UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

Programación semidefinida para resolver el Problema de Momentos Generalizado y aplicación a minimizar polinomios

Leonard Ehrhorn

Director: Santiago Laplagne

Marzo de 2025

Índice general

Agradecimientos

A Santiago, mi director de tesis, por aceptar la tarea, el tiempo que dedicó a acompañarme en el proceso y sus muchos consejos para facilitar el trabajo con el editor LaTeX.

A los jurados, Teresa Krick y Daniel Perruci, por su predisposición, el esfuerzo dedicado a leer y evaluar la tesis y todo lo que me enseñaron a través de sus observaciones.

A mi familia, que me bancó los estudios todo el tiempo que hizo falta y me apoya incondicionalmente.

A mis seres queridos en general, que me escuchan pacientemente hablando de problemas multidimensionales mientras se preguntan cuántas dimensiones tiene el mundo en el que vivo.

A todos los compañeros y compañeras con quienes, durante más o menos años, compartimos cursadas, charlas de café y resoluciones de ejercicios y tps: Federico, Nicolás, Teo, Nicole, Camila, Sol, Nahuel, Luca, Fabrizio, y varios más que no me vienen a la mente ahora.

A Melina, Alina y mis compañeros del FEM, actuales y anteriores, que marcaron mi vida y no solamente mi paso por la Facultad.

A la militancia: una escuela de vida, gente con la que comparto intereses e ideales, y un refugio donde protegernos entre todos de la tristeza y construir la alternativa.

A la educación pública y en particular la universidad pública, que a pesar de todo siguen dando ejemplo de calidad y resiliencia.

Resumen

En este trabajo estudiamos soluciones para el problema de minimización de polinomios multivariados en K, donde K es $\mathbb{R}^n$ o un conjunto semialgebraico básico compacto. Abordamos este problema como caso particular del Problema de Momentos Generalizado (GMP), que consiste en encontrar una medida boreliana, no negativa y finita que minimice la integral de una función sobre K mientras cumple restricciones. Investigamos algoritmos para resolver de forma exacta o asintótica instancias del GMP con relajaciones semidefinidas, estudiando la teoría al respecto elaborada por Lasserre y Parrilo. Estas soluciones se encuentran a través de obtener secuencias de momentos óptimas y medidas representativas soportadas en finitos puntos de K, que también pueden ser localizados si aseguramos que existen. Encontramos condiciones de dualidad y de convergencia de las relajaciones semidefinidas a la solución buscada. Llegamos a que los puntos que soportan las medidas representativas, bajo ciertas hipótesis, son los mínimos globales del polinomio estudiado en K. En este camino pasamos por los certificados de positividad de Putinar y Schmüdgen, que se expresan en términos de sumas de cuadrados. Estudiamos cómo se vinculan estos certificados con la programación semidefinida. También examinamos un funcional lineal y matrices simétricas que nos ayudan a caracterizar soluciones del problema de momentos completo y el truncado y los vinculamos con las soluciones del GMP.

Introducción

Los problemas de optimización plantean encontrar el mínimo de una función f en un conjunto K definido por algunas restricciones (o sin ellas).

$$f^* = \inf_{x \in K} f(x) \tag{1}$$

Como este problema es demasiado general, se construyó a lo largo del tiempo mucha teoría sobre tipos de problemas de optimización, donde se asume que la función f y/o el conjunto K tienen propiedades concretas para poder abordar mejor el problema en cuestión y generar teoría acorde.

En esta tesis vamos a estudiar algunos tipos concretos de problemas de optimización y las formas en que se conectan. Vamos a conocer el Problema de Momentos Generalizado o *Generalized Moment Problem* (GMP), que apunta a encontrar la medida boreliana que minimiza la integral sobre un conjunto de una función en particular, sujeto a restricciones sobre otras integrales de funciones que esa medida debe cumplir.

El Problema de Momentos Generalizado es muy amplio, y de esa amplitud se derivan múltiples aplicaciones dadas por casos particulares [11]. Algunos ejemplos se dan en probabilidad, finanzas, teoría de control, resolución de ecuaciones polinomiales o el que nos interesa en este caso: optimización de polinomios multivariados.

Esta será la forma de los problemas de optimización que queremos resolver, de forma exacta o al menos aproximada: dentro del problema general (1), vamos a suponer que $f \in \mathbb{R}[x_1, \ldots, x_n]$ y que K es un conjunto semialgebraico básico cerrado. Estudiaremos esencialmente dos casos: cuando este conjunto es compacto, y cuando es $\mathbb{R}^n$ (es decir, un problema de optimización polinomial sin restricciones). En este trabajo consideramos $\mathbb{N}$ como el conjunto de números naturales con el 0.

Vamos a examinar la construcción de algoritmos que resuelven instancias del GMP en el caso particular donde $f \in \mathbb{R}[x_1, \ldots, x_n]$ y determinar cómo se conectan estas soluciones con la búsqueda de mínimos de f. Una herramienta en esta búsqueda será encontrar certificados de positividad de polinomios, y estos certificados se expresarán

en términos de sumas de cuadrados.

Esto nos lleva al otro problema de optimización que vamos a estudiar para acercarnos al GMP y la optimización polinomial: la programación semidefinida. Este es un problema de optimización convexa, esto es, un problema con función objetivo convexa (más aún, lineal en este caso) y conjunto factible convexo. Los problemas de optimización convexa tienen una ventaja: todos sus mínimos locales son globales [1, Proposición 1.2]. O sea, cuando encontramos un mínimo local el problema está resuelto, lo que no vale en general (y muchas veces no se puede encontrar algo mejor que un mínimo local).

En esta tesis vamos a estudiar teóricamente algoritmos eficientes de resolución, pero no vamos a dar detalles sobre su complejidad computacional. Omitimos algunas demostraciones debido a que las teorías correspondientes no son centrales para este trabajo y estudiarlas haría la tesis mucho más extensa de lo que ya es. Damos por hecho que los problemas de programación semidefinida se pueden resolver (encontrar el ínfimo o probar que no existe) en tiempo polinomial. Por otra parte, restringir el estudio de la optimización polinomial a conjuntos semialgebraicos básicos compactos o $\mathbb{R}^n$ tiene la finalidad de que se pueda encontrar soluciones con programación semidefinida de dimensiones lo bastante pequeñas para ser útiles en la práctica.

Durante los años '90 y los primeros 2000, Curto y Fialkow trabajaron con matrices de momentos y crearon las matrices localizadoras para lidiar con el problema de momentos completo y el truncado; lo hicieron inicialmente en variable compleja y con el tiempo generalizaron estos conceptos a variable real y problemas multivariados [5]. En 2001, Lasserre [10] aprovecha estas herramientas para crear la familia de relajaciones semidefinidas que vamos a presentar en este trabajo, a fin de poder resolver de forma exacta o asintótica el problema de optimización polinomial para $\mathbb{R}^n$ o cuando K es semialgebraico básico compacto y un módulo cuadrático relacionado es arquimediano. De esta forma, Lasserre es uno de los primeros en construir soluciones exactas o asintóticas a la minimización polinomial mediante programación semidefinida, y por lo tanto de forma computacionalmente eficiente, en los casos mencionados. En 2003 Parrilo [17] presenta en forma independiente una secuencia de relajaciones semidefinidas similar a la de Lasserre con el mismo objetivo de optimización de polinomios, con la diferencia de no requerir que el módulo cuadrático sea arquimediano si K es compacto.

En este trabajo vamos a explorar entonces el GMP, el método de relajaciones semidefinidas de Lasserre y su aplicación a optimizar polinomios, utilizando su libro [11] como principal fuente de información y referencias. También mencionaremos cómo se modifican los resultados cuando no contamos con la arquimedianidad del módulo cuadrático relacionado.

Capítulo 1

Preliminares

1.1. Bases de Geometría Algebraica

En esta sección vamos a dar definiciones y demostraciones básicas para llegar a construir los conceptos de conjunto semialgebraico básico cerrado, preordering y módulo cuadrático, que aparecerán muchas veces a lo largo de este trabajo.

1.1.1. Conceptos iniciales

Definición 1.1.1 *Sean $\mathbb{K}$ un cuerpo y $\leq$ una relación de orden total en $\mathbb{K}$ que para todos $x, y \in \mathbb{K}$ cumple:*

1. *$x \leq y \Rightarrow x + z \leq y + z, \forall z \in \mathbb{K}$.*

2. *Si $0 \leq x, 0 \leq y$, entonces se tiene $0 \leq xy$.*

Entonces, $(\mathbb{K}, \leq)$ se dice ***cuerpo ordenado****.*

Definición 1.1.2 *Dado un cuerpo $\mathbb{K}$, llamamos* ***preordering*** *de $\mathbb{K}$ a un subconjunto $P \subset \mathbb{K}$ no vacío que cumple:*

1. *Dados $x, y \in P$, se verifica que $x + y \in P$ y que $xy \in P$.*

2. *$x^2 \in P$ para todo $x \in \mathbb{K}$.*

Si $-1 \notin P$, decimos que P es un ***preordering propio****.*

Proposición 1.1.3 *Sean $\mathbb{K}$ un cuerpo con característica distinta a 2 y P un preordering de $\mathbb{K}$. Entonces, P es propio si y sólo si $P \neq \mathbb{K}$.*

Demostración: Si $-1 \notin F$, trivialmente $P \neq \mathbb{K}$. En cambio, si $-1 \in P$ veamos que $P = \mathbb{K}$. En efecto, dado $x \in \mathbb{K}$, se tiene

$$x = \left(\frac{x+1}{2}\right)^2 + (-1)\left(\frac{x-1}{2}\right)^2$$

y se sigue que $x \in P$ por la definición de preordering. □

Proposición 1.1.4 *Dado un cuerpo ordenado $(\mathbb{K}, \leq)$, sea $\sum \mathbb{K}^2$ el conjunto de sumas finitas de cuadrados de elementos de $\mathbb{K}$. Entonces, $\sum \mathbb{K}^2$ es un preordering de $\mathbb{K}$, y todo preordering de $\mathbb{K}$ lo contiene.*

Demostración: Veamos primero la contención. Sea P un preordering de $\mathbb{K}$. Dados $x_1, \ldots, x_n \in \mathbb{K}$, sabemos por definición que $x_i^2 \in P, i = 1, \ldots, n$. Además, $x_1^2 + x_2^2 \in P$ nuevamente por definición de P, y se puede ver inductivamente que $x_1^2 + \cdots + x_n^2 \in P$. Como los $x_1, \ldots, x_n$ eran arbitrarios, concluimos que $\sum \mathbb{K}^2 \subset P$.
Para ver que $\sum \mathbb{K}^2$ es un preordering, primero notamos que la suma de dos sumas finitas de cuadrados de $\mathbb{K}$ es suma finita de cuadrados de $\mathbb{K}$. Por último, dados $x_1, \ldots, x_n, y_1, \ldots, y_k \in \mathbb{K}$, tenemos:

$$\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{j=1}^{k} y_j^2\right) = \sum_{i=1}^{n}\sum_{j=1}^{k}(x_i^2 y_j^2) = \sum_{i=1}^{n}\sum_{j=1}^{k}(x_i y_j)^2$$

por las propiedades distributiva, asociativa y conmutativa del cuerpo $\mathbb{K}$, de manera que el producto de sumas finitas de cuadrados de $\mathbb{K}$ es suma finita de cuadrados de $\mathbb{K}$. □

Dado $\mathbb{K}$ un cuerpo, en adelante notamos $\mathbb{K}[\boldsymbol{x}]$ al anillo conmutativo de polinomios de n variables con coeficientes en $\mathbb{K}$, donde $\boldsymbol{x} = (x_1, \ldots, x_n)$, es decir que lo escribiremos de esta forma cuando funcione como variable de evaluación de un polinomio de varias variables. Por otra parte, escribiremos x como variable en $\mathbb{R}$ o como vector en $\mathbb{R}^p$ en otras situaciones.

Definición 1.1.5 *Dado un anillo conmutativo R, un **ideal** de R es un conjunto $I \subset R$ no vacío, cerrado por sumas y tal que además, para todo $b \in R$, se tiene $a \in I \Rightarrow ab \in I$.*

Proposición 1.1.6 *Dados un cuerpo $\mathbb{K}$ y una familia finita $F = \{f_1, \ldots, f_m\} \subset \mathbb{K}[\boldsymbol{x}]$, el siguiente conjunto es un ideal de $\mathbb{K}[\boldsymbol{x}]$. Lo llamamos **ideal generado por** F.*

$$\langle f_1, \ldots, f_m \rangle := \left\{ \sum_{i=1}^{m} g_i f_i : g_i \in \mathbb{K}[\boldsymbol{x}], i = 1, \ldots, m \right\} \tag{1.1}$$

Demostración: Sea $g = \sum_{i=1}^{m} g_i f_i$ con $g_1, \ldots, g_m \in \mathbb{K}[\boldsymbol{x}]$. Entonces, para $h \in \mathbb{K}[\boldsymbol{x}]$ se tiene que $hg_i \in \mathbb{K}[\boldsymbol{x}], i = 1, \ldots, m$, de forma que hg se puede expresar como suma de polinomios multiplicados por los f_i, quedando

$$hg = \sum_{i=1}^{m} (hg_i) f_i \in \langle f_1, \ldots, f_m \rangle$$

Por otro lado, sean $g, h \in \langle f_1, \ldots, f_m \rangle$. Entonces,

$$g + h = \sum_{i=1}^{m} (g_i + h_i) f_i \in I$$

□

Proposición 1.1.7 *Sea I un ideal en $\mathbb{R}[\boldsymbol{x}]$. Sea $a \in \mathbb{R}^n$ tal que $x_i - a_i \in I$ para todo $i = 1, \ldots, n$. Entonces, $f - f(a) \in I$ para todo $f \in \mathbb{R}[\boldsymbol{x}]$.*

Demostración: Escribimos $f(\boldsymbol{x}) = \sum_{j=0}^{d_1} x_1^j f_{1,j}(x_2, \ldots, x_n)$, con $f_{1,j} \in \mathbb{R}[x_2, \ldots, x_n], 0 \leq j \leq d_1$, siendo d_1 el mayor exponente con que aparece x_1 en f. De esta forma, tenemos

$$f(\boldsymbol{x}) - f(a_1, x_2, \ldots, x_n) = \sum_{j=1}^{d_1} f_{1,j}(x_2, \ldots, x_n)(x_1^j - a_1^j) = q_1(\boldsymbol{x})(x_1 - a_1)$$

donde la última igualdad se verifica porque cada sumando es divisible por $x_1 - a_1$, con $q_1 \in \mathbb{R}[\boldsymbol{x}]$. Esta divisibilidad proviene de que $x_1^j - a_1^j = (x_1 - a_1) \sum_{t=0}^{j-1} x_1^t p^{j-1-t}$. Continuamos con

$$f(\boldsymbol{x}) - f(a_1, a_2, x_3, \ldots, x_n) = q_1(\boldsymbol{x})(x_1 - a_1) + f(a_1, x_2, \ldots, x_n) - f(a_1, a_2, x_3, \ldots, x_n)$$

Podemos continuar con $f(a_1, x_2, \ldots, x_n) = h_1(x_2, \ldots, x_n)$ con $h_1 \in \mathbb{R}[x_2, \ldots, x_n]$, $h_1(x_2, \ldots, x_n) = \sum_{j=0}^{d_2} x_2^j f_{2,j}(x_3, \ldots, x_n)$, con $f_{2,j} \in \mathbb{R}[x_3, \ldots, x_n]$, siendo d_2 el mayor exponente con que aparece x_2 en h_1. De esta forma, nos queda

$$f(a_1, x_2, \ldots, x_n) - f(a_1, a_2, x_3, \ldots, x_n) = h_1(x_2, \ldots, x_n) - h_1(a_2, x_3, \ldots, x_n)$$

$$= \sum_{j=0}^{d_2} f_{2,j}(x_3, \ldots, x_n)(x_2^j - a_2^j) = q_2(x_2, \ldots, x_n)(x_2 - a_2)$$

donde $q_2 \in \mathbb{R}[x_2, \ldots, x_n]$. Deducimos entonces que

$$f(\boldsymbol{x}) - f(a_1, a_2, x_3, \ldots, x_n) = q_1(\boldsymbol{x})(x_1 - a_1) + q_2(x_2, \ldots, x_n)(x_2 - a_2)$$

Inductivamente, si tenemos

$$f(\boldsymbol{x}) - f(a_1, \ldots, a_k, x_{k+1}, \ldots, x_n) = \sum_{t=1}^{k} q_t(x_t, \ldots, x_n)(x_t - a_t)$$

para un natural $1 \leq k < n$, con $q_t \in \mathbb{R}[x_t, \ldots, x_n], 1 \leq t \leq k$, avanzamos como sigue: definimos $h_k \in \mathbb{R}[x_{k+1}, \ldots, x_n]$, $h_k(x_{k+1}, \ldots, x_n) = f(a_1, \ldots, a_k, x_{k+1}, \ldots, x_n), h_k(x_{k+1}, \ldots, x_n) = \sum_{j=0}^{d_{k+1}} x_{k+1}^j f_{k+1,j}(x_{k+2}, \ldots, x_n)$, con $f_{k+1,j} \in \mathbb{R}[x_{k+2}, \ldots, x_n]$, $0 \leq j \leq k+1$, siendo d_{k+1} el mayor exponente con que aparece x_{k+1} en h_k. Planteamos ahora que

$$f(a_1, \ldots, a_k, x_{k+1}, \ldots, x_n) - f(a_1, \ldots, a_k, a_{k+1}, x_{k+2}, \ldots, x_n)$$

$$= \sum_{j=0}^{d_{k+1}} (x_{k+1}^j - a_{k+1}^j) f_{k+1,j}(x_{k+2}, \ldots, x_n) = q_{k+1}(x_{k+1}, \ldots, x_n)(x_{k+1} - a_{k+1})$$

donde $q_{k+1} \in \mathbb{R}[x_{k+1}, \ldots, x_n]$. La hipótesis inductiva nos permite afirmar que

$$f(\boldsymbol{x}) - f(a_1, \ldots, a_{k+1}, x_{k+2}, \ldots, x_n) = \sum_{t=1}^{k+1} q_t(x_t, \ldots, x_n)(x_t - a_t)$$

Como hemos completado el razonamiento inductivo, podemos establecer que

$$f(\boldsymbol{x}) - f(a) = \sum_{t=1}^{n} q_t(x_t, \ldots, x_n)(x_t - a_t)$$

con $q_t \in \mathbb{R}[x_t, \ldots, x_n], 1 \leq t \leq n$. Ahora definimos polinomios $\tilde{q}_t \in \mathbb{R}[\boldsymbol{x}], 1 \leq t \leq n$, donde $q_t(\boldsymbol{x}) = q_t(x_t, \ldots, x_n), \forall 1 \leq t \leq n$, es decir, definimos polinomios con todas las variables pero sin cambiar las evaluaciones. Por lo tanto, tenemos que $f(\boldsymbol{x}) - f(a) = \sum_{t=1}^{n} \tilde{q}_t(\boldsymbol{x})(x_t - a_t)$. Luego, $f - f(a) \in I$ porque $x_t - a_t \in I, \forall 1 \leq t \leq n$. □

Dados un anillo conmutativo R y $T \subset R$, notamos $T + T$ al conjunto de sumas de dos elementos de T, y $T.T$ al conjunto de productos de dos elementos de T. Notamos T^2 y $\sum T^2$ como ya lo hacíamos en la Proposición 1.1.4 para un cuerpo ordenado.

Definición 1.1.8 *Sean R un anillo conmutativo con 1 y $T \subset R$. Entonces, T se dice:*

1. ***módulo cuadrático** en R si $T + T \subset T, R^2.T \subset T, 1 \in T$;*
2. ***preordering** de R si cumple las propiedades de la definición 1.1.2, reemplazando $\mathbb{K}$ por R.*

Proposición 1.1.9 *Sea R un anillo conmutativo. Entonces se tiene:*

1. *Todo preordering de R es módulo cuadrático de R.*
2. *$\sum R^2$ es preordering de R, y además está contenido en todo módulo cuadrático de R y en todo preordering de R.*

Demostración:

1. Sea T un preordering de R. Entonces, $T + T \subset T$, y como $R^2 \subset T$ tenemos $R^2.T \subset T.T \subset T$. Por último, $1^2 = 1 \Rightarrow 1 \in R^2$, y T resulta módulo cuadrático.

2. Debemos ver que $\sum R^2$ es preordering y que es el menor módulo cuadrático. Con esto, es inmediatamente el menor preordering porque todo preordering es módulo cuadrático por el item 1. Se puede ver directamente que $\sum R^2$ cumple las tres propiedades que definen un preordering. Por otro lado, sea Q un módulo cuadrático de R. Como $1 \in Q$, tenemos que $R^2 = R^2.\{1\} \subset R^2.Q \subset Q$, y como $Q + Q \subset Q$ se puede probar inductivamente que toda suma finita de cuadrados de R está en Q.

□

Ahora introducimos algunos conjuntos que van a aparecer en nuestro estudio de los polinomios positivos en el siguiente capítulo. En adelante llamamos $\Sigma[\boldsymbol{x}]$ a $\sum R^2$ cuando $R = \mathbb{R}[\boldsymbol{x}]$.

Proposición 1.1.10 *Dados una familia finita $F = \{f_1, \ldots, f_m\} \subset \mathbb{R}[\boldsymbol{x}]$ y $J \subset \{1, \ldots, m\}$, notamos $f_J(x) = \prod_{j \in J} f_j(x)$ con la convención $f_\emptyset \equiv 1$. El conjunto*

$$P(f_1, \ldots, f_m) := \left\{ \sum q_J f_J : q_J \in \Sigma[\boldsymbol{x}], J \subset \{1, \ldots, m\} \right\} \tag{1.2}$$

es un preordering. Lo llamamos el ***preordering generado por*** *F. También cumple que $f_i \in P(f_1, \ldots, f_m)$ para todo $1 \leq i \leq m$.*

Demostración: La última parte es inmediata porque podemos elegir $q_{\{i\}} = 1$, $q_J = 0$ para todo $J \neq \{i\}$. Abreviamos $P(f_1, \ldots, f_m) = P$. Si tenemos $f, h \in P$, como sumas de $q_J f_J, h_J f_J$ respectivamente, $f + h \in P$ es sumatoria de los $(q_J + h_J) f_J$. Además, $\mathbb{R}[\boldsymbol{x}]^2 \subset P$. En efecto, dado $g \in \mathbb{R}[\boldsymbol{x}]^2$, basta elegir $q_\emptyset = g$, $q_J = 0$ para todo otro $J \subset \{1, \ldots, m\}$. Falta ver que $P.P \subset P$. Cuando tomamos $f, g \in P$, fg queda desarrollado como suma de expresiones de la siguiente forma, para $J, T \subset \{1, \ldots, m\}$:

$$q_J q_T f_J f_T \text{ con } q_J, q_T \in \Sigma[\boldsymbol{x}]$$

Notamos que $f_J f_T = (f^2_{J \cap T}) f_B$, donde $B \subset \{1, \ldots, m\}$ es la diferencia simétrica de J, T. Así queda $fg \in P$ puesto que distribuir el producto $q_J q_T$ de sumas de cuadrados también da una suma de cuadrados. □

Proposición 1.1.11 *Sea $F = \{f_1, \ldots, f_m\} \subset \mathbb{R}[\boldsymbol{x}]$. Entonces el conjunto*

$$Q(f_1, \ldots, f_m) := \left\{ q_0 + \sum_{i=1}^{m} q_i f_i : q_i \in \Sigma[\boldsymbol{x}], i = 0, \ldots, m \right\} \tag{1.3}$$

es un módulo cuadrático, y lo llamamos el ***módulo cuadrático generado por*** *F.*

Demostración: Abreviamos $Q(f_1,\ldots,f_m) = Q$. Primero, $1 \in Q$ porque podemos elegir $q_0 = 1, q_j = 0, j = 1,\ldots,m$. También se tiene $Q+Q \subset Q$ porque $q_0+\sum_{j=1}^m q_j f_j + g_0 + \sum_{j=1}^m g_j f_j = (q_0+g_0) + \sum_{j=1}^m (q_j+g_j) f_j$ con $q_j + g_j \in \Sigma[\boldsymbol{x}]$. Por último, dado $q \in \mathbb{R}[\boldsymbol{x}]$, se tiene $q^2(q_0 + \sum_{j=1}^m q_j f_j) = q^2 q_0 + \sum_{j=1}^m (q^2 q_j) f_j$ con $q^2 q_j \in \Sigma[\boldsymbol{x}]$. □

Proposición 1.1.12 *Dados $f_1,\ldots,f_m \in \mathbb{R}[\boldsymbol{x}]$, $f_i \in Q(f_1,\ldots,f_m)$ para todo $1 \le i \le m$. Además, si nombramos $Q(f_1,\ldots,f_m) = Q$, entonces Q está contenido en todo módulo cuadrático de $\mathbb{R}[\boldsymbol{x}]$ al que pertenezcan $f_1,\ldots,f_m$.*

Demostración: La primera parte es directa porque podemos elegir $q_i = 1, q_j = 0$ para todo $0 \le j \le m, j \ne i$. Para la segunda, sea M el módulo cuadrático mencionado en la proposición. Sea $h \in Q, h = q_0 + \sum_{i=1}^m q_i f_i, q_i \in \Sigma[\boldsymbol{x}]$ para todo $i = 0,\ldots,m$. Por definición de módulo cuadrático, q_0 y todos los sumandos $q_i f_i$ están en M, y por lo tanto h también por ser suma de estos. □

Reuniendo las Proposiciones 1.1.9, 1.1.10 y 1.1.12 se deduce que $Q(f_1,\ldots,f_m) \subset P(f_1,\ldots,f_m)$ para todo $\{f_1,\ldots,f_m\} \subset \mathbb{R}[\boldsymbol{x}]$. Terminamos esta sección enunciando los conceptos necesarios para definir los conjuntos semialgebraicos básicos cerrados, con los cuales vamos a trabajar más adelante.

