O. Nelles

Nonlinear System Identification

Springer

Contents

1. Introduction 1
1.1 Relevance of Nonlinear System Identification 1
1.1.1 Linear or Nonlinear? 1
1.1.2 Prediction 2
1.1.3 Simulation 3
1.1.4 Optimization 4
1.1.5 Analysis 4
1.1.6 Control 4
1.1.7 Fault Detection 5
1.2 Tasks in Nonlinear System Identification 6
1.2.1 Choice of the Model Inputs 8
1.2.2 Choice of the Excitation Signals 9
1.2.3 Choice of the Model Architecture 10
1.2.4 Choice of the Dynamics Representation 11
1.2.5 Choice of the Model Order 11
1.2.6 Choice of the Model Structure and Complexity 11
1.2.7 Choice of the Model Parameters 12
1.2.8 Model Validation 13
1.2.9 The Role of Fiddle Parameters 13
1.3 White Box, Black Box, and Gray Box Models 15
1.4 Outline of the Book and Some Reading Suggestions 16
1.5 Terminology 18

Part I. Optimization Techniques

2. Introduction to Optimization 23
2.1 Overview of Optimization Techniques 25
2.2 Kangaroos 25
2.3 Loss Functions for Supervised Methods 28
2.3.1 Maximum Likelihood Method 30
2.3.2 Maximum A-Posteriori and Bayes Method 32
2.4 Loss Functions for Unsupervised Methods 34

3. Linear Optimization 35
3.1 Least Squares (LS) 36
3.1.1 Covariance Matrix of the Parameter Estimate 44
3.1.2 Errorbars 45
3.1.3 Orthogonal Regressors 48
3.1.4 Regularization / Ridge Regression 49
3.1.5 Noise Assumptions 54
3.1.6 Weighted Least Squares (WLS) 55
3.1.7 Least Squares with Equality Constraints 57
3.1.8 Smoothing Kernels 58
3.2 Recursive Least Squares (RLS) 60
3.2.1 Reducing the Computational Complexity 63
3.2.2 Tracking Time-Variant Processes 64
3.2.3 Relationship between the RLS and the Kalman Filter 65
3.3 Linear Optimization with Inequality Constraints 66
3.4 Subset Selection 67
3.4.1 Methods for Subset Selection 68
3.4.2 Orthogonal Least Squares (OLS) for Forward Selection 72
3.4.3 Ridge Regression or Subset Selection? 75
3.5 Summary 77

4. Nonlinear Local Optimization 79
4.1 Batch and Sample Adaptation 81
4.2 Initial Parameters 83
4.3 Direct Search Algorithms 86
4.3.1 Simplex Search Method 86
4.3.2 Hooke-Jeeves Method 88
4.4 General Gradient-Based Algorithms 90
4.4.1 Line Search 91
4.4.2 Finite Difference Techniques 92
4.4.3 Steepest Descent 93
4.4.4 Newton's Method 96
4.4.5 Quasi-Newton Methods 98
4.4.6 Conjugate Gradient Methods 100
4.5 Nonlinear Least Squares Problems 102
4.5.1 Gauss-Newton Method 104
4.5.2 Levenberg-Marquardt Method 105
4.6 Constrained Nonlinear Optimization 107
4.7 Summary 110

5. Nonlinear Global Optimization 113
5.1 Simulated Annealing (SA) 116
5.2 Evolutionary Algorithms (EA) 120
5.2.1 Evolution Strategies (ES) 123
5.2.2 Genetic Algorithms (GA) 126

5.2.3 Genetic Programming (GP) 132
5.3 Branch and Bound (B&B) 133
5.4 Tabu Search (TS) 135
5.5 Summary 135

6. Unsupervised Learning Techniques 137
6.1 Principal Component Analysis (PCA) 139
6.2 Clustering Techniques 142
6.2.1 K-Means Algorithm 143
6.2.2 Fuzzy C-Means (FCM) Algorithm 146
6.2.3 Gustafson-Kessel Algorithm 148
6.2.4 Kohonen's Self-Organizing Map (SOM) 149
6.2.5 Neural Gas Network 152
6.2.6 Adaptive Resonance Theory (ART) Network 153
6.2.7 Incorporating Information about the Output 154
6.3 Summary 155

