Mathematical Statistics
Stockholm University
Research Report **2020:2**,
http://www.math.su.se

Spectral analysis of large reflexive generalized inverse and Moore-Penrose inverse matrices

Taras Bodnar[a] and Nestor Parolya[b]

[a]Department of Mathematics, Stockholm University, Stockholm, Sweden

[b]Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands

April 2020

Abstract

A reflexive generalized inverse and the Moore-Penrose inverse are often confused in statistical literature but in fact they have completely different behaviour in case the population covariance matrix is not a multiple of identity. In this paper, we study the spectral properties of a reflexive generalized inverse and of the Moore-Penrose inverse of the sample covariance matrix. The obtained results are used to assess the difference in the asymptotic behaviour of their eigenvalues.

Keywords: reflexive generalized inverse; Moore-Penrose inverse; random matrix theory; free probability.

1 Introduction

Let $\mathbf{Y}_n = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ be the $p \times n$ data matrix which consists of n column vectors of dimension p with $E(\mathbf{y}_i) = \mathbf{0}$ and $Cov(\mathbf{y}_i) = \mathbf{\Sigma}$ for $i \in 1, ..., n$. We assume that $p/n \to c \in (1, +\infty)$ as $n \to \infty$. This type of limiting behavior is also referred to a "large dimensional asymptotics" or "the Kolmogorov asymptotics". In this case, the traditional estimators perform very poorly and tend to over/underestimate the unknown parameters of the asset returns, i.e., the mean vector and the covariance matrix.

Throughout this paper it is assumed that there exists a $p \times n$ random matrix $\mathbf{X}_n$ which consists of independent and identically distributed (i.i.d.) real random variables with zero mean and unit variance such that

$$\mathbf{Y}_n = \mathbf{\Sigma}^{\frac{1}{2}} \mathbf{X}_n \,, \tag{1}$$