Definición 1.1.13 *Dado un $\mathbb{R}-$espacio vectorial V, Un conjunto no vacío $C \subset V$ es un **cono convexo** de V si cumple $C + C \subset C$ y $f \in C \Rightarrow \lambda.f \in C$ para todo $\lambda \ge 0$.*

Proposición 1.1.14 *Dados $f_1,\ldots,f_m \in \mathbb{R}[\boldsymbol{x}]$, $P(f_1,\ldots,f_m)$ y $Q(f_1,\ldots,f_m)$ son conos convexos de $\mathbb{R}[\boldsymbol{x}]$.*

Demostración: Como $P(f_1,\ldots,f_m) = P$ y $Q(f_1,\ldots,f_m) = Q$ son preordering y módulo cuadrático respectivamente por la Proposición 1.1.10 y la Proposición 1.1.11, sabemos que $P + P \subset P, Q + Q \subset Q$. Además, en ambos casos es inmediato verificar la otra propiedad porque $\lambda \ge 0 \Rightarrow \lambda = (\sqrt{\lambda})^2$. □

1.1.2. Conjuntos semialgebraicos básicos cerrados

Teorema 1.1.15 *Sea $\mathbb{K}$ un cuerpo. Entonces las siguientes afirmaciones son equivalentes:*

1. *Existe un orden de $\mathbb{K}$ de acuerdo a la Definición 1.1.1.*
2. *$\mathbb{K}$ tiene un preordering propio.*
3. *$-1 \notin \sum \mathbb{K}^2$.*
4. *Para todos $x_1,\ldots,x_n \in \mathbb{K}$, se tiene $\sum_{i=1}^n x_i^2 = 0 \Rightarrow x_1 = \cdots = x_n = 0$.*

Demostración: 2. $\Leftrightarrow$ 3.) Dado T un preordering propio, sabemos que $\sum \mathbb{K}^2 \subset T$ por la Proposición 1.1.4, y por lo tanto $-1 \notin \sum \mathbb{K}^2$. Para la otra implicación, el mismo $\sum \mathbb{K}^2$ es un preordering propio por hipótesis, dado que es un preordering por la Proposición 1.1.4.
4. $\Rightarrow$ 3.) Por contrarrecíproco, si $-1 \in \sum \mathbb{K}^2$, como $1 = 1^2$ resulta que $1 + (-1)$ es una suma de cuadrados que da 0 con un primer sumando no nulo.
3. $\Rightarrow$ 4.) Por contrarrecíproco, sean $x_1, \ldots, x_n \in \mathbb{K}$ tales que $\sum_{i=1}^n x_i^2 = 0$ pero $x_1 \neq 0$. Luego, podemos plantear

$$\left(\frac{1}{x_1}\right)^2 \sum_{i=1}^n x_i^2 = 0 = 1 + \sum_{i=2}^n \left(\frac{x_i}{x_1}\right)^2 = 0$$

restamos 1 de ambos lados y resulta que $-1 \in \sum \mathbb{K}^2$.
1. $\Rightarrow$ 3.) Primero veamos que $1 > 0$. En efecto, por la definición del orden en $\mathbb{K}$ sabemos que $a \geq 0 \Rightarrow a^2 \geq 0$. Por otro lado, si $a < 0$ entonces $-a > 0$, y tenemos $a^2 = (-a)^2 \geq 0$. Luego, $1 = 1^2 > 0$, y por lo tanto $-1 < 0$. Además, por la definición del orden en $\mathbb{K}$, toda suma de cuadrados es no negativa y por lo tanto estrictamente mayor a -1. Luego, $-1 \notin \sum \mathbb{K}^2$.
3. $\Rightarrow$ 1.) Verla en [18, Teorema 1.8]. □

Definición 1.1.16 *Sean dos cuerpos $\mathbb{K}_1 = (C_1, +_1, *_1), \mathbb{K}_2 = (C_2, +_2, *_2)$ tales que $C_1 \subset C_2$, y para todos $x, y \in C_1$ se tiene $x +_1 y = x +_2 y, x *_1 y = x *_2 y$. Se dice en este caso que $\mathbb{K}_2$ es una **extensión** de $\mathbb{K}_1$. Si todo elemento de $\mathbb{K}_2$ es raíz de algún polinomio no nulo con coeficientes en $\mathbb{K}_1$, $\mathbb{K}_2$ se dice **extensión algebraica** de $\mathbb{K}_1$.*

Definición 1.1.17 *Un cuerpo con las cuatro propiedades del Teorema 1.1.15 se llama **cuerpo real**. Un **cuerpo real cerrado** es un cuerpo real $\mathbb{K}$ que no tiene ninguna extesión algebraica real $\mathbb{K}_2$ tal que $\mathbb{K} \neq \mathbb{K}_2$.*

Por ejemplo, $\mathbb{R}$ es un cuerpo real cerrado, puesto que su extensión algebraica es $\mathbb{C} \neq \mathbb{R}$, pero $\mathbb{C}$ no es un cuerpo real. A su vez, $\mathbb{Q}$ es un cuerpo real pero no es cerrado, puesto que los números algebraicos reales forman una extensión algebraica real de $\mathbb{Q}$.

Definición 1.1.18 *Sea $\mathbb{K}$ un cuerpo real cerrado. Llamamos **conjunto semialgebraico básico cerrado** a un conjunto $K \subset \mathbb{K}^n$ de la forma*

$$K = \{x \in \mathbb{K}^n : f_j(x) \geq 0, j = 1, \ldots, m\} \tag{1.4}$$

para alguna familia $\{f_1, \ldots, f_m\} \subset \mathbb{K}[\boldsymbol{x}]$.

Por ejemplo, el conjunto $\{x \in \mathbb{R}^2 : x_1 x_2^3 + 4 \geq 0\}$ es semialgebraico básico cerrado.

Proposición 1.1.19 *$\mathbb{Q}$ no es un conjunto semialgebraico básico cerrado en $\mathbb{R}$.*

Demostración: Supongamos que lo es. Sean $f_1, \ldots, f_m \in \mathbb{R}[x]$ tales que $\mathbb{Q} = \{x \in \mathbb{R} : f_1(x) \geq 0, \ldots, f_m(x) \geq 0\}$. Esto implica que para $z \in \mathbb{R} \setminus \mathbb{Q}$ existe $1 \leq i \leq m$ tal que $f_i(z) < 0$. Pero como $f_1, \ldots, f_m$ son polinomios, en particular son continuos. Por lo tanto existen números racionales donde f_i es negativa, y esto contradice la definición supuesta de $\mathbb{Q}$ como conjunto semialgebraico básico cerrado. □

1.2. Programación Semidefinida

Un problema de programación semidefinida es un problema de optimización que involucra matrices semidefinidas positivas en su definición, como vamos a detallar en esta sección. Recordamos que una matriz $A \in \mathbb{R}^{p\times p}$ es semidefinida positiva si es simétrica y además cumple que $x^tAx \geq 0$ para todo $x \in \mathbb{R}^p$. Si la desigualdad es estricta para todo $x \neq 0$, A se dice definida positiva. Recordamos también que toda matriz simétrica en $\mathbb{R}^{p\times p}$ tiene una base ortonormal de autovectores con autovalores reales; y que es semidefinida positiva si y sólo si todos sus autovalores son no negativos. También es semidefinida positiva si y sólo si todos sus menores principales son no negativos, donde se define un menor principal como seleccionar las filas y columnas de índices en S, con $S \subset \{1, \ldots, p\}, S \neq \emptyset$, y calcular el determinante de esta submatriz. Procedemos a probar una última equivalencia.

Proposición 1.2.1 *Una matriz simétrica $A \in \mathbb{R}^{p\times p}$ es semidefinida positiva si y sólo si existe $B \in \mathbb{R}^{p\times m}$ tal que $A = BB^t$ para algún $m \in \mathbb{N}_{>0}, m \leq p$.*

Demostración: $\Leftarrow$) Dado $x \in \mathbb{R}^p$, tenemos $x^tAx = x^tBB^tx$. Nombramos $z = B^tx$, y notamos que $x^tB = z^t$. Luego, se tiene

$$x^tAx = x^tBB^tx = z^tz = \sum_{j=1}^{p} z_j^2 \geq 0$$

y por lo tanto A es semidefinida positiva.
$\Rightarrow$) Como A es semidefinida positiva, tiene una base ortonormal de autovectores y todos sus autovalores son reales. Sean $D, Q \in \mathbb{R}^{p\times p}$, donde D es una matriz diagonal con los autovalores de A en su diagonal principal, y Q es una matriz ortogonal tal que $A = Q^tDQ$. Como A es semidefinida positiva, todos sus autovalores son no negativos, y por lo tanto podemos definir $D^{1/2}$ colocando en la diagonal las raíces cuadradas de los autovalores que aparecen en D. Luego, tenemos $A = Q^tD^{1/2}D^{1/2}Q$. Nombramos $B = Q^tD^{1/2}$ y se sigue que $A = BB^t$. Por construcción, la cantidad de columnas de B es $m = p$. □

1.2.1. Presentación del problema

Definición 1.2.2 *Sean $\mathcal{S}_p \subset \mathbb{R}^{p\times p}$ el espacio de matrices simétricas de p filas y columnas y $A \in \mathcal{S}_p$. Escribimos $A \succeq 0$ (resp. $A \succ 0$) si A es semidefinida positiva (resp. definida positiva). Dada otra matriz $B \in \mathcal{S}_p$, escribimos $A \succeq B$ (resp. $A \succ B$) si $A - B \succeq 0$ (resp. $A - B \succ 0$).*

Proposición 1.2.3 *El conjunto de matrices semidefinidas positivas de $\mathbb{R}^{p\times p}$ es un cono convexo de $\mathbb{R}^{p\times p}$.*

Demostración: Dadas $A, B \in \mathcal{S}_p$ semidefinidas positivas, $A + B$ también lo es. En efecto,

$$x^tAx \geq 0, x^tBx \geq 0 \Rightarrow x^t(A+B)x = \cdots = x^tAx + x^tBx \geq 0$$

y es inmediato verificar que $\lambda \geq 0 \Rightarrow \lambda.A \succeq 0$, puesto que una matriz simétrica es semidefinida positiva si y sólo si sus autovalores son todos no negativos. □

Proposición 1.2.4 *La función $\langle *, * \rangle : \mathbb{R}^{p\times p} \times \mathbb{R}^{p\times p}, \langle A, B\rangle = \sum_{i=1}^{p}\sum_{j=1}^{p} A_{ij}B_{ij}$ es un producto interno en $\mathbb{R}^{p\times p}$. Se llama **producto interno de Frobenius**. Si $A, B \in \mathcal{S}_p$, este producto resulta igual a tr(AB).*

Demostración: Las propiedades del producto interno pueden verificarse directamente, y lo mismo la igualdad con la traza dado que $\mathrm{tr}(AB) = \sum_{i=1}^{p}\sum_{j=1}^{p} A_{i,j}B_{j,i} = \sum_{i=1}^{p}\sum_{j=1}^{p} A_{i,j}B_{i,j}$. □

Definición 1.2.5 *Dados $c \in \mathbb{R}^n, F_0, F_1, \ldots, F_n \in \mathcal{S}_p$ para algún $p \in \mathbb{N}_{>0}$, un problema de Programación Semidefinida (SDP) es de la forma*

$$\mathbf{P} : \rho_{sup} = \sup_{x\in\mathbb{R}^n} \left\{ -c^tx : F_0 + \sum_{i=1}^{n} x_iF_i \succeq 0 \right\} \tag{1.5}$$

*Se define su **conjunto factible** como $\{x \in \mathbb{R}^n : F_0 + \sum_{i=1}^{n} x_iF_i \succeq 0\}$. Al conjunto factible de un problema de esta forma lo llamamos **espectraedro**. El problema $\mathbf{P}$ se dice **factible** si su conjunto factible es no vacío. $x \in \mathbb{R}^n$ se dice **solución factible** si pertenece al conjunto factible. Un **valor factible** es $-c^tx \in \mathbb{R}$ para una solución factible x. ρ_{sup}, si es finito, se llama **valor óptimo**. Una solución factible que lo realiza se llama **solución óptima**. Cuando sólo nos interese verificar si el conjunto factible es vacío o no, llamaremos a esto un **problema SDP de factibilidad**.*

Por ejemplo, se puede plantear el siguiente problema SDP:

$$\rho_{sup} = \sup_{x\in\mathbb{R}^2} \left\{ -x_1 - 2x_2 : \begin{pmatrix} x_1 & x_2 - 3 \\ x_2 - 3 & x_1 + x_2 \end{pmatrix} \succeq 0 \right\}$$

En este caso, tenemos $c = (1, 2)$, $F_0 = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$, $F_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $F_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Su correspondiente problema de factibilidad se define como

$$\text{hallar } x \in \mathbb{R}^2 \text{ tal que } \begin{pmatrix} x_1 & x_2 - 3 \\ x_2 - 3 & x_1 + x_2 \end{pmatrix} \succeq 0$$

Acabamos de definir los problemas SDP con restricciones sobre una función matricial F; pero en algunas ocasiones se requiere que varias funciones matriciales simétricas sean semidefinidas positivas al mismo tiempo. Vamos a ver que los problemas de este tipo también son SDP.

Lema 1.2.6 *Sean $A_1 \in \mathcal{S}_{p_1}, \ldots, A_k \in \mathcal{S}_{p_k}$. Definimos la matriz simétrica*

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix}$$

Entonces, $A \succeq 0$ si y sólo si $A_i \succeq 0$ para todo $i = 1, \ldots, k$.

Demostración: Eligiendo $1 \leq i, j \leq p_1 + \ldots, + p_k$ se puede comprobar por verificación directa que $A_{i,j} = A_{j,i}$. Por otro lado, tomamos $v = (v_1, \ldots, v_k) \in \mathbb{R}^{p_1 + \cdots + p_k}$, con $v_i \in \mathbb{R}^{p_k}$.
$\Rightarrow$) Dado $1 \leq j \leq k$, usamos que $A \succeq 0$ para probar que $A_j \succeq 0$. Dado $v_j \in \mathbb{R}^{p_j}$, tomamos $v = (0 \ldots 0,\ v_j,\ 0 \ldots 0)$, y tenemos $0 \leq v^t A v = v_j^t A_j v_j$.
$\Leftarrow$) se tiene $v^t A v = \sum_{i=1}^{k} v_i^t A_i v_i \geq 0$ porque cada sumando es no negativo. □

Proposición 1.2.7 *Sean $c \in \mathbb{R}^n$, y sean $F_j : \mathbb{R}^n \to \mathcal{S}_{k_j}, 1 \leq j \leq k$ funciones matriciales simétricas de la forma $F_j(x) = F_{j,0} + \sum_{t=1}^{k} x_k F_{j,k}, 1 \leq j \leq k$. Entonces,*

$$\mathbf{P_k} : \rho_{sup} = \sup_{x \in \mathbb{R}^n} \{-c^t x : F_1(x) \succeq 0, \ldots, F_k(x) \succeq 0\} \tag{1.6}$$

es un problema SDP.

Demostración: Por el Lema1.2.6, basta escribir

$$F(x) = \begin{pmatrix} F_1(x) & 0 & \cdots & 0 \\ 0 & F_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_k(x) \end{pmatrix}$$

□

Observamos que la Proposición anterior nos permite sumar restricciones en $\mathbb{R}$ de la forma $\sum_{j=1}^{n} a_j x_j \geq 0, a_j \in \mathbb{R}, j = 1, \ldots, n$. Se puede definir F_1 función matricial de 1×1, $F_1(x) = (\sum_{j=1}^{n} a_j x_j)$ como matriz de 1×1. En efecto, la no negatividad de la sumatoria equivale a que esta matriz de dimensión 1 sea semidefinida positiva.

Proposición 1.2.8 *El problema* **P** *de la forma (1.1), donde $F : \mathbb{R}^n \to \mathcal{S}_p$ es la función matricial definida en (1.5), tiene función objetivo lineal y su conjunto factible es convexo. En particular, es un problema de optimización convexa.*

Demostración: Lo primero es inmediato porque la función objetivo es $-c^t x$, una transformación lineal. Para la segunda afirmación, sean $x, y \in \mathbb{R}^n$ valores factibles. Entonces, dados $t \in (0,1), z = tx + (1-t)y$, tenemos

$$F(z) = F(tx + (1-t)y) = F_0 + \sum_{i=1}^{k} (tx_i + (1-t)y_i)F_i$$

$$= tF_0 + (1-t)F_0 + t\sum_{i=1}^{k} x_i F_i + (1-t)\sum_{i=1}^{k} y_i F_i = tF(x) + (1-t)F(y)$$

Como $F(x), F(y) \succeq 0$ y sabemos por la Proposición 1.2.3 que el conjunto de matrices semidefinidas positivas es convexo, resulta que $F(z) \succeq 0$. □

1.2.2. Dualidad en Programación Semidefinida

Definición 1.2.9 *Llamamos* ***problema dual*** *asociado a* **P** *al problema*

$$\mathbf{P}^* : \rho_{inf} = \inf_{Z \in \mathcal{S}_p} \{\langle F_0, Z\rangle : \langle F_i, Z\rangle = c_i, i = 1, \ldots, n; Z \succeq 0\} \tag{1.7}$$

Una matriz $Z \in \mathbb{R}^{p \times p}$ semidefinida positiva que cumpla $\langle F_i, Z\rangle = c_i, \forall 1 \leq i \leq n$ es una ***solución factible****. El valor $\langle F_0, Z\rangle \in \mathbb{R}$ para una solución factible Z se llama* ***valor factible****. ρ_{inf}, si es finito, se llama* ***valor óptimo****. Una solución factible que lo realiza se llama* ***solución óptima****. El problema se dice* ***factible*** *si tiene al menos una solución factible.*

Proposición 1.2.10 $\mathbf{P}^*$ *es un problema SDP.*

Demostración: Vamos a ver que el conjunto factible de $\mathbf{P}^*$ se puede escribir con la forma de $\mathbf{P}$. Dado el conjunto

$$S = \{Z \in \mathcal{S}_p : \langle F_i, Z\rangle = c_i, i = 1, \ldots, n\}$$

notamos que esto nos da un sistema de ecuaciones lineales. En efecto, $Z \in \mathcal{S}_p$ es una matriz desconocida con $\binom{p}{2} + p$ variables y cada producto interno es una combinación lineal de estas variables. De esta forma, los resultados requeridos c_i nos dan un sistema lineal no homogéneo de n ecuaciones con $\binom{p}{2} + p$ incógnitas. Si tiene solución, sea $s \in \mathbb{N}_{>0}$ la dimensión del espacio de soluciones del sistema homogéneo asociado. Luego, tenemos $z_1, \ldots, z_s$ las variables libres de la solución del sistema homogéneo, y tenemos que las variables restantes son combinaciones lineales de las libres. Así, vemos que $Z \in S$ se escribe como $Z = G(z) = G_0 + \sum_{j=1}^{s} G_j z_j$, donde G_0 es una solución particular del sistema original. Además, notamos que si $z_1 = \cdots = z_s = 0$ entonces $Z = G_0$, y si dado $1 \leq j \leq s$ tenemos $z_j = 1, z_k = 0, \forall k \neq j$, entonces $Z = G_0 + G_j$. De esto se sigue que $G_0, G_1, \ldots, G_s$ son matrices simétricas ya que Z es simétrica.
De esta forma, pedir $Z \succeq 0$ equivale a pedir $G_0 + \sum_{j=1}^{s} G_j z_j \succeq 0$. Por último, tenemos

$$\langle F_0, Z\rangle = \langle F_0, G_0 + \sum_{j=1}^{s} G_j z_j\rangle = \langle F_0, G_0\rangle + \sum_{j=1}^{s} z_j \langle F_0, G_j\rangle = \langle F_0, G_0\rangle + d^t z,$$

$$d_j = \langle F_0, G_j\rangle, \forall\ 1 \leq j \leq s$$

por lo que hallar $Z \succeq 0$ que realice el ínfimo del primer valor equivale a encontrar $z \in \mathbb{R}^s$ que realice el ínfimo del último sujeto a $G(z) \succeq 0$. Como este problema tiene la forma de $\mathbf{P}$, esto completa la demostración. □

Lema 1.2.11 *Dado $x \in \mathbb{R}^n, xx^t \in \mathbb{R}^{n\times n}$ es simétrica y semidefinida positiva.*

Demostración: La simetría se deduce de que $(xx^t)_{ij} = x_i x_j$. De esto mismo se sigue que la $j-$ésima columna de xx^t es $x_j x \in \mathbb{R}^n$. Luego, xx^t tiene rango 1 (o rango 0 si $x = 0$) así que tiene autovalor 0 con multiplicidad al menos $n-1$. El autovalor restante es entonces la suma de autovalores $tr(xx^t) = \sum_{i=1}^{n} x_i^2 \geq 0$. Luego, no hay autovalores negativos, así que xx^t es semidefinida positiva. □

Lema 1.2.12 *Una matriz $A \in \mathcal{S}_p$ es semidefinida positiva si y sólo si $\langle A, X\rangle \geq 0$ para toda matriz $X \in \mathbb{R}^{p\times p}, X \succeq 0$.*

Demostración: $\Rightarrow$) Como A es semidefinida positiva existe $B \in \mathbb{R}^{p\times p}$ tal que $A = B^t B$ (podemos elegir B cuadrada por la demostración de la Proposición 1.2.1), y lo mismo pasa para X. Sea $Z \in \mathbb{R}^{p\times p}$ tal que $ZZ^t = X$. Entonces, tenemos

$$\langle A, X\rangle = tr(B^t B X) = tr(BXB^t) = tr(BZZ^t B^t) \geq 0$$

donde la última desigualdad vale porque BZZ^tB^t es semidefinida positiva, y la traza de una matriz simétrica es la suma de sus autovalores.
$\Leftarrow$) Dado $x \in \mathbb{R}^p$, por el Lema 1.2.11 y la implicación anterior tenemos

$$0 \leq \langle A, xx^t \rangle = tr(Axx^t) = tr(x^tAx) = x^tAx$$

donde la anteúltima igualdad se verifica desarrollando a mano cada traza. Como x era arbitrario, resulta que A es semidefinida positiva. □

Teorema 1.2.13 $\mathbf{P}$ *y* $\mathbf{P}^*$ *cumplen* ***dualidad débil****, es decir, el supremo* ρ_{sup} *en (1.5) y el ínfimo* ρ_{inf} *en (1.7) satisfacen* $\rho_{sup} \leq \rho_{inf}$*. Si no son iguales, decimos que hay una* ***brecha de dualidad****.*

Demostración: Sean $x \in \mathbb{R}^n$ y $Z \succeq 0$ soluciones factibles de $\mathbf{P}$ y $\mathbf{P}^*$ respectivamente. Entonces, se tiene

$$-c^tx = -\sum_{i=1}^{n} \langle F_i, Z \rangle x_i = -\left\langle \sum_{i=1}^{n} F_i x_i, Z \right\rangle \leq \langle F_0, Z \rangle$$

donde la última desigualdad vale porque $F(x)$ dada por (1.5) es semidefinida positiva y usando el Lema 1.2.12. Como esta desigualdad vale para cualquier par de soluciones factibles, en particular vale para el supremo y el ínfimo. □

Una consecuencia inmediata es que si el problema primal no está acotado superiormente, el dual no es factible, y viceversa: si el dual no está acotado inferiormente, el primal no es factible. En particular, en ambos casos se tiene $\rho_{inf} = \rho_{sup}$. Usando este hecho y tomando el resultado de dualidad fuerte de [13, Teorema 3.4.1], podemos enunciar lo siguiente. El item 3 sale de combinar los dos anteriores.

Teorema 1.2.14 *Sean el problema* $\mathbf{P}$ *de (1.5) con* F *la función matricial definida en la misma ecuación y el problema* $\mathbf{P}^*$ *de (1.7). Entonces, se verifican los siguientes items:*

1. *Si existe* $x \in \mathbb{R}^n$ *tal que* $F(x) \succ 0$*, entonces* $\rho_{inf} = \rho_{sup}$*, y además si* ρ_{inf} *es finito se realiza; es decir, existe* $Z^* \succeq 0$ *tal que* $\langle F_0, Z^* \rangle = \rho_{inf}$*.*

2. *Si existe* $Z \succ 0$ *factible para* $\mathbf{P}^*$*, entonces* $\rho_{inf} = \rho_{sup}$*, y además si* ρ_{sup} *es finito se realiza; es decir, existe* $x^* \in \mathbb{R}^n$ *tal que* $F(x^*) \succeq 0$ *y* $-c^tx^* = \rho_{sup}$*.*

3. *Si existen* $Z \succ 0$ *factible para* $\mathbf{P}^*$ *y* $x \in \mathbb{R}^n$ *tal que* $F(x) \succ 0$*, entonces* $\rho_{inf} = \rho_{sup}$*, y además los dos son finitos y se realizan.*

Cuando $\rho_{inf} = \rho_{sup}$, se dice que los problemas $\mathbf{P}$ y $\mathbf{P}^*$ cumplen **dualidad fuerte**.

Veamos un ejemplo de problemas primal y dual con dualidad fuerte. Se define el primal como

$$\mathbb{P} : \rho_{sup} = \sup\{x : \begin{pmatrix} 2-x & 1 \\ 1 & -x \end{pmatrix} \succeq 0\}$$

Como una matriz es semidefinida positiva si y sólo si todos sus menores principales son no negativos, tienen que valer $2 - x \geq 0, x \leq 0, -x(2-x) - 1 \geq 0$. Se puede deducir que el valor óptimo es $\rho_{sup} = 1 - \sqrt{2}$, que se realiza con $x^* = \rho_{sup}$.
El problema dual se plantea entonces como

$$\rho_{inf} = \text{ínf} \left\{ \langle F_0, Z \rangle = 2z_{11} + 2z_{12} : \langle Z, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \rangle = -1; Z = \begin{pmatrix} z_{11} & z_{12} \\ z_{12} & z_{22} \end{pmatrix} \succeq 0 \right\}$$

Se puede obtener que el valor óptimo es $1 - \sqrt{2}$, con solución óptima

$$Z = \begin{pmatrix} \frac{2-\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{4} & \frac{2+\sqrt{2}}{4} \end{pmatrix}$$

Es decir, vale en este caso la dualidad fuerte y los valores óptimos de ambos problemas se realizan. Se verifican las hipótesis del item 3 del teorema de dualidad fuerte. En efecto, con $z_{11} = z_{22} = \frac{1}{2}$ y $z_{12} = 0$ se obtiene $Z \succ 0$ solución factible del problema dual; y con $x = -1$ resulta $F(x) \succ 0$ en el problema primal.

Capítulo 2

Polinomios positivos y representaciones

2.1. Sumas de cuadrados y programación semidefinida

Los algoritmos de programación semidefinida que estudiaremos en los capítulos 4 y 5 se pueden expresar en términos de sumas de cuadrados. A su vez, verificar que un polinomio sea suma de cuadrados es un problema de programación semidefinida. Eso se desarrollará en esta sección.

2.1.1. Polinomios positivos vs. sumas de cuadrados

Definición 2.1.1 *Dado $\alpha \in \mathbb{N}^n$, se define el **grado** del monomio en $\mathbb{R}[\boldsymbol{x}]$ con $c \in \mathbb{R}, c \neq 0$ dado por $c\boldsymbol{x}^\alpha = cx_1^{\alpha_1} \dots x_n^{\alpha_n}$ como $|\alpha| = \sum_{i=1}^n \alpha_i$. Para un polinomio $f \in \mathbb{R}[\boldsymbol{x}]$, se define su grado como el mayor grado de alguno de sus monomios con coeficiente no nulo.*

Varios de los problemas que estudiamos en esta tesis serán formulados en términos de sumas de cuadrados. Por eso, queremos estudiar ahora si todos los polinomios positivos lo son. Comenzamos viendo que esto vale en una variable.

