

Contents

Contributors	xi
Preface	xiii

1. Introduction	1
<i>Federico Marini</i>	
1 Another Book on the Wall	1
2 Organisation of the Book	2
References	4

Part I Theory

2. Experimental Design	9
<i>Riccardo Leardi</i>	
1 Introduction	9
2 Full Factorial Design 2^k	12
3 Plackett–Burman Designs	24
4 Central Composite Design	29
5 Doehlert Design	38
6 D-Optimal Designs	40
7 Qualitative Variables at More than Two Levels	43
8 Mixture Designs	46
9 Conclusions	52
References	53

3. Exploratory Data Analysis	55
-------------------------------------	-----------

Mario Li Vigni, Caterina Durante, and Marina Cocchi

1 The Concept (Let Your Data Talk)	55
2 Descriptive Statistics	58
2.1 Frequency Histograms	59
2.2 Box and Whisker Plots	61
3 Projection Techniques	62
3.1 Principal Component Analysis	64
3.2 Other Projection Techniques	108
4 Clustering Techniques	114

5	Remarks	119
	References	120
4.	Regression	127
	<i>Frank Westad, Marta Bevilacqua, and Federico Marini</i>	
1	Introduction	127
2	Multivariate Calibration	129
3	Theory	131
3.1	Univariate Linear Regression: Introducing the Least Squares Concept	132
3.2	Multivariate Generalization of the Ordinary Least Squares Approach	134
3.3	Principal Component Regression	136
3.4	PLS Regression	138
3.5	Principal Covariate Regression	143
4	Validation	144
4.1	Test-set Validation	145
4.2	Cross-Validation	145
5	Diagnostics and Error Measures	147
5.1	Diagnostics	147
5.2	Error Measures	154
6	Model Interpretation	157
6.1	Interpretation of the Structured Part	157
6.2	Plots used to Detect Outliers	159
7	Variable Selection	161
7.1	Using Model Parameter and Diagnostics	163
7.2	Model-Based Variable Importance	163
7.3	iPLS	165
7.4	Genetic Algorithms	165
7.5	Re-sampling Methods: Bootstrap, Jackknifing and Cross-Validation	167
7.6	Cross Model Validation	168
	References	168
5.	Classification and Class-Modelling	171
	<i>Marta Bevilacqua, Remo Bucci, Andrea D. Magri, Antonio L. Magri, Riccardo Nescatelli, and Federico Marini</i>	
1	Introduction	171
1.1	Classification of Classification Methods	172
2	Discriminant Classification Methods	176
2.1	Linear and QDA	177
2.2	Extended Canonical Variates Analysis	188
2.3	Partial Least Squares Discriminant Analysis	195
2.4	k Nearest Neighbours	205
2.5	Density-Based Methods (Potential Functions)	208

2.6 Other Discriminant Classification Methods	215
3 Class-Modelling Methods	215
3.1 Soft Independent Modelling of Class Analogies	215
3.2 Unequal Class-Modelling	224
3.3 Potential Functions as Class-Modelling Methods	225
4 Conclusions	231
References	231
 6. Multivariate Curve Resolution Methods for Food Chemistry	 235
<i>Anna de Juan and Silvia Mas</i>	
1 Introduction	235
2 MCR: The Basics	236
3 MCR Applied to Qualitative and Quantitative Analysis of Compounds in Food Samples	241
4 MCR and Food Fingerprinting	247
5 MCR for Food Processes	253
6 Conclusions	258
References	259
 7. Multiway Methods	 265
<i>José Manuel Amigo and Federico Marini</i>	
1 Introduction: Why Multiway Data Analysis?	266
2 Nomenclature and General Notation	266
3 Parallel Factor Analysis	270
3.1 The General PARAFAC Model	270
3.2 PARAFAC Iterations. Convergence to the Solution. Alternating Least Squares	271
3.3 Properties of PARAFAC Model	272
3.4 Model Validation. Selection of the Number of Factors	273
3.5 Imposing Constraints to the Model	275
3.6 PARAFAC in Practice	276
4 Parallel Factor Analysis 2	281
4.1 PARAFAC2 General Model	282
4.2 Resemblances and Dissimilarities Between PARAFAC and PARAFAC2	283
4.3 Application of PARAFAC2 in Food Research	285
5 Tucker Models	287
5.1 Mathematical Formulation of the Tucker3 Model	287
5.2 Properties of the Tucker3 Model	288
5.3 Other Tucker Models	289
5.4 Some Considerations on the Core Array	290
5.5 Calculating a Tucker3 Model	291
5.6 Tucker3 in Practice	292
6 Multiway Regression	298

