

PRIMAL-DUAL INTERIOR-POINT METHODS

PRIMAL-DUAL INTERIOR-POINT METHODS

Stephen J. Wright

Argonne National Laboratory
Argonne, Illinois

siam

Society for Industrial and Applied Mathematics
Philadelphia

Copyright © 1997 by the Society for Industrial and Applied Mathematics.

10 9 8 7 6 5 4

All rights reserved. Printed in the United States of America. No part of this book may be reproduced, stored, or transmitted in any manner without the written permission of the publisher. For information, write to the Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, PA 19104-2688.

Library of Congress Cataloging-in-Publication Data

Wright, Stephen J., 1960-

Primal-dual interior-point methods / Stephen J. Wright.

p. cm.

Includes bibliographical references and index.

1. Linear programming. 2. Mathematical optimization. I. Title.

T57.74.W75 1997

519.7'2 -- DC20

96-42071

siam is a registered trademark.

ISBN-13: 978-0-898713-82-4 (pbk.)

ISBN-10: 0-89871-382-X (pbk.)

For Nick, Angela, and Charlotte

Contents

Preface	xiii
Notation	xvii
1 Introduction	1
Linear Programming	2
Primal-Dual Methods	4
The Central Path	7
A Primal-Dual Framework	8
Path-Following Methods	9
Potential-Reduction Methods	10
Infeasible Starting Points	11
Superlinear Convergence	12
Extensions	13
Mehrotra's Predictor-Corrector Algorithm	14
Linear Algebra Issues	16
Karmarkar's Algorithm	17
2 Background: Linear Programming and Interior-Point Methods	21
Standard Form	22
Optimality Conditions, Duality, and Solution Sets	23
More on Duality	24
The $\mathcal{B} \cup \mathcal{N}$ Partition and Strict Complementarity	27
A Strictly Interior Point	29
Rank of the Matrix A	31
Bases and Vertices	32

Farkas's Lemma and a Proof of the Goldman–Tucker Result	34
The Central Path	36
Background: Primal Methods	40
Primal-Dual Methods: Development of the Fundamental Ideas . .	43
Notes and References	45
3 Complexity Theory	49
Polynomial Versus Exponential, Worst Case Versus Average Case .	50
Storing the Problem Data: Dimension and Size	51
The Turing Machine and Rational Arithmetic	52
Primal-Dual Methods and Rational Arithmetic	52
Linear Programming and Rational Numbers	54
Moving to a Solution from an Interior Point	55
Complexity of Simplex, Ellipsoid, and Interior-Point Methods . . .	57
Polynomial and Strongly Polynomial Algorithms	58
Beyond the Turing Machine Model	59
More on the Real-Number Model and Algebraic Complexity	60
A General Complexity Theorem for Path-Following Methods	61
Notes and References	62
4 Potential-Reduction Methods	65
A Primal-Dual Potential-Reduction Algorithm	67
Reducing Φ_ρ Forces Convergence	68
A Quadratic Estimate of Φ_ρ Along a Feasible Direction	70
Bounding the Coefficients in the Quadratic Approximation	72
An Estimate of the Reduction in Φ_ρ and Polynomial Complexity .	75
What About Centrality?	78
Choosing ρ and α	79
Notes and References	80
5 Path-Following Algorithms	83
The Short-Step Path-Following Algorithm	86
Technical Results	88
The Predictor-Corrector Method	91
A Long-Step Path-Following Algorithm	96
Limit Points of the Iteration Sequence	100
Proof of Lemma 5.3	103
Notes and References	104

6	Infeasible-Interior-Point Algorithms	107
	The Algorithm	109
	Convergence of Algorithm IPF	111
	Technical Results I: Bounds on $\nu_k \ (x^k, s^k)\ $	113
	Technical Results II: Bounds on $(D^k)^{-1} \Delta x^k$ and $D^k \Delta s^k$	115
	Technical Results III: A Uniform Lower Bound on α_k	118
	Proofs of Theorems 6.1 and 6.2	120
	Limit Points of the Iteration Sequence	121
7	Superlinear Convergence and Finite Termination	127
	Affine-Scaling Steps	128
	An Estimate of $(\Delta x, \Delta s)$: The Feasible Case	129
	An Estimate of $(\Delta x, \Delta s)$: The Infeasible Case	133
	Algorithm PC Is Superlinear	135
	Nearly Quadratic Methods	136
	Convergence of Algorithm LPF+	139
	Convergence of the Iteration Sequence	144
	$\epsilon(A, b, c)$ and Finite Termination	145
	A Finite Termination Strategy	146
	Recovering an Optimal Basis	150
	More on $\epsilon(A, b, c)$	152
	Notes and References	154
8	Extensions	157
	The Monotone LCP	157
	Mixed and Horizontal LCP	160
	Strict Complementarity and LCP	162
	Convex QP	163
	Convex Programming	164
	Monotone Nonlinear Complementarity and Variational Inequalities	167
	Semidefinite Programming	168
	Proof of Theorem 8.4	171
	Notes and References	173
9	Detecting Infeasibility	177
	Self-Duality	178
	The Simplified HSD Form	179
	The HSD Form	183
	Identifying a Solution-Free Region	185