Teorema 2.1.2 *Sea $f \in \mathbb{R}[x]$. Entonces, $f(x) \geq 0, \forall x \in \mathbb{R} \Leftrightarrow f \in \Sigma[x]$.*

Demostración: Si f es suma de cuadrados es trivialmente positiva. Para probar la otra implicación, asumimos que f tiene grado mayor o igual a 2 ya que el problema se vuelve trivial si f es constante, y no puede ser positiva si tiene grado impar. Más aún, toda raíz real tiene multiplicidad par, porque de lo contrario el signo de f cambiaría en un entorno de esta raíz. Además, las raíces complejas no reales aparecen por pares

conjugados, y ambos elementos del par tienen el mismo grado. Así, si x_i son las raíces reales y z_j las complejas no reales, podemos escribir

$$f(x) = a\prod_{i=1}^{k}(x-x_i)^2\prod_{j=1}^{m}(x-z_j)(x-\bar{z}_j)$$

donde las raíces reales y los pares conjugados pueden repetirse. Por otra parte, si $z_j = a_j + ib_j$, tenemos que $(x-z_j)(x-\bar{z}_j) = (x-a_j)^2 + b_j^2$, que es una suma de cuadrados. Luego, como f es un producto de sumas de cuadrados, cuando se aplica propiedad distributiva queda también una suma de cuadrados. □

No obstante, esta propiedad no funciona para todos los polinomios en varias variables. Veamos un ejemplo.

Lema 2.1.3 ***Desigualdad aritmético-geométrica:*** *Sean $x_1, \ldots, x_k \geq 0$. Entonces, se tiene $\frac{x_1+\cdots+x_k}{k} \geq \sqrt[k]{x_1 \ldots x_k}$.*

Lema 2.1.4 *Sean $r \in \mathbb{N}_{>0}$ y $g_k \in \mathbb{R}[\boldsymbol{x}], 1 \leq k \leq m$, tales que $f = \sum_{k=1}^{m} g_k^2$. Si f tiene grado $2r$, entonces g_k tiene grado a lo sumo r, para todo $1 \leq k \leq m$.*

Demostración: Entre todos los $g_k, 1 \leq k \leq m$, consideramos los monomios con grado máximo $b_{\text{máx}}$. Entre estos, tomamos los de menor grado en x_1. De estos últimos, tomamos los de menor exponente en x_2, y así siguiendo hasta x_n. Sea x^γ uno de estos monomios extremales, y sea $1 \leq j \leq m$ tal que x^γ aparece en g_j. Sean $\alpha, \beta \in \mathbb{N}^n$ tales que x^α, x^β aparecen en g_j y $\alpha + \beta = 2\gamma$. Veamos que $\alpha = \beta = \gamma$.
Supongamos que no. Sea $1 \leq t \leq n$ el menor índice tal que $\alpha_t \neq \beta_t$. Entonces, para todo $1 \leq p < t$ resulta $\alpha_p = \beta_p = \gamma_p$. Pero como $\alpha_t + \beta_t = 2\gamma_t$, se tiene $\alpha_t < \gamma_t$ o $\beta_t < \gamma_t$. Entonces x^α (o x^β respectivamente) es un monomio de g_j menor a x^γ con el ordenamiento que definimos, lo que contradice la definición de x^γ como monomio extremal. Luego, $\alpha = \beta = \gamma$.
Acabamos de probar que, en g_j^2, el monomio $x^{2\gamma}$ se obtiene únicamente como cuadrado de x^γ y no como producto cruzado de monomios diferentes de grado $|\gamma|$. Tampoco se puede obtener de otros monomios donde alguno tenga grado menor a $|\gamma|$. Luego, su coeficiente será positivo. Como x^γ es un monomio extremal entre todos los $g_k, 1 \leq k \leq m$, no existe un g_k^2 donde $x^{2\gamma}$ aparezca con coeficiente negativo. De esta forma, $x^{2\gamma}$ aparece en f. Dado que f tiene grado $2r$, deducimos que $|\gamma| = r$, y el resultado se sigue. □

Proposición 2.1.5 *Se define el* ***polinomio de Motzkin*** $M : \mathbb{R}^2 \to \mathbb{R}$ *como $M(x_1, x_2) = x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2 + 1$. Entonces, vale $M(x) \geq 0, \forall x \in \mathbb{R}^2$ pero M no es suma de cuadrados.*

Demostración: Por la Desigualdad aritmético-geométrica, tenemos

$$\frac{x_1^4x_2^2 + x_1^2x_2^4 + 1}{3} \geq \sqrt[3]{x_1^6x_2^6} = x_1^2x_2^2$$

y de esto se deduce inmediatamente que $M \geq 0$.
Vamos a ver que una suma de cuadrados g del mismo grado que M tiene necesariamente una forma distinta a la de M. Para empezar, si $g = g_1^2 + \cdots + g_m^2$, veamos que en todo monomio de un $g_i, 1 \leq i \leq m$ aparece x_1 con grado a lo sumo 2. En efecto, si consideramos los g, g_i como polinomios en la variable x_1 con x_2 constante, sigue valiendo que $g = \sum_{i=1}^m g_i^2$, y estos g_i tienen grado a lo sumo 2 en x_1 por el Lema 2.1.4. El mismo razonamiento vale si consideramos la variable x_2 con x_1 constante, de forma que entre todos los g_i (nuevamente con las dos variables) los monomios posibles son $1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^2x_2, x_1x_2^2, x_1^2x_2^2$.
Entre los monomios de los g_i, si alguno fuera $x_1^2x_2^2$ no se cancelaría por aparecer siempre con coeficiente positivo o nulo entre los g_i^2 (ya que quedaría x^4y^4 y la única forma de obtenerlo teniendo grado menor o igual a 2 en x_1, x_2 es con $(x_1^2x_2^2)^2$). Pero M no tiene $x_1^4x_2^4$, de forma que $x_1^2x_2^2$ no aparece entre los g_i. Luego, si aparece $x_1^2x_2^a$ entre los g_i, se tiene $a \leq 1$. Veamos que los g_i tampoco pueden tener x_1^2, x_2^2, x_1, x_2. En efecto, si aparece el monomio x_1^2 en un g_i, se tiene x_1^4 con coeficiente positivo en g_i^2 puesto que la única forma de obtenerlo es con $(x_1^2)^2$ (ya que vimos que la variable x_1 no puede aparecer con grado mayor a 2), pero x_1^4 no está en M. Por un razonamiento análogo, no aparece el monomio x_2^2 entre los g_i. Sabiendo esto, por un razonamiento similar no aparecen los monomios x_1, x_2 entre los g_i. Por lo tanto, se tiene

$$g(x_1, x_2) = \sum_{i=1}^{m} (a_ix_1x_2^2 + b_ix_1^2x_2 + c_ix_1x_2 + d_i)^2$$

Entre los g_i así formulados, la única forma de obtener el monomio $x_1^2x_2^2$ en g es con $(x_1x_2)^2$, por lo que este monomio aparece en g con coeficiente no negativo. En conclusión, $g \neq M$. □

Pero no se trata solamente de un ejemplo. Sin detenernos en esto, mencionamos que la equivalencia entre ser positivo y ser suma de cuadrados (para polinomios no homogéneos) sólo vale para polinomios de una variable, de grado 2, o de 2 variables y grado 4 [15, Teorema 1.2.6].

2.1.2. Positividad y suma de cuadrados como problemas de factibilidad en SDP

Vamos a estudiar cómo se puede decidir si un polinomio es suma de cuadrados planteando un problema de programación semidefinida. Para eso debemos comenzar determinando la dimensión de este problema.

Proposición 2.1.6 *Sea $v_d(\boldsymbol{x})$ el vector de monomios en $\mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo $d \in \mathbb{N}$:*

$$v_d(\boldsymbol{x}) = (1, x_1, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^2, \ldots, x_1^d, \ldots, x_n^d) \tag{2.1}$$

Entonces, la longitud de $v_d(\boldsymbol{x})$, que notamos $s(d)$, es $\binom{n+d}{d}$.

Demostración: Un monomio de grado $0 \leq k \leq d$ se construye eligiendo cómo distribuir k unos entre los exponentes de las n variables y definir que $d-k$ unos no van a ningún exponente. Es decir, entre $n+1$ elementos diferentes (las n variables y la elección de ninguna variable) debemos elegir un total de d elementos, donde estos pueden repetirse y no importa el orden. De esta forma, la cantidad de monomios posibles es la de combinaciones con repetición para d elecciones entre $n+1$ tipos de elementos diferentes, que da $\binom{(n+1)-1+d}{d} = \binom{n+d}{d}$. Ver [7, Tabla 1.8]. □

Notamos $\mathbb{N}_p^n$ al conjunto $\{\alpha \in \mathbb{N}^n : |\alpha| \leq p\}, p \in \mathbb{N}$. Vamos a definir formalmente el orden intuitivo de los monomios de $v_d(\boldsymbol{x})$ que nombramos en el resultado anterior. Consideramos los monomios en $\mathbb{R}[\boldsymbol{x}]$ con el orden lexicográfico graduado con $x_1 < x_2 < \ldots < x_n$. De todos los monomios tomamos sus exponentes en $\mathbb{N}^n$ en el mismo orden, y así definimos $ord : \mathbb{N}^n \to \mathbb{N}_{>0}$, comenzando por $ord(0, \ldots, 0) = 1$.

Ahora vamos a construir la verificación de si un polinomio es suma de cuadrados como un problema SDP de factibilidad. Primero vamos a requerir una formulación relacionada con matrices semidefinidas positivas. Vamos a considerar sólo polinomios de grado par porque, como ya vimos en la sección anterior, sólo estos pueden ser positivos, y por lo tanto es posible que sean sumas de cuadrados. El siguiente resultado es fundamental porque relaciona el problema de las sumas de cuadrados con el álgebra lineal. Se puede encontrar en [2, Teorema 2.4].

Teorema 2.1.7 *Un polinomio $g \in \mathbb{R}[\boldsymbol{x}]$ de grado $2d$ es suma de cuadrados si y sólo si existe $Q \in \mathbb{R}^{s(d)\times s(d)}, Q \succeq 0$ tal que vale lo siguiente:*

$$g(\boldsymbol{x}) = v_d(\boldsymbol{x})^t Q v_d(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathbb{R}^n \tag{2.2}$$

Demostración: $\Rightarrow$) Sean $h_j \in \mathbb{R}[\boldsymbol{x}], 1 \leq j \leq k$ de grado a lo sumo d, tales que $g(\boldsymbol{x}) = \sum_{j=1}^k h_j(\boldsymbol{x})^2$. En efecto, los h_i tienen grado a lo sumo d por el Lema 2.1.4. Por cada $1 \leq j \leq k$, sea $z_j \in \mathbb{R}^{s(d)}$ el vector de coeficientes de h_j, ordenados de acuerdo al orden de monomios de $v_d(\boldsymbol{x})$. Así, tenemos $h_j(\boldsymbol{x}) = z_j^t v_d(\boldsymbol{x})$. Podemos escribir $g(\boldsymbol{x}) = \sum_{j=1}^k v_d(\boldsymbol{x})^t z_j z_j^t v_d(\boldsymbol{x}) = v_d(\boldsymbol{x})^t Q v_d(\boldsymbol{x})$, con $Q = \sum_{j=1}^k z_j z_j^t$. Luego, resulta que $Q \in \mathbb{R}^{s(d)\times s(d)}, Q \succeq 0$ por la Proposición 1.2.3 y el Lema 1.2.11.
$\Leftarrow$) Dada Q la matriz de la hipótesis, sea $H \in \mathbb{R}^{s(d)\times m}$ tal que $Q = HH^t$, la cual

existe por la Proposición 1.2.1. Sea $z(\boldsymbol{x}) = H^t v_d(\boldsymbol{x})$, que resulta entonces un vector de polinomios de grado a lo sumo d. Finalmente, tenemos

$$g(\boldsymbol{x}) = v_d(\boldsymbol{x})^t H H^t v_d(\boldsymbol{x}) = z(\boldsymbol{x})^t z(\boldsymbol{x}) = \sum_{j=1}^{m} z_j(\boldsymbol{x})^2, \forall \boldsymbol{x} \in \mathbb{R}^n$$

y esto es una formulación de g como suma de cuadrados. □

Para $\alpha \in \mathbb{N}^n_{2d}$ definimos en $\mathbb{R}^{s(d)\times s(d)}$ las matrices B_α dadas por

$$(B_\alpha)_{ij} = \begin{cases} 1 \text{ si } v_d(\boldsymbol{x})_i v_d(\boldsymbol{x})_j = \boldsymbol{x}^\alpha \\ 0 \text{ si no} \end{cases}$$

Por ejemplo, para $d = n = 2$ tenemos $v_2(\boldsymbol{x}) = (1, x_1, x_2, x_1^2, x_1x_2, x_2^2)$, así que serán 6 matrices construidas así:

$$B_{(0,0)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; B_{(1,0)} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; B_{(0,1)} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B_{(2,0)} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; B_{(1,1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; B_{(0,2)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Observamos que por definición B_α es simétrica. Con esto tenemos los elementos necesarios para conectar la programación semidefinida y las sumas de cuadrados.

Proposición 2.1.8 *Sea $g \in \mathbb{R}[\boldsymbol{x}]$ de grado $2d$, $g(\boldsymbol{x}) = \sum_{\alpha\in\mathbb{N}^n_{2d}} g_\alpha \boldsymbol{x}^\alpha$. Entonces, g es suma de cuadrados si y sólo si el siguiente problema de factibilidad tiene soluciones:*

$$\textit{hallar } Q \succeq 0 : \langle Q, B_\alpha \rangle = g_\alpha, \forall \alpha \in \mathbb{N}^n_{2d} \tag{2.3}$$

Demostración: Sabemos por el Teorema 2.1.7 que g es suma de cuadrados si y sólo si existe $Q \in \mathcal{S}_{s(d)}, Q \succeq 0$ que cumpla (2.2) para g. Falta ver que estas condiciones equivalen a que Q sea solución de (2.3). Ambas fórmulas requieren que Q sea semidefinida

positiva, y resta ver la equivalencia con las igualdades de productos internos. En efecto, se tiene

$$v_d(\boldsymbol{x})^t Q v_d(\boldsymbol{x}) = \sum_{i=1}^{s(d)} \sum_{j=1}^{s(d)} v_d(\boldsymbol{x})_i v_d(\boldsymbol{x})_j Q_{i,j} = \sum_{\alpha \in \mathbb{N}^n_{2d}} g_\alpha \boldsymbol{x}^\alpha$$

$$= \sum_{\alpha \in \mathbb{N}^n_{2d}} \boldsymbol{x}^\alpha \left(\sum_{i,j : v_d(\boldsymbol{x})_i v_d(\boldsymbol{x})_j = \boldsymbol{x}^\alpha} Q_{i,j} \right)$$

La segunda sumatoria en el último término es $\langle Q, B_\alpha \rangle$ por la Proposición 1.2.4 y la definición de B_α, y el hecho de que estas igualdades valgan para todo $\boldsymbol{x}$ determina la igualdad coeficiente a coeficiente, de forma que $\langle Q, B_\alpha \rangle = g_\alpha$. Esto da la equivalencia buscada. □

La proposición anterior permite estudiar como problema SDP si un polinomio de coeficientes conocidos es suma de cuadrados o no. Notamos ahora que (2.3) es la versión como problema de factibilidad de (1.7), y por la Proposición 1.2.10 resulta un problema SDP de factibilidad. Luego, ya contamos con una formulación de este tipo para la verificación de si un polinomio de grado par es suma de cuadrados.

¿Por qué construimos la verificación de sumas de cuadrados como problema SDP y no directamente la verificación de positividad? La respuesta va a ser que este último también es un problema SDP, pero sin utilidad práctica porque tiene cotas excesivamente grandes en los grados de los polinomios involucrados. Vamos a estudiar esto ahora.

Ya vimos en la sección anterior que existen polinomios positivos que no son suma de cuadrados. Sin embargo, se tiene el siguiente teorema probado por Artin, que era el Problema 17 de Hilbert:

Teorema 2.1.9 *Un polinomio $p \in \mathbb{R}[\boldsymbol{x}]$ es no negativo si y sólo si se puede escribir como suma de cuadrados de funciones racionales.*

Notamos que una suma de cuadrados de funciones racionales se puede escribir como cociente de sumas de cuadrados. En efecto, para polinomios $t, w, h, p \in \mathbb{R}[\boldsymbol{x}]$, tenemos

$$\left(\frac{t}{w}\right)^2 + \left(\frac{h}{p}\right)^2 = \frac{t^2}{w^2} + \frac{h^2}{p^2} = \frac{(tp)^2 + (wh)^2}{(wp)^2}$$

Inductivamente, vale la misma propiedad para cuaquier cantidad finita de cuadrados de funciones racionales. De esta forma, se puede plantear la pregunta por la positividad de un polinomio como

$$f \in \mathbb{R}[\boldsymbol{x}], f(\boldsymbol{x}) \geq 0, \forall \boldsymbol{x} \in \mathbb{R}^n \Leftrightarrow fh = g; h, g \in \Sigma[\boldsymbol{x}]$$

A su vez, el Teorema 2.1.7 nos permite reescribir esto de la siguiente manera, para algún $d \in \mathbb{N}$:

$$f \in \mathbb{R}[\boldsymbol{x}], f(\boldsymbol{x}) \geq 0, \forall \boldsymbol{x} \in \mathbb{R}^n \Leftrightarrow \exists Q_1, Q_2 \succeq 0 :$$
$$f(\boldsymbol{x}) v_d(\boldsymbol{x})^t Q_1 v_d(\boldsymbol{x}) = v_{d+m}(\boldsymbol{x})^t Q_2 v_{d+m}(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathbb{R}^n$$

donde Q_1, Q_2 tienen las dimensiones necesarias y $2m$ es el grado de f. Notamos que en cada lado de la igualdad se plantea que los coeficientes de fg y los de h son funciones lineales de las coordenadas de Q_1 y Q_2 respectivamente; y la igualdad determina ecuaciones lineales que estas coordenadas deben cumplir. La Proposición 1.2.7 nos permite, eligiendo a priori un $d \in \mathbb{N}$ fijo, plantear esto como un problema SDP de factibilidad.

El problema con este método es que el tamaño de las cotas establecidas para los grados de h y g excede a cualquier aplicación práctica. La cota que se conoce hoy en día para d es $2^{2^{2^{(2m)^{4^n}}}}$, tomada de [14, Teorema 1.4.4].

Vamos a dar un paso más, directamente vinculado con los certificados de positividad de la próxima sección y con los algoritmos para el problema de momentos que veremos en el Capítulo 4.

Proposición 2.1.10 *Sean $f, g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$. Dados los problemas:*

1. *Verificar si $f \in Q(g_1, \ldots, g_m)$, es decir, si existen $f_0, \ldots, f_m \in \Sigma[\boldsymbol{x}]$ tales que*

$$f = f_0 + \sum_{i=1}^{m} f_i g_i$$

2. *Verificar si $f \in P(g_1, \ldots, g_m)$, es decir, si existen $f_J \in \Sigma[\boldsymbol{x}]$ para cada $J \subset \{1, \ldots, m\}$, con $g_J = \prod_{i \in J} g_i, g_\emptyset \equiv 1$, tales que*

$$f = \sum_{J \subset \{1,\ldots,m\}} f_J g_J$$

Si todas las funciones $f_i, i = 0, \ldots, m$ en el item 1 y todas las $f_J, J \subset \{1, \ldots, m\}$ en el item 2 tienen cotas a priori en sus grados, entonces los dos items representan problemas SDP.

Demostración: Lo hacemos para $Q(g_1, \ldots, g_m)$. Por el Teorema 2.1.7 sabemos que esto equivale a ver si existen $Q_0, \ldots, Q_m \succeq 0$, cada una de dimensión $s(p_i)$ dada por la cota $2p_i$ del grado de las f_i, tales que

$$f(\boldsymbol{x}) = v_{p_0}(\boldsymbol{x})^t Q_0 v_{p_0}(\boldsymbol{x}) + \sum_{i=1}^{m} g_i(\boldsymbol{x})(v_{p_i}(\boldsymbol{x})^t Q_i v_{p_i}(\boldsymbol{x})), \forall \boldsymbol{x} \in \mathbb{R}^n$$

Esto se puede plantear como una igualdad de coeficiente a coeficiente, donde f tiene los coeficientes dados y se le agregan coeficientes nulos hasta grado

$$p = \max_{i=0,\dots,m} \deg(g_i v_{p_i}^t Q_i v_{p_i}), g_0 \equiv 1$$

De esta forma obtenemos un sistema de ecuaciones lineal con los coeficientes de las matrices Q_i como incógnitas. Reunimos todas las matrices Q_i en

$$Q = \begin{pmatrix} Q_0 & 0 & \cdots & 0 \\ 0 & Q_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_m \end{pmatrix}$$

y así nos queda un problema de buscar $Q \succeq 0$ (equivalente por el Lema 1.2.6) donde las coordenadas nulas (los ceros que quedan fuera de los bloques Q_i) se representan con $\langle Q, E_{i,j} \rangle = 0$, donde $E_{i,j}$ es una matriz canónica; y la ecuación de cada coeficiente de f se representa como $f_\alpha = \langle Q, C_\alpha \rangle, \alpha \in \mathbb{N}_p^n$, con matrices C_α adecuadas que se construyen utilizando los coeficientes de las funciones $g_i, i = 0, \dots, m$. Resulta entonces una versión de (1.7) como problema de factibilidad, que por lo tanto es un problema SDP de factibilidad por la Proposición 1.2.10. El razonamiento para $P(g_1, \dots, g_m)$ es análogo, solamente en lugar de las matrices Q_i definimos una matriz Q_J por cada $J \subset \{1, \dots, m\}$. □

2.2. Certificados de positividad y teoremas de representación

A continuación nos dedicamos a examinar los certificados de positividad que vamos a utilizar en los algoritmos de resolución del GMP: el de Schmüdgen y el de Putinar, y en qué se diferencian sus contextos de aplicación. Para eso nos va a interesar estudiar la positividad de polinomios en conjuntos semialgebraicos básicos compactos. Como el dominio donde la función debe ser positiva se restringe, vamos a encontrar condiciones suficientes más laxas que ser suma de cuadrados o cociente de sumas de cuadrados. Recordamos que se define $P(f_1, \dots, f_m)$ como en (1.2) y comenzamos citando el siguiente resultado, tomado de [21, Teorema 4].

Teorema 2.2.1 ***(Certificado de Positividad de Stengle)*** *Sean $\mathbb{K}$ un cuerpo cerrado real, $f_j \in \mathbb{K}[\boldsymbol{x}]$ para $j = 1, \dots, m$ y $K = \{x \in \mathbb{K}^n : f_j(x) \geq 0, j = 1, \dots, m\}$. Entonces, $f \geq 0$ en K si y sólo si existen $p \in \mathbb{N}$ y g, h en $P(f_1, \dots, f_m)$ dado por (1.2) tales que $fg = f^{2p} + h$.*

Podemos ver que el teorema anterior no le pide compacidad a K, y se podría construir verificaciones mediante SDP ya que existen cotas para p y los grados de g, h. Sin embargo, así como ocurre con la verificación de positividad global de un polinomio, estas cotas son demasiado grandes. Por eso se suele buscar soluciones fijando cotas a priori [11, Sección 2.4]. Para buscar formulaciones con las que se pueda trabajar mejor vamos a imponer algunas condiciones extra, por ejemplo en el siguiente resultado, del cual se da una prueba en [20, Teorema 12.24].

Teorema 2.2.2 ***(Certificado de Positividad de Schmüdgen)*** *Sean $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ tales que $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_j(\boldsymbol{x}) \geq 0, j = 1, \ldots, m\}$ es compacto. Si $f \in \mathbb{R}[\boldsymbol{x}]$ es estrictamente positiva en K, entonces $f \in P(g_1, \ldots, g_m)$.*

Notamos que en este caso pasamos de un cuerpo cerrado real genérico a $\mathbb{R}$, y que pasamos a pedirle compacidad al conjunto semialgebraico básico, además de que sólo obtenemos un certificado de positividad estricta. A cambio conseguimos una representación que, como ya vimos, se puede buscar mediante un problema SDP. Pero surge otro problema: el número de sumandos es exponencial en la cantidad de funciones que definen K.

Podemos obtener una cantidad de términos lineal en m suponiendo algo más. Dadas $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$, recordamos que se define el módulo cuadrático $Q(g_1, \ldots, g_m)$ como en (1.3).

Definición 2.2.3 *Dadas $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$, $Q(g_1, \ldots, g_m)$ se dice* ***arquimediano*** *si existe $N \in \mathbb{N}_{>0}$ tal que $h \in Q(g_1, \ldots, g_m)$, donde $h(\boldsymbol{x}) = N - ||\boldsymbol{x}||_2^2$.*

Proposición 2.2.4 *Sean $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ y $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, \forall i = 1, \ldots, m\}$. Si $Q = Q(g_1, \ldots, g_m)$ es arquimediano, entonces K es compacto.*

Demostración: Sea $h \in Q, h(\boldsymbol{x}) = N - ||\boldsymbol{x}||_2^2$ para algún $N \in \mathbb{N}_{>0}$. Luego,

$$h(\boldsymbol{x}) = N - ||\boldsymbol{x}||_2^2 = f_0(\boldsymbol{x}) + \sum_{i=1}^{m} f_i(\boldsymbol{x}) g_i(\boldsymbol{x}) \text{ con } f_i \in \Sigma[\boldsymbol{x}], i = 0, \ldots, m$$

De esta formulación se sigue que si $\boldsymbol{x} \in K$, entonces $h(\boldsymbol{x}) \geq 0$. Luego, $K \subset \bar{B}(0, \sqrt{N})$, y como K es cerrado resulta compacto.

□ Procedemos a dar una caracterización de los módulos cuadráticos arquimedianos. Este resultado se enuncia en [11, Teorema 2.15].

Teorema 2.2.5 *Sean $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$. Entonces, las siguientes afirmaciones son equivalentes:*

1. *$Q = Q(g_1, \ldots, g_m)$ es arquimediano.*
2. *Existe $u \in Q$ tal que $\{\boldsymbol{x} \in \mathbb{R}^n : u(\boldsymbol{x}) \geq 0\}$ es compacto.*
3. *Existen $p_1, \ldots, p_k \in \mathbb{R}[\boldsymbol{x}]$ tales que $K = \{\boldsymbol{x} \in \mathbb{R}^n : p_j(\boldsymbol{x}) \geq 0, j = 1, \ldots, k\}$ es compacto y además $P(p_1, \ldots, p_k) \subseteq Q$.*
4. *Para todo $p \in \mathbb{R}[\boldsymbol{x}]$, existe $N \in \mathbb{N}_{>0}$ tal que $N + p, N - p \in Q$.*

Demostración: 4. $\Rightarrow$ 1) es inmediato tomando $p(\boldsymbol{x}) = ||\boldsymbol{x}||_2^2$.
1. $\Rightarrow$ 2) $u(\boldsymbol{x}) = N - ||\boldsymbol{x}||_2^2$ dado por la definición de arquimedianidad tiene como conjunto de no negatividad a $\bar{B}(0, \sqrt{N})$, que es compacto, y por hipótesis $u \in Q$.
2. $\Rightarrow$ 3) Alcanza elegir u como única función p_1. K definido de esta forma es compacto por hipótesis y además, dado $h \in P(u)$, se tiene que $h = f_0 + uf_1$, donde $f_0, f_1 \in \Sigma[\boldsymbol{x}]$, y por lo tanto $h \in Q$.
3. $\Rightarrow$ 4) Dado un polinomio p, como K es compacto, p alcanza imágenes mínima y máxima J_1, J_2 en K. Luego, $-p$ tiene imagen mínima $-J_2$ y máxima $-J_1$ en K. Eligiendo $N \in \mathbb{N}_{>0}, N > \text{máx}\{|J_1|, |J_2|\}$, $N + p, N - p$ resultan estrictamente positivos en K. Por el Teorema 2.2.2, se sigue que $N + p, N - p \in P(p_1, \ldots, p_k) \subseteq Q$, y listo. $\square$

Ahora estamos en condiciones de dar el último resultado de esta sección. Se obtuvo originalmente en [19, Teorema 1.4]. Se enuncia en [11, Teorema 2.14] y se da una prueba después del Teorema 2.16 del mismo libro. Esta demostración se puede seguir con definiciones y resultados ya descriptos en este trabajo.

