Estimating the principal component scores through maximum likelihood estimation under normality assumption

Yücel Tandoğdu and Övgü Çıdar

Eastern Mediterranean University, North Cyprus

Abstract

Functional principal component analysis(FPCA) methodology is used in trajectory estimation. Principal component scores ξ_k play an important role in FPCA. ξ_k is a random variable with zero expectation and variance $E(\xi_k^2) = \lambda_k$, λ_k being the eigenvalues of the data covariance matrix. In the estimation of a trajectory, the estimates for ξ_k has to be determined. Since the distribution of ξ_k is unknown, it has to be estimated from available data. In the case of difficulty in determining the distribution of ξ_k, a transformation to normality will facilitate the robust estimation of the principal component scores. Further the maximum likelihood estimators of the statistics for transformed ξ_k will have some important asymptotic properties. Transformed scores are then used in the estimation of trajectories. Different approaches are available for employing ξ_k in FPCA. The methodology developed in this study is compared with the principal component analysis through conditional expectation (PACE) which is an accepted method in estimation problems through FPCA. Estimated trajectories using transformed scores have produced better results compared to those obtained from PACE method.

Keywords

Smoothing, Principal component analysis, Maximum-likelihood, Functional principal component scores, Transformation to normal.

References

Cheng, T. and Feser, V.M.P. (2002). High breakdown estimation of multivariate mean and covariance with missing observaitons. *British J. Math. Statist. Psychology 55*, 317–335.

Li, P. (2005). *Box-Cox Transformations: an Overview*. Department of Statistics, University of Connecticut.

Manly, B.F.J. (1976). Exponential data transformations. *Statistician 25*, 37–42.

Müller, H.G. (2005). Functional modelling and classification of longitudinal data. *Scand. J. Statist. 32*, 223–240.

Park, H.M. (2008). *Univariate Analysis and Normality Test Using SAS, Stata, and SPSS, working paper.* The UITS center for Statistical and Mathematical Computing, Indiana University.

Rubin, D.B. and Dempster, A.P. (1977). Maximum-likelihood from incomplete data via EM algorithm. *J. R. Stat. Soc. Ser. B Stat. Methodol. 39*, 1–38.

Yao, F., Müller, H.G., and Wang, J.L. (2005). Functional data analysis for sparse longitudinal data. *J. Amar. Statist. Assoc. 100*, 577–590.

Yeo, I.K. and Johnson, R. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika 87*, 954–959.