Jour. Ind. Soc. Ag. Statistics
46(2), 1994 : 223–230

Estimation of Production Function using Quadratic Programming

S.P. Mukherjee and U. Dasgupta*
Calcutta University
(Received : August, 1991)

Summary

A two-input Cobb-Douglas production function model is considered where it is just natural to require that the marginal productivity coefficients (parameters) are non-negative and that, in some situations, these add up to less than unity. A quadratic programming approach to maximise a suitably defined objective function involving the productivity coefficients as decision variables and their ordinary least square estimates as parameters has been used to derive estimates of the parameters in the restricted space. Performance of these estimates relative to least absolute error, restricted least square, restricted maximum likelihood and ridge estimates has been examined through an example.

Key words : Restricted parameter space, kuhn tucker conditions, J Jack-Knife

Introduction

In the estimation of Cobb-Douglas (CD) production function from agricultural data, an important problem (though quite usual) is to arrive at estimates of marginal productivities which are non-negative. Two different approaches have been followed for estimating CD parameters in the restricted parameter space. In the first approach ordinary least squares and maximum likelihood estimates are accepted if all the estimates come out to be non-negative, while some adjustments are proposed if some of the traditional estimate (s) turns (turn) out to be negative (Dasgupta [2]). In the second approach, estimates have been derived primarily through the constrained minimisation of a suitably chosen objective function in which the parameters appear as the decision variables.

In LAE based estimation we minimise the sum of absolute deviations and get the estimates through a linear programming problem which is adjusted to yield unbiased estimates and is further modified to account for the non-negativity constraints (see Mukherjee & Dasgupta [5]). Ridge Estimation has been extended to the problem of estimating CD parameters by minimising the MSE with respect to one ridge parameter (Mukherjee & Dasgupta [6]).

* St. Xavier's College, Calcutta

The problem incorporating non-negativity constraint has been reformulated as an exercise in quadratic programming.

In the present work, estimates of parameters in a two-input CD-model have been derived by the use of quadratic programming, taking into account non-negativity as well as 'decreasing returns to scale constraints. Some ideas about biases and mean square errors (MSE's) of these estimates have been given in terms of a numerical illustration.

2. *The Model and the Solution Approach*

Consider a two-input Cobb-Douglas model spacified as

$$Y = X_1^{\beta_1} X_2^{\beta_2} U \tag{1}$$

where Y is the output, X_1 and X_2 are inputs and U is the disturbance.

Thus, under log-transformation, model (1) reduces to the two-variable linear model

$$y = \beta_1 x_1 + \beta_2 x_2 + u$$

where $y = \log Y$, $x_i = \log X_i$ $(i = 1, 2)$ and $u = \log U$.

The following assumptions are made.

i. Output and Inputs are measured in terms of their ratios to the respective geometric means.

ii. U has a log-normal distribution.

iii. $\beta_1 \geq 0, \beta_2 \geq 0$.

iv. For the sake of simplicity, further assume that

$$\sigma_u^2 = 1.$$

Incorporate subsequenty the usual restriction of diminishing returns to scale in terms of the inequality

$$\beta_1 + \beta_2 \leq 1.$$

Following Barlow, etal [1] maximum likelihood estimate of (β_1, β_2) in the non-negative quadrant can be obtained through the minimisation of

$$S* = (\hat{\beta}_1 - \beta_1)^2 s_{11} + (\hat{\beta}_2 - \beta_2)^2 S_{22} + 2(\hat{\beta}_1 - \beta_1)(\hat{\beta}_2 - \beta_2) S_{12} \tag{2}$$

$$\beta_1 \geq 0, \beta_2 \geq 0.$$

where $\hat{\beta}_1, \hat{\beta}_2$ are least square estimates of β_1 and β_2 and S_{ij} = sum of squares/products of x_1 and x_2.

3. *Use of Quadratic Programming*

Problem Formulation

Rewrite S* in (2) and reformulate the problem of minimising S* subject to the constrains $\beta_1 \geq 0, \beta_2 \geq 0$ as a quadratic programming problem. Now

$$S* = \sum_{i=1}^{2} \sum_{j=1}^{2} S_{ij} \beta_i \beta_j + \sum_{j=1}^{2} C_j \beta_j + \text{a part not involving } (\beta_1, \beta_2)$$

$$= -2S** + \text{a part not involving } (\beta_1 \beta_2) \quad (3)$$

where $S** = -1/2 (\beta_1, \beta_2) \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} - \sum_{j=1}^{2} C_j \beta_j$

and $C_j = -(S_{j1} \hat{\beta}_1 + S_{j2} \hat{\beta}_2), j = 1, 2.$

Thus the problem reduces to

Maximise S**

subject to $\beta_1 \geq 0, \ \beta_2 \geq 0$.

The objective function S** involves a quardratic and linear part and since the quadratic part is negative definite, both the parts are strictly concave and hence S** is strictly concave. Thus the above maximisation problem either has no feasible solution or has a unique optimal solution. Kuhn-Tucker conditions (Hadley, [4]) provide a necessary and sufficient condition for an optimal solution and we can apply Wolfe's [8] algorithm to solve the corresponding (derived) linear programming problem.

