A Computational Method for the Karush-Kuhn-Tucker Test of Convexity of Univariate Observations and Certain Economic Applications

Sofia A. Georgiadou and Ioannis C. Demetriou

Abstract **- The problem of convexity runs deeply in economic theory. For example, increasing returns or upward slopes (convexity) and diminishing returns or downward slopes (concavity) of certain supply, demand, production and utility relations are often assumed in economics. Quite frequently, however, the observations have lost convexity (or concavity) due to errors of the measuring process. We derive the Karush-Kuhn-Tucker test statistic of convexity, when the convex estimator of the data minimizes the sum of squares of residuals subject to the assumption of non-decreasing returns. Testing convexity is a linear regression problem with linear inequality constraints on the regression coefficients, so generally the work of Gouriéroux, Holly and Monfort (1982) as well as Hartigan (1967) apply. Convex estimation is a highly structured quadratic programming calculation that is solved very efficiently by the Demetriou and Powell (1991) algorithm. Certain applications that test the convexity assumption of real economic data are considered, the results are briefly analyzed and the interpretation capability of the test is demonstrated. Some numerical results illustrate the computation and present the efficacy of the test in small, medium and large data sets. They suggest that the test is suitable when the number of observations is very large.**

Index terms **- Cobb-Douglas, convexity, concavity, data fitting, diminishing return, divided difference, Gini coefficient, infant mortality, least squares, money demand, quadratic programming, statistical test**

I. INTRODUCTION

The problem of convexity runs deeply in economic theory [24]. For example, increasing returns or upward slopes (convexity) and diminishing returns or downward slopes (concavity) of certain supply, demand, production and utility relations are often assumed in economics. Similar situations are familiar in fields, like decision-making [18], behavioral sciences [27], biology [11] etc.

The purpose of this paper is to present a procedure for testing the hypothesis of convexity of a set of univariate observations. This procedure relies on the Karush-Kuhn-Tucker multipliers of the following optimization calculation [7]. Let $\{(x_i,\varphi_i): i=1,2,...,n\}$ be data, where the abscissae $x_i,\ i=1,2,...,n$ are in strictly increasing order, and φ_i is the measurement of an underlying (unknown) function $f(x)$ at x_i contaminated by random error. If $f(x)$ is convex but convexity has been lost due to errors in the measuring process, we seek numbers $\{y_i: i=1,2,...,n\}$ that minimize the objective function

$$\sum_{i=1}^{n}(y_i-\phi_i)^2 \qquad (1)$$

subject to the *convexity* constraints

$$y[x_{i-1},x_i,x_{i+1}]\ge 0\ ,\ i=2,3,...,n-1, \qquad (2)$$

where

$$y[x_{i-1},x_i,x_{i+1}]=\frac{y_{i-1}}{(x_{i-1}-x_i)(x_{i-1}-x_{i+1})}+\frac{y_i}{(x_i-x_{i-1})(x_i-x_{i+1})}+\frac{y_{i+1}}{(x_{i+1}-x_{i-1})(x_{i+1}-x_i)}, \qquad (3)$$

is the i th second divided difference (see, for example, [22]) of the required numbers. We refer to this problem as the *convex estimation* problem and we regard the measurements and the smoothed values as *n*-vectors $\underline{\varphi}$ and $\underline{y}$ respectively. We see that the constraints on $\underline{y}$ are linear and, in order to simplify our notation, we denote the constraint normals with respect to $\underline{y}$ by $\underline{a}_i,\ i=1,2,...,n-2$, so $y[x_i,x_{i+1},x_{i+2}]=\underline{y}^T\underline{a}_i$, for $i=1,2,...,n-2$. Since each divided difference depends on only 3 adjacent components of $\underline{y}$, it follows that the constraints have linearly independent normals. Further, since the second derivative matrix of (1) with respect to $\underline{y}$ is twice the unit matrix, the problem of minimizing (1) subject to (2) is a strictly convex quadratic programming problem that has a unique solution. We refer to this solution as the *convex estimator* of the data.

The calculation of the convex estimator depends on the Karush-Kuhn-Tucker conditions (see, for example, [19]) for the minimization of (1) subject to the constraints (2). They state that $\underline{y}=\underline{y}^*$ is the convex estimator if and only if the constraints (2) are satisfied and there exist nonnegative Karush-Kuhn-Tucker multipliers $\{\lambda_i^*: i\in \mathrm{S}\}$ such that the first order condition

$$\underline{y}^*-\underline{\varphi}=\tfrac{1}{2}\sum_{i\in \mathrm{S}}\lambda_i^* a_i \qquad (4)$$

This work was supported in part by the University of Athens under research grant ELKE 70/4/4729. Professor I. C. Demetriou is with the Department of Economics, University of Athens, 8 Pesmazoglou Str, Athens 10559, Greece (e-mail: demetri@ econ.uoa.gr). S. A. Georgiadou is a PhD student in the same department (e-mail: georgiad@ econ.uoa.gr).

(Advance online publication: 19 February 2008)

holds, where S is the subset $\{i: y^*[x_{i-1},x_i,x_{i+1}]=0\}$. We define $\lambda_i=0$, for all integers $i\in[2,n-1]\setminus \mathrm{S}$. Thus, $\underline{\lambda}$ is a $(n-2)$-vector.

Several general quadratic programming algorithms are available [10], but for the convex estimator the algorithm of [7] takes account of the constraint structure, least squares, *B*-splines (see, for example, [26]) and active set methodology providing a very efficient calculation. Because sometimes it would be better to employ non-positive instead of nonnegative divided differences, this algorithm may well be applied after a change of sign of the components of $\underline{\varphi}$, thus providing a *concave estimator* to $\underline{\varphi}$ by minimizing (1) subject to the (concavity) constraints

$$y[x_{i-1},x_i,x_{i+1}]\le 0,\ i=2,3,...,n-1. \quad (5)$$

It might help to state an alternative form of the constraints (2), which is given by [17]. Specifically, since the i th second divided difference is expressed in a form involving the difference of two consecutive first divided differences (see, for example, [3])

$$y\left[x_{i-1},x_i,x_{i+1}\right]=\frac{1}{x_{i+1}-x_{i-1}}\left[\frac{y_{i+1}-y_i}{x_{i+1}-x_i}-\frac{y_i-y_{i-1}}{x_i-x_{i-1}}\right], \quad (6)$$

the constraints (2) imply the inequalities (increasing rates of change)

$$\frac{y_{i+1}-y_i}{x_{i+1}-x_i}\ge\frac{y_i-y_{i-1}}{x_i-x_{i-1}},\ i=2,3,...,n-1, \quad (7)$$

and similarly the constraints (5) imply the inequalities (decreasing rates of change)

$$\frac{y_{i+1}-y_i}{x_{i+1}-x_i}\le\frac{y_i-y_{i-1}}{x_i-x_{i-1}},\ i=2,3,...,n-1. \quad (8)$$

Therefore, the constraints of our problem now are that we require increasing rates of change in the convexity case and decreasing rates of change in the concavity case. For example, conditions (8) might be derived from the assumption of diminishing marginal productivity of inputs in several production relations [24]. Also, conditions (7) might be derived from the assumption of increasing marginal utility of a utility curve in decision-making [18]. Moreover, these conditions are able to describe a variety of underlying functions without relaying on parametric expressions.

The paper is organized as follows. In Section II we specify the type of a test statistic for convexity. Specifically, we test linearity against convexity, which requires the calculation of the Lagrange and the Karush-Kuhn-Tucker multipliers of the corresponding constrained optimization problems. In Section III and IV we give the numerical method for the linearity and convexity case calculation. In Section V we give the numerical performance of the test. In Section VI we present applications of the test for the assumption of convexity on certain economic data and reveal important underlying properties. Finally, some concluding remarks are presented in Section VII. The Fortran program that implements the convexity test consists of about 1800 lines including comments, which gives an idea of the size of the required calculation.

