

# Mathematical Programming

## Theory and Algorithms

M. MINOUX

*Professor, Ecole Nationale Supérieure  
de Techniques Avancées and  
Université Paris IX-Dauphine  
France*

*Translated by*

STEVEN VAJDA

*University of Sussex, Falmer,  
Brighton, Sussex, UK*

*A Wiley-Interscience Publication*

JOHN WILEY AND SONS

Chichester · New York · Brisbane · Toronto · Singapore

# Contents

<i>Foreword</i> .....	xv
<i>Preface</i> .....	xvii
<i>Acknowledgements and Remerciements</i> .....	xxv
<i>Notation</i> .....	xxv
<b>1 Fundamental Concepts</b> .....	<b>1</b>
1.1 Mathematical Programming. Definitions .....	1
1.2 Elements of Topology .....	3
1.2.1 Convergence of sequences in $\mathbb{R}^n$ or in $\mathbb{R}$ .....	4
1.2.2 Open sets. Closed sets .....	4
1.2.3 Compact sets. Theorem of Weierstrass .....	6
1.2.4 Remark concerning the notation Min and Max .....	8
1.3 Elements of Convex Analysis .....	8
1.3.1 Convex sets .....	8
1.3.2 Convex functions .....	10
1.3.3 Convex programmes .....	11
1.3.4 Extended convex functions .....	12
1.3.5 Subgradient. Subdifferential .....	14
1.3.6 Subgradients and directional derivatives .....	15
1.4 Study of Convergence. Global and Asymptotic Convergence ...	17
1.4.1 The concept of an algorithm in mathematical programming .....	17
1.4.2 A general model of algorithms: point-to-set maps .....	17
1.4.3 The concept of global convergence .....	18
1.4.4 Closed point-to-set maps .....	18
1.4.5 The global convergence theorem .....	20
1.4.6 Asymptotic convergence. Rate of convergence .....	22
References .....	25
<b>2 Linear Programming</b> .....	<b>27</b>
2.1 Definitions and Fundamental Results .....	27
2.1.1 Standard form of a linear programme .....	27
2.1.2 Solutions of a linear programme and convex polyhedra	28

2.1.3	Bases, feasible bases, basic solutions .....	29
2.1.4	Algebraic characterization of vertices. ....	30
2.1.5	Theorem 2.2 (Optimality at a vertex). ....	33
2.1.6	Characterization of optimal bases and basic solutions. ....	34
2.2	Solution of Linear Programmes .....	36
2.2.1	The primal 'Simplex' method (revised form). ....	36
2.2.2	Theorem 2.4 (finite convergence). ....	37
2.2.3	Problems arising from degeneracy .....	37
2.2.4	Algorithmic complexity and practical efficiency of the Simplex method. ....	38
2.2.5	Elementary matrices and the product form of the inverse ....	39
2.2.6	The initial basis .....	39
2.2.7	Canonical form and the Simplex tableau .....	40
2.2.8	An example. ....	43
2.3	The Concept of Duality .....	46
2.3.1	The dual of a linear programme in standard form ....	46
2.3.2	Definition of the dual in the general case .....	47
2.3.3	The duality theorem .....	48
2.3.4	Dual variables and marginal costs .....	50
2.4	Dual and Primal-dual Algorithms. ....	50
2.4.1	The dual method .....	50
2.4.2	The primal-dual algorithm. ....	53
	References .....	55
<b>3</b>	<b>One-dimensional Optimization. ....</b>	<b>57</b>
3.1	Methods Using Derivatives .....	57
3.1.1	The Newton-Raphson method .....	57
3.1.2	The secant method .....	59
3.1.3	The dichotomy method with derivatives. ....	60
3.2	Methods Without Derivatives .....	61
3.2.1	Quadratic Interpolation. ....	61
3.2.2	Unimodal functions .....	63
3.2.3	The dichotomy method without derivatives. ....	65
3.2.4	The Fibonacci search method. ....	66
3.2.5	The golden section search method .....	68
3.2.6	'Economical' one-dimensional search. ....	69
3.3	One-dimensional Optimization Algorithms and Closed Point-to-set Maps. ....	73
3.3.1	Exact one-dimensional optimization .....	73
3.3.2	Approximate one-dimensional optimization. ....	75
3.3.3	One-dimensional constrained optimization. ....	76
	References .....	79
<b>4</b>	<b>Nonlinear, Unconstrained Optimization .....</b>	<b>80</b>
4.1	Introduction. Optimality Conditions .....	80

