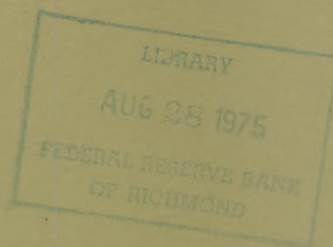


Federal Reserve Bank of Chicago, Staff  
Memoranda

A Series of Occasional Papers in Draft Form Prepared by Members  
of the Research Department for Review and Comment.

75-1



# Autoregressive Form of Time-Polynomial Extrapolation

Thomas A. Gittings

Federal Reserve Bank of Chicago



Research Paper No. 75-1

AUTOREGRESSIVE FORM OF  
TIME-POLYNOMIAL EXTRAPOLATION

By

Thomas A. Gittings  
Department of Research  
Federal Reserve Bank of Chicago

This paper is being circulated for purposes of discussion and comment. The contents should be regarded as preliminary and not for citation or quotation without permission of the author. The views expressed are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

## 1. Introduction

Some of the standard methods of forecasting future values of a time series that has a trend include treating the series as an A.R.I.M.A. process, using general exponential smoothing, or extrapolating least-squares fitted time-polynomials. This article primarily investigates some properties of this last method. Given the current and some past values of a single time series, the procedure of estimating the coefficients of a polynomial in time and extrapolating into the future is equivalent to a specific type of autoregressive process. The lag weights in the autoregressive equation are independent of the data and are functions only of the number of observations, the degree of the estimated time-polynomial, the lead time of the forecast, and the weighting function in the least-squares estimation.

This article is concerned with forecasting a time series which can be represented by a simple polynomial or exponential trend without a seasonal component. One method of forecasting such a variable is to form a stationary series from the data, usually by taking the difference between successive observations (or logarithms of the observations), and then estimate the coefficients for an autoregressive model. In a general A.R.I.M.A. model the equation also can include the deviations of some past observations from their estimated values. This approach is advocated especially by Wold [6], Whittle [5], and Box and Jenkins [1].

For cases where the time series has such a simple trend, an alternative approach is to estimate the time trend and extrapolate to get the forecasts for some future dates. This forecasting technique is equivalent to using certain deterministic autoregressive equations where the coefficients are independent of the data. The primary advantage of this later approach is its computational simplicity; forecasts can be made by just taking a weighted sum of a set of observations. The process of estimating the coefficients of a time-polynomial is embodied in these particular autoregressive coefficients.

In section 2 the autoregressive form for extrapolating a trend estimated by ordinary least-squares is derived and some interesting properties of these coefficients are noted. Section 3 presents the autoregressive schemes which correspond to single period forecasts using exponentially discounted least-squares and double exponential smoothing. Section 4 shows what constraints on the autoregressive coefficients are necessary if one wants to extrapolate any  $N$  degree time-polynomial trend. The constraints can be used to reduce the number of coefficients that must be specified or estimated in a autoregressive equation. Section 5 examines some properties of forecast functions which extrapolate trend and use the first difference of the observations. Section 6 considers two methods of extending the results for any forecast lead time and shows how to extrapolate any exponential trend.

## 2. ORDINARY LEAST-SQUARES

Extrapolation of trend in a single time series is appropriate when the variable satisfies the usual assumptions of a least-squares model. Given  $T$  observations from discrete, equispaced intervals of time, the time-polynomial model of degree  $N$  is

$$z_t = \beta_0 + \beta_1 t + \dots + \beta_N t^N + \varepsilon_t, \quad t = 1, 2, \dots, T. \quad (2.1)$$

The time origin has been set so that the first observation is for time period one. The residuals are assumed to be independently distributed with a constant variance and zero expectation. The number of observations must be greater than the degree of the time-polynomial.

In matrix notation the model is

$$Z = X\beta + \varepsilon \quad (2.2)$$

Where  $Z$  is the  $T \times 1$  vector whose  $t^{\text{th}}$  element is the observation for period  $t$ ,  $\beta$  is the  $(N+1) \times 1$  vector of coefficients that are to be estimated, and  $X$  is the  $T \times (N+1)$  matrix with  $X_{tn} = t^{n-1}$ . The ordinary least-squares estimates of the vector of coefficients is

$$\hat{\beta} = (X'X)^{-1} X'Z. \quad (2.3)$$

Using these estimated coefficients to extrapolate trend, the expected value of this variable in the next time period ( $T+1$ ) is

$$z_{T+1}^* = \tau \hat{\beta}, \quad (2.4)$$

where  $\tau$  is the  $1 \times (N+1)$  vector with  $\tau_n = (T+1)^{n-1}$ .

