International Journal of Statistics and Applied Mathematics 2018; 3(1): 65-78

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2018; 3(1): 65-78

www.mathsjournal.com
Received: 12-11-2017
Accepted: 14-12-2017

AK Nagaraju
Assistant Professor in Statistics, Erithriya, South Africa

C Dwarakanatha Reddy
Research Scholar, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

R Abbaiah
Professor, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

N Ramesh Kumar
Lecturer in Statistics, S.V. Arts College, Tirupati, Andhra Pradesh, India

J Prabhakara Naik
Lecturer in Statistics, Loyola Degree College (YSRR), Pulivendula, Kadapa, Andhra Pradesh, India

P Srivyshnavi
Assistant Professor, Department of CSE, SPMVV Engineering College, Tirupat, Andhra Pradesh, India

P Balasiddamuni
Rtd. Professor, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

Correspondence
AK Nagaraju
Assistant Professor in Statistics, Erithriya, South Africa

Testing linear restrictions on parameters of linear model under heteroscedasticity

AK Nagaraju, C Dwarakanatha Reddy, R Abbaiah, N Ramesh Kumar, J Prabhakara Naik, P Srivyshnavi and P Balasiddamuni

Abstract
A large number of problems in Applied Regression Analysis are concerned with the estimation of regression coefficients in the linear regression models subject to linear restrictions. Efficient estimation of linear models under linear restrictions on parameters has received little attention. Generally linear equality or inequality restrictions may be incorporated in the process of estimation of linear statistical models. These restrictions may be either exact or stochastic models. These restrictions may be either exact or stochastic in nature and they give additional or extraneous information about regression coefficients. Objective of the paper is to describe the inferential problems in linear regression models subject to some linear equality restrictions. A generalized linear model with non-spherical disturbances has been specified with exact linear restrictions about the parameters. Later, the maximum likelihood estimation has applied to estimate unknown covariance matrix of disturbances and then estimated the parameters of linear regression model. MSE criterion has been proposed to test for the better performance of the restricted estimated GLS (EGLS) estimators over that of unrestricted estimators of the regression coefficients.

Keywords: Restrictions, parameters, heteroscedasticity

1. Introduction

A large number of problems in 'Applied Regression analysis' are concerned with the estimation of regression coefficients in the linear regression models subject to linear restrictions. Efficient estimation of linear models under linear restrictions on parameters has received little attention. The increasing demand for restricted estimation in multiple regression analysis coupled with greater use of many-parameter systems, necessitates the re-examination of traditional estimation techniques with respect to precision in computation.

Presently, econometricians are showing interest of linear statistical models under a set of equality constraints on the regression coefficients. To improve the efficiency of the estimates of the regression coefficients, the econometricians may re-estimate the linear statistical by incorporating the restrictions in the estimation procedure. For example, under the assumption of constant elasticity of substitution, Cobb-Douglas production function may be estimated by incorporating the restriction namely, the sum of the estimates of the elasticities equal to one.

Generally linear equality or inequality restrictions may be incorporated in the process of estimation of linear statistical models. These restrictions may be either exact or stochastic in nature and they give additional or extraneous information about the regression coefficients. The superiority of restricted least squares estimators over ordinary least squares estimators was studied by Toro-Vizearrondo and Wallace (1968) [10] by using the Mean Square Error (MSE) criterion.

For the last five decades, there has been considerable growth in the research about the inferential aspects of linear statistical models under linear restrictions about the regression coefficients. In this chapter, the various inferential aspects of linear statistical models under linear equality constraints have been explained. Further, the use of dummy variables in testing the equality between sets of regression coefficients in the linear statistical models has been discussed under linear equality restrictions about the parameters of the linear model.

Generally, three types of tests namely Likelihood Ratio, Lagrange Multiplier and Wald tests are applied to test exact linear restrictions on parameters of linear model. Mixed restricted least square estimator is used for sample and stochastic prior information on regression coefficients of linear model. Gourieroux, Holly and Monfort (1982) [5] have discussed the one-sided versions of the Likelihood Ratio, Lagrange Multiplier and Wald tests.

2. Inference in Linear Statistical Model

Suppose there exists a linear regression relationship between an explained variable Y and a set of k-explanatory variables X_1, $X_2,\ldots,X_k$ and is given by

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \in_i, \qquad i = 1, 2, \ldots, n \qquad \ldots(2.1)$$

Sometimes one may impose a condition, $x_{1i}=1$, $\forall$ i and some $1 \leq i \leq k$. i.e., one of the explanatory variable is a constant term. The model (2.1) can be written matrix form as

$$Y_{n\times 1} = X_{n\times k}\beta_{k\times 1} + \in_{n\times 1} \qquad \ldots(2.2)$$

Here, Y is n×1 and X is n×k matrices;
β is vector of regression coefficients;
$\in$ is n×1 vector of disturbances;
n is number of observations and k is number of explanatory variables.
The full ideal conditions (this term is due to Anscombe and tukey (1963)) or assumptions on linear statistical model (2.2) are given by:

1. $E(\in) = 0$
2. $E(\in \in') = \sigma^2 I_n$
3. Rank of X is k, k<n,
4. X is a nonstochastic matrix;
5. X is independent of measuremental error
6. $\lim_{n\to\infty}\left[\frac{X'X}{n}\right] = Q$, Q is a finite non-singular matrix and
7. $\in$ follows a multivariate normal distribution.

The model (2.2) with these assumptions is known as the standard or classical linear statistical model.

Inferential Results

a) The ordinary least squares (OLS) estimator $\hat{\beta} = (X'X)^{-1} X'Y$ is the best linear unbiased estimator (BLUE) of β and its dispersion matrix is given by $\text{var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$. Also, the BLUE of any linear combination P'β is $P'\hat{\beta}$.

b) $\hat{\beta}$ is consistent estimator of β.

c) An unbiased estimator of σ^2 is $\hat{\sigma}^2 = \frac{e'e}{n-k}$ ...(2.3)

Where $e = Y - X\hat{\beta}$ =OLS residual vector.

a) $\hat{Y}_0 = X_0\hat{\beta}$ is the BLUE of $X_0\beta$, which is referred to as the 'best linear unbiased predictor' of Y_0. The mean square error of $\hat{Y}_0$ is given by $\sigma^2\left[I + X_0 (X'X)^{-1} X_0'\right]$.

b) $\hat{\beta}$ has a multivariate normal distribution with mean β and dispersion matrix $\sigma^2 (X'X)^{-1}$.

c) $\frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k}$ ...(2.4)

d) The variance of $\hat{\sigma}^2$ is given by $2\sigma^4/(n-k)$.

