A Distributed Lag Estimator with Piecewise Monotonic Coefficients

I. C. Demetriou and E. E. Vassiliou *

Abstract—**Linearly distributed-lag models as a time series tool have very useful applications in many disciplines. In these models, the dependent variable depends on one independent variable and its lags. The specification of the lag coefficients is a crucial question to the efficacy of a model. A new approach is proposed for the estimation of lag coefficients subject to the condition that the sequence of the coefficient estimates consists of a certain number of monotonic sections, where the positions of the extrema are also unknowns. The underlying algorithm is iterative, each iteration taking a descent direction, then forming an estimate of the coefficients and finally adjusting this estimate to satisfy the given constraints. An immediate advantage of this approach is that the algorithm avoids inverting an ill-conditioned matrix that frequently occurs in practice. Moreover, the constraints provide a realistic representation of the prior knowledge and the calculation results in a highly efficient time series estimation. The algorithm is described, a proof of convergence is given and an application of the algorithm on real annual macroeconomic data concerning the personal consumption expenditures against the GDP for the U.S.A. during 1929 - 2006 is presented.**

Keywords: approximation, consumption, distributed lag model, piecewise monotonic, regression, smoothing, time series

1 Introduction

The purpose of distributed-lag models is to estimate, from time series data, values y that incorporate prior information of the independent variable x. These models have useful applications in many fields such as econometrics (see, for instance, [18], [21]), engineering (see, for instance, [7], [13], [8]) etc. For example, in econometrics, if y_t denotes consumption expenditures and x_t income, at time period t, a change in x_t will affect not only current consumer expenditures y_t, but also future expenditures y_{t+1}, y_{t+2}, etc. Therefore we assume that y_t depends not only on x_t but also on q past values of x_t, giving the linearly distributed-lag model

$$y_t = \sum_{i=0}^{q} \beta_i x_{t-i} + \epsilon_t, \qquad (1)$$

where q is a prescribed positive number representing the lag length, $\{\beta_i : i = 0, 1, \ldots, q\}$ are the unknown lag coefficients and ϵ_t is a random variable with zero mean and constant variance. The issue of the q selection depends on the data and may be decided with statistical means (see, for example, [15]:p.119). Adopting matrix notation, the unconstrained lag-distribution problem is to determine a vector $\beta = (\beta_0, \beta_1, \ldots, \beta_q)^T$ that minimizes

$$F(\beta) = (y - X\beta)^T (y - X\beta), \qquad (2)$$

where $y = (y_{q+1}, y_{q+2}, \ldots, y_{q+n})^T$ is the n-vector whose components are time series observations and the $n \times (q+1)$ matrix X of current and lagged values of x_t is defined as

$$X = \begin{pmatrix} x_{q+1} & x_q & x_{q-1} & \cdots & x_1 \\ x_{q+2} & x_{q+1} & x_q & \cdots & x_2 \\ x_{q+3} & x_{q+2} & x_{q+1} & \cdots & x_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{q+n} & x_{q+n-1} & x_{q+n-2} & \cdots & x_n \end{pmatrix}.$$

Note that the components of y in (2) correspond to the last n observations of the time series data $y_t, t = 1, 2, \ldots, q, q+1, \ldots, y_{q+n}$, because we lose q degrees of freedom due to (1).

The unconstrained estimate of β, for a full rank X, is

$$\tilde{\beta} = (X^T X)^{-1} X^T y. \qquad (3)$$

The main drawback with this direct least-squares estimation of β is that often there is high multicollinearity among the x_t's giving a notoriously ill-posed inverse problem, which results in imprecise estimation for the β. If, however, avoid severe distortions in the calculation of the true lag distribution, then there appear discernible patterns in the unconstrained estimate, which are affected by the nature of the observations.

So far there have been several suggestions in the literature to put some structure on the β_i's in (1). They all

*Manuscript received April 20, 2008. Professor I. C. Demetriou is partially supported by research grand ELKE 70/4/4729 of the University of Athens. E. E. Vassiliou is a Ph.D. student supported by research grand Heraklitus 2003 of the University of Athens. The authors feel grateful to them. Unit of Mathematics and Informatics, Department of Economics, University of Athens, 8 Pesmazoglou street, Athens 10559, Greece. Email: demetri@econ.uoa.gr, evagvasil@econ.uoa.gr

impose some a priori structure on the form of the lag, in order to combine prior and sample information in the estimation of the regression coefficients. A popular approach is the Almon polynomial lag distribution [1]. In this technique the $q+1$ coefficients of the lagged variables are assumed to lie on a polynomial whose order is predetermined. Shiller's method [19], as a variant of this model, assumes that the coefficients of the lagged variable lie close to, rather than on, a polynomial. An alternative approach, as for example the Koyck's geometric distributed lag [16], represents the lag operator (1) by the ratio of two polynomials. Here, the coefficients are constrained to decline exponentially as the length of the lag increases. More models are found in [5], [17], [20], [12], [11], [18] etc. All these models assume rather arbitrarily that the underlying function of the lag coefficients can be approximated closely by a form that depends on a few parameters. However, over the years, literature on the subject agrees that some weak representation of the lag coefficients is a sensible requirement for a satisfactory model estimation (see, for example, [11], [18] and references therein).

