - K_i\right)_+ - \max_{i=1,\ldots,n} \xi_i \left(K_0 - K_i\right)_+ \leq \left(e^T x - K_0\right)_+$$

as

$$\sum_{i=1, i\neq i_0}^{n} \xi_i x_i + \xi_{i_0} \left(x_{i_0} - K_i\right)_+ - \max_{i=1,\ldots,n} \xi_i \left(K_0 - K_i\right)_+ \leq e^T x - K_0,$$

which holds since $0 \leq \xi \leq e$ and

$$\xi_{i_0} \left(\left(x_{i_0} - K_{i_0}\right)_+ - \left(K_0 - K_{i_0}\right)_+\right) \leq x_{i_0} - K_0.$$

The above, together with (2.19), proves the inequality (2.18). This shows (2.15) directly.

2.3 Relaxation for the general case

2.3.1 An integral transform

Let us come back to the original problem, for $p \in \mathbf{R}_+^m$, $K \in \mathbf{R}_+^m$, $w_0 \in \mathbf{R}^n$, $w_i \in \mathbf{R}^n$, $i = 1, ..., m$ and $K_0 \geq 0$. We consider the problem of computing upper and lower bounds on the price of an European call basket option with strike K_0 and weight vector w_0:

$$\mathbf{E}_\pi (w_0^T x - K_0)_+,$$

with respect to all probability distributions $\pi \in \mathcal{K}$ on the asset price vector x, consistent with a given set of m observed prices p_i of options on other baskets and forward prices q_i, that is, given

$$\mathbf{E}_\pi (w_i^T x - K_i)_+ = p_i, \quad i = 1, \ldots, m \text{ and } \mathbf{E}_\pi x_j = q_j, \quad j = 1, \ldots, n.$$

If we write, for some $\pi \in \mathcal{K}$:

$$\begin{aligned} C(w, K) &= \mathbf{E}_\pi (w^T x - K)_+ \\ &= \int_{\mathbf{R}_+^n} (w^T x - K)_+ d\pi(x), \end{aligned}$$

we can think of $C_\pi(w, K)$ as a particular integral transform of the measure π. We can compute the inverse of this integral transform. If we assume that the measure π is absolutely continuous with respect to the Lebesgue measure with density $\pi(x)$, then for almost all K we have:

$$\hat{f}(w, K) := \frac{\partial^2 C(w, K)}{\partial K^2} = \int_{\mathbf{R}_+^n} \delta(w^T x - K) \pi(x) dx,$$

where $\delta(x)$ is the Dirac Delta function. This means that $\hat{f}(w, K)$ is the Radon transform (see Helgason (1999) or Ramm & Katsevich (1996)) of the measure π.

The general pricing problem above can then be rewritten as the following infinite dimensional problem:

$$\begin{array}{ll} \text{minimize/maximize} & f(w_0, K_0) \\ \text{subject to} & f(w_i, K_i) = p_i, \quad i = 1, \ldots, m \\ & f(w, K) \in \mathcal{R}_C, \end{array}$$

where $\mathcal{R}_C$ is the range of the (linear) integral transform

$$\begin{array}{ll} C: & \mathcal{K} \to \mathcal{R}_C \\ & \pi \to C(w, K) = \int_{\mathbf{R}^n_+} (w^T x - K)_+ d\pi(x). \end{array}$$

Thus, the problem of finding all possible arbitrage-free option prices becomes equivalent to that of characterizing the range of the Radon transform on the set of nonnegative measures $\mathcal{K}$. This has been done by Henkin & Shananin (1990) in the context of production functions (which can be thought of as put options). Using call-put parity, we can directly derive from Henkin & Shananin (1990, Theorem 3.2) the following result:

Proposition 10 *A function $C(w, K)$, with $w \in \mathbf{R}^n_+$ and $K > 0$ belongs to $\mathcal{R}_C$, i.e. it can be represented in the form*

$$C(w, K) = \int_{\mathbf{R}^n_+} (w^T x - K)_+ d\pi(x),$$

where π is a nonnegative measure on a compact of $\mathbf{R}^n_+$, if and only if the following conditions hold.

- *$C(w, K)$ is convex and homogenous of degree one;*
- *for every $w \in \mathbf{R}^n_{++}$, we have*

$$\lim_{K \to \infty} C(w, K) = 0 \text{ and } \lim_{K \to 0^+} \frac{\partial C(w, K)}{\partial K} = -1;$$

- *the function*

$$F(w)=\int_0^{\infty} e^{-K} d\left(\frac{\partial C(w,K)}{\partial K}\right)$$

belongs to $C_0^{\infty}(\mathbf{R}_+^n)$ *and for some* $\tilde{w}\in\mathbf{R}_+^n$ *the inequalities:*

$$(-1)^{k+1} D_{\xi_1}...D_{\xi_k}F(\lambda\tilde{w})\geq 0$$

hold for all positive integers k *and* $\lambda\in\mathbf{R}_{++}$ *and all* $\xi_1,...,\xi_k$ *in* $\mathbf{R}_+^n$*.*

2.3.2 Linear programming relaxation

The conditions above are not tractable in the general case but we can formulate a relaxation of the original program by simply dropping the last condition, and replacing it with a (necessary) linearity condition on $C(w,0)$ with respect to w. We get an upper bound on the upper bound (resp. a lower bound on the lower bound) solution by computing:

$$\begin{array}{ll}
\sup/\inf & C(w_0,K_0)\\
\text{subject to} & C(w,K) \text{ convex in } (K,w)\\
& C(w,K) \text{ homogeneous of degree 1}\\
& -1\leq \partial C(w,K)/\partial K\leq 0 \text{ and } C(w,K) \text{ nondecreasing in } w\\
& C(w_i,0)=w_i^Tq,\ i=1,...,m\\
& C(w_i,K_i)=p_i,\ i=1,...,m.
\end{array}\tag{2.20}$$

This is an infinite dimensional linear program in the variable $C(w,K)\in C(\mathbf{R}^{n+1})$. As we have seen in proposition (1), this infinite program can be reduced to a finite LP if we define $p_i=w_i^Tq$ and $K_i=0$ for $i=m+1,...,m+n$ and $p_{m+n+1}=w^Tq$ with $K_{m+n+1}=0$.

Proposition 11 *If the following finite LP in the variables* $p_0\in\mathbf{R}_+$ *and* $g_i\in\mathbf{R}^{n+1}$

for $i = 0, ..., m + n + 1$*:*

$$\begin{array}{ll} \textit{max./min.} & p_0 \\ \textit{subject to} & \langle g_i, (w_j, K_j) - (w_i, K_i) \rangle \leq p_j - p_i, \quad i, j = 0, ..., m + n + 1 \\ & g_{i,j} \geq 0, -1 \leq g_{i,n+1} \leq 0, \quad i = 0, ..., m + n + 1, \quad j = 1, ..., n \\ & \langle g_i, (w_i, K_i) \rangle = p_i, \quad i = 0, ..., m + n + 1, \end{array} \tag{2.21}$$

is strictly feasible and its optimal value is finite (hence it is attained), the infinite program (2.20) and its discretization (2.21) have the same optimal value. Furthermore, an optimal point of (2.20) can be constructed from the optimal solution to (2.21).

As in proposition (1), let us construct a solution to program (2.20). We first notice that as a discretization of the infinite program (2.20), the finite LP will compute a lower (or upper) bound on its optimal value. Let us now show that this bound is in fact equal to the optimal value of (2.20). If we note $z^{\text{opt}} = \left[p_0^{\text{opt}}, g_0^{\text{opt}T}, \ldots, g_k^{\text{opt}T} \right]^T$ the optimal solution to the LP problem above and if we define:

$$s(w, K) = \max_{i=0,\ldots,m+n+1} \left\{ p_i^{\text{opt}} + \left\langle g_i^{\text{opt}}, (w, K) - (w_i, K_i) \right\rangle \right\},$$

$s(x)$ satisfies

$$s(x_i) = p_i, \quad i = 1, \ldots, m + n + 1.$$

Again, by construction, $s(x_0)$ attains the lower bound p_0 computed in the finite LP. Also, $s(x)$ is convex as the pointwise maximum of affine functions and is piecewise affine with gradient g_i, which implies that it also verifies the convexity and monotonicity conditions in (2.20) and it is a feasible point of the infinite dimensional problem. This means that both problems have the same optimal value and $s(x)$ is an optimal solution to the Infinite Linear Program in (2.20).

2.4 Some cases of perfect duality

In this section, we prove that the bounds we obtained before are tight in some special cases.

2.4.1 Upper bound without forwards

We first compute the optimal probability measures corresponding to the upper and lower price bounds, when forward prices are ignored. We thus consider the problem examined in §2.2.2. Based on (2.13), we can recover an optimal distribution, or a sequence of distributions which achieve the bound in the limit. This provides a direct proof of the fact that $p^{\text{sup}} = d^{\text{sup}}$ in the case when we ignore the forward price information.

If $w^T K \geq K_0$, we choose a distribution π of asset prices such that $x = p + K$ with probability one. Then, constraints (2.2) are trivially satisfied, and the objective (2.1) becomes

$$\mathbf{E}_\pi (w^T x - K_0)_+ = (w^T (p + K) - K_0)_+ = w^T p + w^T K - K_0 = d^{\text{sup}}.$$

If $w^T K < K_0$, we have $d^{\text{sup}} = w^T p$, and the upper bound is only attained in the limit. For a given $\epsilon > 0$, we define a probability distribution $\pi(\epsilon)$ on the asset prices as follows:

$$x = \begin{cases} \epsilon^{-1} p + K & \text{with probability } \epsilon, \\ 0 & \text{with probability } 1 - \epsilon. \end{cases} \tag{2.22}$$

Then, we have

$$\mathbf{E}_{\pi(\epsilon)} (x - K)_+ = \epsilon(\epsilon^{-1} p + K - K))_+ + (1 - \epsilon)(-K)_+ = p,$$

while the objective becomes

$$\begin{aligned}\mathbf{E}_{\pi(\epsilon)}(w^T x - K_0)_+ &= \epsilon(w^T(\epsilon^{-1}p + K) - K_0)_+ + (1-\epsilon)(-K_0)_+ \\ &= (w^T p + \epsilon(w^T K - K_0))_+.\end{aligned}$$

When $\epsilon \to 0$, the above quantity goes to $w^T p = d^{\sup}$, as claimed.

2.4.2 Upper bound with forwards

We now consider the upper bound result with option and forward price constraints, obtained in §2.2.1. Without loss of generality, we assume $e^T w = 1$. In (2.11) we obtained:

$$d^{\sup} = \sup_{0 \le \beta \le 1} : w^T p + \sum_i w_i \min(q_i - p_i, \beta K_i) - \beta K_0,$$

which can be rewritten (the min is taken elementwise):

$$\sup_{0 \le \beta \le 1} \; w^T \min(q - \beta K_0 e, p + \beta\,(K - K_0)),$$

or again:

$$\sup_{0 \le \beta \le 1} \inf_{t \in [0,1]^m} \; w^T((1-t)(q - \beta K_0 e) + t(p + \beta\,(K_i - K_0))).$$

Using LP duality we know that this is also equal to (with $e^T w = 1$):

$$\inf_{t \in [0,1]^m} \sup_{0 \le \beta \le 1} \; \beta\left(w^T t K - K_0\right) + w^T(1-t)q + w^T t p.$$

We express the above as

$$\inf_{t \in [0,1]^m} \; w^T(1-t)q + w^T t p + \left(w^T t K - K_0\right)_+.$$

This problem can be solved exactly as a finite linear program, and we obtain t^{opt} such that:

$$d^{\text{sup}} = w^T \left((1-t^{\text{opt}})q + t^{\text{opt}}p\right) + \left(w^T t^{\text{opt}} K - K_0\right)_+ . \tag{2.23}$$

We recognize here the expression of the upper bound on the price of a basket, where we are only given the following option price constraints (see 2.2.2):

$$\mathbf{E}_\pi (x_i - \hat{K}_i)_+ = \hat{p}_i, \quad i = 1, \ldots, n,$$

where $\hat{K} := t^{\text{opt}} K$ and $\hat{p} := (1-t^{\text{opt}})q + t^{\text{opt}}p$. This means that we can directly recover the upper bound probability as in (2.22), substituting $(\hat{p}, \hat{K})$ with (p, K), setting $\pi(\epsilon)$ such that:

$$x = \begin{cases} \epsilon^{-1}\hat{p} + \hat{K} & \text{with probability } \epsilon, \\ 0 & \text{with probability } 1-\epsilon, \end{cases}$$

and taking the limit when $\epsilon \to 0$.

2.4.3 Lower bound without forwards

We consider the problem examined in §2.2.2. The linear programming expression (2.16) allows us to recover a sequence of distributions that are optimal in the limit, as follows.

Let ν be an optimal vector for problem (2.16). We remark that ν can be interpreted as a probability distribution. Let $\mathcal{I}$ be the set of indices i such that $K_0 > w_i K_i$. We note that $i \notin \mathcal{I}$ implies $\nu_i = 0$. For simplicity we assume that $\mathcal{I} = \{1, \ldots, m\}$, where $0 \le m \le n$ (the choice $m = 0$ corresponding to empty $\mathcal{I}$).

First we examine the case when $m = 0$, that is, $\mathcal{I}$ is empty. In other words, $\min_i w(i)K_i \ge K_0$, and therefore $d^{\text{inf}} = p^T w$. For a given $\epsilon > 0$, we choose the probability distribution on the asset prices given by (2.22), and follow the same steps taken before, for the upper bound. We obtain that d^{inf} is attained as $\epsilon \to 0$.

Next, we assume $m \geq 1$. Let $\alpha = (n-m)/m$. For ϵ such that

$$\epsilon < \alpha^{-1} \min_{1 \leq i \leq m} \nu_i (\neq 0),$$

we define the vector $\nu(\epsilon)$ by

$$\nu_i(\epsilon) = \begin{cases} \nu_i - \alpha\epsilon & \text{if } 1 \leq i \leq m, \\ \epsilon & \text{otherwise.} \end{cases}$$

Since ϵ is small enough, vector $\nu(\epsilon)$ satisfies the constraints of problem (2.16).

We now define a distribution $\pi(\epsilon)$ on the asset price vector x as follows.

$$x = x^\epsilon(i) \text{ with probability } \nu_i(\epsilon),$$

where

$$x_j^\epsilon(i) = \begin{cases} \dfrac{p_j}{\nu_j(\epsilon)} + K_j & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $x_j^\epsilon(i)$ is always well-defined, since $\nu_j(\epsilon) > 0$ for every j.

Let us check that the distribution $\pi(\epsilon)$ of asset prices satisfies the constraints (2.2). For every j, $1 \leq j \leq n$, we have

$$\begin{aligned} \mathbf{E}(x_j - K_j)_+ &= \textstyle\sum_{i=1}^n \nu_i(\epsilon)(x_j^\epsilon(i) - K_j)_+ \\ &= \nu_j(\epsilon)(x_j^\epsilon(j) - K_j)_+ \\ &= p_j. \end{aligned}$$

Let us now check that with this choice of asset price distribution, the objective

(2.1) attains the lower bound d^{inf}, when we let $\epsilon \to 0$. We have

$$\begin{aligned}
\mathbf{E}_{\pi(\epsilon)}(w^T x - K_0)_+ &= \textstyle\sum_{i=1}^n \nu_i (w^T x^\epsilon(i) - K_0)_+ \\
&= \textstyle\sum_{i=1}^n \nu_i(\epsilon)(\sum_{j=1}^n w_j x_j^\epsilon(i) - K_0)_+ \\
&= \textstyle\sum_{i=1}^n \nu_i(\epsilon)(w_i x_i^\epsilon(i) - K_0)_+ \\
&= \textstyle\sum_{i=1}^n \nu_i(\epsilon)(w_i(\frac{p_i}{\nu_i(\epsilon)} + K_i) - K_0)_+ \\
&= \textstyle\sum_{i=1}^n (w_i p_i - \nu_i(\epsilon)(K_0 - w_i K_i))_+ .
\end{aligned}$$

Letting $\epsilon \to 0$, we obtain

$$\begin{aligned}
\lim_{\epsilon \to 0} \mathbf{E}_{\pi(\epsilon)}(w^T x - K_0)_+ &= \textstyle\sum_{i=1}^m (w_i p_i - \nu_i (K_0 - w_i K_i))_+ + \sum_{i=m+1}^n w_i p_i \\
&= \textstyle\sum_{i=1}^n (w_i p_i - \nu_i (K_0 - w_i K_i)_+)_+ \\
&= d^{\text{inf}},
\end{aligned}$$

as claimed. This concludes our proof that $d^{\text{inf}} = p^{\text{inf}}$ in the absence of constraints on forward prices.

2.4.4 Tightness of the linear programming upper bound relaxation

We now show that for the special case considered in §2.2.1, namely when we have option and forward price constraints on individual assets, and we seek to compute the upper bound, the linear programming relaxation devised in (2.21) yields a tight result. In order for our problem to be feasible, we have assumed $0 \leq p \leq q \leq p + K$.

In this case, the LP (2.21) is feasible and its feasible set is compact, which ensures that there exist an optimal solution. Indeed, we can form a piecewise affine function that is feasible for (2.20) by taking $C(w_0, k_0) = E_\pi(w_0^T x - K_0)_+$, where π is the probability measure defined in (2.8), precisely

$$E_\pi(w_0^T x - K_0)_+ = \max\left\{w_0^T q - K_0, w_0^T p - K_0/2, w_0^T(q - p) - K_0/2, 0\right\}.$$

This function also turns out to correspond to a feasible point of (2.21); the variables g_i in (2.21) are simply the subgradients of $C(w_0, k_0)$ at the data points. Finally, the LP in (2.21) is finite, since we always have $0 \leq E_\pi(w_0^T x - K_0)_+ \leq w_0^T q$ and the feasible set of (2.21) is compact. This means that the optimum in (2.21) is attained.

First, we prove tightness of the LP relaxation in the case when forward price information is ignored. The setting of (2.2.2) assumes that $m = n$, and $w_0 \in \mathbf{R}_+^n$. We note e_i, the i-th unit vector. Without loss of generality, we set $w_0^T e = 1$. Since the function $C(w_0, K_0) = w_0^T p + \left(w_0^T K - K_0\right)_+$ is a feasible point of the infinite LP (2.20), if we call V^{LP} the upper bound computed by the linear program (2.21), we must have:

$$V^{\text{LP}} \geq w_0^T p + \left(w_0^T K - K_0\right)_+.$$

Now, using the necessary conditions in (2.20) and the convexity of

$$\mathbf{E}_{\pi_\varepsilon}\left(w^T x - K\right)_+$$

in (w, K) we can write

$$\begin{aligned}
\mathbf{E}_{\pi_\varepsilon}\left(w_0^T x - K_0\right)_+ &= \mathbf{E}_{\pi_\varepsilon}\left(w_0^T x - \left(w_0^T K + \left(K_0 - w_0^T K\right)\right)\right)_+ \\
&\leq \sum_{i=1}^n w_{0,i} \mathbf{E}_{\pi_\varepsilon}\left(x_i - \left(K_i + \left(K_0 - w_0^T K\right)\right)\right)_+ \\
&= \sum_{i=1}^n w_{0,i} C\left(e_i, K_i + \left(K_0 - w_0^T K\right)\right).
\end{aligned}$$

The conditions on the slope of the function $C(w, K)$ imply

$$\sum_{i=1}^n w_{0,i} C\left(e_i, K_i + \left(K_0 - w_0^T K\right)\right) \leq w_0^T p + \left(w_0^T K - K_0\right)_+.$$

Hence, $V^{\mathrm{LP}} \leq w_0^T p + \left(w_0^T K - K_0\right)_+$ and finally

$$V^{\mathrm{LP}} = w_0^T p + \left(w_0^T K - K_0\right)_+ , \tag{2.24}$$

where we recover the expression found in (2.13). This means that the upper bound computed by the LP relaxation is tight in the particular case considered above.