Solution Algorithm

Introducing slack variables S_1^2, S_2^2 the constraints can be written as

$$-\beta_1 + S_1^2 = 0$$

$$-\beta_2 + S_2^2 = 0 \quad (4)$$

Then we construct the Lagrangian function

$$L (\beta_1, \beta_2 ; S_1^2, S_2^2, \lambda_1, \lambda_2)$$

$$= S^{**} - \lambda_1 (-\beta_1 + S_1^2) - \lambda_2 (-\beta_2 + S_2^2) \quad (5)$$

Thus the Kuhn-Tucker conditions can be written as

$$\partial L/\partial \beta_1 = S_{11}\beta_1 + S_{12}\beta_2 - C_1 + \lambda_1 = 0$$

$$\partial L/\partial \beta_2 = S_{12}\beta_1 + S_{22}\beta_2 - C_2 + \lambda_2 = 0$$

$$\partial L/\partial S_1 = -2\lambda_1 S_1 = 0$$

$$\partial L/\partial S_2 = -2\lambda_2 S_2 = 0$$

$$\partial L/\partial \lambda_1 = \beta_1 - S_1^2 = 0$$

$$\partial L/\partial \lambda_2 = \beta_2 - S_2^2 = 0 \quad (6)$$

After some necessary simplifications we get

$$S_{11}\beta_1 + S_{12}\beta_2 + \lambda_1 = C_1$$

$$S_{12}\beta_2 + S_{22}\beta_2 + \lambda_2 = C_2$$

and $$\beta_1\lambda_1 + \beta_2\lambda_2 = 0$$

$$\beta_1, \beta_2, \lambda_1, \lambda_2 \geq 0. \quad (7)$$

Then introducing non-negative artificial variable w_1 and w_2, the problem reduces to

$$\text{minimise } z = (w_1 + w_2)$$

subject to

$$S_{11}\beta_1 + S_{12}\beta_2 + \lambda_1 + w_1 = C_1$$

$$S_{12}\beta_2 + S_{22}\beta_2 + \beta_2 + w_2 = C_2$$

$$\beta_1\lambda_1 + \beta_2\lambda_2 = 0 \quad (8)$$

and $$\beta_1, \beta_2, \lambda_1, \lambda_2, w_1, w_2 \geq 0.$$

Subsequently one can use the two phase simplex method to obtain an optimum solution to the above linear programming problem, the solution satisfying the complementary slackness condition. The optimum solution obtained this way provides an optimal solution to the original quadratic programming problem.

4. *Properties of the Estimates*

Since the restricted estimates of β_1 and β_2 obtained by the use of quadratic programming, as explained in this section, are obtained only numberically from a given sample (and not analytically in terms of smaple observations) the question of estimating the bias and the variance of any estimate does arise. One way to get some idea about such properties would be a simulation exercise. But a better approach to this problem seems to lie in the jackknife and bootstrap methods for estimating bias and variance.

The jacknifed estimate of the bias of an estimate $\hat{\theta} = \hat{\theta}(X_1, X_2, \ldots, X_n)$ where $X_1, X_2, \ldots, X_n$ constitute a simple random sample of size n from an unknown probability distribution F is given by (Quenouille, [7]

$$\text{Est. Bias} = (n-1)(\hat{\theta}_{(.)} - \hat{\theta})$$

where $\hat{\theta}_{(.)} = 1/n \, \Sigma \, \hat{\theta}_{(i)}$ and $\hat{\theta}_{(i)}$ = estimate of θ from the sample deleting the ith sample observation. Then a bias-corrected estimate of θ can be obtained as

$$\hat{\theta} = \hat{\theta} - \text{Bias}$$
$$= n\,\hat{\theta} - (n-1)\hat{\theta}_{(.)}$$

Further with this reasampling process, Jackknife estimate of variance of θ will be

$$\sigma_j^2 = \frac{(n-1)}{n} \sum [\hat{\theta}_{(i)} - \hat{\theta}_{(.)}]^2 \tag{11}$$

The bootstrap generalise (9) in an apparently different way. Let $(X_1^*, X_2^*, \ldots, X_n^*)$ be a random sample drawn with replacement from the given set of sample observations. Denoting the estimate of θ based on $(X_1^*, X_2^*, \ldots, X_n^*)$ as $\theta^* = \theta(X_1^*, X_2^* \ldots, X_n^*)$, the bootstrap estimate of the bias and variance of θ^* is based on θ^* (Efron [3]).

Standard computer programs for the Wolfe's method being reasonably accessible, it is quite possible to estimate the sampling vriance in these two ways, given a set of i.i.d sample observations.

5. *An Illustrative Example*

The Data

Assume that in the model

$$Y = \beta_1 x_1 + \beta_2 x_2 + u$$

i. $\beta_1 = 0.05, \beta_2 = 0.75$

ii. u ~ N (0, 0.1)

iii. Combinations of (x_1, x_2) values are as specified in the first two columns of Table 1.

A single randomly selected value of u from the N (0, 1) population will generate a set of observations on $(x_1, x_2\ y)$. Several u values were chosen randomly (and corresponding 8 sets of values were obtained) in such a way that the following results were found. A little trial and error was involved in this exercise.