In this paper a procedure is suggested for estimating lag coefficients by minimizing (2) subject to the conditions that the lag coefficients $\beta_0, \beta_1, \ldots, \beta_q$ have at most k monotonic sections, where k is a prescribed positive number. An advantage of this approach to lag-coefficient estimation is that we are making the least change to the coefficients that gives properties that occur to a wide range of underlying models. The user may try several values of k if a particular choice does not suggest itself. The constraints on $\beta_0, \beta_1, \ldots, \beta_q$ avoid any parameterization and provide a rather weak though systematic representation of the prior knowledge, as we are going to explain in Section 3.

In the case when $k=1$ the constraints on $\beta_0, \beta_1, \ldots, \beta_q$ are all linear, and they are

$$\beta_0 \geq \beta_1 \geq \cdots \geq \beta_q, \tag{4}$$

if we require monotonically decreasing coefficients, and

$$\beta_0 \leq \beta_1 \leq \cdots \leq \beta_q, \tag{5}$$

if we require monotonically increasing coefficients. Hence and in view of the quadratic function (2), the calculation of β is a convex quadratic programming problem. Thus several general algorithms are available for obtaining the solution (see, for example, [6]). It is worth mentioning that the problem subject to the monotonic decreasing constraints (4) generalizes the method of [5], where the coefficients β_i are imposed to decline arithmetically.

When $k>1$ it is usually quite difficult to develop efficient optimization algorithms for calculating an optimal β. One of the main difficulties is the combinatorial nature of the constraints that defines a nonconvex calculation with very many local minima. However, we address an alternative form of the problem and develop an iterative algorithm that implements a descent method with piecewise monotonicity constraints on the lag coefficients, which attempts to minimize (2). The iterative algorithm and its convergence are presented in Section 2. The piecewise monotonicity problem and its use in distributed lag modelling are discussed in Section 3. An example of an application of our method on real data is presented in Section 4. Some concluding remarks are given in Section 5. The Fortran program that implements our algorithm for distributed-lag estimation consists of about 3000 lines including comments, which gives an idea of the size of the required calculation.

2 The algorithm and its convergence

We develop an algorithm that processes the lag coefficients iteratively. It starts from an initial estimate $\beta^{(0)}$ of β that satisfies the constraints and generates a sequence of estimates $\{\beta^{(j)} : j = 1, 2, 3, \ldots\}$ to β in two phases. In the first phase it takes a descent direction from the current estimate to a new estimate of β. In the second phase it conveys "prior knowledge" to the calculation through the replacement of the new estimate by its best piecewise monotonic approximation. The contraction mapping theorem is used as a basis for establishing convergence.

In the first phase, the algorithm calculates a new estimate of the form

$$\beta^{(j+1)} = \beta^{(j)} + \alpha_j d^{(j)}, \tag{6}$$

where α_j is a step-length and $d^{(j)}$ is the search direction

$$d^{(j)} = X^T (y - X\beta^{(j)}). \tag{7}$$

It is to be noted that the search direction calculation involves matrix X only multiplicatively, so ill-conditioning of X is irrelevant here. The step-length α_j with exact line search is calculated to minimize the convex function of one variable $F(\beta^{(j)} + \alpha d^{(j)})$.

Having calculated $\beta^{(j+1)}$, the algorithm proceeds to the second phase, which calculates a $(q+1)$-vector β that minimizes

$$g(\beta_0, \beta_1, \ldots, \beta_q) = \sum_{i=0}^{q} (\beta_i^{(j+1)} - \beta_i)^2 \tag{8}$$

subject to at most k monotonic sections in the components of β. Specifically the piecewise monotonicity constraints are

$$\left.\begin{array}{ll} \beta_{t_{m-1}} \leq \beta_{t_{m-1}+1} \leq \cdots \leq \beta_{t_m}, & \text{if } m \text{ is odd} \\ \beta_{t_{m-1}} \geq \beta_{t_{m-1}+1} \geq \cdots \geq \beta_{t_m}, & \text{if } m \text{ is even} \end{array}\right\}, \tag{9}$$

while the integers $\{t_m : m = 0, 1, \ldots, k\}$ satisfy the conditions

$$0 = t_0 \leq t_1 \leq \cdots \leq t_k = q. \tag{10}$$

The integers $\{t_m : m = 1, 2, \ldots, k-1\}$, namely the indices of the turning points of the estimated components of β, are not known in advance and they are variables in the optimization calculation of the second phase. This raises the number of combinations of integer variables to about $O(q^k)$, but fortunately the piecewise monotonic problem is solved by [4] in only $O(q^2 + kq\, \log_2 q)$ computer operations. The main properties for this excellent complexity of the calculation are presented in Section 3.