Now we turn to the case when forward price constraints $\mathbf{E}_\pi x_i = q_i$ for $i = 1, \ldots, n$, are included. As already observed in §2.2.1, the function

$$d^{\mathrm{sup}}(w_0, K_0) = \max_{0 \leq j \leq n+1} : w_0^T p + \sum_i w_{0,i} \min(q_i - p_i, \beta_j K_i) - \beta_j K_0,$$

is convex in (w_0, K_0). Also, when $w_0 = e_i$, and $K_0 = K_i$, we obtain $d^{\mathrm{sup}} = p_i$, while for $K_i = 0$, we obtain $d^{\mathrm{sup}} = q_i$. This means that $d^{\mathrm{sup}}(w, K)$ is a feasible point of the infinite program (2.20) and hence $V^{\mathrm{LP}} \geq d^{\mathrm{sup}}(w_0, K_0)$.

Since the finite LP (2.21) is attained, at a point denoted by

$$z^{\mathrm{opt}} = \left[p_0^{\mathrm{opt}}, g_0^{\mathrm{opt}T}, \ldots, g_k^{\mathrm{opt}T}\right]^T,$$

we can define the call price function

$$d^{\mathrm{LP}}(w, K) = \max_{i=0,\ldots,m+n+1} \left\{p_i^{\mathrm{opt}} + \left\langle g_i^{\mathrm{opt}}, (w, K) - (w_i, K_i)\right\rangle\right\},$$

corresponding to the strike prices $\hat{K} = t^{\mathrm{opt}} K$ and option prices $\hat{p} = (1 - t^{\mathrm{opt}})q + t^{\mathrm{opt}} p$, as in §2.4.2. By convexity of $d^{\mathrm{LP}}(w, K)$, we have $d^{\mathrm{LP}}(e_i, \hat{K}) \leq \hat{p}_i$ for $i = 1, \ldots, n$. We know then from (2.24) that $d^{\mathrm{LP}}(w_0, K_0) = V^{\mathrm{LP}} \leq d^{\mathrm{sup}}(w_0, K_0)$, hence finally $d^{\mathrm{LP}}(w_0, K_0) = d^{\mathrm{sup}}(w_0, K_0)$. This shows that the LP relaxation of the upper bound is tight when the input is composed of options and forward prices as in (2.7).

2.5 Summary

We are ready to summarize our results.

Theorem 12 *Tight upper and lower bounds on the price p_{basket} of an European basket call option involving n assets, with weight vector $w > 0$ and strike K_0, given the n prices p_i of individual European call options with strike $K_i > 0$, are given by*

$$p^{\mathrm{inf}} \leq p_{\mathrm{basket}} = \mathbf{E}_\pi (w^T x - K_0)_+ \leq p^{\mathrm{sup}} = \sum_{i=1}^n p_i w_i + (\sum_{i=1}^n w_i K_i - K_0)_+,$$

where

$$\begin{aligned} p^{\mathrm{inf}} = & \sum_{i \,:\, K_i w_i \geq K_0} p_i w_i \\ & + \max_{j \,:\, K_j w_j < K_0} \left(\sum_{i \,:\, K_i w_i < K_0} p_i w_i \min(1, \frac{K_0 - K_j w_j}{K_0 - K_i w_i}) - K_0 + w_j K_j \right)_+ . \end{aligned}$$

When one includes the forward contract prices information $\mathbf{E}_\pi x = q$, then the problem is feasible if and only if $p \leq q \leq p + K$. The tight upper bound then becomes

$$p^{\mathrm{sup}} = \max_{0 \leq j \leq n+1} w^T p + \sum_i w_i \min(q_i - p_i, \beta_j K_i) - \beta_j K_0,$$

with the convention $\beta_0 = 0$, $\beta_{n+1} = 1$, and $\beta_j := (q_j - p_j)/K_j$, $j = 1, \ldots, n$.

The lower bound p^{inf} is given by the solution of the linear program defined in (2.12). This bound is tight when forward prices are ignored.

In the general version of the problem, the linear programming relaxation (2.21) provides bounds in polynomial-time. The upper bound is tight in the special cases considered above.

Note that we have not proven the tightness of the lower bound in the case when individual option prices are given, and forward price constraints are included. We conjecture that the lower bound computed by (2.12) is tight, and we leave this topic for further research.

We observe that the results pertaining to our special cases (those involving individual option and forward prices only) are readily extended to a situation where we only have upper and lower bounds on these prices: simply replace the prices p_i by their upper bound in the expression for the upper bound of the basket price, and by their lower bound to compute the lower bound on the basket price.

2.6 Numerical results

We test here the various bounds obtained above on a simulated arbitrage-free data set. We first evaluate by Monte-Carlo simulation the following option prices:

$$C(w_0, K_0) = \mathbf{E}\left(w_0^T x - K_0\right)_+,$$

where $x_{i,T} = S_i \exp\left(g_i\sqrt{T} - \frac{1}{2}V_{i,i}T\right)$ for $i = 1, \ldots, 5$, with g a centered multivariate Gaussian variable with given covariance matrix V. The x_i are the simulated Black &

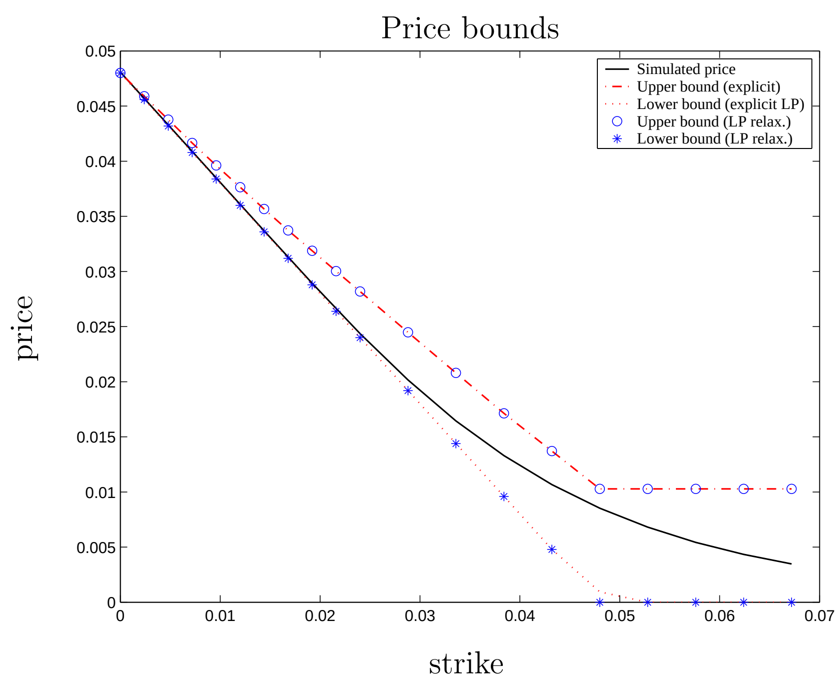

Figure 2.1: Upper and lower price bounds obtained for various strikes using both the explicit bounds and the LP relaxation method.

Scholes (1973) lognormal asset prices at maturity, with S the initial stock values. The numerical values used here are $S = \{0.7, 0.5, 0.4, 0.4, 0.4\}$, $w_0 = \{0.2, 0.2, 0.2, 0.2, 0.2\}$,

$T = 5$ years and the covariance matrix is given by:

$$V = \frac{11}{100}\begin{pmatrix} 0.64 & 0.59 & 0.32 & 0.12 & 0.06 \\ 0.59 & 1 & 0.67 & 0.28 & 0.13 \\ 0.32 & 0.67 & 0.64 & 0.29 & 0.14 \\ 0.12 & 0.28 & 0.29 & 0.36 & 0.11 \\ 0.06 & 0.13 & 0.14 & 0.11 & 0.16 \end{pmatrix}.$$

All individual options are at-the-money, hence $K = \{0.7, 0.5, 0.4, 0.4, 0.4\}$. We get $p = \{0.0161, 0.0143, 0.0093, 0.0070, 0.0047\}$. In figure (2.1), we plot the upper and lower bounds obtained for various strikes using both the explicit bounds and the LP relaxation methods. We can notice that the lower bound computed using (2.12) is tighter than that provided by the LP relaxation in (2.21). We also observe that, as showed in §2.4.4, the two upper bounds coincide.

Chapter 3

Semidefinite relaxations

In the previous two chapters, we have seen how linear programs on the values and variations of a convex functions could be used to provide tractable relaxations of some hard optimization problems. In particular we focused on the following option pricing problem:

$$\begin{array}{ll}\text{max./min.} & C(\omega_0^T, K_0)\\ \text{subject to} & C(\omega_i^T, K_i) = p_i, \quad i = 1, \ldots, m,\\ & C(\omega, K) = \mathbf{E}_\nu \left(\omega^T x - K\right)_+,\end{array}$$

which is an infinite linear program in the function $C(\omega, K)$ and the measure ν. In the first chapter, we formulated a relaxation of this problem as a shape constrained problem. In the second chapter, we studied some particular cases where this relaxation was tight. In this chapter, we look at ways of improving these relaxations by imposing progressively stricter constraints beyond the convexity requirement on $C(\omega, K)$ used in the first chapter. In general terms, we are interested in the following type of moment problem:

$$
\begin{array}{ll}
\text{minimize} & c^T z \\
\text{subject to} & Az \preceq b,\ Cz = d \\
& z = [f(x_1), \ldots, f(x_m)]^T \\
& f(x) = \mathbf{E}(g(y,x)), \quad x \in \mathbf{R}^n,
\end{array}
\tag{3.1}
$$

in the variables $f \in \mathcal{C}(\mathbf{R}^n)$, $z \in \mathbf{R}^m$, with parameters $A \in \mathbf{R}^{p \times m}$, $C \in \mathbf{R}^{q \times m}$, $c \in \mathbf{R}^m$, $b \in \mathbf{R}^p$, $d \in \mathbf{R}^q$. Problems of this type are NP-hard in general since they include for example all multivariate moment problems:

$$
\begin{array}{ll}
\text{minimize} & c^T z \\
\text{subject to} & Az \preceq b,\ Cz = d \\
& z = [f(x_1), \ldots, f(x_m)]^T \\
& f(x_i) = \mathbf{E}(y^{x_i}), \quad x_i \in \mathbf{N}_+^n,\ i = 1, \ldots, m,
\end{array}
$$

in the variables $f \in \mathcal{C}(\mathbf{R}^n)$, $z \in \mathbf{R}^m$. We will see that the dual of a multivariate moment problem is a polynomial optimization problem and problem class (3.1) also includes all problems of the form:

$$
\begin{array}{ll}
\text{minimize} & g_0(x) \\
\text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m.
\end{array}
$$

in the variable $x \in \mathbf{R}^n$, where $g_i(x) \in \mathbf{R}[x]$ are multivariate polynomials.

Note that in the case where $f(x) = \mathbf{E}(g(y,x))$ is a convex function of x, we can directly formulate a relaxation of (3.1) as a program of the form (1.1):

$$
\begin{array}{ll}
\text{minimize} & c^T z \\
\text{subject to} & Az \preceq b,\ Cz = d \\
& z = [f(x_1), \ldots, f(x_m)]^T \\
& f(x) \text{ convex.}
\end{array}
$$

This is the case when f is a polynomial moment function $f(x) = \mathbf{E}(y^x)$ or when f is the price of a call option $f(x) = \mathbf{E}((y - x)_+)$. This relaxation formulates (3.1) as a shape constrained problem of the form (1.1), which can be solved efficiently using the results in the first chapter. Here, we will then study problem (3.1) as a natural extension of the relaxation techniques discussed in the previous two chapters.

We begin in §3.1 by discussing standard semidefinite programming based relaxation techniques for nonconvex quadratically constrained quadratic programs (QCQP). These programs are very generic and include in particular all combinatorial and polynomial problems. These polynomial problems are the object of §3.2, where we discuss ways to enhance some classical relaxations using recent results in semialgebraic geometry. Finally, in §3.3, we extend the relaxation results of the previous sections to generalized moment problems. As a direct application, we revisit the static basket option arbitrage problem introduced in §1.4.3.

3.1 Semidefinite relaxations

While some special classes of nonconvex problems can be efficiently solved, most nonconvex problems are very difficult to solve (at least, globally). Here, we show how convex optimization can be used to find bounds on the optimal value of a hard polynomial problem, and can also be used to find good (but not necessarily optimal) feasible points. We first focus on Lagrangian relaxations, *i.e.*, using weak duality and the convexity of duals to get bounds on the optimal value of nonconvex problems. In a second section, we show how randomization techniques provide near optimal feasible points with, in some cases, bounds on their suboptimality.

Nonconvex QCQPs

In this note, we will focus on a specific class of problems: nonconvex quadratically constrained quadratic programs, or nonconvex QCQP (see also §4.4 in Boyd & Vandenberghe (2003)). We will see that the range of problems that can be formulated as nonconvex QCQP is vast, and we will focus on some specific examples throughout the notes. We write a nonconvex QCQP as:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array} \tag{3.2}$$

with variable $x \in \mathbf{R}^n$, and parameters $P_i \in \mathbf{S}^n$, $q_i \in \mathbf{R}^n$, and $r_i \in \mathbf{R}$. In the case where all the matrices P_i are positive semidefinite, the problem is convex and can be solved efficiently. Here we will focus on the case where at least one of the P_i is not positive semidefinite. Note that the formulation above implicitly includes problems with equality constraints, which are equivalent to two opposing inequalities.

The nonconvex QCQP is NP-hard: it is at least as hard as a large number of other problems that also seem to be hard. While no one has proved that these problems really are hard, it is widely suspected that they are, and as a practical matter, all known algorithms to solve them have a complexity that grows exponentially with problem dimensions. So it's reasonable to consider them hard to solve (globally).

Examples and applications

We list here some examples of nonconvex QCQPs.

Boolean least squares The problem is:

$$\begin{array}{ll} \text{minimize} & \|Ax-b\|^2 \\ \text{subject to} & x_i \in \{-1,1\}, \quad i=1,\ldots,n, \end{array} \tag{3.3}$$

in the variable $x \in \mathbf{R}^n$. This is a basic problem in digital communications (maximum likelihood estimation for digital signals). A brute force solution is to check all 2^n possible values of x. The problem can be expressed as a nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T A^T Ax - 2b^T Ax + b^T b \\ \text{subject to} & x_i^2 - 1 = 0, \quad i=1,\ldots,n. \end{array} \tag{3.4}$$

Minimum cardinality problems The problem is to find a minimum cardinality solution to a set of linear inequalities:

$$\begin{array}{ll} \text{minimize} & \mathbf{Card}(x) \\ \text{subject to} & Ax \preceq b, \end{array} \tag{3.5}$$

in the variable $x \in \mathbf{R}^n$, with $\mathbf{Card}(x)$ the cardinality of the set $\{i \mid x_i \neq 0\}$. We assume that the feasible set $Ax \preceq b$ is included in an Euclidean ball centered in zero with radius $R > 0$. We reformulate this problem as:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T v \\ \text{subject to} & Ax \preceq b \\ & -Rv \preceq x \preceq Rv \\ & v \in \{0,1\}^n, \end{array} \tag{3.6}$$

in the variables $x, v \in \mathbf{R}^n$, and we then turn this into a nonconvex QCQP by replacing the constraints $v_i \in \{0,1\}$ by $v_i^2 - v_i = 0$. The problem then becomes:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T v \\ \text{subject to} & Ax \preceq b \\ & -Rv \preceq x \preceq Rv \\ & v_i^2 - v_i = 0, \quad i = 1, \ldots, n. \end{array} \tag{3.7}$$

This problem has many applications in engineering and finance, including for example low-order controller design and portfolio optimization with fixed transaction costs.

Partitioning problems We consider here the two-way partitioning problem described in §5.1.4 and exercise 5.39 of Boyd & Vandenberghe (2003):

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{array} \tag{3.8}$$

with variable $x \in \mathbf{R}^n$, where $W \in \mathbf{S}^n$ satisfies $W_{ii} = 0$. This problem is directly a nonconvex QCQP of the form (3.2). A feasible x corresponds to the partition

$$\{1, \ldots, n\} \;=\; \{i \mid x_i = -1\} \;\cup\; \{i \mid x_i = 1\},$$

and the matrix coefficient W_{ij} can be interpreted as the cost of having the elements i and j in the same partition, with $-W_{ij}$ the cost of having i and j in different partitions. The objective in (3.8) is the total cost, over all pairs of elements, and problem (3.8) seeks to find the partition with least total cost.

MAXCUT MAXCUT is a classic problem in network optimization and a particular case of the partitioning problem above. Here $W \in \mathbf{S}^n$ is a matrix with nonnegative coefficients and $W_{ij} = 0$ if no arc connects nodes i and j in the network. The problem is formulated as:

$$\begin{array}{ll} \text{maximize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{array} \tag{3.9}$$

with variable $x \in \mathbf{R}^n$. The objective here is to find a partition of the set such that the sum of the coefficients W_{ij} of the nodes linking the two partitions is maximized (hence the name MAXCUT).

Polynomial problems A polynomial problem seeks to minimize a polynomial over a set defined by polynomial inequalities:

$$\begin{array}{ll} \text{minimize} & p_0(x) \\ \text{subject to} & p_i(x) \leq 0, \quad i = 1, \ldots, m. \end{array}$$

While seemingly much more general than simple nonconvex QCQPs, all polynomial problems can be turned into nonconvex QCQPs. Let us briefly detail how. First, we notice that we can reduce the maximum degree of an equation by adding variables. For example, we can turn the constraint

$$y^{2n} + (\ldots) \leq 0$$

into

$$u^n + (\ldots) \leq 0, \qquad u = y^2.$$

We have reduced the maximum degree of the original inequality by introducing a new variable and a quadratic equality constraint. We can also get rid of product terms; this time

$$xyz + (\ldots) \leq 0$$

becomes

$$ux + (\ldots) \leq 0, \qquad u = yz.$$

Here, we have replaced a product of three variables by a product of two variables (quadratic) plus an additional quadratic equality constraint. By applying these transformations iteratively, we can transform the original polynomials into quadratic objective and constraints, thus turning the original polynomial problem into a nonconvex QCQP, with additional variables.

