# Examples on lag distributed models subject to nonnegative divided differences of orders 2, 3 and 4

*Evangelos E. Vassiliou*
Department of Economics, University of Athens

**Keywords:** approximation, divided differences, lag distributed models, least squares, quadratic programming, r-convexity.
**AMS:** 60J80

**Abstract**

We consider noisy measurements from a time series that follow a linearly distributed lag model. It is usual to assume that the lag coefficients lie on some curve and then specify the curve by a least squares calculation. However, we define the r-th order smoothness priors by requiring nonnegative divided differences of order r for the lag coefficients. Such priors do not imply any parameterization of the lag curve and provide a more accurate representation of the prior knowledge. For the calculation of the solution we propose an algorithm that gives the least squares change to the data subject to nonnegative divided differences of the lag coefficients of order r, where r is a prescribed positive integer. The problem is a strictly convex quadratic programming calculation, where each of the constraints functions depends on r+1 adjacent components of the smoothed values of the lag coefficients. We take account of this special structure and use a special active set method that is more efficient than general quadratic programming algorithms. In fact we construct a basis that reduces the equality-constrained minimization calculations that appear during the quadratic programming iterations to unconstrained minimization ones, which depend on much fewer variables. Finally, we present an example that illustrates our approach.

## 1. Introduction

A time series $y_t$ is said to follow a linearly distributed lag model on time time series $x_t$ if

$$y_t = \sum_{i=0}^{\lambda-1} \beta_i x_{t-i} + \epsilon_t, \tag{1}$$

where $\lambda$ a known constant representing the lag length, $\beta_i$ the unknown coefficients and $\epsilon_t$ a stationary process with zero mean.

The problem is, given $y = \{y_t : i = 1, 2, \ldots, n\}$ and given a matrix $X$ of current and lagged values of $x_t$ to obtain an estimate of $\beta = \{\beta_i : i = 0, 1, \ldots, \lambda - 1\}$

There are a number of ways of tackling this problem. However, they all have the same objective, which is to impose some a priori structure on the form of the lag, thereby reducing the number of parameters to be estimated. The most popular approach is the Almon polynomial lag distribution [1]. In this technique the $\lambda$ coefficients of the lagged explanatory variables are assumed to lie on a polynomial of order $r$. Shiller's distributed lag [3] is a variant of this in which the restrictions are stochastic, incorporated via the mixed estimation technique. In this method it is assumed that the coefficients of the lagged explanatory variable lie close to, rather than on, a polynomial. An alternative approach, which fits more naturally into a time series framework, is to represent the lag structure by the ratio of two polynomials in the lag operator. This is known as a rational distributed lag. One of the simplest examples of a rational lag structure is the Koyck or geometric distributed lag. In this model the coefficients are constrained to decline exponentially as the length of the lag increases.

In this paper we assume that the underlying function of the lag coefficients is $r$-times differentiable with nonnegative $r$th derivative (r-convex). Thus the $r$th order divided differences of the $\beta_i$'s are also nonnegative. However, we define the $r$th order smoothness priors by requiring nonnegative divided differences of order $r$ for the lag coefficients. Such priors do not imply any parameterization of the lag curve and provide a more accurate representation of the prior knowledge. Therefore, it seems appropriate to modify the data so that the $r$th order divided differences of the lag coefficients allow no sign changes. This condition when $r = 1$ implies monotonicity, when $r = 2$ implies convexity and when $r = 3$ implies an inflection point away from which the underlying function is concave and convex.

## 2. Estimation of Coefficients

In the general case, $r$ is a positive integer, much smaller than $\lambda$, and we address the problem of calculating numbers $\{\beta_i : i = 0, 1, \ldots, \lambda - 1\}$ from the data that minimize the objective function

$$q(\beta) = \frac{1}{2}\beta^T 2X^T X\beta - 2y^T X\beta \tag{2}$$

subject to the constraints that the divided differences of order $r$ for the lag coefficients are nonnegative. For the calculation of the solution we propose an algorithm that gives the least squares change to the data subject to nonnegative divided differences of the lag coefficients of order $r$, where $r$ is a prescribed positive integer.