Table 1. Generated values of Y given x_1 and x_2

x_1	x_2	Y
−1	−1	−0.8305
+1	+1	0.7679
−1	−1	−0.6100
+1	+1	0.7222
−1	−0.8	−0.5883
+1	+0.8	0.5070
0.8	1	0.7807
−0.8	−1	−0.6922

Results

For quadratic programming approach, maximisation of the objective function S^{**} in (3) was achieved through a relatively simple grid method by trying out values of $\beta_1, \beta_2 \geq 0$.

Estimates obtained by applying four estimation schemes (excluding the LAE scheme) are presented in Table 2. For the sake of comparability, estimates of bias and MSE in each case are calculated by applying Jackknife technique described in section 4).

It should be pointed out, however, that in the first two schemes since the LSE of β_1 came out to be negative, the proposed estimate was taken as zero, the Jackknife estimate in each repetition was zero and the estimation of bias and MSE becomes meaningless.

Comment

From the results presented in Table 2, it is observed that estimates obtained through different schemes do not differ much among themselves. However, biases of these estimates as well as their MSE's exhibit significant differences. These results are purely illustrative, rather than confirmatory and that no general conclusions regarding the relative efficiencies of the methods can be decided on the basis of these limited results.

Table 2. Estimastes Obtained Under Different Schemes Together with Bias and MSE

	β_1			β_2		
Schemes	Estimate	Bias	MSE	Estimate	Bias	MSE
I. Restrict least	0	*	*	.923	–.0416	.0135
II. Restricted maximum Likelihood	0	*	*	.7335	.0028	.0010
III. Using Quadratic Programming	.0110	.0049	.0008	.7110	.0024	.0003
IV. Ridge	.1391	.0051	.0036	.7997	.0032	.0054

6. *Case of Further Restrictions*

In Indian agricultural context, we generally come across decreasing return to scale. Thus in a two-input model (1), the restrictions to be considered while estimating the parameters can be written as

$$\beta_1 \geq 0, \beta_2 \geq 0, \beta_1 + \beta_2 \leq 1 \qquad (12)$$

Taking account of a new restriction on parameters will be equivalent to including a new constraint in the problem formulation and subsequently adjusting the solution algorithm.

The set of equations in (4) will now include the additional equation

$$(\beta_1 + \beta_2) - 1 + S_3^2 = 0$$

and the Lagrangian function will take the form

$$L(\beta_1, \beta_2, S_1^2, S_2^2, S_3^2, \lambda_1, \lambda_2, \lambda_3)$$

$$= S^{**} - \lambda_1(-\beta_1 + S_1^2) + \lambda_2(-\beta_2 + S_2^2) - \lambda_3(\beta_1 + \beta_2 - 1 + S_3^2)$$

The equivalent Kuhn-Tucker condition can be deduced as

$$S_{11}\beta_1 + S_{12}\beta_2 + \lambda_1 - \lambda_3 = C_1$$

$$S_{12}\beta_1 + S_{22}\beta_2 + \lambda_2 - \lambda_3 = C_2$$

$$\beta_1\lambda_1 + \beta_2\lambda_2 = 0$$

and

$$\lambda_3(\beta_1 + \beta_2 - 1) = 0$$

$$\beta_1, \beta_2, \lambda_1, \lambda_3 \geq 0.$$

Then the reformulated problem will be

$$\text{Minimize } Z = (w_1 + w_2)$$

subject to

$$S_{11}\beta_1 + S_{12}\beta_2 + \lambda_1 - \lambda_3 + w_1 = C_1$$

$$S_{12}\beta_1 + S_{22}\beta_2 + \lambda_2 - \lambda_3 + w_2 = C_2$$

$$\beta_1\lambda_1 + \beta_2\lambda_2 = 0.$$

$$\lambda_3(\beta_1 + \beta_2 - 1) = 0.$$

Again two phase simplex method can give an optimal solution.

REFERENCES

[1] Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D., 1972. Statistical inference Under Order Restrictions, John Wiley and Sons, London.

[2] Dasgupta, U., 1980. Estimation of Cobb-Douglas production function with non-negative marginal productivities, IAPQR *Transactions*, **5**, 39–47.

[3] Efron, B., 1982. The Jacknife, the Bootstrap and Other Resampling Plans, Philadelphia Society of Industrial and Applied Mathematics.

[4] Hadley, G., 1964. Non-linear and Dynamic programming, Addison-Wesley Publishing Company, Inc.

[5] Mukherjee, S.P. and Dasgupta, U., 1986. Unbiasedness and least absolute error estimation, *Agri. Stat.*, **38**, 218–222.

[6] Mukherjee, S.P. and Dasgupta, U., 1987. Ridge Estimation of Restricted Parameters in Production Techniques, *Comm. in Stat., Theory and Methods*, **16**, 3439–3446.

[7] Quenouille, M., 1956. Notes on Bias in Estimation, *Biometrika*, **43**, 353–360.

[8] Wolfe, P., 1959. The simplex method for quadratic programming, *Econometrica*, **27**, 382–398.