Example. Let's work out a specific example. Suppose that we want to solve the following polynomial problem:

$$\begin{array}{ll} \text{minimize} & x^3 - 2xyz + y + 2 \\ \text{subject to} & x^2 + y^2 + z^2 - 1 = 0, \end{array}$$

in the variables $x, y, z \in \mathbf{R}$. We introduce two new variables $u, v \in \mathbf{R}$ with

$$u = x^2, \quad v = yz.$$

The problem then becomes:

$$\begin{array}{ll} \text{minimize} & xu - 2xv + y + 2 \\ \text{subject to} & x^2 + y^2 + z^2 - 1 = 0 \\ & u - x^2 = 0 \\ & v - yz = 0, \end{array}$$

which is a nonconvex QCQP of the form (3.2), in the variables $x, y, z, u, v \in \mathbf{R}$.

3.1.1 Convex relaxations

In this section, we begin by describing some direct relaxations of (3.2) using semidefinite programming (cf. Vandenberghe & Boyd (1996)). We then detail how Lagrangian duality can be used as an "automatic" procedure to get lower bounds on the optimal value of the nonconvex QCQP described in (3.2). Note that both techniques provide lower bounds on the optimal value of the problem but give only a minimal hint on how to find an approximate solution (or even a feasible point . . .), this will be the object of the next section.

Semidefinite relaxations

Starting from the original nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array}$$

using $x^T P x = \mathbf{Tr}(P(xx^T))$, we can rewrite it:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \\ & X = xx^T. \end{array} \tag{3.10}$$

We can directly relax this problem into a convex problem by replacing the last nonconvex equality constraint $X = xx^T$ with a (convex) positive semidefiniteness constraint $X - xx^T \succeq 0$. We then get a lower bound on the optimal value of (3.2) by solving the following convex problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \\ & X \succeq xx^T. \end{array}$$

The last constraint $X \succeq xx^T$ is convex and can be formulated as a Schur complement (see §A.5.5 in Boyd & Vandenberghe (2003)):

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad \text{i=1,...,m}, \\ & \begin{bmatrix} X & x^T \\ x & 1 \end{bmatrix} \succeq 0, \end{array} \tag{3.11}$$

which is an SDP. This is called the *SDP relaxation* of the original nonconvex QCQP. Its optimal value is a lower bound on the optimal value of the nonconvex QCQP. Since it's an SDP, it's easy to solve, so we have a cheaply computable lower bound on the optimal value of the original nonconvex QCQP.

Lagrangian relaxations

We now study another method for getting a cheaply computable lower bound on the optimal value of the nonconvex QCQP. We take advantage of the fact that the dual of a problem is always convex, hence efficiently solvable. Again, starting from the original nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array}$$

we form the Lagrangian,

$$L(x, \lambda) = x^T \left(P_0 + \sum_{i=1}^m \lambda_i P_i \right) x + \left(q_0 + \sum_{i=1}^m \lambda_i q_i \right)^T x + r_0 + \sum_{i=1}^m \lambda_i r_i.$$

To find the dual function, we minimize over x, using the general formula (see example 4.5 in Boyd & Vandenberghe (2003)):

$$\inf_{x \in \mathbf{R}} x^T P x + q^T x + r = \begin{cases} r - \frac{1}{4} q^T P^\dagger q, & \text{if } P \succeq 0 \text{ and } q \in \mathcal{R}(P) \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual function is then:

$$
\begin{aligned}
g(\lambda) &= \inf_{x \in \mathbf{R}^n} L(x,\lambda) \\
&= -\frac{1}{4}\left(q_0 + \sum_{i=1}^m \lambda_i q_i\right)^T \left(P_0 + \sum_{i=1}^m \lambda_i P_i\right)^\dagger \left(q_0 + \sum_{i=1}^m \lambda_i q_i\right) \\
&\quad + \sum_{i=1}^m \lambda_i r_i + r_0.
\end{aligned}
$$

We can form the dual of (3.2), using Schur complements (cf. §A.5.5):

$$
\begin{array}{ll}
\text{maximize} & \gamma + \sum_{i=1}^m \lambda_i r_i + r_0 \\
\text{subject to} & \begin{bmatrix} \left(P_0 + \sum_{i=1}^m \lambda_i P_i\right) & \left(q_0 + \sum_{i=1}^m \lambda_i q_i\right)/2 \\ \left(q_0 + \sum_{i=1}^m \lambda_i q_i\right)^T/2 & -\gamma \end{bmatrix} \succeq 0 \\
& \lambda_i \geq 0, \quad i = 1, \ldots, m,
\end{array}
\tag{3.12}
$$

in the variable $\lambda \in \mathbf{R}^m$. As the dual to (3.2), this is a convex program, it is in fact a semidefinite program. This SDP is called the *Lagrangian relaxation* of the nonconvex QCQP. It's easy to solve, and gives a lower bound on the optimal value of the nonconvex QCQP.

An interesting question is, what is the relation between the Lagrangian relaxation and the SDP relaxation? They are both SDPs, and they both provide lower bounds on the optimal value of the nonconvex QCQP. In particular, is one of the bounds better than the other? The answer turns out to be simple: (3.11) and (3.12) are dual of each other, and so (assuming a constraint qualification holds) the bounds are exactly the same.

Perfect duality

Weak duality implies that the optimal value of the Lagrangian relaxation is a lower bound on that of the original program. In some particular cases, even though the original program is not convex, this duality gap is zero and the convex relaxation produces the optimal value.

QCQP with only one constraint is a classic example (see Appendix B in Boyd & Vandenberghe (2003), or Feron (1999) for others), based on the fact that the numerical range of two quadratic forms is a convex set. This means that, under some technical conditions, the programs:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_1 x + q_1^T x + r_1 \leq 0, \end{array} \tag{3.13}$$

and

$$\begin{array}{ll} \text{maximize} & \gamma + \lambda r_1 + r_0 \\ \text{subject to} & \left[\begin{array}{cc} (P_0 + \lambda P_1) & (q_0 + \lambda q_1)/2 \\ (q_0 + \lambda q_1)^T/2 & -\gamma \end{array} \right] \succeq 0 \\ & \lambda \geq 0, \end{array} \tag{3.14}$$

in the variables $x \in \mathbf{R}^n$ and $\lambda \in \mathbf{R}$ respectively, produce the same optimal value, even if the first one is nonconvex. This result is also known as the S-procedure in control theory. The key implication here of course is that while the original program is possibly nonconvex and numerically hard, its dual is a semidefinite program and is easy to solve.

Examples

Let us now work out the Lagrangian relaxations of the examples detailed above.

MINCARD relaxation Let's first consider the MINCARD problem detailed in (3.5):

$$\begin{array}{ll} \text{minimize} & \mathbf{Card}(x) \\ \text{subject to} & Ax \preceq b. \end{array}$$

Using the problem formulation in (3.7), the relaxation given by (3.11) is then:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T v \\ \text{subject to} & Ax \preceq b \\ & -Rv \preceq x \preceq Rv \\ & \mathbf{Tr}(e_i e_i^T X) - e_i^T x = 0, \quad i = 1, \ldots, n \\ & \left[\begin{array}{cc} X & x^T \\ x & 1 \end{array}\right] \succeq 0, \end{array}$$

where e_i is the Euclidean basis in $\mathbf{R}^n$. Both Boyd, Fazel & Hindi (2000) and Lemaréchal & Oustry (1999, Th. 5.2) show that this relaxation produces the same lower bound as the direct linear programming relaxation:

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T v \\ \text{subject to} & Ax \preceq b \\ & -Rv \preceq x \preceq Rv \\ & v \in [0,1]^n. \end{array} \tag{3.15}$$

It is also related to the classical ℓ_1 heuristic described in Boyd et al. (2000), which replaces the function $\mathbf{Card}(x)$ with its largest convex lower bound (over a ball in ℓ_∞) $\|x\|_1$:

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax \preceq b. \end{array} \tag{3.16}$$

Boolean least squares The original boolean least squares problem in (3.3) is written:

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|^2 \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{array}$$

We can relax its QCQP formulation (3.4) as an SDP:

$$\begin{array}{ll}\text{minimize} & \mathbf{Tr}(AX) + 2b^T Ax + b^T b \\ \text{subject to} & \begin{bmatrix} X & x^T \\ x & 1 \end{bmatrix} \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \ldots, n,\end{array} \tag{3.17}$$

in the variables $x \in \mathbf{R}^n$ and $X \in \mathbf{S}_+^n$. This program then produces a lower bound on the optimal value of the original problem.

Partitioning and MAXCUT The partitioning problem defined above reads:

$$\begin{array}{ll}\text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n.\end{array} \tag{3.18}$$

Here, the problem is directly formulated as a nonconvex QCQP and the variable x disappears from the relaxation, which becomes:

$$\begin{array}{ll}\text{minimize} & \mathbf{Tr}(WX) \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \ldots, n.\end{array} \tag{3.19}$$

MAXCUT corresponds to a particular choice of matrix W.

3.1.2 Randomization

The Lagrangian relaxation techniques developed in §3.1.1 provided lower bounds on the optimal value of the program in (3.2), but did not however give any particular hint on how to compute good feasible points. The semidefinite relaxation in (3.11) produces a positive semidefinite or covariance matrix together with the lower bound on the objective. In this section, we exploit this additional output to compute good approximate solutions with, in some cases, hard bounds on their suboptimality.

Randomization

In the last section, the original nonconvex QCQP:

$$\begin{array}{ll} \text{minimize} & x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \end{array}$$

was relaxed into:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XP_0) + q_0^T x + r_0 \\ \text{subject to} & \mathbf{Tr}(XP_i) + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m, \\ & \begin{bmatrix} X & x^T \\ x & 1 \end{bmatrix} \succeq 0. \end{array} \tag{3.20}$$

The last (Schur complement) constraint being equivalent to $X - xx^T \succeq 0$, if we suppose x and X are the solution to the relaxed program in (3.20), then $X - xx^T$ is a covariance matrix.

If we pick x as a Gaussian variable with $x \sim \mathcal{N}(x, X - xx^T)$, x will solve the nonconvex QCQP in (3.2) "on average" over this distribution, meaning:

$$\begin{array}{ll} \text{minimize} & \mathbf{E}(x^T P_0 x + q_0^T x + r_0) \\ \text{subject to} & \mathbf{E}(x^T P_i x + q_i^T x + r_i) \leq 0, \quad i = 1, \ldots, m, \end{array}$$

and a "good" feasible point can then be obtained by sampling x a sufficient number of times, then simply keeping the best feasible point.

Feasible points

Of course the direct sampling technique above does not guarantee that a feasible point will be found. In particular, if the program includes an equality constraint, then this method will certainly fail. However, it is sometimes possible to directly project the random samples onto the feasible set. This is the case, for example, in the partitioning problem, where we can discretize the samples by taking the $\mathbf{sgn}(x)$ function. In this case, the randomization procedure then looks like this. First, sample points x_i with a normal distribution $\mathcal{N}(0, X)$, where X is the optimal solution of (3.19). Then get feasible points by taking $\hat{x}_i = \mathbf{sgn}(x_i)$.

Bounds on suboptimality

In certain particular cases, it is also possible to get a hard bound on the gap between the optimal value and the relaxation result. A classic example is that of the MAXCUT bound described in Goemans & Williamson (1995) or Ben-Tal & Nemirovski (2001, Th. 4.3.2). The MAXCUT problem (3.9) reads:

$$\begin{array}{ll} \text{maximize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{array} \tag{3.21}$$

its Lagrangian relaxation was computed in (3.19):

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(WX) \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \ldots, n. \end{array} \tag{3.22}$$

We then sample feasible points $\hat{x}_i$ using the procedure described above. Crucially, when $\hat{x}$ is sampled using that procedure, the expected value of the objective $\mathbf{E}(\hat{x}^T W x)$ can be computed explicitly:

$$\mathbf{E}(\hat{x}^T W x) = \frac{2}{\pi} \sum_{i,j=1}^{n} W_{ij} \arcsin(X_{ij}) = \frac{2}{\pi} \mathbf{Tr}(W \arcsin(X)).$$

We are guaranteed to reach this expected value $2/\pi\, \mathbf{Tr}(W \arcsin(X))$ after sampling a few (feasible) points $\hat{x}$, hence we know that the optimal value OPT of the MAXCUT problem is between $2/\pi\, \mathbf{Tr}(W \arcsin(X))$ and $\mathbf{Tr}(WX)$.

Furthermore, with $\arcsin(X) \succeq X$ (see Ben-Tal & Nemirovski (2001, p. 174)), we can simplify (and relax) the above expression to get:

$$\frac{2}{\pi}\,\mathbf{Tr}(WX) \leq OPT \leq \mathbf{Tr}(WX).$$

This means that the procedure detailed above guarantees that we can find a feasible point that is at most $2/\pi$ suboptimal (after taking a certain number of samples from a Gaussian distribution).

3.1.3 Linearization and convex restriction

The relaxation techniques detailed in §3.1.1 produce lower bounds on the optimal value but no feasible points. Here, we work on the complementary approach and try to find "good" feasible points corresponding to a local minimum. Let $x^{(0)}$ be an initial feasible point which might be found using the results in the last section or by a phase I procedure; see the discussion on phase I problems in §11.4 of Boyd & Vandenberghe (2003)).

Linearization

We start by leaving all convex constraints unchanged, linearizing the nonconvex ones around the original feasible point $x^{(0)}$. Consider for example the constraint:

$$x^T P x + q^T x + r \leq 0,$$

we decompose the matrix P into its positive and negative parts:

$$P = P_+ - P_-, \qquad P_+,\ P_- \succeq 0.$$

The original constraint can be rewritten as

$$x^T P_+ x + q_0^T x + r_0 \leq x^T P_- x,$$

and both sides of the inequality are now convex quadratic functions. We linearize the right hand side around the point x_0 to obtain

$$x^T P_+ x + q_0^T x + r_0 \leq x^{(0)T} P_- x^{(0)} + 2x^{(0)T} P_-(x - x^{(0)}).$$

The right hand side is now an affine lower bound on the original function $x^T P_- x$ (see §3.1.3 in Boyd & Vandenberghe (2003)). This means that the resulting constraint is convex and more conservative than the original one, hence the feasible set of the new problem will be a convex subset of the original feasible set. Thus, by linearizing the concave parts of all constraints, we obtain a set of convex constraints that are tighter than the original nonconvex ones. In other words, we form a *convex restriction* of the problem.

Iteration

The new problem, formed by linearizing all the nonconvex constraints using the method described above, is a convex QCQP and can be solved efficiently to produce a new feasible point $x^{(1)}$ with a lower objective value. If we linearize again the problem around $x^{(1)}$ and repeat the procedure, we get a sequence of feasible points with decreasing objective values.

This simple idea has been discovered and rediscovered several times. It is sometimes called the *convex-concave procedure*. It doesn't work for Boolean problems, since the only convex subsets of the feasible set are singletons.

3.1.4 Numerical examples

In this section, we work out some numerical examples.

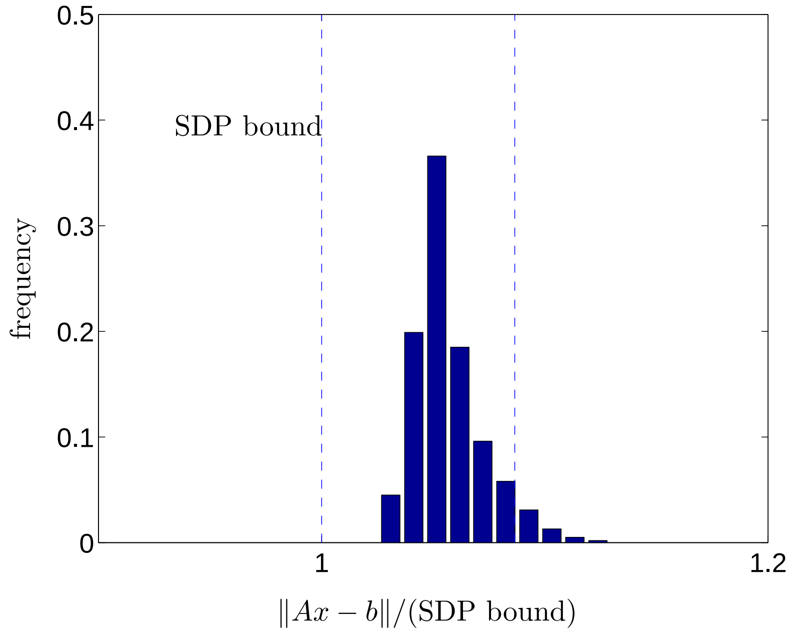

Figure 3.1: Distribution of objective values for points sampled using the randomization technique in §3.1.2.

Boolean least-squares

We use (randomly chosen) parameters $A \in \mathbf{R}^{150\times 100}$, $b \in \mathbf{R}^{150}$ and $x \in \mathbf{R}^{100}$, the feasible set has $2^{100} \approx 10^{30}$ points. In figure (3.1.4), we plot the distribution of the objective values reached by the feasible points found using the randomized procedure above. Our best solution comes within 2.6% of the SDP lower bound.

Partitioning

We consider here the two-way partitioning problem described in (3.8) and compare various methods.

A simple heuristic for partitioning One simple heuristic for finding a good partition is to solve the SDPs above, to find $X^\star$ (and the bound $d^\star$). Let v denote an eigenvector of $X^\star$ associated with its largest eigenvalue, and let $\hat{x} = \mathbf{sgn}(v)$. The vector $\hat{x}$ is our guess for a good partition.

A randomized method We generate independent samples using the procedure described in §3.1.2.

Greedy method We can improve these results a little bit using the following simple greedy heuristic. Suppose the matrix $Y = \hat{x}^T W \hat{x}$ has a column j whose sum $\sum_{i=1}^n y_{ij}$ is positive. Switching $\hat{x}_j$ to $-\hat{x}_j$ will decrease the objective by $2\sum_{i=1}^n y_{ij}$. If we pick the column y_{j_0} with largest sum, switch $\hat{x}_{j_0}$ and repeat until all column sums $\sum_{i=1}^n y_{ij}$ are negative, we further decrease the objective.

Numerical Example For our example, the optimal SDP lower bound $d^\star$ is equal to -1641 and the $\mathbf{sgn}(x)$ heuristic gives a point (partition) with total cost -1348. Extracting a solution from the SDP solution using the simple heuristic above gives a solution with cost -1280, while applying the greedy method pushes that cost down to -1372. Exactly what the optimal value is, we can't say; all we can say at this point is that it is between -1641 and -1372.