In this paper we consider the case where the abscissae are equally spaced, that is the most common case in time series data, and obtain some computational advantages. Indeed now the constraints are

$$\mathbf{A}_r^T\beta \geq 0, \tag{3}$$

where $\mathbf{A}_r$ is the $\lambda \times (\lambda - r)$ rectangular matrix, whose elements $(a_r)_{ij}$ are defined by the relation

$$(a_r)_{ij} = \begin{cases} (-1)^{r+j-i} \begin{pmatrix} r \\ i-j \end{pmatrix}, & j \leq i \leq j+r, \quad j = 1, 2, \ldots, \lambda - r, \\ 0, & \text{otherwise}. \end{cases} \tag{4}$$

The problem is a strictly convex quadratic programming calculation, where each of the constraints functions depends on $r+1$ adjacent components of the smoothed values of the lag coefficients. We take account of this special structure and we use a special active set method, suggested by Demetriou, Lipitakis and Vassiliou [2] that is more efficient than general quadratic programming algorithms. In fact, we construct a linear transformation in the linear space of variables, that reduces the equality-constrained minimization calculations that appear during the quadratic programming iterations to unconstrained minimization ones, which depend on much fewer variables. Associated to this calculation is the calculation of Lagrange multipliers. They are uniquely defined by the first order optimality condition

$$2X^T(X\beta - y) = \sum_{i \in \mathcal{A}} l_i a_i \tag{5}$$

where $\mathcal{A}$ is a subset of the constraints indices $\{1, 2, \ldots, \lambda - r\}$, $a_i$ is the $i$th column of $\mathbf{A}_r$ and $l_i$ are the corresponding Lagrange multipliers.

Actually, we suggest an alternative way for estimating the lag coefficients in a lag distributed model. What our nonnegative divided differences method gives is a piecewise polynomial fit, where the joins are found automatically by the process. As a consequence, we relax the constraints and gain flexibility in comparison to Almon approach.

## 3. An illustrative example

To illustrate our method we present an application on real quarterly macro data derived from Eurostat for the period 1995-2006. The dependent variable is the Consumption and the independent variable is the Gross Domestic Product (GDP) for EU25 countries, both in constant prices. We assume that a change in the GDP will affect not just the current consumption but also the future consumption for six time periods.

We estimate the coefficients of the lag distributed model with lag length $\lambda = 6$, by requiring the divided differences of order $r = 2, 3, 4$ of lag coefficients to be nonnegative. In Figs. 1, 3, 5 we let $r = 2, 3, 4$ respectively, and display the piecewise linear interpolant to the estimated values of consumption, denoted in all the figures by a red curve. In addition, in Figs. 2, 4, 6 we see the pattern of the estimated lag coefficients for $r = 2, 3, 4$ respectively.

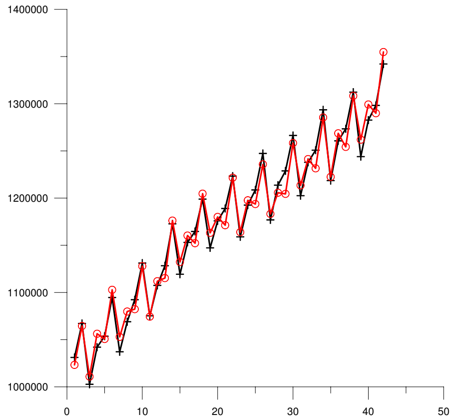


Figure 1: Best least square approximation for $r = 2$. The data are denoted by "+" and the estimated values by "$o$".

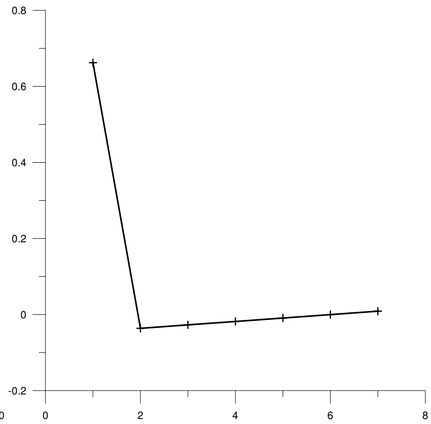


Figure 2: Estimated lag coefficients for $r = 2$.

