International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

A Note on the Estimation of Polynomial Distributed Lags

Ledolter, J.

A NOTE ON THE ESTIMATION

OF POLYNOMIAL DISTRIBUTED LAGS

Johannes Ledolter

March 1976

A Note on the Estimation of Polynomial Distributed Lags

Johannes Ledolter

Methods of estimation of polynomial distributed lags in econometrics and procedures relating tree ring growth data to climatic and hydrologic data are shown to be equivalent to a method first described by R.A. Fisher in 1924 [3].

1. Estimation of Polynomial Distributed Lags:

Almon [1] considers the following model:

$$Y_i = x_i' \beta + \varepsilon_i \qquad p+1 \leq i \leq n \qquad (1.1)$$

$$\beta = T\alpha \qquad (1.2)$$

where:

$$x_i' = (x_i, x_{i-1}, \ldots, x_{i-p})$$

$$\beta' = (\beta_1, \beta_2, \ldots, \beta_{p+1})$$

$$\alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_{m+1}) \qquad \text{with } m << p$$

$$T = \begin{bmatrix} T_0(1) \cdots\cdots T_m(1) \\ T_0(2) \cdots\cdots T_m(2) \\ \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \\ T_0(p+1) \cdots T_m(p+1) \end{bmatrix}$$

where the $T_i(\tau)$ $0 \leq i \leq m$ are chosen such that

$$\beta_\tau = \alpha_1 + \alpha_2 \tau + \cdots + \alpha_{m+1} \tau^m \quad .$$

In matrix notation we get

$$y = X\beta + \varepsilon$$
$$= XT\alpha + \varepsilon \qquad (1.3)$$

where

$$\underset{\sim}{y}' = (y_{p+1}, y_{p+2}, \ldots, y_n)$$

are deviations from the mean

$$\underset{\sim}{\varepsilon}' = (\varepsilon_{p+1}, \varepsilon_{p+2}, \ldots, \varepsilon_n)$$

$$\underset{\sim}{X}' = [\underset{\sim}{x}_{p+1}, \underset{\sim}{x}_{p+2}, \ldots, \underset{\sim}{x}_n]$$

The disturbances $\underset{\sim}{\varepsilon}$ are assumed to be independent, centered around zero and with constant variance.

The least squares estimates for the parameters in the polynomial distributed lag model are given by

$$\underset{\sim}{\hat{\alpha}} = (T'X'XT)^{-1}T'X'\underset{\sim}{y} \tag{1.4}$$

$$\underset{\sim}{\hat{\beta}} = T\underset{\sim}{\hat{\alpha}} = T(T'X'XT)^{-1}T'X'\underset{\sim}{y}$$

It was shown by Cooper [2] that this direct approach of estimating the distributed lag model is equivalent to Almon's [1] approach using Langrangian interpolation techniques.

2. Fisher's Model

Fisher considered a similar problem as presented in section 1. He studied the influence of rainfall on the yield of wheat. More than 60 yearly observations on wheat yield at the same experimental plot and weekly averages of rainfall for this time span were available.

Fisher considered the model:

$$y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta} + \varepsilon_i \qquad 1 \leq i \leq n \tag{2.1}$$

$$\underset{\sim}{\beta} = T\underset{\sim}{\alpha} \tag{2.2}$$

$$\underset{\sim}{x}_i = T\underset{\sim}{\rho}_i \qquad 1 \leq i \leq n \tag{2.3}$$

where

$$\underset{\sim}{x}_i' = (x_{i1}, x_{i2}, \ldots, x_{ip+1})$$

are the rainfall data for the ith year.

$$\underset{\sim}{\beta}' = (\beta_1, \beta_2, \ldots, \beta_{p+1})$$

$$\underset{\sim}{\alpha}' = (\alpha_1, \alpha_2, \ldots, \alpha_{m+1}) \qquad \text{with } m << p$$

$$\underset{\sim}{\rho}_i' = (\rho_{i1}, \rho_{i2}, \ldots, \rho_{im+1}) \qquad 1 \le i \le n$$

$$T = \begin{bmatrix} T_0(1) & \cdot\ \cdot & T_m(1) \\ T_0(2) & \cdot\ \cdot & T_m(2) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ T_0(p+1) & \cdot\ \cdot & T_m(p+1) \end{bmatrix}$$