The algorithm finishes if

$$\|\beta - \beta^{(j)}\|_2 / \|\beta\|_2 \leq \varepsilon \tag{11}$$

where ε is a small positive tolerance. This test is applied at every estimate $\beta^{(j+1)}$ including the first iteration as well. When the test (11) fails, then the algorithm replaces $\beta^{(j+1)}$ by its best piecewise monotonic vector β, increases j by one and branches to the beginning of the first phase in order to calculate at least one new vector in the sequence $\{\beta^{(j)} : j = 1, 2, 3, \ldots\}$. This gives the following algorithm.

Algorithm 1 **($k > 1$)**
Step 0 Set $j = 0$ and $\beta^{(0)} = 0$.
Step 1 Set $d^{(j)} = X^T(y - X\beta^{(j)})$.
Step 2 Calculate α_j and set $\beta^{(j+1)} = \beta^{(j)} + \alpha_j d^{(j)}$.
Step 3 By employing Algorithm 2 of [4] calculate β, namely a least squares approximation with k monotonic sections to $\beta^{(j+1)}$.
Step 4 If criterion (11) is satisfied then quit, otherwise replace $\beta^{(j+1)}$ by β, increase j by one and branch to Step 1. ■

Theorem 1 Algorithm 1 meets the termination condition (11) for some finite integer j.

Proof An outline of the proof of the convergence of Algorithm 1 is as follows. At Step 0 the starting vector $\beta^{(0)} = 0$ is not restrictive. If a more appropriate initial guess to β that satisfies the constraints (9) is available, then set this guess to $\beta^{(0)}$. Step 1 calculates the search direction. Step 2 calculates the step-length and obtains the estimate $\beta^{(j+1)}$. Step 3 calculates a least squares approximation with k monotonic sections to $\beta^{(j+1)}$. The algorithm either terminates at Step 4 or it sets this approximation vector to $\beta^{(j+1)}$ and then branches to Step 1. If we drop Step 3, which provides the piecewise monotonic approximation to current β, the remaining steps provide an iteration that converges to the minimum norm solution of (2). Since this problem has already been solved for an appropriate choice of α (see, for example, [9], [14]), and since the piecewise monotonicity constraint further restricts the solution, Algorithm 1 has to converge. ■

3 The piecewise monotonicity model

In this section we discuss some properties of the piecewise monotonicity model that is employed by Step 3 of Algorithm 1. Besides that the calculation of the unconstrained minimum of (2) due to (3) is highly ill-conditioned, the unconstrained minimum allows so much freedom in the calculation of β, that model (1) is almost useless in any estimation process.

We, instead, take the view that the calculation should make the smallest change to the current estimate of β that is necessary to satisfy constraints (9). The rationale for this choice is as follows. The sequence of the coefficients $\{\beta_i^{(j)} : i = 0, 1, \ldots, q\}$ may be attended as measurements of an unknown function. Due to errors of measurement in the time series data $y_t, t = 1, 2, \ldots, y_{q+n}$, it is possible that the number of turning points in $\beta_i^{(j)}$ is substantially larger than the the number of turning points in the true function values. Then the number of turning points in $\beta_i^{(j)}$ may suggest that it would be advantageous to smooth these estimates by requiring a certain number of monotonic sections. Therefore, given a positive integer $k < q$, we seek a $(q+1)$-vector β that is closest to $\beta^{(j)}$ in the least squares sense, subject to the condition that the components of β consist of at most k monotonic sections. By specifying that the first monotonic section is increasing, we obtain the constraints (9), but the user may well define it to be decreasing.

