

INTRODUCTION TO DERIVATIVE-FREE OPTIMIZATION

Andrew R. Conn
Katya Scheinberg

IBM
Yorktown Heights, New York

Luis N. Vicente
Universidade de Coimbra
Coimbra, Portugal



Society for Industrial and Applied Mathematics
Philadelphia



Mathematical Programming Society
Philadelphia

Contents

Preface	xi
1 Introduction	1
1.1 Why derivative-free optimization	1
1.2 Examples of problems where derivatives are unavailable	3
1.3 Limitations of derivative-free optimization	5
1.4 How derivative-free algorithms should work	7
1.5 A short summary of the book	11
I Sampling and modeling	13
2 Sampling and linear models	15
2.1 Positive spanning sets and positive bases	15
2.2 Gradient estimates used in direct search	21
2.3 Linear interpolation and regression models	24
2.4 Error bounds for linear interpolation and regression	26
2.5 Other geometrical concepts	29
2.6 Simplex gradients	32
2.7 Exercises	33
3 Interpolating nonlinear models	35
3.1 Basic concepts in interpolation	36
3.2 Lagrange polynomials	39
3.3 Λ -poisedness and other measures of well poisedness	42
3.4 Condition number as a measure of well poisedness	48
3.5 Exercises	56
4 Regression nonlinear models	57
4.1 Basic concepts in polynomial least-squares regression	58
4.2 Lagrange polynomials in the regression sense	60
4.3 Λ -poisedness in the regression sense	62
4.4 Condition number as a measure of well poisedness	68
4.5 Notes and references	70
4.6 Exercises	72

5	Underdetermined interpolating models	73
5.1	The choice of an underdetermined model	74
5.2	Lagrange polynomials and Λ -poisedness for underdetermined interpolation	76
5.3	Minimum Frobenius norm models	80
5.4	Notes and references	87
5.5	Exercises	87
6	Ensuring well poisedness and suitable derivative-free models	89
6.1	Fully linear and fully quadratic models	90
6.2	Ensuring well poisedness using Lagrange polynomials	93
6.3	Ensuring well poisedness using pivotal algorithms	99
6.4	Practical considerations of geometry improvement algorithms	107
6.5	Ensuring well poisedness for regression and minimum Frobenius norm models	108
6.6	Other notes and references	110
6.7	Exercises	112
II	Frameworks and algorithms	113
7	Directional direct-search methods	115
7.1	The coordinate-search method	115
7.2	A directional direct-search framework	118
7.3	Global convergence in the continuously differentiable case	120
7.4	Global convergence in the nonsmooth case	124
7.5	Simple decrease with integer lattices	127
7.6	The mesh adaptive direct-search method	132
7.7	Imposing sufficient decrease	134
7.8	Other notes and references	135
7.9	Exercises	139
8	Simplicial direct-search methods	141
8.1	The Nelder–Mead simplex method	141
8.2	Properties of the Nelder–Mead simplex method	148
8.3	A globally convergent variant of the Nelder–Mead method	149
8.4	Other notes and references	161
8.5	Exercises	161
9	Line-search methods based on simplex derivatives	163
9.1	A line-search framework	163
9.2	Global convergence for first-order critical points	165
9.3	Analysis for noise	167
9.4	The implicit-filtering algorithm	168
9.5	Other simplex derivatives	169
9.6	Other notes and references	170
9.7	Exercises	171

10	Trust-region methods based on derivative-free models	173
10.1	The trust-region framework basics	173
10.2	Conditions on the trust-region models	179
10.3	Derivative-free trust-region methods (first order)	181
10.4	Global convergence for first-order critical points	185
10.5	Derivative-free trust-region methods (second order)	191
10.6	Global convergence for second-order critical points	194
10.7	Model accuracy in larger concentric balls	200
10.8	Trust-region subproblem	202
10.9	Other notes and references	204
10.10	Exercises	205
11	Trust-region interpolation-based methods	207
11.1	Common features and considerations	207
11.2	The “DFO” approach	208
11.3	Powell’s methods	211
11.4	Wedge methods	215
11.5	Other notes and references	225
III	Review of other topics	227
12	Review of surrogate model management	229
12.1	Surrogate modeling	229
12.2	Rigorous optimization frameworks to handle surrogate models	235
12.3	Exercises	240
13	Review of constrained and other extensions to derivative-free optimization	241
13.1	Directional direct-search methods	242
13.2	Trust-region interpolation-based methods	248
13.3	Derivative-free approaches for global optimization, mixed-integer programming, and other problems	249
Appendix	Software for derivative-free optimization	251
	Bibliography	255
	Index	271