7. Model Complexity Optimization 157
7.1 Introduction 157
7.2 Bias/Variance Tradeoff 158
7.2.1 Bias Error 160
7.2.2 Variance Error 161
7.2.3 Tradeoff 164
7.3 Evaluating the Test Error and Alternatives 167
7.3.1 Training, Validation, and Test Data 168
7.3.2 Cross Validation 169
7.3.3 Information Criteria 171
7.3.4 Multi-Objective Optimization 172
7.3.5 Statistical Tests 174
7.3.6 Correlation-Based Methods 176
7.4 Explicit Structure Optimization 176
7.5 Regularization: Implicit Structure Optimization 179
7.5.1 Effective Parameters 179
7.5.2 Regularization by Non-Smoothness Penalties 180
7.5.3 Regularization by Early Stopping 182
7.5.4 Regularization by Constraints 184
7.5.5 Regularization by Staggered Optimization 186
7.5.6 Regularization by Local Optimization 187
7.6 Structured Models for Complexity Reduction 189
7.6.1 Curse of Dimensionality 190
7.6.2 Hybrid Structures 192
7.6.3 Projection-Based Structures 195
7.6.4 Additive Structures 196
7.6.5 Hierarchical Structures 197
7.6.6 Input Space Decomposition with Tree Structures 198

7.7 Summary 200

8. Summary of Part I 203

Part II. Static Models

9. Introduction to Static Models 209
9.1 Multivariable Systems 209
9.2 Basis Function Formulation 210
9.2.1 Global and Local Basis Functions 211
9.2.2 Linear and Nonlinear Parameters 212
9.3 Extended Basis Function Formulation 215
9.4 Static Test Process 216
9.5 Evaluation Criteria 216

10. Linear, Polynomial, and Look-Up Table Models 219
10.1 Linear Models 219
10.2 Polynomial Models 221
10.3 Look-Up Table Models 224
10.3.1 One-Dimensional Look-Up Tables 225
10.3.2 Two-Dimensional Look-Up Tables 227
10.3.3 Optimization of the Heights 229
10.3.4 Optimization of the Grid 231
10.3.5 Optimization of the Complete Look-Up Table 232
10.3.6 Incorporation of Constraints 232
10.3.7 Properties of Look-Up Table Models 235
10.4 Summary 237

11. Neural Networks 239
11.1 Construction Mechanisms 242
11.1.1 Ridge Construction 242
11.1.2 Radial Construction 244
11.1.3 Tensor Product Construction 245
11.2 Multilayer Perceptron (MLP) Network 246
11.2.1 MLP Neuron 247
11.2.2 Network Structure 249
11.2.3 Backpropagation 252
11.2.4 MLP Training 253
11.2.5 Simulation Examples 256
11.2.6 MLP Properties 260
11.2.7 Multiple Hidden Layers 261
11.2.8 Projection Pursuit Regression (PPR) 262
11.3 Radial Basis Function (RBF) Networks 264
11.3.1 RBF Neuron 264

11.3.2 Network Structure 267
11.3.3 RBF Training 269
11.3.4 Simulation Examples 277
11.3.5 RBF Properties 279
11.3.6 Regularization Theory 281
11.3.7 Normalized Radial Basis Function (NRBF) Networks 283
11.4 Other Neural Networks 286
11.4.1 General Regression Neural Network (GRNN) 286
11.4.2 Cerebellar Model Articulation Controller (CMAC) 288
11.4.3 Delaunay Networks 292
11.4.4 Just-in-Time Models 293
11.5 Summary 296