Teorema 2.2.6 ***(Certificado de Positividad de Putinar)*** *Sean $g_1, \ldots, g_m$ en $\mathbb{R}[\boldsymbol{x}]$ tales que $Q = Q(g_1, \ldots, g_m)$ es arquimediano. Si $f \in \mathbb{R}[\boldsymbol{x}]$ es estrictamente positivo en $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, \forall i = 1, \ldots, m\}$, entonces $f \in Q$.*

Vemos que ahora la cantidad de sumandos es lineal, lo que mejora mucho la complejidad de los programas destinados a verificar positividad. Por eso va a ser el certificado que más usemos en lo que sigue.

Por otra parte, como ya se observó en el Capítulo 1, tenemos que $Q(g_1, \ldots, g_m) \subseteq P(g_1, \ldots, g_m)$. Es decir, el Certificado de positividad de Putinar ubica f estrictamente positiva en K en un subconjunto del conjunto donde está f según el certificado de Schmüdgen. Así que, pidiendo arquimedianidad, el último certificado mejora el anterior.

Capítulo 3

Problemas de momentos, sucesiones y medidas representativas

3.1. Problema de Momentos Generalizado y teoría de dualidad

El Problema de Momentos Generalizado (GMP, por su denominación en inglés) es un problema de optimización en Teoría de la Medida. Tiene muchas aplicaciones, y en esta tesis nos vamos a centrar en la optimización de polinomios. Es decir, dado un polinomio $f \in \mathbb{R}[\boldsymbol{x}]$, vamos a buscar un mínimo valor de f en un dominio $K \subset \mathbb{R}^n$ y, si es posible, vamos a encontrar uno o varios puntos que lo realizan. En este trabajo llamaremos medida boreliana en K a una medida que se aplica a subconjuntos de K, tal que todos los subconjuntos borelianos son medibles. El Problema de Momentos Generalizado se define como sigue:

Definición 3.1.1 *Sean K un subconjunto boreliano de $\mathbb{R}^n$, Γ un conjunto de índices y funciones $f, h_j : K \to \mathbb{R}, j \in \Gamma$. Sea $\{\gamma_j\}_{j\in\Gamma} \subset \mathbb{R}$ un conjunto de números. Se define $\mathcal{M}(K)$ como la familia de medidas borelianas, no negativas y finitas sobre K. Entonces, se define el **Problema de Momentos Generalizado** como encontrar $\rho_{mom} \in \mathbb{R}$ dado de esta forma:*

$$\begin{aligned} \textit{GMP: } \rho_{mom} = & \sup_{\mu\in\mathcal{M}(K)} \int_K f d\mu \\ & \textit{sujeto a } \int_K h_j d\mu \leq \gamma_j, \forall j \in \Gamma \end{aligned} \tag{3.1}$$

*En este problema, una **solución factible** es una medida $\mu \in \mathcal{M}(K)$ que cumple las restricciones enunciadas para todos los $j \in \Gamma$. Para una solución factible μ, $\int_K f d\mu \in \mathbb{R}$ es un **valor factible**. ρ_{mom}, si es finito, se dice **valor óptimo**. Una solución factible*

*$\mu^* \in \mathcal{M}(K)$ que realiza el valor óptimo se llama **solución óptima**. El problema se dice **factible** si tiene al menos una solución factible.*

Como vemos, se busca una medida no negativa finita que maximice la integral y este valor máximo, con un conjunto de restricciones. Notar que algunas restricciones pueden ser igualdades. En efecto, supongamos que existen $j \in \Gamma, \gamma_j \in \mathbb{R}$ tales que se requiere $\int_K h_j d\mu = \gamma_j$. Esto se puede expresar como dos restricciones con el formato de la Definición 3.1.1: si definimos $h_{j_1} = h_j, h_{j_2} = -h_j$, la restricción de igualdad equivale a

$$\int_K h_{j_1} d\mu \leq \gamma_j \wedge \int_K h_{j_2} d\mu \leq -\gamma_j,$$

es decir que reemplazamos una restricción de igualdad por dos restricciones de desigualdad. Estudiamos el GMP y sus propiedades asumiendo Γ contable, pero los casos de aplicación que veremos van a tener una cantidad finita de restricciones; esto es, Γ será finito.

Vamos a formular a continuación el problema dual y estudiar resultados de dualidad:

Definición 3.1.2 *Dado el problema (3.1), se define su problema dual como encontrar ρ_{pop} dado por*

$$\begin{aligned} \rho_{pop} = & \inf_{\lambda \in \mathbb{R}^\Gamma} \sum_{j \in \Gamma} \lambda_j \gamma_j \\ \text{sujeto a } & \sum_{j \in \Gamma} \lambda_j h_j(\boldsymbol{x}) \geq f(\boldsymbol{x}), \forall \boldsymbol{x} \in K \\ & \lambda_j \geq 0, \forall j \in \Gamma_+ \end{aligned} \tag{3.2}$$

*donde el ínfimo se define sobre los elementos $\lambda \in \mathbb{R}^\Gamma$ y $\Gamma_+ \subset \Gamma$ es el subconjunto de índices para cuyas funciones h_j hay restricción de desigualdad y no se puede deducir una de igualdad. Un elemento $\lambda \in \mathbb{R}^\Gamma$ que cumple las restricciones enunciadas se llama **solución factible**. La sumatoria o serie $\sum_{j \in \Gamma} \lambda_j \gamma_j$ para una solución factible $\lambda \in \mathbb{R}^\Gamma$ se llama **valor factible**. ρ_{pop}, si es finito, se llama **valor óptimo**, y una solución factible que lo realiza se llama **solución óptima**. El problema se dice **factible** si tiene al menos una solución factible.*

Vamos a ver un ejemplo de problemas primal y dual: sean las funciones en $\mathbb{R}[x]$ dadas por $f(x) = x^2 - 3x + 1$, $\Gamma = \mathbb{N}_{>0}$, $\{\gamma_j\}_{j \in \Gamma}$ dado por $\gamma_j = j$, y $h_j(x) = x^j$. Entonces, el problema primal queda definido por

$$\begin{aligned} \rho_{mom} = & \sup_{\mu \in \mathcal{M}(K)} \int_K (x^2 - 3x + 1) d\mu \\ \text{sujeto a } & \int_K x^j d\mu \leq j, \forall j \in \mathbb{N}_{>0} \end{aligned}$$

y su correspondiente dual queda formulado de la forma que sigue. En este caso, $\Gamma_+ = \Gamma$ porque todos los h_j tienen restricciones sobre ser menores o iguales a valores γ_j y no mayores o iguales (lo que redundaría en tener restricciones de igualdad).

$$\begin{aligned} \rho_{pop} = & \inf_{\lambda \in \mathbb{R}^{\mathbb{N}_{>0}}} \sum_{j \in \mathbb{N}_{>0}} \lambda_j . j \\ \text{sujeto a } & \sum_{j \in \mathbb{N}_{>0}} \lambda_j x^j \geq x^2 - 3x + 1, \forall x \in K \\ & \lambda_j \geq 0, \forall j \in \mathbb{N}_{>0} \end{aligned}$$

Seguimos esta sección con los resultados de dualidad que vamos a utilizar, comenzando por la **dualidad débil**.

Teorema 3.1.3 *Sean z_0 un valor factible del problema (3.1) y λ una solución factible del problema (3.2). Entonces, se tiene $z_0 \leq \sum_{j \in \Gamma} \lambda_j \gamma_j$. En particular, $\rho_{mom} \leq \rho_{pop}$.*

Demostración: Por definición de z_0, para alguna medida μ boreliana, no negativa y finita en K se tiene

$$z_0 = \int_K f d\mu \leq \int_K \left(\sum_{j \in \Gamma} \lambda_j h_j \right) d\mu = \sum_{j \in \Gamma} \lambda_j \int_K h_j d\mu \leq \sum_{j \in \Gamma} \lambda_j \gamma_j$$

donde la primera desigualdad sale de la restricción del problema dual, la siguiente igualdad es por linealidad de la integral y la última desigualdad es por las restricciones del problema primal. □

Continuamos ahora con la dualidad fuerte, enunciando un resultado tomado parcialmente de [11, Teorema 1.3].

Teorema 3.1.4 *Sean $K \subset \mathbb{R}^n$ compacto, f acotada y semicontinua superiormente en K, h_j continua en K para todo $j \in \Gamma$. Suponemos que además hay $k \in \Gamma$ tal que $h_k(\boldsymbol{x}) > 0$ para todo $\boldsymbol{x} \in K$. Entonces, $\rho_{mom} = \rho_{pop}$. Además, si (3.1) tiene una solución factible entonces tiene una solución óptima.*

Observamos que, en los casos que nos interesan, las funciones f y $h_j, j \in \Gamma$, van a ser polinomios. Luego, las hipótesis de acotación en K compacto y continuidad siempre se van a cumplir.

Procedemos a definir casos particulares de GMP de interés para este trabajo.

Definición 3.1.5 *Dada una medida boreliana finita μ sobre $K \subset \mathbb{R}^n$, se define su **soporte** como $\mathbb{R}^n \setminus G$, donde G es la unión de todos los conjuntos abiertos que miden 0 en μ.*

Definición 3.1.6 *Sean $g_1, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$, una sucesión $\{y_\alpha\}_{\alpha \in \mathbb{N}^n}$ y el conjunto semialgebraico básico cerrado $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, i = 1, \ldots, m\}$. Se define el* ***problema de momentos completo*** *como la pregunta: ¿existe una medida μ boreliana, no negativa y finita con soporte contenido en K que cumpla*

$$y_\alpha = \int_K \boldsymbol{x}^\alpha d\mu, \forall \alpha \in \mathbb{N}^n? \tag{3.3}$$

A la misma pregunta, tomando sólo los α con $|\alpha| \leq k \in \mathbb{N}$, se la llama ***problema de momentos truncado****:*

$$y_\alpha = \int_K \boldsymbol{x}^\alpha d\mu, \forall \alpha \in \mathbb{N}^n_k \tag{3.4}$$

Notamos que ambos problemas son casos particulares del problema (3.1), donde los γ_j son los y_α (todos con restricciones de igualdad), los h_j son los monomios de n variables, pero no se pide un valor óptimo sino sólo factibilidad.

Definición 3.1.7 *Tanto en el problema completo como en el truncado, una medida μ que cumpla todas las igualdades se llama* ***medida representativa*** *de la secuencia finita o infinita $\{y_\alpha\}$. Si la medida es única, se dice* ***determinada****, e* ***indeterminada*** *en caso contrario.*

Definición 3.1.8 *Dado $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n}$, se define el funcional $L_{\boldsymbol{y}} : \mathbb{R}[\boldsymbol{x}] \to \mathbb{R}$ como sigue: dado $f \in \mathbb{R}[\boldsymbol{x}], f(\boldsymbol{x}) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha \boldsymbol{x}^\alpha$, formamos*

$$L_{\boldsymbol{y}}(f) = \sum_{\alpha \in \mathbb{N}^n} y_\alpha f_\alpha \tag{3.5}$$

Si en cambio $\{y_\alpha\}$ es una secuencia finita, el funcional se define extendiendo la secuencia con $y_\beta = 0$ para todos los índices restantes $\beta \in \mathbb{N}^n$ y recurriendo al caso anterior.

Se puede comprobar de forma inmediata que $L_{\boldsymbol{y}}$ es lineal. Además está bien definido: como f es un polinomio, tiene finitos coeficientes no nulos y por lo tanto la serie converge. El funcional es la sumatoria, para todos los coeficientes del polinomio f, de esos coeficientes multiplicados por los elementos de igual índice de la secuencia de momentos y. El siguiente teorema viene a dar una vinculación entre las medidas representativas y los polinomios no negativos en un conjunto. En [8, Teorema] (no tiene número) se da una demostración para 2 dimensiones por conveniencia para la notación, pero puede generalizarse a n variables.

Teorema 3.1.9 ***(Riesz-Haviland)*** *Dados $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n}$, $K \subset \mathbb{R}^n$ cerrado, la existencia de una medida representativa $\mu \in \mathcal{M}(K)$ para la secuencia $\boldsymbol{y}$ es equivalente a tener $L_{\boldsymbol{y}}(f) \geq 0$ para todo polinomio f no negativo en K.*

3.2. Matriz de momentos y extensiones

Recordamos que, para ciertos $n, d \in \mathbb{N}, n \geq 1$, se define

$$v_d(\boldsymbol{x}) = (1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1^d, \ldots, x_n^d)$$

con orden de los monomios dado por aplicar *ord* a sus exponentes. Por ejemplo, si $n = d = 2$, tenemos $v_2(x_1, x_2) = (1, x_1, x_2, x_1^2, x_1x_2, x_2^2)$. Luego, un polinomio en $\mathbb{R}^n$ de grado a lo sumo d dado por $p(x) = \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \boldsymbol{x}^\alpha$ puede escribirse como $p^t v_d(\boldsymbol{x})$, donde p es el vector de coeficientes del polinomio indexados por *ord*. También vimos antes que $v_d(\boldsymbol{x})$ tiene longitud $s(d)$, con $s(d) = \binom{n+d}{d}$.

A continuación definiremos la matriz de momentos.

Definición 3.2.1 *Dados $n, d \in \mathbb{N}, n \geq 1$, sea $\boldsymbol{y} = \{y_\gamma\}_{\gamma \in \mathbb{N}^n_{2d}}$. Se define en $\mathbb{R}^{s(d)\times s(d)}$ la **matriz de momentos** de $\boldsymbol{y}$, construida de la siguiente forma: dados $\alpha, \beta \in \mathbb{N}^n_d$, se define $M_d(\boldsymbol{y})_{\alpha,\beta} = L_{\boldsymbol{y}}(\boldsymbol{x}^{\alpha+\beta})$, con $L_{\boldsymbol{y}}$ el funcional definido en (3.5). Las filas y columnas de la matriz tienen sus índices ordenados de acuerdo a ord.*

Observar que $M_d(\boldsymbol{y})$ es una matriz simétrica ya que la suma de índices de $\mathbb{N}^n$ es conmutativa.

Proposición 3.2.2 *Sea $L_{\boldsymbol{y}}$ el funcional definido en (3.5) con $\boldsymbol{y} = \{y_\gamma\}_{\gamma \in \mathbb{N}^n_{2d}}$. Entonces, dados $\gamma \in \mathbb{N}^n_{2d}$; $p, q \in \mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo d y $\tilde{p}, \tilde{q}$ sus vectores de coeficientes indexados por la función ord, se tiene:*

1. $L_{\boldsymbol{y}}(\boldsymbol{x}^\gamma) = y_\gamma$
2. $L_{\boldsymbol{y}}(pq) = \tilde{p}^t M_d(\boldsymbol{y})\tilde{q}$

Demostración: El primer item es inmediato por la definición del funcional $L_{\boldsymbol{y}}$. Para el segundo, desarrollamos el producto matricial por definición. Si tenemos $z = M_d(\boldsymbol{y})\tilde{q}$, para $\alpha \in \mathbb{N}^n_d$ se sigue

$$z_\alpha = \sum_{\beta \in \mathbb{N}^n_d} M_d(\boldsymbol{y})_{\alpha,\beta}\tilde{q}_\beta = \sum_{\beta \in \mathbb{N}^n_d} y_{\alpha+\beta}\tilde{q}_\beta$$

Desarrollando $\tilde{p}^t z$ de forma similar, queda

$$\tilde{p}^t z = \sum_{\alpha,\beta \in \mathbb{N}^n_d} \tilde{p}_\alpha \tilde{q}_\beta y_{\alpha+\beta}$$

Es decir, cada elemento $y_\gamma, \gamma \in \mathbb{N}^n_{2d}$, se multiplica por todos los posibles productos de coeficientes de p, q tales que la suma de sus índices respectivos da γ. Pero esta suma

corresponde al coeficiente de pq con índice γ. Luego, la sumatoria completa que obtuvimos es $L_{\boldsymbol{y}}(pq)$. □

Como ejemplo de matriz de momentos, vamos a construir $M_1(\boldsymbol{y})$ con $n = 2$. Recordamos que las tuplas α indexadas por *ord* con $|\alpha| \leq 1$ son $(0,0), (1,0), (0,1)$.

$$M_1(\boldsymbol{y}) = \begin{pmatrix} y_{(0,0)+(0,0)} & y_{(0,0)+(1,0)} & y_{(0,0)+(0,1)} \\ y_{(1,0)+(0,0)} & y_{(1,0)+(1,0)} & y_{(1,0)+(0,1)} \\ y_{(0,1)+(0,0)} & y_{(0,1)+(1,0)} & y_{(0,1)+(0,1)} \end{pmatrix} = \begin{pmatrix} y_{(0,0)} & y_{(1,0)} & y_{(0,1)} \\ y_{(1,0)} & y_{(2,0)} & y_{(1,1)} \\ y_{(0,1)} & y_{(1,1)} & y_{(0,2)} \end{pmatrix}$$

Observamos que para construir $M_d(\boldsymbol{y})$ se necesita términos de $\boldsymbol{y}$ hasta orden $2d$, porque es el mayor grado posible de un $\boldsymbol{x}^{\alpha+\beta}, \alpha, \beta \in \mathbb{N}^n_d$. Ahora comenzamos a conectar las medidas representativas con la matriz de momentos de forma útil para abordar problemas mediante programación semidefinida.

Proposición 3.2.3 *Sea $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2d}} \subset \mathbb{R}$. Si $\boldsymbol{y}$ tiene una medida μ representativa en $K \subset \mathbb{R}^n$ (problema truncado), entonces para todo polinomio $p \in \mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo $2d$ se tiene $L_{\boldsymbol{y}}(p) = \int_K p d\mu$. Lo mismo vale para todo $q \in \mathbb{R}[\boldsymbol{x}]$ si μ es una medida representativa de $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n}$ (problema completo).*

Demostración: Desarrollando, tenemos $L_{\boldsymbol{y}}(p) = L_{\boldsymbol{y}}(\sum_{\alpha \in \mathbb{N}^n_{2d}} h_\alpha \boldsymbol{x}^\alpha) = \sum_{\alpha \in \mathbb{N}^n_{2d}} h_\alpha y_\alpha = \sum_{\alpha \in \mathbb{N}^n_{2d}} h_\alpha \int_K \boldsymbol{x}^\alpha d\mu = \sum_{\alpha \in \mathbb{N}^n_{2d}} \int_K h_\alpha \boldsymbol{x}^\alpha d\mu = \int_K p d\mu$, y vale un razonamiento análogo para $q \in \mathbb{R}[\boldsymbol{x}]$ en el caso del problema completo. □

Corolario 3.2.4 *Sea $\{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2d}} \subset \mathbb{R}$. Si $\boldsymbol{y}$ tiene una medida representativa en $K \subset \mathbb{R}^n$ (problema truncado), entonces $M_d(\boldsymbol{y}) \succeq 0$.*

Demostración: Dado $q \in \mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo d, si $\tilde{q}$ es su vector de coeficientes indexados por *ord*, por las Proposiciones 3.2.2 y 3.2.3 tenemos

$$\tilde{q}^t M_d(\boldsymbol{y}) \tilde{q} = L_{\boldsymbol{y}}(q^2) = \int_K q^2 d\mu \geq 0$$

donde la última desigualdad vale porque μ es una medida no negativa por definición de medida representativa y q^2 es no negativo. Como q era arbitrario, se sigue que $M_d(\boldsymbol{y}) \succeq 0$. □

Hablamos ahora de extensiones, un concepto necesario para los algoritmos que vamos a ver más adelante.

Definición 3.2.5 *Sea $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2d}} \subset \mathbb{R}$. Si tenemos $M_d(\boldsymbol{y}) \succeq 0$, se define el **problema de extensión de momentos** como: extender la secuencia con nuevos valores $y_\kappa, 2d < |\kappa| \leq 2(d+1)$ de modo que $M_{d+1}(\boldsymbol{y}) \succeq 0$. Si esta extensión existe, se llama **extensión positiva**. Si además verifica $\operatorname{rg}(M_d(\boldsymbol{y})) = \operatorname{rg}(M_{d+1}(\boldsymbol{y}))$, se llama **extensión plana**.*

Definición 3.2.6 *Dado $k \in \mathbb{N}_{>0}$, una combinación lineal positiva de k medidas de Dirac se llama* ***medida* k$-$*atómica***. *Los puntos donde se concentran estas medidas de Dirac se llaman* ***átomos***.

Como las medidas $k-$atómicas son combinaciones lineales finitas de medidas de Dirac, en particular están en $\mathcal{M}(K)$.

Cerramos esta sección con una forma computable de caracterizar la existencia de medidas representativas atómicas. Se da una demostración en [5, Teorema 2.19].

Teorema 3.2.7 ***(de la Extensión Plana)*** *Sea $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}_{2d}} \subset \mathbb{R}$. Entonces, esta secuencia admite una medida representativa $\mathrm{rg}(M_d(\boldsymbol{y}))$-atómica en $\mathbb{R}^n$ si y sólo si $M_d(\boldsymbol{y}) \succeq 0$ y además la secuencia $\boldsymbol{y}$ admite una extensión plana.*

3.3. Dominio semialgebraico y matriz localizadora

En la sección anterior vimos medidas representativas en $\mathbb{R}^n$ y formas de comprobar su existencia. Pasamos ahora a considerar algunos dominios restringidos, en este caso por la región de no negatividad de un polinomio.

Definición 3.3.1 *Sean $u \in \mathbb{R}[\boldsymbol{x}]$ y la secuencia $\boldsymbol{y} = \{y_\gamma\}_{\gamma \in \mathbb{N}^n_{2d+\deg(u)}} \subset \mathbb{R}$ indexada por ord, donde $\deg(u)$ es el grado de u. Se define en $\mathbb{R}^{s(d)\times s(d)}$ la* ***matriz localizadora*** *de $u, \boldsymbol{y}$ como*

$$M_d(u, \boldsymbol{y})_{\alpha,\beta} = L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{\alpha+\beta}), \forall \alpha, \beta \in \mathbb{N}^n_d \tag{3.6}$$

Las filas y columnas de la matriz tienen sus índices ordenados de acuerdo a ord.

Es decir, la matriz localizadora se construye multiplicando los monomios por el polinomio u y aplicando el funcional a cada uno de estos productos. La multiplicación por u aumentará los grados de los monomios, generando cambios de lugar de los términos de y en la matriz como vamos a ver en la siguiente proposición. Notamos que efectivamente para construir esta matriz alcanza con términos de la secuencia $\boldsymbol{y}$ a lo sumo hasta orden $2d + \deg(u)$, ya que este es el mayor grado posible de un $u(\boldsymbol{x})\boldsymbol{x}^{\alpha+\beta}, \alpha, \beta \in \mathbb{N}^n_d$.

Nuevamente, vemos que la matriz localizadora es simétrica. Notar que si $u \equiv 1$, la matriz localizadora es la de momentos.

Proposición 3.3.2 *Con los elementos de la definición anterior, donde los u_γ son los coeficientes de u indexados por ord, se tiene*

$$M_d(u, \boldsymbol{y})_{\alpha,\beta} = \sum_{\gamma \in \mathbb{N}^n} u_\gamma y_{\gamma+\alpha+\beta}$$

Demostración: Sale por desarrollo directo y definición de $L_{\boldsymbol{y}}$:

$$u(\boldsymbol{x})\boldsymbol{x}^{\alpha+\beta} = \sum_{\gamma\in\mathbb{N}^n} u_\gamma x^{\gamma+\alpha+\beta} \Rightarrow L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{\alpha+\beta}) = M_d(u,\boldsymbol{y})_{\alpha,\beta} = \sum_{\gamma\in\mathbb{N}^n} u_\gamma y_{\gamma+\alpha+\beta}$$

□

Por ejemplo, dados $n = 2, u(x_1, x_2) = 4 - 3x_1$, vamos a construir $M_1(u, \boldsymbol{y})$. Nuevamente, las tuplas α indexadas por *ord* con $|\alpha| \leq 1$ son: $(0,0), (1,0), (0,1)$.

$$M_1(u,\boldsymbol{y}) = \begin{pmatrix} L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(0,0)+(0,0)}) & L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(0,0)+(1,0)}) & L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(0,0)+(0,1)}) \\ L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(1,0)+(0,0)}) & L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(1,0)+(1,0)}) & L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(1,0)+(0,1)}) \\ L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(0,1)+(0,0)}) & L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(0,1)+(1,0)}) & L_{\boldsymbol{y}}(u(\boldsymbol{x})\boldsymbol{x}^{(0,1)+(0,1)}) \end{pmatrix}$$

$$= \begin{pmatrix} L_{\boldsymbol{y}}(4-3x_1) & L_{\boldsymbol{y}}((4-3x_1)x_1) & L_{\boldsymbol{y}}((4-3x_1)x_2) \\ L_{\boldsymbol{y}}((4-3x_1)x_1) & L_{\boldsymbol{y}}((4-3x_1)x_1^2) & L_{\boldsymbol{y}}((4-3x_1)x_1x_2) \\ L_{\boldsymbol{y}}((4-3x_1)x_2) & L_{\boldsymbol{y}}((4-3x_1)x_1x_2) & L_{\boldsymbol{y}}((4-3x_1)x_2^2) \end{pmatrix}$$

$$= \begin{pmatrix} 4y_{(0,0)} - 3y_{(1,0)} & 4y_{(1,0)} - 3y_{(2,0)} & 4y_{(0,1)} - 3y_{(1,1)} \\ 4y_{(1,0)} - 3y_{(2,0)} & 4y_{(2,0)} - 3y_{(3,0)} & 4y_{(1,1)} - 3y_{(2,1)} \\ 4y_{(0,1)} - 3y_{(1,1)} & 4y_{(1,1)} - 3y_{(2,1)} & 4y_{(0,2)} - 3y_{(1,2)} \end{pmatrix}$$

También vamos a dar condiciones suficientes para que la matriz localizadora sea semidefinida positiva. Esto se usará para minimizar polinomios con restricciones de dominio.