We then try the randomized method, applying the greedy method to each sample, and plot in figure (3.1.4) a histogram of the objective obtained over 1000 samples. Many of these samples have an objective value larger than the original one above, but some have a lower cost. For our implementation, we found the minimum value -1392. The evolution of the minimum value found as a function of the sample size is shown in figure (3.1.4). Note that our best partition was found in around 100 samples. We're not sure what the optimal cost is, but now we know it's between -1641 and -1392.

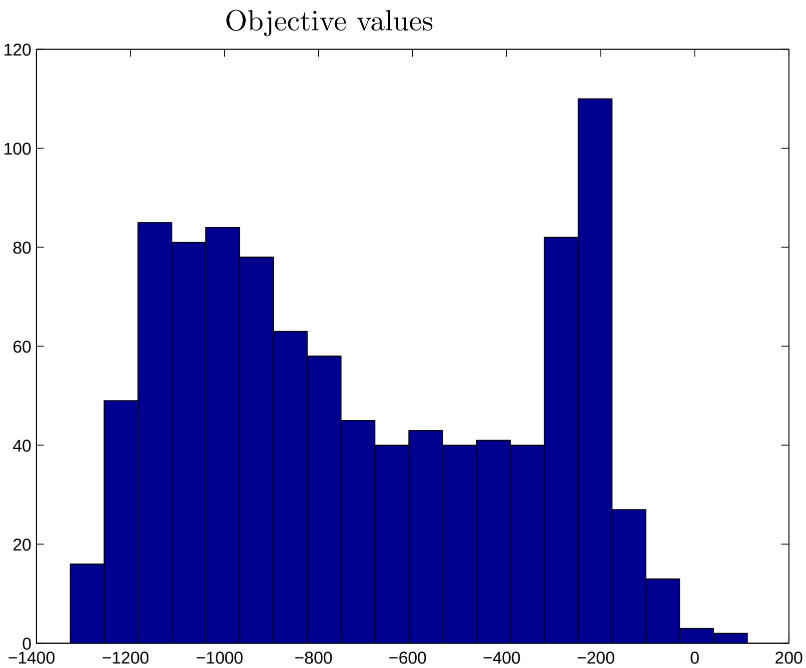

Figure 3.2: Histogram of the objective values attained by the random sample partitions.

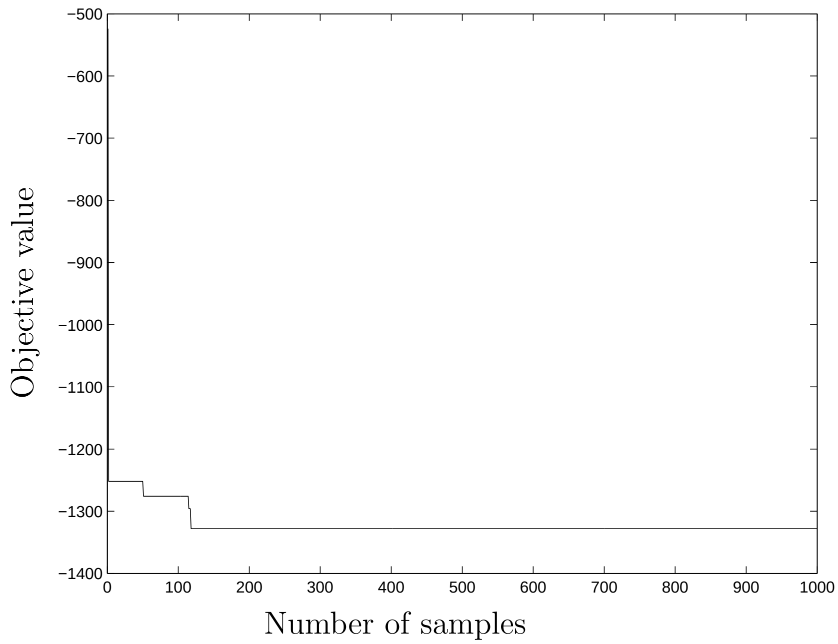

Figure 3.3: Best objective value versus number of sample points.

3.2 Polynomial problems

Here, we discuss direct extensions of the results detailed in the last section, using techniques developed in Putinar (1993), Lasserre (2001) or Parillo & Sturmfels (2001). As an application, we show that some minimum cardinality problems subject to linear inequalities can be represented as finite sequences of semidefinite programs. In particular, we provide a semidefinite representation of the minimum rank problem on positive semidefinite matrices.

Notation

We note $\mathbf{R}[x_1, ..., x_n]$ (or $\mathbf{R}[x]$ when there is no ambiguity) the ring of multivariate polynomials $p(x) = p(x_1, ..., x_n)$ on a variable $x \in \mathbf{R}^n$. We say that $p(x) \in \mathbf{R}[x]$ is *SOS* when $p(x)$ is a sum of squares of polynomials in $\mathbf{R}[x]$. For $x \in \mathbf{R}^n$, $\mathbf{Card}(x)$ will be the cardinal of the set $\{i : x_i \neq 0\}$. We note $\mathbf{S}^n$ the set of $n \times n$ symmetric matrices.

For multivariate polynomials, we adopt the multiindex notation $p(x) = \sum_\alpha p_\alpha x^\alpha$, where $x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n}$, and we note $d = \sum_{i=1}^n \alpha_i$ the degree of $p(x)$. $\mathbf{R}_d[x]$ is the set of polynomials of degree at most d. Finally, $C(p)$ will be the *Newton polytope* of the polynomial $p(x)$, with $C(p) = \mathbf{Co}\left(\{\alpha : p_\alpha \neq 0\}\right)$.

3.2.1 Introduction

Given a convex set $\mathcal{C} \subset \mathbf{R}^n$, we are interested in solving the following problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{Card}(x) \\ \text{subject to} & x \in \mathcal{C}, \end{array} \tag{3.23}$$

in the particular case where $\mathcal{C}$ is described by a set of linear inequalities. Except in certain rare instances, this problem is very hard to solve (see Vandenberghe & Boyd (1996)). Excellent heuristics exist however, a classical one (see Hassibi, How & Boyd (1999) for example) replacing the function $\mathbf{Card}(x)$ by $\|x\|_1$, its largest convex lower bound on the unit cube.

A related problem is that of minimizing the rank of a p.s.d. matrix subject to LMI constraints:

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}(X) \\ \text{subject to} & X \in \mathcal{C}, \end{array} \tag{3.24}$$

where $\mathcal{C}$ is here an affine subset of the semidefinite cone (a LMI). In this case also, minimizing the nuclear norm $\|X\|_{\text{opt}}$ of X will produce excellent approximate solutions (see Boyd et al. (2000)).

In this paper, using results by Cassier (1984), Shor (1987), Putinar (1993), Choi, Lam & Reznick (1995), Nesterov (2000), Lasserre (2001), Parillo & Sturmfels (2001) and Lasserre (2002), we show that the $\mathbf{Min\,Card}(x)$ and $\mathbf{Min\,Rank}(X)$ problems in (3.23) and (3.24) are equivalent to large semidefinite programs (see Nesterov & Nemirovskii (1994)). To be precise, based on a reformulation à la Shor (1987) of problems (3.23) and (3.24), we use the technique in Lasserre (2002) to produce a finite (possibly exponential) sequence of increasingly tighter semidefinite relaxations.

The rest of the paper is organized as follows. In §3.2.2, we recall some key definitions and properties on semidefinite representability and the sum of squares representation of positive polynomials. We also summarize the application of these representations to semialgebraic problems. In §3.2.3, we show that both the $\mathbf{Min\,Card}(x)$ and the $\mathbf{Min\,Rank}(X)$ problems are equivalent to large scale semidefinite programs. Based on the work by Putinar (1993), Nesterov (2000) and Lasserre (2002) we explicitly construct in §3.2.4 a sequence of semidefinite programs solving problems (3.23) and (3.24). We also show how the problem of finding optimal convex lower bounds on the objective function can be represented in a similar way. Finally, in §3.2.5, we discuss the complexity of these techniques.

3.2.2 Sums of squares and semidefinite programming

We quickly recall here some key definitions and properties linking semidefinite and semialgebraic problems.

Hilbert's 17^{th} problem (see Reznick (1996) for an overview), which asked if all positive polynomials could be written as sums of squares of other polynomials, has a positive answer in dimension one. Nesterov (2000) provides an efficient way of computing the SOS representation of a given positive univariate polynomial in the following result from Nesterov (2000).

Proposition 13 *Let $p(x) \in \mathbf{R}[x]$ be a univariate polynomial of degree d. Then $p(x)$ for all $x \in \mathbf{R}$ iff there exists a matrix $X \in \mathbf{S}^v$ such that:*

$$p(x) = y_x^T X y_x, \text{ with } X \succeq 0, \quad \text{for all } x \in \mathbf{R} \tag{3.25}$$

with $v = \lceil d/2 \rceil$ and $y_x = (1, x, x^2, ..., x^v)$ is the list of univariate monomials up to degree v.

The coefficients of the polynomials in the representation are then computed as the eigenvectors of the matrix X.

In the general multivariate case, that representation property of positive polynomials is lost. It can be shown (see Berg (1980)) that the set of multivariate SOS polynomials is dense in the set of positive polynomials, but there are simple examples of positive polynomials that are not SOS. However, recent results in semialgebraic geometry (see Cassier (1984), Schmüdgen (1991), Putinar (1993) or Putinar & Vasilescu (1999)) bridge the gap between positive and SOS polynomials on compact semialgebraic sets. We cite here the result in Putinar (1993). Let $g_k(x) \in \mathbf{R}[x_1, ..., x_n]$ for $j = 1, ..., r$, and we note K, the semialgebraic set defined by

$$K = \{x \in \mathbf{R}^n : g_k(x) \geq 0, k = 1, ..., r\}.$$

We suppose that K is compact and that there exists $u(x) \in \mathbf{R}[x_1, ..., x_n]$ such that $\{u(x) \geq 0\}$ is compact with

$$u(x) = u_0(x) + \sum_{k=1}^{r} g_k(x) u_k(x), \quad \text{for all } x \in \mathbf{R}^n \tag{3.26}$$

where the polynomials $u_k(x) \in \mathbf{R}[x_1, ..., x_n]$ are SOS for $k = 1, ..., r$. Under this assumption, we can represent all polynomials positive on K using SOS polynomials as in Putinar (1993) or Putinar & Vasilescu (1999).

Proposition 14 *Suppose (3.26) holds. A polynomial $p(x) \in \mathbf{R}[x_1, ..., x_n]$ is positive on K iff:*

$$p(x) = q_0(x) + \sum_{k=1}^{r} g_k(x) q_k(x), \quad \textit{for all } x \in \mathbf{R}^n \tag{3.27}$$

where the polynomials $q_k(x) \in \mathbf{R}[x_1, ..., x_n]$ are SOS for $k = 0, ..., r$.

Now, as in Parillo (2000) or Lasserre (2001), we can write the multivariate version of the relation (3.25) mapping SOS polynomials to the semidefinite cone.

Proposition 15 *Let $p(x) \in \mathbf{R}[x_1, ..., x_n]$ be a polynomial and K a semialgebraic set* *defined by* $K = \{x \in \mathbf{R}^n : g_k(x) \geq 0, k = 1, ..., r\}$*, satisfying assumption (3.26). Then* $p(x) \geq 0$ *on* K *iff there is an integer* $m \in \mathbf{Z}_+$ *and matrices* $X_k \in \mathbf{S}^N$*, with* $X_k \succeq 0$ *for* $k = 0, ..., r$ *such that:*

$$p(x) = y_x^T X_0 y_x + \sum_{k=1}^{r} \left(y_x^T X_k y_x\right) g_k(x), \quad \textit{for all } x \in \mathbf{R}^n, \tag{3.28}$$

where $N = \left\lceil \binom{n+m-1}{m} / 2 \right\rceil$ *and*

$$y_x = \left(1, x_1, ..., x_n, x_1^2, x_1 x_2, ..., x_1 x_n, x_2^2, x_2 x_3, ..., x_n^2, ..., x_1^m, ..., x_n^m\right)$$

is the vector of all monomials in $\mathbf{R}[x_1, ..., x_n]$*, up to degree* m*, listed in graded lexicographic order.*

The result on polynomials above shows that testing the positivity of a multivariate polynomial on a semialgebraic set K satisfying the assumption (3.26) can be cast as a semidefinite program. In general, the result in Lasserre (2001) shows that all compact semialgebraic problems, i.e. problems seeking to minimize a polynomial over a compact semialgebraic set, are equivalent to large-scale semidefinite programs. This provides a positive answer in the compact multivariate case to all the open questions in §4.10.2 in Ben-Tal & Nemirovski (2001). A converse result is also true (and much simpler). Because the positive semidefiniteness of a matrix is equivalent to that of all its principal minors, all semidefinite programs are semialgebraic programs, with additional convexity and invariance properties.

The central result of moment theory exploited in Lasserre (2001) sets polynomial positivity problems and moment problems as duals (see e.g. Berg (1980)). Let s be a positive semidefinite sequence $s \in \mathbf{R}^N$, we have

$$\begin{gathered} s \text{ is p.s.d.} \\ \Updownarrow \\ \langle s, p_\alpha \rangle \geq 0, \quad \text{for all } p(x) \in \mathbf{R}_m[x] \text{ with } p(x)\ \mathit{SOS}, \end{gathered}$$

and

$$s \text{ is a moment sequence}$$
$$\Updownarrow$$
$$\langle s, p_\alpha \rangle \geq 0, \quad \text{for all } p(x) \in \mathbf{R}_m[x] \text{ with } p(x) \geq 0 \text{ on } \mathbf{R}^n,$$

hence the cone of coefficients of SOS polynomials and that of p.s.d. sequences are polar, and so are the cones of moment sequences and positive polynomials.

From Putinar (1993) then, we know that the problem of testing if a sequence y is the moment sequence of some measure μ with support in a compact semialgebraic set

$$K = \{x \in \mathbf{R}^n : g_k(x) \geq 0, k = 1, ..., r\}$$

and the problem of representing positive polynomials on K as $g_k(x)$ weighted sums of SOS polynomials are dual of each other and both representable as linear matrix inequalities.

3.2.3 Semidefinite representations of MinCard(x) and MinRank(X)

As above K is the semialgebraic set defined by

$$K = \{x \in \mathbf{R}^n : g_k(x) \geq 0, k = 1, ..., r\}$$

and we assume that (3.26) holds. Let $p(x) \in \mathbf{R}[x_1, ..., x_n]$ and as K is compact, we also note $t^{\text{opt}} = \min_K p(x)$. Of course, $p(x) - t^{\text{opt}} \geq 0$ on K, hence there are SOS polynomials $q_k(x) \in \mathbf{R}[x_1, ..., x_n]$ for $k = 1, ..., r$ such that (3.27) holds for $p(x) - t^{\text{opt}}$ on K. We first show that the $\mathbf{Min\,Card}(x)$ problem can be cast as a semialgebraic program, hence a semidefinite program, using the results from §3.2.2.

Proposition 16 *Let $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$. There are polynomials $g_k(x) \in$*

$\mathbf{R}[x_1, ..., x_n]$, *for* $k = 0, ..., r$ *such that the optimum values of:*

$$\begin{array}{ll} \textit{minimize} & \mathbf{Card}(x) \\ \textit{subject to} & Ax \geq b, \end{array} \tag{3.29}$$

and

$$\begin{array}{ll} \textit{minimize} & g_0(x) \\ \textit{subject to} & g_k(x) \succeq 0, \quad \textit{for } k = 1, ..., r, \end{array}$$

are equal.

Proof. First, as in Shor (1987) we notice that:

$$\begin{array}{rll} \mathbf{Card}(x) = & \min & \sum_{i=1}^n v_i \\ & \text{s.t.} & (v_i - 1)x_i = 0 \\ & & v_i \geq 0, \quad \text{for } i = 1, ..., n, \end{array}$$

hence the $\mathbf{Min\,Card}(x)$ problem in (3.29) can be written:

$$\begin{array}{rll} \mathbf{Min\,Card}(x) \equiv & \text{min.} & \sum_{i=1}^n v_i \\ & \text{s.t.} & (v_i - 1)x_i = 0 \\ & & v_i \geq 0, \quad \text{for } i = 1, ..., n \\ & & Ax \geq b. \end{array}$$

which is a semialgebraic problem. ■

We now show a similar result on the $\mathbf{Min\,Rank}(X)$, a minimum cardinality problem on the eigenvalues of the matrix X.

Proposition 17 *Let* $A_i \in \mathbf{S}^n$, *for* $i = 1, ..., p$ *and* $b \in \mathbf{R}^p$. *There are polynomials*

$g_k(x) \in \mathbf{R}[x_1, ..., x_n]$, *for* $k = 0, ..., r$ *such that the optimum values of:*

$$\begin{array}{ll} \textit{minimize} & \mathbf{Rank}(X) \\ \textit{subject to} & Tr\,(A_i X) = b_i, \quad \textit{for } i = 1, ..., p \\ & X \succeq 0 \end{array} \tag{3.30}$$

and

$$\begin{array}{ll} \textit{minimize} & g_0(x) \\ \textit{subject to} & g_k(x) \succeq 0, \quad \textit{for } i = 1, ..., M \end{array}$$

are equal.

Proof. We note $(\lambda_1, ..., \lambda_n)$ the eigenvalues of the matrix $X \succeq 0$ and

$$\sigma_k(X) = \sum_{\alpha \subset \{1,...,n\}, |\alpha| = k} \lambda_\alpha$$

the symmetric functions. We note $\chi_t(X)$ the characteristic polynomial of the matrix X, with $\chi_t(X) = \sum_{i=1}^n (-1)^i \sigma_i(X) t^i$. Because the matrix X is semidefinite positive, we have $\sigma_k(X) = 0$ iff $\mathbf{Rank}\,(X) < k$, hence the $\mathbf{Min}\,\mathbf{Rank}(X)$ problem can be expressed as a minimum cardinality problem on the coefficients of the characteristic polynomial. With

$$\begin{array}{rll} \mathbf{Rank}(X) = & \min & v_i \\ & \text{s.t.} & \sigma_i(X)(v_i - 1) = 0 \\ & & v_i \geq 0, \quad \text{for } i = 1, ..., n, \end{array}$$

we enforce the remaining constraints and the **Min Rank**(X) problem becomes:

$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^{n} v_i \\ \text{s.t.} & (v_i-1)\sigma_i(X)=0 \\ & v_i \geq 0, \quad \text{for } i=1,...,n \\ & Tr\,(A_iX)=b_i, \quad \text{for } i=1,...,p \\ & d_I(X)\geq 0, \quad \text{for } I\subset\{1,\ldots,n\},\end{array}$$

where $d_I(X)$ is the principal minor with index set $I \subset \{1,\ldots,n\}$. This is a semialgebraic program in the coefficients of the matrix X. ■

These two results together with the results cited in §3.2.2 show that the two problems considered are equivalent to very large scale semidefinite programs.