Implementations of the HSD Formulations	188
Notes and References	190
10 Practical Aspects of Primal-Dual Algorithms	193
Motivation for Mehrotra's Algorithm	194
The Algorithm	195
Superquadratic Convergence	198
Second-Order Trajectory-Following Methods	199
Higher-Order Methods	202
Further Enhancements	204
Notes and References	207
11 Implementations	209
Three Forms of the Step Equation	210
The Cholesky Factorization	211
Sparse Cholesky Factorization: Minimum-Degree Orderings	212
Other Orderings	215
Small Pivots in the Cholesky Factorization	216
Dense Columns in A	219
The Augmented System Formulation	220
Factoring Symmetric Indefinite Matrices	221
Starting Points	224
Termination	225
Alternative Formulations for the Linear Program	226
Free Variables	228
Presolving	230
Primal-Dual Codes	232
Notes and References	232
A Basic Concepts and Results	239
Order Notation	239
Convex Sets and Functions	240
KKT Conditions	240
Convexity and Global Solutions	242
Self-Duality and a Proof of Theorem 9.1	243
The Natural log Function and a Proof of Lemma 4.1	245
Singular Value Decomposition, Matrix Rank, and QR Factorization	246
Hoffman's Lemma	248
The Stewart-Todd Result and a Proof of Lemma 7.2	249

The Sherman–Morrison–Woodbury Formula	250
Asymptotic Convergence	251
Storing Sparse Matrices	253
The Turing Machine	254
B Software Packages	257
BPMPD	258
CPLEX/Barrier	259
HOPDM	259
LIPSOL	260
LOQO	261
Newton Barrier XPRESS-MP	261
OSL/EKKBSLV	262
PCx	263
Bibliography	265
Index	281

Preface

Linear programming has been the dominant paradigm in optimization since Dantzig's development of the simplex method in the 1940s. In 1984, the publication of a paper by Karmarkar started a wave of research into a new class of methods known as interior-point methods, and in the decade since then, *primal-dual algorithms* have emerged as the most important and useful algorithms from this class.

On the theoretical side, the properties of primal-dual methods for linear programming have been quite well understood by researchers since approximately 1994. On the computational side, most interior-point software written since 1990 has been based on a single primal-dual algorithm: Mehrotra's predictor-corrector algorithm. The interesting results are, however, widely scattered in the literature, and most of the relevant papers and reports assume a significant amount of background knowledge on the part of the reader. In this book, we use a simple, unified framework to describe the major results, and we provide a straightforward, self-contained account of the underlying theory.

Primal-dual methods have excellent theoretical properties, good practical performance, and pleasing relationships to earlier fundamental ideas in mathematical programming. They can be extended to wider classes of problems, including convex quadratic programming, monotone linear complementarity, and semidefinite programming, without losing their attributes of simplicity and good practical behavior. Extensions to more difficult classes of problems (even to nonlinear programming) are under way, but the issues become more complicated with the loss of linearity and convexity. These extensions are keeping many researchers busy.

Chapter 1 presents a broad overview of primal-dual methods and should be read first by all nonspecialist readers. In the next two chapters, we fill in some of the background for primal-dual methods with results from the theory of linear programming (Chapter 2) and complexity theory (Chapter 3).

Chapter 2 also contains some general historical background on interior-point methods. Descriptions and full analyses of the most interesting algorithms appear in Chapters 4, 5, and 6. An algorithm based on the primal-dual potential function is the subject of Chapter 4. Chapter 5 describes three algorithms of the path-following genre, while Chapter 6 discusses an infeasible-interior-point method that generalizes the long-step path-following method from the preceding chapter. Issues of fast local convergence and finite termination at an optimal solution are discussed in Chapter 7.

