Discussion Paper #538

IMPOSING CURVATURE RESTRICTIONS ON FLEXIBLE FUNCTIONAL FORMS

by

A. Ronald Gallant* and Gene H. Golub**

November 1982

*A. Ronald Gallant, North Carolina State University, Raleigh, North Carolina 27650 USA, is currently on leave at the Department of Economics and the J. L. Kellogg Graduate School of Management, Northwestern University, Evanston, Illinois 60201, USA.

**Gene H. Golub, Stanford University, Stanford, California 94305 USA

Thanks are due Margaret H. White for helpful discussions. This research was supported by North Carolina Agricultural Experiment Station Project NC03641 and by National Science Foundation Grant SES8014239.

Imposing Curvature Restrictions on
Flexible Functional Forms

A. Ronald Gallant and Gene H. Golub

ABSTRACT

A general computational method for estimating the parameters of a flexible functional form subject to convexity, quasi-convexity, concavity, or quasi-concavity at a point, at several points, or over a region are set forth and illustrated with an example.

IMPOSING CURVATURE RESTRICTIONS ON FLEXIBLE FUNCTIONAL FORMS

by

A. Ronald Gallant and Gene H. Golub

1. Introduction

The generality of parametric statistical inference is inevitably limited by model induced augmenting hypotheses. The validity of an inference requires that the true data generating model belongs to the assumed, parametric family of models. In hopes of increasing the generality of implications, richer parametric families of models called flexible functional forms are finding increasing use in empirical economic research. The idea is to assign to a consumer, to a firm, to an industry (or to whomever), a utility function, an indirect utility function, a cost function, a profit function (or whatever), that is parametrically rich enough so as not to impose extraneous behavioral restrictions that limit the generality of an inference. As instances one has the Translog, generalized Leontief, Box-Cox, Almost Ideal, Laurent, and Fourier flexible forms. Of these, only the Fourier flexible form has been shown to be Sobelev-flexible. A form that is Sobelev-flexible asymptotically removes all model induced augmenting hypotheses (Gallant, 1982).

Unfortunately, all of these forms are too flexible in the sense that curvature properties that the approximated function are known to possess are not possessed, of necessity, by the approximating flexible function form. For example, a consumer's indirect utility function must be quasi-convex but a Fourier, Box-Cox, Laurent, etc., approximation to it with parameters estimated statistically need not be quasi-convex. As seen later, quasi-convexity is the idealized form of the general curvature problem in that only slight modification is required to use a solution of that problem to impose convexity, concavity, and quasi-concavity. Therefore, we shall restrict discussion to the quasi-convexity constraint and, to avoid vague generalities, we shall specialize to the problem of imposing quasi-convexity on the consumer's indirect utility function.

To varying degrees, depending on the particular form, one can impose quasi-convexity on an ad hoc basis by finding explicit parametric constraints that imply quasi-convexity (see, e.g., Jorgenson, Lau, Stoker, 1981). Usually, it is easier to impose sufficient conditions (Gallant, 1982; Barnett, 1982). But if the conditions are sufficient, not necessary and sufficient, then the flexibility of the form can be lost. Regarding Sobelev-flexibility this is a critical consideration as the desirable asymptotic properties of estimators and test statistics are thereby destroyed. At any rate, attacking the curvature problem form by form is tedious and the parametric conditions derived thereby are often very difficult to implement. Trying alternate forms to look at comparative results for a given data set becomes a prohibitively onerous task. One should prefer a general method of imposing quasi-convexity that works for any flexible functional form and that is easily implemented. To this, we turn our attention.

2. Statement of the Problem

Let $g(x,\theta)$ denote an approximation to the consumer's indirect utility function where $x = (x_{(1)}', x_{(2)}')'$ is a vector with leading elements $x_{(1)}$ corresponding to income normalized prices $x_{(1)} = p/y$ and the remaining elements $x_{(2)}$ of x corresponding to taste and demographic variables. The p-dimensional parameter θ is to be determined statistically, subject to the constraint that $g(x,\theta)$ be quasi-convex in the argument $x_{(1)}$ for all x in some region C. Denote the data space by X and the parameter space by Θ so that $(x,\theta) \varepsilon X \times \Theta$ are the admissable data/parameter value pairs.