3.2.4 Semidefinite relaxations

In practice, the exact representations obtained in the last section can be exponentially large and in general, we cannot expect these problems to be tractable. Hence, the central contribution of these representations is not to reduce the complexity of these problems, but to provide a sequence of successively sharper relaxations covering the entire complexity spectrum, thus allowing the complexity/sharpness tradeoff to be tuned. This is what we intend to describe in this section.

We begin by recalling the construction of moment matrices as detailed in Curto & Fialkow (2000), Lasserre (2001) and Lasserre (2002). Again, we let

$$y_x=\left(1,x_1,...,x_n,x_1^2,x_1x_2,...,x_1x_n,x_2^2,x_2x_3,...,x_n^2,...,x_1^m,...,x_n^m\right)$$

be the vector of all monomials in $\mathbf{R}[x_1,...,x_n]$, up to degree m, listed in increasing graded lexicographic order. We note $s(m)$ the size of the vector y_x. Let $y\in\mathbf{R}^{s(2m)}$ be the vector of moments (indexed according to y_x) of some probability measure μ with support $K=\{x\in\mathbf{R}^n : g(x)\geq 0\}$, we note $M_m(y)\in\mathbf{S}^{s(m)}$, for the moment matrix defined by

$$M_m(y)_{i,j} = \int_K (y_x)_i (y_x)_j \, \mu(dx), \quad \text{for } i,j = 1,...,s(m) \tag{3.31}$$

i.e. the (symmetric) matrix of moments with rows and columns indexed as in y_x. We note $\beta(i)$ the exponent of the monomial $(y_x)_i$ and conversely, we note $i(\beta)$ the index of the monomial x^β in y_x. For a given moment vector $y \in \mathbf{R}^{s(m)}$ ordered as in (3.2.4), the first row and columns of the matrix $M_m(y)$ are then equal to y. The rest of the matrix is then constructed following:

$$M_m(y)_{i,j} = y_{\alpha+\beta} \text{ if } M_m(y)_{1,i} = y_\alpha \text{ and } M_m(y)_{j,1} = y_\beta.$$

Similarly, let $g(x) \in \mathbf{R}[x_1, ..., x_n]$, we derive the moment matrix for the measure $g(x)\mu(dx)$ on K (called the localizing matrix), noted $M_m(gy) \in \mathbf{S}^{s(m)}$, from the matrix of moments $M_m(y)$ by:

$$M_m(gy)_{i,j} = \int_K (y_x)_i (y_x)_j \, g(x)\mu(dx) \tag{3.32}$$

for $i,j = 1,...,s(m)$. The coefficients of the matrix $M_m(gy)$ are then given by:

$$M_m(gy)_{i,j} = \sum_\alpha g_\alpha M_m(y)_{i(\beta(i)+\beta(j)+\alpha)} \tag{3.33}$$

We can remark as in Lasserre (2001) that if the measure μ has its support included in $K = \{x \in \mathbf{R}^n : g(x) \geq 0\}$, then for all coefficient vectors $v \in \mathbf{R}^{s(m)}$:

$$\langle v, M_m(gy)v \rangle = \int_K v(x)^2 g(x)\mu(dx) \geq 0$$

hence $M_m(gy) \succeq 0$.

In dimension one, for a given vector $y \in \mathbf{R}^{s(2m)}$, $M_m(y) \succeq 0$ (which is a LMI) is also a sufficient condition in order for y to the moment sequence of a probability measure. In $\mathbf{R}^n$, this equivalence does not hold in general. The compact semialgebraic case is called the K-moment problem and is dual to the compact SOS problem in (3.28). Following Lasserre (2001), we now exploit this duality to compute a sequence of semidefinite relaxations for the $\mathbf{Min\,Card}(x)$ and $\mathbf{Min\,Rank}(X)$ problems.

The MinCard(x) problem

In §3.2.3, we saw that the optimum value of the $\mathbf{Min\,Card}(x)$ problem can be computed as the optimum value of the semialgebraic program:

$$\begin{array}{ll}\text{min.} & \sum_{i=1}^{n} v_i \\ \text{s.t.} & (v_i-1)x_i = 0 \\ & v_i \geq 0, \quad \text{for } i=1,...,n \\ & a_j^T x \geq b_j, \quad \text{for } j=1,...,m.\end{array} \tag{3.34}$$

As in Lasserre (2001), to ensure compactness, we impose the additional constraint $x_1^2 + ... + x_n^2 \leq \alpha$ for some constant $\alpha > 1$. It is easy to check that the program above, together with this additional bound on the feasible set, satisfies the constraints qualification assumption (3.26). For $N \geq 1$, a lower bound l_N on the optimal value of the above problem is then computed as:

$$\begin{array}{lll} l_N := & \inf & \sum_{i=1}^{n} y_i \\ & \text{s.t.} & M_N(y) \succeq 0 \\ & & M_{N-1}\left(x_i(v_i-1)y\right) = 0 \\ & & M_{N-1}\left((\alpha - x^T x - v^T v)y\right) \succeq 0 \\ & & M_{N-1}\left(v_i y\right) \succeq 0, \quad \text{for } i=1,...,n \\ & & M_{N-1}\left(\left(a_j^T x - b_j\right)y\right) \succeq 0, \quad \text{for } j=1,...,m,\end{array} \tag{3.35}$$

in the variable $y \in \mathbf{R}^{s(2n)}$. Theorem 3.2 in Lasserre (2002) then states that there exists some N^{opt} such that

$$l_N = \mathbf{Min\,Card}(x), \quad \text{for all } N \geq N^{\text{opt}},$$

and the optimum is achieved whenever the rank of the matrices $M_N((...)\,y)$ stabilizes.

The MinRank(X) problem

In §3.2.3, for $X \in \mathbf{S}^n$, we saw that the optimum of the $\mathbf{Min\,Rank}(X)$ problem can be computed as the optimum value of the semialgebraic program:

$$\begin{array}{ll}\text{min.} & \sum_{i=1}^n v_i \\ \text{s.t.} & (v_i - 1)\sigma_i(X) = 0 \\ & v_i \geq 0, \quad \text{for } i = 1, ..., n \\ & Tr\,(A_j X) = b_j, \quad \text{for } j = 1, ..., p \\ & d_I(X) \geq 0, \quad I \subset [1, n],\end{array}$$

To further simplify this program, we can substitute to the 2^n constraints on the principal minors a more economical semialgebraic constraint. The modified program then reads:

$$\begin{array}{ll}\text{min.} & \sum_{i=1}^n v_i \\ \text{s.t.} & (v_i - 1)\sigma_i(X) = 0 \\ & v_i \geq 0, \quad \text{for } i = 1, ..., n \\ & Tr\,(A_j X) = b_j, \quad j = 1, ..., p \\ & u^T X u \geq 0, \quad u \in \mathbf{R}^n,\end{array}$$

and again, to ensure compactness, we impose $X^T X + v^T v + u^T u \leq \alpha$ for some constant $\alpha > 1$. If we set the variable $x = (u, X, v)$, for N sufficiently large, a lower bound l_N

on the optimal value of the above problem is computed as:

$$
\begin{array}{lll}
l_N := & \inf & \sum_{i=1}^{n} u_i \\
& \text{s.t.} & M_N(y) \succeq 0 \\
& & M_{N-\lceil \frac{i+1}{2} \rceil}\left((v_i - 1)\sigma_i(X) y\right) = 0 \\
& & M_{N-1}\left(v_i y\right) \succeq 0, \quad i = 1, \ldots, n \\
& & M_{N-1}\left(\left(Tr\left(A_j X\right) - b_j\right) y\right) = 0, \quad j = 1, ..., p \\
& & M_{N-1}\left(\left(\alpha - X^T X + v^T v + u^T u\right) y\right) \succeq 0 \\
& & M_{N-2}\left(\left(u^T X u\right) y\right) \succeq 0,
\end{array}
\tag{3.36}
$$

in the variable $y \in \mathbf{R}^{s(2n)}$, where the matrices $M(q(x)y)$ are computed as in (3.33). Theorem 3.2 in Lasserre (2002) then states that there exists some N^{opt} such that

$$
l_N = \mathbf{Min\,Rank}(X), \quad \text{for all } N \geq N^{\text{opt}},
$$

and the optimum is reached whenever the rank of the matrices $M_N((...)\,y)$ stabilizes. Alternatively, one could use the fact that if we note $\chi_t(X)$ the characteristic polynomial of X, then $X \succeq 0$ is equivalent to $\chi_{-t^2}(X)$ being SOS as a univariate polynomial in t.

Convex envelope

Suppose that instead of having only one $\mathbf{Min\,Card}(x)$ or $\mathbf{Min\,Rank}(X)$ problem to solve, we need to solve a (long) sequence of these problems with only some variation in the constraints. Here, instead of computing an exact relaxation for every instance of the problem, we are interested in finding an efficient heuristic method for approximating the solution to all the problems to be solved. The complexity of the first "bound design" program will be high, but that of the subsequent programs will then be much lower. The heuristics in Boyd et al. (2000) replaced the $\mathbf{Card}(x)$ (resp. $\mathbf{Rank}(X)$) functions by their convex envelope on the sets $0 \leq x \leq 1$ (resp.

$0 \preceq X \preceq I$), i.e. the largest convex function $f(x)$ such that $f(x) \leq \textbf{Card}(x)$ if $0 \leq x \leq 1$ (resp. $f(X) \leq \textbf{Rank}(X)$ if $0 \preceq X \preceq I$). In this section, we extend these bounds to semialgebraic sets with more complex shapes.

Of course, a function and its convex envelope share the same global minimum, so solving for this optimal lower bound is at least as hard as finding the global minimum. Here however we look for a convex lower bound for the problem in (3.29) inside the set of polynomials of degree at most d. This becomes a semialgebraic program:

$$\begin{array}{ll} \text{maximize} & \int_{[0,1]^n} p(x)dx \\ \text{subject to} & t - p(x) \geq 0 \text{ on } K \\ & \sum_{i=1}^n v_i - p(x) \geq 0 \text{ on } K \\ & u^T \nabla^2 p(x) u \geq 0 \text{ on } \|u\|^2 = 1 \end{array} \tag{3.37}$$

in the coefficients p_α of the polynomial $p(x) \in \mathbf{R}_d[x_1, ..., x_n]$, where K is the compact semialgebraic set given by:

$$K = (v, x) \in \mathbf{R}^{2n} : \begin{array}{l} (v_i - 1)\, x_i = 0 \\ Ax - b \geq 0 \\ x, v \geq 0. \end{array} \tag{3.38}$$

Again, this can be cast as a LMI using the technique in Lasserre (2002). We notice that the l_1 heuristic is a particular case when the constraint $Ax - b \geq 0$ is dropped.

3.2.5 Complexity

Of course, the two semidefinite programs detailed in the last section are far from tractable if the dimension n and the relaxation order N grow beyond textbook example sizes. The $\textbf{Min}\,\textbf{Card}(x)$ problem is equivalent to solving 2^n linear programs, so it is right to ask whether the programs above provide any benefit over, for example, branch-and-bound methods?

Even if these two methodologies have similar worst-case complexities, the semidefinite relaxations in (3.35) and (3.36) do sometimes produce the global optimum for low order N (see Lasserre (2002)) and because the objective is integer valued here, they only need to be solved up to an absolute precision of 1/2. We quickly detail below some other possible simplifications.

But the results above have to be considered first as representations, providing an insight of the relative complexity of minimum cardinality problems versus that of tractable convex optimization problems.

Structure, sparsity and symmetry

The first element that can be used to simplify the programs in (3.35) and (3.36) is *structure*. In Lasserre (2002) for example, the constraints $x \in \{-1, 1\}$, that translate into $x^2 - x = 0$ and $M_m((x^2 - x)\, y) = 0$ also imply that the variables y_α, for $\alpha \in \mathbf{Z}_+^n$, can be replaced by $y_{\mathbf{Card}(\alpha)}$ in the program. Secondly, if the constraints in (3.35) and (3.36) include some *symmetry*, we could use the results in Gatermann (2000) and Gatermann & Parillo (2002) to preprocess and simplify the original semialgebraic program. The simplest example of these symmetries is of course when the constraints are invariant with respect to a change of basis, in which case $\mathbf{Min\,Rank}(X)$ reduces to a $\mathbf{Min\,Card}(x)$ problem and as in Gatermann & Parillo (2002) p. 31, the asymptotic complexity goes from being exponential in n to being exponential in $\sqrt{n}$.

In general, the complexity of algorithms in semialgebraic geometry grows at least exponentially with the dimension. Efficiency is then very often measured by the ability of a method to maintain *sparsity*. Some results on Newton polytopes and SOS polynomials can then be used to efficiently handle sparsity in the SOS representations. In particular, a result of Reznick (1978) shows that if $p(x), h_i(x) \in \mathbf{R}[x]$, for $i = 1, ..., r$, with $p(x) = \sum_{i=1}^r h_i^2(x)$, then $C(h_i) \subseteq \frac{1}{2}C(p)$. That result however does not hold as is for the representation in (3.27).

Finally, a lower bound on the optimal value can be obtained by simply dropping some of the constraints in (3.35) and (3.36).

3.2.6 Numerical example

In this section, we present a brief example of how the relaxation technique from (3.2.4) can be used to "spice up" the l_1 norm heuristic on a sample **Min Card**(x) problem. We look at the following problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{Card}(x) \\ \text{subject to} & Ax \geq b \\ & 0 \leq x \leq \mathbf{1}, \end{array} \tag{3.39}$$

where

$$A = \begin{bmatrix} 0.922 & 0.152 & 0.606 & 0.0513 & 0.318 \\ 0.232 & 0.988 & 0.393 & 0.223 & 0.547 \\ 0.236 & 0.987 & 0.821 & 0.111 & 0.158 \\ 0.998 & 0.279 & 0.899 & 0.258 & 0.39 \\ 0.0877 & 0.206 & 0.692 & 0.657 & 0.00404 \end{bmatrix}$$

and

$$b = \begin{bmatrix} 1.18 \\ -1.2 \\ 0.715 \\ 0.94 \\ 1.48 \end{bmatrix}.$$

In this case the classic l_1 relaxation gives a lower bound on the optimal cardinality of 2.9193 with a solution of cardinal 4, while the order 2 SDP relaxation detailed in (3.2.4) gives a lower bound of 3.7508 with a solution of cardinal 4 hence produces a globally optimal solution. However, the computing time (20 sec.) is far from being competitive with that of $MILP$ packages (MOSEK took less than a second to solve globally this example).

One of the central contributions of semidefinite programs to the optimization toolbox is their ability to efficiently solve a wide class of convex eigenvalue problems.

In this chapter, we have illustrated how the method described in Lasserre (2001) for solving semialgebraic programs, by lifting them to semidefinite programs, can also be used to represent some semialgebraic eigenvalue problems and convex envelope relaxations. This contribution is centered around semidefinite representations and the insight they can provide on the theoretical complexity of these problems. Wether or not they also improve the practical complexity of computing relaxations to these problems remains to be explored.

3.3 Generalized moment problems

We focus here on generalized moment problems. Using results by Berg & Maserick (1984), Putinar & Vasilescu (1999) and Lasserre (2001) on harmonic analysis on semigroups, the $\mathbb{K}$-moment problem and its applications to optimization, we revisit the problem of computing upper and lower bounds on the price of a European basket call option, given prices on other similar baskets. We formulate necessary and sufficient conditions for the absence of static (or buy-and-hold) arbitrage between basket straddles, hence on basket calls and puts. These allow us to produce a sequence of increasingly tight relaxations of this problem, in the form of increasingly large semidefinite programs.

3.3.1 Introduction

As in §2.1.1, we let $p \in \mathbf{R}_+^m$, $K \in \mathbf{R}^{m+1}$, $w_i \in \mathbf{R}^n$, $i = 0, \ldots, m$ and we consider the problem of computing upper and lower bounds on the price of an European basket call option with strike K_0 and weight vector w_0:

$$\begin{array}{ll} \text{maximize/minimize} & p_0 := \mathbf{E}_\nu (w_0^T x - K_0)_+ \\ \text{subject to} & \mathbf{E}_\nu (w_i^T x - K_i)_+ = p_i, \quad i = 1, \ldots, m, \end{array} \tag{3.40}$$

with respect to all probability measures ν on the asset prices $x \in \mathbf{R}_+^n$, consistent with the (given) set of observed prices p_i of options on other baskets.

We implicitly assume that all the options have the same maturity, and that, without loss of generality, the risk-free interest rate is zero (we compare prices in the forward market). We seek non-parametric bounds, *i.e.* we do not assume any specific model for the underlying asset prices, our only assumption is the absence of a static arbitrage today (*i.e.* the absence of an arbitrage that only requires trading today and at maturity).

Here, we interpret (3.40) as a *generalized moment problem*. This approach was successfully used in Bertsimas & Popescu (2002) to get tractable bounds in dimension one and to show the NP-hardness of the multivariate problem (3.40). NP-hardness means that we have no chance of finding a direct and efficient method for detecting all arbitrage opportunities, our objective here instead is to look for a sequence of successively tighter price bounds. This would mean that outlandish arbitrage opportunities can be detected at little numerical cost while detecting finer price discrepancies has a higher theoretical complexity.

The motivation for this work lies in the fact that most of the interest rate market volatility information is concentrated in caps and swaptions, which can be interpreted as basket options. Most interest rate models are flexible enough to calibrate on a vast range of market conditions, but the question still arises of how to interpret cases where an exact fit to the market data is impossible. This can be caused by one of two scenarios. The first one occurs when the market data is arbitrage free but the model dynamics are not rich enough to calibrate on them. In the second one, the market data is simply not arbitrage free. This can be a simple technical discrepancy, *i.e.* a missing data point or an outlier in the data set, or it can be an actual static arbitrage opportunity. It is then important to be able to detect where the mispricing is occurring and how to correct it.

We use here recent results on multivariate moment problems (see Schmüdgen (1991), Putinar & Vasilescu (1999) or Curto & Fialkow (2000)), semidefinite programming (see Nesterov & Nemirovskii (1994), Vandenberghe & Boyd (1996) and Nesterov (2000)) and harmonic analysis on semigroups (see Berg, Christensen & Ressel (1984) and Roman (2003)) allow us to derive static arbitrage price bounds on a set of products linked by a semigroup structure. The resulting constraints can be formulated as successively tighter linear matrix inequalities, hence we can compute increasingly sharp bounds on the solution to problem (3.40) as solutions of increasingly large semidefinite programs (linear programs on the cone of positive semidefinite matrices). Semidefinite programming has been the object of intensive research since the seminal work of Nesterov & Nemirovskii (1994) and several numerical packages (see for example SEDUMI by Sturm (1999)) are now available to solve these problems very efficiently.

The core of our argument is to substitute to the classical duality between the cones of probability measures and positive portfolios, the conic duality between positive definite functions on one hand and sums of squares on the other. These last two cones have the advantage of being numerically tractable and lead to exploitable formulations of the static portfolio super/sub-replication problems.

