

Table of Contents

Acknowledgment	xi
Introduction	xiii
I.1. Types and representation of data	xiii
I.1.1. Binary data	xiv
I.1.2. Categorical data	xiv
I.1.3. Continuous data	xv
I.1.4. Contingency table	xvii
I.1.5. Data representations	xix
I.2. Simultaneous analysis	xx
I.2.1. Data analysis	xx
I.2.2. Co-clustering	xxii
I.2.3. Applications	xxiii
I.3. Notation	xxvii
I.4. Different approaches	xxviii
I.4.1. Two-mode partitioning	xxviii
I.4.2. Two-mode hierarchical clustering	xxxvii
I.4.3. Direct or block clustering	xxxix
I.4.4. Biclustering	xxxix
I.4.5. Other structures and other aims	xliv
I.5. Model-based co-clustering	xlvi
I.6. Outline	xlix

Chapter 1. Cluster Analysis	1
1.1. Introduction	1
1.2. Miscellaneous clustering methods	4
1.2.1. Hierarchical approach	4
1.2.2. The k -means algorithm	5
1.2.3. Other approaches	7
1.3. Model-based clustering and the mixture model	11
1.4. EM algorithm	15
1.4.1. Complete data and complete-data likelihood	16
1.4.2. Principle	17
1.4.3. Application to mixture models	18
1.4.4. Properties	19
1.4.5. EM: an alternating optimization algorithm	19
1.5. Clustering and the mixture model	20
1.5.1. The two approaches	20
1.5.2. Classification likelihood	21
1.5.3. The CEM algorithm	22
1.5.4. Comparison of the two approaches	22
1.5.5. Fuzzy clustering	24
1.6. Gaussian mixture model	26
1.6.1. The model	26
1.6.2. CEM algorithm	28
1.6.3. Spherical form, identical proportions and volumes	29
1.6.4. Spherical form, identical proportions but differing volumes	30
1.6.5. Identical covariance matrices and proportions	31
1.7. Binary data	32
1.7.1. Binary mixture model	32
1.7.2. Parsimonious model	33
1.7.3. Examples of application	35
1.8. Categorical variables	36
1.8.1. Multinomial mixture model	36

1.8.2. Parsimonious model	38
1.9. Contingency tables	41
1.9.1. MNDKI2 algorithm	41
1.9.2. Model-based approach	43
1.9.3. Illustration	47
1.10. Implementation	49
1.10.1. Choice of model and of the number of classes	51
1.10.2. Strategies for use	51
1.10.3. Extension to particular situations	52
1.11. Conclusion	53
Chapter 2. Model-Based Co-Clustering	55
2.1. Metric approach	55
2.2. Probabilistic models	57
2.3. Latent block model	59
2.3.1. Definition	59
2.3.2. Link with the mixture model	61
2.3.3. Log-likelihoods	62
2.3.4. A complex model	63
2.4. Maximum likelihood estimation and algorithms	67
2.4.1. Variational EM approach	69
2.4.2. Classification EM approach	72
2.4.3. Stochastic EM-Gibbs approach	73
2.5. Bayesian approach	75
2.6. Conclusion and miscellaneous developments	76
Chapter 3. Co-Clustering of Binary and Categorical Data	79
3.1. Example and notation	80
3.2. Metric approach	82
3.3. Bernoulli latent block model and algorithms	84
3.3.1. The model	84
3.3.2. Model identifiability	85
3.3.3. Binary LBVEM and LBCEM algorithms	86

3.4. Parsimonious Bernoulli LBMs	90
3.5. Categorical data	91
3.6. Bayesian inference	93
3.7. Model selection	96
3.7.1. The integrated completed log-likelihood (ICL)	96
3.7.2. Penalized information criteria	97
3.8. Illustrative experiments	98
3.8.1. Townships	98
3.8.2. Mero	101
3.9. Conclusion	105

Chapter 4. Co-Clustering of Contingency

Tables	107
4.1. Measures of association	108
4.1.1. Phi-squared coefficient	109
4.1.2. Mutual information	111
4.2. Contingency table associated with a couple of partitions	113
4.2.1. Associated distributions	113
4.2.2. Associated measures of association	116
4.3. Co-clustering of contingency table	119
4.3.1. Two equivalent approaches	119
4.3.2. Parameter modification of criteria	121
4.3.3. Co-clustering with the phi-squared coefficient	124
4.3.4. Co-clustering with the mutual information	129
4.4. Model-based co-clustering	131
4.4.1. Block model for contingency tables	133
4.4.2. Poisson latent block model	137
4.4.3. Poisson LBVEM and LBCEM algorithms	138
4.5. Comparison of all algorithms	140
4.5.1. CROKI2 versus CROINFO	142
4.5.2. CROINFO versus Poisson LBCEM	142
4.5.3. Poisson LBVEM versus Poisson LBCEM	144

4.5.4. Behavior of CROKI2, CROINFO, LBCEM and LBVEM	147
4.6. Conclusion	149
Chapter 5. Co-Clustering of Continuous Data . . .	151
5.1. Metric approach	152
5.1.1. Measure of information	153
5.1.2. Summarized data associated with partitions	153
5.1.3. Objective function	156
5.1.4. CROEUC algorithm	157
5.2. Gaussian latent block model	159
5.2.1. The model	159
5.2.2. Gaussian LBVEM and LBCEM algorithms	160
5.2.3. Parsimonious Gaussian latent block models	161
5.3. Illustrative example	163
5.4. Gaussian block mixture model	168
5.4.1. The model	169
5.4.2. GBEM algorithm	170
5.5. Numerical experiments	173
5.5.1. GBEM versus CROEUC and EM	174
5.5.2. Effect of the size of data	175
5.6. Conclusion	175
Bibliography	177
Index	199