To justify the econometric and statistical methodology employed,

$g(x,\theta)$ must be twice continuously differentiable in $x_{(1)}$ and θ over $X \times \Theta$. Typically, explicit formulas are avaiable to compute

$$\nabla g(x,\theta) = (\partial/\partial x_{(1)})g(x,\theta)$$

$$\nabla^2 g(x,\theta) = (\partial^2/\partial x_{(1)}\partial x_{(1)}')g(x,\theta)$$

and to compute $(\partial/\partial\theta_i)\nabla g(x,\theta)$, $(\partial/\partial\theta_i)\nabla^2 g(x,\theta)$, for $i=1,2,\ldots,p$. Under these assumptions, a necessary and sufficient condition for the quasi-convexity of $g(x,\theta)$ in $x_{(1)}$ over C at θ is

$$0 \leqslant \min_{x\varepsilon C} \min_{z} \{z'\nabla^2 g(x,\theta)z\colon\ z'\nabla g(x,\theta) = 0,\quad z'z = 1\}$$

(Lau, 1973). Of the many characterizations of quasi-convexity (Diewert, Avriel and Zang, 1977) this form seems to be most useful in the present context. Define

$$h(x,\theta) = \min_{z} \{z'\nabla^2 g(x,\theta)z\colon\ z'\nabla g(x,\theta) = 0,\quad z'z = 1\}$$

whence $h(x,\theta)$ is negative when the quasi-convexity constraint is violated and zero or positive when satisfied. We shall term $h(x,\theta)$ the constraint indicator.

All statistical estimation procedures that are commonly used in econometric research can be formulted as an optimization problem of the following type (Burquete, Gallant, and Souza, 1982):

$$\hat{\theta} \text{ minimizes } s(\theta) \text{ over } \Theta$$

with $s(\theta)$ twice continuously differentiable in θ. For example, if the Seemingly Unrelated Nonlinear Regressions method is used to estimate θ from share data, $s(\theta)$ has the form

$$s(\theta) = (1/n) \sum_{t=1}^{n} [s_t - f(x_t,\theta)]'\hat{\Sigma}^{-1}[s_t - f(x_t,\theta)]$$

where, from Roy's identity,

$$f(x,\theta) = \operatorname{diag}(x_{(1)})\nabla g(x,\theta)/x'_{(1)}\nabla g(x,\theta)$$

and $\hat{\Sigma}$ is an estimate of the error variance. To impose quasi-convexity, the optimization problem is modified to read as follows:

$$\text{minimize } s(\theta)$$

$$\text{subject to } \min_{x \in C} h(x,\theta) \geq 0$$

The motivations for imposing this inequality contraint are two. The first is the more pressing and it is that reported results and policy recommendations at least appear reasonable. It is common practice to choose some representative point x_0 in the data and report estimated price, income, and substitution elasticities at that point. If $g(x,\hat{\theta})$ is not quasi-convex at the point $x = x_0$ then the reported substitution elasticity matrix will not be negative semi-definite. Reported results, subsequent computations and policy recommendations can appear nonsensical as a result. The situation is corrected by imposing quasi-convexity at the prediction point x_0. In this case C is a singelton set. If predictions are made at more than one point or if the constraint is imposed at every data point in the sample then C is a set with a finite number of elements x_j $j=0,1,2,...,J$. When C is a finite set the optimization problem is of the form

$$\text{minimize} \quad s(\theta)$$

$$\text{subject to} \quad \begin{aligned} &h(x_0,\theta) \geq 0, \\ &h(x_1,\theta) \geq 0, \\ &\vdots \\ &h(x_J,\theta) \geq 0. \end{aligned}$$

We anticipate that it is this version of the problem that will arise most often in application.

The second motivation is to gain statistical efficiency in the estimation of θ. Since it is known a priori that $g(x,\theta)$ must be quasi-convex in $x_{(1)}$ every where over the region X, imposing the constraint

$$\min_{x \in X} h(x,\theta) \geq 0$$

will, under typical regularity conditions, reduce the variance of the estimator $\hat{\theta}$ without affecting consistency so that improved efficiency obtains. Typically X is a rectangle whence the optimization problem becomes

$$\text{minimize} \quad s(\theta)$$

$$\text{subject to} \quad g(\theta) \geq 0$$

where $g(\theta)$ is the solution to

$$\text{minimize} \quad h(x,\theta)$$

$$\text{subject to} \quad a_j \leq x_j \leq b_j \quad j=1,2,\ldots,k.$$

We shall refer to the first problem as the outer minimization problem and the second problem as the inner problem. The inner problem is an unconstrained optimization problem save that the variables are bounded. That function $g(\theta)$ that equals the solution of the inner problem is then used as the constraining function in the outer minimization problem. The outer problem has one inequality constraint.

Whichever form of the problem arises, the efficient, stable computation of

$$h(x,\theta) = \min_z \left\{z'\nabla^2 g(x,\theta)z\colon\ z'\nabla g(x,\theta) = 0,\quad z'z = 1\right\}$$

is the critical issue. As the performance of optimization algorithms is enhanced when analytic derivatives are available, it is helpful to have an explicit formula for $(\partial/\partial\theta)h(x,\theta)$ when C is a finite set and for $(\partial/\partial x)h(x,\theta)$ when C = X. To this we turn our attention.