e) $\hat{\beta}$ and $\frac{(n-k)\hat{\sigma}^2}{\sigma^2}$ are independently distributed as N[β, $\sigma^2 (X'X)^{-1}$] and χ^2_{n-k} respectively.

f) $\hat{\beta}$ and $(e'e)$ are joint sufficient statistics for β and σ^2.

g) $\hat{\beta}$ and $\hat{\sigma}^2$ are efficient estimators.

h) The Cramer-Rao lower bounds for the variance of $\hat{\beta}$ and $\hat{\sigma}^2$ are $\sigma^2 (X'X)^{-1}$ and $\frac{2\sigma^4}{n}$ respectively.

i) The maximum likelihood estimators of β and σ^2 are $\hat{\beta} = (X'X)^{-1} X'Y$ and $\frac{e'e}{n}$ respectively.

j) $\hat{\beta}$ and $\hat{\sigma}^2$ are asymptotically efficient. The asymptotic variances of $\hat{\beta}$ and $\hat{\sigma}^2$ equal to the Cramer – Lower bounds.

k) $\sqrt{n}\left(\hat{\beta}-\beta\right) AsyN\left(0,\sigma^2 Q^{-1}\right)$, where $Q = \lim_{n\to\infty}\left[\frac{X'X}{n}\right]$ ...(2.5)

l) $\sqrt{n}\left(\hat{\sigma}^2-\sigma^2\right) AsyN\left(0,2\sigma^4\right)$ ...(2.6)

3. Estimation of Linear Statistical Model Under Linear Equality Constraints

Consider a standard linear statistical model as

$$Y_{n\times 1} = X_{n\times k}\beta_{k\times 1} + \in_{n\times 1} \quad ...(3.1)$$

Such that $\in \sim N\left(0,\sigma^2 I_n\right)$.

Write a set of q(≤k) linear equality constraints on the regression coefficients of the model (3.1) as

$$R_{q\times k}\beta_{k\times 1} = r_{q\times 1} \quad ...(3.2)$$

Where R is q×k matrix of known constants with rank q

r is q×1 vector of known constants.

Consider the log likelihood function as

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^{*2}\right) - \frac{1}{2\sigma^{*2}}\left(Y - X\beta^*\right)^{|}\left(Y - X\beta^*\right) \quad ...(3.3)$$

Where σ^{*2} and β^* are the estimates of σ^2 and β to be obtained by maximizing this log likelihood function

$$\frac{\partial \log L}{\partial \sigma^{*2}} = 0 \Rightarrow -\frac{n}{2\sigma^{*2}} + \frac{1}{2\sigma^{*2}}\left(Y - X\beta^*\right)^{|}\left(Y - X\beta^*\right) = 0$$

Or $\sigma^{*2} = \frac{1}{n}\left(Y - X\beta^*\right)^{|}\left(Y - X\beta^*\right)$ ...(3.4)

From (3.3) and (3.4), we have,

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\frac{1}{n}\left(Y - X\beta^*\right)^{|}\left(Y - X\beta^*\right)\right) - \frac{n}{2} \quad ...(3.5)$$

Now, one wishes to maximize log L in (3.5) with respect to β^* subject to the set of linear restrictions Rβ=r.

Since, the maximization of log L is the same as the minimization of the error sum of squares $\left(Y - X\beta^*\right)^{|}\left(Y - X\beta^*\right)$, we write the function for constrained minimization as

$$\phi = \left(Y - X\beta^*\right)^{|}\left(Y - X\beta^*\right) - 2\lambda^{|}\left(R\beta^* - r\right) \quad ...(3.6)$$

Or $\phi = Y^{|}Y 2\beta^{*|}X^{|}Y + \beta^{*|}X^{|}X\beta^* - 2\lambda^{|}\left(R\beta^* - r\right)$ ...(3.7)

Where λ is q×1 vector of q Lagrange multipliers.

Taking the partial derivatives of φ w.r.t. β^* and X gives the first order condition as

1. $\frac{\partial \phi}{\partial \beta^*} = 0 \Rightarrow -2X^{|}Y + 2X^{|}X\beta^* - 2R^{|}\lambda = 0$ ...(3.8)

2. $\frac{\partial \phi}{\partial \lambda} = 0 \Rightarrow -2\left(R\beta^* - r\right) = 0$ ...(3.9)

Or $X^{|}X\beta^* - X^{|}Y - R^{|}\lambda = 0$ ...(3.10)

$R\beta^* - r = 0$ ...(3.11)

Premultiplying equation (3.10) by $R\left(X^{|}X\right)^{-1}$ gives

$$R\beta^* - R\left(X^{|}X\right)^{-1} X^{|}Y - R\left(X^{|}X\right)^{-1} R^{|}\lambda = 0$$

$$\Rightarrow R\beta^* - R\hat{\beta} - R\left(X^{|}X\right)^{-1} R^{|}\lambda = 0$$

Where $\hat{\beta} = \left(X'X\right)^{-1} X'Y$ is the OLS estimator of β.

$$\Rightarrow \lambda = \left[R\left(X^{|}X\right)^{-1} R^{|}\right]^{-1}\left(r - R\hat{\beta}\right) \qquad ...(3.12)$$

Substituting back in equation (3.10) yields

$$\beta^* = \hat{\beta} + \left(X^{|}X\right)^{-1} R^{|}\left[R\left(X^{|}X\right)^{-1} R^{|}\right]^{-1}\left(r - R\hat{\beta}\right) \qquad ...(3.13)$$

This β^* is the Restricted Least Squares (RLS) estimator satisfying the set of q constraints embodied in $R\beta^* = r$.
The mean vector and dispersion matrices of RLS estimator β^* are given by

1. $E\left(\beta^*\right) = \beta + \left(X^{|}X\right)^{-1} R^{|}\left[R\left(X^{|}X\right)^{-1} R^{|}\right]^{-1}\left(r - R\beta\right)$ ...(3.14)

2. $\operatorname{var}\left(\beta^*\right) = \sigma^2\left(X^{|}X\right)^{-1}\left(X^{|}X\right)^{-1} R^{|}\left[R\left(X^{|}X\right)^{-1} R^{|}\right]^{-1} R\left(X^{|}X\right)^{-1}$ ...(3.15)

Under H_0: Rβ = r, the RLS estimator β^* has mean vector β. Thus, RLS estimator β^* is biased estimator of β.