Extensions of primal-dual methods beyond linear programming are outlined in Chapter 8. We discuss monotone linear and nonlinear complementarity problems, convex programming, and semidefinite programming. Techniques for reliable detection of infeasible linear programs, including use of the homogeneous self-dual formulation, are the subject of Chapter 9.

The last two chapters turn to topics of more immediate practical interest. Most current software is based on Mehrotra's predictor-corrector algorithm, which enhances the algorithms of Chapters 4, 5, and 6 with a higher-order search direction and some effective heuristics. These enhancements are motivated and described in Chapter 10. In Chapter 11, we examine the computational issues involved in implementing primal-dual methods. The agenda here is dominated by large sparse systems of linear equations, which must be solved to obtain search directions at each iteration. We also discuss the particular features of some primal-dual codes that have appeared recently in the public domain.

Although this book contains a good deal of analysis, it has a somewhat practical bias. We tend to focus on algorithms that are closely related to practical methods. Primal-dual affine-scaling algorithms are omitted for this reason, despite their fascinating theoretical properties. Within each chapter, the intuitive/descriptive sections of the text are well separated from the more technical sections; the latter can be skipped safely and left to a rainy day.

This book is intended to fulfill a number of needs. It can be used as a text for a course on interior-point methods for linear programming. Selected chapters can be used for a short lecture series within a wider course on optimization or linear programming. The prerequisites are a basic understanding of real analysis and linear algebra and a familiarity with the elements of duality theory for linear programming.

Optimization practitioners and researchers who are not interior-point specialists can use the book as an introduction to some of the major ideas and algorithms. Experts can use it as a reference for the main theoretical results and analytical techniques and also as a guide to aspects of the area

with which they are not so familiar.

The field of interior-point methods is an active and competitive one, and the origin of some of the main ideas is sometimes a topic of contention. We have not tried to taxonomize the literature or to trace the development of the various ideas, since these issues are of limited interest to students and nonspecialists. Still, we have tried to point out the key publications, including papers with comprehensive citation lists. No doubt there are some omissions, for which we apologize.

Two sites on the World Wide Web will interest readers of this book. *Interior-Point Methods Online* collects new technical reports in the area, together with other relevant information. Its URL is

<http://www.mcs.anl.gov/home/otc/InteriorPoint/>

The other site is the home page for this book, which contains up-to-date information on software packages, a list of misprints, a feedback column, and so on. The URL is

<http://www.siam.org/books/swright/>

Other Web sites are mentioned in the text as well. The Web is proving to be a valuable adjunct to research activity in this and other fast-moving areas.

I thank Mihai Anitescu, Uri Ascher, Larry Biegler, Joe Czyzyk, Sharon Filipowski, Jean-Pierre Haerberly, Irv Lustig, David Mayne, Michael Overton, Florian Potra, Barry Smith, Yin Zhang, and the anonymous referees of the February 1996 draft for their interest, advice, and extremely valuable comments, which filled numerous gaps in my knowledge. Gail Pieper's assiduous proofreading turned up many infelicities of grammar and logic. As always, her help was invaluable. I am also grateful to Paul Plassmann and Jorge Nocedal for their encouragement during the early stages of this project. Finally, I thank Susan Ciabrano and the staff at SIAM, who made the publication process a sheer delight.

STEPHEN J. WRIGHT
ARGONNE, ILLINOIS

Notation

We summarize the notation and terminology of the book here to avoid confusion and save page ruffling.

