Sindh. Univ. Res. Jour. (Sci. Sr) Vol: 30 (1) 10-12 (1998)

GENERALIZED LEAST SQUARES ESTIMATION IN LINEAR MODELS.

Raja M. Ilyas

Department of Statistics University of Sindh, Jamshoro.

Key Word: Least Square estimation, linear models.

Introduction

While working with linear models in experimental designs one has to deal frequently with several "nuisance" parameters. This is especially the case with "block" designs where the block effect are usually of no interest. For example, in an experiment with two matched samples (Y_{i1} , Y_{i2}) with expected values (n_i , $n_i + \beta$), i= 1,2,….,m, the only parameter of interest is β. The "natural" way to simplify the problem is to work with a variables $Z_i = Y_{i2} - Y_{i1}$ with expected value β for i=1,2,….. n, in place of the original n-2m variables.

More generally, consider a random vector $Y_1 = (Y_1 , Y_2 , \ldots . Y_n)$ with E (Y)=μ , cov (Y)=σ^2 v. where v is assumed to be known positive-definite matrix, and the linear model.

$$A\mu = X\beta$$

Where A, X are mxn and mxr known matrics of ranks m and r, respectively, $r \leq m \leq n$. the classical Gauss-Markoff linear model considers the case m=n with A=1 (or any matrix). We are interested in estimating and also in testing linear hypotheses concerning β. One way of dealing with this problem is to work directly in terms of Y.

An alternative way is to work in terms of the "natural" random vector Z-AY. Then

$$S^2(\beta)=(Z-X\beta)` (AVA')^{-1} (X-X\beta) \qquad (2)$$

is the generalized sum of squares (and product) of residuals in terms of Z, the normal equations are

$$X` (AVA')^{-1} X\beta = X` (AVA')^{-1}z, \qquad (3)$$

and is the solution is

$$\beta = [X` (AVA')^{-1} X]^{-1} X` (AVA')^{-1}Z \qquad (4)$$

β will be termed the GLS castimator of β Observe that

$$E(\beta)=\beta \; ; \; Cov\,(\beta)=\sigma^2[X` (AVA`)^{-1}X]^{-1} \qquad (5)$$

Now under the model (1), m independent functions of μ are expressed in terms of β ; thus there are still n-m "nuisance" parameters from μ besides r parameters β if we work directly in term of Y. the alternative way in term of Z appears simpler since it deals with only , which contains the parameters of interest.

The equivalence in some sense of these two approaches is discussed in this paper. The main interest behind this paper is to present explicitly some results which are possibly in use by statisticians in the area of linear models, but are not adequately documented in the literature.

Estimation:

The basic result concerning the equivalence of the two methods of estimation is the following:

Theorem:- β minimizes $S^2(\beta)$ if any only if it minimizes (Y-)' V^{-1} subject to the restriction Aμ=Xβ. Also if S^2 is the classical error sum of squares, based on Y under the model (l), then $S^2(\beta)=S^2$

Proof: Let L' = (A',B') be nonsingular such that AVB'=0 if U-LY=[Z/K]; n-Lμ then $(Y-\mu)' V^{-1} (Y-\mu) = (U-n)` (LVL)^{-1} (U-n)$

$$= (Z-x\beta)' (AVA')^{-1} (Z-x\beta) + (K-B\mu)` (BVB')^{-1} (K-B\mu)$$

Where the model Aμ = Xβ. the left hand side will attain (at μ= μ*) its minimum value equal to that of the first term on the right (at β=β), since the second term on the right is non-

negative and can attain the value zero by taking A $\mu^*=x\beta$ and B $\mu^*=K$. this proves the theorem

In the theorem it is not necessary to assume that x is of full rank r. if rank (x)=r` <r, the modification in section 1 would be as in the Gauss-Markoff case. [see Rao, (1973)]. For example the GLS estimate would not be unique and not all a linear functions of β would be estimable. Thus the two procedures to estimate (estimable) linear functions of β lead to identical results. Similarly, to estime σ^2 if we use the Gauss-Markoff estimators.

$$\hat{\sigma}^2 = \frac{1}{m-r'} S^2(\hat{\beta}), \hat{\sigma}^2 = \frac{1}{m-r'} S^2$$

based on Z and Y respectively, then $\hat{\sigma}=\sigma^*$. Thus, estimators of (estimable) linear functions of β and of σ^2 are preserved under the of data from Y to Z=AY even when m<n.