12. Fuzzy and Neuro-Fuzzy Models 299
12.1 Fuzzy Logic 299
12.1.1 Membership Functions 300
12.1.2 Logic Operators 302
12.1.3 Rule Fulfillment 303
12.1.4 Accumulation 303
12.2 Types of Fuzzy Systems 304
12.2.1 Linguistic Fuzzy Systems 304
12.2.2 Singleton Fuzzy Systems 307
12.2.3 Takagi-Sugeno Fuzzy Systems 309
12.3 Neuro-Fuzzy (NF) Networks 310
12.3.1 Fuzzy Basis Functions 311
12.3.2 Equivalence between RBF and Fuzzy Models 312
12.3.3 What to Optimize? 313
12.3.4 Interpretation of Neuro-Fuzzy Networks 316
12.3.5 Incorporating and Preserving Prior Knowledge 320
12.3.6 Simulation Examples 321
12.4 Neuro-Fuzzy Learning Schemes 323
12.4.1 Nonlinear Local Optimization 323
12.4.2 Nonlinear Global Optimization 325
12.4.3 Orthogonal Least Squares Learning 325
12.4.4 Fuzzy Rule Extraction by a Genetic Algorithm 327
12.4.5 Adaptive Spline Modeling of Observation Data 337
12.5 Summary 339

13. Local Linear Neuro-Fuzzy Models: Fundamentals 341
13.1 Basic Ideas 342
13.1.1 Illustration of Local Linear Neuro-Fuzzy Models 343
13.1.2 Interpretation of the Local Linear Model Offsets 346
13.1.3 Interpretation as Takagi-Sugeno Fuzzy System 347
13.1.4 Interpretation as Extended NRBF Network 349
13.2 Parameter Optimization of the Rule Consequents 351

13.2.1 Global Estimation 351
13.2.2 Local Estimation 352
13.2.3 Global Versus Local Estimation 356
13.2.4 Data Weighting 361
13.3 Structure Optimization of the Rule Premises 362
13.3.1 Local Linear Model Tree (LOLIMOT) Algorithm 365
13.3.2 Structure and Parameter Optimization 372
13.3.3 Smoothness Optimization 374
13.3.4 Splitting Ratio Optimization 376
13.3.5 Merging of Local Models 378
13.3.6 Flat and Hierarchical Model Structures 380
13.3.7 Principal Component Analysis for Preprocessing 383
13.3.8 Models with Multiple Outputs 385
13.4 Summary 389

14. Local Linear Neuro-Fuzzy Models: Advanced Aspects 391
14.1 Different Input Spaces 391
14.1.1 Identification of Direction Dependent Behavior 395
14.2 More Complex Local Models 397
14.2.1 From Local Neuro-Fuzzy Models to Polynomials 397
14.2.2 Local Quadratic Models for Input Optimization 400
14.2.3 Different Types of Local Models 402
14.3 Structure Optimization of the Rule Consequents 404
14.4 Interpolation and Extrapolation Behavior 408
14.4.1 Interpolation Behavior 408
14.4.2 Extrapolation Behavior 411
14.5 Global and Local Linearization 416
14.6 Online Learning 420
14.6.1 Online Adaptation of the Rule Consequents 421
14.6.2 Online Construction of the Rule Premise Structure .. 428
14.7 Errorbars and Design of Excitation Signals 430
14.7.1 Errorbars 431
14.7.2 Detecting Extrapolation 434
14.7.3 Design of Excitation Signals 435
14.7.4 Active Learning 436
14.8 Hinging Hyperplanes 437
14.8.1 Hinging Hyperplanes 438
14.8.2 Smooth Hinging Hyperplanes 439
14.8.3 Hinging Hyperplane Trees (HHT) 441
14.8.4 Comparison with Local Linear Neuro-Fuzzy Models.. 443
14.9 Summary and Conclusions 444

