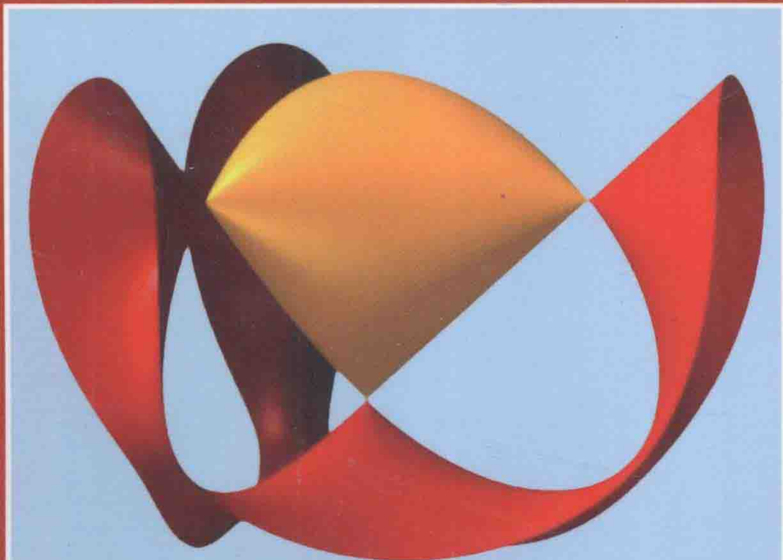
CAMBRIDGE TEXTS IN APPLIED MATHEMATICS

An Introduction to Polynomial and Semi-Algebraic Optimization



JEAN BERNARD LASSERRE

An Introduction to Polynomial and Semi-Algebraic Optimization

JEAN BERNARD LASSERRE
LAAS-CNRS and Institut de Mathématiques, Toulouse, France

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107060579

First published 2015

Printed in the United Kingdom by Clays, St Ives plc

A catalog record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data
Lasserre, Jean-Bernard, 1953- author.
An introduction to polynomial and semi-algebraic optimization / Jean Bernard Lasserre, LAAS-CNRS and Institut de mathématiques, Toulouse, France.
pages cm. – (Cambridge texts in applied mathematics)
Includes bibliographical references and index.
ISBN 978-1-107-06057-9 (Hardback) – ISBN 978-1-107-63069-7 (Paperback)
1. Polynomials. 2. Mathematical optimization. I. Title.
QA161.P59L37 2015
512.9′422–dc23 2014031786

ISBN 978-1-107-06057-9 Hardback
ISBN 978-1-107-63069-7 Paperback

An Introduction to Polynomial and Semi-Algebraic Optimization

This is the first comprehensive introduction to the powerful moment approach for solving global optimization problems (and some related problems) described by polynomials (and even semi-algebraic functions). In particular, the author explains how to use relatively recent results from real algebraic geometry to provide a systematic numerical scheme for computing the optimal value and global minimizers. Indeed, among other things, powerful positivity certificates from real algebraic geometry allow one to define an appropriate hierarchy of semidefinite (sum of squares) relaxations or linear programming relaxations whose optimal values converge to the global minimum. Several specializations and extensions to related optimization problems are also described.

Graduate students, engineers and researchers entering the field can use this book to understand, experiment and master this new approach through the simple worked examples provided

JEAN BERNARD LASSERRE is Directeur de Recherche at the LAAS-CNRS laboratory in Toulouse and a member of the Institute of Mathematics of Toulouse (IMT). He is a SIAM Fellow and in 2009 he received the Lagrange Prize, awarded jointly by the Mathematical Optimization Society (MOS) and the Society for Industrial and Applied Mathematics (SIAM).

Cambridge Texts in Applied Mathematics

All titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing, visit www.cambridge.org/mathematics.

Nonlinear Dispersive Waves
MARK J. ABLOWITZ

Flow, Deformation and Fracture
G. I. BARENBLATT

Hydrodynamic Instabilities
FRANÇOIS CHARRU

The Mathematics of Signal Processing
STEVEN B. DAMELIN & WILLARD MILLER, JR

An Introduction to Stochastic Dynamics
JINQIAO DUAN

Singularities: Formation, Structure and Propagation
J. EGGERS & M. A. FONTELOS

A First Course in Continuum Mechanics
OSCAR GONZALEZ & ANDREW M. STUART

A Physical Introduction to Suspension Dynamics
ÉLISABETH GUAZZELLI & JEFFREY F. MORRIS

Applied Solid Mechanics
PETER HOWELL, GREGORY KOZYREFF & JOHN OCKENDON

A First Course in the Numerical Analysis of Differential Equations (2nd Edition)
ARIEH ISERLES

Iterative Methods in Combinatorial Optimization
LAP CHI LAU, R. RAVI & MOHIT SINGH

An Introduction to Computational Stochastic PDEs
GABRIEL J. LORD, CATHERINE E. POWELL & TONY SHARDLOW

An Introduction to Parallel and Vector Scientific Computation
RONALD W. SHONKWILER & LEW LEFTON

To my daughter Julia, and to Carole ...

