Improved Inference on the Rank of a Matrix

Qihui Chen
School of Management and Economics
The Chinese University of Hong Kong, Shenzhen
qihuichen@cuhk.edu.cn

Zheng Fang
Department of Economics
Texas A&M University
zfang@tamu.edu

May 6, 2018

Abstract

This paper develops a general framework for conducting inference on the rank of an unknown matrix $\Pi_0$. A defining feature of our setup is the null hypothesis of the form $H_0 : \text{rank}(\Pi_0) \leq r$. We argue that the problem is of first order importance because the previous literature instead focuses on $H'_0 : \text{rank}(\Pi_0) = r$ by implicitly assuming away $\text{rank}(\Pi_0) < r$, which may lead to over-rejections for some data generating processes and under-rejections for others (both having $\text{rank}(\Pi_0) < r$). In particular, limiting distributions of test statistics under $H'_0$ may be stochastically dominated by those under $\text{rank}(\Pi_0) < r$. A multiple test on the nulls $\text{rank}(\Pi_0) = 0, \ldots, r$, though valid for $H_0$, may be substantially conservative. We employ a testing statistic whose limiting distributions under $H_0$ are highly nonstandard due to the inherent irregular natures of the problem, and then construct bootstrap critical values that deliver size control and improved power. Since our procedure relies on a tuning parameter, a two-step procedure is designed to mitigate concerns on this nuisance. We additionally argue that our setup is also important for estimation. Empirical relevance of our results is illustrated through a series of examples including testing identification in linear IV models, inference on cointegration rank, estimation of the number of types in finite mixture models, and inference on sorting dimensions in a two-sided matching model with transferrable utility.

Keywords: Matrix rank, Bootstrap, Two-step test, Rank estimation, Identification, Cointegration, Finite mixtures, Matching dimension

*We would like to thank Yonghong An, Brendan Beare, Jiaying Gu, Yingyao Hu, Ivana Komunjer, Qi Li, Peter Phillips, Andres Santos, Yixiao Sun, Xun Tang and seminar participants for helpful discussions and comments, and Frank Kleibergen for generously sharing the data with us.