15. Summary of Part II 451

Part III. Dynamic Models

16. Linear Dynamic System Identification 457
16.1 Overview of Linear System Identification 458
16.2 Excitation Signals 459
16.3 General Model Structure 462
16.3.1 Terminology and Classification 465
16.3.2 Optimal Predictor 471
16.3.3 Some Remarks on the Optimal Predictor 474
16.3.4 Prediction Error Methods 476
16.4 Time Series Models 478
16.4.1 Autoregressive (AR) 479
16.4.2 Moving Average (MA) 480
16.4.3 Autoregressive Moving Average (ARMA) 481
16.5 Models with Output Feedback 482
16.5.1 Autoregressive with Exogenous Input (ARX) 482
16.5.2 Autoregressive Moving Average with Exogenous Input 492
16.5.3 Autoregressive Autoregressive with Exogenous Input. 496
16.5.4 Output Error (OE) 499
16.5.5 Box-Jenkins (BJ) 503
16.5.6 State Space Models 505
16.5.7 Simulation Example 506
16.6 Models without Output Feedback 509
16.6.1 Finite Impulse Response (FIR) 510
16.6.2 Orthonormal Basis Functions (OBF) 512
16.6.3 Simulation Example 520
16.7 Some Advanced Aspects 524
16.7.1 Initial Conditions 524
16.7.2 Consistency 526
16.7.3 Frequency-Domain Interpretation 526
16.7.4 Relationship between Noise Model and Filtering 528
16.7.5 Offsets 529
16.8 Recursive Algorithms 531
16.8.1 Recursive Least Squares (RLS) Method 532
16.8.2 Recursive Instrumental Variables (RIV) Method 532
16.8.3 Recursive Extended Least Squares (RELS) Method .. 533
16.8.4 Recursive Prediction Error Methods (RPEM) 534
16.9 Determination of Dynamic Orders 536
16.10 Multivariable Systems 537
16.10.1 P-Canonical Model 539
16.10.2 Matrix Polynomial Model 540
16.10.3 Subspace Methods 541
16.11 Closed-Loop Identification 541

16.11.1 Direct Methods 542
16.11.2 Indirect Methods 544
16.11.3 Identification for Control 545
16.12 Summary 546

17. Nonlinear Dynamic System Identification 547
17.1 From Linear to Nonlinear System Identification 547
17.2 External Dynamics 549
17.2.1 Illustration of the External Dynamics Approach 550
17.2.2 Series-Parallel and Parallel Models 555
17.2.3 Nonlinear Dynamic Input/Output Model Classes 557
17.2.4 Restrictions of Nonlinear Input/Output Models 562
17.3 Internal Dynamics 563
17.4 Parameter Scheduling Approach 564
17.5 Training Recurrent Structures 564
17.5.1 Backpropagation-Through-Time (BPTT) Algorithm . 565
17.5.2 Real Time Recurrent Learning 567
17.6 Multivariable Systems 568
17.7 Excitation Signals 569
17.8 Determination of Dynamic Orders 574
17.9 Summary 576

18. Classical Polynomial Approaches 579
18.1 Properties of Dynamic Polynomial Models 580
18.2 Kolmogorov-Gabor Polynomial Models 581
18.3 Volterra-Series Models 582
18.4 Parametric Volterra-Series Models 583
18.5 NDE Models 583
18.6 Hammerstein Models 584
18.7 Wiener Models 585

19. Dynamic Neural and Fuzzy Models 587
19.1 Curse of Dimensionality 587
19.1.1 MLP Networks 588
19.1.2 RBF Networks 588
19.1.3 Singleton Fuzzy and NRBF Models 588
19.2 Interpolation and Extrapolation Behavior 589
19.3 Training 591
19.3.1 MLP Networks 592
19.3.2 RBF Networks 592
19.3.3 Singleton Fuzzy and NRBF Models 592
19.4 Integration of a Linear Model 593
19.5 Simulation Examples 594
19.5.1 MLP Networks 595
19.5.2 RBF Networks 597