4. Comparing The Risk Functions For Restricted And Unrestricted Least Squares Estimators Using Wallace Weak Mean Square Error Criterion

Consider with K-parameter linear regression model

$$Y_{n\times 1} = X_{n\times k}\beta_{k\times 1} + \in_{n\times 1} \qquad ...(4.1)$$

Under q linear restrictions or hypotheses
H_0: $H^{|}\beta - h = 0$ ...(4.2)

Wallace weak mean square error (MSE) criterion is given by

$$\lambda \le \frac{1}{2}\mu k \operatorname{tr} S^{-1} H\left(HS^{-1}H\right)^{-1} H^{|}S^{-1} = \lambda_0 \qquad ...(4.3)$$

Where (say) $S = \left(X^{|}X\right)$ and μk is the smallest characteristic root of $S = X^{|}X$. The parameter $\lambda = \left(H^{|}\beta - h\right)^{|}\left(H^{|}S^{-1}H\right)^{-1}\left(H^{|}\beta - h\right)/2\sigma^2$ (or) $\delta^{|}\left(H^{|}S^{-1}H\right)^{-1}\delta/2\sigma^2$ and is the non centrality parameter in the $F_{(\lambda.J,n-k)}$ distribution of the test statistic,

$$u=\left(H^{\prime}\hat{\beta}-h\right)^{\prime}\left(H^{\prime}S^{-1}H\right)^{-1}\left(H^{\prime}\hat{\beta}-h\right)/J\hat{\sigma}^{2}$$

Here $\hat{\beta}$ is the OLS least squares estimator of β.

Yancey, Judge and Bock (1973) [2, 12] have compared the risk functions for restricted and unrestricted least squares estimators, with reference to Wallace's (1972) [11] results.

Now using a non-singular matrix W=QP such that Wβ=θ and $Z=XW^{-1}$, the original model (4.1) & (4.2) becomes

$$Y=Z\theta+\in \qquad ...(4.4)$$

And $\left[I_J 0\right]\theta-h_0=0$ ...(4.5)

Where P is such $P^{-1}SP^{-1}=I$. Q is an orthogonal matrix such that

$$QBQ^{\prime}=\begin{bmatrix} I_J & 0 \\ 0 & 0 \end{bmatrix}$$

Since $B=P^{-1^{\prime}}H\left(H^{\prime}S^{-1}H\right)^{-1}H^{\prime}P^{-1}$ is idempotent, and $h_0=G_J^{-1}h$ from $H^{\prime}P^{-1}Q^{\prime}QP^{-1^{\prime}}H=H^{\prime}S^{-1}H=G_JG_J^{\prime}$ where $H^{\prime}P^{-1}Q^{\prime}=\left[G_J 0\right]$, $Z^{\prime}Z=I_k$ and any estimator $\hat{\theta}$ implies an estimator $W^{-1}\tilde{\theta}=\tilde{\beta}$ for β.

The least squares estimator, restricted estimator, test statistic and non-centrality of the model (4.4) and (4.5) are

$$\hat{\theta}=Z^{\prime}Y\sim N\left(0,\sigma^2 I_k\right),\ \hat{\theta}^{\prime}=\left(h_0^{\prime},\hat{\theta}_{K-J}\right),\ u=\left(\hat{\theta}_J-h_0\right)^{\prime}\left(\hat{\theta}_J-h_0\right)/J\hat{\sigma}^2$$

And $\lambda=\left(\theta_J-h_0\right)^{\prime}\left(\theta_J-h_0\right)/2\hat{\sigma}^2$

Respectively where

$$\hat{\theta}_{K-J}=\left[0I_{K-J}\right]\hat{\theta}$$

The weighted risk functions for the least squares unrestricted and restricted estimators which are the same as the un-weighted risk functions of $\hat{\beta}$ and $\hat{\beta}_R$, are

$$E\left(\hat{\beta}-\beta\right)^{\prime}\left(\hat{\beta}-\beta\right)=E\left(\hat{\theta}-\theta\right)^{\prime}W^{-1^{\prime}}W^{-1}\left(\hat{\theta}-\theta\right) \qquad ...(4.6)$$

$$=\sigma^2\mathrm{trW}^{-1^{\prime}}=\sigma^2\mathrm{trS}^{-1}$$

$$=\sigma^2\left(\sum_{i=1}^{J}di+\sum_{i=J+1}^{K}di\right)$$

Where $d_1,\ldots,d_J$ and $d_{J+1},\ldots,d_K$ are characteristic roots of A = $\left[I_J 0\right]W^{-1^{\prime}}W^{-1}\left[I_J 0\right]^{\prime}$ and $\left[0I_{K-J}\right]W^{-1^{\prime}}W^{-1}\left[0I_{K-J}\right]^{\prime}$ respectively and

$$E\left(\hat{\beta}_R-\beta\right)^{\prime}\left(\hat{\beta}_R-\beta\right)=E\left(\hat{\hat{\theta}}-\theta\right)^{\prime}W^{-1^{\prime}}W^{-1}\left(\hat{\hat{\theta}}-\theta\right) \qquad ...(4.7)$$

$$=E\left(\hat{\theta}-\theta\right)^{\prime}W^{-1^{\prime}}W^{-1}\left(\hat{\theta}-\theta\right)+E\left(h_0-\theta_J\right)^{\prime}A\left(h_0-\theta_J\right)-E\left(\hat{\theta}_J-\theta_J\right)^{\prime}A\left(\hat{\theta}_J-\theta_J\right)$$

In order to determine conditions under which the risk functions for b exceeds that for $\hat{\beta}$, we subtract (4.7) from (4.6) giving

$$E\left(\hat{\beta}-\beta\right)^{|}\left(\hat{\beta}-\beta\right)-E\left(\hat{\beta}_R-\beta\right)^{|}\left(\hat{\beta}_R-\beta\right)=-\left(\theta_J-h_0\right)^{|}A\left(\theta_J-h_0\right)+E\left(\hat{\theta}_J-\theta_J\right)^{|}A\left(\hat{\theta}_J-\theta_J\right)$$

Which is non-negative iff

$$E\left(\hat{\theta}_J-\theta_J\right)^{|}A\left(\hat{\theta}_J-\theta_J\right)=\sigma^2\mathrm{trA}\geq-\left(\theta_J-h_0\right)^{|}A\left(\theta_J-h_0\right) \quad ...(4.9)$$

From a theorem on extreme of quadratic forms (C.R.Rao, 1965)

$$\underset{(\theta_J-h_0)}{Sup}\frac{\left(\theta_J-h_0\right)^{|}A\left(\theta_J-h_0\right)}{\left(\theta_J-h_0\right)^{|}\left(\theta_J-h_0\right)}=d_L$$