This approximation process is a projection, because if $\beta^{(j)}$ satisfies the constraints then $\beta = \beta^{(j)}$. Therefore if $\beta^{(j)}$ consists of more than k monotonic sections, as it is usually expected in practice, then the constraints prevent the equation $\beta = \beta^{(j)}$, so $\{t_m : m = 1, 2, \ldots, k-1\}$ are all different. Further at the turning points of a best fit, the equations $\beta_{t_m} = \beta_{t_m}^{(j)}, m = 1, 2, \ldots, k-1$ hold, which allows the search for the t_m's among the $\beta_i^{(j)}$ indices and reduces the amount of required computation. The most important property, however, is that the monotonic sections in a best piecewise monotonic fit are distinct. Indeed, the components $\{\beta_i : i = t_{m-1}, t_{m-1}+1, \ldots, t_m\}$ on $[t_{m-1}, t_m]$ minimize the sum of the squares $\sum_{i=t_{m-1}}^{t_m} (\beta_i^{(j)} - \beta_i)^2$ subject to the constraints $\beta_i \leq \beta_{i+1}$, $i = t_{m-1}, \ldots, t_m - 1$, if m is odd, and subject to the constraints $\beta_i \geq \beta_{i+1}$, $i = t_{m-1}, \ldots, t_m - 1$, if m is even. In the former case the sequence $\{\beta_i : i = t_{m-1}, t_{m-1}+1, \ldots, t_m\}$ is the best monotonic increasing fit to $\{\beta_i^{(j)} : i = t_{m-1}, t_{m-1}+1, \ldots, t_m\}$ and on the latter case the best monotonic decreasing one. Therefore, provided that $\{t_m : m = 1, 2, \ldots, k-1\}$ are known, the components of β can be generated by solving a separate monotonic problem on each section $[t_{m-1}, t_m]$ in the cost of only $O(t_m - t_{m-1})$ computer operations. It follows that an optimal fit β associated with the integer variables $\{t_m : m = 1, 2, \ldots, k-1\}$ can split at t_{k-1} into two optimal sections. One section that provides a best fit on $[t_0, t_{k-1}]$, which in fact is similar to β with one monotonic section less, and one section on $[t_{k-1}, t_k]$ that is a single monotonic fit to the remaining data. Hence the optimization problem can be replaced by a problem, whose variables are the positions of the turning points

PCE	GDP	$k = 1$	$k = 2$	$k = 4$	PCE	GDP	$k = 1$	$k = 2$	$k = 4$
661.4	865.2				2310.5	3652.7	2382.3	2361.6	2362.5
626.1	790.7				2396.4	3765.4	2486.1	2465.6	2466.2
606.9	739.9				2451.9	3771.9	2560.0	2543.8	2543.6
553.0	643.7				2545.5	3898.6	2656.7	2641.3	2641.7
541.0	635.5				2701.3	4105.0	2770.8	2764.7	2766.3
579.3	704.2				2833.8	4341.5	2893.6	2901.2	2903.7
614.8	766.9				2812.3	4319.6	2952.9	2962.5	2960.2
677.0	866.6	573.5	594.5	594.4	2876.9	4311.2	3008.3	3004.6	3000.8
702.0	911.1	584.8	595.2	594.2	3035.5	4540.9	3111.7	3099.0	3099.5
690.7	879.7	585.2	584.3	583.0	3164.1	4750.5	3218.1	3219.8	3222.8
729.1	950.7	612.8	599.9	600.4	3303.1	5015.0	3352.4	3367.1	3369.8
767.1	1034.1	655.0	644.4	645.9	3383.4	5173.4	3469.3	3482.4	3480.7
821.9	1211.1	727.2	724.0	727.1	3374.1	5161.7	3540.0	3542.3	3538.3
803.1	1435.4	818.8	818.3	821.1	3422.2	5291.7	3629.3	3611.4	3610.5
826.1	1670.9	924.4	921.3	923.8	3470.3	5189.3	3670.4	3656.3	3653.8
850.2	1806.5	1014.1	1000.0	1001.2	3668.6	5423.8	3790.2	3788.0	3789.9
902.7	1786.3	1070.6	1039.3	1038.5	3863.3	5813.6	3949.5	3966.6	3969.8
1012.9	1589.4	1083.3	1040.7	1038.0	4064.0	6053.7	4083.4	4122.1	4123.1
1031.6	1574.5	1123.5	1079.4	1080.0	4228.9	6263.6	4208.0	4233.3	4231.5
1054.4	1643.2	1178.1	1159.6	1162.4	4369.8	6475.1	4336.6	4332.2	4329.4
1083.5	1634.6	1205.7	1220.7	1222.0	4546.9	6742.7	4494.7	4478.4	4479.6
1152.8	1777.3	1255.5	1281.4	1281.4	4675.0	6981.4	4654.9	4634.1	4636.5
1171.2	1915.0	1297.6	1323.1	1322.0	4770.3	7112.5	4811.5	4797.7	4797.4
1208.2	1988.3	1323.6	1341.5	1341.0	4778.4	7100.5	4924.8	4915.0	4911.5
1265.7	2079.5	1360.6	1359.4	1359.1	4934.8	7336.6	5073.1	5060.4	5060.1
1291.4	2065.4	1390.8	1381.0	1379.5	5099.8	7532.7	5211.1	5215.4	5216.0
1385.5	2212.8	1461.9	1448.2	1449.8	5290.7	7835.5	5375.0	5395.9	5397.6
1425.4	2255.8	1512.0	1503.3	1504.0	5433.5	8031.7	5518.3	5544.9	5542.8
1460.7	2301.1	1566.2	1564.6	1564.8	5619.4	8328.9	5682.2	5700.6	5699.1
1472.3	2279.2	1596.7	1595.4	1593.6	5831.8	8703.5	5869.7	5878.1	5878.8
1554.6	2441.3	1662.8	1659.6	1660.4	6125.8	9066.9	6071.1	6073.7	6075.2
1597.4	2501.8	1709.3	1715.2	1715.7	6438.6	9470.3	6308.6	6307.7	6309.4
1630.3	2560.0	1753.1	1759.8	1759.7	6739.4	9817.0	6543.2	6533.8	6534.0
1711.1	2715.2	1826.4	1829.7	1829.6	6910.4	9890.7	6719.1	6704.1	6701.0
1781.6	2834.0	1891.0	1893.6	1893.9	7099.3	10048.8	6900.6	6871.2	6869.0
1888.4	2998.6	1972.5	1973.2	1974.7	7295.3	10301.0	7103.3	7080.5	7080.5
2007.7	3191.1	2069.2	2063.8	2064.9	7577.1	10703.5	7340.9	7345.1	7348.3
2121.8	3399.1	2184.5	2178.7	2180.2	7841.2	11048.6	7565.9	7593.6	7594.4
2185.0	3484.6	2272.1	2261.5	2261.7	8091.4	11415.3	7795.1	7826.0	7824.1