In the previous chapter, we focused on the interpretations of problem (3.40) as an *integral transform inversion problem* or a *linear semi-infinite program*, *i.e.* a linear program with a finite number of linear constraints on an infinite dimensional variable, and used the related theories to compute closed-form solutions for some particular cases and a linear programming relaxation for the general case.

This section is organized as follows. In section two, we describe the static market structure and start with a brief introduction on harmonic analysis on semigroups. Based on these results, we then derive necessary and sufficient conditions for the absence of arbitrage in the static market, formulated as semidefinite programs. In section three, we describe the conic duality between positive definite functions and sums of squares and use it to show how a super/sub-replicating portfolio can be constructed from the solution to the programs of the preceding section. Finally, in section four we discuss the numerical complexity of the arbitrage conditions and describe in details two applications to spread and FOREX option pricing.

3.3.2 Static arbitrage constraints

Market structure

We work in a one period framework and suppose that the market is composed of cash and n underlying assets x_i for $i = 1, \ldots, n$ with $x \in \mathbf{R}^n_+$. We suppose that the forward prices of the assets are known and given by p_i, for $i = 1, \ldots, n$, hence w_i is the Euclidean basis and $K_i = 0$ for $i = 1, \ldots, n$. In addition to these basic products, there are $m+1$ basket *straddles* on the assets x, with payoff given by $|w_{n+i}^T x - K_{n+i}|$,

$i = 1, \ldots, m$. Because a straddle is obtained as the sum of a call and a put, we get the market price of straddles from those of basket calls and forward contracts by call-put parity.

These payoff functions will be denoted by e_i, for $i = 0, \ldots, m+n$, with $e_i(x) = x_i$ for $i = 1, \ldots, n$ and $e_{(n+j)}(x) = |w_j^T x - K_j|$ for $j = 0, \ldots, m$. In what follows, we will focus on the Abelian (commutative) semigroup $(\mathbb{S}, \cdot)$ generated by the payoffs $e_i(x)$ for $i = 0, \ldots, m+n$, the cash $1_{\mathbb{S}}$ and their products

$$\mathbb{S} = \{1, x_1, \ldots, |w_m^T x - K_m|, x_1^2, \ldots, x_i |w_j^T x - K_j|, \ldots \}$$

In this one period setting, we will look for conditions that guarantee the absence of *static arbitrages*, *i.e.* arbitrage opportunities that only involve trading today and at maturity, assuming that there are no transaction costs.

Harmonic analysis on semigroups

We start by a brief introduction on harmonic analysis on semigroups, for a complete treatment see Berg et al. (1984) and the references therein. Unless otherwise specified, all measures are supposed to be positive.

Definition 18 *A function $\rho : \mathbb{S} \to \mathbf{R}$ is called a semicharacter iff it satisfies $\rho(st) = \rho(s)\rho(t)$ for all $s, t \in \mathbb{S}$ and $\rho(1_{\mathbb{S}}) = 1$.*

In Berg et al. (1984) an involution operation is defined on the semigroup $(\mathbb{S}, \cdot)$, here and in the rest of the paper we suppose that involution to be the identity, which means in particular that we take all semicharacters to be real valued. The dual semigroup of $\mathbb{S}$, *i.e.* the set of semicharacters on $\mathbb{S}$ is called $\mathbb{S}^*$. In this context, we call a function $f : \mathbb{S} \to \mathbf{R}$ a *moment function on* $\mathbb{S}$ iff $f(1_{\mathbb{S}}) = 1$ and f can be represented as:

$$f(s) = \int_{\mathbb{S}^*} \rho(s) d\nu(\rho), \quad \text{for all } s \in \mathbb{S}, \tag{3.41}$$

where ν is a Radon measure on $\mathbb{S}^*$.

When $\mathbb{S}$ is the semigroup defined in (3.3.2) as an enlargement of the semigroup of monomials on $\mathbf{R}^n$, its dual $\mathbb{S}^*$ is the set of applications $\rho_x : \mathbb{S} \to \mathbf{R}$ such that $\rho_x(s) = s(x)$ for all $s \in \mathbb{S}$ and all $x \in \mathbf{R}^n$. The measure ν is then assimilated to a probability measure on $\mathbf{R}^n$ and the representation above becomes:

$$f(s) = \mathbf{E}_\nu\left[s(x)\right], \quad \text{for all } s \in \mathbb{S}. \tag{3.42}$$

Our objective below is to find tractable conditions for a set of prices $p_0, \ldots, p_{n+m}$ to be represented as $\mathbf{E}_\nu\left[|w_i^T x - K_i|\right] = p_i$ for $i = 0, \ldots, n+m$ and some positive measure ν.

The compact case

In this section we assume the asset distribution has a compact support K. We treat the compact case independently as it is rather simple yet captures many of the key features of the general result. We begin by a few definitions along the lines of Berg & Maserick (1984) and Berg et al. (1984). An *absolute value* on $\mathbb{S}$ is a function $|\cdot| : \mathbb{S} \to \mathbf{R}_+$ satisfying

$$|s^2| \leq |s|^2, \quad \text{for all } s \in \mathbb{S}$$

and

$$|1_\mathbb{S}| \geq 1.$$

A function $f : \mathbb{S} \to \mathbf{R}$ is said to be bounded with respect to an absolute value $|\cdot|$ iff there exists some $M > 0$ such that

$$|f(s)| \leq M|s|, \quad \text{for all } s \in \mathbb{S}.$$

Furthermore, f is called *exponentially bounded* iff f is bounded with respect to some absolute value. Remark that if the measure ν in (3.41) has its support contained in the compact K then the moment function $f(s) = \int_{\mathbb{S}^*} \rho(s) d\nu(\rho)$ is bounded with respect to the following absolute value:

$$|s|_K = \sup_{\rho \in K} \rho(s)$$

for $s \in \mathbb{S}$.

Finally, we define the key notion of *positive definite function.*

Definition 19 *A function $f : \mathbb{S} \to \mathbf{R}$ is called positive semidefinite iff for all finite families $\{s_i\}$ of elements of $\mathbb{S}$, the matrix with coefficients $f(s_i s_j)$ is positive semidefinite.*

We remark that moment functions are necessarily positive semidefinite. Necessary and sufficient conditions on $f(s)$ for the existence of a measure ν in (3.42) were derived in Henkin & Shananin (1990), they were however difficult to exploit numerically. Here, based on the results in Berg et al. (1984), Putinar & Vasilescu (1999) and Roman (2003), we look for exploitable conditions for representation (3.42) to hold.

Let α be an absolute value, the central result in Berg et al. (1984, Th. 2.6) states that the set of α-bounded positive semidefinite functions $f : \mathbb{S} \to \mathbf{R}$ such that $f(1_{\mathbb{S}}) = 1$ is a Bauer simplex whose extreme points are given by the set of α-bounded semicharacters. Hence a function f is positive semidefinite and exponentially bounded if and only if it can be represented as $f(s) = \int_{\mathbb{S}^*} \rho d\nu(\rho)$ with the support of ν included in some compact subset of $\mathbb{S}^*$.

Based on these results, we derive below a set of tractable necessary and sufficient conditions allowing a function f to be represented as in (3.42). For s, u in $\mathbb{S}$, we denote by E_s the shift operator such that for $f : \mathbb{S} \to \mathbf{R}$, we have $E_s(f(u)) = f(su)$ and we let $\mathcal{E}$ be the commutative algebra generated by the shift operators on $\mathbb{S}$. Finally, we let $\beta = \sup_{x \in K}\{\sum_{i=0}^{n+m} e_i(x)\}$.

Theorem 20 *Suppose the asset distribution has compact support K and $\mathbb{S}$ is the payoff semigroup defined in (3.3.2), with β is defined as above. A function $f(s) : \mathbb{S} \to$*

$\mathbf{R}$ *can be represented as*

$$f(s) = \mathbf{E}_\nu[s(x)], \quad \text{for all } s \in \mathbb{S}, \tag{3.43}$$

for some measure ν on K, and satisfies the price constraints in (3.40) if and only if:

(i) *f is positive semidefinite,*

(ii) *$E_{e_i} f$ is positive semidefinite for $i = 0, \ldots, n+m$,*

(iii) *$\left(\beta I - \sum_{i=0}^{n+m} E_{e_i}\right) f$ is positive semidefinite,*

(iv) *$f(e_i) = p_i$ for $i = 1, \ldots, n+m$.*

Furthermore, for each function f satisfying conditions (i) to (iv), the measure ν in representation (3.43) is unique.

Proof. The family of shift operators $\tau = \{\{E_{e_i}\}_{i=0,\ldots,n+m}, \left(\beta I - \sum_{i=0}^{n+m} E_{e_i}\right)\} \subset \mathcal{E}$ is such that $I - T \in \operatorname{span}^+ \tau$ for each $T \in \tau$ and $\operatorname{span} \tau = \mathcal{E}$, hence τ is linearly admissible in the sense of Berg & Maserick (1984, Corollary 2.5) or Maserick (1977), which states that (ii) and (iii) are equivalent to f being τ-positive. Then, Maserick (1977, Th. 2.1) means that f is τ-positive if and only if there is a measure ν such that $f(s) = \int_{\mathbb{S}^*} \rho(s) d\nu(\rho)$, whose support is a compact subset of the τ-positive semicharacters. This means in particular that for a semicharacter $\rho_x \in \operatorname{supp}(\nu)$ we must have $\rho_x(e_i) \geq 0$, for $i = 1, \ldots, n$ hence $x \geq 0$. The set of τ-positive semicharacters is then included in the nonnegative orthant and includes both the simplex $\{x \geq 0 : \|x\|_1 \leq \beta\}$ and K, hence f being τ-positive is equivalent to f admitting a representation of the form $f(s) = \mathbf{E}_\nu[s(x)]$, for all $s \in \mathbb{S}$ with ν having a compact support $K \subset \mathbf{R}^n_+$. ■

The unbounded case

The conditions derived in the last part do not describe all possible arbitrage free prices as they cannot account for unbounded asset distributions. Here, we use results from

Putinar & Vasilescu (1999) and Roman (2003) to derive intrinsic characterizations of viable multivariate straddle prices.

We denote by $\mathcal{A}(\mathbb{S})$ the $\mathbf{R}$-algebra generated by the functions $\chi_s : \mathbb{S}^* \to \mathbf{R}$ such that $\chi_s(\rho) = \rho(s)$ for all $s \in \mathbb{S}$. By construction, $\chi_s(\rho) = E_s\rho(1_{\mathbb{S}})$, and for a polynomial $p \in \mathcal{A}(\mathbb{S})$ with $p = \sum_k q_k \chi_{g_k}$ and for $\rho \in \mathbb{S}^*$ we have $p\rho(s) = \sum_k q_k \rho(s g_k)$ for all $s \in \mathbb{S}$. When $\mathbb{S}$ is the payoff semigroup defined in (3.3.2), we naturally have $\chi_s(\rho_x) = s(x)$, for all $x \in \mathbf{R}^n$, $s \in \mathbb{S}$ and $\rho \in \mathbb{S}^*$.

We shall denote by $\mathcal{A}_\theta(\mathbb{S})$ the $\mathbf{R}$-algebra generated by $\mathcal{A}(\mathbb{S})$ and θ where

$$\theta(\rho) = \left(1 + \sum_{i=0}^{m+n} \chi_{e_i^2}(\rho)\right)^{-1}, \quad \text{for all } \rho \in \mathbb{S}^*, \tag{3.44}$$

we also denote by $\mathcal{A}(\mathbb{S}, y)$ the algebra generated by $\mathcal{A}(\mathbb{S})$ and $\mathbf{R}[y]$. We first simplify the equality constraints on 2^n variables in Roman (2003, Th. A) to recover an additive formulation as in Putinar & Vasilescu (1999). We begin by proving the following lemma.

Lemma 21 *The kernel of the algebra homomorphism* Φ*:*

$$\begin{array}{lcl} \mathcal{A}(\mathbb{S}, y) & \to & \mathcal{A}_\theta(\mathbb{S}) \\ p(\rho, y) & \mapsto & \Phi p = p(\rho, \theta(\rho)) \end{array} \tag{3.45}$$

is the ideal generated by $\sigma \in \mathcal{A}(\mathbb{S}, y)$ *such that* $\sigma(\rho, y) = y(1 + \sum_{i=1}^{m+n+1} \chi_{e_i^2}(\rho)) - 1$.

Proof. We adapt the proof of Putinar & Vasilescu (1999, lemma 2.3) and let $p \in \mathcal{A}(\mathbb{S}, y)$ be such that $p(\rho, \theta(\rho)) = 0$, we write $p(\rho, y) = \sum_k q_k(\rho) y^k$ with $q_k \in \mathcal{A}(\mathbb{S})$. We have:

$$\begin{aligned} p(\rho, y) &= p(\rho, y) - p(\rho, \theta(\rho)) = \textstyle\sum_{k>0} q_k(\rho)(y^k - \theta(\rho)^k) \\ &= (y - \theta(\rho)) l(\rho, y, (\theta(\rho)), \end{aligned}$$

where l is a polynomial. Let $\kappa = max\{k : q_k \neq 0\}$ and

$$\tau(\rho) = \left(1 + \sum_{i=1}^{m+n+1} \chi_{e_i^2}(\rho)\right)^{\kappa}, \quad \text{for all } \rho \in \mathbb{S}^*,$$

we then have

$$\tau(\rho)p(\rho, y) = \sigma(\rho, y)r(\rho, y), \tag{3.46}$$

with $r(\rho, y) \in \mathcal{A}(\mathbb{S}, y)$. The case $\kappa = 0$ is trivial hence we can assume $\kappa \neq 0$. Using the fact that the polynomials $\tau(z)$ and $\sigma(z)$ have no common zeroes in $\mathbf{C}^{m+n+2}[\text{z}]$, Hilbert's *Nullstellensatz* (see Bochnak, Coste & Roy (1998) for example) states that there must be $\tilde{\tau}, \tilde{\sigma} \in \mathbf{C}^{m+n+2}[z]$ such that

$$\tau\tilde{\tau} + \sigma\tilde{\sigma} = 1.$$

Multiplying this last identity by p yields, together with (3.46):

$$p = \sigma(r\tilde{\tau} + p\tilde{\sigma})$$

hence the desired result. $\blacksquare$

The next proposition is adapted from the dimensional extension method in Putinar & Vasilescu (1999, Th. 2.5) and Roman (2003, Th. 4), our objective is to replace the exponential number of equality constraints in Roman (2003, Th. A) with an additive formulation as in Putinar & Vasilescu (1999). Remember that the function $\theta(\rho)$ is defined as in (3.44) and $\mathcal{A}_\theta(\mathbb{S})$ is the $\mathbf{R}$-algebra generated by $\mathcal{A}(\mathbb{S})$ and θ.

Proposition 22 *With $\mathbb{S}$ being the payoff semigroup defined in (3.3.2), let Λ be a positive semidefinite linear form on $\mathcal{A}_\theta(\mathbb{S})$ such that $\Lambda\left(x_i r^2\right) \geq 0$ for all $r \in \mathcal{A}_\theta(\mathbb{S})$ and $i = 1, \ldots, n$, then Λ has a unique representing measure ν with support in $\mathbf{R}_+^n$ and $\mathcal{A}_\theta(\mathbb{S})$ is dense in $L^2(\nu)$.*

Proof. We recall that the linear form Λ is positive semidefinite iff $\Lambda(r^2) \geq$, for all

$r \in \mathcal{A}_\theta(\mathbb{S})$. As in Roman (2003), we define a bilinear form on $r \in \mathcal{A}_\theta(\mathbb{S})$ by:

$$\langle r_1, r_2 \rangle := \Lambda(r_1 r_2), \quad \text{for all } r_1, r_2 \in \mathcal{A}_\theta(\mathbb{S})$$

We let $\mathcal{N}$ be the set $\{r \in \mathcal{A}_\theta(\mathbb{S}) : \Lambda(r^2) = 0\}$. The bilinear form above then defines a scalar product on $\mathcal{A}_\theta(\mathbb{S})/\mathcal{N}$, and we denote by $\mathcal{H}$ the completion of this space. We define in $\mathcal{H}$ the operators:

$$T_i(r + \mathcal{N}) = \chi_{e_i} r + \mathcal{N}, \quad \text{for all } r \in \mathcal{A}_\theta(\mathbb{S})/\mathcal{N} \text{ and } i = 0, \ldots, n+m,$$

which are symmetric and densely defined in $\mathcal{H}$. We also define the operator $(D(B), B)$ by:

$$D(B) = \mathcal{A}_\theta(\mathbb{S})/\mathcal{N} \text{ and } B = \sum_{i=0}^{m+n} T_i^2.$$

The operator B is positive as a sum of squares of operators and, by construction, the domain $D(B)$ is dense in $\mathcal{H}$ and invariant by B. Let $\tau = \sum_{i=0}^{m+n} \chi_{e_i^2}(\rho)$ and $r \in \mathcal{A}_\theta(\mathbb{S})/\mathcal{N}$, then $u = r\theta$ is such that $(1+\tau)u = r$, hence the operator $I + B$ is bijective on $D(B)$. This means that B satisfies the hypothesis of Putinar & Vasilescu (1999, Lemma 2.2) and is essentially self-adjoint. Roman (2003, Prop. 1) then implies that the operators T_i for $i = 0, \ldots, n+m$ are essentially normal and that their canonical closures commute, meaning that there exists a common spectral measure H for the operators $\bar{T}_i$ for $i = 0, \ldots, n+m$. With $T = (T_i)_{i=0,\ldots,n+m}$ and $r \in \mathcal{A}_\theta(\mathbb{S})/\mathcal{N}$, we define the operator $r(T)$ by:

$$\begin{array}{lcl} \mathcal{A}_\theta(\mathbb{S})/\mathcal{N} & \to & \mathcal{A}_\theta(\mathbb{S})/\mathcal{N} \\ w + \mathcal{N} & \mapsto & r(T)(w + \mathcal{N}) = rw + \mathcal{N}. \end{array} \tag{3.47}$$

With $\gamma(x) = \left(\sum_{i=0}^{n+m} x_i^2\right)^{-1}$, there is an element q of $\mathbf{R}_\gamma[x]$, the $\mathbf{R}$-algebra generated by $\mathbf{R}[x]$ and $\gamma(x)$ such that $r(\rho) = q\left((\chi_{e_i})_{i=0,\ldots,n+m}(\rho), \theta(\rho)\right)$ for all $\rho \in \mathbb{S}^*$. We then have:

$$\Lambda(r) = \langle r(T)1, 1\rangle = \langle q(\bar{T})1, 1\rangle = \int_{\mathbf{R}^{n+m+1}} q(x) dH_{1+\mathcal{N},1+\mathcal{N}}(x),$$

The homomorphism f:

$$\begin{array}{rcl} \mathbf{R}_\gamma[x] & \rightarrow & \mathcal{A}_\theta(\mathbb{S}) \\ p(x) & \mapsto & f(p) = p((\chi_{e_i}(\rho))_{i=0,\dots,n+m}, \theta(\rho)) \end{array} \tag{3.48}$$

satisfies the hypothesis of Roman (2003, Lemma 2) hence there is a (positive) Radon measure ν on such that:

$$\Lambda(r) = \int_{\mathbb{S}^*} r(\rho) d\nu(\rho),$$

which, if $\mathbb{S}$ is defined as in section (3.3.2), is also:

$$\Lambda(r) = \int_{\mathbf{R}^n} r(x) d\nu(x).$$

Uniqueness and density follow from the argument in Roman (2003). Now, because the operators T_i for $i = 1, \dots, n$ are essentially self-adjoint with $\Lambda\left(x_i r^2\right) \geq 0$ for $r \in \mathcal{A}_\theta(\mathbb{S})$ and $i = 1, \dots, n$, we know that the T_i are positive for all i. The spectral measure F_i of T_i is given by $F_i(X) = H\left(\bar{T}_i^{-1}(X)\right)$ for all Borel sets $X \subset \mathbf{R}$ and F_i must be concentrated in $\mathbf{R}_+$ for all $i = 1, \dots, n$ hence the spectral measure H of $\bar{T}$ is concentrated in $\mathbf{R}^n_+$ and so is the representing measure ν. ■

We can now formulate a general moment theorem that describes all the price systems that admit a representation as in (3.42).

