

Asymptotic distributions of principal components based on robust dispersions

Hengjian Cui, Xuming He and Kai W. Ng

Department of Mathematics, Beijing Normal University, China

Department of Statistics, University of Illinois at Urbana-Champaign

Department of Actuarial Science and Statistics, The University of Hong Kong

Jan. 23, 2002

Abstract

Principal components are routinely used in multivariate data analysis. Algebraically, they can be defined as the eigenvalues and eigenvectors of a covariance or correlation matrix, but they are statistically meaningful as successive projections of the multivariate data in the direction of maximal variability. An attractive alternative in the robust principal component analysis is to replace the classical variability measure, i.e. variance, by a robust dispersion measure. This projection-pursuit approach was first proposed in Li and Chen (1985) as a method of constructing a robust scatter matrix. Recent work of Croux and Ruiz-Gazen (2001) provided the influence functions of the resulting principal components. The present paper focuses on the asymptotic distributions of robust principal components. In particular, we obtain the asymptotic normality of the principal components that maximize a robust dispersion measure. We also explain the need to use a dispersion functional with a continuous influence function.

Some key words: Asymptotic normality; Dispersion; Principal component; Projection pursuit; Robustness.

PLEASE INDICATE HERE THE NUMBER FROM THE LIST OF TOPICS BELOW WHICH BEST FITS TO YOUR ABSTRACT .

23. Other (Robust principal components)

List of Topics:

1. Algorithms
2. Applications
3. Biostatistics
4. Computing and graphics
5. Data analysis
6. Data mining
7. Economics, finance
8. Efficiency and robustness
9. Functionals and bias
10. Fuzzy statistics
11. Geostatistics
12. Inference for robust methods, model testing
13. Location depth and regression depth
14. Multivariate methods
15. Neural networks
16. Rank-based methods
17. Regression quantiles, trimming
18. Robust covariance
19. Robust designs
20. Robust regression
21. Time series analysis
22. Wavelets
23. Other (please specify)

Corresponding address:

Hengjian Cui

Department of Mathematics

Beijing Normal University

Beijing, 100875

PR. China

Email: hjcui@bnu.edu.cn or cuihj@hotmail.com

URL: <http://math.bnu.edu.cn/~chj>