The subject of convex estimation has raised some interest over the past years because of its applications to economics, statistics and engineering [5], [6], [7], [8], [14], [15], [17], [23], [25]. In [15] especially the consistency of the convex estimator has been proved, thus providing a good reason for using it in statistical and economic analyses. The work of [1], [9], [21], [28] and [16] may be further inspiring for alternative test statistics and their distribution for the convexity problem, but it is beyond the scope of this paper.

II. THE TEST FOR CONVEXITY

In view of the above data, we wish to develop a procedure for testing the null hypothesis of linearity of the data

$$H_0:\ \varphi[x_{i-1},x_i,x_{i+1}]=0,\ i=2,3,...,n-1, \quad (9)$$

against the alternative hypothesis of convexity of the data

$$H_1:\ \varphi[x_{i-1},x_i,x_{i+1}]\ge 0,\ i=2,3,...,n-1. \quad (10)$$

Under H_0, the estimator of $\underline{\varphi}$, say it is $\underline{\tilde{y}}$, is obtained by minimizing (1) subject to the equality constraints

$$y[x_{i-1},x_i,x_{i+1}]=0,\ i=2,3,...,n-1. \quad (11)$$

Under H_1, the estimator of $\underline{\varphi}$ is $\underline{y}^*$. By using matrix notation, the constraints for the null hypothesis estimate may be expressed as $A^T\underline{y}=\underline{0}$ and for the alternative as $A^T\underline{y}\ge\underline{0}$, where A is the $n\times(n-2)$ matrix whose columns are the constraint normals $\{\underline{a}_i: i=1,2,...,n-2\}$.

The null estimate is on the boundary of the constraint region of the alternative estimate, because it satisfies immediately as equalities all the constraints required by the latter estimate. Therefore, the null hypothesis requires an estimate to the data that cannot be better than the convex estimator, because the latter one allows some inequalities that are amply satisfied.

Clearly, $\underline{\tilde{y}}$ is the unique solution to the corresponding equality constrained problem defined by H_0, because equations $A^T\underline{y}=\underline{0}$ are consistent and the constraint normals are linearly independent. We see in section III that the null hypothesis estimate is the line fit to the data. Moreover, if $\underline{\tilde{\lambda}}$ is the vector of Lagrange multipliers associated with $\underline{\tilde{y}}$, then it satisfies the first order conditions

$$2(\underline{\tilde{y}}-\underline{\varphi})=A\underline{\tilde{\lambda}}. \quad (12)$$

An efficient method to obtain the Lagrange multipliers from (12) is provided by [7] and will be presented in Section III.

In order to test the null hypothesis, we follow [13] and define the Karush-Kuhn-Tucker test statistic ξ_{KKT-LM}, namely

$$\xi_{KKT-LM} = \frac{1}{4\sigma^2}(\underline{\lambda}^* - \underline{\tilde{\lambda}})^T A^T A(\underline{\lambda}^* - \underline{\tilde{\lambda}}) , \tag{13}$$

as the least value of the objective function

$$\frac{1}{4\sigma^2}\left(\underline{\lambda} - \underline{\tilde{\lambda}}\right)^T A^T A\left(\underline{\lambda} - \underline{\tilde{\lambda}}\right), \tag{14}$$

subject to the conditions $\lambda_i \geq 0, i = 2,3,...,n-1$. In practice, however, we make use of an expression of ξ_{KKT-LM} that depends on the following lemma.

Lemma 1 Let $\mathrm{S} \subset \{2,3,...,n-1\}$, let the function

$$\left(\underline{\varphi} + \tfrac{1}{2}\sum_{i=2}^{n-1}\lambda_i \underline{a}_i\right)^T \left(\underline{\varphi} + \tfrac{1}{2}\sum_{i=2}^{n-1}\lambda_i \underline{a}_i\right), \tag{15}$$

and let $\underline{\lambda}^*$ be a value of $\underline{\lambda}$ that minimizes expression (14) subject to the constraints. Then $\underline{\lambda}^*$ is unique and the vector

$$\underline{y}^* = \underline{\varphi} + \tfrac{1}{2}\sum_{i \in \mathrm{S}} \lambda_i^* \underline{a}_i , \tag{16}$$

minimizes the function (1) subject to the constraints $y[x_{i-1}, x_i, x_{i+1}] = 0$, $i \in \mathrm{S}$.

Proof: The proof is based on [7].

In view of this lemma and (12) we let $|\cdot|$ denote cardinality and express (13) in the form

$$\xi_{KKT-LM} = \frac{\left((\underline{y}^* - \underline{\varphi}) - (\underline{\tilde{y}} - \underline{\varphi})\right)^T \left((\underline{y}^* - \underline{\varphi}) - (\underline{\tilde{y}} - \underline{\varphi})\right)}{\sigma^2 (n - 2 - |\mathrm{S}|)}, \tag{17}$$

which has certain computational advantages that are directly relevant to the methods of calculation of $\underline{y}^*$ and $\underline{\tilde{y}}$.

If σ^2 is known, then ξ_{KKT-LM} follows the asymptotic chi squared distribution with $n-2-|\mathrm{S}|$ degrees of freedom [28]. If, however, σ^2 is unknown then we replace σ^2 in (17) with the standard unbiased estimate, s^2 say, whose value is $s^2 = (\underline{\phi} - \underline{y}^*)^T(\underline{\phi} - \underline{y}^*)/|\mathrm{S}|$. In this case, ξ_{KKT-LM} follows the asymptotic $F(n-2-|\mathrm{S}|, |\mathrm{S}|)$ distribution [16].

III. LINEAR REGRESSION IN TERMS OF ZEROED SECOND DIVIDED DIFFERENCES

This section considers a useful interpretation of the minimization of (1) subject to the equality constraints (11), as defined by the null hypothesis. Since $\tilde{y}[x_{i-1}, x_i, x_{i+1}] = 0$, the points $\{(x_k, \tilde{y}_k), k = i-1, i, i+1\}$, lie on a straight line. Then, in view of constraints (11), it follows that all points $\{(x_i, \tilde{y}_i) : i = 1,2,...,n\}$, lie on a straight line. Due to uniqueness, this remark implies that $\{\tilde{y}_i : i = 1,2,...,n\}$ may also be obtained from the linear regression model

$$\varphi_i = a x_i + b_i + \varepsilon_i, \ \varepsilon_i \sim N(0, \sigma^2), i = 1,2,...,n , \tag{18}$$

where a and b are parameters to be estimated given the observations $\{\phi_i : i = 1,2,...,n\}$ at the abscissae $\{x_i : i = 1,2,...,n\}$. Besides that the equality constrained minimization calculation (1)-(11) avoids the serious loss of accuracy that usually occurs in the linear regression model (18), there is an important reason to be preferred in our calculation. Because it employs constraints, it is particularly informative about those divided differences that have an impact on the line fit due to the size of the Lagrange multipliers. Indeed, since the Lagrange multipliers can be interpreted as a sensitivity measure of $\Phi(\underline{\tilde{y}})$ to the associated constraint, some $\tilde{\lambda}_i$ may suggest the possibility of further improvement of the linear fit, and they indeed do so when the data follow a convex trend. Further, if the Lagrange multipliers pinpoint important constraints due to the underlying convexity, some breaks may be allowed in the sequence (11) leading to the inequality constraints (2). This of course makes the convex estimator more flexible than the linear estimator of the data. The discussion suggests that the Lagrange multipliers of the linear estimate and the Karush-Kuhn-Tucker multipliers of the convex estimate may prove valuable in the construction of some statistics that test the convexity of the data.