This forecast can be written as

$$z_{T+1}^* = \tau (X'X)^{-1} X'Z \quad (2.5)$$

or

$$z_{T+1}^* = \alpha Z = \alpha_1 z_1 + \alpha_2 z_2 + \dots + \alpha_T z_T \quad (2.6)$$

Where  $\alpha$  is the  $1 \times T$  vector  $\tau(X'X)^{-1}X'$ . Since the normal autoregressive form is

$$z_{T+1}^* = \theta_1 z_T + \theta_2 z_{T-1} + \dots + \theta_T z_1, \quad (2.7)$$

the  $t^{\text{th}}$  lag weight  $\theta_t$  must equal  $\alpha_{T+1-t}$ ,  $t = 1, 2, \dots, T$ .

In order to determine some interesting properties of these autoregressive coefficients, let the square matrix  $Y$  be equal to  $X'X$ . This matrix has elements which are sums of the first  $T$  positive integers raised to certain powers.

$$Y_{ij} = \sum_{t=1}^T t^{i+j-2}, \quad i, j, = 1, 2, \dots, N+1. \quad (2.8)$$

Since sums of positive powers can be represented as a function of only the variables  $T$  and  $N$  (and the corresponding Bernoulli numbers), the inverse matrix  $(X'X)^{-1}$  is also a function of only  $T$  and  $N$ . Therefore, the equations which generate the autoregressive terms are specific  $N$  degree polynomials with constant coefficients that are functions of  $T$  and  $N$ . For example, the polynomial which generates  $\alpha_t$  is

$$\alpha_t = \tau (X'X)^{-1} \begin{pmatrix} t^0 \\ t^1 \\ \vdots \\ t^N \end{pmatrix} = a_0 + a_1 t + \dots + a_N t^N. \quad (2.9)$$

Since  $\phi_t = \alpha_{T+1-t}$ ,  $t = 1, 2, \dots, T$ , the lag weights are independent of the data when one uses these autoregressive schemes to extrapolate a time-polynomial estimated by ordinary least-squares.

Another interesting property of these weights can be determined by post-multiplying the equation for  $\alpha$  by the matrix  $X$ .

$$\alpha X = \tau, \quad (2.10)$$

or

$$\sum_{t=1}^T \alpha_t t^n = (T+1)^n, \quad n = 0, 1, \dots, N. \quad (2.11)$$

By virtue of the binomial theorem, this implies that

$$\sum_{t=1}^T \phi_t = 1, \quad (2.12)$$

and

$$\sum_{t=1}^T \phi_t t^n = 0, \quad n = 1, 2, \dots, N. \quad (2.13)$$

The explicit form of the polynomial that determines the lag weights has been derived for the cases where  $n = 1, 2$ , or  $3$  and  $t = 1, 2, \dots, T$ .

$$N = 1, \quad \phi_t = \frac{2(2T+1 - 3t)}{T(T-1)} \quad (2.14)$$

$$N = 2, \quad \phi_t = \frac{3(3T^2+3T+2 - 6(2T+1)t + 10t^2)}{T(T-1)(T-2)} \quad (2.15)$$

$$N = 3, \quad \phi_t = \frac{4(T-4)!}{T!} [4T^3+6T^2+14T+6 - 5(6T^2+6T+5)t + 30(2T+1)t^2 - 35t^3] \quad (2.16)$$

Some special properties which hold at least for these three cases but which have not been proven for all  $N$  cases are

$$\phi_1 = (N+1)^2/T; \quad (2.17)$$

$$\phi_T = (-1)^N(N+1)/T; \quad (2.18)$$

$$\phi_t = (X'X)^{-1} \begin{pmatrix} t^0 \\ t^1 \\ \vdots \\ t^N \end{pmatrix}, \quad t = 1, 2, \dots, T \quad (2.19)$$

This last property implies that the coefficients of the polynomial which generates the autoregressive weights are the elements of the first row of the inverse matrix  $(X'X)^{-1}$ .