3. Computation of the Constraint Indicator and Its Derivatives

Let A denote an N by N symmetric matrix that depends on a real valued parameter t. Let $\dot{A}$ denote the matrix with typical element $(d/dt)a_{ij}(t)$ where $a_{ij}(t)$ denotes a typical element of A. Similarly let a denote an N-vector that depends on t and let $\dot{a}$ denote its derivative with respect to t. Put

$$\lambda = \min_z \{z'Az: a'z = 0, \quad z'z = 1\}.$$

We seek to compute λ and its derivative $\dot{\lambda}$.

Let Q be any N by N-1 matrix with $Q'a = 0$ and $Q'Q = I_{N-1}$. Since

$$\{z: a'z = 0, z'z = 1\} = \{z: z = Qw, w'w = 1\}$$

where w is an (N-1) vector we have

$$\lambda = \min_w \{w'Q'AQw: w'w = 1\}$$

$$= \text{smallest eigenvalue of } Q'AQ.$$

Then what is needed is a stable, efficient means to compute $Q'AQ$. Given that, λ can be computed using any standard routine for finding the eigenvalues of a real symmetric matrix. To compute $Q'AQ$ we borrow from the ideas in Golub and Underwood (1970).

A Householder matrix is a real, symmetric, orthogonal matrix of the form (Golub, 1965)

$$H = I + \beta^{-1} uu'$$

where

$$\beta = -u'u/2.$$

The first column of H will be proportional to a if u is chosen as

$$u = \binom{a_1 - \|a\|}{a_{(2)}}$$

where a has been partitioned as $a = (a_1, a'_{(2)})'$ and $\|a\| = (\sum_{i=1}^{N} a_i^2)^{1/2}$. With this choice, the matrix $Q'AQ$ that we seek can be obtained by deleting the first row and column of $K = HAH$. Using a device of Wilkinson (1960), put

$$\alpha = (\beta^{-2}u'Au)$$

$$w = -\beta^{-1}Au$$

$$\nu = \frac{\alpha}{2} u - w$$

whence

$$K = A + u\nu' + \nu u'.$$

Delete the first row and column of K and let K_{22} denote the N-1 by N-1 matrix thereby obtained.

One next computes the smallest eigenvalue λ of K_{22}. Let $\hat{w}$ denote the corresponding eigenvector which has length N-1. As noted above, λ is the solution of

$$\lambda = \min_z \{z'Az: a'z = 0, z'z = 1\}.$$

The optimizing value of z is obtained from $\hat{w}$ by appending a leading zero to

$\hat{w}$ to obtain the N-vector $(0,\hat{w}')'$ and then computing

$$z = H\binom{0}{\hat{w}}$$

$$= \binom{0}{\hat{w}} + \beta^{-1}\left(\sum_{i=2}^{N} u_i \hat{w}_{i-1}\right)u.$$

Let

$$\psi(z,\lambda,\mu) = z'Az - \lambda(z'z - 1) + 2\mu a'z$$

denote the Lagrangian for the problem

$$\min_z \left\{z'Az: \ a'z = 0, \ z'z = 1\right\}.$$

The first order conditions, obtained by differentiating $\psi(z,\lambda,\mu)$ with respect to (z,λ,μ), are:

$$Az - \lambda z + \mu a = 0,$$

$$a'z = 0,$$

$$z'z = 1.$$

Left-multiplying the first equation by z' we have

$$z'Az - \lambda' z'z + \mu z'a = 0$$

whence, using $z'z = 1$ and $z'a = 0$,

$$\lambda = z'Az.$$

We see that the first Lagrange multiplier λ is indeed the solution of the problem

$$\lambda = \min_{z} \{z'Az\colon\ a'z = 0,\ z'z = 1\}$$

which justifies our choice of notation. Left-multiplying by a' we have

$$a'Az + \mu a'a = 0$$

whence

$$\mu = -a'Az/a'a$$

which fact we shall use later. Arranging the first order conditions in matrix form

$$\begin{pmatrix} A - \lambda I & a \\ a' & 0 \end{pmatrix} \begin{pmatrix} z \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and then differentiating with respect to t, we have

$$\begin{pmatrix} \dot{A} - \dot{\lambda} I & \dot{a} \\ \dot{a}' & 0 \end{pmatrix} \begin{pmatrix} z \\ \mu \end{pmatrix} + \begin{pmatrix} A - \lambda I & a \\ a' & 0 \end{pmatrix} \begin{pmatrix} \dot{z} \\ \dot{\mu} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Left-multiplying by (z',u') we have