Where d_L is the largest characteristic root of A, we have

$$\sigma^2\mathrm{tr}\left(A\right)\geq\left(\theta_J-h_0\right)^{|}A\left(\theta_J-h_0\right)\leq d_L\left(\theta_J-h_0\right)^{|}\left(\theta_J-h_0\right)=d_L 2\lambda\sigma^2 \quad ...(4.10)$$

From (4.10) we conclude that (4.8) is non negation if

$$\lambda\leq\frac{1}{2}d_L^{-1}\mathrm{trA} \quad ...(4.11)$$

To explore the structure of the difference of the two risk functions using (4.8), we use an orthogonal transformation C to obtain the characteristic roots of A, so

$$E\left(\hat{\beta}-\beta\right)^{|}\left(\hat{\beta}-\beta\right)-E\left(\hat{\beta}_R-\beta\right)^{|}\left(\hat{\beta}_R-\beta\right)=\sigma^2\mathrm{trA}-\left(\theta_J-h_0\right)^{|}C^{|}CAC^{|}C\left(\theta_J-h_0\right) \quad ...(4.8a)$$

$$=\sigma^2\mathrm{trA}-\xi^{|}D\xi=\sigma^2\sum_{i=1}^{J}di-\sum_{i=1}^{J}\xi_i^2 di$$

Where D = $CAC^{|}$ is a diagonal matrix whose elements are characteristic roots, d_i, of A and ξ is a vector incorporating the specification error in the exact restrictions. To find the smallest value of λ for which we are sure that the difference in (4.8a) is positive, one can perform the conceptual experiment of varying the values of the elements in ξ in $\xi_i^2 di$ while $\lambda=\left(\sum\xi_i^2\right)\big/2\sigma^2$ constant, so that all ξ_i are zero except the one associated with d_L. The non-zero ξ_i is replaced by $\xi_L^2=\sum\xi_i^2$ thus $2\lambda\sigma^2=\xi_i^2$

And $E\left(\hat{\beta}-\beta\right)^{|}\left(\hat{\beta}-\beta\right)-E\left(\hat{\beta}_R-\beta\right)^{|}\left(\hat{\beta}_R-\beta\right)\geq\sigma^2\sum_{i=1}^{J}di 2\lambda\sigma^2 d_L$ ...(4.8b)

Since we do not know the values of ξ, we can only be sure the risk function of $\hat{\beta}$ is greater than or equal to that $\hat{\beta}_R$ if $\lambda\leq\sum_{i=1}^{J}di\big/2d_L$ and without information on the values of ξ no tighter bound is possible. A similar conceptual experiment associated with the smallest characteristic root of A, d_S, shows that the risk of $\hat{\beta}_R$ will exceed that the b whenever $\lambda\leq\sum_{i=1}^{J}di\big/2d_S$. Given the value of λ, if no other information about $(\theta_J\text{-}h_0)$ is known, then nothing can be said in a given problem situation about which risk function is larger when $\lambda\leq\left(\sum di\right)\big/2d_S$. For the risk function of $\hat{\beta}_R$, the inequality

$$\sigma^2\left[\sum_{i=J+1}^{K} di+2\lambda d_S\right]\le E\left(\hat{\beta}_R-\beta\right)^{|}\left(\hat{\beta}_R-\beta\right)\le\sigma^2\left[\sum_{i=J+1}^{K} di+2\lambda d_L\right] \quad ...(4.12)$$

Holds for each value of λ with the equalities holding at λ=0 where the risk function of $\hat{\beta}_R$ is smallest. The terms in (4.12) are also equal if all the roots of S are equal, as is the case for an idempotent matrix. Since H and S are known, the characteristic roots of $S^{-1}H\left(H^{|}S^{-1}H\right)^{-1}H^{|}S^{-1}$, di, are easily obtainable.

For a comparison with the Wallace results for the risk function case we observe that $S^{-1}H\left(H^{|}S^{-1}H\right)^{-1}H^{|}S^{-1}$ can be written as

$$P^{-1}Q^{|}QP^{-1}H\left(H^{|}S^{-1}H\right)^{-1}H^{|}P^{-1}Q^{|}QP^{-1^{|}}=W^{-1}\left[I_J 0\right]^{|}\left[I_J 0\right]W^{-1}$$

The non-zero characteristic root of $W^{-1}\left[I_J 0\right]^{|}\left[I_J 0\right]W^{-1}$ and

$$\begin{pmatrix} I_J & 0 \\ 0 & 0 \end{pmatrix} W^{-1}W^{-1}\begin{pmatrix} I_J & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} A. & 0 \\ 0 & 0 \end{pmatrix}$$

are the same. Since the order of matrices does not affect the values of the characteristic roots of their products. For the same reason, S^{-1} and $W^{-1^{|}}W^{-1}$ have the same characteristic roots. In addition, $\text{tr}W^{-1}\left[I_J 0\right]^{|}\left[I_J 0\right]W^{-1^{|}}=\text{trA}$, and $\text{trS}^{-1}=\text{tr}W^{-1^{|}}W^{-1}$. To see that d_L is no larger than the largest characteristic roots of S^{-1}, d_L^*,
We note that

$$\underset{x}{Sup}\frac{x^{|}W^{-1^{|}}Wx}{x^{|}x}=d_L^*$$

is greater than

$$\underset{\left(x_j^{|}0^{|}\right)}{Sup}\frac{\left(x_j^{|}0^{|}\right)W^{-1^{|}}W\left(x_j^{|}0^{|}\right)^{|}}{\left(x_j^{|}0^{|}\right)\left(x_j^{|}0^{|}\right)^{|}}=d_L.$$

Hence the largest root of S^{-1} is greater than or equal to that of $S^{-1}H\left(H^{|}S^{-1}H\right)^{-1}H^{|}S^{-1}$.

Thus is a sufficient condition, and the smallest one that will always hold, for the risk of $\hat{\beta}$ to be at least as large as $\hat{\beta}_R$ such that

$$\lambda\le\frac{1}{2d_L}\text{trA}=\frac{1}{2d_L}\text{tr}S^{-1}H\left(H^{|}S^{-1}H\right)^{-1}H^{|}S^{-1} \quad ...(4.13)$$

This expression is the same as (4.3) with µk, the reciprocal of the largest root of S^{-1}, replaced by $1/d_L$. The lower bound obtained for λ in (4.13) must be greater than or equal to that in the Wallace result.