Preface

This book has benefited from several stimulating discussions, especially with M. Anjos, G. Blekherman, R. E. Curto, E. de Klerk, L. A. Fialkow, J. W. Helton, M. Kojima, S. Kuhlmann, M. Laurent, M. Marshall, T. Netzer, J. Nie, P. Parrilo, D. Pasechnich, D. Plaumann, V. Powers, M. Putinar, F. Rendl, B. Reznick, C. Riener, M.-F. Roy, C. Scheiderer, K. Schmüdgen, M. Schweighofer, B. Sturmfels, T. Theobald, K. Toh, and V. Vinnikov, during several visits and/or workshops in particular at the *Mathematical Sciences Research Institute* (MRSI, Berkeley), the *Institute of Pure and Applied Mathematics* (IPAM, Los Angeles), the *American Institute of Mathematics* (AIM, Palo Alto), the *Fields Institute for Research in Mathematical Sciences* (Toronto), the *Institute for Mathematics and its Applications* (IMA, Minneapolis), the *Oberwolfach Institute*, the *Isaac Newton Institute* (Cambridge, UK), the *Institute for Computational and Experimental Research in Mathematics* (ICERM, Providence) and the *Institute for Mathematical Sciences* (IMS, Singapore). I gratefully acknowledge financial support from the above institutes which made these visits possible.

I also want to thank G. Blekherman, F. Bugarin, M. Ghasemi, N. Gravin, R. Laraki, M. Laurent, M. Mevissen, T. Netzer, M. Putinar, S. Robins, P. Rostalski, C. Savorgnan, and T. Phan Thanh, as some of the results presented in this book have been obtained with their collaboration. I would like to thank Cynthia Vinzant who kindly provided the picture on the book cover. Finally, special thanks are also due to my colleague Didier Henrion at LAAS (especially for our collaboration on the software GloptiPoly which he made user friendly with free access), as well as to J.-B. Hiriart-Urruty, J. Renegar, S. Sorin, and M. Todd for their constant and friendly support.

For part of this research, financial support from the (French) ANR (GEOLMI project), the Simone and Cino Del Duca foundation (OPTIGACOM project), and the Gaspar Monge Program for Optimization and Operations Research (from the Fondation Mathématique Jacques Hadamard (FMJH)) is gratefully acknowledged. And last but not least, I want to thank the CNRS institution for providing a very nice and pleasant working environment at the LAAS-CNRS laboratory in Toulouse.

Toulouse, June 2014 *Jean B. Lasserre*

Symbols

$\mathbb{N}$, the set of natural numbers
$\mathbb{Z}$, the set of integers
$\mathbb{Q}$, the set of rational numbers
$\mathbb{R}$, the set of real numbers
$\mathbb{R}_+$, the set of nonnegative real numbers
$\mathbb{C}$, the set of complex numbers