$$(z',\mu')\begin{pmatrix}\dot{A}-\dot{\lambda}I & \dot{a} \\ \dot{a}' & 0\end{pmatrix}\begin{pmatrix}z \\ \mu\end{pmatrix} = \begin{pmatrix}0 \\ 0\end{pmatrix}$$

whence

$$z'\dot{A}z - \dot{\lambda}z'z + 2\mu\dot{a}'z = 0.$$

Using $\mu = -a'Az/a'a$ and $z'z = 1$ we have

$$\dot{\lambda} = z'[\dot{A} - 2(a'a)^{-1}\, Aa\dot{a}']z.$$

To use these results to compute

$$h(x,\theta) = \min_{z} \{z'\nabla^2 g(x,\theta)z\colon\ z'\nabla g(x,\theta) = 0,\ z'z = 1\}$$

one puts

$$a = \nabla g(x,\theta),$$

$$A = \nabla^2 g(x,\theta).$$

To compute $(\partial/\partial\theta_i)h(x,\theta)$ one puts

$$\dot{a} = (\partial/\partial\theta_i)\ \nabla g(x,\theta)$$

$$\dot{A} = (\partial/\partial\theta_i)\ \nabla^2 g(x,\theta)$$

and to compute $(\partial/\partial x_i)h(x,\theta)$ one puts

$$\dot{a} = (\partial/\partial x_i)\ \nabla g(x,\theta)$$

$$\dot{A} = (\partial/\partial x_i)\ \nabla^2 g(x,\theta).$$

4. An Example: Fitting the Log Fourier Cost Function to KLEM Data

We shall illustrate using data on the U.S. manufacturing sector from 1947 to 1971 from Berndt and Wood (1975) and Berndt and Khaled (1979). To these data, Berndt and Khaled (1979) fit (a factor demand system corresponding to) a generalized Box-Cox cost function. Using a nonhomoethetic, nonneutral technical change specification, they reported an estimated elasticity of substitution matrix at (prices and output prevailing in) the year 1959 that was not negative semi-definite. This implies that the concavity restriction that a cost function must obey in theory was violated by their estimated cost function. To these same data, Gallant (1982) fit a nonhomoethetic Fourier log cost function with the same outcome. Here we shall re-fit the Fourier log cost function subject to the constraint that the cost function be concave at the year 1959.

Total input cost (C), input prices of capital (K), labor (L), energy (E), and materials (M), and the corresponding cost shares are taken from Tables 1 and 2 of Berndt and Wood (1975). The output series (Y) is taken from Table 1 of Berndt and Khaled (1979). Following the protocal set forth in Gallant (1982), these data are transformed as shown in Table 1.

In the notations set forth in Table 1, the Fourier log cost function is written as

Table 1. Data and Scaling Factors

Endogeneous Variables	Exogeneous Variables	Scaling Factors
$s_0 = \ell n(C)$	$x_1 = \ell_1 = \ell n(P_K) - \ell n(.74371) + \varepsilon$	$\varepsilon = 10^{-5}$
s_1 = K cost share	$x_2 = \ell_2 = \ell n(P_L) + \varepsilon$	$\lambda_5 = \dfrac{\ell n(2.76025) + \varepsilon}{\ell n(466.82965/182.82936) + \varepsilon}$
s_2 = L cost share	$x_3 = \ell_3 = \ell n(P_E) + \varepsilon$	$\lambda = \dfrac{6}{\ell n(2.76025) + \varepsilon}$
s_3 = E cost share	$x_4 = \ell_4 = \ell n(P_M) + \varepsilon$	
s_4 = M cost share	$x_5 = v = \lambda_5\, [\ell n(Y) - \ell n(182.82936) + \varepsilon]$	

$$g(x,\theta) = g_K(x,\theta)$$

$$= u_0 + b'x + \tfrac{1}{2}x'Cx + \textstyle\sum_{\alpha=1}^{A}\{u_{0\alpha} + 2\sum_{j=1}^{J}[u_{j\alpha}\cos(j\lambda k_\alpha'x) - v_{j\alpha}\sin(j\lambda k_\alpha'x)]\}$$

$$\nabla g(x,\theta) = (\partial/\partial \ell)g_K(x,\theta)$$

$$= c - \lambda\textstyle\sum_{\alpha=1}^{A}\{u_{0\alpha}\lambda k_\alpha'x + 2\sum_{j=1}^{J} j[u_{j\alpha}\sin(j\lambda k_\alpha'x) + v_{j\alpha}\cos(j\lambda k_\alpha'x)]\}r_\alpha$$