Proposición 3.3.3 *Dados $p, q \in \mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo d con $\tilde{p}, \tilde{q}$ sus vectores de coeficientes indexados por ord, la matriz localizadora verifica $L_{\boldsymbol{y}}(upq) = \tilde{p}^t M_d(u,\boldsymbol{y})\tilde{q}$.*

Demostración: Por desarrollo directo del producto matricial, se verifica que por cada término y_σ, el valor por el que se multiplica es la suma de todos los productos $\tilde{p}_\alpha \tilde{q}_\beta u_\gamma$ tales que $\alpha + \beta + \gamma = \sigma$. Esto coincide con el coeficiente de índice σ de upq. Como al desarrollar el producto matricial quedan sumados estos valores para todos los σ, el resultado coincide con $L_{\boldsymbol{y}}(upq)$. □

Corolario 3.3.4 *Dados $u \in \mathbb{R}[\boldsymbol{x}], \boldsymbol{y} = \{y_\gamma\}_{\gamma\in\mathbb{N}^n_{2d+\deg(u)}} \subset \mathbb{R}$, si $\boldsymbol{y}$ tiene una medida representativa μ con soporte contenido en $K \subseteq \{\boldsymbol{x} \in \mathbb{R}^n : u(\boldsymbol{x}) \geq 0\}$, entonces $M_d(u,\boldsymbol{y}) \succeq 0$.*

Demostración: Sean $q \in \mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo d y $\tilde{q}$ su vector de coeficientes indexado por *ord*. Tenemos $\tilde{q}^t M_d(u,\boldsymbol{y})\tilde{q} = L_{\boldsymbol{y}}(uq^2) = \int_K uq^2 d\mu$, donde la primera igualdad sale de la Proposición anterior y la segunda de la Proposición 3.2.3. Como μ es una medida no negativa y su soporte está contenido en K dado en la hipótesis, esta integral es no negativa. Como q era arbitrario, se sigue que $M_d(u,\boldsymbol{y}) \succeq 0$. □

A continuación damos dos resultados que utilizan los certificados de positividad de Schmüdgen y Putinar para dar propiedades equivalentes a la existencia de una medida representativa para una secuencia de momentos $\boldsymbol{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ (solución del problema de momentos completo). Se demuestran en [11, Teorema 3.8].

Teorema 3.3.5 *Sean $\boldsymbol{y} = \{y_\gamma\}_{\gamma\in\mathbb{N}^n} \subset \mathbb{R}$ una sucesión, sean $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ y $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, i = 1, \ldots, m\}$, al cual suponemos compacto. Para cada $J \subset \{1, \ldots, m\}$, se define $g_J = \prod_{k\in J} g_k$, donde $g_\emptyset \equiv 1$. Entonces, las siguientes afirmaciones son equivalentes:*

1. *$\boldsymbol{y}$ tiene una medida representativa con soporte contenido en K;*
2. *$M_d(g_J, \boldsymbol{y}) \succeq 0, \forall J \subset \{1, \ldots, m\}, \forall d \in \mathbb{N}$;*
3. *$L_{\boldsymbol{y}}(f^2 g_J) \geq 0, \forall J \subset \{1, \ldots, m\}, \forall f \in \mathbb{R}[\boldsymbol{x}]$.*

Demostración: 2. $\Rightarrow$ 3.) Fijando $J \subset \{1, \ldots, m\}$ y $d \in \mathbb{N}$ arbitrario, se obtiene la desigualdad de 3. para todo f de grado a lo sumo d por la Proposición 3.3.3.
3. $\Rightarrow$ 2.) Como la desigualdad vale para todo f, en particular vale para todo f tal que $\deg(f) \leq d \in \mathbb{N}$, y de esto se deduce 2. por la Proposición 3.3.3.
1. $\Rightarrow$ 2.) Sea μ la medida representativa. Dado $J \subset \{1, \ldots, m\}$, como $g_J \geq 0$ en K (o bien $J = \emptyset$ y $g_J = 1 \geq 0$ en $\mathbb{R}^n$), del Corolario 3.3.4 se deduce que $M_d(g_J, \boldsymbol{y}) \succeq 0, \forall d \in \mathbb{N}$ (porque una medida representativa para la sucesión $\{y_\alpha\}$ lo es en particular para los términos con $|\alpha| \leq 2d + \deg(g_J)$).
3. $\Rightarrow$ 1.) Dado $f \in \mathbb{R}[\boldsymbol{x}]$ no negativo en K, vamos a probar que $L_{\boldsymbol{y}}(f) \geq 0$, y el resultado se sigue del Teorema 3.1.9. Para empezar, sea $f > 0$ en K. Entonces, por el Teorema 2.2.2 resulta que $f \in P(g_1, \ldots, g_m)$. Luego, por hipótesis y por linealidad del funcional $L_{\boldsymbol{y}}$, se sigue que $L_{\boldsymbol{y}}(f) \geq 0$.
Ahora, suponemos $f \geq 0$ en K. Entonces, dado $\varepsilon > 0$, resulta $f + \varepsilon > 0$ en $K \Rightarrow L_{\boldsymbol{y}}(f + \varepsilon) = L_{\boldsymbol{y}}(f) + y_0\varepsilon \geq 0$, donde la última desigualdad sale del caso anterior. Como ε es arbitrario, se deduce $L_{\boldsymbol{y}}(f) \geq 0$. □

Teorema 3.3.6 *Sean $\boldsymbol{y} = \{y_\gamma\}_{\gamma\in\mathbb{N}^n} \subset \mathbb{R}$ una sucesión, $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ y $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, i = 1, \ldots, m\}$. Suponemos que $Q(g_1, \ldots, g_m)$ es arquimediano. Entonces, las siguientes afirmaciones son equivalentes:*

1. *$\boldsymbol{y}$ tiene una medida representativa con soporte contenido en K;*
2. *$\forall d \in \mathbb{N} : M_d(g_j, \boldsymbol{y}) \succeq 0, \forall 1 \leq j \leq m, M_d(\boldsymbol{y}) \succeq 0$;*
3. *$L_{\boldsymbol{y}}(f^2) \geq 0, L_{\boldsymbol{y}}(f^2 g_j) \geq 0, \forall 1 \leq j \leq m, \forall f \in \mathbb{R}[\boldsymbol{x}]$.*

Demostración: Análoga al razonamiento de la demostración anterior, sólo que probamos 1. $\Leftrightarrow$ 2.) fijando j entre 1 y m en lugar de $J \subset \{1, \ldots, m\}$; y en 3. $\Rightarrow$ 1.) usamos que $f > 0 \Rightarrow f \in Q(g_1, \ldots, g_m)$. $\square$

Es decir, si suponemos K compacto podemos encontrar una caracterización de la existencia de medida representativa con soporte en K. Cerramos el capítulo con un teorema que da una condición suficiente computable para la existencia de medidas representativas atómicas. Se encuentra una prueba en [12, Teorema 1.6].

Teorema 3.3.7 *Sean* $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$ *y* $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_1(\boldsymbol{x}) \geq 0, \ldots, g_m(\boldsymbol{x}) \geq 0\}$. *Sean* $v_j \in \mathbb{N}$ *tales que* g_j *tiene grado* $2v_j$ *o* $2v_j - 1$ *para cada* $1 \leq j \leq m$. *Sea* $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2d}}$, *y sea* $v = \max_{1 \leq j \leq m} v_j$. *Entonces,* $\boldsymbol{y}$ *tiene una medida representativa* $\operatorname{rg}(M_{d-v}(\boldsymbol{y}))$*-atómica* μ *con soporte contenido en* K *si y sólo si se verifican:*

1. $M_d(\boldsymbol{y}) \succeq 0$ *y* $M_{d-v}(g_j, \boldsymbol{y}) \succeq 0, \forall 1 \leq j \leq m$;
2. $\operatorname{rg}(M_d(\boldsymbol{y})) = \operatorname{rg}(M_{d-v}(\boldsymbol{y}))$.

Además, si se cumplen estas condiciones, μ *es única, y para cada* $1 \leq j \leq m$, μ *tiene* $\operatorname{rg}(M_d(\boldsymbol{y})) - \operatorname{rg}(M_{d-v}(g_j, \boldsymbol{y}))$ *átomos* $x \in \mathbb{R}^n$ *que cumplen* $g_j(x) = 0$.

Notamos que este teorema no pide que K sea compacto. Por otro lado, calcular el rango es sensible a imprecisiones numéricas.

Capítulo 4

Algoritmos para el Problema de Momentos Generalizado

A continuación vamos a estudiar algoritmos para resolver los problemas de momentos vistos en el capítulo anterior. Más adelante vamos a aplicarlos puntualmente a la minimización de polinomios. Vamos a comenzar describiendo el GMP por su secuencia de momentos. Recordamos que buscamos resolver el Problema (3.1). En este caso consideramos Γ finito. Recordamos que $\Gamma_+ \subset \Gamma$ es el subconjunto de índices para los que hay restricción de desigualdad y no se puede deducir una de igualdad. Vamos a reformular (3.1) en términos del funcional lineal $L_{\boldsymbol{y}}$, como se define en (3.5). En el problema descripto a continuación, $f, h_j : K \to \mathbb{R}$ son polinomios, $h_j(\boldsymbol{x}) = \sum_{\alpha \in \mathbb{N}^n} h_{j,\alpha} \boldsymbol{x}^\alpha$. Además, $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_1(\boldsymbol{x}) \geq 0, \ldots, g_m(\boldsymbol{x}) \geq 0\}$ (puede ser $\mathbb{R}^n$, si sólo tenemos $g_1 \equiv 1$). Como las funciones mencionadas son polinomiales, este problema equivale a (3.1), como explicamos luego de presentarlo.

$$\begin{aligned} \rho_{mom} = & \sup_{\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}^n}} L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & L_{\boldsymbol{y}}(h_j) = \sum_{\alpha \in \mathbb{N}^n} h_{j,\alpha} y_\alpha \leq \gamma_j, \forall j \in \Gamma \\ & \exists \mu \in \mathcal{M}(K) / y_\alpha = \int_K \boldsymbol{x}^\alpha d\mu, \forall \alpha \in \mathbb{N}^n \end{aligned} \tag{4.1}$$

Efectivamente, por la Proposición 3.2.3 tenemos que $L_{\boldsymbol{y}}(f) = \int_K f d\mu$, y lo mismo para los polinomios h_j, así que la definición de ρ_{mom} equivale a hallar el supremo de las integrales de f y las restricciones sobre $L_{\boldsymbol{y}}(h_j)$ equivalen a que estos polinomios cumplan las restricciones respectivas en sus integrales. No obstante, para que el funcional $L_{\boldsymbol{y}}$ aplicado a estos polinomios sea realmente igual a las integrales respectivas, $\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}^n}$ debe ser la secuencia de momentos de una medida boreliana, no negativa y finita μ. Esto es exactamente lo que pide la segunda restricción. De esta manera, (3.1) y (4.1) equivalen cuando f y los h_j son polinomios, y de hecho esto es necesario para que la

integral se pueda calcular como igual al funcional. Así, hemos formulado buscar un supremo sobre medidas como buscarlo sobre secuencias de momentos. Recordamos que al problema (3.1) le corresponde el problema dual (3.2).

4.1. Relajaciones semidefinidas

Comenzamos definiendo una **relajación primal** del problema (4.1) y explicando por qué es una relajación, es decir, toda solución factible del problema original cumple las restricciones del nuevo y no vale la recíproca. Se arma entonces la relajación primal de la siguiente forma: sean $v_k, k = 1, \ldots, m$ y $w_j, j \in \Gamma$ tales que g_k tiene grado $2v_k$ o $2v_k - 1$ y h_j tiene grado $2w_j$ o $2w_j - 1$; también sea t tal que f tiene grado $2t$ o $2t - 1$. Nombramos $i_0 = \max\{t, \max_{k=0,\ldots,m} v_k, \max_{j\in\Gamma} w_j\}$, y fijamos $i \geq i_0$. Tenemos entonces el siguiente problema de programación semidefinida donde las variables de decisión son los y_α (debemos tomar los momentos hasta orden $2i \geq 2i_0$ para que la matriz de momentos, las localizadoras, $L_{\boldsymbol{y}}(f)$ y los $L_{\boldsymbol{y}}(h_j)$ estén bien definidos). En el caso particular donde $i = v_k$, $M_{i-v_k}(g_k, \boldsymbol{y})$ es una matriz de 1×1.

$$\begin{aligned} \rho_i = \sup_{\boldsymbol{y}\in\mathbb{N}^{\mathbf{n}}_{\mathbf{2i}}} & L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & L_{\boldsymbol{y}}(h_j) \leq \gamma_j, j \in \Gamma \\ & M_i(\boldsymbol{y}) \succeq 0 \\ & M_{i-v_k}(g_k, \boldsymbol{y}) \succeq 0, k = 1, \ldots, m \end{aligned} \tag{4.2}$$

Las matrices $M_i(\boldsymbol{y}), M_{i-v_k}(g_k, \boldsymbol{y})$ son respectivamente la matriz de momentos y las matrices localizadoras, dadas por la Definición 3.2.1 y la Definición 3.3.1. Notamos que se acaba de construir una familia de relajaciones, con parámetro a elegir $i \geq i_0$: este regula la dimensión de las matrices mencionadas. Vemos que (4.2) es en efecto un problema SDP. La primera restricción es lineal con finitos coeficientes (puesto que los h_j son polinomios), por lo que se puede expresar como restricción semidefinida mediante una matriz diagonal; para incluir las otras dos restricciones como parte de un problema SDP invocamos la Proposición 1.2.7. Veamos que es efectivamente una relajación del problema original.

Proposición 4.1.1 *(4.2) da condiciones necesarias y no suficientes para una solución factible de (4.1).*

Demostración: La primera condición es copiada del problema anterior. Además, la última restricción del problema original indica que $\boldsymbol{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ debe ser la secuencia de momentos de una medida boreliana finita con soporte contenido en K. Esto implica las restricciones segunda y tercera de la relajación por el Corolario 3.3.4. Resta ver que las condiciones de la relajación no son suficientes. En efecto, por el Teorema 3.3.6 las

matrices localizadoras y la de momentos deberían ser semidefinidas positivas para todo $i \geq i_0$ para que $\boldsymbol{y}$ tenga una medida representativa, pero sólo lo pedimos en la relajación para un valor de i en particular. □

Por otro lado, a una relajación primal suele corresponder un problema dual con restricciones más fuertes que las del original ([11, Capítulo 4]). Para el problema (4.2), que se puede reformular como un problema con una sola restricción matricial de acuerdo al formato (1.6), tomamos su dual SDP de [11, Ecuación (4.6)].

$$\begin{aligned} \rho_i^* = & \inf_{\lambda \in \mathbb{R}^\Gamma} \sum_{j \in \Gamma} \lambda_j \gamma_j \\ \text{sujeto a } & -\langle X, B_\alpha \rangle - \sum_{k=1}^{m} \langle Z_k, C_{k,\alpha} \rangle + \sum_{j \in \Gamma} \lambda_j h_{j,\alpha} = f_\alpha, \forall \alpha \in \mathbb{N}^n_{2i} \\ & X, Z_k \succeq 0, \forall k = 1, \ldots, m \\ & \lambda_j \geq 0, \forall j \in \Gamma_+ \end{aligned} \tag{4.3}$$

Las matrices B_α ya fueron definidas en el Capítulo 2. Recordamos que $M_i(\boldsymbol{y}) = \sum_{\alpha \in \mathbb{N}^n_{2i}} B_\alpha y_\alpha$. Dado $1 \leq k \leq m$, las matrices $C_{k,\alpha}$ verifican que $\sum_{\alpha \in \mathbb{N}^n_{2i}} y_\alpha C_{k,\alpha} = M_{i-v_k}(g_k, \boldsymbol{y})$. Vamos a verificar que la formulación dual que acabamos de dar es equivalente a la nueva que sigue.

$$\begin{aligned} \rho_i^* = & \inf_{\lambda \in \mathbb{R}^\Gamma} \sum_{j \in \Gamma} \lambda_j \gamma_j \\ \text{sujeto a } & \sum_{j \in \Gamma} \lambda_j h_j - f = f_0 + \sum_{k=1}^{m} f_k g_k \\ & f_k \in \Sigma[\boldsymbol{x}], \deg(f_k) \leq 2(i - v_k), k = 0, \ldots, m \\ & \lambda_j \geq 0, \forall j \in \Gamma_+ \end{aligned} \tag{4.4}$$

Lema 4.1.2 *Dado $p \in \mathbb{N}_{>0}, A \in \mathcal{S}_p \succeq 0$ si y sólo si existen $z_d \in \mathbb{R}^p, 1 \leq d \leq p$ tales que $A = \sum_{d=1}^{p} z_d z_d^t$.*

Demostración: $\Leftarrow$) Se deduce inmediatamente de la Proposición 1.2.3 y el Lema 1.2.11.
$\Rightarrow$) Sean $O \in \mathbb{R}^{p \times p}$ una matriz ortogonal y $D \in \mathbb{R}^{p \times p}$ diagonal tales que $A = ODO^t$. Dado $1 \leq d \leq p$, llamamos $O_d \in \mathbb{R}^{p \times p}$ a la matriz con $d-$ésima columna igual a la de O y el resto ceros (llamamos z_d a los respectivos vectores columna), y $D_d \in \mathbb{R}^{p \times p}$ a la matriz que tiene el $d-$ésimo elemento de la diagonal de D, al que llamamos q_d, en el mismo lugar y el resto ceros. Como $A \succeq 0$, toda la diagonal de D es no negativa. Lo

siguiente se puede verificar a mano:

$$A = ODO^t = \sum_{d=1}^{p} OD_dO^t = \sum_{d=1}^{p} (\sqrt{q_d}O_d)(\sqrt{q_d}O_d^t) = \sum_{d=1}^{p} \sqrt{q_d}z_d\sqrt{q_d}z_d^t$$

y el resultado se sigue. □

Proposición 4.1.3 *Los problemas (4.3) y (4.4) son equivalentes.*

Demostración: En el problema primal (4.2) se pide

$$\begin{pmatrix} D & 0 & \cdots & 0 & 0 \\ 0 & M_i(\boldsymbol{y}) & \cdots & 0 & 0 \\ \vdots & \vdots & M_{i-v_1}(g_1, \boldsymbol{y}) & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & M_{i-v_m}(g_m, \boldsymbol{y}) \end{pmatrix} \succeq 0$$

donde D es una matriz diagonal de $\#\Gamma$ filas y columnas, con $j-$ésimo elemento de la diagonal igual a $\gamma_j - L_{\boldsymbol{y}}(h_j)$. Usando el formato del problema dual SDP (1.7) se obtiene la función objetivo y las primeras dos restricciones del problema (4.3).
Multiplicando ambos miembros de la primera restricción por $\boldsymbol{x}^\alpha$ para cada $\alpha \in \mathbb{N}^n_{2i}$, se deduce

$$\sum_{j\in\Gamma} \lambda_j h_j(\boldsymbol{x}) = \left\langle X, \sum_{\alpha\in\mathbb{N}^n_{2i}} B_\alpha \boldsymbol{x}^\alpha \right\rangle + \sum_{k=1}^{m} \left\langle Z_k, \sum_{\alpha\in\mathbb{N}^n_{2i}} C_{k,\alpha}\boldsymbol{x}^\alpha \right\rangle$$

Recordamos ahora que $L_{\boldsymbol{y}}(p)$ se define tomando los coeficientes del polinomio p y reemplazando los $\boldsymbol{x}^\alpha$ por los y_α. Entonces, para volver a obtener p se puede revertir la construcción, cambiando los y_α por $\boldsymbol{x}^\alpha$. Por cómo se definen las coordenadas de la matriz de momentos y la localizadora (Definiciones 3.2.1 y 3.3.1), haciendo esa reversión se puede deducir que $\sum_{\alpha\in\mathbb{N}^n_{2i}} C_{k,\alpha}\boldsymbol{x}^\alpha = g_k(\boldsymbol{x})v_{i-v_k}(\boldsymbol{x})v_{i-v_k}(\boldsymbol{x})^t, 1 \le k \le m$, y que $\sum_{\alpha\in\mathbb{N}^n_{2i}} B_\alpha\boldsymbol{x}^\alpha = v_i(\boldsymbol{x})v_i(\boldsymbol{x})^t$.
Ahora, sean las descomposiciones $X = \sum_{d=1}^{s(i)} z_{0,d}z_{0,d}^t$, $Z_k = \sum_{d=1}^{s(i-v_k)} z_{k,d}z_{k,d}^t, 1 \le k \le m$ que existen por el Lema 4.1.2. Dado $0 \le k \le m$, nombrando $v_0 = 0$ se tiene

$$\left\langle \sum_{d=1}^{s(i-v_k)} z_{k,d}z_{k,d}^t, v_{i-v_k}v_{i-v_k}^t \right\rangle = \sum_{d=1}^{s(i-v_k)} \langle z_{k,d}z_{k,d}^t, v_{i-v_k}v_{i-v_k}^t\rangle = \sum_{d=1}^{s(i-v_k)} tr(z_{k,d}z_{k,d}^t v_{i-v_k}v_{i-v_k}^t)$$

$$= \sum_{d=1}^{s(i-v_k)} tr(v_{i-v_k}^t z_{k,d}z_{k,d}^t v_{i-v_k})$$

La última igualdad se verifica a mano, y la última fórmula es una suma de cuadrados de polinomios. Así, Llamamos $f_k, k = 0, \ldots, m$ a estas sumas de cuadrados, y concluimos así las restricciones primera y segunda de (4.4).
Todo el razonamiento realizado es reversible, lo que permite construir una solución factible de (4.3) desde una de (4.4). La restricción $\lambda_j \geq 0, j \in \Gamma_+$ es igual en ambos problemas. □

Además de ser equivalente a un dual de un problema SDP, notamos que (4.4) también es un problema SDP por la Proposición 2.1.10. Este problema es más fuerte que (3.2) puesto que se reemplaza la no negatividad de $\sum_{j\in\Gamma} \lambda_j h_j(\boldsymbol{x}) - f(\boldsymbol{x})$ para todo $\boldsymbol{x}$ en K por una condición suficiente para que esto ocurra. Puntualmente, estamos escribiendo $\sum_{j\in\Gamma} \lambda_j h_j - f$ con la forma del Certificado de Positividad de Putinar. En conclusión, (4.4) puede ser tomado como problema SDP dual de (4.2) por ser equivalente a (4.3), y así lo consideramos en lo que sigue.

A continuación vamos a desarrollar condiciones de convergencia de las relajaciones.

Lema 4.1.4 *Sean $d, D \in \mathbb{N}, d \leq D$. Sean $g \in \mathbb{R}[\boldsymbol{x}], \boldsymbol{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n_{2D+\deg(g)}}$. Si $M_D(g, \boldsymbol{y}) \succeq 0$, entonces $M_d(g, \boldsymbol{y}) \succeq 0$.*

Demostración: Observamos que, por construcción de las matrices localizadoras, seleccionar las primeras $s(d)$ filas y columnas de $M_D(g, \boldsymbol{y})$ da la matriz $M_d(g, \boldsymbol{y})$. De esta forma, todo menor principal de $M_d(g, \boldsymbol{y})$ es un menor principal de $M_D(g, \boldsymbol{y})$, y el resultado se sigue de que una matriz simétrica es semidefinida positiva si y sólo si todos sus menores principales son no negativos. □

Teorema 4.1.5 *Sean $f; h_j, j \in \Gamma; g_k, 1 \leq k \leq m$ polinomios en $\mathbb{R}[\boldsymbol{x}]$, $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}) \geq 0\}$. Suponemos que Γ es finito, $Q = Q(g_1, \ldots, g_m)$ es arquimediano, $h_0 > 0$ en K y (4.1) tiene valor óptimo finito ρ_{mom}. Para $k = 1, \ldots, m$ sean v_k tales que g_k tiene grado $2v_k$ o $2v_k - 1$ y $v = \max_{1\leq k\leq m} v_k$. Consideramos todas las relajaciones primales (4.2) para $i \geq i_0$ y sus respectivos problemas duales (4.4), con valores óptimos ρ_i, ρ_i^* respectivamente. Entonces se tiene:*

1. *$\rho_i^* \searrow \rho_{mom}$ y $\rho_i \searrow \rho_{mom}$ cuando $i \to \infty$.*
2. *Si el problema (4.2) para un $i \geq i_0$ tiene una solución óptima $\boldsymbol{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n_{2i}}$ tal que $\operatorname{rg}(M_s(\boldsymbol{y})) = \operatorname{rg}(M_{s-v}(\boldsymbol{y}))$ para algún $i_0 \leq s \leq i$, entonces $\rho_i = \rho_{mom}$ y además el problema (4.1) tiene una solución óptima $\tilde{\boldsymbol{y}} = \{\tilde{y}_\beta\}_{\beta\in\mathbb{N}^n}$ con una medida representativa $\operatorname{rg}(M_s(\boldsymbol{y}))$-atómica con soporte contenido en K.*

Demostración: Primero notamos que $\rho_{i+1} \leq \rho_i, \forall i \geq i_0$, puesto que toda solución factible de (4.2) para $i + 1$ lo es para i (ya que $M_{i+1-v_k}(g_k, \boldsymbol{y}) \succeq 0 \Rightarrow M_{i-v_k}(g_k, \boldsymbol{y}) \succeq$

$0, M_{i+1}(\boldsymbol{y}) \succeq 0 \Rightarrow M_i(\boldsymbol{y}) \succeq 0$ por el Lema 4.1.4). Veamos también que $\rho_i \geq \rho_{mom}$. Suponiendo lo contrario, existe z valor factible para (4.1) tal que $\rho_i < z \leq \rho_{mom}$. Este valor está dado por una solución factible $\boldsymbol{y'} = \{y'_\alpha\}_{\alpha \in \mathbb{N}^n}$. Recordamos que $i \geq i_0 \geq t$, donde f tiene grado $2t$ o $2t-1$. Luego, truncar la secuencia $\boldsymbol{y'}$ tomando sólo los índices con $|\alpha| \leq 2i$ no altera la evaluación de $L_{\boldsymbol{y'}}(f)$, puesto que el resto de términos de $\boldsymbol{y'}$ se multiplican por coeficientes nulos de f. Así, la subsecuencia que acabamos de tomar es solución factible de (4.2), donde también toma valor z. Pero entonces ρ_i no era el supremo de valores factibles de (4.2), lo que contradice su definición.
Como Q es arquimediano, K es compacto (Proposición 2.2.4), vale $h_0 > 0$ en K y además f y los h_j son polinomios (en particular, semicontinuos superiormente y acotados en K por el Teorema de Weierstrass). Del Teorema 3.1.4 se sigue que hay dualidad fuerte, es decir $\rho_{mom} = \rho_{pop}$. Además, por el Teorema 1.2.13 sabemos que $\rho_i \leq \rho_i^*$.
Dado $\varepsilon > 0$, sea $\lambda \in \mathbb{R}^\Gamma$ una solución factible de (3.2) con valor asociado ρ_λ que cumple

$$\rho_{pop} \leq \rho_\lambda \leq \rho_{pop} + \varepsilon$$

Como es solución factible, se tiene $\sum_{j \in \Gamma} \lambda_j h_j - f \geq 0$ en K.
Sea $\hat{\lambda}$ igual a λ excepto por $\hat{\lambda}_0 = \lambda_0 + \varepsilon$. Luego, $\hat{\lambda}$ resulta factible en (3.2) puesto que

$$\sum_{j \in \Gamma} \hat{\lambda}_j h_j = \sum_{j \in \Gamma} \lambda_j h_j + h_0 \varepsilon \geq h_0 \varepsilon > 0$$

ya que por hipótesis tenemos $h_0 > 0$ en K. Además, sigue valiendo $\hat{\lambda}_j \geq 0, \forall j \in \Gamma_+$ porque ya valía para los λ_j respectivos y el único cambio es que a λ_0 se le suma un positivo. Como Γ es finito, $\sum_{j \in \Gamma} \hat{\lambda}_j h_j$ está bien definido, y por el Teorema 2.2.6 pertenece a $Q(g_1, \ldots, g_m)$, es decir $\sum_{j \in \Gamma} \hat{\lambda}_j h_j = f_0 + \sum_{b=1}^m f_b g_b$, con $f_b \in \Sigma[\boldsymbol{x}], b = 0, \ldots, m$. Además, si $2i \geq \max_{k=0,\ldots,m} \deg(f_k g_k)$ con $g_0 \equiv 1$ (alcanza pedir i lo bastante grande porque estamos probando convergencia), $\hat{\lambda}$ resulta una solución factible de (4.4). Se pide esto para tener

$$\deg(f_b g_b) \leq 2i \Rightarrow \deg(f_b) \leq 2(i - v_b)$$

donde esta implicación sale de la definición de los v_b y da la condición que debe verificar la solución dada.
Como $\hat{\lambda}$ es solución factible, se tiene

$$\rho_{mom} = \rho_{pop} \leq \rho_i^* \leq \rho_{\hat{\lambda}} = \sum_{j \in \Gamma} \hat{\lambda}_j \gamma_j = \sum_{j \in \Gamma} \lambda_j \gamma_j + \varepsilon \gamma_0 \leq \rho_{pop} + \varepsilon(1 + \gamma_0) = \rho_{mom} + \varepsilon(1 + \gamma_0)$$

Como para conseguir esto sólo se requiere i lo bastante grande y ε es arbitrario, existe una sucesión $(\rho_{\hat{\lambda}_j})_i$ que converge a ρ_{mom}, y por lo tanto los respectivos ínfimos ρ_i^* también convergen a ρ_{mom}. Como $\rho_{mom} \leq \rho_i \leq \rho_i^*$, se deduce lo mismo para los ρ_i.

