ONLINE ISSN : 1348-6365
PRINT ISSN : 1882-2754

JOURNAL OF THE JAPAN STATISTICAL SOCIETY
Vol. 37 (2007) , No. 1 pp.53-86

[PDF (268K)] [References]

Multivariate Theory for Analyzing High Dimensional Data

M. S. Srivastava[1)]

1) Department of Statistics, University of Toronto

Abstract: In this article, we develop a multivariate theory for analyzing multivariate datasets that have fewer observations than dimensions. More specifically, we consider the problem of testing the hypothesis that the mean vector $\boldsymbol{\mu}$ of a p-dimensional random vector $\boldsymbol{x}$ is a zero vector where N, the number of independent observations on $\boldsymbol{x}$, is less than the dimension p. It is assumed that $\boldsymbol{x}$ is normally distributed with mean vector $\boldsymbol{\mu}$ and unknown nonsingular covariance matrix $\sum$. We propose the test statistic $F^{+} = n^{-2} (p - n + 1) N \bar{\boldsymbol{x}}' S^{+} \bar{\boldsymbol{x}}$, where $n = N - 1 < p$, $\bar{\boldsymbol{x}}$ and S are the sample mean vector and the sample covariance matrix respectively, and S^{+} is the Moore-Penrose inverse of S. It is shown that a suitably normalized version of the F^{+} statistic is asymptotically normally distributed under the hypothesis. The asymptotic non-null distribution in one sample case is given. The case when the covariance matrix $\sum$ is singular of rank r but the sample size N is larger than r is also considered. The corresponding results for the case of two-samples and k samples, known as MANOVA, are given.

Key words: distribution of test statistics, DNA microarray data, fewer observations than dimension, multivariate analysis of variance, singular Wishart

[PDF (268K)] [References]

Download Meta of Article[Help]
RIS
BibTeX

To cite this article:

M. S. Srivastava; “Multivariate Theory for Analyzing High Dimensional Data”, *JOURNAL OF THE JAPAN STATISTICAL SOCIETY*, Vol. **37**, pp.53-86 (2007) .

JOI JST.JSTAGE/jjss/37.53