4.1.1	Necessary conditions for local optimality.....	81
4.1.2	Sufficient conditions for local optimality.....	82
4.1.3	The case of convex functions: necessary and sufficient condition for global optimality.....	83
4.1.4	Case of arbitrary functions. Difficulty of the general problem.....	83
4.2	Numerical Methods for the Optimization of Differentiable Functions.....	84
4.2.1	Gradient methods. Gradient with predetermined steps	84
4.2.2	Steepest descent method.....	85
4.2.3	'Accelerated' steepest descent methods.....	87
4.2.4	'Second order' methods.....	89
4.2.5	Conjugate direction methods: the general principle...	89
4.2.6	The conjugate gradient method for quadratic functions	91
4.2.7	The case of arbitrary functions.....	93
4.2.8	Newton's method.....	94
4.2.9	Quasi-Newton methods. The general principle.....	96
4.2.10	Algorithm of Davidon-Fletcher-Powell (DFP).....	99
4.2.11	The algorithm of Broyden, Fletcher, Goldfarb, Shanno (BFGS).....	102
4.2.12	Comparison of the various methods. Results concerning the rate of convergence.....	104
4.3	Optimization of not Everywhere Differentiable (Non-smooth) Convex Functions.....	107
4.3.1	Example: duality in discrete programming.....	107
4.3.2	Finding a subgradient. Example.....	108
4.3.3	Use of generalized linear programming.....	109
4.3.4	A class of subgradient algorithms.....	109
4.3.5	Convergence of method 1 (method with fixed step sizes)	113
4.3.6	Convergence of method 2 (convergent series method)	117
4.3.7	Convergence of method 3 (relaxation method).....	120
4.3.8	The space dilatation methods of Shor.....	123
4.3.9	Projection onto the subspace generated by the successively encountered gradients.....	126
4.3.10	Space dilatation in the direction of the difference of two successive subgradients.....	127
4.3.11	Space dilatation in the direction of a subgradient at the current point.....	128
4.3.12	Cut-off method with space dilatation.....	129
4.3.13	Recent developments in optimization of not everywhere differentiable functions.....	133
4.4	Optimization Methods Without Derivatives.....	134
4.4.1	Methods of cyclic relaxation.....	134
4.4.2	The algorithm of Powell (1964).....	135
	References.....	138