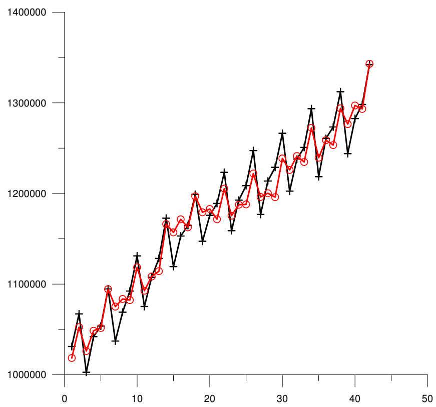


Figure 3: As in Fig. 1, but $r = 3$.

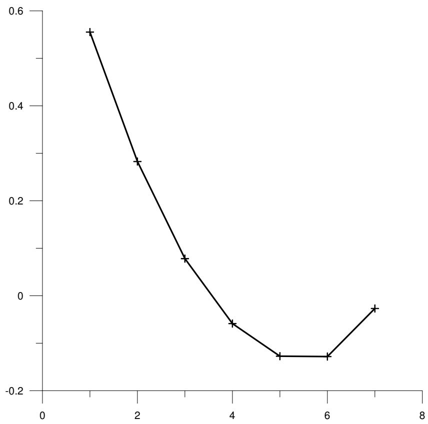


Figure 4: As in Fig. 2, but $r = 3$.

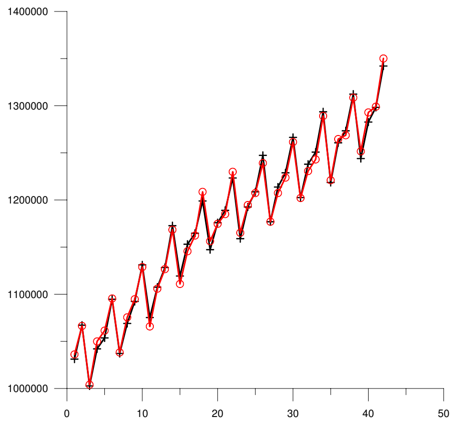


Figure 5: As in Fig. 1, but $r = 4$.

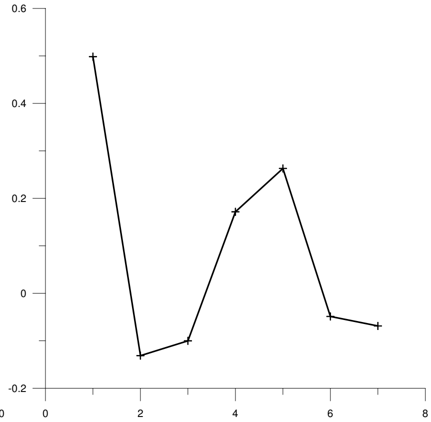


Figure 6: As in Fig. 2, but $r = 4$.

**Acknowledgements:** The work of the author is supported by research grand 70/7/ 7145 (Heracleitus) of the university of Athens, which is co-financed within Op. Education by the European Social Fund and National Resources. We feel grateful to them all. Finally, the author wishes to sincerely thank I. C. Demetriou for his strong support and guidance during this work.

## 4. Bibliography

[1] Almon, S. (1965). *The distributed lag between capital appropriations and expenditures.* Econometrica, Vol. 33, No. 1., 178-196.

[2] Demetriou I. C., Lipitakis E. A. & Vassiliou E. E. (2006). *Efficient use of Toeplitz Matrices for Least Squares Data Fitting by Nonnegative Differences.* Hermis-$\mu\pi$, Volume 6, 135-153.

[3] Shiller, R. J. (1973). *A distributed lag estimator derived from smothness priors.* Econometrica, Vol. 41, No. 4., 775-788.