Table 1: The values of GDP and PCE for U.S.A. during the years 1929-2006 and the least squares estimates to PCE from GDP when the lag coefficients consist of one, two and four monotonic sections.

of the required fit and which is amenable to dynamic programming. The implementation of this idea includes several options that are considered by [4] and [2], while a versatile computer code has been written by [3].

4 An example on consumption data

To illustrate our method we present an application on real annual macroeconomic data derived from the Bureau of Economic Analysis of the U.S. Department of Commerce for the period 1/1/1929 - 1/1/2006. The dependent variable is the Real Personal Consumption Expenditures (PCE) and the independent variable is the Real Gross Domestic Product (GDP) for U.S.A., both measured in billions of chained 2000 dollars. The data of our application are given explicitly in the relevant columns of Table 1. We assume that a change in the GDP will affect not only current consumption, but also future consumption for seven time periods. Therefore we estimate the coefficients of the distributed-lag model with lag length $q = 7$ subject to the piecewise monotonicity constraints (9) on the components of β by considering separately the

ISBN:978-988-17012-3-7

cases $k = 1$ and $k > 1$.

	$k = 1$	$k = 2$	$k = 4$	$\tilde{\beta}$
β_0	0.2527	0.2472	0.2606	0.3105
β_1	0.0688	0.0981	0.0871	0.0037
β_2	0.0688	0.0321	0.0282	0.0862
β_3	0.0688	0.0056	0.0019	-0.0455
β_4	0.0688	0.0631	0.0713	0.1112
β_5	0.0688	0.0938	0.0978	0.0915
β_6	0.0688	0.0938	0.0883	0.0707
β_7	0.0688	0.1095	0.1068	0.1127

Table 2: The estimated ($k = 1$, 2 and 4) and the unconstrained lag coefficients in Section 4.

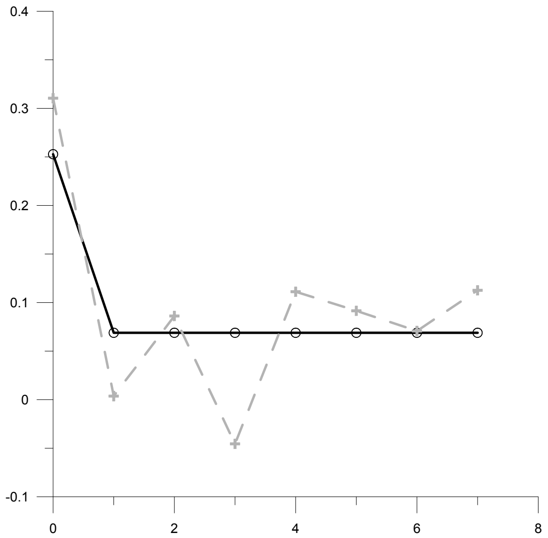

Figure 1: The unconstrained (+) and the monotonically decreasing (o) lag coefficients of Table 1.