**5 Nonlinear Optimization With Constraints. I Direct (or Primal)**

<b>Methods.</b>	<b>142</b>
5.1 Necessary Optimality Conditions	142
5.1.1 Admissible directions and constraint qualification	142
5.1.2 The necessary conditions of Kuhn and Tucker	146
5.1.3 Geometric interpretation of the Kuhn and Tucker conditions	148
5.1.4 Extension to problems with equality as well as inequality constraints. Conditions of Lagrange	148
5.2 Sufficient Optimality Conditions: 'Saddle-points' and Lagrange Functions	150
5.2.1 Theorem 5.2 (Characteristic property of saddle-points)	151
5.2.2 Theorem 5.3 (Sufficiency of the saddle-point condition)	151
5.2.3 Example (a non-convex optimization problem)	152
5.2.4 A necessary and sufficient condition for the existence of a saddle-point: perturbation functions	152
5.2.5 Existence of a saddle-point in the convex case—Lagrange multipliers as subgradients of the perturbation function	158
5.2.6 Connection with the Kuhn and Tucker conditions in the convex differentiable case	159
5.2.7 Sufficient conditions for local optimality in the non-convex case	160
5.2.8 Lagrange multipliers and post-optimal analysis	160
5.3 Constrained Optimization: Primal Methods	161
5.3.1 Method of changing the variables	161
5.3.2 Methods of feasible directions	162
5.3.3 The gradient projection method	167
5.3.4 The reduced gradient method	172
5.3.5 The generalized reduced gradient method (GRG)	174
5.3.6 Linearization methods	177
5.3.6.1 The method of Frank and Wolfe (1956)	179
5.3.6.2 The cutting plane method of Kelley (1960)	180
5.3.6.3 The column generation method of Dantzig (1959)	182
5.3.7 Analysis of the convergence. Evaluation of the algorithms	184
5.4 Constrained Optimization by Solving the Kuhn and Tucker Equations	186
5.4.1 Newton's method	186
5.4.2 Extension of Newton's method: method of Wilson. The case of inequality constraints	188
5.4.3 Connection with the feasible directions methods	189
5.4.4 Connection with the dual methods	190
References	191

<b>6 Nonlinear Constrained Optimization. II Dual Methods</b>	<b>194</b>
6.1 Introduction: Penalty Function Methods	194
6.1.1 General principle of penalty function methods	194
6.1.2 Method of exterior penalties	195
6.1.3 Interior penalty methods	198
6.1.4 Approximation to the Kuhn and Tucker multipliers at the optimum	200
6.1.5 Combination of the penalty methods with other methods	201
6.2 Classical Lagrangean Duality	202
6.2.1 Property 6.1 (Weak duality theorem)	203
6.2.2 Property 6.2 (Concavity of the dual function)	203
6.2.3 Property 6.3 (Duality theorem)	204
6.2.4 Example 1	205
6.2.5 Example 2 (Duality in linear programming)	206
6.2.6 Example 3	207
6.2.7 Property 6.4 (Subgradient)	209
6.2.8 Differentiability of the dual function	211
6.2.9 Property 6.5	212
6.2.10 Duality gaps and non-differentiability of the dual function at the optimum	213
6.2.11 Obtaining approximate primal solutions by solving the dual problem	213
6.3 Classical Lagrangean Methods	214
6.3.1 Methods of Uzawa (1958) and of Arrow-Hurwicz (1958)	214
6.3.2 The method of Dantzig (1959)	215
6.3.3 Use of Lagrangean methods in the non-convex case	217
6.4 Generalized Lagrangeans and Saddle-points in Non-convex Programming	218
6.4.1 Augmented Lagrangeans: introduction	218
6.4.2 Lagrangean representations of a constrained optimization problem	220
6.4.3 Perturbational representations of a problem	222
6.4.4 Example 1: The 'ordinary' Lagrangean	227
6.4.5 Example 2: The 'augmented Lagrangean' of Rockafellar	227
6.4.6 Example 3: Multiplier methods	228
6.4.7 Saddle-points in non-convex programming	229
6.5 Comparative study of algorithms. Convergence	231
6.5.1 Geometrical interpretation and comparison of the various methods	231
6.5.2 Convergence analysis	237
6.5.2.1 Penalty methods	237
6.5.2.2 Methods using the 'ordinary' Lagrangean	238
6.5.2.3 Methods using an augmented Lagrangean	240
References	241