The procedure for obtaining $\underline{\tilde{\lambda}}$ depends on the first order conditions (10), which are written as

$$2(\underline{y} - \underline{\varphi}) = \sum_{i=2}^{n-1}\lambda_i \underline{a}_i . \tag{19}$$

Since the constraint normals are linearly independent, it follows in theory that the overdetermined vector equation (19) is consistent and that it defines a unique vector $\underline{\tilde{\lambda}}$. The calculation of $\underline{\tilde{\lambda}}$ requires further attention. For each $i = 2,3,...,n-1$, we pick a scaled row of the vector equation (19) as follows. We choose the ith row multiplied by $(x_{i-1} - x_{i+1})$, so the first and last rows are never chosen, which helps numerical stability. Thus, we form a square system of equations in the required multipliers $\{\lambda_i : i = 2,3,...,n-1\}$, whose right hand sides have the values $\{2(x_{i-1} - x_{i+1})(y_i - \varphi_i) : \ i = 2,3,...,n-1\}$. This linear system is tridiagonal, the equation that is picked for λ_i having two nonzero off-diagonal elements, because the integers $i-1$ and $i+1$ also participate in this construction. Thus, the coefficient matrix is

$$D = \begin{bmatrix} \frac{(x_3-x_1)}{(x_2-x_1)(x_3-x_2)} & \frac{1}{(x_3-x_2)} & & & O \\ \frac{1}{(x_3-x_2)} & \frac{(x_4-x_2)}{(x_3-x_2)(x_4-x_3)} & \frac{1}{(x_4-x_3)} & & \\ & \frac{1}{(x_4-x_3)} & \ddots & \ddots & \\ & & \ddots & & \frac{1}{(x_{n-1}-x_{n-2})} \\ O & & & \frac{1}{(x_{n-1}-x_{n-2})} & \frac{(x_n-x_{n-2})}{(x_{n-1}-x_{n-2})(x_n-x_{n-1})} \end{bmatrix}.$$

Lemma 2 *Matrix D is positive definite*.

Proof. We immediately see that the diagonal elements of D are positive, the off-diagonal elements are negative and equal, each diagonal element is at least the sum of the moduli of the off-diagonal elements (either in each row, or in each column of the matrix), and suitable rows possess strict diagonal dominance. Thus, matrix D is positive definite. ■

It follows that the linear system in $\{\lambda_i : i = 2,3,...,n-1\}$ can be solved efficiently and stably by a Cholesky factorization in only $O(n)$ computer operations (see, for example, [12]). The calculation of $\underline{\tilde{y}}$ is carried out by the method of Section IV.

IV. AN EFFICIENT NUMERICAL METHOD FOR THE CONVEX ESTIMATION

In this section we outline the quadratic programming method of [7] for the calculation of the convex estimator $\underline{y}^*$. It is an elaborate technique that is by far faster than a general quadratic programming algorithm, because it takes into account the structure of the constraints. A large part of its efficiency is due to a *B*-spline representation of the solution and the banded matrices that occur. For proofs on the quadratic programming, one may consult this reference. This section, however, is not directed to a precise description of the algorithm, but to emphasize the facts that are needed in defining the test statistic ξ_{KKT-LM}.

The method begins by calculating an initial approximation to the convex estimator in only $O(n)$ computer operations, which is an advantage to the quadratic programming calculation, because either it identifies the optimal active set or it comes quite close to it. Quadratic programming generates a sequence of sets $\{S_k : k = 1,2,...\}$, where S_k is a subset of the constraint indices $\{2,3,...,n-1\}$ with the property

$$\underline{y}^T \underline{a}_i = 0 \ , \ i \in S_k . \tag{20}$$

We call *active*, the set of the constraints whose indices are in S_k and for each k, we denote by $\underline{y}^{(k)}$ the vector that minimizes (1) subject to the equations (20). Since the constraint normals are linearly independent, unique Karush-Kuhn-Tucker multipliers $\{\lambda_i^{(k)} : i \in S_k\}$ are defined by the first order optimality condition

$$2(\underline{y}^{(k)} - \underline{\phi}) = \sum\nolimits_{i \in S_k} \lambda_i^{(k)} \underline{a}_i \ . \tag{21}$$

The method starts by deleting constraints if necessary from the active set derived by the $O(n)$ approximation until all the remaining active constraints have nonnegative Karush-Kuhn-Tucker multipliers. This gives S_1. If $S_k, k \geq 1,$ is not the final set of the active set sequence, then the quadratic programming algorithm adds to the active set the most violated constraint and deletes constraints with negative multipliers alternately, until the Karush-Kuhn-Tucker conditions (see Section I) are satisfied.

Related to each active set S_k, this process requires the calculation of $\underline{y}^{(k)}$ and $\underline{\lambda}^{(k)}$, the latter being uniquely determined by an extension of the method that gives $\underline{\tilde{\lambda}}$. Specifically, for each integer i in S_k we pick the ith row of (21) multiplied by $(x_{i-1} - x_{i+1})$, which gives a block diagonal positive definite system of equations. Thus $\underline{\lambda}^{(k)}$ is derived efficiently and stably by a Cholesky factorization in only $O(|S_k|)$ computer operations, where $|S_k|$ is the number of elements of S_k.

The equality constrained minimization problem of $\underline{y}^{(k)}$ forms an important part of the calculation, because it is solved very efficiently by a reduction to an equivalent unconstrained one with fewer variables due to linear *B*-splines. If $y(x)$, $x_1 \leq x \leq x_n$, is the piecewise linear interpolant to the values $\{(x_i, y_i^{(k)}) : \ i = 1,2,...,n\}$, then $y(x)$ is a linear spline, with knots in the set $\{x_i : i \in \{1,2,...,n\} \setminus S_k\}$. Indeed, $y^{(k)}[x_{i-1}, x_i, x_{i+1}] = 0$, whenever $i \in S_k$, which implies colinearity of $(x_{i-1}, y_{i-1}^{(k)})$, $(x_i, y_i^{(k)})$ and $(x_{i+1}, y_{i+1}^{(k)})$, but if $y^{(k)}[x_{i-1}, x_i, x_{i+1}] \neq 0$, then i is the index of a knot in $y(x)$. Thus the knots of $y(x)$ are determined from the data points due to the active set. So let $j = n - 1 - |S_k|$, let $\{\xi_p : p = 1,2,..., j-1\}$ be the interior knots of $y(x)$ in ascending order, let also $\xi_{-1} = \xi_0 = x_1$ and $\xi_j = \xi_{j+1} = x_n$, and let $\{B_p : p = 0,1,..., j\}$ be a basis of normalized linear *B*-splines defined on the abscissae $\{x_i : i = 1,2,...,n\}$ and satisfying the equations $B_p(\xi_p) = 1$ and $B_p(\xi_q) = 0$, $p \neq q$ [26] :

$$B_p(x) = \begin{cases} (x - \xi_{p-1})/(\xi_p - \xi_{p-1}), \ \xi_{p-1} \leq x < \xi_p \\ (\xi_{p+1} - x)/(\xi_{p+1} - \xi_p), \ \xi_p \leq x < \xi_{p+1} \\ \qquad 0, \text{otherwise.} \end{cases} \tag{22}$$

Then $y(x)$ may be written uniquely in the form

$$y(x) = \sum\nolimits_{p=0}^{j} \sigma_p B_p(x), \ \ x_1 \leq x \leq x_n, \tag{23}$$

where the coefficients $\{\sigma_p : p = 0,1,..., j\}$ are the values of $y(x)$ at $\{\xi_p : p = 0,1,..., j\}$ and are calculated by solving the normal equations associated with the minimization of $\sum\nolimits_{i=1}^{n} [y(x_i) - \phi_i]^2$. To be specific the normal equations can be written in the form of a square system of equations

$$\sum\nolimits_{p=0}^{j} \left[\sum\nolimits_{i=1}^{n} B_k(x_i) B_p(x_i)\right] \sigma_i = \sum\nolimits_{i=1}^{n} B_k(x_i) \phi_i \ , k = 0,1,..., j . \tag{24}$$