$\leq$, less than or equal to
$\leqq$, inequality "$\leq$" or equality "$=$"
$\mathbf{A}$, matrix in $\mathbb{R}^{m\times n}$
$\mathbf{A}_j$, column j of matrix $\mathbf{A}$
$\mathbf{A} \succeq 0$ ($\succ 0$), $\mathbf{A}$ is positive semidefinite (definite)
x, scalar $x \in \mathbb{R}$
$\mathbf{x}$, vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$
$\boldsymbol{\alpha}$, vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$
$|\boldsymbol{\alpha}| = \sum_{i=1}^{n} \alpha_i$ for $\boldsymbol{\alpha} \in \mathbb{N}^n$
$\mathbb{N}^n_d$, $\subset \mathbb{N}^n$, the set $\{\, \boldsymbol{\alpha} \in \mathbb{N}^n : |\boldsymbol{\alpha}| \leq d \,\}$
$\mathbf{x}^{\boldsymbol{\alpha}}$, monomial $\mathbf{x}^{\boldsymbol{\alpha}} = (x_1^{\alpha_1} \cdots x_n^{\alpha_n})$ $\mathbf{x} \in \mathbb{C}^n$ or $\mathbf{x} \in \mathbb{R}^n$, $\boldsymbol{\alpha} \in \mathbb{N}^n$
$\mathbb{R}[x]$, ring of real univariate polynomials
$\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$, ring of real multivariate polynomials
$(\mathbf{x}^{\boldsymbol{\alpha}})$, $\boldsymbol{\alpha} \in \mathbb{N}^n$, canonical monomial basis of $\mathbb{R}[\mathbf{x}]$
$V_{\mathbb{C}}(I) \subset \mathbb{C}^n$, the algebraic variety associated with an ideal $I \subset \mathbb{R}[\mathbf{x}]$
$\sqrt{I}$, the radical of an ideal $I \subset \mathbb{R}[\mathbf{x}]$
$\sqrt[\mathbb{R}]{I}$, the real radical of an ideal $I \subset \mathbb{R}[\mathbf{x}]$
$I(V_{\mathbb{C}}(I)) \subset \mathbb{C}^n$, the vanishing ideal $\{\, f \in \mathbb{R}[\mathbf{x}] : f(\mathbf{z}) = 0,\ \forall \mathbf{z} \in V_{\mathbb{C}}(I) \,\}$
$V_{\mathbb{R}}(I) \subset \mathbb{R}^n$ (equal to $V_{\mathbb{C}}(I) \cap \mathbb{R}^n$), the real variety associated with an ideal $I \subset \mathbb{R}[\mathbf{x}]$
$I(V_{\mathbb{R}}(I)) \subset \mathbb{R}[\mathbf{x}]$, the real vanishing ideal $\{\, f \in \mathbb{R}[\mathbf{x}] : f(\mathbf{x}) = 0, \forall \mathbf{x} \in V_{\mathbb{R}}(I) \,\}$

$\mathbb{R}[\mathbf{x}]_t \subset \mathbb{R}[\mathbf{x}]$, vector space of real multivariate polynomials of degree at most t
$\sum[\mathbf{x}]_t \subset \mathbb{R}[\mathbf{x}]_{2t}$, the convex cone of SOS polynomials of degree at most $2t$
$\mathbb{R}[\mathbf{x}]^*$, vector space of linear forms on $\mathbb{R}[\mathbf{x}]$
$\mathbb{R}[\mathbf{x}]_t^*$, vector space of linear forms on $\mathbb{R}[\mathbf{x}]_t$
$\mathbf{y} = (y_\alpha)$, $\boldsymbol{\alpha} \in \mathbb{N}^n$, real moment sequence indexed in the canonical basis of $\mathbb{R}[\mathbf{x}]$
$\mathbf{M}_d(\mathbf{y})$, moment matrix of order d associated with the sequence $\mathbf{y}$
$\mathbf{M}_d(g\,\mathbf{y})$, localizing matrix of order d associated with the sequence $\mathbf{y}$ and $g \in \mathbb{R}[\mathbf{x}]$
$P(g) \subset \mathbb{R}[\mathbf{x}]$, preordering generated by the polynomials $(g_j) \subset \mathbb{R}[\mathbf{x}]$
$Q(g) \subset \mathbb{R}[\mathbf{x}]$, quadratic module generated by the polynomials $(g_j) \subset \mathbb{R}[\mathbf{x}]$
$\mathrm{co}\,\mathbf{X}$, convex hull of $\mathbf{X} \subset \mathbb{R}^n$
$B(\mathbf{X})$, space of bounded measurable functions on $\mathbf{X}$
$C(\mathbf{X})$, space of bounded continuous functions on $\mathbf{X}$
$\mathcal{M}(\mathbf{X})$, vector space of finite signed Borel measures on $\mathbf{X} \subset \mathbb{R}^n$
$\mathcal{M}(\mathbf{X})_+ \subset \mathcal{M}(\mathbf{X})$, space of finite (nonnegative) Borel measures on $\mathbf{X} \subset \mathbb{R}^n$
$\mathcal{P}(\mathbf{X}) \subset \mathcal{M}(\mathbf{X})_+$, space of Borel probability measures on $\mathbf{X} \subset \mathbb{R}^n$
$L_1(\mathbf{X}, \mu)$, Banach space of functions on $\mathbf{X} \subset \mathbb{R}^n$ such that $\int_{\mathbf{X}} |f| d\mu < \infty$
$L_\infty(\mathbf{X}, \mu)$, Banach space of measurable functions on $\mathbf{X} \subset \mathbb{R}^n$ such that $\|f\|_\infty := \text{ess sup } |f| < \infty$
$\sigma(\mathcal{X}, \mathcal{Y})$, weak topology on $\mathcal{X}$ for a dual pair $(\mathcal{X}, \mathcal{Y})$ of vector spaces
$\mu_n \Rightarrow \mu$, weak convergence for a sequence $(\mu_n)_n \subset \mathcal{M}(\mathbf{X})_+$
$\nu \ll \mu$, ν is absolutely continuous with respect to to μ (for measures)
$\uparrow$, monotone convergence for nondecreasing sequences
$\downarrow$, monotone convergence for nonincreasing sequences

