Contents

1 What is chemometrics? 1

1.1 The computer-based laboratory 2
1.2 Statistics and data interpretation 9
1.3 Computer-based information systems and artificial intelligence 10
1.4 General reading 11

2 Basic statistics 13

2.1 Descriptive statistics 14
2.2 Statistical tests 25
2.3 Analysis of variance 41
2.4 General reading 48

3 Signal processing and time-series analysis 51
3.1 Signal processing 51
3.2 Time-series analysis 72
3.3 General reading 78

4 Optimization and experimental design 81

4.1 Objective functions and factors 83
4.2 Experimental design and response surface methods 90
4.2.1 Fundamentals 90
4.2.2 Two-level designs: screening designs 93
4.2.3 Three-level designs: response surface designs 100
4.3 Sequential optimization: the simplex method 110
4.4 General reading 116

5 Pattern recognition and classification 119

5.1 Preprocessing of data 121
5.2 Unsupervised methods 124
5.2.1 Factorial methods 124
5.2.2 Cluster analysis 148
5.2.3 Graphical methods 158
5.3 Supervised methods 160
5.3.1 Linear learning machine 160
5.3.2 Discriminant analysis 162
5.3.3 *k*-nearest neighbor method 168
5.3.4 SIMCA 169
5.4 General reading 172

6 Modeling 175

6.1 Univariate linear regression 176
6.2 Multiple linear regression 192
6.2.1 Ordinary least squares regression 192
6.2.2 Biased parameter estimations: PCR and PLS 196
6.2.3 Applications in multicomponent analysis 200
6.2.4 Regression diagnostics 207
6.3 Nonlinear methods 215
6.3.1 Nonlinear regression analysis 216
6.3.2 Nonparametric methods 220
6.4 General reading 224

7 Analytical databanks 227

7.1 Representation of analytical information 228
7.2 Library search 237
7.3 Simulation of spectra 243
7.4 General reading 244

8 Knowledge processing and soft computing 245

8.1 Artifical intelligence and expert systems 245
8.2 Neural networks 253
8.3 Fuzzy theory 267
8.4 Genetic algorithms 276
8.5 General reading 280

9 Quality assurance and good laboratory practice 283

9.1 Validation and quality control 283
9.2 Accreditation and good laboratory practice 288
9.3 General reading 289

Appendix 291

Statistical distributions 291
Digital filters 297
Experimental designs 299
Matrix algebra 303
Software 307

Index 309