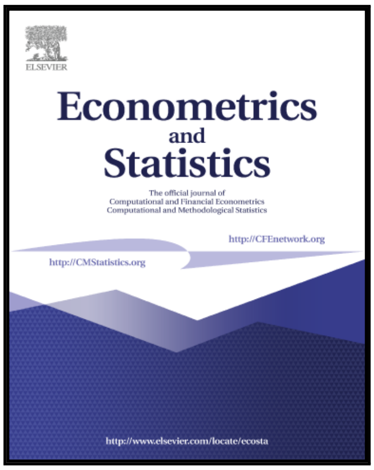
Accepted Manuscript



Preliminary Test Estimation for Multi-sample Principal Components

Davy Paindaveine, Rondrotiana Joséa Rasoafaraniaina,
Thomas Verdebout

PII: S2452-3062(17)30006-0
DOI: 10.1016/j.ecosta.2017.01.004
Reference: ECOSTA 41

To appear in: *Econometrics and Statistics*

Received date: 19 February 2016
Revised date: 9 January 2017
Accepted date: 10 January 2017

Please cite this article as: Davy Paindaveine, Rondrotiana Joséa Rasoafaraniaina, Thomas Verdebout, Preliminary Test Estimation for Multi-sample Principal Components, *Econometrics and Statistics* (2017), doi: 10.1016/j.ecosta.2017.01.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Preliminary Test Estimation for Multi-sample Principal Components

Davy Paindaveine, Rondrotiana Joséa Rasoafaraniaina and Thomas Verdebout

ECARES and Département de Mathématique
Université libre de Bruxelles (ULB)
Avenue F.D. Roosevelt, 50 - ECARES, CP 114/04
and Campus Plaine, CP 210, bvd du triomphe
B-1050 Bruxelles, BELGIUM

Abstract

Point estimation is considered in a multi-sample principal components setup, in a situation where it is suspected that the hypothesis of common principal components (CPC) holds. Preliminary test estimators of the various principal eigenvectors are proposed. Their asymptotic distributions are derived (i) under the CPC hypothesis, (ii) under sequences of hypotheses that are contiguous to the CPC hypothesis, and (iii) away from the CPC hypothesis. A Monte-Carlo study shows that the proposed estimators perform well, particularly so in the Gaussian case.

Keywords: Preliminary test estimation, Common Principal Components

1. Introduction

Principal Component Analysis (PCA) is arguably one of the most popular multivariate methods. In this paper, we consider PCA in a multi-sample context. Consider $m(> 1)$ mutually independent samples of p-vectors $\mathbf{X}_{i1}, \ldots, \mathbf{X}_{in_i}$, $i = 1, \ldots, m$, with respective sample sizes $n_1, \ldots, n_m$, such that for any i, the $\mathbf{X}_{ij}$'s form a random sample from a distribution with mean $\boldsymbol{\theta}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$. In the ith population, the rth principal component scores are

$$(\boldsymbol{\beta}_i^{(r)})'\mathbf{X}_{i1}, \ldots, (\boldsymbol{\beta}_i^{(r)})'\mathbf{X}_{in_i}, \tag{1.1}$$

where $\boldsymbol{\beta}_i^{(r)}$ is the unit eigenvector associated with the rth largest eigenvalue of $\boldsymbol{\Sigma}_i$. In other words, $\boldsymbol{\beta}_i^{(r)}$ is the rth column vector in the matrix $\boldsymbol{\beta}_i$ from the factorization $\boldsymbol{\Sigma}_i = \boldsymbol{\beta}_i \boldsymbol{\Lambda}_i \boldsymbol{\beta}_i'$, where $\boldsymbol{\beta}_i \in \mathcal{SO}_p := \{\mathbf{O} \in \mathbb{R}^{p\times p} : \det(\mathbf{O}) = 1 \text{ and } \mathbf{O}^{-1} = \mathbf{O}\}$ and $\boldsymbol{\Lambda}_i := \mathrm{diag}(\lambda_{i1}, \ldots, \lambda_{ip})$ (with $\lambda_{i1} \geq \lambda_{i2} \geq \ldots \geq \lambda_{ip}$).

If no extra assumptions are adopted, then m eigenvectors matrices, namely $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$, are to be estimated, each on the basis of the observations from the corresponding population,

Download English Version:

https://daneshyari.com/en/article/8919530

Download Persian Version:

https://daneshyari.com/article/8919530

Daneshyari.com