<b>7</b>	<b>Integer Programming</b>	<b>245</b>
7.1	Introduction	245
7.2	Tree-Search Methods (Branch and Bound)	248
7.2.1	Reduction to a problem in binary (0-1) variables	249
7.2.2	Enumeration?	250
7.2.3	Definition of the tree: the concept of branching	250
7.2.4	Bounding	252
7.2.5	Practical implementation	252
7.2.6	Example	254
7.2.7	The Tradeoff between cost and quality of the evaluation	258
7.3	Lower Bounds and Approximate Solutions by Lagrangean Relaxation and Solution of a Dual Problem	258
7.3.1	The traveling salesman problem. The directed and non-directed cases	259
7.3.2	Location problems. Clustering problems	264
7.3.3	The Steiner Tree problem	269
7.3.4	Set packing and set partitioning problems	272
7.3.5	Shortest path problems with additional constraint(s) and related combinatorial problems	275
7.3.6	The general problem of the intersection of two families of combinatorial objects and its solution by Lagrangean relaxation	280
7.3.7	The generalized assignment problem	283
7.3.8	Other examples of application of Lagrangean relaxation in combinatorial optimization	285
7.4	Cutting Plane Methods	285
7.4.1	Principle of cutting plane methods	285
7.4.2	The Gomory cuts	287
7.4.3	The dual algorithm of Gomory. Principles and variations	289
7.4.4	The method of decreasing congruences	291
7.4.5	Example	293
7.4.6	Recent developments in the theory of cutting planes	296
7.5	Integer Programmes and Shortest Paths. Representation by Finite Groups	306
7.5.1	Equivalence with a shortest path problem	306
7.5.2	Homomorphic images of a problem	308
7.5.3	An equivalent formulation of the relaxed problem $P(G)$	311
7.5.4	An important special case: the Gomory group	312
7.5.5	Equivalent images of the same problem	314
7.5.6	Lagrangean relaxation and definition of a dual problem	317
7.5.7	Algorithm 1. Solution of the asymptotic problem	318
7.5.8	Algorithm 2 (Bell, 1977)	319
7.5.9	Algorithm 3 (Bell, 1976)	320
7.5.10	Algorithm 4 (Bell and Shapiro, 1977)	321
	References	323

<b>8</b>	<b>Solution of Large-scale Programming Problems. Generalized Linear Programming and Decomposition Techniques</b>	<b>330</b>
8.1	Generalized Linear Programming (Column Generation)	330
8.1.1	Statement of the problem	330
8.1.2	Example 1 (Dantzig-Wolfe decomposition)	331
8.1.3	Example 2 (minimum cost flows in a graph)	332
8.1.4	Solution by generalized linear programming: a column generation method	333
8.1.5	Implementation and storage space	334
8.1.6	Application to the optimization of convex or concave functions by the cutting plane method	335
8.1.7	Other applications of generalized linear programming	337
8.2	Lagrangian Relaxation and Price-decomposition	338
8.2.1	Comments on the structure of large linear programmes	338
8.2.2	Statement of the problem	339
8.2.3	Lagrangian relaxation of coupling constraints. Computation of the dual function by decomposition	340
8.2.4	Solution of the dual problem by a subgradient algorithm	342
8.2.5	Formulation of the dual problem as a linear programme	342
8.2.6	The Dantzig-Wolfe decomposition method (1961)	344
8.2.7	Economic interpretation	346
8.3	Decomposition by Right-hand side Allocation (Decomposition by Resources)	346
8.3.1	Principle of the method	346
8.3.2	An equivalent formulation of the problem	347
8.3.3	Finding a subgradient of the master objective function	349
8.3.4	Solution of the master programme	349
8.3.5	Extension of the method	350
8.4	Decomposition by Partitioning of the Variables (Benders')	351
8.4.1	Statement of the problem	351
8.4.2	Use of the necessary and sufficient conditions of Farkas and Minkowski	352
8.4.3	Equivalence to a large linear programme in variables $y$	353
8.4.4	Solution by generalized linear programming: Benders' method	355
8.4.5	Convergence	357
8.4.6	Connection with the Dantzig-Wolfe decomposition method	358
8.4.7	Other applications of the Benders method	358
8.5	Examples of the Application of Decomposition Methods: Large Scale Network Optimization Problems	360
8.5.1	Feasibility problems for multi-commodity flows	360
8.5.2	Problems of feasible multi-commodity flows at minimum cost	366
8.5.3	Optimal synthesis of networks with non-simultaneous multi-commodity flow requirements	368