\mathbb{R}^n	The space of real n -dimensional vectors.
\mathbb{R}_+^n	The nonnegative orthant $\{x \in \mathbb{R}^n \mid x \geq 0\}$.
$\mathbb{R}^{m \times n}$	The set of real matrices with dimensions $m \times n$.
$\ \cdot\ , \ \cdot\ _2$	Euclidean norm. For $u \in \mathbb{R}^n$, $\ u\ = (\sum_{i=1}^n u_i^2)^{1/2}$. For a matrix M , $\ M\ = \max_{\ u\ =1} \ Mu\ $.
$\ \cdot\ _1$	The 1-norm. For $u \in \mathbb{R}^n$, $\ u\ _1 = \sum_{i=1}^n u_i $.
$\ \cdot\ _\infty$	The ∞ -norm. For $u \in \mathbb{R}^n$, $\ u\ _\infty = \max_{i=1,2,\dots,n} u_i $. (Note that $\ u\ _\infty \leq \ u\ _2 \leq \ u\ _1$ for any vector u .)
e	$(1, 1, \dots, 1)^T$
$\log(\cdot)$	The <i>natural</i> logarithm: \log_e
x	\mathbb{R}^n -vector of primal variables.
λ	\mathbb{R}^m -vector of dual variables, that is, Lagrange multipliers for the equality constraints $Ax = b$.
s	\mathbb{R}^n -vector of dual slacks, that is, Lagrange multipliers for the bound constraints $x \geq 0$.
A	$\mathbb{R}^{m \times n}$ -coefficient matrix for the linear program.
A_i	i th column of A .
problem data	The triple (A, b, c) , which completely defines the linear programming problem.
L	Number of bits required to store the problem data when all components of (A, b, c) are integers or rational numbers.
\mathcal{F}	Primal-dual feasible set:

	$\{(x, \lambda, s) \mid Ax = b, A^T \lambda + s = c, (x, s) \geq 0\}$.
\mathcal{F}^o	Primal-dual strictly feasible set: $\{(x, \lambda, s) \mid Ax = b, A^T \lambda + s = c, (x, s) > 0\}$.
duality gap	If (x, λ, s) is a feasible point, the duality gap is $c^T x - b^T \lambda$, which coincides with $x^T s$.
μ	Duality measure, defined as $x^T s/n$.
r_b	Primal residual, defined as $Ax - b$.
r_c	Dual residual, defined as $A^T \lambda + s - c$.
Z^*	Optimal objective at a solution of the linear program. If (x^*, λ^*, s^*) is a primal-dual solution, then $c^T x^* = Z^* = b^T \lambda^*$. If (x, λ, s) is any feasible point, then $c^T x \geq Z^* \geq b^T \lambda$.
Ω_P	Set of primal solutions x^* .
Ω_D	Set of dual solutions (λ^*, s^*) .
Ω	Set of primal-dual solutions (x^*, λ^*, s^*) ; $\Omega = \Omega_P \times \Omega_D$.
\mathcal{B}	Defined by $i \in \mathcal{B} \subset \{1, 2, \dots, n\}$ if $x_i^* > 0$ for some primal solution x^* .
\mathcal{N}	Defined by $i \in \mathcal{N} \subset \{1, 2, \dots, n\}$ if $s_i^* > 0$ for some dual slack solution s^* . $(\mathcal{B} \cup \mathcal{N} = \{1, 2, \dots, n\}$ and $\mathcal{B} \cap \mathcal{N} = \emptyset)$
$ \mathcal{B} , \mathcal{N} $	The number of elements in \mathcal{B} and \mathcal{N} , respectively.
$A_{\mathcal{B}}, A_{\mathcal{N}}$	Partition of A into columns corresponding to the index sets \mathcal{B} and \mathcal{N} .
$x_{\mathcal{B}}, x_{\mathcal{N}}$	Subvectors of x that correspond to \mathcal{B} and \mathcal{N} , respectively. (Similarly for $s_{\mathcal{B}}, s_{\mathcal{N}}, x_{\mathcal{B}}^*$, etc.)
strictly complementary solution	A solution (x^*, λ^*, s^*) such that $x_{\mathcal{B}}^* > 0$ and $s_{\mathcal{N}}^* > 0$.
$\epsilon(A, b, c)$	$\min(\min_{i \in \mathcal{B}} \sup_{x^* \in \Omega_P} x_i^*, \min_{i \in \mathcal{N}} \sup_{(\lambda^*, s^*) \in \Omega_D} s_i^*)$.
\mathcal{M} , basis, B	A subset of m indices from $\{1, 2, \dots, n\}$ such that the submatrix $B = [A_{\cdot j}]_{j \in \mathcal{M}}$ is nonsingular and $B^{-1}b \geq 0$.
k (superscript or subscript)	Iteration counter for a sequence, $k = 0, 1, 2, \dots$
i, j (subscript)	Usual notation for indices of particular components of a vector or matrix.