a) Monotonic lag coefficients ($k = 1$)
We require the estimated lag coefficients to be monotonically decreasing, so the problem is to minimize (2) subject to the constraints (4). We have developed a special quadratic programming method for this problem that takes account of the fact that each of the constraint functions depends on only two adjacent components of β, but we do not enter into the details of our computation. The optimal lag coefficients are shown in the second column ($k = 1$) of Table 2, while the unconstrained lag coefficients due to (3) are shown in the fifth column ($\tilde{\beta}$) of the table. In Fig. 1 we display all these coefficients. On the one hand, the first unconstrained coefficient seems to be more significant than the others, but the fluctuation of the unconstrained coefficients make them inadequate to the estimation problem. On the other hand, the optimal monotonic decreasing estimates follow the main trend of the unconstrained coefficients and maintain the importance of the first coefficient. Therefore, with the monotonicity assumption, the resultant estimated values of PCE are given in the third column ($k = 1$) of Table 1 and displayed in Fig. 2 together with the provided GDP values. It is remarkable that the estimated PCE values via formula (1), which actually involves the GDP values, fall close to the observed PCE values. In particular, the current GDP value rather than past ones affects mainly the associated PCE value. Indeed, in view of the monotonically decreasing lag coefficients, GDP affects strongly consumption at the beginning of the lags, while its action subsequently is reduced and kept at a low level. Thus the constraints (4) provide a plausible choice for the lag coefficients that leads to a satisfactory estimation of the true PCE values.

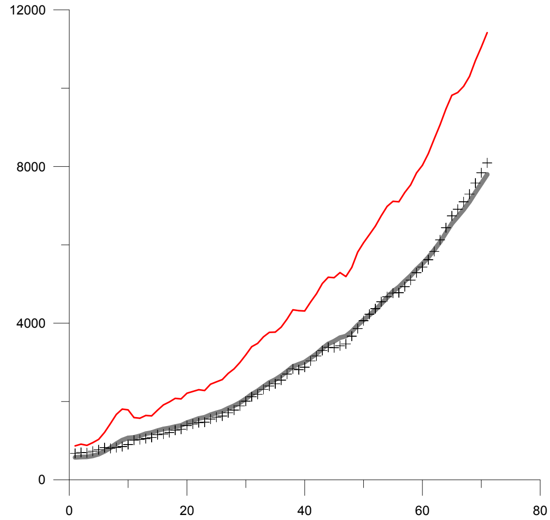

Figure 2: Least squares estimation (grey line) to the PCE values (+) with the monotonically decreasing lag coefficients ($k = 1$) of Table 2 on the GDP values (thin line).

b) Piecewise monotonic lag coefficients ($k > 1$)
In order to illustrate some features of Algorithm 1 we performed two experiments. In the first experiment we calculated the lag coefficients by employing Algorithm 1 with $k = 2$ and $k = 4$, while the first monotonic section was let to be decreasing. The tolerance for the termination criterion (11) in Step 4 has been set to 10^{-4} and the estimated values of β are shown in the third and fourth column of Table 2 and displayed in Fig. 3 together with the unconstrained lag coefficients. As can be seen, not only these estimates are smoother than the unconstrained lag coefficients, but also they provide satisfactory fits to $\tilde{\beta}$, although $\tilde{\beta}$ is not explicitly available. The $k = 2$ case has introduced one turning point at β_3, but the extra turning points of the $k = 4$ case that were added at the 5th and 6th data point seem to add little to the fit obtained when $k = 2$ in combating the unconstrained coefficients trend. Therefore it would have been sensible to require just two monotonic sections for the lag coefficients estimation, but this was realized only after the $k = 4$ case had been considered. The result suggests that the user may apply Algorithm 1 with increasing values

ISBN:978-988-17012-3-7

of k until a satisfactory estimation is obtained. Further, the corresponding estimated values of PCE are shown in columns $k = 2$ and $k = 4$ of Table 1 and displayed in Fig. 4 together with the GDP values. Again, it is remarked that the estimated PCE values fall close to the observed PCE values providing very satisfactory estimations of the true PCE values.

In the second experiment we computed successive approximations to the lag coefficients by employing Algorithm 1 with $k = 2$, for $\varepsilon = 10^{-3}, 10^{-4}$ and 10^{-5}. The calculated values of β are shown in Table 3 and displayed in Fig. 5. All three approximations capture the pattern of the unconstrained lag coefficients. Moreover, the smaller the value of ε, the closer the approximation components are to the unconstrained coefficients, while these components cannot be worse than the unconstrained lag coefficients, due to the employed piecewise monotonicity constraints. Thus the user may decide to monitor the smoothing performance of the method by means of the tolerance magnitude in (11). This is a helpful consideration, because due to degeneracy or near degeneracy of matrix X, a line search, sometimes, may have to choose a tiny step-length α_j, which implies that the algorithm may make slow progress to the solution.