5. Testing Linear Restrictions on Parameters of Linear Model Under Heteroscedasticity

Consider the linear regression model

$$Y_{n\times1}=X_{n\times k}\beta_{k\times1}+\in_{n\times1} \quad ...(5.1)$$

Such that $E(\in)=0$

$$E\left(\in\in^{|}\right)=\sigma^2\Omega$$

and $\in$ follows normal distribution $N\left(0,\sigma^2\Omega\right)$

where Y is (n×1) vector, X is (n×k) fixed known data matrix with full rank k, k<n;
β is (k×1) fixed and unknown vector of parameters;
$\in$ is (n×1) vector of jointly normal distributions with mean vector zero and positive definite covariance matrix Ω.
The q linear restrictions on β, with respect to the extraneous information, under consideration are:

$$R_{q\times k}\beta_{k\times 1} = r_{q\times 1} \qquad \text{...(5.2)}$$

Where R is (q×k) known matrix of full rank q, q<k;
r is (q×1) known vector.
Since Ω is positive definite matrix, there exists a non-singular matrix P such that PP$^{|}$=Ω or $P^{-1}\Omega P^{|^{-1}} = I$ or $\left(P^{|}\Omega P\right)^{-1} = I$ or

$P^{|}\Omega P = I$.

Premultiplication of (5.1) of P^{-1} gives,

$$P^{-1}Y = P^{-1}X\beta + P^{-1}\in$$

Or $Y^{*} = X^{*}\beta + \in^{*}$ ...(5.3)

Such that $E\left(\in^{*}\right) = 0$

And $E\left(\in^{*}\in^{*|}\right) = \sigma^{2}I$

Since $\in\sim N\left(0,\sigma^{2}\Omega\right)$

$$\in^{*} = P^{-1}\in\sim N\left(0,\sigma^{2}I_{n}\right)$$

The unrestricted and restricted least squares estimators of β is (5.3) are identical to the unrestricted and restricted generalized least squares estimators of β in (5.1).
When Ω is known, all the Mean Square Error (MSE) tests for least squares estimators can be applied for the generalized least squares estimators.
Consider the system of generalized linear equations along with the linear restrictions as

$$\begin{bmatrix} Y^{*} \\ r \end{bmatrix} = \begin{bmatrix} X^{*} \\ R \end{bmatrix}\beta + \begin{bmatrix} \in^{*} \\ 0 \end{bmatrix} \qquad \text{...(5.4)}$$

Since the matrix R is of full rank, there exists at least one non-singular (q×q) matrix in R. Thus, we may write

$$\begin{bmatrix} R \\ X^{*} \end{bmatrix} = \begin{bmatrix} R_{1} & R_{2} \\ X_{1}^{*} & X_{2}^{*} \end{bmatrix}$$

Where R_1 is non-singular matrix of order q.

Write β as $\begin{bmatrix} \beta_{1} \\ \beta_{2} \end{bmatrix}$

Here, X^* and β are conformably partitioned.
Thus (5.4) can be expressed as

$$\begin{bmatrix} Y^{*} \\ r \end{bmatrix} = \begin{bmatrix} X_{1}^{*} & X_{2}^{*} \\ R_{1} & R_{2} \end{bmatrix}\begin{bmatrix} \beta_{1} \\ \beta_{2} \end{bmatrix} + \begin{bmatrix} \in^{*} \\ 0 \end{bmatrix} \qquad \text{...(5.5)}$$

$$\Rightarrow X_{1}^{*}\beta_{1} + X_{2}^{*}\beta_{2} + \in^{*} = Y^{*} \qquad \text{...(5.6)}$$

and $R_1\beta_1 + R_2\beta_2 = r$ ...(5.7)

From (5.7), we have, $\beta_1 = R_1^{-1}\left[r - R_2\beta_2\right]$ ...(5.8)

(5.6) and (5.8) $\Rightarrow X_1^* R_1^{-1}\left(r - R_2\beta_2\right) + X_2^*\beta_2 + \in^* = Y^*$

or $\left[Y^* - X_1^* R_1^{-1} r\right] = \left[X_2^* - X_1^* R_1^{-1} R_2\right]\beta_2 + \in^*$

The GLS estimator of β_2 is given by

$$\tilde{\beta}_2 = \left\{\left[X_2^* - X_1^* R_1^{-1} R_2\right]^{|}\left[X_2^* - X_1^* R_1^{-1} R_2\right]\right\}^{-1}\left[X_2^* - X_1^* R_1^{-1} R_2\right]\left[Y^* - X_1^* R_1^{-1} r\right] \quad ...(5.9)$$

Since $PP^{|} = \Omega$, we have,

$$\tilde{\beta}_2 = \left\{\left[X_2 - X_1 R_1^{-1} R_2\right]^{|}\Omega^{-1}\left[X_2 - X_1 R_1^{-1} R_2\right]\right\}^{-1}\left[X_2 - X_1 R_1^{-1} R_2\right]\Omega^{-1}\left[Y - X_1 R_1^{-1}\Omega\right] \quad ...(5.10)$$

By substituting (5.10) for β_2 into (5.8) gives the generalized least squares estimate of β_1.

i.e., $\tilde{\beta}_1 = R_1^{-1}\left[r - R_2\tilde{\beta}_2\right]$ ...(5.11)

It can be easily shown that $\tilde{\beta}_1$ and $\tilde{\beta}_2$ unbiased estimators of β_1 and β_2 respectively.

The dispersion matrix of $\tilde{\beta} = \begin{bmatrix} \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{bmatrix}$ is given by

$$\operatorname{var}\left(\tilde{\beta}\right) = \sigma^2 \begin{bmatrix} R_1^{-1}R_2\left(Z^{*|}Z^*\right)^{-1} R_2^{|}R_1^{-1} & R_1^{-1}R_2\left(Z^{*|}Z^*\right)^{-1} \\ \left[R_1^{-1}R_2\left(Z^{*|}Z^*\right)^{-1}\right]^{|} & \left(Z^{*|}Z^*\right)^{-1} \end{bmatrix} \quad ...(5.12)$$

Where $Z^* = \left[X_2^* - X_1^* R_1^{-1} R_2\right] = P^{-1}\left[X_2 - X_1 R_1^{-1} R_2\right]$

and $Z^{*|}Z^* = \left[X_2 - X_1 R_1^{-1} R_2\right]^{|}\Omega^{-1}\left[X_2 - X_1 R_1^{-1} R_2\right] = Z^{|}\Omega^{-1}Z$

