

Contents

About the Authors	ix
Preface	xi
Symbols and Abbreviations	xiii
1 Introduction	1
1.1 Elements of System Identification	1
1.2 Traditional Identification Criteria	3
1.3 Information Theoretic Criteria	4
1.3.1 MEE Criteria	6
1.3.2 Minimum Information Divergence Criteria	7
1.3.3 Mutual Information-Based Criteria	7
1.4 Organization of This Book	8
Appendix A: Unifying Framework of ITL	9
2 Information Measures	13
2.1 Entropy	13
2.2 Mutual Information	19
2.3 Information Divergence	21
2.4 Fisher Information	23
2.5 Information Rate	24
Appendix B: α -Stable Distribution	26
Appendix C: Proof of (2.17)	26
Appendix D: Proof of Cramer–Rao Inequality	27
3 Information Theoretic Parameter Estimation	29
3.1 Traditional Methods for Parameter Estimation	29
3.1.1 Classical Estimation	29
3.1.2 Bayes Estimation	31
3.2 Information Theoretic Approaches to Classical Estimation	34
3.2.1 Entropy Matching Method	34
3.2.2 Maximum Entropy Method	35
3.2.3 Minimum Divergence Estimation	37
3.3 Information Theoretic Approaches to Bayes Estimation	40
3.3.1 Minimum Error Entropy Estimation	40
3.3.2 MC Estimation	51

3.4	Information Criteria for Model Selection	56
	Appendix E: EM Algorithm	57
	Appendix F: Minimum MSE Estimation	58
	Appendix G: Derivation of AIC Criterion	58
4	System Identification Under Minimum Error Entropy Criteria	61
4.1	Brief Sketch of System Parameter Identification	61
4.1.1	Model Structure	62
4.1.2	Criterion Function	65
4.1.3	Identification Algorithm	65
4.2	MEE Identification Criterion	72
4.2.1	Common Approaches to Entropy Estimation	73
4.2.2	Empirical Error Entropies Based on KDE	76
4.3	Identification Algorithms Under MEE Criterion	82
4.3.1	Nonparametric Information Gradient Algorithms	82
4.3.2	Parametric IG Algorithms	86
4.3.3	Fixed-Point Minimum Error Entropy Algorithm	91
4.3.4	Kernel Minimum Error Entropy Algorithm	93
4.3.5	Simulation Examples	95
4.4	Convergence Analysis	104
4.4.1	Convergence Analysis Based on Approximate Linearization	104
4.4.2	Energy Conservation Relation	106
4.4.3	Mean Square Convergence Analysis Based on Energy Conservation Relation	111
4.5	Optimization of ϕ -Entropy Criterion	122
4.6	Survival Information Potential Criterion	129
4.6.1	Definition of SIP	129
4.6.2	Properties of the SIP	131
4.6.3	Empirical SIP	136
4.6.4	Application to System Identification	139
4.7	Δ -Entropy Criterion	143
4.7.1	Definition of Δ -Entropy	145
4.7.2	Some Properties of the Δ -Entropy	148
4.7.3	Estimation of Δ -Entropy	152
4.7.4	Application to System Identification	157
4.8	System Identification with MCC	161
	Appendix H: Vector Gradient and Matrix Gradient	164
5	System Identification Under Information Divergence Criteria	167
5.1	Parameter Identifiability Under KLID Criterion	167
5.1.1	Definitions and Assumptions	168
5.1.2	Relations with Fisher Information	169

5.1.3	Gaussian Process Case	173
5.1.4	Markov Process Case	176
5.1.5	Asymptotic KLID-Identifiability	180
5.2	Minimum Information Divergence Identification with Reference PDF	186
5.2.1	Some Properties	188
5.2.2	Identification Algorithm	196
5.2.3	Simulation Examples	198
5.2.4	Adaptive Infinite Impulsive Response Filter with Euclidean Distance Criterion	201
6	System Identification Based on Mutual Information Criteria	205
6.1	System Identification Under the MinMI Criterion	205
6.1.1	Properties of MinMI Criterion	207
6.1.2	Relationship with Independent Component Analysis	211
6.1.3	ICA-Based Stochastic Gradient Identification Algorithm	212
6.1.4	Numerical Simulation Example	214
6.2	System Identification Under the MaxMI Criterion	216
6.2.1	Properties of the MaxMI Criterion	217
6.2.2	Stochastic Mutual Information Gradient Identification Algorithm	222
6.2.3	Double-Criterion Identification Method	227
	Appendix I: MinMI Rate Criterion	238
	References	239