### 3. DISCOUNTED LEAST-SQUARES

A logical extension of this approach is to estimate the coefficients of the time-polynomial by minimizing the discounted sum of squared residuals. This approach has been proposed by Brown [2,3] and Box and Jenkins [1]. If this discounted sum is of the form

$$S = \sum_{t=1}^T w_t \varepsilon_t^2, \quad (3.1)$$

then the vector of coefficients estimated by discounted least-squares is

$$\hat{\beta} = (X'WX)^{-1}X'WZ. \quad (3.2)$$

The  $T \times T$  matrix  $W$  is the product of an identity matrix multiplied by the vector of discount factors  $w_t$ ,  $t = 1, 2, \dots, T$ . Now the autoregressive coefficients are a function of  $T, N$ , and the discount factors.

Consider the case where  $N = 1$  and the discount weights are determined by

$$w_t = \delta^{T+1-t}, \quad t = 1, 2, \dots, T, \quad (3.3)$$

where  $\delta$  is some positive fraction less than one. By using the explicit forms for a finite sum of an arithmetic-geometric series, one can derive the corresponding autoregressive weights for one period extrapolation. The general equation for the  $\alpha$  vector is now

$$\alpha = \tau(X'WX)^{-1}XW. \quad (3.4)$$

In this example the autoregressive coefficient  $\phi_t$  ( $\alpha_{T+1-t}$ ) is determined by

$$\phi_t = a_0 \delta^t (a_1 - a_2 t), \quad t = 1, 2, \dots, T \quad (3.5)$$

where

$$a_0 = \frac{(1 - \delta)}{\delta(\delta - \delta^T(T^2(1-\delta)^2 + \delta(2 - \delta^T)))}; \quad (3.6)$$

$$a_1 = (1 + \delta)(1 - \delta^T) - T\delta^T(1 - \delta)(2 + T(1 - \delta)); \quad (3.7)$$

$$a_2 = (1 - \delta)(1 - \delta^T(1 + T(1 - \delta))). \quad (3.8)$$

In the limit as  $\delta$  approaches one, the equation (3.5) for the autoregressive coefficients is equivalent to equation (2.14).

As proven by D'Esopo [4], the polynomial of degree  $N$  obtained by multiple exponential smoothing is equivalent to fitting a  $N$  degree polynomial using exponentially discounted least-squares where the sample size is infinite. Therefore, the autoregressive form of the forecast equation for double exponential smoothing ( $N = 1$ ) can be determined from equation (3.5)

by evaluating the limit as  $T$  approaches infinity. For one period forecasts using double exponential smoothing, the corresponding linear autoregressive forecast function has coefficients which are generated by

$$\phi_t^* = \delta^{t-2}((1-\delta^2) - (1-\delta)^2 t), \quad t = 1, 2, \dots, \infty. \quad (3.9)$$

#### 4. CONSTRAINTS FOR EXTRAPOLATING POLYNOMIAL TRENDS

Another approach to extrapolating polynomial trends is to consider what are the properties of a  $T$  order linear homogeneous difference equation which has a  $N$  order polynomial solution equation. These properties can be used as constraints if one wishes to estimate autoregressive coefficients that will insure the predicted value of a variable will equal its extrapolated value whenever  $T$  successive observations lie on any  $N$  degree time-polynomial. In other words, if

$$z_t = \beta_0 + \beta_1 t + \dots + \beta_N t^N, \quad t = 1, 2, \dots, T, \quad (4.1)$$

and

$$z_{T+1}^* = \phi_1 z_T + \phi_2 z_{T-1} + \dots + \phi_T z_1, \quad (4.2)$$

then

$$z_{T+1}^* = \beta_0 + \beta_1 (T+1) + \dots + \beta_N (T+1)^N. \quad (4.3)$$

If this condition holds for any value of the  $\beta$  coefficients, simple substitutions yield the following constraints:

$$\sum_{t=1}^T \phi_t = 1; \quad (4.4)$$

$$\sum_{t=1}^T t^n \phi_t = 0, \quad n = 1, 2, \dots, N \quad (4.5)$$

Notice that these constraints imply that at least one of the autoregressive coefficients must be negative in order for the forecast function to extrapolate a time-polynomial trend.

The results derived in the section on extrapolating a time-polynomial fitted by ordinary least squares are equivalent to imposing these extrapolation constraints and the assumption that a  $N$  degree polynomial generates the lag weights. Alternative autoregressive coefficients that will extrapolate a  $N$  degree time-polynomial trend can be determined by specifying a different generating function and/or by estimating some lag weights subject to the extrapolation constraints. In order for a solution to exist there must be at least  $N+1$  parameters in the function which is assumed to generate the autoregressive coefficients.