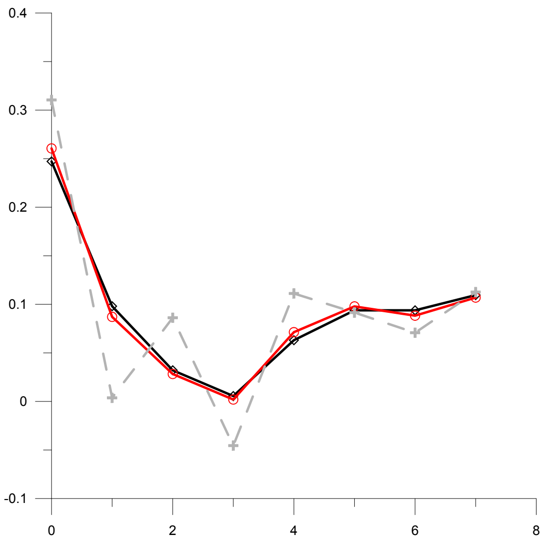

Figure 3: The unconstrained (+) and the piecewise monotonic lag coefficients with $k = 2$ ($\diamond$) and $k = 4$ ($\circ$) of Table 2.

5 Conclusions and future work

We have developed a new method for estimating distributed-lag coefficients in time series data subject to the condition that the coefficient estimates are composed of a certain number of monotonic sections.

The method seems to be both effective in computation and competent to its modelling task. Three distinctive features of this process are to be noted: (1) The process is designed to overcome the multicollinearity problem that frequently occurs in practice, (2) the piecewise monotonicity model provides a rather weak, nonetheless realistic representation of the lag coefficients and, (3) the calculation benefits from the excellent complexity of the piecewise monotonicity method. In particular, the choice of the prior knowledge parameter k gives the time series estimation valuable flexibility.

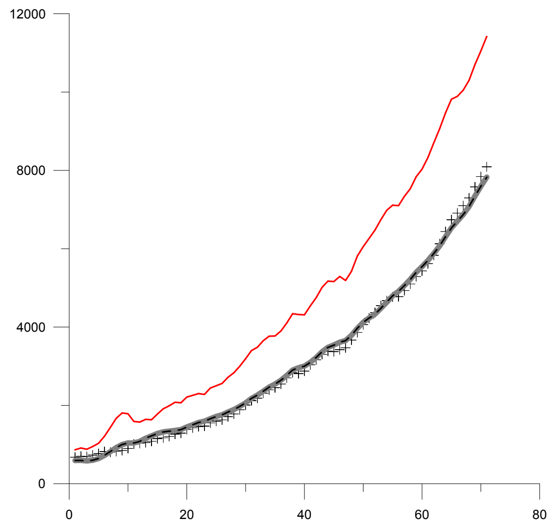

Figure 4: Least squares estimations (grey and dashed line) to the PCE values (+) when the lag coefficients on the GDP values (thin line) allow $k = 2$ and $k = 4$ monotonic sections respectively.

	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-5}$
β_0	0.2129	0.2472	0.2964
β_1	0.1310	0.0981	0.0417
β_2	0.0595	0.0321	0.0417
β_3	0.0182	0.0056	-0.0096
β_4	0.0344	0.0631	0.0893
β_5	0.0690	0.0938	0.0893
β_6	0.0979	0.0938	0.0893
β_7	0.1198	0.1095	0.1030

Table 3: The lag coefficients with $k = 2$ monotonic sections, for $\varepsilon = 10^{-3}, 10^{-4}$ and 10^{-5} in Step 4 of Algorithm 1.

For the piecewise monotonic model we have used a Fortran package that has been developed recently [3], which indeed is a major part of our calculation. For the special problem that minimizes (2) subject to the monotonic constraints (4) we have used Fortran codes developed by one of the authors (EEV).