Since

$$\sum\nolimits_{i=1}^{n} B_k(x_i) B_p(x_i) = 0, \text{ for } |k - p| > 1 , \tag{25}$$

system (24) simplifies to the system

$$\begin{bmatrix} e_0 & d_1 & & & & \\ d_1 & e_1 & d_2 & & O & \\ & d_2 & e_2 & d_3 & & \\ & & \ddots & \ddots & \ddots & \\ & O & & d_{j-1} & e_{j-1} & d_j \\ & & & & d_j & e_j \end{bmatrix} \begin{bmatrix} \sigma_0 \\ \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_{j-1} \\ \sigma_j \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{j-1} \\ b_j \end{bmatrix}, \quad (26)$$

where, for $p = 0,1,...,j$, we define

$$e_p = \sum_{x_i \in [\xi_{i-1}, \xi_{i+1}]} \left[B_p(x_i) \right]^2 \quad (27)$$

and

$$b_p = \sum_{x_i \in [\xi_{i-1}, \xi_{i+1}]} B_p(x_i)\phi_i , \quad (28)$$

and, for $p = 1,2,...,j$, we define

$$d_p = \sum_{x_i \in [\xi_{i-1}, \xi_{i+1}]} B_p(x_i) B_{p-1}(x_i) . \quad (29)$$

We remark that the *hat* function $B_p(x)$ overlaps only with its closest neighbors $B_{p-1}(x)$ and $B_{p+1}(x)$, and the coefficient matrix of the equations (26) is positive definite (see, for example, [22]). Thus, Cholesky factorization is applied to give $\{\sigma_p = y(\xi_p): \ p = 0,1,..., j = n-1-|S_k|\}$ in $O(j)$ computer operations. The intermediate components of $\underset{\sim}{y}^{(k)}$ are found by linear interpolation to the coefficients σ_p's due to (23) and (22). Further, suitable updating procedures of $y(x)$ during the quadratic programming calculation as a knot is inserted into the spline basis (that is when a constraint is dropped from the active set) or a knot is deleted from the spline basis (that is when a constraint is added to the active set) have been employed that require no more than $O(n)$ operations. Finally, it is clear that the calculation of $\underset{\sim}{\tilde{y}}$ requires only B_0 and B_1 in (23).

V. NUMERICAL PERFORMANCE OF THE CONVEXITY TEST ALGORITHM

In this section we present numerical results that illustrate the efficiency of the method of Section IV and the performance of the test of Section II on several data sets. Conclusions concerning the efficiency of the computation and the efficacy of the tests are also presented here.

The data $\{\varphi_i : i = 1,2,...,n\}$ were random perturbations of two underlying convex functions, namely,

$$f(x) = \exp(-\tfrac{1}{2}\ln x + 2), \quad x \in (0.1, 1.1) \quad (30)$$

and

$$f(x) = \exp(-1.4x + 2), \ x \in (-0.728, 0.034) . \quad (31)$$

The number of data ranges from 500 to 2000, and for each n the abscissae have equally spaced values on the given intervals. For each of the two underlying functions and each n, three values of a nonnegative parameter σ were chosen. Each φ_i was generated by adding to $f(x_i)$ a number from the uniform distribution in $(-\sigma,\sigma)$. The actual values of σ and n and some calculated results are given in Table 1 for each of the functions (30) and (31). The third column displays the CPU time required to perform the calculation in double precision arithmetic on an Intel Pentium 466 MHz personal computer operating in a Windows 98 environment. The fourth and fifth columns of this table show the number of active constraints ($|S|$) and the number of active set changes (constraint additions to or deletions from the active set) required by the program to calculate the convex estimator. The remaining columns present the values of the objective function at the convex estimator and the linear estimator, the maximum and minimum of the Karush-Kuhn-Tucker and the Lagrange multipliers and the statistic ξ_{KKT-LM}. The actual values of the $F(n-2-|S|,|S|)$ statistic are not given here, because they are by far smaller than the presented values of ξ_{KKT-LM}. Therefore in all the experiments we reject H_0. Notice that (cf. Table 1) the smaller the error σ is, the larger ξ_{KKT-LM} becomes. Indeed, as σ tends to zero, measurements tend to the function values, in which case the convex estimator tends to interpolate convex data values, (17) becomes $\xi_{KKT-LM} = (\underset{\sim}{\tilde{y}} - \underset{\sim}{\varphi})^T (\underset{\sim}{\tilde{y}} - \underset{\sim}{\varphi}) / s^2 (n-2-|S|)$ and s^2 tends to zero. Hence ξ_{KKT-LM} obtains very large values and the null is rejected. Of course this is true only if $n > 2$. If $n = 2$, then ξ_{KKT-LM} is zero, thus the test suggests trivially the linearity of the convex estimator. If $n = 1$, then we trivially define ξ_{KKT-LM} to be zero. Further, the negative signs of the Lagrange multipliers and the large values of the objective function (1) at $\underset{\sim}{\tilde{y}}$, are also grounds for rejecting the null. The sizes of the Karush-Kuhn-Tucker multipliers provide a measure of the sensitivity of the problem to the constraints.

The ratio of the *CPU time* over the *active set changes* shows about the work required by one active set change. It also indicates the work required to obtain the linear estimator by the equality constrained optimization of Section II, which includes not only the calculation of $\underset{\sim}{\tilde{y}}$, but also the calculation of the Lagrange multipliers. The results on the active set changes show that the algorithm of [7] is quite efficient in obtaining the convex estimator and the associated multipliers terminating in a number of active set changes that is only a fraction of the number n. As one active set change requires $O(n)$ computer operations, column five suggests that this algorithm performs in practice almost linearly with respect to n.

VI. APPLICATIONS OF THE CONVEXITY TEST

In this section we apply the convexity test on a set with artificial data and on some sets with real data from money demand in the U.S.A. for the period 1919–1964, the GNP per capita and infant mortality in 147 countries and the Gini coefficients in the U.S.A. for the period 1947–1996. We present our conclusions for each case separately.

Table 1 Numerical results and performance for the Karush-Kuhn-Tucker convexity test statistic