Here, $Z = \left[X_2 - X_1 R_1^{-1} R_2\right]$

$$\therefore \operatorname{var}\left(\tilde{\beta}\right) = \sigma^2 \begin{bmatrix} R_1^{-1}R_2\left(Z^{|}\Omega^{-1}Z\right)^{-1} R_2^{|}R_1^{-1} & R_1^{-1}R_2\left(Z^{|}\Omega^{-1}Z\right)^{-1} \\ \left[R_1^{-1}R_2\left(Z^{|}\Omega^{-1}Z\right)^{-1}\right]^{|} & \left(Z^{|}\Omega^{-1}Z\right)^{-1} \end{bmatrix} \quad ...(5.13)$$

Further, the dispersion matrix of the OLS estimator $\hat{\beta}$ of β is given by

$$\operatorname{var}\left(\hat{\beta}\right) = \sigma^2 \begin{bmatrix} X_1^{|}X_1 & X_1^{|}X_2 \\ X_2^{|}X_1 & X_2^{|}X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1^{|}\Omega X_1 & X_1^{|}\Omega X_2 \\ X_2^{|}\Omega X_1 & X_2^{|}\Omega X_2 \end{bmatrix} \begin{bmatrix} X_1^{|}X_1 & X_1^{|}X_2 \\ X_2^{|}X_1 & X_2^{|}X_2 \end{bmatrix}^{-1} \quad ...(5.14)$$

It can be shown that

$$\operatorname{var}\left(\sum C_i\tilde{\beta}_i\right) \le \operatorname{var}\left(\sum C_i\hat{\beta}_i\right), \text{for } 1 \le \text{i} \le \text{k}$$

Where ($C_1, C_2, \dots, C_k$) are known coefficients in the linear combination of estimators.

6. Estimation Of Unknown Error Covariance Matrix Ω

Consider the linear regression model

$$Y = X\beta + \in \text{ with } \in \sim N\left(0, \sigma^2\Omega\right)$$

Where Ω is unknown covariance matrix.
It is assumed that Ω depends upon a finite number of unknown parameters $\theta_1, \theta_2, \dots, \theta_m$.
If the parameters $\theta_1, \theta_2, \dots, \theta_m$ are known, then the Generalized Least Squares (GLS) estimator would be the 'Best Linear Unbiased Estimator (BLUE)' and the maximum likelihood estimator.
Generally ($\theta_1, \theta_2, \dots, \theta_m$) are unknown and the parameters ($\beta_1, \beta_2, \dots, \beta_k$) will be estimated along with ($\theta_1, \theta_2, \dots, \theta_m$). For this the maximum likelihood estimation may be used.
Given Ω is a positive definite matrix, whose elements are twice differential functions of a finite and constant number of parameters ($\theta_1, \theta_2, \dots, \theta_m$).
Thus, Ω = Ω(θ), θ belongs to Θ.
The parameters β_i's are independent from those in θ.
Write the logarithmic likelihood function of $\in$.

$$\ln L\left(\beta, \sigma^2, \Omega(\theta)/Y_1X\right) = -\frac{n}{2}\ln\sigma^2 - \frac{1}{2}\ln\left|\Omega(\theta)\right| - \frac{1}{2\sigma^2}(Y - X\beta)^{|}\Omega^{-1}(\theta)(Y - X\beta) \qquad \text{...(6.1)}$$

Maximization of ln L with respect to β and σ^2 conditional on θ, gives the maximum likelihood estimators for β and σ^2 as

$$\bar{\beta}(\theta) = \left[X^{|}\Omega^{-1}(\theta)X\right]^{-1}\left[X^{|}\Omega^{-1}(\theta)Y\right] \qquad \text{...(6.2)}$$

$$\text{and } \bar{\sigma}^2(\theta) = \frac{\left[Y - X\bar{\beta}(\theta)\right]^{|}\Omega^{-1}(\theta)\left[Y - X\bar{\beta}(\theta)\right]}{n} \qquad \text{...(6.3)}$$

Using (6.2) and (6.3) in (6.1) and ignoring the constants, yields the concentrated logarithmic likelihood function as

$$\ln L(\theta) = -n\ln\left\{\left[Y - X\bar{\beta}(\theta)\right]^{|}\Omega^{-1}(\theta)\left[Y - X\bar{\beta}(\theta)\right]\right\} - \ln\left|\Omega(\theta)\right| \qquad \text{...(6.4)}$$

Maximization of ln L(θ) with respect to θ or equivalently minimization of

$$\phi(\theta) = \left|\Omega(\theta)\right|^{1/n}\left[Y - X\bar{\beta}(\theta)\right]^{|}\Omega^{-1}(\theta)\left[Y - X\bar{\beta}(\theta)\right] \qquad \text{...(6.5)}$$

with respect to θ gives the maximum likelihood estimator for θ as $\bar{\theta}$.
Now, for $\Omega^* = \Omega\left(\bar{\theta}\right)$, the maximum likelihood estimator for β and σ^2 are given by

$$\hat{\bar{\beta}} = \bar{\beta}\left(\bar{\theta}\right) = \left[X^{|}\Omega^{*^{-1}}X\right]^{-1}\left[X^{|}\Omega^{*^{-1}}Y\right] \qquad \text{...(6.6)}$$

$$\text{and } \hat{\bar{\sigma}}^2 = \bar{\sigma}^2\left(\bar{\theta}\right) = \frac{\left[Y - X\hat{\bar{\beta}}\right]^{|}\Omega^{*^{-1}}\left[Y - X\hat{\bar{\beta}}\right]}{n} \qquad \text{...(6.7)}$$

In general, the maximization of (6.4) does not give a closed form solution for $\bar{\theta}$. A two-step iterative procedure has been suggested by:

1. First choose $\theta = \theta_0$ belongs to Θ, the class of all admissible values of θ.
2. Compute $\Omega_0^{-1} = \Omega^{-1}\left(\theta_0\right)$, $\tilde{\beta}_0 = \left(X^{|}\Omega_0^{-1}X\right)^{-1}\left(X^{|}\Omega_0^{-1}X\right)$ and hence the residual vector $e_0 = \left(Y - X\tilde{\beta}_0\right)$.
3. Substituting e_0 into the first order condition for the maximization of ln L, gives

$$\text{Trace}\left[\frac{\partial\Omega^{-1}}{\partial\theta_i}\Omega\right]_{\theta=\theta_0} = e_0^{|}\left[\frac{\partial\Omega^{-1}}{\partial\theta_i}\right]_{\theta=\theta_0} e_0 \quad \text{...(6.8)}$$

i=1,2,...,m.

Where $\left[\frac{\partial\Omega^{-1}}{\partial\theta_i}\right]$ is (n×n) matrix of partial derivatives of each element in Ω^{-1} with respect to θ_i.