$$\nabla^2 g(x,\theta) = (\partial^2/\partial \ell\partial \ell')g_K(x,\theta)$$

$$= -\lambda^2\textstyle\sum_{\alpha=1}^{A}\{u_{0\alpha} + 2\sum_{j=1}^{J} j^2[u_{j\alpha}\cos(j\lambda k_\alpha'x) - v_{j\alpha}\sin(j\lambda k_\alpha'x)]\}r_\alpha r_\alpha'$$

where

$\{k_a\}$ = a sequence of elementary multi-indexes (Table 2)

$$\theta_{(0)} = b' = (c', b_5)$$

$$\theta_{(\alpha)} = (u_{0\alpha}, u_{1\alpha}, v_{1a}, \ldots, u_{J\alpha}, v_{J\alpha})'$$

$$\theta = (u_0, \theta_{(0)}', \theta_{(1)}', \ldots, \theta_{(A)}')'$$

and

$$C = -\textstyle\sum_{\alpha=1}^{A} u_{0\alpha}\,\lambda^2 k_\alpha k_\alpha'.$$

The restriction of linear homogeneity is imposed as a maintained hypothesis

$$
R_0' \begin{cases} \sum_{i=1}^{4} b_i = 1, \text{ and} \\ u_{j\alpha} = v_{j\alpha} = 0 \text{ if } \sum_{i=1}^{4} k_{i\alpha} \neq 0. \end{cases}
$$

The set of elementary multi-indexes that satisfy $\sum_{i=1}^{4} k_{i\alpha} = 0$ and have norm $|k_\alpha|^* \leq 3$ are displayed in Table 2. For this set A = 19, and we take J = 1, whence θ is a vector of nominal length 63. The effective number of parameters is 53 due to the following restrictions.

The nonhomoegeneous restriction $\sum_{i=1}^{4} b_i = 1$ reduces the number of effective parameters by one. The remaining restrictions are due to overparameterization of the matrix C. The matrix C is a 5 x 5 symmetric matrix which satisfies five linearly independent homogeneous restrictions $\sum_{j=1}^{4} c_{ij} = 0$ (i = 1,2,3,4,5). Thus C can have at most ten free parameters and in the parameterization $C = -\sum_{\alpha=1}^{19} u_{0\alpha} \lambda^2 k_\alpha k_\alpha'$, ten of the $u_{0\alpha}$ are free parameters and nine must be set to zero. These nine are α = 10,11,12,14,15, 16,17,18,19 and were identified numerically as described in Gallant (1982). To impose these restrictions, let

$$
\underset{63 \times 1}{r} = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ \hline \underset{\sim}{0} \end{array}\right]
\qquad
\underset{63 \times 53}{R} = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \underset{\sim}{0} \\ 0 & -1 & -1 & -1 & \\ \hline & & \underset{\sim}{0} & & R_{22} \end{array}\right]
$$

where R_{22} is obtained by deleting columns 10, 11, 12, 14, 15, 16, 17, 18, 19 from the identity matrix of order 58. Then by writing

Table 2. The Sequence $\{k_\alpha\}_{\alpha=1}^{19}$

	Main Effects							Interactions											
		Input ratios						Ratios x low level of u						Ratios x high level of u					
	u	K/L	K/E	L/E	K/M	L/M	E/M	K/L	K/E	L/E	K/M	L/M	E/M	K/L	K/E	L/E	K/M	L/M	E/M
K	0	1	1	0	1	0	0	1	1	0	1	0	0	1	1	0	1	0	0
L	0	-1	0	1	0	1	0	-1	0	1	0	1	0	-1	0	1	0	1	0
E	0	0	-1	-1	0	0	1	0	-1	-1	0	0	1	0	-1	-1	0	0	1
M	0	0	0	0	-1	-1	-1	0	0	0	-1	-1	-1	0	0	0	-1	-1	-1
u	1	0	0	0	0	0	0	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1
α	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$\|k_\alpha\|^*$	1	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3

$$\theta = \phi(\rho)$$

with

$$\phi(\rho) = r + R\rho$$

one can express the vector θ in terms of 53 independent parameters contained in ρ of length 53.

Assuming additive errors and using Shepard's lemma, the data follow the statistical model

$$s_{0t} = g_K(x_t|\theta) + e_{0t}$$

$$s_{1t} = (\partial/\partial \ell_1)\ g_K(x_t|\theta) + e_{1t}$$

$$s_{2t} = (\partial/\partial \ell_2)\ g_K(x_t|\theta) + e_{2t}$$

$$s_{3t} = (\partial/\partial \ell_3)\ g_K(x_t|\theta) + e_{3t}$$

where the share equation for s_{4t} is discarded due to the restriction that $\sum_{i=1}^{4} s_{it} = 1$. See Gallant (1982) for details.