The calculations performed so far on real data show that our method overcomes some severe shortcomings of traditional lag estimation techniques. Still, there is plenty of room for much empirical analysis on this method. The

ISBN:978-988-17012-3-7

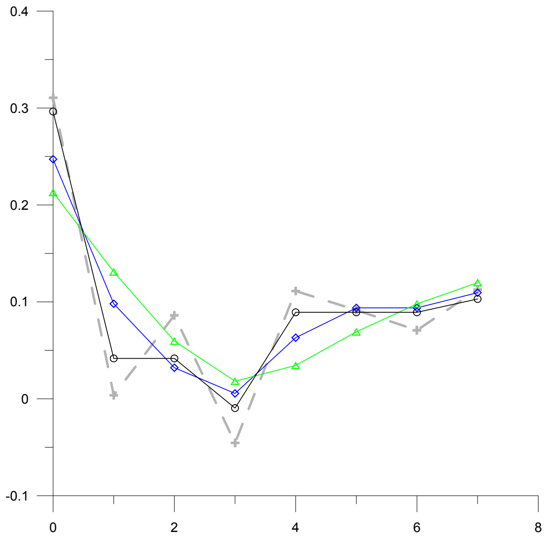

Figure 5: Successive approximations to lag coefficients with $k = 2$ monotonic sections, for $\varepsilon = 10^{-3}$ ($\triangle$), 10^{-4} ($\diamond$) and 10^{-5} ($\circ$). The unconstrained coefficients are denoted by "+".

algorithm is sufficiently fast for interactive computation, but a possible improvement is to introduce some features of the conjugate gradient technique. It is expected that the algorithm will find useful applications to real problems, so work is under way in order to provide a Fortran package that is suitable for submission to an international software library.

References

[1] Almon, S., "The Distributed Lag between Capital Appropriations and Expenditures", *Econometrica*, V33, N1, pp. 178-196, 1965.

[2] Demetriou, I. C., "Discrete Piecewise Monotonic Approximation by a Strictly Convex Distance Function", *Mathematics of Computation*, V64, N209, pp. 157-180, 1/95.

[3] Demetriou, I. C., "Algorithm 863: L2WPMA, a Fortran 77 Package for Weighted Least-squares Piecewise Monotonic Data Approximation", *ACM Transactions on Mathematical Software*, V33, N1, pp. 1-19, 3/07.

[4] Demetriou, I. C., Powell, M. J. D., "Least Squares Smoothing of Univariate Data to Achieve Piecewise Monotonicity", *IMA Journal of Numerical Analysis*, V11, pp. 411-432, 1991.

[5] Fisher, I., "Note on a Short-Cut Method for Calculating Distributed Lags", *International Statistical Institute Bulletin*, pp. 323-327, 1937.

[6] Fletcher, R., *Practical Methods of Optimization, Second Edition*, Chichester: J. Wiley and Sons, 2003.

[7] Gershenfeld, N., *The Nature of Mathematical Modelling*, Cambridge: Cambridge University Press, 1999.

[8] Gray, R. M., *Toeplitz and Circulant Matrices: A Review*, Stanford: Department of Engineering, http://ee.stanford.edu/~gray/toeplitz.pdf, 2006.

[9] Golub, G., van Loan, C. F., *Matrix Computations, Second Edition*, Baltimore and London: The John Hopkins University Press, 1989.

[10] Griliches, Z., "Distributed Lags: A Survey", *Econometrica*, V35, N1, pp. 16-49, 1/67.

[11] Hannan, E. J., "The Estimation of Relations Involving Distributed Lags", *Econometrica*, V33, N1, pp. 206-224, 1/65.

[12] Jorgenson, D., "Rational Distributed Lag Functions", *Econometrica*, V34, N1, pp. 135-149, 1/66.

[13] Kailath, T. (editor), *Linear Least-Squares Estimation*, Stroudsburg, Pennsylvamia: Dowden, Hutchinson and Ross, Inc., Benchmark Papers in Electrical Engineering and Computer Science, V17, 1977.

[14] Katsaggelos, A. K., "Iterative Image Restoration Algorithms", *Optical Engineering*, V28, N7, pp. 735-748, 1989.

[15] Koop, G., *Analysis of Economic Data*, Chichester: J. Wiley and Sons, 2000.

[16] Koyck, L., *Distributed Lags and Investment Analysis*, Amsterdam: North-Holand Pub. Co., 1954.

[17] DeLeeuw, F., "The Demand for Capital Goods by Manufactures: A Study of Quartely Time Series", *Econometrica*, 1962.

[18] Maddala, G. S., *Econometrics*, London: McGraw-Hill Book Co., 1977.

[19] Shiller, R. J., "A Distributed Lag Estimator Derived from Smothness Priors", *Econometrica*, V41, N4, pp. 775-788, 1973.

[20] Solow, R. M., "On a Family of Lag Distributions", *Econometrica*, V28, N2, pp. 393-406, 4/60.

[21] Tsay, R. S., *Analysis of Financial Time Series*, New York: J. Wiley and Sons, 2002.

[22] Zellner, A., Geisel, M. S., "Analysis of Distributed Lag Models with Applications to Consumption Function Estimation", *Econometrica*, V38, N6, pp. 865-888, 11/70.