σ	n	CPU time	$\|S\|$	Active set changes	*Value of* (1) Convex fit	Linear fit	*Kuhn-Tucker multipliers* max	min	*Lagrange multipliers* max	min	*KKT-LM statistic*
					Function (30)						
	500	0.05	133	0	3.11E-07	1.25E+03	5.03E-07	4.90E-11	-2.30E-03	-3.86E+02	1468884775,64
.0001	1000	0.05	479	6	1.25E-06	2.48E+03	3.36E-07	1.97E-10	-4.07E-03	-1.94E+02	1832425092,77
	2000	0.68	1318	63	4.01E-06	4.95E+03	3.42E-07	6.63E-10	-8.15E-03	-9.73E+01	2389260255,01
	500	0.10	469	66	1.43E+00	1.24E+03	9.27E-03	4.51E-06	-8.11E-03	-9.68E+01	14034,46
.1	1000	0.32	962	147	3.03E+00	2.48E+03	9.79E-03	4.12E-07	-4.20E-03	-1.94E+02	21864,70
	2000	1.53	1951	316	6.45E+00	4.95E+03	1.04E-02	9.98E-08	-2.05E-03	-3.86E+02	31794,63
	500	0.10	482	56	1.52E+02	1.30E+03	4.28E-01	7.25E-05	-7.69E-03	-9.31E+01	228,91
1	1000	0.27	985	146	3.12E+02	2.77E+03	3.70E-01	4.34E-05	-5.41E-03	-1.95E+02	597,32
	2000	0.82	1979	162	3.12E+02	5.57E+03	1.75E-01	3.71E-05	-2.24E-03	-3.85E+02	1754,51
					Function (31)						
	500	0.00	197	0	3.68E-07	1.37E+02	1.66E-07	1.27E-09	-3.20E-03	-2.79E+01	244172861,15
.0001	1000	0.05	624	3	1.65E-06	2.74E+02	1.93E-07	7.27E-11	-1.60E-03	-5.55E+01	277071837,35
	2000	0.98	1571	81	4.85E-06	5.46E+02	3.38E-07	2.35E-10	-7.99E-04	-1.11E+02	414581399,90
	500	0.05	479	60	1.49E+00	1.35E+02	6.96E-03	3.73E-06	-3.16E-03	-2.75E+01	2271,92
.1	1000	0.27	974	137	3.09E+00	2.77E+02	5.79E-03	2.93E-06	-1.70E-03	-5.56E+01	3591,30
	2000	1.31	1970	285	6.51E+00	5.52E+02	7.89E-03	1.28E-06	-8.15E-04	-1.11E+02	5886,80
	500	0.10	489	74	1.53E+02	2.59E+02	2.52E-01	8.33E-06	-6.62E-04	-2.48E+01	37,99
1	1000	0.21	989	98	3.14E+02	5.91E+02	2.69E-01	4.34E-05	-4.19E-04	-5.66E+01	96,78
	2000	0.93	1986	218	6.58E+02	1.20E+03	4.73E-01	6.61E-06	2.05E-04	1.10E+02	135,89

A. Artificial data from use of the Cobb-Douglas function

To illustrate the test we present a numerical example, which is a fit to measurements of $f(x) = (27/x^{0.3})^{1/0.7}$ that is obtained from the Cobb-Douglas curve $x^{0.3}y^{0.7} = 27$. An advantage of having a known underlying function is that we can see immediately whether the convex fit is more accurate than the data, and it is. The reason for employing a Cobb-Douglas function is because of its wide use in economic computations under convexity assumptions. One hundred data at equally spaced abscissae x_k in the interval [1,100] were chosen and we simulated data errors by adding to $f(x_k)$ a number sampled from the uniform distribution over the interval $[-5,5]$.

A Fortran 77 program is provided by [6] that calculates the least squares convex estimator $\underline{y}^*$ and the Karush-Kuhn-Tucker multipliers $\underline{\lambda}^*$. The authors have extended this program so as to include the calculation of the equality constrained minimization that gives the linear fit $\underline{\tilde{y}}$, the Lagrange multipliers $\underline{\tilde{\lambda}}$ and the test statistic (17).

The Lagrange and the Karush-Kuhn-Tucker multipliers yield an indication of the impact of a constraint in both the linear and the convex model. Indeed, the Lagrange multipliers are all negative with $\min_{1\le i\le n-2}\tilde{\lambda}_i = -8968.52$ and $\max_{1\le i\le n-2}\tilde{\lambda}_i = -18.92$ implying that the linear fit is too restraining for the data and providing, in some sense, a measure of the evidence against H_0.

Table 2 Convex fit estimates (y_i^*) to data (φ_i) from a Cobb-Douglas contour and rates of change (first divided differences)

x_i	φ_i	y_i^*	*First divided differences*
1	109,35	109,35	
2	79,21	79,21	-30,1452
3	64,98	64,98	-14,2281
4	56,34	56,34	-8,6459
5	50,78	51,71	-4,6221
7	43,98	44,91	-3,4009
8	40,77	42,75	-2,1632
15	33,35	34,69	-1,1513
34	21,21	23,29	-0,5999
36	21,09	22,13	-0,5805
65	13,76	18,85	-0,1130
99	10,96	16,09	-0,0811
100	17,93	17,93	1,8331

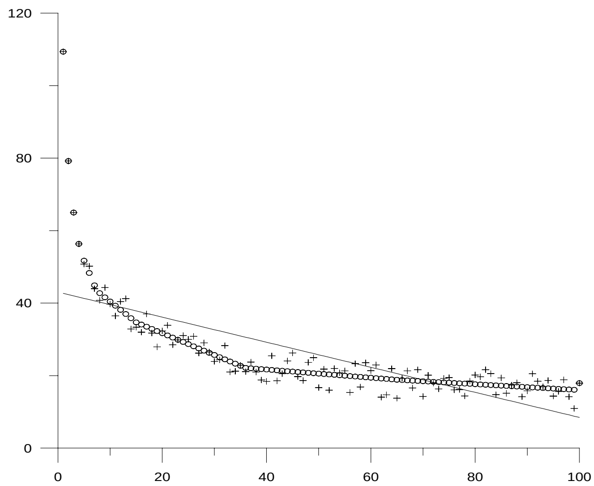

Figure 1 Best least squares convex and linear estimation to 100 data generated by the Cobb-Douglas contour of Section VI. The data are denoted by (+) and the smoothed values by (○). The straight line gives the least squares linear regression to the data. Note the piecewise linearity of the best convex estimator.

By relaxing constraints (11) so as to obtain (2), the convex fit is a polygonal line with 11 interiors knots (at those x_i, where $y^*[x_{i-1}, x_i, x_{i+1}]$ >0), thus allowing 11 zero Karush-Kuhn-Tucker multipliers (one per knot). The remaining multipliers vary between $\min_{i\in S} \lambda_i = 0.06$ and $\max_{i\in S} \lambda_i =$ 118.36. As for the extreme values of the multipliers, it is interesting to note that the absolute Lagrange multipliers are about two orders of magnitude larger than the Karush-Kuhn-Tucker multipliers. It is not surprising then, that the 11 interior knots of the convex fit implied the inequalities $R^2(\underline{\tilde{y}}) = 0.5079 < R^2(\underline{y}^*) = 0.9662$ and $\Phi(\underline{y}^*) = 663.32 <$ $9647.71 = \Phi(\underline{\tilde{y}})$, where $R^2(.)$ is the determination coefficient, both inequalities being in favour of H_1. In Table 2 we have tabulated x_i, φ_i, y_i^* and the first divided differences of y_i^* at the knots and at the end points. The first differences show the in-between the knots rates of change of the convex estimate. In Fig. 1 we plot both the convex and the linear estimator.

However, the validity of the convex model would require the test of Section II. Since the variance is equal to $\sigma^2 =$ $(5-(-5))^2/12 = 8.33$, due to the uniform errors that we have used in order to generate the data, the value of the test statistic (17) is ξ_{KKT-LM} =1078.56~ $\chi^2(11)$, where $\chi^2(11) = 19.68$ for α =0.05 and 24.73 for α =0.01. Since at both levels of significance we have $\xi_{KKT-LM} > \chi^2(11)$, we reject the linear model in favour of the convex one. Furthermore, assuming that σ^2 is unknown (as it is usual in practice), we replace σ^2 in formula (17) by its estimate $s^2 = 7.62$ and obtain ξ_{KKT-LM} =107.13 that is larger than $F(11,87) = 1.90$ for α =0.05 and 2.46 for α =0.01. So, we again reject H_0.

B. *Money Demand*

Given 44 observations from money demand in the USA for the period 1919–1964, we wish to provide an estimate of the demand function and to test the underlying assumption of convexity. The source of our data is the U.S. Department of Commerce and the time series include the quantity of money *M* (money stock in currency and demand deposits in billions of \$ in 1958 prices), the interest rate *r* (U.S. Treasury bill yield with a 3 to 6 month maturity in %) and the national product *Y* (GNP in billions of \$ in 1958 prices). The values of *r* and *M*/*Y* are presented in the first two columns of Table 3.