The system (6.7) gives m nonlinear equations in m unknowns $\theta_1, \theta_2, \ldots, \theta_m$. If one writes this θ-equation explicitly as θ=θ(e), then one can put $\theta_1=\theta(e_0)$. If an explicit solution does not exists then one may find more than one solution. In this case one may select the solution θ_1 with the highest likelihood.

Now, calculate $\Omega_1^{-1} = \Omega^{-1}(\theta_1)$,

$$\tilde{\beta}_1 = \left(X^{|}\Omega_1^{-1}X\right)^{-1}\left(X^{|}\Omega_1^{-1}Y\right) \quad \text{...(6.9)}$$

and hence $e_1 = \left(Y - X\tilde{\beta}_1\right)$ ...(6.10)

This iterative procedure may be continued until converge. The two step generalized least squares estimators $\Omega_1 = \Omega(\theta_1)$ and $\tilde{\beta}_1$ may be used in practice as the estimated generalized least squares (EGLS) for Ω and β.

7. Mean Squared Error Criteria Based On Egls Estimators

Consider unrestricted GLS estimator of β as

$$\beta^* = \left(X^{|}\Omega^{-1}X\right)^{-1}\left(X^{|}\Omega^{-1}Y\right) \quad \text{...(7.1)}$$

The sampling distribution of β^* is given by

$$\sqrt{n}\left(\beta^* - \beta\right) \sim N\left(0, \sigma^2 n\left(X^{|}\Omega^{-1}X\right)^{-1}\right) \quad \text{...(7.2)}$$

The Mean Square Error (MSE) is given by

$$\text{MSE}\sqrt{n}\left(\beta^*\right) = \sigma^2 n\left(X^{|}\Omega^{-1}X\right)^{-1} \quad \text{...(7.3)}$$

Also, we have

$$\frac{\left(Y - X\beta^*\right)^{|}\Omega^{-1}\left(Y - X\beta^*\right)}{\sigma^2} \sim \chi^2_{(n-k)} \quad \text{...(7.4)}$$

Replacing Ω with the two step generalized least squares estimator $\Omega_1 = \Omega(\theta_1)$ gives the estimated MSE of $\sqrt{n}\left(\beta^*\right)$ as

$$\text{MSE}\left(\sqrt{n}\beta^*\right) = \sigma^2 n\left(X^{|}\Omega_1^{-1}X\right)^{-1} \quad \text{...(7.5)}$$

Consider the restricted GLS estimator of β as $\tilde{\beta}$, which is given by

$$\tilde{\beta} = \beta^* - \left(X^{|}\Omega^{-1}X\right)^{-1}R^{|}\left[R\left(X^{|}\Omega^{-1}X\right)^{-1}R^{|}\right]^{-1}\left(R\beta^* - \Omega\right) \quad \text{...(7.6)}$$

We have

$$\sqrt{n}\left(\tilde{\beta}-\beta\right) \sim N\{[-\left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\left[R\left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\right]^{-1}\left(R\beta^{*}-r\right),$$
$$\sigma^{2}n\left(X^{\prime}\Omega^{-1}X\right)^{-1}\left[I-R^{\prime}\left[R\left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\right]^{-1}R\left(X^{\prime}\Omega^{-1}X\right)^{-1}\right]]\} \quad ...(7.7)$$

The MSE of $\sqrt{n}\tilde{\beta}$ is given by

$$\sqrt{n}\tilde{\beta} = \left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\left[R\left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\right]^{-1}n\left(R\beta-r\right)\left(R\beta-r\right)^{\prime}$$
$$\left[R\left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\right]^{-1}R\left(X^{\prime}\Omega^{-1}X\right)^{-1}+\sigma^{2}n\left(X^{\prime}\Omega^{-1}X\right)^{-1} \quad (7.8)$$
$$\left[I-R^{\prime}\left[R\left(X^{\prime}\Omega^{-1}X\right)^{-1}R^{\prime}\right]^{-1}R\left(X^{\prime}\Omega^{-1}X\right)^{-1}\right]$$

Where, subject to $\sqrt{n}\left(R\beta-r\right)=0$, $\tilde{\beta}$ minimizes the sum of squared transformed residuals

$$\mathrm{SSE}\left(\tilde{\beta}\right)=\left(Y-X\tilde{\beta}\right)^{\prime}\Omega^{-1}\left(Y-X\tilde{\beta}\right) \quad ...(7.9)$$

By replacing Ω with two step GLS estimator Ω_1 gives

$$\mathrm{MSE}\left(\sqrt{n}\tilde{\beta}\right)=\{\left(X^{\prime}\Omega_1^{-1}X\right)^{-1}R^{\prime}\left[R\left(X^{\prime}\Omega_1^{-1}X\right)^{-1}R^{\prime}\right]^{-1}n\left(R\beta-\Omega\right)\left(R\beta-\Omega\right)^{\prime}$$
$$\left[R\left(X^{\prime}\Omega_1^{-1}X\right)^{-1}R^{\prime}\right]^{-1}R\left(X^{\prime}\Omega_1^{-1}X\right)^{-1}+\sigma^{2}n\left(X^{\prime}\Omega_1^{-1}X\right)^{-1} \quad ...(7.10)$$
$$\left[I-R^{\prime}\left[R\left(X^{\prime}\Omega_1^{-1}X\right)^{-1}R^{\prime}\right]^{-1}R\left(X^{\prime}\Omega_1^{-1}X\right)^{-1}\right]\}$$

Defining $\tilde{\delta}_n=\sqrt{n}\left(R\beta-r\right)$,

$V_1=n\left(X^{\prime}\Omega^{-1}X\right)^{-1}$ and

$$\tilde{\lambda}_n=\frac{\tilde{\delta}_n^{\prime}\left(RV_1R^{\prime}\right)^{-1}\tilde{\delta}_n}{2\sigma^2} \quad ...(7.11)$$

For all n, the various test criteria for testing different hypotheses are given by:

a) H_0 (i): Rβ = r or $\sqrt{n}\left(R\beta-r\right)=0$

H_0 is true iff $\tilde{\lambda}_n=0$ ...(7.12)

b) H_0 (ii): The restricted GLS estimator $\tilde{\beta}$ is better than unrestricted GLS estimator β^*.