Testing of Hypothesis:

Assume for simplicity that X is of full rank r. if we further assume a normal distribution so that Y is $N_n(\mu, \sigma^2 V)$, it will follow that β is $N_r\{\beta, \sigma^2[x` (AVA`)^{-1} x]^{-1}\}$ and S^2/σ^2 is distributed as a chi-squared variable with m-r degrees of freedom under the model $A\mu = x\beta$. then, if σ is known, the test rejects the hypothesis

$$H_o : A\mu = x\beta$$

if $S^2(\beta) / \sigma^2 > Y_{m-r\ 1-\alpha}$ provides a "natural" goodness-of-fit test, confining the attention only to Z.

Alternatively, the "natural" test of Ho in terms of the original variable Y is to reject Ho if $S^2/\sigma^2 > Y^2_{m-r,\ 1-\alpha}$ Indeed, this test is known to be optimal [see Lehmann (1959)]. In view of the result in the previous section that $S^2=S^2(\beta)$, it is seen that we are led to Z-AY. This result corresponds to the analogous result proved in section 2 for estimation of β without making the normality assumptions. Furthermore, if this fit of the model Ho is acceptable, a test of

$$H_1 : B\beta = b$$

Where B is and sxr known matrix of rank s and b is a given vector, under the model Ho, is available from the statistic.

$$T^2 = (B\beta\text{-}b)` [B\{x`(AVA`)^{-1}x\}^{-1}B`] (B\beta\text{-}b)B$$

which is distributed as $Y^2\sigma^2$ with s degrees o freedom. If σ is known, the test rejecting H1 if $T^2/\sigma^2 >$ $Y2s_{,1-\alpha}$ is optimal while, with σ unknown, the classical F-test rejecting H_1 if

is optimal [see Lehmann (1959)]. Thus, we are again led to optimal tests even if we confine attention to Z-AY. This would be seen to be true even in the general case r`≤ r with obvious modifications.

$$\frac{T^2/S}{S^2/m-r} > F_{s,m-r,1-\beta}$$

Asymptotic Normal Case

Suppose now we have a sequence of n-dimensional random vectors $[Y_N]$, N=1,2,……., such that

i.e.Ys have a limiting normal distribution as N-->^{oo}. Suppose also that V_N is consisitent esimator of the positive definite matrix V.

$$N^{\frac{1}{2}}(Y_N - \mu) \xrightarrow{L} > N_n(0, V)$$

Denote by (β_N), (S^2_N) and (T^2_N) the sequences corresponding to N, of statistics in section 3 with σ=1 and V replaced by V_N. corresponding to the optimality of the Guss-markoff estimators and test statistics S^2 and T^2 in section 3 for testing H_o and H_1, respectively. We have corresponding optimality in the asymptotic sense [wald (1943)] for estimator β_N and test-statistics S^2_N and T_{N2} as N-->oo. As in the earlier sections these statistics can be produced by the GLS technique using Z_n = Ayn. One important area of application of this GLS technique in the asymptotic case on the analysis of contingency tables, or, more. Generally, data arising from one or more multinomial distributions [Grizzle, etat (1961)].

CamScanner

References

1. Bhapkar, V.P (1961) some tests for categorical Date. Annals of Mathematics. Statistics. 32. P 72-83.

2. Grizzle. James. Starmer. Frank. and Kock. (1969) "Analysis of Categorical Data by Linear Models. Biometries. 25 p-489-500.

3. Lehmann. E.L: (1959) Testing Statistical Hypotheses. New York. John Wiley & Sons, Inc.

4. Rao. C.R. (1973) Linear Statistical Inference and its application new York. John Wiley & Sons. Inc

5. Wald. A (1943). Tests of statistical Hypotheses Concerning several parameters when the Number of Qssevations if lage. Transactions of American Math Society. 54. p- 426-82.