(x^k, λ^k, s^k) or (x, λ, s)	Generic iterate of a primal-dual interior-point method. (Iteration counter k often omitted for clarity.)
$(\Delta x, \Delta \lambda, \Delta s)$	Generic primal-dual step, satisfying (1.12) or (1.20), depending on the context.
(x^+, λ^+, s^+)	Generic “next point”—the new iterate obtained by taking a step from the current point (x, λ, s) .
$F(\cdot)$	A function of (x, λ, s) consisting of the equalities from the Karush–Kuhn–Tucker conditions.
$J(\cdot)$	Jacobian of F .
α, α_k	Step length parameter.
unit step	A step with $\alpha = 1$.
$(x(\alpha), \lambda(\alpha), s(\alpha))$	$(x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s)$.
$(x^k(\alpha), \lambda^k(\alpha), s^k(\alpha))$	$(x^k, \lambda^k, s^k) + \alpha(\Delta x^k, \Delta \lambda^k, \Delta s^k)$.
$\mu(\alpha)$	$x(\alpha)^T s(\alpha) / n$.
$\mu_k(\alpha)$	$x^k(\alpha)^T s^k(\alpha) / n$.
X, X^k	$n \times n$ diagonal matrix constructed from the vector x or x^k : $X = \text{diag}(x_1, x_2, \dots, x_n)$.
S, S^k	$n \times n$ diagonal matrix constructed from the vector s or s^k : $S = \text{diag}(s_1, s_2, \dots, s_n)$.
D	$X^{1/2} S^{-1/2}$.
$\Delta S, \Delta X$	Diagonal matrices constructed from Δs and Δx , respectively.
v	Vector in \mathbb{R}^n with components $(x_i s_i)^{1/2}$.
v_{\min}	$\min_i v_i$.
V	$\text{diag}(v_1, v_2, \dots, v_n)$.
r	$-v + (n/\rho)\mu V^{-1}e$, for some $\rho > n$.
affine-scaling direction	The pure Newton direction obtained by setting $\sigma = 0$ in (1.12 or (1.20).
centering direction	The direction obtained by setting $\sigma = 1$ in (1.12) or (1.20).

σ	Centering parameter for step computation (1.12), (1.20); restricted to the range $[0, 1]$.
$\text{Range}(B)$	Range space of the matrix B .
$\text{Null}(B^T)$	Null space of the matrix B^T . (The fundamental theorem of algebra states that $\text{Null}(B^T) \oplus \text{Range}(B) = \mathbb{R}^n$, where B has n rows.)
$P_{\mathcal{U}}$	Projection into the subspace \mathcal{U} .
$\text{dist}(x, \mathcal{T})$	Distance from x to the set \mathcal{T} , defined as $\inf_{y \in \mathcal{T}} \ x - y\ $.
vertex	If \mathcal{T} is a polyhedron in \mathbb{R}^n , a point $x \in \mathcal{T}$ is a vertex if it is the unique minimizer in \mathcal{T} of some linear function.
$\Phi_{\rho}(x, s) = \rho \log x^T s - \sum_{i=1}^n \log x_i s_i$	for some $\rho \geq n$ The Tanabe–Todd–Ye potential function.
$\mathcal{N}_2(\theta)$	The 2-norm central path neighborhood, defined by $\{(x, \lambda, s) \in \mathcal{F}^o \mid \ XSe - \mu e\ \leq \theta \mu\}$ for given $\theta \in (0, 1)$.
$\mathcal{N}_{-\infty}(\gamma)$	The ∞ -norm central path neighborhood, defined by $\{(x, \lambda, s) \in \mathcal{F}^o \mid x_i s_i \geq \gamma \mu, \forall_i\}$ for given $\gamma \in (0, 1)$.
$\mathcal{N}_{-\infty}(\gamma, \beta)$	The infeasible ∞ -norm neighborhood, defined by $\{(x, \lambda, s) \mid \ (r_b, r_c)\ \leq [\ (r_b^0, r_c^0)\ /\mu_0] \beta \mu, (x, s) > 0,$ $x_i s_i \geq \gamma \mu, \forall_i\}$ for given $\gamma \in (0, 1)$, $\beta \geq 1$, and initial point (x^0, λ^0, s^0) .
LCP	Monotone linear complementarity problem.
mLCP	Mixed monotone linear complementarity problem.
hLCP	Horizontal monotone linear complementarity problem.
symmetric, skew-symmetric	M is symmetric if $M^T = M$, skew-symmetric if $M^T = -M$.
\mathcal{S}^n	The set of $n \times n$ symmetric matrices.
\mathcal{D}_+	The set of diagonal matrices with strictly positive diagonal elements.
u	Unit roundoff error (about 10^{-16} for 8-byte floating-point arithmetic).