References .....	372
<b>9 Dynamic Programming .....</b>	<b>375</b>
9.1 Introduction and Examples .....	375
9.1.1 An example: solution of a 'knapsack' problem by dynamic programming .....	376
9.1.2 To find an optimal solution $x^*$ .....	377
9.1.3 Analysis of the complexity and of the limitations of dynamic programming .....	378
9.1.4 Dynamic programming and shortest (longest) path pro- blems in graphs .....	380
9.2 The Theoretical Foundations of Dynamic Programming .....	381
9.2.1 The optimality theorem. Unconstrained case .....	381
9.2.2 Extension of the optimality theorem to the constrained case .....	383
9.2.3 The concept of state in dynamic programming .....	383
9.2.4 The functional equation of dynamic programming .....	386
9.2.5 The principle of optimality .....	388
9.2.6 Examples of decomposable functions .....	390
9.2.7 Dynamic programming with a finite number of states and the search for a shortest (or longest) path in a graph .....	391
9.2.8 Second dynamic programming algorithm. 'Forward' procedure .....	392
9.3 Techniques for Reducing Computations in Dynamic Programming .....	393
9.3.1 The method of Lagrange multipliers .....	393
9.3.2 Combination of dynamic programming with Branch and Bound methods .....	395
9.3.3 Admissible search method. Algorithm A* .....	399
9.3.4 State space relaxation methods .....	401
9.4 Examples of Dynamic Programming Applications .....	403
9.4.1 Optimization of dynamic systems .....	403
9.4.2 Shortest path in a graph .....	405
9.4.3 Problem of minimum length Hamiltonian circuits .....	406
9.4.4 Problems of matching, of partitioning, and of covering a hypergraph of intervals .....	407
9.4.5 Dynamic programming and stochastic systems: an application to the filtering problem for Markov processes .....	410
References .....	413
<b>10 Optimization in Infinite Dimension and Applications .....</b>	<b>416</b>
10.1 Introduction and Examples .....	416
10.1.1 Example 1: Calculus of variations .....	416
10.1.2 Example 2: Optimal control of systems governed by differential equations .....	418

10.1.3	Systems governed by elliptic partial differential equations	421
10.2	Banach Spaces and Hilbert Spaces	424
10.2.1	Normed vector spaces	425
10.2.2	Banach spaces	425
10.2.3	Dual of a normed vector space	427
10.2.4	The bidual. Reflexive Banach spaces	428
10.2.5	Weak convergence	429
10.2.6	Theorem of weak compactness for reflexive Banach spaces	431
10.2.7	Scalar product. Hilbert spaces	431
10.3	Optimization of Functionals. Existence of a Minimum. Necessary Optimality Conditions	434
10.3.1	Existence of a solution. Theorem of Weierstrass	434
10.3.2	Semi-continuity and extension of the Weierstrass theorem	435
10.3.3	Derivation in the sense of Gateaux. Gradient of a functional	437
10.3.4	Convex G-differentiable functionals. Convex coercive functionals	439
10.3.5	Necessary optimality conditions	442
10.3.6	Application to the calculus of variations. Equation of Euler-Lagrange	443
10.3.7	Application to optimal control. Pontryagin's theorem of the minimum	446
10.3.8	Application to the projection onto a closed convex set in a Hilbert space	449
10.4	Optimization Methods in Infinite Dimension	450
10.4.1	Unconstrained optimization: steepest descent method	450
10.4.2	Application of steepest descent methods to the solution of optimal control problems	454
10.4.3	Constrained optimization: fixed step gradient with projection	454
10.4.4	Constrained optimization: penalty methods	456
10.4.5	Constrained optimization: duality methods	457
10.4.6	Approximation of problems in infinite dimension by problems in finite dimension. Method of Galerkin	459
	References	463
Appendix 1	Separation of Convex Sets. Theorem of Farkas and Minkowski. Theorem of Gordan	465
Appendix 2	Existence of Saddle-points in Convex Mathematical Programming	469
Appendix 3	Solution of Integer Linear Systems	471
Index		482