6.1 Multilinear PLS (N-PLS)	298
6.2 Multiway Covariate Regression	305
7 Future Perspectives	308
References	309
 8. Robust Methods in Analysis of Multivariate Food Chemistry Data	 315
<i>Ivana Stanimirova, Michał Daszykowski, and Beata Walczak</i>	
1 Introduction	316
2 Basic Concepts in Robust Statistics	317
2.1 Classic and Robust Estimators of Data Location and Scale	319
2.2 Robust Estimates of Covariance and Multivariate Location and Scatter	322
3 Robust Modelling of Data Variance	323
3.1 Spherical Principal Component Analysis	324
3.2 Robust PCA using PP with the Qn Scale	326
3.3 ROBPCA: A Robust Variant of PCA	328
4 Classic and Robust Calibration	329
4.1 Partial Robust M-Regression	330
4.2 RSIMPLS and RSIMCD: Robust Variants of SIMPLS	331
4.3 Spatial Sign Preprocessing and Robust PLS	331
4.4 Identification of Outlying Samples using a Robust Model	332
5 Discrimination and Classification	333
5.1 Classic and Robust Discrimination	333
5.2 Classic and Robust Classification	334
6 Dealing with Missing Elements in Data Containing Outliers	334
7 Further Reading and Software	336
References	337

Part II Applications

9. Hyperspectral Imaging and Chemometrics: A Perfect Combination for the Analysis of Food Structure, Composition and Quality	343
---	-----

José Manuel Amigo, Idoia Martí, and Aoife Gowen

1 Introduction	344
1.1 Quality Assessment	344
1.2 The Role of Hyperspectral Image in Food Quality Assessment	345
1.3 The Need for Chemometrics	345
1.4 Objective of the Book Chapter	350

2 Structure of a Hyperspectral Image	350
3 Hyperspectral Analysis and Chemometrics: Practical Examples	352
3.1 Overview of HSI Data Analysis	352
3.2 Pre-processing Methods	353
3.3 Unsupervised Techniques to Explore the Image: PCA	355
3.4 Supervised Techniques for Classification of Features	358
3.5 Regression Modelling for Obtaining Quantitative Information from Hyperspectral Images	362
4 Final Remarks	364
References	366
 10. The Impact of Chemometrics on Food Traceability	 371
<i>Lucia Bertacchini, Marina Cocchi, Mario Li Vigni, Andrea Marchetti, Elisa Salvatore, Simona Sighinolfi, Michele Silvestri, and Caterina Durante</i>	
1 Introduction	372
1.1 Authenticity and Traceability: The European Union Point of View	373
1.2 Authenticity and Traceability: A Scientific Point of View	374
2 Food Traceability Applications	376
2.1 Chemometrics Approaches for Soil Sampling Planning in Traceability Studies	376
2.2 Geographical Traceability of Raw Materials for PDO and PGI Oenological Products	384
3 Food Authenticity Applications	390
3.1 Study of Grape Juice Heating Process in a Context of Quality Control of Food	390
3.2 Study of Sensory and Compositional Profiles During the Ageing Process of ABTM	397
3.3 Characterisation and Classification of Ligurian Extra Virgin Olive Oil	403
References	407
 11. NMR-Based Metabolomics in Food Quality Control	 411
<i>Alberta Tomassini, Giorgio Capuani, Maurizio Delfini, and Alfredo Miccheli</i>	
1 Introduction	411
2 Methodology	412
2.1 NMR Sample Preparation	413
2.2 NMR Acquisition and Processing Parameters	415
2.3 Targeted Analysis and Pattern Recognition	418
3 NMR-Based Metabolomics Applications	420
3.1 Food Quality Control	420
3.2 Quality Control: Geographical Origin and Authentication	426
3.3 Quality Control, Adulteration, and Safety	433

3.4 Quality Control and Processing	437
References	443

12. Interval-Based Chemometric Methods in NMR Foodomics 449

Francesco Savorani, Morten Arendt Rasmussen, Åsmund Rinnan, and Søren Balling Engelsen

1 Introduction	449
1.1 The Interval Approach	450
2 Interval-Based Methods	456
2.1 Differences Between Interval Approach and Binning (Bucketing)	456
2.2 Different Ways of Defining Intervals	458
2.3 Combinations of Intervals for Model's Performance Improvement	458
2.4 Comparison Between Interval Models and Global Model Performances	460
2.5 Requirements for Bilinear Models	461
2.6 <i>icoshift</i> for Data Alignment	461
2.7 <i>iPCA</i>	465
2.8 <i>iPLS</i>	469
2.9 <i>iPLS-DA</i>	474
2.10 <i>iECVA</i>	477
2.11 Block Chemometric Methods	481
3 Concluding Remarks	482
References	484