LEAST SQUARES THEORY USING AN ESTIMATED DISPERSION MATRIX AND ITS APPLICATION TO MEASUREMENT OF SIGNALS

C. RADHAKRISHNA RAO
INDIAN STATISTICAL INSTITUTE

1. Introduction

In this paper are considered some problems in the estimation and inference on unknown parameters in a linear model under various assumptions on the error term. We write the linear model in the form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{e} \tag{1}$$

where $\mathbf{Y}$ is a $p \times 1$ vector of observable random variables, $\mathbf{X}$ is $p \times m$ matrix of known coefficients, $\boldsymbol{\tau}$ is a $p \times 1$ vector of unknown (nonstochastic) parameters, and $\mathbf{e}$ is a $p \times 1$ vector of errors. If $\boldsymbol{\Sigma}$, the dispersion matrix of $\mathbf{e}$, is known, then there is no problem, as the method of least squares (for the correlated case) can be applied to estimate and draw inferences on linear parametric functions of $\boldsymbol{\tau}$. We shall consider the case where $\boldsymbol{\Sigma}$ is unknown but an estimate $\hat{\boldsymbol{\Sigma}}$ of $\boldsymbol{\Sigma}$ is available, which may be computed from previous data or from the present data without making any assumption on $\boldsymbol{\tau}$, and discuss how this information can be used. In other words, we will discuss the theory of *least squares using an estimated dispersion matrix*. It is shown that the estimator of $\boldsymbol{\tau}$, obtained by merely substituting $\hat{\boldsymbol{\Sigma}}$ for $\boldsymbol{\Sigma}$ in the least squares estimator of $\boldsymbol{\tau}$ when $\boldsymbol{\Sigma}$ is known, is not necessarily the best. Certain improvements can be made depending on the *known or inferred structure* of $\boldsymbol{\Sigma}$.

Let us denote by E, D, and C the operators for expectation, dispersion, and covariance respectively. We consider the following specific structures for $\boldsymbol{\Sigma}$.

Case 1. The matrix $D(\mathbf{Y}) = \boldsymbol{\Sigma}$ is an unknown arbitrary positive definite matrix.

Case 2. The matrix $\boldsymbol{\Sigma} = \mathbf{X}\boldsymbol{\Gamma}\mathbf{X}' + \mathbf{Z}\boldsymbol{\Theta}\mathbf{Z}' + \sigma^2\mathbf{I}$, where $\boldsymbol{\Gamma}$, $\boldsymbol{\Theta}$, and σ^2 are unknown, and $\mathbf{Z}$ is a matrix such that $\mathbf{X}'\mathbf{Z} = \mathbf{0}$. Such a situation arises when we consider the mixed model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{X}\boldsymbol{\gamma} + \mathbf{Z}\boldsymbol{\xi} + \mathbf{e} \tag{2}$$

where $\boldsymbol{\gamma}$, $\boldsymbol{\xi}$, and $\mathbf{e}$ are all uncorrelated random vectors such that $E(\boldsymbol{\gamma}) = \mathbf{0}$, $D(\boldsymbol{\gamma}) = \boldsymbol{\Gamma}$, $E(\boldsymbol{\xi}) = \mathbf{0}$, $D(\boldsymbol{\xi}) = \boldsymbol{\Theta}$, and $D(\mathrm{e}) = \sigma^2\mathbf{I}$.