Theorem 23 *Let $\mathbb{S}$ be defined as in (3.3.2). A sequence $f(s) : \mathbb{S} \rightarrow \mathbf{R}$ is a moment sequence and can be represented as in (3.42):*

$$f(s) = \mathbf{E}_\nu[s(x)], \quad \textit{for all } s \in \mathbb{S},$$

for some measure ν with support in $\mathbf{R}^n_+$, if and only if there is a sequence $p(s,k)$:

$(\mathbb{S}, \mathbf{N}) \to \mathbf{R}$ *such that:*

(i) $p(s,0) = f(s)$ *for all* $s \in \mathbb{S}$,

(ii) $p(s,k)$ *is positive semidefinite on* $(\mathbb{S}, \mathbf{N})$,

(iii) $p(e_i s, k)$ *is positive semidefinite on* $(\mathbb{S}, \mathbf{N})$ *for* $i = 1, \ldots, n$,

(iv) $p(s,k) = p(s,k+1) - \sum_{i=0}^{n+m} p(e_i^2 s, k+1)$ *for all* $(s,k) \in (\mathbb{S}, \mathbf{N})$.

Furthermore, the representing measure for sequence f *is unique if and only if the sequence* p *is unique.*

Proof. First we show that conditions (i)-(iv) are necessary. With $\mathbb{S}$, the payoff semigroup defined in (3.3.2), we recall that $\mathbb{S}^*$ can be identified with $\mathbf{R}_+^n$, hence $\chi_s(\rho_x) = s(x)$, for all $x \in \mathbf{R}_+^n$, $s \in \mathbb{S}$ and $\rho \in \mathbb{S}^*$. Suppose that f can be represented as:

$$f(s) = \int_{\mathbf{R}_+^n} s(x) d\nu(x), \quad \text{for all } s \in \mathbb{S},$$

we let

$$p(s,k) = \int_{\mathbf{R}_+^n} s(x) \left(1 + \sum_{i=0}^{m+n} e_i^2(x)\right)^{-k} d\nu(x), \quad \text{for all } (s,k) \in (\mathbb{S}, \mathbf{N}),$$

which satisfies (i) and (iv) by construction, $p(s,k)$ is then a moment sequence on the product semigroup $((\mathbb{S}, \cdot) \times (\mathbf{N}, +))$ and as such must be positive semidefinite, hence condition (ii). Then, because for $i = 1, .., n$ we have

$$p(e_i s,k) = \int_{\mathbf{R}_+^n} s(x) \left(1 + \sum_{i=0}^{m+n} e_i^2(x)\right)^{-k} e_i(x) d\nu(x), \quad \text{for all } (s,k) \in (\mathbb{S}, \mathbf{N}),$$

we know that $p(e_i s, k)$ is a moment sequence for the measure $e_i(x)d\nu$, hence condition (iii).

Conversely, let's assume that we are given a sequence $p(s,k)$ satisfying (i)-(iv). We let $\mathcal{A}_\theta(\mathbb{S})$ and $\mathcal{A}(\mathbb{S},y)$ be the **R**-algebras described at the beginning of the section. We define a linear function Λ on $\mathcal{A}(\mathbb{S},y)$ by:

$$L\left(\sum_{j,k} a_j \chi_{s_j} y^k\right) = \sum_{j,k} a_j p(s_j,k)$$

and as in lemma 21, we can define the following algebra homomorphism Φ:

$$\begin{array}{rcl} \mathcal{A}(\mathbb{S},y) & \rightarrow & \mathcal{A}_\theta(\mathbb{S}) \\ p(\rho,y) & \mapsto & \Phi p = p(\rho,\theta(\rho)) \end{array} \tag{3.49}$$

whose kernel $\mathcal{N}$ has been computed in lemma 21, and $\mathcal{A}_\theta(\mathbb{S})$ is isomorphic to the quotient $\mathcal{A}(\mathbb{S},y)/\mathcal{N}$. Condition (iv) implies that $L(\mathcal{N})=0$ and we can then define a linear form Λ on $\mathcal{A}_\theta(\mathbb{S})$ by:

$$\Lambda(r) = L(q), \quad \text{where } r(\rho) = q(\rho,\theta(\rho)), \quad \text{for all } \rho \in \mathbb{S}^*,$$

with $r \in \mathcal{A}_\theta(\mathbb{S})$ and $q \in \mathcal{A}(\mathbb{S},y)$. Because of (i)-(iv), the form Λ satisfies the hypothesis of proposition 22 and has a unique representing measure ν. ■

3.3.3 Price bounds and static hedging

In this section, we show how the classic duality between the existence of a pricing measure and that of a replicating portfolio transposes into the moment framework described in the previous section. In particular, we detail how an optimal static super/sub-replicating portfolio can be constructed using the solution to the dual of to the moment problem in (3.40). In particular, in a result that is consistent with the dynamic framework (see Avellaneda, Levy & Paras (1995)), the replicating portfolio only involves options in the data set and no other option is needed to "complete the grid".

Price bounds via semidefinite programming

Here, we show how one can compute bounds on the solution of problem (3.40) using a subset of the moment conditions imposed by theorem 23. These conditions cast (3.42) as a semidefinite program (see Nesterov & Nemirovskii (1994) or Vandenberghe & Boyd (1996)), for which there are numerical solvers such as SEDUMI by Sturm (1999).

Asset distributions with compact support As before, we denote by $\mathcal{A}(\mathbb{S})$ the $\mathbf{R}$-algebra generated by the functions $\chi_s : \mathbb{S}^* \to \mathbf{R}$ such that $\chi_s(\rho) = \rho(s)$ for all $s \in \mathbb{S}$ and $\rho \in \mathbb{S}^*$. For a polynomial $p \in \mathcal{A}(\mathbb{S})$ with $p = \sum_i q_i \chi_{g_i}$ where $g_i \in \mathbb{S}$, and for $s \in \mathbb{S}$ we set

$$p\rho(s) = \sum_i q_i \rho(s g_i).$$

With $\mathbb{S}$ the payoff semigroup defined in (3.3.2), we recall that $\mathbb{S}^*$ can be identified with $\mathbf{R}^n$, hence $\chi_s(\rho_x) = s(x)$, for all $x \in \mathbf{R}^n$, $s \in \mathbb{S}$ and $\rho_x \in \mathbb{S}^*$. This means that $p \in \mathcal{A}(\mathbb{S})$ can be rewritten

$$p(x) = \sum_i q_i s(x) g_i(x), \quad \text{for all } x \in \mathbf{R}^n_+.$$

We now recall the construction of moment matrices as in Curto & Fialkow (2000) and Lasserre (2001). We adopt the following multiindex notation for monomials in $\mathcal{A}(\mathbb{S})$:

$$e^\alpha(x) := e_0^{\alpha_0}(x) e_1^{\alpha_1}(x) \cdots e_{m+n}^{\alpha_{m+n}}(x),$$

for $\alpha \in \mathbf{N}^{n+m+1}$ and we let

$$y^{\mathrm{e}} = (1, e_0, \ldots, e_{m+n}, e_0^2, e_0 e_1, \ldots, e_0^d, \ldots, e_{m+n}^d) \tag{3.50}$$

be the vector of all monomials in $\mathcal{A}(\mathbb{S})$, up to degree d, listed in graded lexicographic order. We denote by $s(d)$ the size of the vector y^{e}. Let $y \in \mathbf{R}^{s(2d)}$ be the vector of moments (indexed as in y^{e}) of some probability measure ν with support in $\mathbf{R}^n_+$, we denote by $M_d(y) \in \mathbf{R}^{s(d)\times s(d)}$, the symmetric matrix:

$$M_d(y)_{i,j} = \int_{\mathbf{R}^n_+} y^{\mathrm{e}}_i(x) y^{\mathrm{e}}_j(x) d\nu(x), \quad \text{for } i,j = 1,...,s(d)$$

In the rest of the paper, we will always implicitly assume that $y_1 = 1$. With $\beta(i)$ the exponent of the monomial y^{e}_i and conversely, $i(\beta)$ the index of the monomial e^β in y^{e}. We notice that for a given moment vector $y \in \mathbf{R}^{s(d)}$ ordered as in (3.53), the first row and columns of the matrix $M_d(y)$ are then equal to y. The rest of the matrix is then constructed according to:

$$M_d(y)_{i,j} = y_{i(\alpha+\beta)} \;\text{ if }\; M_d(y)_{i,1} = y_{i(\alpha)} \;\text{ and }\; M_d(y)_{1,j} = y_{i(\beta)}.$$

Similarly, let $g \in \mathcal{A}(\mathbb{S})$, we derive the moment matrix for the measure $g(x)d\nu$ on $\mathbf{R}^n_+$ (called the localizing matrix in Curto & Fialkow (2000)), denoted by $M_d(gy) \in \mathbf{S}^{s(d)}$, from the matrix of moments $M_d(y)$ by:

$$M_d(gy)_{i,j} = \int_{\mathbf{R}^n_+} (y_e)_i(x) (y_e)_j(x) g(x) d\nu(x)$$

for $i,j = 1,...,s(d)$. The coefficients of the matrix $M_m(gy)$ are given by:

$$M_d(gy)_{i,j} = \sum_\alpha g_\alpha y_{i(\beta(i)+\beta(j)+\alpha)} \tag{3.51}$$

We can then form a semidefinite program to compute a lower bound on the optimal solution to (3.40) using a subset of the moment constraints in theorem 3.43, taking only monomials and moments in y up to a certain degree.

Corollary 24 *Let N be a positive integer and $y \in \mathbf{R}^{s(2N)}$, a lower bound on the* *optimal value of:*

$$\begin{array}{ll} \textit{minimize} & p_0 := \mathbf{E}_\nu[e_0(x)] \\ \textit{subject to} & \mathbf{E}_\nu[e_i(x)] = p_i, \quad i = 1, \ldots, n+m, \end{array}$$

can be computed as the solution of the following semidefinite program:

$$\begin{array}{ll} \textit{minimize} & y_2 \\ \textit{subject to} & M_N(y) \succeq 0 \\ & M_N(e_j y) \succeq 0, \quad \textit{for } j = 1, \ldots, n, \\ & M_N\left((\beta - \sum_{k=0}^{n+m} e_k)y)\right) \succeq 0 \\ & y_{(j+2)} = p_j, \quad \textit{for } j = 1, \ldots, n+m \textit{ and } s \in \mathbb{S} \end{array} \tag{3.52}$$

where s is such that $i(s) \leq s(2N)$. The optimal value of (3.52) converges to the optimal value of the original program as $N \to \infty$.

Unbounded distributions Here we work on the product semigroup $(\mathbb{S}, \cdot) \times (\mathbf{N}, +)$. Its dual is the set of functions $\rho_x : (\mathbb{S}, \mathbf{N}) \to \mathbf{R}$ such that $\rho_x((s,k)) = s(x)x^k$ for all $s \in \mathbb{S}$, $k \in \mathbf{N}$ and $x \in \mathbf{R}^n$. As before, we denote by $\mathcal{A}(\mathbb{S}, \mathbf{N})$ the $\mathbf{R}$-algebra generated by the functions $\chi_s : (\mathbb{S}, \mathbf{N})^* \to \mathbf{R}$ such that $\chi_{(s,k)}(\rho) = \rho((s,k))$ for all $s \in \mathbb{S}$ and $\rho \in (\mathbb{S}, \mathbf{N})^*$. With $\mathbb{S}$, the payoff semigroup defined in (3.3.2), here $(\mathbb{S}, \mathbf{N})^*$ can again be identified with $\mathbf{R}^n$, hence $\chi_{(s,k)}(\rho_x) = s(x)x^k$, for all $x \in \mathbf{R}^n$, $(s,k) \in (\mathbb{S}, \mathbf{N})$ and $\rho_x \in (\mathbb{S}, \mathbf{N})^*$. By construction, we have

$$\left(\chi_{(s,k)}\right)^2 = \chi_{(s^2,2k)}, \quad \text{for all } (s,k) \in (\mathbb{S}, \mathbf{N}),$$

and for a polynomial $p \in \mathcal{A}(\mathbb{S}, \mathbf{N})$ with $p = \sum_i q_i \chi_{g_i} x^{k_i}$ where $(g_i, k_i) \in (\mathbb{S}, \mathbf{N})$, and for $(s,l) \in (\mathbb{S}, \mathbf{N})$ we set

$$p((s,l))(x) = \sum_i q_i s(x) g_i(x) x^{k_i + l},$$

for all $x \in \mathbf{R}^n$. We adopt here the multiindex notation for monomials in $\mathcal{A}(\mathbb{S}, \mathbf{N})$:

$$e^\alpha := (e_0, 0)^{\alpha_0} (e_1, 0)^{\alpha_1} \cdots (e_{m+n}, 0)^{\alpha_{m+n}} (1, 1)^{\alpha_{m+n+1}}.$$

We then let

$$y^{\mathrm{e}} = (1, (e_0, 0), \ldots, (e_{m+n}, 0), (1, 1), (e_0, 0)^2, (e_0, 0)(e_1, 0), \ldots, (e_0, 0)^d, \ldots, (1, 1)^d) \tag{3.53}$$

be the vector of all monomials in $\mathcal{A}(\mathbb{S}, \mathbf{N})$, up to degree d, listed in graded lexicographic order. We denote by $s(d)$ the size of the vector y^{e}. The matrices $M_d(y)$ and $M_d(gy)$ are defined as in the compact case above.

We can again form a semidefinite program, this time using a subset of the moment constraints in theorem 23, taking only moments up to a certain degree.

Corollary 25 *Let N be a positive integer and $y \in \mathbf{R}^{s(2N)}$, a lower bound on the optimal value of:*

$$\begin{array}{ll} \textit{minimize} & p_0 := \mathbf{E}_\nu[e_0(x)] \\ \textit{subject to} & \mathbf{E}_\nu[e_i(x)] = p_i, \quad i = 1, \ldots, n+m, \end{array}$$

can be computed as the solution of the following semidefinite program:

$$\begin{array}{ll} \textit{minimize} & y_2 \\ \textit{subject to} & M_N(y) \succeq 0 \\ & M_N((e_j, 0)y) \succeq 0, \quad \textit{for } j = 1, \ldots, n, \\ & y_{i(s,k)} = y_{i(s,k+1)} - \sum_{i=0}^{n+m} y_{i(e_i^2 s, k+1)} \\ & y_{(j+1)} = p_j, \quad \textit{for } j = 1, \ldots, n+m \textit{ and } (s,k) \in (\mathbb{S}, \mathbf{N}) \end{array} \tag{3.54}$$

where (s, k) are taken such that $i(s, k) \leq s(2N)$. The optimal value of (3.54) converges to the optimal value of the original program as $N \to \infty$.

Static hedging portfolios and sums of squares

We let here $\Sigma \subset \mathcal{A}(\mathbb{S})$ be the set of polynomials that are sums of squares of polynomials in $\mathcal{A}(\mathbb{S})$, and $\mathcal{P}$ the set of positive semidefinite sequences on $\mathbb{S}$. The central argument of this paper is to replace the conic duality between probability measures and positive portfolios:

$$p(x) \geq 0 \Leftrightarrow \int p(x) d\nu \geq 0, \quad \text{for all measures } \nu,$$

by the conic duality between positive semidefinite sequences $\mathcal{P}$ and sums of squares polynomials Σ:

$$\langle f, p \rangle \geq 0 \text{ for all } p \in \Sigma \text{ iff } f \in \mathcal{P}.$$

for $p \in \mathcal{A}(\mathbb{S})$ with $p = \sum_i q_i \chi_{s_i}$ and $f : \mathbb{S} \to \mathbf{R}$ having defined $\langle f, p \rangle = \sum_i q_i f(s_i)$. The previous section used positive semidefinite sequences to characterize viable price sets, in this section, we detail the dual point of view and use sum of squares polynomials to characterize super/sub-replicating portfolios.

From the initial price problem (3.40) written in terms of straddles:

$$\begin{array}{ll} \text{minimize} & p_0 := \int_{\mathbf{R}^n_+} e_0(x) d\nu(x) \\ \text{subject to} & \int_{\mathbf{R}^n_+} e_i(x) d\nu(x) = p_i, \quad i = 1, \ldots, n+m, \\ & \int_{\mathbf{R}^n_+} d\nu(x) = 1, \end{array} \tag{3.55}$$

in the variable ν, a positive measure on $\mathbf{R}^n_+$. We can form the Lagrangian:

$$L(\nu, \lambda) = \lambda_{n+m+1} + \sum_{i=1}^{n+m} \lambda_i p_i + \int_{\mathbf{R}^n_+} \left(e_0(x) - \sum_{i=1}^{n+m} \lambda_i e_i(x) - \lambda_{n+m+1} \right) d\nu(x)$$

with variables ν and $\lambda \in \mathbf{R}^{n+m+1}$. We obtain the classic dual as a portfolio replication problem:

$$\begin{array}{ll} \text{maximize} & \lambda_{n+m+1} + \sum_{i=1}^{n+m} \lambda_i p_i \\ \text{subject to} & e_0(x) - \sum_{i=1}^{n+m} \lambda_i e_i(x) - \lambda_{n+m+1} \geq 0, \quad \text{for all } x \in \mathbf{R}_+^n. \end{array} \tag{3.56}$$

in the variable $\lambda \in \mathbf{R}^{n+m+1}$.

Unfortunately, the problem formulations above are numerically intractable except in certain particular cases (see Bertsimas & Popescu (2002) and d'Aspremont & El Ghaoui (2003)). On the other hand, as we have seen in the previous section, the conditions of theorem 3.43 turn problem (3.40) into an infinite dimensional semidefinite program which can be relaxed to produce tractable bounds on the solution of (3.40). Here, we detail the accompanying duality theory to exhibit a static hedging portfolios corresponding to these bounds.