An interesting feature of Table 3 is that the Lagrange multipliers $\tilde{\lambda}_i$ occur in two groups. The first 14 multipliers are positive and the last 28 multipliers are negative, the latter group indicating that the line fit is rather restricting the associated range of data. Therefore we may require that a certain number of indices of inactive constraints separate adjacent indices of the constraints with negative multipliers. This idea suggests allowing the inequality constraints (2) instead of the equality constraints (11) in the calculation of an estimate. Thus, we obtain the convex fit, whose components are in column 3, and it is interesting to note that it has only one interior knot located at $x_{37} = 4.38$ (see Fig. 2). Moreover, the first 38 components of the convex fit decrease with respect to *r* with a rate of change equal to -0.0338, while the last 8 components (now *r* exceeds 4%) increase with a rate equal to 0.0079. Therefore our method reveals what economic theory suggests: if interest rates rise, one prefers to hold money than bonds, thus leading *M*/*Y* to higher values. Further, the inequalities $R^2(\underline{\tilde{y}})$ =0.5531< $R^2(\underline{y}^*)$ =0.6600 and $\Phi(\underline{y}^*) = 0.0489 <$ $0.0643 = \Phi(\underline{\tilde{y}})$ are in favour of the convex fit, but the discrepancy of these values is so small that it justifies the one-knot convex model of Fig. 2.

Table 3 The convex fit, the line fit and the Lagrange multipliers to data from money demand in the U.S.A. between 1919-1965

x_i (*r*)	φ_i (*M/Y*)	y_i^*	$\tilde{y}_i$	λ_i^*	$\tilde{\lambda}_i$	x_i (*r*)	φ_i (*M/Y*)	y_i^*	$\tilde{y}_i$	λ_i^*	$\tilde{\lambda}_i$
0.54	0.366	0.396	0.382	-	-	2.64	0.314	0.325	0.332	0.1295	-0.429
0.56	0.388	0.395	0.382	0.0012	0.001	2.73	0.357	0.322	0.330	0.1173	-0.465
0.59	0.369	0.394	0.381	0.0034	0.001	2.97	0.275	0.314	0.324	0.0679	-0.572
0.66	0.346	0.392	0.379	0.0120	0.004	3.26	0.261	0.304	0.317	0.0305	-0.674
0.69	0.374	0.391	0.379	0.0184	0.007	3.31	0.324	0.302	0.316	0.0283	-0.686
0.73	0.403	0.389	0.378	0.0284	0.012	3.55	0.255	0.294	0.310	0.0073	-0.747
0.75	0.352	0.389	0.377	0.0328	0.014	3.59	0.282	0.293	0.309	0.0070	-0.753
0.76	0.352	0.388	0.377	0.0357	0.015	3.81	0.310	0.285	0.304	0.0097	-0.773
0.81	0.429	0.387	0.376	0.0542	0.023	3.85	0.280	0.284	0.303	0.0083	-0.777
0.94	0.336	0.382	0.373	0.0910	0.031	3.97	0.271	0.280	0.300	0.0048	-0.783
1.02	0.329	0.380	0.371	0.1211	0.042	3.98	0.282	0.280	0.300	0.0047	-0.783
1.03	0.484	0.379	0.370	0.1259	0.044	4.02	0.277	0.278	0.299	0.0040	-0.782
1.44	0.436	0.365	0.361	0.2363	0.041	4.11	0.273	0.275	0.297	0.0027	-0.774
1.45	0.401	0.365	0.360	0.2376	0.040	4.34	0.266	0.267	0.291	0.0003	-0.745
1.49	0.434	0.364	0.359	0.2399	0.030	4.38	0.236	0.266	0.290	0.0000	-0.737
1.58	0.357	0.361	0.357	0.2324	-0.005	4.52	0.295	0.267	0.287	0.0074	-0.697
1.73	0.351	0.356	0.354	0.2211	-0.063	4.85	0.272	0.270	0.279	0.0067	-0.606
2.16	0.364	0.341	0.343	0.1929	-0.226	5.07	0.270	0.271	0.274	0.0051	-0.542
2.18	0.336	0.340	0.343	0.1906	-0.234	5.37	0.274	0.274	0.267	0.0036	-0.453
2.33	0.362	0.335	0.339	0.1753	-0.295	5.85	0.256	0.278	0.255	0.0012	-0.317
2.46	0.309	0.331	0.336	0.1551	-0.354	6.62	0.319	0.284	0.237	0.0299	-0.102
2.52	0.351	0.329	0.335	0.1483	-0.378	7.50	0.274	0.291	0.216	-	-

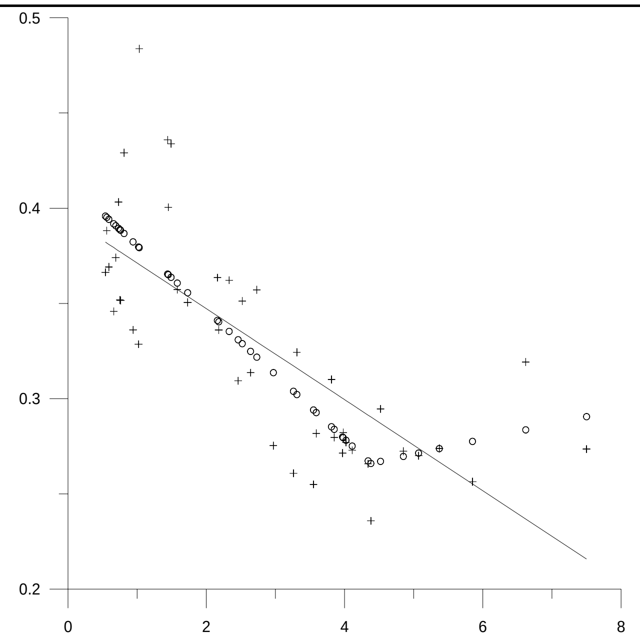

Figure 2 As in Fig. 1, but the 44 data of Table 2 (money demand) are used.

In order to test the assumption of linearity against convexity, we substitute the estimate s^2=0.00119 of σ^2 into the formula (17) and we obtain ξ_{KKT-LM} = 12.89, which is larger than $F(1,41) = 4.08$ for $\alpha = 5\%$ and 7.30 for $\alpha = 1\%$. We reject linearity in both cases.

C. GNP versus Infant Mortality Rate

This example is concerned with the relationship between the per capita GNP and the infant mortality rate (IMR). The data consist of the GNP per capita in thousands of $ and the infant mortality rate per 1000 births for the year 1995 for 149 countries, the data being obtained from the World Bank Database. The actual values are not presented here, but in Fig. 3 we see their graphical representation together with the convex and the line fit. Also in Table 4 we present the knots (*data index* column), the GNP per capita, the convex fit (*convex estimate of IMR* column) and the rates of change (*first divided differences* column) of the convex estimator. The convex fit contains 7 interior knots and the calculation has given $\min \tilde{\lambda}_i = -10\,224\,500$, $\max \tilde{\lambda}_i = -1215$, $\min_{i\in S} \lambda_i^* = 262$ and $\max_{i\in S} \lambda_i^* = 50\,867$. Further, we obtain the inequalities $R^2(\tilde{y})$=0.30< $R^2(\underline{y}^*)$=0.74 and $\Phi(\underline{y}^*) = 69753 < 190981 = \Phi(\underline{\tilde{y}})$, both in favour of the convex fit. We are interested in the convex relation between the two variables, which explains the argument that for high values of IMR, a small rise in GNP will provoke a relatively larger decrease in infant mortality, whereas after an adequate value of GNP has been achieved infant mortality will not decrease significantly any further and after some GNP level, IMR will be almost flat. The convex estimator determines the rates of change of this relationship, while it reveals the linearity between any two adjacent knots. Indeed, we see that the IMR estimated components decrease rapidly between the first three knots with negative though increasing rates of change and subsequently, after the fourth interior knot (GNP=10142.20), they are close to zero. However, it has been observed by the theoreticians [2], [4] that IMR increases for very high GNP levels. It is remarkable that our method has revealed such an increase between the last two knots. Indeed, we see in Table 4 that the convex IMR estimates decrease from 134.66 down to 5.11 and then increase up to 7.56. Moreover the method estimated that the rates of increase between the last three knots are 0.000013 and 0.000281 (which means that as GNP increases to the highest level, the rate of increase of IMR is about 21 times faster than in the previous GNP level). To explain this behaviour 'urbanization' theories are employed [2], which suggest that in highly developed countries with high GNP pc values not only poverty increases, but also living conditions are not favourable for measures such as life expectancy at birth etc.

