

ZMATH 2015b.00825

Koch, Inge

Analysis of multivariate and high-dimensional data.

Cambridge Series in Statistical and Probabilistic Mathematics 32. Cambridge: Cambridge University Press (ISBN 978-0-521-88793-9/hbk; 978-1-139-02580-5/ebook). xxv, 504 p. (2014).

This book covers many important material in analysis of multivariate and high-dimensional data, mainly in the framework of component analysis. It is well-written and well-organized. It not only contains important theories with proofs or references but also presents well-chosen real data examples with colorful figures, which greatly helps readers to learn theories, methodologies and their applications. It has also included some latest developments in analysis of multivariate and high-dimensional data. In addition, algorithms, notations, and data index provided in the book are very helpful. Furthermore, the author has made generic MATLAB code for most of the examples and all the data sets are available on the Cambridge University Press website with helpful comments for R users, which is definitely another highlight. This book has three parts: Classical methods, Factors and groupings, and Non-Gaussian analysis. Each part begins with a preliminary chapter and ends with problems for that part. Part I mainly introduces three fundamental multivariate analysis methods including principal component analysis, canonical correlation analysis and discriminant analysis. It contains four chapters. The later sections of each of Chapters 2–4 “contain more advanced or more recent ideas and results, such as principal component analysis for high-dimension low sample size data and principal component regression.” Part II, consisting of four chapters, deals with another three fundamental multivariate analysis methods including cluster analysis, factor analysis, and multidimensional scaling. These three methods “paved the way for non-Gaussian component analysis and in particular for independent component analysis and projection pursuit.” Part III “gives an overview of more recent and current ideas and developments in component analysis methods and links these to statistical learning ideas and research directions for high-dimensional data.” It has five chapters. Three of them focused on independent component analysis, projection pursuit, and kernel independent component analysis. “The final chapter returns to the beginning – principal component analysis – but focuses on current ideas and research directions: feature selection, component analysis of high-dimension low sample size data, decision rules for such data, asymptotics and consistency results when the dimension increases faster than the sample size.” This book can serve as a textbook for a graduate course on analysis of multivariate and high-dimensional data. How much to cover and the prerequisites required for such a course are clearly given in the book preface. This book is definitely an excellent reference book for faculty and professionals. It is also an excellent reference book for researchers working in sciences, social sciences, education, finance, and economic, biological, medical, and health disciplines.

Yuehua Wu (Toronto)

Classification: K45 K85 K75

Keywords: multidimensional data; data analysis

doi:10.1017/CBO9781139025805