SOS, sum of squares
LP, linear programming (or linear program)
SDP, semidefinite programming (or semidefinite program)
GMP, generalized moment problem (or GPM, generalized problem of moments)
SDr, semidefinite representation (or semidefinite representable)
KKT, Karush–Kuhn–Tucker
CQ, constraint qualification
LMI, linear matrix inequality
b.s.a., basic semi-algebraic
b.s.a.l., basic semi-algebraic lifting
l.s.c., lower semi-continuous
u.s.c., upper semi-continuous

Contents

1

Introduction and message of the book

1.1 Why polynomial optimization?

Consider the global optimization problem:

$$\mathbf{P}: \qquad f^* := \inf_{\mathbf{x}} \{ f(\mathbf{x}) \; : \; \mathbf{x} \in \mathbf{K} \} \tag{1.1}$$

for some feasible set

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^n \; : \; g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m \}, \tag{1.2}$$

where $f, g_j : \mathbb{R}^n \to \mathbb{R}$ are some continuous functions.

If one is only interested in finding a *local* (as opposed to *global*) minimum then **P** is a Nonlinear Programming (NLP) problem for which several methods and associated algorithms are already available.

But in this book we insist on the fact that **P** is a *global* optimization problem, that is, f^* is the *global* minimum of f on **K**. In full generality problem (1.1) is very difficult and there is no general purpose method, even to approximate f^*.

However, and this is one of the messages of this book, if one now restricts oneself to *Polynomial Optimization*, that is, optimization problems **P** in (1.1) with the restriction that:

f and $g_j : \mathbb{R}^n \to \mathbb{R}$ are all polynomials, $j = 1, \ldots, m$,

then one may approximate f^* as closely as desired, and sometimes solve **P** exactly. (In fact one may even consider *Semi-Algebraic Optimization*, that is, when f and g_j are semi-algebraic functions.) That this is possible is due to the conjunction of two factors.

- On the one hand, *Linear Programming* (LP) and *Semidefinite Programming* (SDP) have become major tools of convex optimization and today's powerful LP and SDP software packages can solve highly nontrivial problems of relatively large size (and even linear programs of extremely large size).
- On the other hand, remarkable and powerful representation theorems (or positivity certificates) for polynomials that are positive on sets like $\mathbf{K}$ in (1.2) were produced in the 1990s by real algebraic geometers and, importantly, the resulting conditions can be checked by solving appropriate semidefinite programs (and linear programs for some representations)!

And indeed, in addition to the usual tools from *Analysis*, *Convex Analysis* and *Linear Algebra* already used in optimization, in Polynomial Optimization *Algebra* may also enter the game. In fact one may find it rather surprising that algebraic aspects of optimization problems defined by polynomials have not been taken into account in a systematic manner earlier. After all, the class of linear/quadratic optimization problems is an important subclass of Polynomial Optimization! But it looks as if we were so familiar with linear and quadratic functions that we forgot that they are polynomials! (It is worth noticing that in the 1960s, Gomory had already introduced some algebraic techniques for attacking (pure) linear integer programs. However, the algebraic techniques described in the present book are different as they come from *Real Algebraic Geometry* rather than pure algebra.)

Even though Polynomial Optimization is a restricted class of optimization problems, it still encompasses a lot of important optimization problems. In particular, it includes the following.

- Continuous convex and nonconvex optimization problems with linear and/or quadratic costs and constraints, for example

$$\inf_{\mathbf{x}} \{ \mathbf{x}^T\mathbf{A}_0\mathbf{x} + \mathbf{b}_0^T\mathbf{x} : \mathbf{x}^T\mathbf{A}_j\mathbf{x} + \mathbf{b}_j^T\mathbf{x} - c_j \geq 0, \quad j = 1, \ldots, m \},$$

  for some scalars c_j, $j = 1, \ldots, m$, and some real symmetric matrices $\mathbf{A}_j \in \mathbb{R}^{n\times n}$ and vectors $\mathbf{b}_j \in \mathbb{R}^n$, $j = 0, \ldots, m$.
- 0/1 optimization problems, modeling a Boolean variable $x_i \in \{0, 1\}$ via the quadratic polynomial constraint $x_i^2 - x_i = 0$. For instance, the celebrated MAXCUT problem is the polynomial optimization problem