We can assume without loss of generality that the payoff functions $e_i(x)$ for $i = 0, \ldots, m+n$, together with the cash $1_{\mathbb{S}}$, are linearly independent. Then Berg et al. (1984, Proposition 6.1.8 and Theorem 6.1.10) hold and we can form a dual to the cone of positive semidefinite functions on $\mathbb{S}$ as follows. For $p \in \mathcal{A}(\mathbb{S})$ with $p = \sum_i q_i \chi_{s_i}$ and $f : \mathbb{S} \to \mathbf{R}$ with:

$$\langle f, p \rangle = \sum_i q_i f(s_i),$$

Berg et al. (1984, Theorem 6.1.10) states that Σ is the polar cone of $\mathcal{P}$ for the above bilinear form, in other words:

$$\langle f, p \rangle \geq 0 \text{ for all } p \in \Sigma \text{ iff } f \in \mathcal{P}. \tag{3.57}$$

We can use this conic duality to compute a dual to program (3.52). Considering the compact case for simplicity, corollary 24 states that the initial pricing problem:

$$\begin{array}{ll} \text{minimize} & p_0 := \mathbf{E}_\nu[e_0(x)] \\ \text{subject to} & \mathbf{E}_\nu[e_i(x)] = p_i, \quad i = 1, \ldots, n+m, \end{array}$$

is equivalent to the following (infinite) semidefinite program:

$$
\begin{array}{ll}
\text{minimize} & y_2 \\
\text{subject to} & M(y) \succeq 0 \\
& M(e_j y) \succeq 0, \quad \text{for } j = 0, \ldots, n+m \\
& M\left((\beta - \sum_{k=0}^{n+m} e_k) y)\right) \succeq 0 \\
& y_{(j+2)} = p_j, \quad \text{for } j = 1, \ldots, n+m \\
& y_1 = 1.
\end{array}
$$

in the variable $y : \mathbb{S} \to \mathbf{R}$. Using the conic duality in (3.57), we form the Lagrangian:

$$
\begin{array}{rcl}
L(y, \lambda, q) & := & y_2 + (1 - y_1)\lambda_{n+m+1} + \sum_{j=1}^{n+m} (p_j - y_{(j+2)})\lambda_j - \langle y, q_0 \rangle \\
& & - \sum_{j=0}^{n+m} \langle e_j y, q_j \rangle - \langle (\beta - \sum_{k=0}^{n+m} e_k) y, q_{n+1} \rangle
\end{array}
$$

or again:

$$
\begin{array}{rcl}
L(y, \lambda, q) & := & y_2 + (1 - y_1)\lambda_{n+m+1} + \sum_{j=1}^{n+m} (p_j - y_{(j+2)})\lambda_j - \langle y, q_0 \rangle \\
& & - \sum_{j=0}^{n+m} \langle y, e_j q_j \rangle - \langle y, (\beta - \sum_{k=0}^{n+m} e_k) q_{n+1} \rangle
\end{array}
$$

in the variables $y : \mathbb{S} \to \mathbf{R}$, $\lambda \in \mathbf{R}^{n+m+1}$ and $q_j \in \Sigma$ for $j = 0, \ldots, (n+1)$. We then get the dual as a portfolio problem:

$$
\begin{array}{ll}
\text{maximize} & \sum_{j=1}^{n+m} p_j \lambda_j + \lambda_{n+m+1} \\
\text{subject to} & e_0(x) - \sum_{j=1}^{n+m} \lambda_j e_j(x) - \lambda_{n+m+1} \\
& = q_0(x) + \sum_{j=1}^{n+m} q_j(x) e_j(x) + (\beta - \sum_{k=0}^{n+m} e_k(x)) q_{n+1}(x)
\end{array}
\tag{3.58}
$$

in the variables $\lambda \in \mathbf{R}^{n+m+1}$ and $q_j \in \Sigma$ for $j = 0, \ldots, (n+1)$.

The key difference between this portfolio problem and the one in (3.56) is that the (intractable) positivity constraint $e_0(x) - \sum_{i=1}^{n+m} \lambda_i e_i(x) - \lambda_{n+m+1} \geq 0$ in (3.56) is replaced by the tractable condition that this portfolio be written as a combination of sums of squares of polynomials in $\mathcal{A}(\mathbb{S})$. Such combinations can be constructed directly from the dual solution to the semidefinite program in (3.52), hence a numerical solution to the program in (3.52) provides both a price bound and a super/sub replicating portfolio.

3.3.4 Applications

In this section, we list common examples of basket options.

Spread options

Spread options are call or put options on the difference of two assets, x_1 and x_2. Their payoff at maturity is given by:

$$(x_1 - x_2 - K)_+.$$

This is a direct example of basket option, which are used as an instrument to hedge correlation among assets.

Index options

Indices are computed as a combination of the values of a certain number of assets x_i, weighted by their market capitalization w_i. The payoff of an index option is then naturally:

$$\left(\sum_{i=1}^{n} x_i - K\right)_+,$$

which is another very common example of basket option.

Swaptions

In interest rate markets, the equivalent of the equity call option is called a *swaption*, which is a call option on a swap rate. As detailed in Rebonato (1998) for example, swaptions can be treated as options on a basket of rates, with the swaption's payoff can approximated by that of a basket call on forward rates.

FOREX options

If we let x_{AB} and x_{BC} be the exchange rate between currencies A & B, and B & C respectively. Then a call option on the exchange rate between currencies A & C can be written:

$$(x_{AB}x_{AC} - K)_+,$$

since $x_{AC} = x_{AB}x_{AC}$. This is not a basket option, but we can adapt the result in (3.52) to get static arbitrage constraints on the price of options on the rate x_{AC} knowing the price of options on the exchange rates x_{AB} and x_{BC}. The only difference here is that the semigroup $\mathcal{S}$ is her generated by $e_i(x)$ for $i = 0, \ldots, m+2$, the cash $1_{\mathbb{S}}$ and their products, where $e_i(x) = x_i$ for $i = 1, \ldots, 2$ and $e_{(n+j)}(x) = |x_1 x_2 - K_i|$ for $j = 0, \ldots, m$. We then have $e^2_{(n+j)}(x) = (x_1 x_2 - K_i)^2$ instead of $e^2_{(n+j)}(x) = (w_j^T x - K_j)^2$.

3.3.5 Numerical Example

We consider a model two assets: x_1, x_2 and look for bounds on the price of the basket $|x_1 + x_2 - K|$. We use a simple discrete model for the assets:

$$x = \{(0,0), (0,3), (3,0), (1,2), (5,4)\}$$

with probability

$$p = (.2, .2, .2, .3, .1)$$

to simulate market prices for the forwards and the following straddles:

$$|x_1 - .9|, |x_1 - 1|, |x_2 - 1.9|, |x_2 - 2|, |x_2 - 2.1|.$$

The results are detailed in figure (3.4).

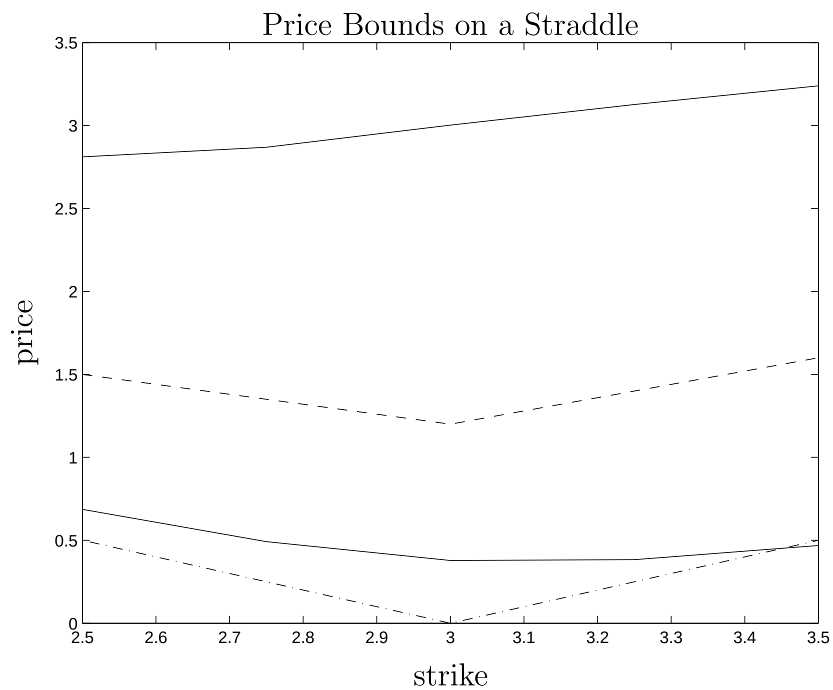

Figure 3.4: Upper and lower price bounds on a straddle (solid lines). The dashed lines represent the payoff function and the actual price.

Bibliography

Anderson, E. J. & Nash, P. (1987), *Linear programming in infinite-dimensional spaces*, Wiley, Chichester. Edward J. Anderson, Peter Nash.

Avellaneda, M., Levy, A. & Paras, A. (1995), 'Pricing and hedging derivative securities in markets with uncertain volatilities', *Applied Mathematical Finance* **2**, 73–88.

Bellman, R. E. (1953), *Stability theory of differential equations*, McGraw-Hill, New York,. 24 cm.

Bellman, R. E. (1957), *Dynamic programming*, Princeton University Press, Princeton,. diagrs. 25 cm. "A Rand Corporation research study.".

Ben-Tal, A. & Nemirovski, A. (2001), *Lectures on modern convex optimization : analysis, algorithms, and engineering applications*, MPS-SIAM series on optimization, Society for Industrial and Applied Mathematics : Mathematical Programming Society, Philadelphia, PA.

Berg, C. (1980), The multidimensional moment problem and semi-groups, *in* H. Landau, ed., 'Moments in Mathematics', AMS, Providence, RI, pp. 110–124.

Berg, C., Christensen, J. P. R. & Ressel, P. (1984), *Harmonic analysis on semigroups : theory of positive definite and related functions*, Vol. 100 of *Graduate texts in mathematics*, Springer-Verlag, New York.

Berg, C. & Maserick, P. H. (1984), 'Exponentially bounded positive definite functions', *Illinois journal of mathematics* **28**(1), 162–179.

Bertsimas, D. & Popescu, I. (2002), 'On the relation between option and stock prices: a convex optimization approach', *Operations Research* **50**(2).

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–659.

Bochnak, J., Coste, M. & Roy, M. F. (1998), *Real algebraic geometry*, Springer, Berlin.

Boyd, S. P., Fazel, M. & Hindi, H. (2000), 'A rank minimization heuristic with application to minimum order system approximation', *American Control Conference* .

Boyd, S. & Vandenberghe, L. (2003), *Convex Optimization*, Cambridge University Press.

Breeden, D. T. & Litzenberger, R. H. (1978), 'Price of state-contingent claims implicit in option prices', *Journal of Business* **51**(4), 621–651.

Buchen, P. & Kelly, M. (1996), 'The maximum entropy distribution of an asset inferred from option prices', *Journal of Financial and Quantitative Anaysis* **31**, 143–159.

Cassier, G. (1984), 'Problème de moments sur un compact de R^n et décomposition de polynômes à plusieurs variables', *Journal of Functional analysis* **58**, 254–266.

Choi, M. D., Lam, T. Y. & Reznick, B. (1995), 'Sums of squares of real polynomials', *Proceedings of Symposia in Pure Mathematics* **58**(2), 103–126.

Curto, R. & Fialkow, A. (2000), 'The truncated complex K-moment problem', *Transactions of the American mathematical society* **352**(6), 2825–2855.

d'Aspremont, A. (2003), 'Interest rate model calibration using semidefinite programming', *Applied Mathematical Finance* **10**(3), 183–213.

d'Aspremont, A. & El Ghaoui, L. (2003), 'Static arbitrage bounds on basket option prices', *ArXiv: math.OC/0302243* .

Duffie, D. (1996), *Dynamic asset pricing theory*, 2nd edn, Princeton University Press, Princeton, N.J.

El Karoui, N. & Quenez, M. (1991), 'Programmation dynamique et évaluation des actifs contingents en marchés incomplets.', *Comptes Rendus de l'Académie des Sciences de Paris, Série I* **313**, 851–854.

El Karoui, N. & Quenez, M. (1995), 'Dynamic programming and pricing of contingent claims in an incomplete market', *Siam Journal of Control and Optimization* **33**, 29–66.

Feron, E. (1999), Nonconvex quadratic programming, semidefinite relaxations and randomization algorithms in information and decision systems, *in* T. E. Djaferis & I. Schick, eds, 'System theory: modeling analysis and control', Kluwer.

Gatermann, K. (2000), *Computer algebra methods for equivariant dynamical systems*, Vol. 1728 of *Lecture notes in mathematics*, Springer-Verlag.

Gatermann, K. & Parillo, P. (2002), Symmetry groups, semidefinite programs, and sums of squares, Technical Report arXiv: math.AC/0211450, ETH Zurich.

Goemans, M. & Williamson, D. (1995), 'Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming', *J. ACM* **42**, 1115–1145.

Grenander, U. (1956), 'On the theory of mortality measurement', *Skandinavisk Aktuarietidskrift* **39**, 70–153.

Groeneboom, P., Jongbloed, G. & Wellner, J. (2001), 'Estimation of a convex function: characterizations and asymptotic theory', *The Annals of Statistics* **29**, 1653–1698.

Hansen, M. B., Gill, R. D. & Baddeley, A. (1996), 'Kaplan-meier type estimators for linear contact distributions', *Scandinavian Journal of Statistics* **23**, 129–55.

Hansen, M. B. & Lauritzen, S. L. (1998), 'Non-parametric bayes inference for concave distribution functions', *Research Report R-98-2015. Department of Mathematics. Aalborg University* .

Hassibi, A., How, J. & Boyd, S. (1999), 'Low-authority controller design via convex optimization', *AIAA Journal of Guidance, Control and Dynamics* **22**(6), 862–872.

Helgason, S. (1999), *The Radon transform*, 2nd edn, Birkhauser, Boston. Progress in mathematics (Boston, Mass.) ; vol. 5.

Henkin, G. & Shananin, A. (1990), 'Bernstein theorems and Radon transform, application to the theory of production functions', *American Mathematical Society: Translation of mathematical monographs* **81**, 189–223.

Hettich, R. & Kortanek, K. O. (1993), 'Semi-infinite programming: Theory, methods and applications', *SIAM Review* **35**(3), 380–429.

Karatzas, I. & Shreve, S. E. (1998), *Methods of mathematical finance*, Vol. 39 of *Applications of mathematics*, Springer, New York.

Keeney, R. L. (1977), 'The art of assessing multiattribute utility functions', *Organizational Behavior and Human Performance* **19**, 267–310.

Keeney, R. L. & Raiffa, H. (1993), *Decisions with multiple objectives : preferences and value tradeoffs*, Cambridge University Press, Cambridge England ; New York, NY. Ralph L. Keeney and Howard Raiffa. ill. ; 24 cm. Originally published: New York : Wiley, 1976.

Lasserre, J. B. (2001), 'Global optimization with polynomials and the problem of moments', *SIAM Journal on Optimization* **11**(3), 796–817.

Lasserre, J. B. (2002), 'An explicit equivalent positive semidefinite program for nonlinear 0-1 programs', *SIAM Journal on Optimization* **12**(3), 756–769.

Laurent, J. & Leisen, D. (2000), Building a consistent pricing model from observed option prices, *in* M. Avellaneda, ed., 'Quantitative Analysis in Financial Markets', World Scientific Publishing.

Lemaréchal, C. & Oustry, F. (1999), 'Semidefinite relaxations and lagrangian duality with application to combinatorial optimization.', *INRIA, Rapport de recherche* **3710**.

Maserick, P. H. (1977), 'Moments of measures on convex bodies', *Pacific journal of mathematics* **68**(1), 135–152.

Meyer, R. F. & Pratt, J. W. (1968), 'The consistent assessment and fairing of preference functions', *IEEE Transactions on Systems Science and Cybernetics* **4**(3), 270–278.

Mirrlees, J. A. (1971), 'An exploration of the theory of optimum income taxation', *Review of economic studies* **38**, 175–208.

Nesterov, I. (2000), Squared functional systems and optimization problems, Technical Report 1472, CORE reprints.

Nesterov, I. & Nemirovskii, A. (1994), *Interior-point polynomial algorithms in convex programming*, Society for Industrial and Applied Mathematics, Philadelphia.

Parillo, P. (2000), Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization, PhD thesis, California Institute of Technology.

Parillo, P. & Sturmfels, B. (2001), Minimizing polynomial functions, Technical Report arXiv: math.OC/0103170, DIMACS, Rutgers University.

Pratt, J. W., Raiffa, H. & Schlaifer, R. (1964), 'The foundation of decision under uncertainty: an elementary exposition', *Journal of the American Statistical Association* **59**, 353–375.

Pullan, M. C. (2000), 'Convergence of a general class of algorithms for separated continuous linear programs', *SIAM Journal on Optimization* **10**(3), 722–731.

Putinar, M. (1993), 'Positive polynomials on compact semi-algebraic sets', *Indiana University Mathematics Journal* **42**(3), 969–984.

Putinar, M. & Vasilescu, F.-H. (1999), 'Solving moment problems by dimensional extension', *Annals of Mathematics* **149**, 1087–1107.

Ramm, A. G. & Katsevich, A. I. (1996), *The radon transform and local tomography*, CRC Press, Boca Raton.

Rebonato, R. (1998), *Interest-Rate Options Models*, Financial Engineering, Wiley.

Reznick, B. (1978), 'Extremal psd forms with few terms', *Duke mathematical journal* **45**, 363–374.

Reznick, B. (1996), Some concrete aspects of hilbert's 17th problem, Technical Report 56, Séminaire de structures algébriques ordonnées: Equipe de logique mathématique Prépublications (F. Delon, M. Dickmann, D. Gondard, eds.), Publ. Math. Univ. Paris VII.

Rochet, J. C. & Chone, P. (1998), 'Ironing, sweeping and multidimensional screening', *Econometrica* **66**(4), 783–826.

Roman, P. (2003), 'Problème des moments', *Semigroup Forum* **66**, 212–230.

Schmüdgen, K. (1991), 'The K-moment problem for semi-algebraic sets', *Mathematische Annalen* **289**, 203–206.

Serra, J. P. (1982), *Image analysis and mathematical morphology.*

Shor, N. (1987), 'Quadratic optimization problems', *Soviet Journal of Computer and Systems Sciences* **25**, 1–11.

Sturm, J. F. (1999), 'Using sedumi 1.0x, a matlab toolbox for optimization over symmetric cones.', *Working paper, Department of Quantitative Economics, Maastricht University, The Netherlands.* .

Vandenberghe, L. & Boyd, S. (1996), 'Semidefinite programming', *SIAM Review* **38**, 49–95.

Wilson, R. (1993), *Non Linear Pricing*, Oxford University Press, Oxford.