In matrix notation the model can be written as

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} = XT\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon} \tag{2.4}$$

$$X = \hat{X} + E$$

where

$$\hat{X}' = [\hat{\underset{\sim}{x}}_1, \hat{\underset{\sim}{x}}_2, \ldots, \hat{\underset{\sim}{x}}_n]$$

$$E' = [\underset{\sim}{x}_1 - \hat{\underset{\sim}{x}}_1, \underset{\sim}{x}_2 - \hat{\underset{\sim}{x}}_2, \ldots, \underset{\sim}{x}_n - \hat{\underset{\sim}{x}}_n]$$

and

$$\hat{\underset{\sim}{x}}_i = T\hat{\underset{\sim}{\rho}}_i = T(T'T)^{-1}T'\underset{\sim}{x}_i \qquad 1 \le i \le n \tag{2.5}$$

Thus, the model becomes

$$\underset{\sim}{y} = XT\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon} = \hat{X}T\underset{\sim}{\alpha} + ET\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$$

and since

$$ET\underset{\sim}{\alpha} = X[I-T(T'T)^{-1}T']T\underset{\sim}{\alpha} = O$$

$$\underset{\sim}{y} = \hat{X}T\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon} \tag{2.6}$$

The least squares estimate of α is given by

$$\hat{\underset{\sim}{\alpha}} = (T'\hat{X}'\hat{X}T)^{-1}T'\hat{X}'\underset{\sim}{y}$$

Substitution of (2.5) gives

$$\hat{\underset{\sim}{\alpha}} = (T'X'XT)^{-1}T'X\underset{\sim}{y} \qquad (2.7)$$

and

$$\hat{\underset{\sim}{\beta}} = T\hat{\underset{\sim}{\alpha}} = T(T'X'XT)^{-1}T'X\underset{\sim}{y}$$

which coincides with the expression in (1.4).

In his paper [3] Fisher also introduced the concept of orthogonal polynomials. Since at the time Fisher wrote this paper computations had to be done by hand, the introduction of orthogonal polynomials was of great help. It is easily shown that by using orthogonal polynomials

$$T'T = I \qquad \sum_{j=1}^{p+1} T_i(j)T_k(j) = \begin{cases} 1 & \text{for} \quad i = k \\ 0 & \text{for} \quad i \neq k \end{cases}$$

the estimate for $\underset{\sim}{\alpha}$ is derived by:

1) regressing $\underset{\sim}{x}_i$ on T for $1 \leq i \leq n$ to derive the co-efficients $\hat{\underset{\sim}{\rho}}_i$

2) regressing the y_i on the $\hat{\underset{\sim}{\rho}}_i$ to get $\hat{\underset{\sim}{\alpha}}$.

3. A Further Comment

Models of the type discussed in sections 1 and 2 are also applied to the study of tree ring data. Biological models for the relationships between variations in tree ring widths and variations in temperature and precipitation are discussed in [4,5]. These transfer function models between the yearly tree ring growth and climatic and hydrologic variables are derived in a method similar to the one described by Fisher.

References

[1] Almon, S. (1965): "The Distributed Lag between Capital
     Appropriations and Expenditures," Econometrica,
     Vol.33, 178.

[2] Cooper, J.P. (1972): "Two Approaches to Polynomial
     Distributed Lags Estimation: An Expository Note
     and Comment," American Statistician, June 1972, 32-35.

[3] Fisher, R.A. (1924): "The Influence of Rainfall on the
     Yield of Wheat at Rothamsted," Philosophical
     Transactions of the Royal Society in London, Series B,
     Vol.213 (1924), 89-142.

[4] Fritts, H.C. (1971): "Dendroclimatology and Dendroecology,"
     Quaternary Research, Vol.1, 419-449.

[5] Fritts, H.C. (1974): "Relationships of Ring Widths in
     Arid-Site Conifers to Variations in Monthly
     Temperature and Precipitation," Ecological Monograph,
     Vol.44, 411-440.