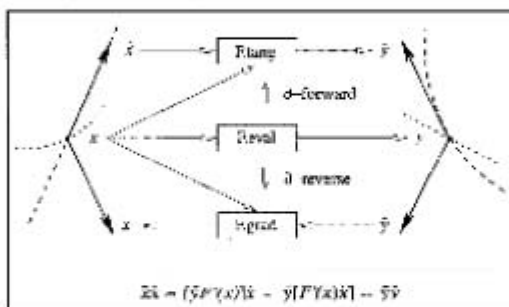


Evaluating Derivatives

Principles and Techniques of Algorithmic Differentiation



Andreas Griewank

Institute of Scientific Computing
Technical Institute of Dresden
Dresden, Germany

siam

Society for Industrial and Applied Mathematics
Philadelphia

Contents

Rules	xiii
Preface	xv
Prologue	xix
Mathematical Symbols	xxiii
1 Introduction	1
I Tangents and Gradients	13
2 A Framework for Evaluating Functions	15
2.1 The Lighthouse Example	16
2.2 Three-Part Evaluation Procedures	18
2.3 Elemental Differentiability	24
2.4 Memory Allocation and Overwrites	26
2.5 A Temporal Complexity Model	29
2.6 Examples and Exercises	36
3 Fundamentals of Forward and Reverse	37
3.1 Forward Propagation of Tangents	38
3.2 Complexity of Tangent Propagation	43
3.3 Reverse Propagation of Gradients	45
3.4 Complexity of Gradient Propagation	56
3.5 Summary and Discussion	60
3.6 Examples and Exercises	65
4 Repeating and Extending Reverse	69
4.1 Adjoining Iterative Assignments	71
4.2 Selective Saves and Restores	76
4.3 Adjoints of Adjoints	79
4.4 Tangents of Adjoints	82
4.5 Complexity of Second-Order Adjoints	84
4.6 Examples and Exercises	87

5	Implementation and Software	91
5.1	Program Transformation and Overloading	93
5.2	Implementing and Using AD	97
5.3	Simple Forward Implementation	98
5.4	Simple Reverse Implementation	103
5.5	Second-Order Adjoints	107
5.6	Propagating Vectors and Higher Derivatives	109
5.7	Source Transformation	111
5.8	Using AD Preprocessors	113
5.9	Examples and Exercises	115
II	Jacobians and Hessians	119
6	Sparse Forward and Reverse	121
6.1	Quantifying Structure and Sparsity	123
6.2	Sparse Forward Derivative Propagation	127
6.3	Sparse Reverse Derivative Propagation	129
6.4	Sparse Second Derivatives, Forward	130
6.5	Sparse Second Derivatives, Reverse	132
6.6	Summarizing Conclusion	134
6.7	Examples and Exercises	135
7	Exploiting Sparsity by Compression	137
7.1	Curtis-Powell-Reid Seeding	140
7.2	Newsam-Ramsdell Seeding	143
7.3	Column Compression	145
7.4	Combined Column and Row Compression	148
7.5	Second Derivatives, Two-Sided Compression	151
7.6	Summary	155
7.7	Examples and Exercises	157
8	Going beyond Forward and Reverse	159
8.1	Jacobians as Schur Complements	161
8.2	Accumulation by Edge-Elimination	166
8.3	Greedy and Other Heuristics	176
8.4	Local Preaccumulations	185
8.5	Hessians and Their Computational Graph	190
8.6	Summary and Outlook	194
8.7	Examples and Exercises	195
9	Observations on Efficiency	199
9.1	Ramification of the Rank-One Example	199
9.2	Remarks on Partial Separability	202
9.3	Advice on Problem Preparation	209
9.4	Examples and Exercises	212

III Advances and Reversals	215
10 Taylor and Tensor Coefficients	217
10.1 Higher Order Derivative Vectors	218
10.2 Taylor Polynomial Propagation	221
10.3 Multivariate Tensor Evaluation	228
10.4 Higher Order Gradients and Jacobians	233
10.5 Special Relations in Dynamical Systems	243
10.6 Summary and Outlook	248
10.7 Examples and Exercises	249
11 Differentiation without Differentiability	251
11.1 Another Look at the Lighthouse Example	253
11.2 Putting the Pieces Together	260
11.3 One-Sided Laurent Expansions	269
11.4 Implicit and Iterative Differentiation	282
11.5 Summary and Outlook	299
11.6 Examples and Exercises	301
12 Serial and Parallel Reversal Schedules	303
12.1 Reversal with Taping or Recalculation	304
12.2 Reversal of Call Trees	306
12.3 Reversals of Evolutions by Checkpointing	319
12.4 Parallel Reversal Schedules	330
12.5 Examples and Exercises	334
Epilogue	337
List of Figures	339
List of Tables	341
Assumptions and Definitions	345
Propositions, Corollaries, and Lemmas	347
Bibliography	349
Index	365