The calculation of the convexity test is performed by considering the estimate s^2=498.24 of σ^2, which we substitute into the formula (17) and obtain the test statistic value ξ_{KKT-LM} = 34.76. Since this value is larger than $F(7,141)$=2.08 for $\alpha = 5\%$ and 2.77 for $\alpha = 1\%$, we reject the null hypothesis.

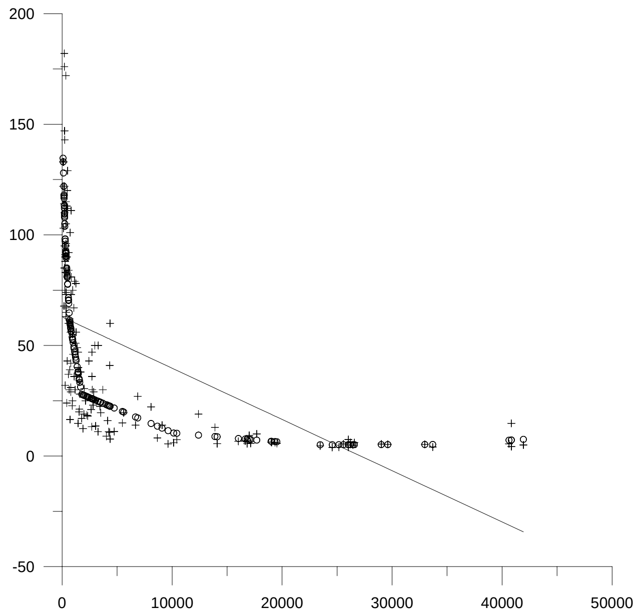

Figure 3 As in Fig. 1, but 147 data of the GNP pc versus the Infant Mortality Rate are used (Source: the World Bank Database for the year 1995).

Table 4 GNP pc and IMR convex estimate for GNP versus Infant Mortality Rate for the year 1995 for 149 countries

Data index	*GNP pc*	*Convex estimate of IMR*	*First divided differences*
1	85.09	134.66	-
18	276.45	98.06	-0.191263
48	696.24	60.41	-0.089688
80	1785.24	27.89	-0.029862
118	10142.20	10.41	-0.002092
132	19522.11	6.46	-0.000421
133	23463.75	5.11	-0.000342
145	33692.07	5.24	0.000013
149	41935.20	7.56	0.000281

D. Gini coefficient time series

We consider the Gini coefficient (see, for example, [20]) and its evolution in the U.S.A. for the time period 1947 to 1996. Fifty data points were retrieved from the World Income Inequality Database of the U.S. Bureau of Census 1997 and presented in the first two columns of Table 5, the remaining columns been explained in Table 3. The interest here does not lie on any theoretical assumption of convexity nor on any underlying relation that has to be validated, but on investigating the macroeconomic trend of the Gini coefficient. The convex estimator of the Gini data yields an estimate of the rate of income inequality change over a long time period perspective. We see in Table 5 that $\min \tilde{\lambda}_i = -535{,}83$, $\max \tilde{\lambda}_i = -5.36$, $\min_{i\in S} \lambda_i^* = 0.133$ and $\max_{i\in S} \lambda_i^* = 13.62$. Also, we find that $R^2(\underline{\tilde{y}}) = 0.40 < R^2(\underline{y}^*) = 0.94$ and $\Phi(\underline{y}^*) = 13.56 < 128.28 = \Phi(\underline{\tilde{y}})$. All these values are in favour of the convex fit, which contains 8 interior knots (where $\lambda_i^* = 0$) and it is illustrated in Fig. 4. Moreover, we can calculate from Table 5 the in-between the knots rates of change of the Gini estimate, namely -0.21, -0.08, -0.08, 0.10, 0.10, 0.25, 0.32 and 0.50, with respect to the periods [1947–53], [53–56], [56–69], [69–75], [75–76], [76–78], [78-79], [79–91] and [91–96]. Thus the Gini coefficients decrease during the first half of the time range (1947–68) and subsequently increase during (1969–96), with a rate of change of income inequality that increases gradually to positive values.

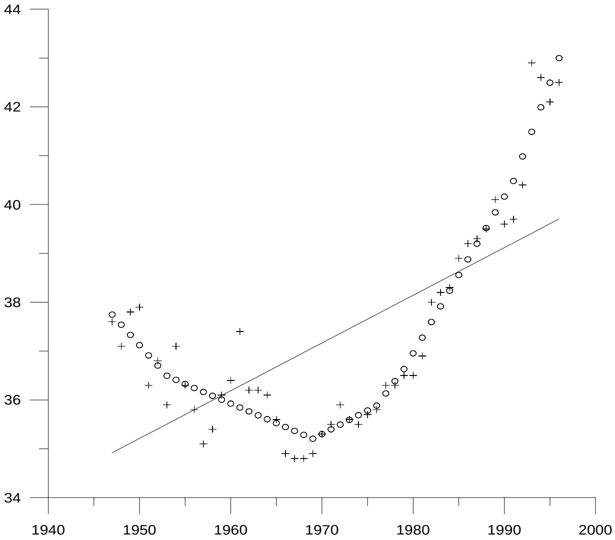

Figure 4 As in Fig. 1, but the 50 data (Gini coefficients) of Table 4 are used.

In order to test linearity against convexity we use the estimate s^2=0.33 of σ^2 in (17) and the value of the statistic is ξ_{KKT-LM} = 49.54 $> F(7,41)$=2.24 for α=0.05 and 3.11 for α=0.01. Therefore we reject linearity over convexity.

Table 5 As in Table 3, but the data are from the Gini coefficients in the U.S.A. between 1947-1990