For example, assume that one wants a set of lag weights that will extrapolate any linear trend ( $N = 1$ ) and which are generated by the second degree polynomial

$$\phi_t = a_0 + a_1 t + a_2 t^2, \quad t = 1, 2, \dots, T \quad (4.6)$$

Using the two extrapolation constraints enables one to solve for  $a_0$  and  $a_1$  as functions of  $a_2$ .

$$a_0 = \frac{2(2T+1)}{T(T-1)} + \frac{(T+1)(T+2)}{6} a_2 \quad (4.8)$$

$$a_1 = \frac{-6}{T(T-1)} - (T+1)a_2 \quad (4.8)$$

The value of  $a_2$  can be estimated if one has more than  $T$  observations of the variable. Alternatively it could be determined by imposing an additional condition, such as an end point constraint.

Another example of an autoregressive process that can extrapolate any linear trend is Brown's double moving average method [3]. At time period  $T$  the moving average of a variable using the  $K$  previous observations is

$$M_t = \frac{1}{K} \sum_{i=1}^K z_{T+1-i} \quad (4.9)$$

The moving average of the moving average (double moving average) is

$$M_T^{[2]} = \frac{1}{K} \sum_{j=1}^K M_{T+1-j} \quad (4.10)$$

Brown's equation to forecast the value in the next period can be written as

$$z_{T+1}^* = \frac{2K}{K-1} M_T - \frac{(K+1)}{(K-1)} M_T^{[2]} \quad (4.11)$$

In order to translate this forecast equation into a linear autoregressive form, first manipulate the equation for the double moving average.

$$M_T^{[2]} = \frac{1}{K^2} \sum_{j=1}^K \sum_{i=1}^K z_{T+1-i-j} \quad (4.12)$$

By expanding and then adding the coefficients for each observation, one finds that

$$M_T^{[2]} = \frac{1}{K^2} \left[ \sum_{k=1}^K k z_{T+1-k} + \sum_{k=K+1}^{2K-1} (2K-k) z_{T+1-k} \right] \quad (4.13)$$

Substituting equations (4.9) and (4.13) into (4.11) yields the autoregressive equation for this forecasting technique.

$$z_{t+1}^* = \sum_{k=1}^{2K-1} \phi_k z_{T+1-k}, \quad (4.14)$$

where

$$\phi_k = \frac{2K^2 - (K+1)k}{(K-1)K^2}, \quad k = 1, 2, \dots, K, \quad (4.15)$$

and

$$\phi_k = \frac{-(K+1)(2K-k)}{(K-1)K^2}, \quad k = K+1, K+2, \dots, 2K-1. \quad (4.16)$$

Notice that the first  $K$  coefficients are positive and decreasing in value whereas that last  $K-1$  coefficients are negative and increasing in value. This implies that there is a relatively large difference between the  $K$  and the  $K+1$  lag weights.



5. EXTRAPOLATING USING FIRST DIFFERENCES  
OF THE OBSERVATIONS

As shown in the previous section, in order for a linear autoregressive function to be able to extrapolate all N degree time-polynomial trends, the coefficients must satisfy the constraints (4.4) and (4.5). An alternative forecast equation form is to use the first differences of the observations so that

$$z_{T+1}^* = z_T + \sum_{t=1}^{T-1} \phi_t^{[1]} (z_{T+1-t} - z_{T-t}). \quad (5.1)$$

The notation  $\phi_t^{[1]}$  indicates that these coefficients are for the first difference between successive observations. Since the first difference of a N degree polynomial is a N-1 degree polynomial, the corresponding constraints on the  $\phi_t^{[1]}$  coefficients that will enable one to extrapolate any time trend are

$$\sum_{t=1}^{T-1} \phi_t^{[1]} = 1 \quad (5.2)$$

and, if  $N > 1$ ,

$$\sum_{t=1}^{T-1} t^n \phi_t^{[1]} = 0, \quad n = 1, 2, \dots, N-1. \quad (5.3)$$