Probamos ahora el segundo item. Sean $i_0 \leq s \leq i$ tales que una solución óptima $\boldsymbol{y}$ de (4.2) con valor asociado ρ_i cumple para s la hipótesis del segundo item de este teorema. Por el Teorema 3.3.7, reemplazando d por s, la hipótesis de este teorema da el segundo item de la equivalencia. El primero se obtiene nuevamente reemplazando d por s, porque al ser $\boldsymbol{y}$ solución factible de (4.2) para i, entonces cumple $M_i(\boldsymbol{y}) \succeq 0, M_{i-v_k}(g_k, \boldsymbol{y}) \succeq 0, \forall k = 1, \ldots, m$. En particular satisface $M_s(\boldsymbol{y}) \succeq 0, M_{s-v_k}(g_k, \boldsymbol{y}) \succeq 0$ por el Lema 4.1.4. De esto se sigue que $M_{s-v}(g_k, \boldsymbol{y}) \succeq 0, k = 1, \ldots, m$ por definición de v. Por el teorema mencionado, $\{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2s}}$ es secuencia de momentos de una medida $\mathrm{rg}(M_s(\boldsymbol{y}))$-atómica μ con soporte contenido en K.
Notamos que la secuencia que tiene medida representativa μ garantizada es la de los y_α hasta orden $2s$, es decir no toda la secuencia definida hasta orden $2i$ que era solución óptima de la relajación primal. Igualmente, con esto alcanza: como $2s \geq 2i_0$, la secuencia de términos hasta orden $2s$ satisface todas las restricciones de valores del funcional $L_{\boldsymbol{y}}$ en el problema (4.1), que además están bien definidos. Los y_α restantes no cambian las evaluaciones del funcional porque todos los f, h_j tienen grados acotados por $2i_0 \leq 2s$, y por lo tanto los términos y_α de órdenes mayores a $2s$ se multiplican por coeficientes nulos. Luego, la secuencia puede redefinirse, extendiendo $\{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2s}}$ con los momentos dados por

$$y_\kappa = \int_K \boldsymbol{x}^\kappa d\mu, \forall \kappa \in \mathbb{N}^n, |\kappa| > 2s$$

donde los y_κ están bien definidos porque los monomios son continuos, μ es una medida finita y K es compacto, de modo que todas las integrales son finitas. Así, la nueva secuencia $\boldsymbol{y}' = \{y_\alpha\}_{\alpha \in \mathbb{N}^n}$ es solución factible de (4.1) con valor asociado $\rho_i \geq \rho_{mom}$. Al mismo tiempo, como ρ_i es valor factible resulta $\rho_i \leq \rho_{mom}$ ya que este último es el supremo de los valores factibles. Finalmente, $\rho_i = \rho_{mom}$, de forma que la secuencia $\boldsymbol{y}'$ redefinida es solución óptima de (4.1). □

Este último resultado se prueba en [11, Teorema 4.1]. Si $Q(g_1, \ldots, g_m)$ no es arquimediano, pidiendo que K sea compacto se puede obtener convergencia pero a cambio de un mayor costo computacional. Se plantea la relajación primal con una cantidad exponencial en m de restricciones matriciales, con la convención $g_\emptyset \equiv 1$:

$$\begin{aligned} \rho_i = \sup_{\boldsymbol{y} \in \mathbb{N}^n_{2i}} & L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & L_{\boldsymbol{y}}(h_j) \leq \gamma_j, j \in \Gamma \\ & M_i(\boldsymbol{y}) \succeq 0 \\ & g_J = \prod_{k \in J} g_k \wedge v_J = \deg(g_J), \forall J \subset \{1, \ldots, m\} \\ & M_{i-v_J}(g_J, y) \succeq 0, \forall J \subset \{1, \ldots, m\} \end{aligned} \tag{4.5}$$

Tomamos $i \geq i_0$, con i_0 construido de forma análoga a la de (4.2) para que todo esté bien definido. Se vuelve a tener una relajación (las nuevas condiciones son necesarias

pero no suficientes) por el mismo razonamiento que en la relajación (4.2); la única diferencia es que para mostrar que las nuevas condiciones no son suficientes invocamos el Teorema 3.3.5. A este nuevo problema corresponde el dual con restricciones más fuertes que sigue, usando las mismas funciones $g_J, v_J, J \subset \{1, \dots, m\}$:

$$\begin{aligned} \rho_i^* = & \inf_{\lambda \in \mathbb{R}^\Gamma} \sum_{j \in \Gamma} \lambda_j \gamma_j \\ \text{sujeto a } & \sum_{j \in \Gamma} \lambda_j h_j - f = \sum_{J \subset \{1,\dots,m\}}^{m} f_J g_J \\ & f_J \in \Sigma[\boldsymbol{x}], \deg(f_J) \leq 2(i - v_J), \forall J \subset \{1, \dots, m\} \\ & \lambda_j \geq 0, \forall j \in \Gamma_+ \end{aligned} \tag{4.6}$$

Ahora escribimos $\sum_{j \in \Gamma} \lambda_j h_j - f$ con la forma del Certificado de Positividad de Schmüdgen. Así, tenemos un resultado de convergencia análogo al Teorema 4.1.5, donde los únicos cambios en la demostración son asumir K compacto, escribir $\sum_{j \in \Gamma} \hat{\lambda}_j h_j \in P(g_1, \dots, g_m)$ y continuar trabajando como en la demostración pero con las funciones f_J, g_J. Nuevamente, este es un problema SDP por la Proposición 2.1.10.

Aunque seguimos teniendo convergencia, con esta nueva versión el costo computacional aumenta mucho por la cantidad de matrices que deben ser semidefinidas positivas. Esta cantidad pasa de ser lineal a exponencial en m. También notamos que podemos contar con el item 2. del Teorema 4.1.5 cambiando la mención a (4.2) por (4.5), porque las matrices que en este teorema deben ser semidefinidas positivas son sólo algunas de las que deben serlo en la relajación dada con el formato asociado a Schmüdgen (las matrices localizadoras de las funciones $g_J, J = \{j\}, j = 1, \dots, m$); y además recordamos que el Teorema 3.3.7 no pedía compacidad de K ni que $Q(g_1, \dots, g_m)$ fuera arquimediano. En otras palabras, en esta versión de las relajaciones también tenemos la misma condición de rango matricial para verificar optimalidad.

Damos ahora la relajación primal en el caso $K = \mathbb{R}^n$. Podemos interpretar K como el conjunto semialgebraico básico cerrado formado sólo por $g_1 \equiv 1$. Luego, como la matriz localizadora de la función constante 1 es igual a la de momentos, se deduce la formulación

$$\begin{aligned} \rho_i = & \sup_{\boldsymbol{y} \in \mathbb{N}^n_{2i}} L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & L_{\boldsymbol{y}}(h_j) \leq \gamma_j, \forall j \in \Gamma \\ & M_i(\boldsymbol{y}) \succeq 0 \end{aligned} \tag{4.7}$$

Esta versión se utilizará luego en la minimización de polinomios sin restricciones.

Damos un ejemplo de cómo comienza un caso exitoso de problema de momentos (en el

sentido de que se cumple la hipótesis de rango y por lo tanto hay una medida atómica representativa). Lo tomamos de [11, Ejemplo 4.1]. Lasserre realizó estos cálculos mediante un software llamado GloptiPoly. Sean $f(x_1, x_2) = (x_1-1)^2+(x_1-x_2)^2+(x_2-3)^2$, con $h_0 \equiv 1, \gamma_0 = 1$ con restricción de igualdad, y definimos

$$K = \{\boldsymbol{x} \in \mathbb{R}^2 : 1 - (x_1 - 1)^2 \geq 0, 1 - (x_1 - x_2)^2 \geq 0, 1 - (x_2 - 3)^2 \geq 0\}$$

Como los tres polinomios que definen K y el f tienen grado 2, se define $i_0 = 1$, y esto da la primera relajación primal a ejecutar. Se obtiene $\rho_1 = 3$ con solución óptima $\boldsymbol{y}^*$. Por otro lado, se tiene $\text{rg}(M_1(\boldsymbol{y}^*)) = 3$. La relajación con $i = 2$ da $\rho_2 = 2$ con solución óptima $\mathbf{z}^*$ y $\text{rg}(M_2(\mathbf{z}^*))$=3. Como se verifica $\text{rg}(M_2(\mathbf{z}^*)) = \text{rg}(M_1(\mathbf{z}^*))$, deducimos que $\rho_{mom} = \rho_2$ y que existe una medida representativa con $\text{rg}(M_2(\mathbf{z}^*)) = 3$ átomos.

Debemos notar que la secuencia $\boldsymbol{y}^*$ obtenida en cada relajación no tiene necesariamente una relación con las anteriores, y por eso en nuestro ejemplo usamos nombres distintos $\boldsymbol{y}^*, \mathbf{z}^*$. Cada vez que terminamos de ejecutar una relajación, tomamos la secuencia de momentos óptima de la última relajación y verificamos la condición de rango para las matrices de momentos dadas por ella misma: $M_s(\mathbf{z}^*), M_{s-v}(\mathbf{z}^*)$.

4.2. Extracción de soluciones

Ya estudiamos las sucesivas relajaciones semidefinidas y en qué circunstancias convergen a la solución buscada, así como condiciones que marcan haber alcanzado el óptimo deseado. Para esta etapa, suponemos que para cierto $i \geq i_0$ se consiguió una secuencia de momentos $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2i}}$ tal que es solución óptima de (4.2) y además $\text{rg}(M_a(\boldsymbol{y})) = \text{rg}(M_{a-v}(\boldsymbol{y}))$, para algún $i_0 \leq a \leq i$. Por el Teorema 4.1.5, se tiene $\rho_i = \rho_{mom}$ y además la subsecuencia $\{y_\beta\}_{\beta \in \mathbb{N}^n_{2a}}$ tiene una medida representativa $\text{rg}(M_a(\boldsymbol{y}))$-atómica μ que realiza ρ_{mom}. Esta medida es solución óptima de (3.1) por el item 2 del Teorema 4.1.5. El algoritmo de extracción tiene como objetivo encontrar esos átomos $\boldsymbol{x}^*(k) \in \mathbb{R}^n, k = 1, \ldots, \text{rg}(M_a(\boldsymbol{y}))$.

Para comenzar, recordamos la definición de integral de una función no negativa sobre una medida abstracta [22, ecuación (10.15)]:

$$\int_E f d\mu = \sup\{\sum_j (\inf_{\boldsymbol{x} \in E_j} f(\boldsymbol{x}))\mu(E_j)\} \tag{4.8}$$

donde el supremo se toma sobre todas las posibles particiones finitas $\{E_j\}_j$ del conjunto E en conjuntos medibles. A continuación se define la integral para funciones medibles en general según [22, ecuación (10.22)]. De esto y de las propiedades de la integral que se demuestran en el mismo capítulo, se puede deducir la integral de una función para una medida $r-$atómica con soporte contenido en K y átomos $\boldsymbol{x}(k), k = 1, \ldots, r$:

Proposición 4.2.1 *Dados $K \subset \mathbb{R}^n$ boreliano, $\mu \in \mathcal{M}(K)$ una medida $r-$atómica en K con $r \in \mathbb{N}_{>0}$, se tiene:*

$$\mu = \sum_{k=1}^{r} z_k \delta_{\boldsymbol{x}(k)}, z_k > 0 \Rightarrow \int_K f d\mu = \sum_{k=1}^{r} z_k f(\boldsymbol{x}(k)) \tag{4.9}$$

donde $\delta_{\boldsymbol{x}}$ es la medida de Dirac concentrada en $\boldsymbol{x} \in \mathbb{R}^n$.

Demostración: Utilizamos la aditividad de la medida y partimos K en los conjuntos $Z_k = \{\boldsymbol{x}(k)\}, K \setminus \bigcup_{1 \leq k \leq r} Z_k$. El ínfimo de f sobre cada conjunto puntual es la imagen en ese punto, y $\mu(K \setminus \bigcup_{1 \leq k \leq r} Z_k) = 0$. Así vemos que la fórmula que queremos demostrar es una de las posibles sumatorias con ínfimos de (4.8). Falta ver que es cota superior de cualquier otra.
Dada una partición alternativa $E = \bigcup_{t=1}^{p} B_t$, algunos de estos B_t contienen a todos los $\boldsymbol{x}(k)$ (puede haber varios puntos en un mismo conjunto). La unión de todos los otros mide 0 porque no tiene ninguno de los $\boldsymbol{x}(k)$. $\mu(B_t)$ es la suma de los z_k para los $\boldsymbol{x}(k)$ que este B_t contiene. Esto nos permite reescribir

$$\sum_{t=1}^{p} \operatorname*{ínf}_{\boldsymbol{x} \in B_t} (f(\boldsymbol{x}))\mu(B_t) = \sum_{k=1}^{r} z_k \operatorname*{ínf}_{\boldsymbol{x} \in B_{t(k)}} (f(\boldsymbol{x}))$$

donde $t(k)$ verifica que $\boldsymbol{x}(k) \in B_{t(k)}$. Como los $\boldsymbol{x}(k)$ están en los $B_{t(k)}$, los $f(\boldsymbol{x}(k))$ son mayores o iguales a los ínfimos respectivos, y el resultado se sigue. □

Recordamos ahora la Proposición 3.2.2, que nos indica que $L_{\boldsymbol{y}}(\boldsymbol{x}^\gamma) = y_\gamma$, donde y era la secuencia de momentos de la medida μ. Es decir, $y_\gamma = \int_K \boldsymbol{x}^\gamma d\mu$. Tomamos ahora $r = \operatorname{rg}(M_a(\boldsymbol{y}))$, y consideramos μ como la medida representativa $r-$atómica de los y_α con $|\alpha| \leq 2a$ que sabemos que existe. Por (4.9), tenemos que $y_\gamma = \sum_{k=1}^{r} z_k(\boldsymbol{x}^*(k))^\gamma$, con $z_k > 0$, para todo $\gamma \in \mathbb{N}^n_{2a}$. En particular, dados $\alpha, \beta \in \mathbb{N}^n_a$, tenemos que $M_a(\boldsymbol{y})_{\alpha,\beta} = \sum_{k=1}^{r} z_k(\boldsymbol{x}^*(k))^{\alpha+\beta}$ y por lo tanto nos queda

$$M_a(\boldsymbol{y}) = V^*(V^*)^t, \text{ con } V^* = (\sqrt{z_1}v_a(\boldsymbol{x}^*(1)) \ \sqrt{z_2}v_a(\boldsymbol{x}^*(2)) \ \ldots \ \sqrt{z_r}v_a(\boldsymbol{x}^*(r))) \tag{4.10}$$

donde V^* tiene sus vectores escritos como columnas y v_a es el vector de monomios definido en el Capítulo 2, al que ahora evaluamos en los átomos en cuestión. Si conociéramos los vectores $v_a(\boldsymbol{x}^*(k))$ ya tendríamos nuestro problema resuelto, porque los puntos que buscamos están formados por las coordenadas 2 a $n+1$ de los vectores $v_a(\boldsymbol{x}^*(k))$ (los monomios de grado 1 evaluados). Sin embargo, no podemos conocer las columnas ni los z_k directamente; sólo sabemos entonces que estas columnas son múltiplos no nulos de los vectores que buscamos.

Por eso, vamos obtener una matriz V de iguales dimensiones cuyas columnas generen el mismo subespacio.

Proposición 4.2.2 *Sean $M \in \mathbb{R}^{m\times m}$ una matriz de rango $h \leq m$ y $V \in \mathbb{R}^{m\times h}$ tal que $M = VV^t$. Entonces, las columnas de V forman una base del subespacio de $\mathbb{R}^m$ generado por las columnas de M.*

Demostración: Si las columnas de V son generadores del subespacio mencionado, entonces forman una base, porque son h columnas que generarían un subespacio de dimesión h (el generado por las columnas de M, que tiene rango h).
Veamos que generan este subespacio. Si nombramos $M(j)$ a la $j-$ésima columna, se tiene:

$$M_{i,j} = \sum_{k=1}^{r} V_{i,k}V^t_{k,j} = \sum_{k=1}^{r} V_{i,k}V_{j,k} \Rightarrow M(j) = \sum_{k=1}^{r} V_{j,k}V(k)$$

es decir, cada columna de M es una combinación lineal de columnas de V. □

En este caso, tomamos la matriz M como $M_a(\boldsymbol{y})$. Realizamos una factorización de la forma $M = VV^t, V \in \mathbb{R}^{s(a)\times r}$. Se puede obtener a partir de [6, Ecuación 4.2.17], donde D_r tiene todos los elementos de la diagonal no negativos puesto que $M_a(\boldsymbol{y}) \succeq 0$, de forma que se les puede tomar raíz cuadrada. Esta V no va a ser la V^* que queríamos, pero sus columnas sí generan el mismo subespacio que las de V^* por la Proposición 4.2.2. Vamos a querer expresar este subespacio a través de las $n - r$ ecuaciones que lo definen. Para simplificar esta tarea recurrimos a operaciones elementales de columnas sobre V para obtener su forma reducida por columnas (reduced column echelon form, RCEF). Se dice que una matriz se encuentra en esta forma si su traspuesta está en row echelon form, ver la definición en [24]. Por ejemplo, damos una matriz y su RCEF:

$$\begin{pmatrix} 3 & 5 \\ 4 & 1 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{22}{17} & -\frac{8}{17} \end{pmatrix}$$

Volviendo a nuestro problema, de la matriz V calculada antes **tomamos su RCEF** y la llamamos U. Las columnas de U son base del mismo subespacio porque se obtuvieron aplicando operaciones elementales de columnas a partir de las de V.

En las matrices de ejemplo, si llamamos S al subespacio generado por las columnas, vemos por la RCEF que un vector $z \in S$ satisface $z_3 = \frac{22}{17}z_1 - \frac{8}{17}z_2$. Esta ecuación define a S; en particular la cumplen las columnas de V, V^* y $M_a(\boldsymbol{y})$. Por otra parte, recordamos que en nuestro problema los vectores columna de V^* son los $v_a(\boldsymbol{x}^*(k)), k = 1, \ldots, r$, donde cada coordenada representa un monomio aplicado a cada uno de los átomos. Como V^* tenía rango r, U tendrá r filas que tienen un 1 y el resto ceros. Esto corresponde a r monomios que, evaluados en los átomos que estamos buscando, generan a todos los demás. Las ecuaciones del subespacio generado por las columnas se convierten así en ecuaciones polinomiales que todos los átomos buscados verifican. Tomamos de [11,

Ejemplo 4.1] un ejemplo con $n = a = 2$, o sea con 2 variables y monomios de grado a lo sumo 2. Las filas corresponden respectivamente a los monomios $1, x_1, x_2, x_1^2, x_1x_2, x_2^2$.

$$U = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 3 & 0 \\ -4 & 2 & 2 \\ -6 & 0 & 5 \end{array}\right) \rightarrow x_1^2 = -2 + 3x_1;\ x_1x_2 = -4 + 2x_1 + 2x_2;\ x_2^2 = -6 + 5x_2$$

En general, si $w(\boldsymbol{x}) = (\boldsymbol{x}^{\alpha_1}, \ldots, \boldsymbol{x}^{\alpha_r})$ es el vector de monomios base (en el ejemplo sería $w(\boldsymbol{x}) = (1, x_1, x_2)$), tenemos que $v_a(\boldsymbol{x}^*(k)) = Uw(\boldsymbol{x}^*(k)), k = 1, \ldots, r$.

Notamos además que los vectores $w(\boldsymbol{x}^*(k))$ son linealmente independientes. En efecto, como cada monomio corresponde a una fila de la matriz V^*, estas son r filas que generan a todas las demás. Por lo tanto, deben ser LI porque $\mathrm{rg}(V) = r$. Resolver las ecuaciones polinomiales planteadas se puede construir como un problema de autovalores. Vamos a armar a partir de U las **matrices de multiplicación** $N_i \in \mathbb{R}^{r\times r}, 1 \leq i \leq n$. Cada matriz N_i se define tomando de U las filas correspondientes a los monomios $x_i\boldsymbol{x}^{\alpha_k}, 1 \leq k \leq r$. En el ejemplo anterior, tenemos

$$N_1 = \left(\begin{array}{ccc} 0 & 1 & 0 \\ -2 & 3 & 0 \\ -4 & 2 & 2 \end{array}\right); N_2 = \left(\begin{array}{ccc} 0 & 0 & 1 \\ -4 & 2 & 2 \\ -6 & 0 & 5 \end{array}\right)$$

donde N_1 tiene las filas de los monomios x_1, x_1^2, x_1x_2 y N_2 tiene las filas de los monomios x_2, x_2x_1, x_2^2. Como cada fila corresponde a los valores que toman estos nuevos monomios en los distintos átomos, se tiene entonces que

$$N_i w(\boldsymbol{x}^*(k)) = x^*(k)_i w(\boldsymbol{x}^*(k)), 1 \leq i \leq n, 1 \leq k \leq r$$

Es decir, cada N_i tiene r autovalores contados con multiplicidad, y tienen todos una misma base de autovectores. Luego, cada N_i resulta diagonalizable.

Notar que la construcción de las matrices de multiplicación depende de elegir filas correspondientes a ciertos monomios. Estas filas no existirían si algún monomio $\boldsymbol{x}^{\alpha_k}$ tuviera grado a, el mayor posible. ¿Cómo asegurar que esto no ocurra? Encontramos la respuesta en una hipótesis del problema original: que $\mathrm{rg}(M_a(\boldsymbol{y})) = \mathrm{rg}(M_{a-v}(\boldsymbol{y}))$. De esto deducimos que todas las filas de $M_a(\boldsymbol{y})$ correspondientes a monomios de grado a son generadas por otras asociadas a monomios de grado menor, y lo mismo ocurre en las filas de V y de U porque en todas estas matrices las columnas generan el mismo subespacio. De esta manera, se puede obtener una RCEF U tal que sus filas con un 1 y ceros en las coordenadas restantes sean de monomios de grados menores a a. De esta

forma se logra que ningún monomio de $w(\boldsymbol{x})$ tenga grado máximo. Con esta hipótesis el algoritmo de extracción siempre funciona según [9], y se puede encontrar en la misma fuente ejemplos donde este paso falla sin la hipótesis sobre el rango.

Ahora tenemos definidas las matrices de multiplicación, sabemos que las coordenadas $x^*(k)_i$ que buscamos son autovalores de las mismas y estos autovalores se pueden calcular. ¿Tenemos el problema resuelto? En realidad no: sólo sabemos que los autovalores de N_i serán las coordenadas $i-$ésimas de los átomos, pero no hay forma de saber a cuál de los r átomos corresponde cada una. Para resolver este problema, vamos a comenzar calculando una combinación lineal aleatoria $N = \sum_{i=1}^{n} \lambda_i N_i$, con $\lambda_i \geq 0$ y $\sum_{i=1}^{n} \lambda_i = 1$. Esta estrategia aporta estabilidad numérica [4]. Por construcción de N notamos que para todo $1 \leq k \leq r$, el vector $w(\boldsymbol{x}^*(k))$ resulta autovector de N, y además su autovalor asociado es $\sum_{i=1}^{n} \lambda_i \boldsymbol{x}^*(k)_i$. Estos r autovalores son todos distintos con probabilidad 1 [4],[9].

Finalmente, a esta matriz nueva le tomamos una factorización de Schur $N = QTQ^t$, con q_k los vectores columna de Q una matriz ortogonal [9] y T una matriz triangular superior con los autovalores de N en su diagonal. Nombramos $T_i = Q^t N_i Q, i = 1, \ldots, n$. Sabemos que las matrices N_i conmutan, es decir, $N_i N_j = N_j N_i, \forall i, j = 1, \ldots, n$ [3, propiedad 5.]. También notamos que N conmuta con todas las matrices N_i.