H_0 is true if there exist non-null q×1 vector c, such that

$$C^{\prime}\left[\mathrm{MSE}\left(\sqrt{n}\beta^{*}\right)-\mathrm{MSE}\left(\sqrt{n}\tilde{\beta}\right)\right]C \equiv C^{\prime}[V_1R^{\prime}\left(RV_1R^{\prime}\right)^{-1}\left(\sigma_1^2RV_1R^{\prime}-\tilde{\delta}_n\tilde{\delta}_n^{\prime}\right)$$
$$\left(RV_1R^{\prime}\right)^{-1}RV_1]C \geq 0 \quad ...(7.13)$$

This holds iff $\tilde{\lambda}_m \leq \frac{1}{2}$

In other words H_0 (ii) is true iff $\tilde{\lambda}_m \le \frac{1}{2}$ ...(7.14)

This criterion in the strong MSE criterion.

c) H_0 (iii): The restricted GLS estimator $\tilde{\beta}$ is better than unrestricted GLS estimator β^*.

The first weak MSE criterion for testing H_0 (iii) is given by

$$tr\left[\text{MSE}\left(\sqrt{n}\beta^*\right) - \text{MSE}\left(\sqrt{n}\tilde{\beta}\right)\right]$$
$$\equiv tr\left[\sigma_1^2 V_1 R^| \left(RV_1R^|\right)^{-1} RV_1 - V_1R^|\left(RV_1R^|\right)^{-1}\tilde{\delta}_n\tilde{\delta}_n^|\left(RV_1R^|\right)^{-1}RV_1\right] \ge 0 \quad ...(7.15)$$

This holds iff

$$\tilde{\lambda}_m \le \frac{1}{2}\mu_1 tr\left[V_1R^|\left(RV_1R^|\right)^{-1}RV_1\right] = \Delta_n\,(\text{say}) \quad ...(7.16)$$

and $\sigma_1^2 = \dfrac{\left(Y - X\tilde{\beta}\right)^| \Omega_1^{-1}\left(Y - X\tilde{\beta}\right)}{n}$

d) H_0 (iv): The restricted GLS estimator $\tilde{\beta}$ is better than unrestricted GLS estimator β^*.

The second weak MSE criterion for testing H_0 (iv) is given by

$$E\left[n\left(\beta^* - \beta\right)^| V_1^{-1}\left(\beta^* - \beta\right) - n\left(\tilde{\beta} - \beta\right)^| V_1^{-1}\left(\tilde{\beta} - \beta\right)\right] \ge 0 \quad ...(7.17)$$

This holds iff

$$\tilde{\lambda}_n \le \frac{q}{2} \quad ...(7.18)$$

e) H_0 (v): $R\beta - r = 0$

The test statistic for testing H_0 (v) is given by

$$F^* = \frac{\left[\left(Y - X\tilde{\beta}\right)^| \Omega_1^{-1}\left(Y - X\tilde{\beta}\right) - \left(Y - X\beta^*\right)^| \Omega_1^{-1}\left(Y - X\beta^*\right)\right] / q}{\left(Y - X\beta^*\right)^| \Omega_1^{-1}\left(Y - X\beta^*\right) / (n-k)} \quad ...(7.19)$$

or $F^* = \dfrac{\left[n\left(R\beta^* - r\right)^|\left(RV_1R^|\right)^{-1}\left(R\beta^* - r\right)\right] / q}{\left(Y - X\beta^*\right)^| \Omega_1^{-1}\left(Y - X\beta^*\right) / (n-k)}$ ...(7.20)

where F^* follows noncentral F-distribution with [q,(n-k)] degrees of freedom and non-centrality parameter is given by

$$\tilde{\lambda}_n \le \frac{1}{2\sigma^2}\delta_n^|\left(RV_1R^|\right)^{-1}\delta_n \quad ...(7.21)$$

Where $\delta_n = \sqrt{n}\left(R\beta - r\right), V = n\left(X^|\Omega_1^{-1}X\right)^{-1}$

H_0 (v) is rejected for too large value of F^*.

In other words

H_0 (v) is rejected if

$$\tilde{\lambda}_n = \frac{1}{2\sigma_1^2}\tilde{\delta}_n^{|}\left(RV_1R^{|}\right)^{-1}\tilde{\delta}_n \geq \frac{1}{2} \qquad ...(7.22)$$

or $\frac{q}{2}$ according to the chosen MSE criterion.

8. Conclusions

The econometric literature has witnessed recently an upsurge of interest in extending inferential procedures to linear regression models with linear equality and inequality constraints. Most of the theoretical and empirical research on testing hypotheses about the parameters of the linear models has been carried out in the context of scalar covariance matrix for disturbances.

A generalized linear model with non-spherical disturbances has been specified with exact linear restrictions about the parameters. Later, the maximum likelihood estimation has applied to estimate unknown covariance matrix of disturbances and then estimated the parameters of linear regression model. Mean Square Error (MSE) criterion has been proposed to test for the better performance of the restricted estimated GLS (EGLS) estimators over that of unrestricted estimators of the regression coefficients.

9. Reference

1. Bera AK. Testing for the Regression Coefficient Stability, Journal of Quantitative Economics. 1997; 13:17-35.
2. Bock ME, Judge GG, Yancey TA. “Some Comments on Estimation in Regression After Preliminary Tests of Significance, 1973; 1:191-200.
3. Farebrother RW. A Survey of Some Recent Econometric Work on Tests for Inequality Constraints. 1988; 4(2):227-238.
4. Firoozi, Fathali. A Comparison of the Procedures for Testing Joint Inequality Hypotheses. 1993; 9(1):73-84.
5. Gourieroux C, Holly A, Monfort A. Likelihood ratio test, Wald test and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters Econometrics. 1982; 50:63-80.
6. Leo, Michelis. The distributions of the J and Cox non-nest tests in regression models with Weakly Correlated Regressors, Journal of Econometrics. 1999; 93:369-401.
7. Ohtani K. On Pooling Disturbance Variances when the Goal is Testing Restrictions on Regression Coefficients, 1987; 35:219-231.
8. Theil H, Goldberger AS. On pure and mixed Statistical estimation in economics, International Economic Review, 1961; 1:65-78.
9. Theil M. On the use of Incomplete Prior Information in Regression Analysis”, Journal of the American Statistical Association, 1963, 401-14.
10. Wallace TD, Toro-Vizcarrondo CE. A test of the mean Square error criterion for restrictions in linear regression, Journal of the American Statistical Association. 1968; 63:558-572.
11. Wallace TD. Weaker Criteria and Tests for Linear Restriction in Regression, Econometrica. 1972; 40:689-698.
12. Yancey TA, Judge GG, Bock ME. Wallace’s Weak mean Square Error Criterion for Testing Linear Restrictions in Regression A Tighter Bound. 1973; 41:1203-1206.
13.