

## Contents

<i>Foreword by Michael W. Berry</i>	<i>page xi</i>
<i>Preface</i>	xiii
<b>1 Introduction and motivation . . . . .</b>	<b>1</b>
1.1 A way to embed ASCII documents into a finite dimensional Euclidean space	3
1.2 Clustering and this book	5
1.3 Bibliographic notes	6
<b>2 Quadratic <math>k</math>-means algorithm . . . . .</b>	<b>9</b>
2.1 Classical batch $k$ -means algorithm	10
2.1.1 Quadratic distance and centroids	12
2.1.2 Batch $k$ -means clustering algorithm	13
2.1.3 Batch $k$ -means: advantages and deficiencies	14
2.2 Incremental algorithm	21
2.2.1 Quadratic functions	21
2.2.2 Incremental $k$ -means algorithm	25
2.3 Quadratic $k$ -means: summary	29
2.3.1 Numerical experiments with quadratic $k$ -means	29
2.3.2 Stable partitions	31
2.3.3 Quadratic $k$ -means	35
2.4 Spectral relaxation	37
2.5 Bibliographic notes	38
<b>3 BIRCH . . . . .</b>	<b>41</b>
3.1 Balanced iterative reducing and clustering algorithm	41

viii	Contents
3.2	BIRCH-like $k$ -means 44
3.3	Bibliographic notes 49
<b>4</b>	<b>Spherical <math>k</math>-means algorithm . . . . . 51</b>
4.1	Spherical batch $k$ -means algorithm 51
4.1.1	Spherical batch $k$ -means: advantages and deficiencies 53
4.1.2	Computational considerations 55
4.2	Spherical two-cluster partition of one-dimensional data 57
4.2.1	One-dimensional line vs. the unit circle 57
4.2.2	Optimal two cluster partition on the unit circle 60
4.3	Spherical batch and incremental clustering algorithms 64
4.3.1	First variation for spherical $k$ -means 65
4.3.2	Spherical incremental iterations–computations complexity 68
4.3.3	The “ping-pong” algorithm 69
4.3.4	Quadratic and spherical $k$ -means 71
4.4	Bibliographic notes 72
<b>5</b>	<b>Linear algebra techniques . . . . . 73</b>
5.1	Two approximation problems 73
5.2	Nearest line 74
5.3	Principal directions divisive partitioning 77
5.3.1	Principal direction divisive partitioning (PDDP) 77
5.3.2	Spherical principal directions divisive partitioning (sPDDP) 80
5.3.3	Clustering with PDDP and sPDDP 82
5.4	Largest eigenvector 87
5.4.1	Power method 88
5.4.2	An application: hubs and authorities 88
5.5	Bibliographic notes 89
<b>6</b>	<b>Information theoretic clustering . . . . . 91</b>
6.1	Kullback–Leibler divergence 91
6.2	$k$ -means with Kullback–Leibler divergence 94
6.3	Numerical experiments 96
6.4	Distance between partitions 98
6.5	Bibliographic notes 99

Cambridge University Press

978-0-521-85267-8 - Introduction to Clustering Large and High-Dimensional Data

Jacob Kogan

Table of Contents

[More information](#)

Contents	ix
<b>7 Clustering with optimization techniques . . . . .</b>	<b>101</b>
7.1 Optimization framework	102
7.2 Smoothing $k$ -means algorithm	103
7.3 Convergence	109
7.4 Numerical experiments	114
7.5 Bibliographic notes	122
<b>8 <math>k</math>-means clustering with divergences . . . . .</b>	<b>125</b>
8.1 Bregman distance	125
8.2 $\varphi$ -divergences	128
8.3 Clustering with entropy-like distances	132
8.4 BIRCH-type clustering with entropy-like distances	135
8.5 Numerical experiments with $(\nu, \mu)$ $k$ -means	140
8.6 Smoothing with entropy-like distances	144
8.7 Numerical experiments with $(\nu, \mu)$ smoka	146
8.8 Bibliographic notes	152
<b>9 Assessment of clustering results . . . . .</b>	<b>155</b>
9.1 Internal criteria	155
9.2 External criteria	156
9.3 Bibliographic notes	160
<b>10 Appendix: Optimization and linear algebra background . . . . .</b>	<b>161</b>
10.1 Eigenvalues of a symmetric matrix	161
10.2 Lagrange multipliers	163
10.3 Elements of convex analysis	164
10.3.1 Conjugate functions	166
10.3.2 Asymptotic cones	169
10.3.3 Asymptotic functions	173
10.3.4 Smoothing	176
10.4 Bibliographic notes	178
<b>11 Solutions to selected problems . . . . .</b>	<b>179</b>
<i>Bibliography</i>	189
<i>Index</i>	203