Proposición 4.2.3 *Las matrices T y T_i arriba definidas conmutan dos a dos.*

Demostración: En efecto, como N_i, N_j conmutan, tenemos

$$T_i T_j = Q^t N_i Q Q^t N_j Q = Q^t N_i N_j Q = Q^t N_j N_i Q = Q^t N_j Q Q^t N_i Q = T_j T_i$$

De esto se deduce además que las T_i conmutan con T porque esta última es una suma de matrices que conmutan con T_i. □

Sabemos que T es triangular superior por definición, y que tiene r autovalores distintos con probabilidad 1 porque es semejante a N. La Proposición 4.2.3 nos permite determinar, utilizando [4, Proposición 6], que las matrices T_i también son triangulares superiores. Estos nos permite concluir la búsqueda gracias al siguiente teorema, que obtenemos de [3, Teorema] (sin número). En ese trabajo se llama 'ceros' a los átomos, son simples porque como la medida representativa es $r-$atómica sabemos que hay exactamente r átomos distintos:

Teorema 4.2.4 *Para cada $1 \leq k \leq r$, se tiene $\boldsymbol{x}^*(k) = ((T_1)_{k,k} \ \ldots \ (T_n)_{k,k})^t$ o, equivalentemente, $\boldsymbol{x}^*(k)_i = q_k^t N_i q_k, \forall 1 \leq k \leq r, 1 \leq i \leq n$.*

La primera formulación se obtiene de la fuente citada, y la segunda desarrollando $(T_i)_{k,k}$. Terminamos esta sección recopilando el algoritmo de extracción que acabamos de construir, el cual se puede ver también en [11, Algoritmo 4.2]. Notar que el valor de la

medida en cada uno de los átomos se pierde porque el algoritmo no depende de ella ni permite obtenerla. Sin embargo, no será necesaria para nuestros fines.

Algoritmo 4.2.5 *Input: Una matriz $M_a(\boldsymbol{y})$ construida a partir de una secuencia de momentos $\boldsymbol{y}$ que verifica las hipótesis del segundo ítem del Teorema 4.1.5.*
Output: los $r = \operatorname{rg}(M_a(\boldsymbol{y}))$ átomos de la medida representativa $r-$atómica de y.

1. *Obtener una factorización $M_a(\boldsymbol{y}) = VV^t$, con V de rango completo.*
2. *Hallar una RCEF U de V.*
3. *Identificar los monomios base que conforman w (los de las filas de U con un solo 1 y el resto ceros).*
4. *Construir las matrices de multiplicación N_i como se describe arriba.*
5. *Obtener una combinación lineal aleatoria N de las N_i con coeficientes no negativos que sumen 1.*
6. *Obtener una factorización de Schur ordenada $N = QTQ^t$.*
7. *Calcular los átomos como $\boldsymbol{x}^*(k)_i = q_k^t N_i q_k, 1 \leq i \leq n; 1 \leq k \leq r$.*

4.3. El algoritmo resultante

Ahora recopilamos las definiciones y resultados ya dados para construir un algoritmo de resolución del problema (3.1) cuando las funciones involucradas son polinomios y K no es todo $\mathbb{R}^n$. Esta condición se pide para que $s \neq s - v$ y así tener una condición de rango no trivial en el paso 3. Este método es una reescritura del que aparece en [11, Algoritmo 4.1].

Algoritmo 4.3.1 *Input: Un polinomio f de grado $2v_0$ o $2v_0 - 1$, un conjunto finito de polinomios $\{h_j\}_{j\in\Gamma}$ de grados $2w_j$ o $2w_j - 1, j \in \Gamma$; Un conjunto de polinomios $\{g_i\}_{1\leq i\leq m}$ de grados $2v_i$ o $2v_i - 1, 1 \leq i \leq m$ que definen el conjunto $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, \forall 1 \leq i \leq m\}$; y $N \in \mathbb{N}_{>0}$ una cota superior del índice de las relajaciones semidefinidas a ejecutar.*
Output: el valor óptimo ρ_{mom} y un conjunto de puntos $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r \in \mathbb{R}^n$ que forman el soporte de una medida óptima μ; o bien un valor $\rho_N \geq \rho_{mom}$.

1. *Calcular $v = \max_{k=1,\ldots,m} v_k$ y $i_0 = \max\{v_0, v, \max_{j\in\Gamma} w_j\}$. Definir $i = i_0$.*
2. *Resolver la relajación semidefinida (4.2) con índice i y valor óptimo ρ_i.*
3. *Sea $\boldsymbol{y}^*$ la solución óptima asociada a ρ_i. Si no existe $i_0 \leq s \leq i$ tal que $\operatorname{rg}(M_s(\boldsymbol{y}^*)) = \operatorname{rg}(M_{s-v}(\boldsymbol{y}^*))$, saltar al paso 5.*

4. *Para s el valor hallado que cumple la condición de rango verificada en el paso 3, retornar $\rho_{mom} = \rho_i$, y dado $r = \mathrm{rg}(M_s(\boldsymbol{y}^*))$ extraer los puntos $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r$ con el Algoritmo 4.2.5. Terminar la ejecución.*

5. *Si $i < N$, aumentar i en 1 y volver al paso 2.*

6. *Retornar ρ_N como cota superior de ρ_{mom}. Terminar la ejecución.*

Ahora que conocemos el algoritmo de extracción, cerramos el ejemplo que habíamos comenzado a desarrollar al final de la Sección 4.2. La column echelon form U de este ejemplo es, luego de redondear, la matriz de 6 filas dada en esa sección. En este caso, con los coeficientes aleatorios obtenidos se define $N = 0{,}6909N_1 + 0{,}3091N_2$, y los átomos obtenidos son $\boldsymbol{x}^*(1) = (1, 2), \boldsymbol{x}^*(2) = (2, 2), \boldsymbol{x}^*(3) = (2, 3)$. Notar además que $w(\boldsymbol{x}^*(1)) = (1, 1, 2), w(\boldsymbol{x}^*(2)) = (1, 2, 2), w(\boldsymbol{x}^*(3)) = (1, 2, 3)$ son autovectores de N_1 y N_2.

Procedemos a justificar los pasos del algoritmo recordando lo ya trabajado. El i_0 se calcula de forma que las restricciones de las relajaciones primales estén bien definidas (que se puedan formar las matrices de momentos y de localización). La definición de v es para que la condición de rango chequeada en el paso 3 implique la existencia de óptimo y medida representativa atómica. En el paso 2, en teoría se podría tener $\rho_i = +\infty$ o $\rho_i < +\infty$ que no se realiza. Ninguno de estos casos se puede detectar numéricamente al ejecutar el algoritmo, así que se genera una solución $\boldsymbol{y}^*$ para la que luego se comprobará la condición de rango. En cuanto al output, sabemos por el Teorema 4.1.5 que si se verifica la condición mencionada, ρ_i puede ser retornado como valor óptimo y existe una medida representativa con la cantidad de átomos indicada. En la sección de extracción de soluciones se justifica que ese algoritmo obtiene los átomos pedidos. Al mismo tiempo, hemos mencionado en la sección anterior que si esta condición de rango no se cumple, el algoritmo de extracción puede fallar y por eso se decide no ejecutarlo.

Se sabe también por el Teorema 4.1.5 que los ρ_i son cotas superiores de ρ_{mom}, y en particular lo será ρ_N si no se encuentra una solución óptima y sus átomos. También vimos que los ρ_i son decrecientes, por lo que establecer el máximo índice N de una relajación, además de garantizar que el programa termine, indica cuánto queremos acercarnos al valor óptimo y cuánta computación estamos dispuestos a invertir. No obstante, debemos notar que el resultado de convergencia que tenemos depende de que K sea compacto, por lo que sólo en este caso aumentar N garantiza acercarnos al óptimo. Más aún, para esta versión del algoritmo necesitamos suponer $Q(g_1, \ldots, g_m)$ arquimediano. Si no lo es, para conseguir convergencia deberíamos ejecutar un método análogo usando (4.5) y asumiendo un costo computacional exponencial en m.

Por otro lado, la extracción de soluciones no depende de la compacidad de K. En el

caso de haber encontrado una solución óptima $\boldsymbol{y}^*$ de la relajación primal que verifica la condición de rango, en particular esta solución es factible (las matrices correspondientes son semidefinidas positivas). deducimos que existe la medida representativa $r-$atómica por el Teorema 3.3.7, del cual recordamos que no requiere compacidad. Más aún, este resultado da un si y sólo si. De esto se sigue que sin la hipótesis de igualdad de rango no existe una medida representativa con la cantidad de átomos dada por el rango de la matriz correspondiente (puede no existir una con finitos átomos o requerir una cantidad mayor, pero la relajación no da información al respecto).

En el caso de minimización de polinomios sin restricciones, y por lo tanto sin la hipótesis de compacidad, buscaremos una variante para alcanzar este óptimo en caso de que exista. En este caso no restringido, daremos una variante de la condición de rango del paso 3 utilizando el Teorema de la Extensión Plana, con vistas a verificar que el algoritmo de extracción puede ejecutarse y dar las soluciones buscadas.

Capítulo 5

Aplicación a optimizar polinomios con y sin restricciones

Llegamos al objetivo del camino que fuimos desarrollando a lo largo de este trabajo: resolver problemas de optimización de polinomios en varias variables, tanto en $\mathbb{R}^n$ como en conjuntos semialgebraicos básicos compactos.

5.1. Las formas primal y dual

Para darle sentido al estudio previo del Problema de Momentos Generalizado, debemos ver primero que la optimización de polinomios es efectivamente una versión del GMP. Para eso sirve el siguiente resultado, cuya demostración también se encuentra en [11, Teorema 1.1].

Proposición 5.1.1 *Sean $f : \mathbb{R}^n \to \mathbb{R}$ y $K \subset \mathbb{R}^n$ no vacío. Sean el problema de minimizar una función en el dominio restringido:*

$$f_K^* = \inf_{\boldsymbol{x} \in K} f(\boldsymbol{x}) \tag{5.1}$$

y el problema de momentos

$$\begin{aligned} \rho_{mom} &= \inf_{\mu \in \mathcal{M}(K)} \int_K f d\mu \\ \text{sujeto a } & \int_K d\mu = 1 \end{aligned} \tag{5.2}$$

Entonces, estos problemas equivalen, es decir $f_K^* = \rho_{mom}$.

Demostración: Como asumimos K no vacío, los dos problemas tienen soluciones factibles: elementos de K en el primero y medidas borelianas que cumplan la restricción en el segundo. Luego, los respectivos ínfimos pueden ser finitos o $-\infty$. Estudiaremos cada caso.
Primero suponemos $f_K^* = -\infty$. En este caso, dado $M < 0$ existe $\boldsymbol{x}_0 \in K$ tal que $f(\boldsymbol{x}_0) \leq M$. Luego, podemos definir la medida de Dirac $\delta_{\boldsymbol{x}_0}$. Esta es una solución factible del problema (5.2) y verifica $\int_K f d\delta_{\boldsymbol{x}_0} = f(\boldsymbol{x}_0) \leq M$ por ecuación (4.9). Como M era arbitrario, resulta $\rho_{mom} = -\infty$.
Ahora, supongamos $f_K^* > -\infty$. Primero veamos que $f_K^* \leq \rho_{mom}$. En efecto, como $f(\boldsymbol{x}) \geq f_K^*$ para todo $\boldsymbol{x} \in K$, por monotonía de la integral tenemos $\int_K f d\mu \geq \int_K f_K^* d\mu = f_K^* \mu(K) = f_K^*$, dado que μ es una medida factible. Como toda integral factible de f está acotada inferiormente por f_K^*, también lo está el ínfimo. Para la otra desigualdad, dado $\boldsymbol{x} \in K$ tomamos la medida $\delta_{\boldsymbol{x}}$, que verifica $\int_K f d\delta_{\boldsymbol{x}} = f(\boldsymbol{x})$ por (4.9). Es decir, todos los valores que toma f son integrales sobre f de medidas factibles, y por lo tanto están acotadas inferiormente por el ínfimo de integrales. Esto es, $f_K^* \geq \rho_{mom}$. □

El problema (5.2) es de momentos con una restricción de igualdad o, como lo expresamos en este trabajo, con dos restricciones de desigualdad. Se puede ver como un GMP con el siguiente formato: $\Gamma = \{0, 1\}, h_0 \equiv 1, h_1 \equiv -1$ y $\gamma_0 = 1, \gamma_1 = -1$. De este modo, nos queda

$$\int_K h_0 d\mu \leq 1 \wedge \int_K h_0 d\mu \geq 1 \Leftrightarrow \int_K h_0 d\mu \leq 1 \wedge \int_K h_1 d\mu \leq -1$$

Hallada la formulación de las restricciones con el formato del GMP definido en el Capítulo 4, falta ver la fórmula a optimizar. Notamos que buscar el ínfimo de f con las restricciones encontradas equivale a buscar el supremo de $-f$ con las mismas restricciones. Es decir, (5.2) adaptado al formato de (3.1) queda:

$$\rho_{mom} = \sup_{\mu \in \mathcal{M}(K)} \int_K (-f) d\mu$$
$$\text{sujeto a } \int_K d\mu \leq 1, \int_K -d\mu \leq -1$$

De esto podemos deducir su problema dual usando el formato de (3.2):

$$\rho_{pop} = \inf_{\lambda_0, \lambda_1 \in \mathbb{R}} \lambda_0 - \lambda_1$$
$$\text{sujeto a } \lambda_0 - \lambda_1 - (-f(\boldsymbol{x})) \geq 0, \forall \boldsymbol{x} \in K$$

En esta última fórmula, notamos que se puede reescribir $\lambda_0 - \lambda_1$ como un único $\lambda \in \mathbb{R}$. Por último, buscar el ínfimo de λ con la restricción dada equivale a buscar el supremo

de $-\lambda$ con la misma restricción. Le cambiamos el signo a λ en la función objetivo y en la restricción, y nos queda el siguiente problema dual:

$$\begin{aligned} \rho_{pop} = & \sup_{\lambda \in \mathbb{R}} \lambda \\ & \text{sujeto a } f(\boldsymbol{x}) - \lambda \geq 0, \forall \boldsymbol{x} \in K \end{aligned} \tag{5.3}$$

Como el intercambio de ínfimo y supremo trajo aparejados cambios de signo de las funciones objetivo, ahora la dualidad débil nos indica que $\rho_{pop} \leq \rho_{mom}$. Estas versiones de los problemas primal y dual sirven para buscar valores mínimos de cualquier función definida en un dominio no vacío. En nuestro caso, vamos a trabajar sólo con polinomios.

Ahora planteamos cómo queda el problema primal de momentos en términos del funcional $L_{\boldsymbol{y}}$. Esto es, adaptamos (4.1) al caso de optimización polinomial. Como $h_0 \equiv 1, h_1 \equiv -1$ y $\gamma_0 = 1, \gamma_1 = -1$, se tiene las siguientes condiciones:

$$L_{\boldsymbol{y}}(h_0) = y_{(0,\ldots,0)} \leq 1$$

$$L_{\boldsymbol{y}}(h_1) = -y_{(0,\ldots,0)} \leq -1$$

Es decir, $y_{(0,\ldots,0)} = 1$, y la nueva formulación queda:

$$\begin{aligned} \rho_{mom} = & \inf_{\boldsymbol{y} \in \mathbb{R}^{\mathbb{N}^n}} L_{\boldsymbol{y}}(f) \\ & \text{sujeto a } y_{(0,\ldots,0)} = 1 \\ & \exists \mu \in \mathcal{M}(K) / y_\alpha = \int_K \boldsymbol{x}^\alpha d\mu, \forall \alpha \in \mathbb{N}^n \end{aligned} \tag{5.4}$$

Mostramos ahora una forma directa de ver que existen medidas atómicas óptimas siempre que f tenga un minimizador global en el dominio considerado. Luego veremos en qué casos podemos encontrarlas mediante esquemas de relajaciones.

Teorema 5.1.2 *Sean $K \subset \mathbb{R}^n$ un conjunto boreliano no vacío, $f \in \mathbb{R}[\boldsymbol{x}]$ que tiene valor mínimo $f_K^* \in \mathbb{R}$ en K. Sea μ una medida $r-$atómica, con $r \in \mathbb{N}_{>0}$, tal que todos sus átomos son mínimos globales de f en K y además $\int_K d\mu = 1$. Entonces, μ es una solución óptima de (5.2).*

Demostración: Como μ es $r-$atómica, en particular es boreliana, no negativa y finita, y la hipótesis sobre la integral la vuelve solución factible de (5.2). Por otro lado, de la Proposición 5.1.1 deducimos que μ es óptima puesto que f_K^* es valor óptimo de (5.1) y además

$$\int_K f d\mu = \sum_{k=1}^{r} \beta_k f(\boldsymbol{x}(k)) = \sum_{k=1}^{r} \beta_k f_K^* = f_K^*$$

□

En particular, si para $f \in \mathbb{R}[\boldsymbol{x}]$ existen un valor mínimo global f_K^* en K y $\boldsymbol{x}_0 \in K$ que lo realice, entonces $\delta_{\boldsymbol{x}_0}$ es una medida óptima $1-$atómica por el Teorema 5.1.2. Estas hipótesis siempre se cumplen si K es compacto y no vacío, es decir, en el caso que examinamos en la Sección 5.3.

5.2. Optimización polinomial sin restricciones

Buscamos ahora optimizar $f \in \mathbb{R}[\boldsymbol{x}]$ en el caso en que $K = \mathbb{R}^n$, y para eso podemos apoyarnos en la Proposición 5.1.1, que nos da un GMP equivalente. Notamos en primer lugar que f se puede asumir no constante (si no, el problema estaría trivialmente resuelto) y de grado par (porque, de lo contrario, se puede elegir $(x_k)_{k\geq 1} \subset \mathbb{R}^n$ tal que $f(x_k) \to -\infty$). Es decir, podemos tomar $r \in \mathbb{N}_{>0}$ tal que $\deg(f) = 2r$.

Como estamos trabajando con el problema de momentos no restringido, para $i \geq r$ aplicamos la relajación primal (4.7) al problema (5.4). Debemos agregar que $y_{(0,\dots,0)} = \int_K d\mu = 1$. De esta manera, la relajación primal resulta

$$\begin{aligned} \rho_i = \inf_{\boldsymbol{y}\in\mathbb{N}^n_{2i}} & L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & M_i(\boldsymbol{y}) \succeq 0 \\ & y_{(0,\dots,0)} = 1 \end{aligned} \tag{5.5}$$

Anotamos por otro lado el problema dual fortalecido (4.4) que corresponde al problema dual de momentos (5.3). Recordamos que $\mathbb{R}^n$ es un conjunto semialgebraico básico, el cual podemos definir como el conjunto de no negatividad de $g_1 \equiv 1$. El problema dual fortalecido nos queda entonces

$$\begin{aligned} \rho_i^* = & \sup_{\lambda\in\mathbb{R}} \lambda \\ \text{sujeto a } & f - \lambda = f_0 \\ & f_0 \in \Sigma[\boldsymbol{x}], \deg(f_0) \leq 2i \end{aligned} \tag{5.6}$$

Vemos que aparece sólo f_0 porque no hay polinomios g_k no triviales que definan $K = \mathbb{R}^n$, por lo que sólo necesitamos imponer una restricción de grado sobre f_0.

Aunque estos problemas modificados vienen definidos en el capítulo anterior como una secuencia con $i \geq i_0$, en el problema sin restricciones nos alcanza con estudiar un solo caso: el de $i = r$. Esto es porque $f-\lambda$ tiene grado $2r$, y por lo tanto ninguno de los f_k^2 que se suman para formarlo, con f_k polinomios, puede tener grado mayor a $2r$ (Lema 2.1.4).

A continuación mostramos que en este contexto vale la dualidad fuerte en los problemas SDP primal y dual.

Proposición 5.2.1 *Los valores ρ_r, ρ_r^* de los problemas (5.5) y (5.6) verifican $\rho_r = \rho_r^*$. Además, si $\rho_r > -\infty$, entonces (5.6) tiene solución óptima.*

Demostración: Debemos buscar $\boldsymbol{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2r}}$ tal que $M_r(\boldsymbol{y}) \succ 0$. Si conseguimos esto, el resultado se sigue del Teorema 1.2.14.
Para ver esto, sea μ la medida de probabilidad de un vector con distribución normal multivariada en n variables $X \sim \mathcal{N}(0, Id)$. Esta medida tiene densidad g estrictamente positiva en $\mathbb{R}^n$ y sus momentos de todos los órdenes son finitos (se calculan en [23, Teorema 1.1]).
Ahora, sean $q \in \mathbb{R}[\boldsymbol{x}]$ de grado a lo sumo r y $\tilde{q}$ su vector de coeficientes indexados por *ord*. Recordamos de la demostración del Corolario 3.2.4 que

$$\tilde{q}^t M_r(y) \tilde{q} = \int_{\mathbb{R}^n} q^2 d\mu = \int_{\mathbb{R}^n} q(\boldsymbol{x})^2 g(\boldsymbol{x}) d\boldsymbol{x} > 0 \text{ si } q \neq 0$$

En este caso, $\boldsymbol{y}$ es la secuencia de momentos truncada de la distribución normal mencionada, y el valor calculado es positivo porque q es continuo, no nulo y $g > 0$. De esto se sigue que $M_r(y) \succ 0$. □

En el caso sin restricciones, para encontrar dualidad fuerte tuvimos que recurrir a una condición de factibilidad estricta porque el dominio no es compacto.

Vamos a encontrar ahora el mayor valor que se le puede restar a f para que sea suma de cuadrados, con la única condición de que un problema SDP sea factible.

Teorema 5.2.2 *Sea $f \in \mathbb{R}[\boldsymbol{x}]$ de grado $2r$. Si (5.6) tiene solución factible $\lambda \in \mathbb{R}$, entonces $\rho_r^* \in \mathbb{R}$; existe $f^* \in \mathbb{R}$ ínfimo de f en $\mathbb{R}^n$, y además $f - \rho_r^* \in \Sigma[\boldsymbol{x}]$.*

Demostración: Como el problema dual es factible, se sigue que $\rho_r^* > -\infty$. Veamos que $\rho_r^* < +\infty$. En efecto, dado $\boldsymbol{x}_0 \in \mathbb{R}^n$, tenemos que $f(\boldsymbol{x}) - (f(\boldsymbol{x}_0) + 1)$ vale -1 en $\boldsymbol{x}_0$, así que $f - \lambda$ no es suma de cuadrados para $\lambda \geq f(\boldsymbol{x}_0) + 1$. Por lo tanto, $\rho_r^* < +\infty$.
Veamos ahora que $f(\boldsymbol{x}) \geq \rho_r^*, \forall \boldsymbol{x} \in \mathbb{R}^n$, y en particular existe el ínfimo de f. Si no, sea $\boldsymbol{x}_0 \in \mathbb{R}^n$ tal que esto no vale y $\lambda \in (f(\boldsymbol{x}_0), \rho_r^*)$. Luego, $f - \lambda$ no es suma de cuadrados por dar negativo en $\boldsymbol{x}_0$. Pero dado $\varepsilon = \frac{\rho_r^* - \lambda}{2}$, existe $\lambda_1 \in (\rho_r^* - \varepsilon, \rho_r^*)$ tal que $f - \lambda_1$ es suma de cuadrados por definición de ρ_r^*, y esto contradice la deducción de que $f - \lambda$ no lo era. Luego, el ínfimo de f es finito. Por último, por la Proposición 5.2.1 tenemos que $\rho_r = \rho_r^*$ es un valor finito y por lo tanto (5.6) tiene solución óptima; pero esta es de hecho $\lambda = \rho_r^*$. En particular, es factible. Es decir, $f - \rho_r^* \in \Sigma[\boldsymbol{x}]$. □

En este último resultado se puede ver que si el problema SDP dual es factible entonces siempre f difiere en una constante de una suma de cuadrados. También notamos que, con esta hipótesis, ρ_r^* no es sólo un supremo en (5.6), sino de hecho un máximo.

El siguiente resultado es muy importante porque da condiciones para que un problema de optimización polinomial sin restricciones tenga solución óptima y una forma de encontrar candidatos a soluciones.

Teorema 5.2.3 *Sea $f \in \mathbb{R}[\boldsymbol{x}]$ de grado $2r$ con ínfimo global $f^* > -\infty$. Si $f - f^* \in \Sigma[\boldsymbol{x}]$, entonces el problema (5.1) equivale a la relajación primal (5.5), es decir, $f^* = \rho_r = \rho_r^*$. Más aún, si existe $\boldsymbol{x}^* \in \mathbb{R}^n$ mínimo global de f, entonces*

$$\boldsymbol{y}^* = v_{2r}(\boldsymbol{x}^*) = (1, x_1^*, \ldots, x_n^*, (x_1^*)^2, x_1^* x_2^*, \ldots, (x_1^*)^{2r}, \ldots, (x_n^*)^{2r})$$

es un minimizador de (5.5) con $i = r$.

Demostración: Primero veamos que $f^* = \rho_r^*$. En efecto, ya vimos en la demostración del Teorema 5.2.2 que $f^* \geq \rho_r^*$, y además se tiene $f^* \leq \rho_r^*$ porque f^* es solución factible de (5.6) por hipótesis. Luego, $f^* = \rho_r^* = \rho_r$, donde la última igualdad sale de la Proposición 5.2.1. Por último, como por hipótesis existe $\boldsymbol{x}^* \in \mathbb{R}^n$ mínimo global de f, entonces $\mu = \delta_{\boldsymbol{x}^*}$ es una medida óptima $1-$atómica para (5.2) como vimos al final de la Sección 5.1. Además, por (4.9) su secuencia de momentos hasta orden $2r$, indexada por *ord* y dada como vector, es $v_{2r}(\boldsymbol{x}^*)$. Luego, esta secuencia es solución factible y óptima de la relajación (5.5) con $i = r$. □

Para empezar, este resultado nos indica que si $f - f^*$ es suma de cuadrados entonces el valor mínimo f^* se calcula de forma exacta resolviendo la relajación semidefinida con $i = r$, y el ejemplo de solución óptima de la relajación nos indica que existe alguna si hay un minimizador global de f. Por otra parte, en este caso nos da una forma de encontrar candidatos a minimizadores globales de f: son los vectores dados por las coordenadas 2 a $n + 1$ de una solución óptima de (5.5), dado que estas coordenadas corresponden a los monomios evaluados $\boldsymbol{x}_i^*, i = 1, \ldots, n$. Fuera de las secuencias de momentos óptimas de la relajación primal no hay minimizadores de f. Notamos también que en el caso de polinomios de una variable (Teorema 2.1.2) y los otros casos mencionados en la Sección 2.1.1, ser no negativo equivale a ser suma de cuadrados, así que siempre se da la condición del teorema que acabamos de ver porque $f - f^* \geq 0$.

Veamos qué ocurre si $f - f^*$ no es suma de cuadrados.

Proposición 5.2.4 *Sea $f \in \mathbb{R}[\boldsymbol{x}]$ de grado $2r$ con ínfimo global $f^* > -\infty$. Si $f - f^* \notin \Sigma[\boldsymbol{x}]$, entonces $f^* > \rho_r$. En particular, la relajación primal (5.5) no resuelve el problema de minimización de f.*

Demostración: Si $\rho_r = -\infty$, la afirmación es trivialmente cierta. Suponemos $\rho_r > -\infty$. Ya sabemos que $\rho_r = \rho_r^*$, y que por lo tanto $f - \rho_r \in \Sigma[\boldsymbol{x}]$. De esto se deduce que $f^* > \rho_r^*$, porque si suponemos lo contrario se tiene

$$f - f^* = f - \rho_r^* + (\rho_r^* - f^*) = f - \rho_r^* + (\sqrt{\rho_r^* - f^*})^2$$

que resulta suma de cuadrados, y esto contradice la hipótesis. □

A continuación damos una condición suficiente tomada de [11, Teorema 5.5] para que un valor óptimo de la relajación primal sea efectivamente el mínimo global de un polinomio.