19.5.3 Singleton Fuzzy and NRBF Models ... 599
19.6 Summary ... 600

20. Dynamic Local Linear Neuro-Fuzzy Models ... 601
20.1 One-Step Prediction Error Versus Simulation Error ... 604
20.2 Determination of the Rule Premises ... 606
20.3 Linearization ... 608
20.3.1 Static and Dynamic Linearization ... 608
20.3.2 Dynamics of the Linearized Model ... 610
20.3.3 Different Rule Consequent Structures ... 612
20.4 Model Stability ... 613
20.4.1 Influence of Rule Premise Inputs on Stability ... 614
20.4.2 Lyapunov Stability and Linear Matrix Inequalities ... 616
20.4.3 Ensuring Stable Extrapolation ... 617
20.5 Dynamic LOLIMOT Simulation Studies ... 618
20.5.1 Nonlinear Dynamic Test Processes ... 618
20.5.2 Hammerstein Process ... 620
20.5.3 Wiener Process ... 624
20.5.4 NDE Process ... 625
20.5.5 Dynamic Nonlinearity Process ... 625
20.6 Advanced Local Linear Methods and Models ... 626
20.6.1 Local Linear Instrumental Variables (IV) Method ... 628
20.6.2 Local Linear Output Error (OE) Models ... 630
20.6.3 Local Linear ARMAX Models ... 631
20.7 Local Linear Orthonormal Basis Functions Models ... 631
20.8 Structure Optimization of the Rule Consequents ... 636
20.9 Summary and Conclusions ... 640

21. Neural Networks with Internal Dynamics ... 645
21.1 Fully Recurrent Networks ... 645
21.2 Partially Recurrent Networks ... 646
21.3 State Recurrent Networks ... 647
21.4 Locally Recurrent Globally Feedforward Networks ... 648
21.5 Internal Versus External Dynamics ... 650

Part IV. Applications

22. Applications of Static Models ... 655
22.1 Driving Cycle ... 655
22.1.1 Process Description ... 656
22.1.2 Smoothing of a Driving Cycle ... 657
22.1.3 Improvements and Extensions ... 658
22.1.4 Differentiation ... 659
22.2 Modeling and Optimization of Combustion Engine Exhaust ... 659

22.2.1 The Role of Look-Up Tables 660
22.2.2 Modeling of Exhaust Gases 663
22.2.3 Optimization of Exhaust Gases 666
22.2.4 Outlook: Dynamic Models 672
22.3 Summary 674

23. Applications of Dynamic Models 677
23.1 Cooling Blast 677
23.1.1 Process Description 677
23.1.2 Experimental Results 679
23.2 Diesel Engine Turbocharger 683
23.2.1 Process Description 684
23.2.2 Experimental Results 685
23.3 Thermal Plant 691
23.3.1 Process Description 692
23.3.2 Transport Process 693
23.3.3 Tubular Heat Exchanger 698
23.3.4 Cross-Flow Heat Exchanger 702
23.4 Summary 707

24. Applications of Advanced Methods 709
24.1 Nonlinear Model Predictive Control 709
24.2 Online Adaptation 713
24.2.1 Variable Forgetting Factor 714
24.2.2 Control and Adaptation Models 715
24.2.3 Parameter Transfer 717
24.2.4 Systems with Multiple Inputs 718
24.2.5 Experimental Results 719
24.3 Fault Detection 723
24.3.1 Methodology 723
24.3.2 Experimental Results 726
24.4 Fault Diagnosis 729
24.4.1 Methodology 729
24.4.2 Experimental Results 731
24.5 Reconfiguration 732

A. Vectors and Matrices 735
A.1 Vector and Matrix Derivatives 735
A.2 Gradient, Hessian, and Jacobian 737

B. Statistics 739
B.1 Deterministic and Random Variables 739
B.2 Probability Density Function (pdf) 741
B.3 Stochastic Processes and Ergodicity 743
B.4 Expectation 745

B.5 Variance 748
B.6 Correlation and Covariance 749
B.7 Properties of Estimators 753

References 757

Index 779