x_i *Year*	φ_i *Gini coef*	y_i^*	$\tilde{y}_i$	λ_i^*	$\tilde{\lambda}_i$	x_i *Year*	φ_i *Gini coef*	y_i^*	$\tilde{y}_i$	λ_i^*	$\tilde{\lambda}_i$
1947	37.6	37.75	34.92	-	-	1972	35.9	35.50	37.36	1.24	-535.83
1948	37.1	37.54	35.02	0.30	-5.36	1973	35.6	35.59	37.46	0.71	-534.95
1949	37.8	37.33	35.11	1.47	-14.89	1974	35.5	35.69	37.56	0.17	-530.37
1950	37.9	37.12	35.21	1.71	-29.79	1975	35.7	35.78	37.65	0.00	-521.66
1951	36.3	36.91	35.31	0.39	-50.07	1976	35.8	35.88	37.75	0.00	-509.06
1952	36.8	36.70	35.41	0.29	-72.33	1977	36.3	36.13	37.85	0.17	-492.54
1953	35.9	36.49	35.50	0.00	-97.37	1978	36.3	36.38	37.95	0.00	-472.93
1954	37.1	36.41	35.60	0.90	-123.21	1979	36.5	36.63	38.04	0.00	-450.03
1955	36.3	36.33	35.70	0.42	-152.03	1980	36.5	36.95	38.14	0.27	-424.04
1956	35.8	36.24	35.80	0.00	-182.06	1981	36.9	37.27	38.24	1.44	-394.76
1957	35.1	36.16	35.90	0.47	-212.09	1982	38.0	37.60	38.34	3.36	-362.80
1958	35.4	36.08	35.99	3.06	-240.54	1983	38.2	37.92	38.44	4.48	-330.17
1959	36.1	36.00	36.09	7.02	-267.79	1984	38.3	38.24	38.53	5.02	-297.07
1960	36.4	35.92	36.19	10.80	-295.06	1985	38.9	38.56	38.63	5.44	-263.50
1961	37.4	35.84	36.29	13.62	-322.76	1986	39.2	38.88	38.73	5.17	-230.47
1962	36.2	35.76	36.38	13.33	-352.68	1987	39.3	39.20	38.83	4.26	-198.38
1963	36.2	35.69	36.48	12.17	-382.23	1988	39.5	39.52	38.92	3.15	-167.24
1964	36.1	35.61	36.58	9.98	-411.23	1989	40.1	39.84	39.02	2.07	-137.25
1965	35.6	35.53	36.68	6.80	-439.26	1990	39.6	40.16	39.12	0.47	-109.42
1966	34.9	35.45	36.77	3.47	-465.14	1991	39.7	40.48	39.22	0.00	-82.55
1967	34.8	35.37	36.87	1.24	-487.26	1992	40.4	40.98	39.31	1.09	-56.65
1968	34.8	35.29	36.97	0.13	-505.25	1993	42.9	41.49	39.41	3.35	-32.92
1969	34.9	35.21	37.07	0.00	-518.89	1994	42.6	41.99	39.51	2.78	-16.16
1970	35.3	35.30	37.17	0.48	-528.20	1995	42.1	42.49	39.61	1.00	-5.59
1971	35.5	35.40	37.26	0.96	-533.77	1996	42.5	43.00	39.71	-	-

VII. DISCUSSION

In this paper we have been concerned with the problem of testing the convexity of a set of univariate observations that include random errors. The convex estimation problem was published first by [17]. It addresses the question of making the least sum of squares change to the observations so that the piecewise linear interpolant to the smoothed data to be convex. This problem is a highly structured quadratic programming calculation that is solved very efficiently by the special algorithm of [7]. We have provided a Karush-Kuhn-Tucker test statistic based on [13], [16] and [28]. Our work has been helped substantially because of a Fortran implementation of the quadratic programming algorithm by [6] that allowed the practical application of the testing procedure. Of course, the algorithm and the Fortran codes have been enhanced so as to include the linearly equality constrained calculation that gives the linear estimate and the test statistic.

The efficacy of the convexity test on artificial measurements from a Cobb-Douglas curve and on real data sets from money demand in the U.S.A., the GNP per capita and infant mortality in 147 countries and the Gini coefficients in the U.S.A. for the period 1947-1996 was demonstrated.

The convexity test was proved capable in practice and consistent with the economic theories that underlie the data that was tested on. It is expected to find important applications in computations in economics, especially in large-scale computations, when the required assumption is that of increasing returns. It should be mentioned, however, that the test is also suitable when it would be better to assume diminishing instead of increasing returns.

REFERENCES

[1] J. Aitchinson and S. D. Silvey, Maximum-Likelihood estimation of parameters subject to restraints. *Ann. Math. Statist*., Vol 29, 1958, pp. 813-828.

[2] L. Bertinelli, *Urbanization and Growth*. London, UK: Duncan Black, 2000.

[3] C. de Boor, *A practical Guide to Splines Revised Edition.* New York: Springer-Verlag, 2001.

[4] I. Chakraborty, Living standard and economic growth: A fresh look at the relationship through the nonparametric approach. WP283.pdf, www.cds.edu, 1997.

[5] A. Cuevas and W. Gonzalez-Manteiga, Data-driven smoothing based on convexity properties. In *Nonparametric Functional Estimation and Related Topics*, G. Roussas, Ed. Netherlands: Kluwer Publishers, 1991, pp. 225-240.

[6] I. C. Demetriou, Algorithm 742: L2CXFT, A Fortran 77 subroutine for least squares data fitting with non-negative second divided differences. *ACM Transactions on Math. Software*, Vol. 21, No 1, 1995, pp. 98-110.

[7] I. C. Demetriou and M. J. D. Powell, The minimum sum of squares change to univariate data that gives convexity. *IMA J. of Numerical Analysis*, 11, 1991, pp. 433-448.

[8] I. C. Demetriou and M. J. D. Powell, Least squares fitting to univariate data subject to restrictions on the signs of the second differences. In *Approximation Theory and Optimization. Tributes to M. J. D. Powell*, M. D. Buhmann and A. Iserles, Eds. Cambridge, UK: Cambridge University Press, 1997, pp. 109-132.

[9] R. W. Farebrother, Testing linear inequality constraints in the standard linear model. *Comm. Statistics* (*A*), 15, 1986, pp. 7-31.

[10] R. Fletcher, *Practical Methods of Optimization*. Chichester, UK: J. Wiley & Sons, 2002.

[11] F. R. Giordano, M. D. Weir and W. P. Fox, *A First Course In Mathematical Modeling*, 2nd ed. New York: Brooks / Cole Pub. Co., 1997.

[12] G. E. Golub and C. F. van Loan *Matrix Computations* 2nd ed. Baltimore and London: The John Hopkins University Press 1989.

[13] C. Gouriéroux, A. Holly and A. Monfort, Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters. *Econometrica*, Vol. 50, No 1, 1982, pp. 63-79.

[14] P. Groeneboom, G. Jongbloed and A. J. Wellner, Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist*., 29, 2001, pp. 1653-1698.

[15] D. L. Hanson and G. Pledger, Consistency in concave regression. *Ann. Statist*., 1976, pp. 1038-1050.

[16] J. A. Hartigan, Distribution of the residual sum of squares in fitting inequalities. *Biometrica*, 54, Vol. 1 and 2, 1967, pp. 69-84.

[17] C. Hildreth, Point estimates of ordinates of concave functions. *J. Amer. Stat. Assoc*., 49, 1954, pp. 598-619.

[18] D. V. Lindley, *Making Decisions*, 2nd ed. London, UK: J. Wiley & Sons, 1985.

[19] J. Nocedal and S. J. Wright *Numerical Optimization*. New York: Springer, 1999.

[20] M. Pemberton and N. Rau, *Mathematics for Economists*, 2nd ed. Manchester: Manchester University Press, 2007.

[21] M. D. Perlman, One Sided Problems in Multivariate Statistics. *Ann. of Math. Statist.*, 40, 1969, pp. 549-567.

[22] M. J. D. Powell, *Approximation Theory and Methods.* Cambridge, UK: Cambridge University Press, 1981.

[23] T. Robertson, F. T. Wright and R. L. Dykstra, *Order Restricted Statistical Inference*. Chichester, UK: J. Wiley & Sons, 1988.

[24] A. Rosenberg, *Economics – Mathematical Politics or Science of Diminishing Returns?* Chicago and London: The University of Chicago Press, 1992.

[25] P. Ruud, Restricted least squares subject to monotonicity and concavity constraints. Technical Report, University of California, Berkeley, California, 1996.

[26] L. L. Schumaker, *Spline Functions: Basic Theory*. New York: J. Wiley & Sons, 1981.

[27] D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavioral Research*. Cambridge, UK: Cambridge University Press, 1986.

[28] F. A. Wolak, An exact test for multiple inequality and equality constraints in the linear regression model. *J. Amer. Stat. Assoc*., Vol. 82, No 399, 1987, pp.782-792.