Statistica Sinica **33** (2023), 1-25
doi:https://doi.org/10.5705/ss.202021.0117

COINTEGRATION RANK ESTIMATION FOR HIGH-DIMENSIONAL TIME SERIES WITH BREAKS

Ngai Hang Chan[1] and Rongmao Zhang[2,3]

[1]*City University of Hong Kong,*
[2]*Minnan Normal University and* [3]*Zhejiang University*

Abstract: We propose an intuitive and simple-to-use procedure for estimating the cointegration rank of a high-dimensional time series system with possible breaks. Based on a similar idea to a principal component analysis, the cointegration rank can be estimated by the number of eigenvalues of a certain nonnegative definite matrix. There are several advantages to the new method: (a) the dimension of the cointegrated time series is allowed to vary with the sample size; (b) it is model free; and (c) it is simple to use and robust against possible breaks in trend. The cointegration rank can be estimated without the need for a priori testing and estimating of the break points. The asymptotic properties of the proposed methods are investigated when the dimension of the time series increases with the sample size, which offers a new alternative to deal with high-dimensional time series. Finally, the proposed precedure is demonstrated by means of simulations.

Key words and phrases: Cointegration, eigenanalysis, high-dimensional time series, nonstationary processes, structural break.

1. Introduction

Cointegration was introduced to visualize the long-run equilibrium between several nonstationary economic series. During the past thirty years, cointegration analysis has attracted increasing attention from both theoretical and empirical researchers in economics and statistics alike. An excellent survey on the early developments in cointegration can be found in Johansen (1995).

Engle and Granger (1987) derived a representation for cointegration in the form of an error correction model (ECM), which reflects the correction of the long-run relationship with short-run dynamics. One of the remarkable features of the ECM is that it clearly identifies the gain in prediction achieved using the cointegrated variables, rather than the standard ARIMA approach. However, the associated inference method, now known as the Engle–Granger method, is designed for bivariate series only. The likelihood inference-based ECM is sys-

Corresponding author: Rongmao Zhang, School of Mathematics, Minnan Normal University, Fujian, China. E-mail: rmzhang@zju.edu.cn.