The Fundamental Theorem of Least Squares: Its Relevance to Experimental Design and Bayesian Inference

By Heng Li and Alan Zaslavsky

University of Rochester and Harvard Medical School

The necessary and sufficient condition for the least squares estimator to be best linear unbiased estimator is stated for a large class of common linear models. It is shown that this condition implies parallelism between the Bayesian and frequentist inferences under a non-informative reference prior for this class of models.

1. Introduction. It is difficult to find a theorem in statistics that has been stated in as many different ways in the literature as the necessary and sufficient condition for the ordinary least squares (OLS) estimator to be best linear unbiased estimator (BLUE). In a comprehensive review, Puntanen and Styan [9] collected at least twenty different statements of the condition, and claimed that seventeen more may easily be obtained. That this condition is a fundamental theorem that is not so well known in statistics is witnessed by its rediscoveries since its first appearance in [1] ([5, 9]). The class of models in terms of which the theorem has usually been stated may have made it sound more restrictive and less relevant than it really is, and contributed to its relative obscurity. In this paper we explicitly specify a large class of realistic models to which the theorem is applicable, and point out its relevance to Bayesian inference for this class of models.

2. The Fundamental Theorem. The necessary and sufficient condition for the OLS estimator to be BLUE is commonly stated, as in Kruskal's influential paper [6], and in Puntanen and Styan's [9] comprehensive review, for the linear models in which the covariance is known up to a multiplicative constant:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon, \quad E(\epsilon) = \mathbf{0}, \quad cov(\epsilon) = \sigma^2\mathbf{V}, \tag{1}$$

where $\mathbf{V}$ is fixed and known. One of the formulations of the condition, labeled as Z1 (where Z is for Zyskind) in [9], can be expressed as: "A subset of the eigenvectors of $\mathbf{V}$ span the column space of the design matrix $\mathbf{X}$". Notice that since the above statement refers only to the eigenvectors, $\mathbf{V}$ does not have to be completely known in order to verify the condition. In particular, the condition can always be verified if $\mathbf{V}$ has known eigenspaces. Hence condition Z1 is also meaningful and valid for the following model:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon, \quad E(\epsilon) = \mathbf{0}, \quad cov(\epsilon) = \mathbf{\Sigma},$$

where

$$\mathbf{\Sigma} = \lambda_1\mathbf{E}_1 + \cdots + \lambda_K\mathbf{E}_K \tag{2}$$