The model may be written in a vector notation

$$y_t = f(x_t|\theta) + e_t \quad t=1,2,\ldots,25$$

with $y_t = (s_{0t}, s_{1t}, s_{2t}, s_{3t})'$ and similarly for f and e_t where we assume that the errors are independently distributed each with mean zero and variancecovariance matrix $\sum$. As $f(x_t|\theta)$ is linear in the parameters,

$$f(x_t|\theta) = z_t'\theta$$

where z_t' is of order 4 by 63, this is a multivariate linear model and can be fitted using the Seemingly Unrelated Regressions method (Zellner, 1962). The method is as follows:

Let

$$s(\theta,\Sigma) = \frac{1}{n}\sum_{t=1}^{n} (y_t - z_t'\theta)'\sum^{-1}(y_t - z_t'\theta).$$

First compute

$$\bar{\rho} \text{ to minimize } s[\phi(\rho),I]$$

Let

$$\bar{\theta} = \phi(\bar{\rho}) = r + R\bar{\rho}$$

Next, estimate Σ by

$$\hat{\Sigma} = \frac{1}{n}\sum_{t=1}^{n} (y_t - z_t'\bar{\theta})(y_t - z_t'\bar{\theta})'.$$

Finally, compute

$$\hat{\rho} \text{ to minimize } s[\phi(\rho), \hat{\Sigma}].$$

As shown in Gallant (1982), a twice continuously differentiable cost function is concave at a point if and only if its corresponding log cost function satisfies:

R5. Concavity. $\nabla^2 g + \nabla g \nabla' g - \mathrm{diag}(\nabla g)$ is a negative semi-definite matrix of rank N-1 with 1 being the eigenvector of root zero.

Then at the value x^* of x obtaining in the year 1959, the constraint to be imposed on the Fourier log cost function is

R_5'. Concavity. $\nabla^2 g[x^*,\phi(\rho)] + \nabla g[x^*,\phi(\rho)]\nabla' g[x^*,\phi(\rho)] - \mathrm{diag}\{\nabla g[x^*,\phi(\rho)]\}$ is a negative semi-definite matrix of rank N-1 with 1 being the eigenvector of root zero.

Using the methods of the previous sections, we propose to impose concavity at the year 1959 on the Fourier log cost function by modifying the last step of the Seemingly Unrelated Regressions estimator to read: Compute

$$\tilde{\rho} \text{ to miminize } s[\phi(\rho),\hat{\Sigma}] \text{ subject to } R_5'.$$

To do this, let

$$-A(\rho) = \nabla^2 g[x^*,\phi(\rho)] + \nabla g[x^*,\phi(\rho)]\nabla' g[x^*,\phi(\rho)] - \mathrm{diag}\{\nabla g[x^*,\phi(\rho)]\}$$

$$a = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

and set

$$h(x^*,\rho) = \min_z \{z'A(\rho)z\colon\ a'z = 0,\ z'z = 1\}.$$

The problem becomes

$$\text{minimize } s[\phi(\rho),\hat{\Sigma}]$$

$$\text{subject to: } h(x^*,\rho) \geq 0$$

which can be solved directly using the methods of the previous sections provided that

$$\dot{A} = (\partial/\partial\rho_k)A(\rho)$$

can be easily computed; note $\dot{a} \equiv 0$. We turn our attention to this task.

A first partial derivative in x of the Fourier log cost function evaluated at the year 1959 is a linear function of the form

$$(\partial/\partial x_i)g_K(x^*,\theta) = g_i'\theta$$

as seen by inspecting the formula for $\nabla g(x,\theta)$ set forth above. Similarly, a second order partial derivative is a linear function of the form

$$(\partial^2/\partial x_i\partial x_j)g_K(x^*,\theta) = h_{ij}'\theta$$

where g_i, h_{ij}, and θ are vectors of length 63. Then

$$(\partial/\partial x_i)g_K[x^*,\phi(\rho)] = g_i'r + g_i'R\rho,$$

$$(\partial^2/\partial x_i \partial x_j)g_K[x^*,\phi(\rho)] = h_{ij}'r + h_{ij}'R\rho.$$

A diagonal element $a_{ii}(\rho)$ of $A(\rho)$ is computed as

$$-a_{ii}(\rho) = (h_{ii}'r + h_{ii}'R\rho) + (g_i'r + g_i'R\rho)^2 - (g_i'r + g_i'R\rho)$$

whence, at sight,

$$-(\partial/\partial\rho')a_{ii}(\rho) = h_{ii}'R + 2(g_i'r + g_i'R\rho)g_i'R - g_i'R.$$

