Gauss-Markov theorem

Canonical name	GaussMarkovTheorem
Date of creation	2013-03-22 15:02:53
Last modified on	2013-03-22 15:02:53
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	8
Author	CWoo (3771)
Entry type	Theorem
Classification	msc 62J05
Synonym	BLUE
Related topic	LinearLeastSquaresFit
Defines	Gauss-Markov linear model
Defines	best linear unbiased estimator

A *Gauss-Markov linear model* is a linear statistical model that satisfies all the conditions of a general linear model except the normality of the error terms. Formally, if $\boldsymbol{Y}$ is an m-dimensional response variable vector, and $\boldsymbol{Z_i} = z_i(\boldsymbol{X})$, $i = 1, \ldots, k$ are the m-dimensional functions of the explanatory variable vector $\boldsymbol{X}$, a Gauss-Markov linear model has the form:

$$\boldsymbol{Y} = \beta_0 \boldsymbol{Z_0} + \cdots + \beta_k \boldsymbol{Z_k} + \boldsymbol{\epsilon},$$

with $\boldsymbol{\epsilon}$ the error vector such that

1. $\mathrm{E}[\boldsymbol{\epsilon}] = \boldsymbol{0}$, and
2. $\mathrm{Var}[\boldsymbol{\epsilon}] = \sigma^2 \boldsymbol{I}$.

In other words, the observed responses Y_i, $i = 1, \ldots, m$ are not assumed to be normally distributed, are not correlated with one another, and have a common variance $\mathrm{Var}[Y_i] = \sigma^2$.

Gauss-Markov Theorem. Suppose the response variable $\boldsymbol{Y} = (Y_1, \ldots, Y_m)$ and the explanatory variables $\boldsymbol{X}$ satisfy a Gauss-Markov linear model as described above. Consider any linear combination of the responses

$$Y = \sum_{i=1}^{m} c_i Y_i, \tag{1}$$

where $c_i \in \mathbb{R}$. If each μ_i is an estimator for response Y_i, parameter θ of the form

$$\theta = \sum_{i=1}^{m} c_i \mu_i, \tag{2}$$

can be used as an estimator for Y. Then, among all unbiased estimators for Y having form (2), the ordinary least square estimator (OLS)

$$\hat{\theta} = \sum_{i=1}^{m} c_i \hat{\mu}_i, \tag{3}$$

yields the smallest variance. In other words, the OLS estimator is the uniformly minimum variance unbiased estimator.

Remark. $\hat{\theta}$ in equation (3) above is more popularly known as the *BLUE*, or the *best linear unbiased estimator* for a linear combination of the responses in a Gauss-Markov linear model.