The relation between these two sets of autoregressive coefficients can be summarized by

$$\phi_t^{[1]} = \sum_{s=1}^t \phi_s - 1, \quad t = 1, 2, \dots, T-1 \quad (5.4)$$

or

$$\begin{aligned} \phi_1 &= 1 + \phi_1^{[1]} \\ \phi_t &= \phi_t^{[1]} - \phi_{t-1}^{[1]}, \quad t = 2, 3, \dots, T-1; \\ \phi_T &= -\phi_{T-1}^{[1]} \end{aligned} \quad (5.5)$$

When one assumes that the  $\phi_t$  coefficients are determined by some specific function, the corresponding function which generates the  $\phi_t^{[1]}$  coefficients can be determined. For example, if one forecasts one period ahead by estimating a linear time-polynomial, the coefficients of the equivalent autoregressive process are given by equation (2.14). Substituting this equation into (5.4) yields the following function for the corresponding  $\phi_t^{[1]}$  coefficients.

$$\phi_t^{[1]} = -1 + \frac{(4T-1)t - 3t^2}{T(T-1)}, \quad t = 1, 2, \dots, T-1. \quad (5.6)$$

Notice that this is a second degree polynomial in t; whereas, the  $\phi_t$  coefficients are determined by a linear equation. Furthermore, this equation is concave from below.

## 6. SOME ADDITIONAL OBSERVATIONS

One of the simplifying assumptions of the preceding analysis is that the forecasts are being made for only one period ahead. There are two established approaches one can use when generalizing the results for any lead time  $L$ . One may estimate the coefficients of a time-polynomial and solve the estimated equation for any desired time. The coefficients of the corresponding linear autoregressive process can be determined from equation (3.4), where the  $\tau$  vector is now defined by

$$\tau_n = (T+L)^{n-1}, \quad n = 1, 2, \dots, N+1 \quad (6.1)$$

The second approach is to consider the autoregressive process to be a  $T$  order difference equation where the observations are the initial conditions. Evaluating this difference equation for successive time periods provides forecasts for any desired lead time. When using an autoregressive process for extrapolating an estimated trend, this second approach is equivalent to estimating the trend in the original  $T$  observations, extrapolating to get the next period forecast, using this value and the  $T-1$  previous observations to reestimate the trend, and extrapolating the revised trend one period. This process of estimating and extrapolating one period is repeated until the forecast of each desired lead time is obtained.

When a variable has a simple exponential time trend, the coefficients can be estimated by least-squares after transforming the observations into a linear logarithmic system. The logarithm of the next period forecast can therefore be determined by the linear autoregressive process

$$\ln(z_{T+1}^*) = \sum_{t=1}^T \theta_t \ln(z_{T+1-t}), \quad (6.2)$$

Where  $\theta_t$  is generated by equation (2.14). The value of this forecast can also be determined by the corresponding linear logarithmic autoregressive process

$$z_{T+1}^* = \prod_{t=1}^T (z_{T+1-t})^{\theta_t}$$

Using this logarithmic form, any simple exponential time trend can be extrapolated provide the exponential weights satisfy the constraints (4.4) and (4.5) for  $N = 1$ .

## 7. CONCLUSIONS

Assume one wishes to forecast future values of a time series that can be represented by a time-polynomial model where the stochastic terms are independently distributed with zero means and the observations are from discrete, equispaced intervals of time. Instead of estimating the coefficients of the time-polynomial by ordinary or discounted least-squares and extrapolating to get the forecasts, one only needs to take a weighted sum of the observations. The equation which generates the autoregressive coefficients is a function of the number of observations used, the degree of the fitted time-polynomial, the lead time of the forecast, and the function which determines the discount weights.

## REFERENCES

- [1] Box, G. E. and Jenkins, G. M., Time Series Analysis: Forecast and Control, San Francisco; Holden-Day, 1970.
- [2] Brown, R. G., Smoothing, Forecasting and Prediction of Discrete Time Series, New Jersey; Prentice-Hall, 1962.
- [3] Brown, R. G. and Meyer, R. F., "The Fundamental Theorem of Exponential Smoothing," Operations Research, 9 (1961), 673-635.
- [4] D'Esopo, D. A., "A Note on Forecasting by the Exponential Smoothing Operator," Operations Research, 9 (1961), 686-687.
- [5] Whittle, P., Prediction and Regulation by Linear Least-Squares Methods, London; English Universities Press, 1963.
- [6] Wold, H., A Study in the Analysis of Stationary Time Series (2nd ed.), Stockholm; Almqvist & Wiksell, 1954.