Mathematical Statistics
Stockholm University
Research Report **2016:16**,
http://www.math.su.se

Optimal Shrinkage Estimator for High-Dimensional Mean Vector

November 2016

Taras Bodnar[a,1], Ostap Okhrin[b] and Nestor Parolya[c]

[a] *Department of Mathematics, Stockholm University, SE-10691 Stockholm, Sweden*
[b] *Chair of Econometrics and Statistics esp. Transportation, Dresden University of Technology, D-01062 Dresden, Germany*
[c] *Institute of Statistics, Leibniz University of Hannover, D-30167 Hannover, Germany*

Abstract

In this paper we derive the optimal linear shrinkage estimator for the large-dimensional mean vector using random matrix theory. The results are obtained under the assumption that both the dimension p and the sample size n tend to infinity such that $n^{-1}p^{1-\gamma} \to c \in (0, +\infty)$ and $\gamma \in [0, 1)$. Under weak conditions imposed on the the underlying data generating process, we find the asymptotic equivalents to the optimal shrinkage intensities, prove their asymptotic normality, and estimate them consistently. The obtained non-parametric estimator for the high-dimensional mean vector has a simple structure and is proven to minimize asymptotically with probability 1 the quadratic loss in the case of $c \in (0, 1)$. For $c \in (1, +\infty)$ we modify the suggested estimator by using a feasible estimator for the precision covariance matrix. At the end, an exhaustive simulation study and an application to real data are provided where the proposed estimator is compared with known benchmarks from the literature.

AMS 2010 Classification: 60B20, 62H12, 62G20, 62G30
Keywords: mean vector estimation, shrinkage estimator, large-dimensional asymptotics, random matrix theory.

1 Introduction

High-dimensional problems found a front place in the modern statistics with the development of high performance and high storage computers. The latter forces the collection of huge amounts

[1*] Corresponding author. E-mail address: taras.bodnar@math.su.se