Teorema 5.2.5 *Sea $f \in \mathbb{R}[\boldsymbol{x}]$ de grado $2r$, de forma que el valor óptimo ρ_r de (5.5) se alcanza en $\boldsymbol{y}^* = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2r}}$. Si se verifica que $\mathrm{rg}(M_r(\boldsymbol{y}^*)) = \mathrm{rg}(M_{r-1}(\boldsymbol{y}^*))$, entonces $f^* = \rho_r$ y además existen al menos $\mathrm{rg}(M_r(\boldsymbol{y}^*))$ mínimos globales diferentes de f en $\mathbb{R}^n$.*

Demostración: Ya sabemos por la demostración del Teorema 5.2.2 y el Teorema 5.2.1 que $\rho_r = \rho_r^* \leq f^*$. Además, por el Teorema 3.2.7 sabemos que $\boldsymbol{y}^*$ tiene una medida representativa μ^* con $t = \mathrm{rg}(M_r(\boldsymbol{y}^*))$ átomos en $\mathbb{R}^n$. Asimismo tenemos que $\rho_r = L_{\boldsymbol{y}^*}(f) = \int_{\mathbb{R}^n} f d\mu^*$, por lo que μ^* es una solución óptima de (5.2). En efecto, sabemos que μ^* es factible porque, como $\boldsymbol{y}^*$ es factible, vale que $y^*_{(0,\dots,0)} = \int_{\mathbb{R}^n} d\mu = 1$. Luego, $\rho_r \geq \rho_{mom} = f^*$ porque ρ_{mom} es el ínfimo de valores factibles de (5.2). Finalmente, $\rho_r = f^*$.
Falta probar que los átomos de μ^* son minimizadores globales de f. Sean $\{\boldsymbol{x}^*(k)\}_{k=1}^t$ estos átomos. Sabemos entonces que

$$\mu^* = \sum_{k=1}^{t} \beta_k \delta_{\boldsymbol{x}^*(k)}, \beta_k > 0, \forall k = 1, \dots, t; \sum_{k=1}^{t} \beta_k = 1$$

donde la última igualdad se deduce de que μ^* es factible. En suma, tenemos que

$$f^* = \rho_{mom} = \int_{\mathbb{R}^n} f d\mu^* = \sum_{k=1}^{t} \beta_k f(\boldsymbol{x}^*(k))$$

Como $f(\boldsymbol{x}^*(k)) \geq f^*$ para todos los k y los coeficientes β_k son positivos y suman 1, necesariamente $f(\boldsymbol{x}^*(k)) = f^*, \forall 1 \leq k \leq t$. □

Pasamos ahora a dar un algoritmo para buscar el valor mínimo de un polinomio f y mínimos globales. Luego repasaremos por qué el algoritmo es correcto y su utilidad.

Algoritmo 5.2.6 *Input: Un polinomio $f \in \mathbb{R}[\boldsymbol{x}]$ de grado $2r$.*
Output: si el ínfimo global f^* de f existe: ese valor y minimizadores globales, o bien una cota inferior $\rho_r \leq f^*$ (puede ocurrir que $\rho_r = -\infty$, en cuyo caso no se obtiene información sobre si f tiene ínfimo global o no).

1. *Buscar una solución factible de (5.6) con $i = r$. Si no existe, el algoritmo no responde si f tiene ínfimo global o no. Terminar la ejecución.*

2. *Resolver el problema (5.5) para $i = r$ con ínfimo $\rho_r = \rho_r^* > -\infty$. Se define $\boldsymbol{y}^* = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2r}}$ como la solución óptima encontrada.*

3. *Si $\operatorname{rg}(M_r(\boldsymbol{y}^*)) = \operatorname{rg}(M_{r-1}(\boldsymbol{y}^*))$, devolver $f^* = \rho_r$, computar $\operatorname{rg}(M_r(\boldsymbol{y}^*))$ mínimos globales de f como los átomos que se obtienen del Algoritmo 4.2.5 y terminar la ejecución.*

4. *Devolver ρ_r como cota inferior de f^*.*

En efecto, ya vimos que la única relajación semidefinida que necesitamos resolver es con $i = r$. También dedujimos que $\rho_r = \rho_r^* \leq f^*$, incluyendo la posibilidad de que $\rho_r^* = \rho_r = -\infty$. Si llegamos al paso 2 es que el problema SDP dual resulta factible, y por lo tanto ρ_r es finito. En teoría el valor ρ_r podría no realizarse, pero el algoritmo dará numéricamente una solución óptima, para la que luego se va a chequear la condición de rango. En el paso 3, ya sabemos del Capítulo anterior que si se da esta hipótesis de rango el algoritmo de extracción nos da la cantidad citada de átomos, y probamos que los átomos son minimizadores globales de f^*. Si encontramos $\boldsymbol{y}^*$ pero no vale la hipótesis del rango, entonces $\boldsymbol{y}^*$ no admite medida representativa $\operatorname{rg}(M_r(\boldsymbol{y}^*))$-atómica porque el Teorema de la Extensión Plana 3.2.7 es una equivalencia. Luego, no podemos ponernos a buscar átomos.

Ya mencionamos la posibilidad de que f tenga ínfimo global pero no mínimos. Damos un ejemplo de que esto efectivamente puede ocurrir. Sea $f(x_1, x_2) = (x_1 x_2 - 1)^2 + x_1^2$. Se puede comprobar de forma directa que f es suma de cuadrados, que su único punto crítico es $(0, 0)$, donde vale 1, que su ínfimo es 0 (basta elegir $(x_{n,1}, x_{n,2}) = (1/n, n)$) y que este no se realiza (ya que $(0, 0)$ era el único candidato a mínimo).

Por último, tener $\rho_r = -\infty$ no asegura que $f^* = -\infty$, aunque sí vale la implicación contraria porque demostramos que $\rho_r \leq f(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathbb{R}^n$. En efecto, si tomamos $f \in \mathbb{R}[x_1, x_2], f(x_1, x_2) = x_1^2 x_2^2 (x_1^2 + x_2^2 - 1)$, este polinomio es de grado 6 y por lo tanto $r = 3$. En este caso, $\rho_3 = -\infty$ pero se tiene $f^* = -\frac{1}{27}$ y 4 mínimos globales [11, Ejemplo 5.2]. Por el resultado de dualidad fuerte visto en este capítulo, tenemos también que $\rho_3^* = -\infty$, es decir que el problema SDP dual no es factible. En otras palabras, no existe $\lambda \in \mathbb{R}$ tal que $f - \lambda$ sea suma de cuadrados.

5.3. Optimización polinomial en conjuntos semialgebraicos básicos compactos

Ahora abordamos el problema (5.1) con $K \subset \mathbb{R}^n$ semialgebraico básico compacto. Lo primero que debemos mencionar es que un polinomio en este tipo de conjuntos, si son no vacíos, siempre posee mínimo y puntos que lo realizan por el Teorema de Weierstrass. En segundo lugar, notamos que ya no se puede asumir que el polinomio tiene grado par: $f \in \mathbb{R}[\boldsymbol{x}]$ tendrá grado $2v_0$ o $2v_0-1, v_0 \in \mathbb{N}_{>0}$. También observamos que el problema de momentos siempre tiene medidas óptimas en este caso: por el Teorema 5.1.2, una medida soportada en mínimos de f en K es óptima, y estos mínimos siempre existen porque K es compacto. Sin embargo, el esquema de relajaciones que estudiamos podría no encontrar ninguna.

Como en el Capítulo 4, vamos a trabajar en las condiciones donde se puede aplicar el Certificado de Positividad de Putinar, y comentar cómo cambiarían los resultados en caso de necesitar el de Schmüdgen.

Vamos a aplicar la relajación primal (4.2) para el problema (5.4); damos también la versión dual más estricta, como se sigue de (4.4). El procedimiento para intercambiar supremo e ínfimo es análogo al que realizamos en la Sección 5.1 para el GMP primal y su dual.

$$\begin{aligned} \rho_i = & \inf_{\boldsymbol{y} \in \mathbb{N}^n_{2i}} L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & y_{(0,\dots,0)} = 1 \\ & M_i(\boldsymbol{y}) \succeq 0 \\ & M_{i-v_k}(g_k, \boldsymbol{y}) \succeq 0, k = 1, \dots, m \end{aligned} \tag{5.7}$$

$$\begin{aligned} \rho_i^* = & \sup_{\lambda \in \mathbb{R}^\Gamma} \lambda \\ \text{sujeto a } & f - \lambda = f_0 + \sum_{k=1}^{m} f_k g_k \\ & f_k \in \Sigma[\boldsymbol{x}], \deg(f_k) \leq 2(i - v_k), k = 0, \dots, m \end{aligned} \tag{5.8}$$

Vemos que la última restricción desapareció porque $\Gamma_+ = \emptyset$.

Ya vimos en el capítulo anterior que (5.7) es una relajación del problema de momentos, la única modificación es que copiamos del problema original la condición de que el primer momento sea 1. Ahora derivamos un resultado de convergencia de lo visto en el Capítulo 4:

Teorema 5.3.1 *Sean $f, g_1, \dots, g_m \in \mathbb{R}[\boldsymbol{x}]$, $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \geq 0, \forall i = 1, \dots, m\}$ tal que $Q(g_1, \dots, g_m)$ es arquimediano; y sean las relajaciones primales (5.7) con $i \geq$*

i_0 (definido cuando dimos la relajación (4.2)). Sea f_K^* el valor mínimo de f en K. Entonces, $\rho_i \nearrow f_K^*$.

Demostración: Observamos que, en (5.2), las funciones de restricción son h_0 y h_1, con $h_0 = 1 > 0$ en $\mathbb{R}^n$, y en particular en K. Por lo tanto, el presente resultado se sigue del item 1 del Teorema 4.1.5. El valor $\rho_{mom} = f_K^*$ (Teorema 5.1.1) resulta finito porque K es compacto (Proposición 2.2.4), y la convergencia pasa de decreciente a creciente por el cambio de signo operado en la función objetivo cuando pasamos de buscar supremo a ínfimo. □

Una pregunta natural que podemos hacernos es si esta convergencia siempre ocurre en finitas iteraciones (o sea, si existe $i \geq i_0$ tal que $\rho_i = f_K^*$), o si por lo menos existen hipótesis extra que garanticen esto. Resulta que sí: estas hipótesis adicionales se derivan de las condiciones KKT de optimalidad global. Presentamos ese resultado, publicado por Nie en 2014 [16, Teorema 1.1], el cual asume $Q(g_1, \ldots, g_m)$ arquimediano. No damos la demostración porque escapa al área temática de este trabajo.

Teorema 5.3.2 *Sean $f \in \mathbb{R}[\boldsymbol{x}]$, $K \subset \mathbb{R}^n$ como en el Teorema 5.3.1. Además, suponemos que para todo $\boldsymbol{x}^* \in \mathbb{R}^n$ mínimo global de f en K valen las siguientes condiciones:*

1. *Los gradientes $\nabla g_j(\boldsymbol{x}^*), j = 1, \ldots, m$ son linealmente independientes (y por lo tanto existen $\sigma_j^*, j = 1, \ldots, m$ tales que $\nabla f(\boldsymbol{x}^*) - \sigma_j^* \nabla g_j(\boldsymbol{x}^*) = 0$ y además $\sigma_j^* g(\boldsymbol{x}^*) = 0, \forall j = 1, \ldots, m$);*
2. *$\forall j = 1, \ldots, m : g_j(\boldsymbol{x}^*) = 0 \Rightarrow \sigma_j^* > 0$ (complementariedad estricta).*
3. *Se cumple la siguiente condición suficiente de segundo orden: $a^t b > 0$, donde $b = \nabla_x^2(f(\boldsymbol{x}^*) - \sum_{j=1}^m \sigma_j^* g_j(\boldsymbol{x}^*))a$ y a es cualquier vector no nulo ortogonal a $\nabla(f(\boldsymbol{x}^*) - \sum_{j=1}^m \sigma_j^* g_j(\boldsymbol{x}^*))$.*

Entonces, $f - f_K^* \in Q(g_1, \ldots, g_m)$.

Falta ver por qué esto implica la convergencia en finitas iteraciones. Notamos que, bajo las condiciones del teorema, f_K^* es una solución factible, y por lo tanto óptima, del problema (5.8) (un valor λ más grande daría $f - \lambda$ negativo en algunos puntos de K, así que la función no podría estar en $Q(g_1, \ldots, g_m)$); sólo debemos tomar algún $i \geq i_0$ lo suficientemente grande para que $\deg(f_k) \leq 2(i - v_k), \forall k = 1, \ldots, m$. Recordamos que por la dualidad débil de los problemas SDP tenemos $f_K^* = \rho_i^* \leq \rho_i$, y además tenemos $\rho_i \leq f_K^* = \rho_{mom}$ puesto que ρ_i es el resultado de una relajación primal. En suma, $\rho_i = f_K^*$, y esta indica que se tiene el valor exacto en la iteración correspondiente.

Ahora pasamos a detallar otro hecho similar a algo visto en la sección anterior: cuando se satisfagan las hipótesis necesarias, se podrá hallar mínimos globales en K como

átomos de una medida representativa. La demostración también se puede ver en [11, Teorema 5.7].

Teorema 5.3.3 *Sea $f \in \mathbb{R}[\boldsymbol{x}]$, $i \geq i_0$ tal que el ínfimo ρ_i de la relajación (5.7) se realiza en una solución óptima $\boldsymbol{y}^*$. Sea $v = \max_{j=1,\ldots,m} v_j$. Si existe $i_0 \leq s \leq i$ tal que $\operatorname{rg}(M_s(\boldsymbol{y}^*)) = \operatorname{rg}(M_{s-v}(\boldsymbol{y}))$, entonces $\rho_r = f_K^*$ y existen al menos $r = \operatorname{rg}(M_s(\boldsymbol{y}))$ mínimos globales diferentes de f en K. Se los puede obtener como output del Algoritmo 4.2.5.*

Demostración: Ya sabemos que $\rho_{mom} = f_K^*$, y por el Teorema 4.1.5 y su demostración vimos que $\rho_{mom} = \rho_i$ (una vez hecho el cambio de signo por pasar de buscar supremo a ínfimo siguen valiendo $\rho_{mom} \leq \rho_i, \rho_{mom} \geq \rho_i$), y además puede aplicarse el algoritmo de extracción. Sólo nos falta probar que los átomos obtenidos son mínimos globales de f en K.
Sean $\boldsymbol{x}^*(k), k = 1, \ldots, r$ los átomos, de los cuales sabemos que están en K porque el soporte de μ^* está contenido en ese conjunto. Entonces, dada μ^* la medida representativa de $\boldsymbol{y}^*$ se tiene

$$\mu^* = \sum_{k=1}^{r} \beta_k \delta_{\boldsymbol{x}^*(k)}, \beta_k > 0, \forall k = 1, \ldots, r, \sum_{k=1}^{r} \delta_k = 1$$

donde la última igualdad se desprende de que μ^* es una medida factible de (5.2) por tener primer momento igual a 1. Utilizando todo esto, la Proposición 3.2.3 y la fórmula (4.9) se deduce lo siguiente:

$$f_K^* \geq \rho_i = L_{\boldsymbol{y}^*}(f) = \int_K f d\mu^* = \sum_{k=1}^{r} \beta_k f(\boldsymbol{x}^*(k)) \geq \sum_{k=1}^{r} \beta_k f_K^* = f_K^*$$

donde la última desigualdad sale de que f_K^* es cota inferior de todos los $f(\boldsymbol{x}^*(k))$ y de que los β_k suman 1. Por lo tanto, necesariamente se tiene $f(\boldsymbol{x}^*(k)) = f_K^*$ para todo $k = 1, \ldots, r$. □

La nueva relajación primal para el caso donde K es compacto pero $Q(g_1, \ldots, g_m)$ no es arquimediano, obtenida a partir de (4.5), queda como sigue:

$$\begin{aligned} \rho_i = & \sup_{\boldsymbol{y} \in \mathbb{N}^n_{2i}} L_{\boldsymbol{y}}(f) \\ \text{sujeto a } & y_{(0,\ldots,0)} = 1 \\ & M_i(\boldsymbol{y}) \succeq 0 \\ & M_{i-v_J}(g_J, \boldsymbol{y}) \succeq 0 \wedge g_J = \prod_{k \in J} g_k, v_J = \deg(g_J), J \subset \{1, \ldots, m\} \end{aligned} \tag{5.9}$$

y el problema dual deducido de (4.6) queda así:

$$\begin{aligned} \rho_i^* &= \sup_{\lambda \in \mathbb{R}^\Gamma} \sum_{j \in \Gamma} \lambda_j \gamma_j \\ \text{sujeto a } f - \lambda &= \sum_{J \subset \{1,\dots,m\}}^{m} f_J g_J \\ & f_J \in \Sigma[\boldsymbol{x}], \deg(f_J) \leq 2(i - v_J), \forall J \subset \{1, \dots, m\} \end{aligned} \tag{5.10}$$

Con los resultados que vimos en esta sección, podemos concluir que si K es semialgebraico básico compacto se puede obtener f_K^* o una cota inferior ρ_N para algún $N \in \mathbb{N}$. Para esto, ejecutamos una variante que vamos a describir a continuación del Algoritmo 4.3.1. Notar que tenemos dualidad fuerte, es decir, $\rho_i = \rho_i^*$ para todo $i \geq i_0$ porque como $h_0 \equiv 1$ se verifican todas las hipótesis del Teorema 3.1.4. Por lo tanto sabemos que $\rho_i = -\infty$ si y sólo si el problema dual correspondiente no es factible. Si se llega a ejecutar el paso 3, en caso de que $\rho_i = f_K^*$ siempre existe una secuencia de momentos que realiza ρ_i porque hay al menos un $\boldsymbol{x}^* \in K$ que realiza f_K^*, y de esto se deduce una medida 1-atómica óptima por el Teorema 5.1.2. En cambio, si $\rho_i < f_K^*$, en teoría la relajación primal puede no tener solución óptima, pero numéricamente se obtendrá una.

Algoritmo 5.3.4 *Input: Un polinomio f de grado $2v_0$ o $2v_0 - 1$, un conjunto finito de polinomios $\{h_j\}_{j \in \Gamma}$ de grados $2w_j$ o $2w_j - 1, j \in \Gamma$; Un conjunto de polinomios $\{g_k\}_{1 \leq k \leq m}$ de grados $2v_k$ o $2v_k - 1, 1 \leq k \leq m$ que definen el conjunto $K = \{\boldsymbol{x} \in \mathbb{R}^n : g_k(\boldsymbol{x}) \geq 0, \forall 1 \leq k \leq m\}$ y tales que $Q(g_1, \dots, g_m)$ es arquimediano; y $N \in \mathbb{N}$ una cota superior del índice de las relajaciones semidefinidas a ejecutar.*
Output: f_K^* el mínimo de f en K y un conjunto de puntos $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r \in \mathbb{R}^n$ que forman el soporte de una medida óptima μ; o bien un valor $\rho_N \leq f_K^*$.

1. *Calcular $v = \max_{k=1,\dots,m} v_k$ y $i_0 = \max\{v_0, v, \max_{j \in \Gamma} w_j\}$. Definir $i = i_0$.*
2. *Buscar una solución factible del problema (5.8). Si no se encuentra, definir $\rho_i = -\infty$ y saltar al paso 6.*
3. *Resolver la relajación semidefinida (5.7) con índice i y valor óptimo ρ_i.*
4. *Sea $\boldsymbol{y}^*$ la solución óptima asociada a ρ_i. Si no existe $i_0 \leq s \leq i$ tal que $\operatorname{rg}(M_s(\boldsymbol{y}^*)) = \operatorname{rg}(M_{s-v}(\boldsymbol{y}^*))$, saltar al paso 6.*
5. *Para s el valor hallado que cumple la condición de rango verificada en el paso 4, retornar $f_K^* = \rho_i$, $r = \operatorname{rg}(M_s(\boldsymbol{y}))$ y retornar los puntos $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r$ extraídos con el Algoritmo 4.2.5 como minimizadores de f en K. Terminar la ejecución.*
6. *Si $i < N$, aumentar i en 1 y volver al paso 2.*

7. *Retornar ρ_N como cota inferior de f_K^*. Terminar la ejecución.*

El algoritmo está construido asumiendo $Q(g_1, \ldots, g_m)$ arquimediano. En caso de que no lo sea, si K es compacto se puede utilizar el algoritmo sin cambios para obtener cotas inferiores o eventualmente encontrar el valor mínimo y mínimos globales (si se verifica la hipótesis de rango). Sin embargo, para tener convergencia de los ρ_r a f^* se requiere adaptar las relajaciones primales y sus problemas duales al formato de Schmüdgen; este cambio es necesario para garantizar que las sucesivas relajaciones realmente se acerquen a encontrar f_K^*.

Conclusiones

En esta tesis examinamos la teoría necesaria para estudiar la optimización de polinomios en $\mathbb{R}^n$ y en conjuntos semialgebraicos básicos compactos como un caso particular del GMP.

En el caso sin restricciones, vimos que se obtiene toda la información posible de una sola relajación, y que la misma resuelve el problema de forma exacta si $f - f^*$ existe y es suma de cuadrados. Incluso cuando esto no ocurre, hemos encontrado el mayor valor $\lambda \in \mathbb{R}$ tal que $f - \lambda$ es suma de cuadrados, con la única condición pedida de que un problema SDP dual sea factible. También hemos encontrado, gracias al Teorema de la Extensión Plana, una condición suficiente de rango matricial, y por lo tanto computable, para garantizar que una solución óptima de la única relajación da una solución exacta y minimizadores para el problema de optimización polinomial sin restricciones.

Para conjuntos semialgebraicos básicos compactos, hemos aplicado la familia de relajaciones semidefinidas planteada en el Capítulo 4. Vimos que, planteando una versión u otra de estas relajaciones dependiendo de una condición sobre $Q(g_1, \ldots, g_m)$, los valores óptimos de las relajaciones convergen de forma creciente a f_K^*. También en este caso existe una condición suficiente de rango matricial para garantizar que una relajación semidefinida obtiene f_K^* de forma exacta y da una forma de encontrar minimizadores.

Probamos también que el GMP aplicado a optimización polinomial tiene siempre alguna medida óptima átomica si el ínfimo de f se realiza, y vimos algunas condiciones adicionales que garantizan encontrar una en finitas iteraciones de la familia de relajaciones.

Un posible trabajo a futuro es examinar en detalle la complejidad computacional de los algoritmos que vimos y cómo se puede mejorar en algunos casos aprovechando propiedades de las matrices involucradas (por ejemplo ser ralas, [11, Subsección 4.6.1]). Además nuestro trabajo se limitó a relajaciones dadas como problemas SDP, pero estas pueden ser lineales si K tiene ciertas propiedades [11, Sección 4.4].

Bibliografía

[1] Sébastien Bubeck et al. "Convex optimization: Algorithms and complexity". En: *Foundations and Trends® in Machine Learning* 8.3-4 (2015), págs. 231-357.

[2] Man-Duen Choi, Tsit Yuen Lam y Bruce Reznick. "Sums of squares of real polynomials". En: *Proceedings of Symposia in Pure mathematics*. Vol. 58. American Mathematical Society. 1995, págs. 103-126.

[3] Robert M Corless. "Gröbner bases and matrix eigenproblems". En: *ACM SIGSAM Bulletin* 30.4 (1996), págs. 26-32.

[4] Robert M Corless, Patrizia M Gianni y Barry M Trager. "A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots". En: *Proceedings of the 1997 international symposium on Symbolic and algebraic computation*. 1997, págs. 133-140.

[5] Raúl E Curto y Lawrence A Fialkow. "Truncated K-moment problems in several variables". En: *Journal of Operator Theory* (2005), págs. 189-226.

[6] Gene H Golub y Charles F Van Loan. *Matrix computations*. JHU press, 2013.

[7] Ralph P Grimaldi. *Matemáticas discretas y combinatoria: una introducción con aplicaciones*. Pearson Educación, 1998.

[8] EK Haviland. "On the momentum problem for distribution functions in more than one dimension. II". En: *American Journal of Mathematics* 58.1 (1936), págs. 164-168.

[9] Didier Henrion y Jean-Bernard Lasserre. "Detecting global optimality and extracting solutions in GloptiPoly". En: *Positive polynomials in control*. Springer, 2005, págs. 293-310.

[10] Jean B Lasserre. "Global optimization with polynomials and the problem of moments". En: *SIAM Journal on optimization* 11.3 (2001), págs. 796-817.

[11] Jean Bernard Lasserre. *Moments, positive polynomials and their applications*. Vol. 1. World Scientific, 2009.

[12] Monique Laurent. "Revisiting two theorems of Curto and Fialkow on moment matrices". En: *Proceedings of the American Mathematical Society* 133.10 (2005), págs. 2965-2976.

[13] Monique Laurent y Frank Vallentin. "Semidefinite optimization". En: *Lecture Notes, available at http://page. mi. fu-berlin. de/fmario/sdp/laurentv. pdf* (2012).

[14] Henri Lombardi, Daniel Perrucci y Marie-Françoise Roy. *An elementary recursive bound for effective Positivstellensatz and Hilbert's 17th problem*. Vol. 263. 1277. American mathematical society, 2020.

[15] Murray Marshall. *Positive polynomials and sums of squares*. 146. American Mathematical Soc., 2008.

[16] Jiawang Nie. "Optimality conditions and finite convergence of Lasserre's hierarchy". En: *Mathematical programming* 146 (2014), págs. 97-121.

[17] Pablo A Parrilo. "Semidefinite programming relaxations for semialgebraic problems". En: *Mathematical programming* 96 (2003), págs. 293-320.

[18] Alexander Prestel. *Lectures on formally real fields*. Vol. 1093. Springer, 2007.

[19] Mihai Putinar. "Positive polynomials on compact semi-algebraic sets". En: *Indiana University Mathematics Journal* 42.3 (1993), págs. 969-984.

[20] Konrad Schmüdgen et al. *The moment problem*. Vol. 9. Springer, 2017.

[21] Gilbert Stengle. "A nullstellensatz and a positivstellensatz in semialgebraic geometry". En: *Mathematische Annalen* 207 (1974), págs. 87-97.

[22] Richard Lee Wheeden y Antoni Zygmund. *Measure and integral*. Vol. 26. Dekker New York, 1977.

[23] Christopher S Withers. "The moments of the multivariate normal". En: *Bulletin of the Australian Mathematical Society* 32.1 (1985), págs. 103-107.

[24] Thomas Yuster. "The reduced row echelon form of a matrix is unique: A simple proof". En: *Mathematics Magazine* 57.2 (1984), págs. 93-94.