The desired $(\partial/\partial\rho_k)a_{ii}(\rho)$ is the k-th element of $(\partial/\partial\rho')a_{ii}(\rho)$. An off-diagonal element $a_{ij}(\rho)$ with $i \neq j$ is computed as

$$-a_{ij}(\rho) = (h_{ij}'r + h_{ij}'R\rho) + (g_i'r + g_i'R\rho)(g_j'r + g_j'R\rho)$$

whence, again at sight

$$-(\partial/\partial\rho')a_{ij}(\rho) = h_{ij}'R + (g_i'r + g_i'R\rho)g_j'R + (g_j'r + g_j'R\rho)g_i'R.$$

Using SUBROUTINE SALQDR of the NPL Library (NPL, 1980)--a quasi-Newton method (Gill, Murray and Wright, 1981)--to solve

$$\text{minimize } s[\phi(\rho),\hat{\Sigma}]$$

$$\text{subject to: } h(x^*,\rho) \geqslant 0,$$

and using SUBROUTINE EIGRS of the IMSL library (IMSL, 1981) for eigenvector/eigenvalue determination we obtain the solution $\tilde{\theta} = \phi(\tilde{\rho})$ reported in Table 3. Shown also is the Seemingly Unrelated Regressions estimate $\hat{\theta} = \phi(\hat{\rho})$ as computed in Gallant (1982). In both instances $\hat{\Sigma}$ is that obtained from residuals from $\bar{\rho}$ minimizing $s[\phi(\rho),I]$ as described earlier so that the values of $s[\phi(\hat{\rho}),\hat{\Sigma}]$ and $s[\phi(\tilde{\rho}),\hat{\Sigma}]$ shown in the last line of the table are strictly comparable.

Using formulae set forth in Gallant (1982) the Allen partial elasticities of substitution and price elasticities at the year 1959 were computed from the estimates shown in Table 3 and are reported in Table 4.

A readily available source of high quality software for inequality constrained optimization is NAg Libraries, 1250 Grace Court, Downers Grove, Illinois 60516, USA; eigenvector/eigenvalue routines are in the NAg library also. A FORTRAN subroutine to compute λ and $\dot{\lambda}$ given A, a, $\dot{A}$ and $\dot{a}$ is available from A. R. Gallant at the cost of reproduction and postage. This offer expires two years from the publication date.

Table 3. Unconstrained and Concavity Constrained Estimates of the Log Fourier Cost Function.

Parameter	Estimates: Unconstrained	Estimates: Constrained
θ_1	5.15225840	5.15162315
θ_2	0.04998733	0.04711753
θ_3	0.25741658	0.26905123
θ_4	0.15489132	0.15903904
θ_5	0.53770477	0.52479220
θ_6	-0.39950191	-0.40411924
θ_7	-0.01540615	-0.01507702
θ_8	-0.01251629	-0.01186372
θ_9	-0.00673627	-0.00672459
θ_{10}	0.00483888	0.00471898
θ_{11}	0.00146290	0.00134721
θ_{12}	-0.00027147	-0.00006238
θ_{13}	-0.00782715	0.00055413
θ_{14}	0.00308733	-0.00091758
θ_{15}	0.00540352	0.00657224
θ_{16}	0.00034498	0.00028351
θ_{17}	-0.00034014	-0.00027510
θ_{18}	-0.00059587	-0.00055996
θ_{19}	-0.00172445	-0.00130695
θ_{20}	-0.00206864	-0.00192279
θ_{21}	0.00458804	0.00454106
θ_{22}	-0.00108728	-0.00167624
θ_{23}	0.00062251	0.00067070
θ_{24}	-0.00008305	-0.00034696
θ_{25}	0.00179617	0.00197237
θ_{26}	-0.00154407	-0.00149405
θ_{27}	0.00041365	0.00050895
θ_{28}	0.00373929	0.00360038
θ_{29}	-0.00085219	-0.00097934
θ_{30}	-0.00037345	-0.00029840
θ_{31}	0.00385884	0.00295936
θ_{32}	-0.00067486	-0.00052931
θ_{33}	-0.00475092	-0.00405092
θ_{34}	0.0	0.0
θ_{35}	0.00056657	0.00028994
θ_{36}	-0.00054179	-0.00038718
θ_{37}	0.0	0.0
θ_{38}	0.00002994	0.00007476
θ_{39}	0.00056254	0.00043954
θ_{40}	0.0	0.0
θ_{41}	0.00108936	0.00094977
θ_{42}	0.00182818	0.00170208
θ_{43}	-0.00018134	-0.00018251
θ_{44}	-0.00004127	-0.00006491
θ_{45}	-0.00062405	-0.00059148
θ_{46}	0.0	0.0
θ_{47}	-0.00029692	-0.00029107
θ_{48}	0.00013566	0.00032772
θ_{49}	0.0	0.0
θ_{50}	0.00007749	0.00009056
θ_{51}	0.00013755	0.00013684
θ_{52}	0.0	0.0
θ_{53}	-0.00026505	-0.00029745
θ_{54}	-0.00043343	-0.00036084
θ_{55}	0.0	0.0
θ_{56}	-0.00045413	-0.00049713
θ_{57}	-0.00002568	0.00003890
θ_{58}	0.0	0.0
θ_{59}	-0.00006449	0.00000642
θ_{60}	-0.00000549	-0.00006614
θ_{61}	0.0	0.0
θ_{62}	0.00091222	0.00092219
θ_{63}	0.00043126	0.00034011
$S(\theta,\hat{\Sigma})$	63.61102751	63.71693345

Table 4. Fourier Flexible Form Estimates of Allen Partial Elasticities of Substitution and Price Elasticities, U.S. manufacturing, 1959.

Elasticity	Estimates: Unconstrained	Estimates: Constrained
σ_{KK}	- 6.5321	- 6.0083
σ_{KL}	.3288	.4890
σ_{KE}	.6613	1.1704
σ_{KM}	.4545	.2935
σ_{LL}	- .2813	- .4163
σ_{LE}	4.5678	3.6353
σ_{LM}	- .2422	- .1289
σ_{EE}	- 28.5133	- 37.1911
σ_{EM}	- .0157	.9837
θ_{MM}	.0642	- .0436
η_{KK}	- .4013	- .3691
η_{KL}	.0908	.0300
η_{KE}	.0300	.0719
η_{KM}	.2806	.0180
η_{LK}	.0202	.1350
η_{LL}	- .0776	- .1149
η_{LE}	.2069	1.0037
η_{LM}	- .1495	- .0356
η_{EK}	.0406	.0530
η_{EL}	1.2608	.1645
η_{EE}	- 1.2917	- 1.6830
η_{EM}	-.0097	.0445
η_{MK}	.0279	.1811
η_{ML}	- .0668	- .0796
η_{ME}	- .0007	.6071
η_{MM}	.0396	- .0269

REFERENCES

Barnett, William A. (1982), "New Indices of Money Supply and the Flexible Laurent Demand System," Journal of Business and Economics, forthcoming.

Berndt, Ernst R. and David O. Wood (1975), "Technology Prices and the Derived Demand for Energy," Review of Economics and Statistics 57, 259-268.

Berndt, Ernst R. and Mohammed A. Khaled (1979), "Parametric Productivity Measurement and Choice Among Flexible Functional Forms," Journal of Political Economy 87, 1220-1245.

Burquete, Jose F., A. Ronald Gallant, and Geraldo Souza (1982), "On the Unification of the Asymptotic Theory of Nonlinear Econometric Models," Econometric Reviews, forthcoming.

Diewert, W. E., M. Avriel, and I. Zang (1977), "Nine Kinds of Quasiconcavity and Concavity," discussion paper 77-31. (University of British Columbia, Vancouver, Canada.)

Gallant, A. Ronald (1981), "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form," Journal of Econometrics 15, 211-245.

Gallant, A. Ronald (1982), "Unbiased Determination of Production Technologies," Journal of Econometrics, forthcoming.

Gill, Philip E., Walter Murray, and Margaret H. White (1981), Practical Optimization. New York: Academic Press, Inc.

Golub, Gene H. (1965), "Numerical Methods for Solving Linear Least Squares Problems," Numerische Mathematik 7, 206-216.

Golub, Gene H. and Richard Underwood (1970), "Stationary Values of the Ratio of Quadratic Forms Subject to Linear Constraints," Z. angew. Math. Phys. 21, 318-326.

International Mathematical and Statistical Libraries, Inc. (1981). IMSL Library, Ninth Edition. Houston: International Mathematical and Statistical Libraries, Inc.

Jorgenson, Dale W., Lawrence J. Lau, and Thomas M. Stoker (1982), "The Transcandental Logarithmic Model of Aggregate Consumer Behavior," in Advances in Econometrics, R. L Basemann and G. Rhodes, eds. Greenwich, Connecticut: J.A.I. Press.

Lau, Lawrence J. (1973), "Testing and Imposing Monotonicity, Convexity and Quasi-Convexity," Technical Report No. 123. (Institute of Mathematical Studies in the Social Sciences, Stanford University, Stanford, California.)

National Physical Laboratoy (1980), NPL Software Library. Teedington, U. K.: National Physical Laboratory.

Wilkinson, J. H. (1960), "Householder's Method for the Solution of the Algebraic Eigenproblem," Computer Journal 3, 23-27.

Zellner, Arnold (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," Journal of The American Statistical Association 57, 348-368.