EXTRAPOLATING TIME SERIES BY DISCOUNTED LEAST SQUARES

by

R. J. Duffin

Report 66-11

November, 1966

University Libraries
Carnegie Mellon University
Pittsburgh PA 15213-3890

Extrapolating Time Series by Discounted Least Squares

Abstract

An approximating function is fitted to a time series, such as daily observation. The fitting is carried out over all past time by weighted least squares with an exponential weight factor. The approximating function is restricted to be a solution of a certain linear differential equation of the mth order having constant coefficients. The solution which minimizes the least square expression can be continued into the future. In particular tomorrow's extrapolated value is defined by this continuation. To obtain an explicit solution of the problem a formula is constructed which gives the extrapolated value as a linear combination of the last m observed values and the last m extrapolated values. The coefficients of this extrapolation formula prove to be simply related to the coefficients of the differential equation. Another extrapolation formula is of vectorial nature. The components of a vector are m independent functionals of the past observations. Then tomorrow's vector is given as a linear function of today's vector and today's scalar observation.

HURT LIBRARY
CARNEGIE-MELLON UNIVERSITY

EXTRAPOLATING TIME SERIES BY DISCOUNTED LEAST SQUARES*

R. J. Duffin

CARNEGIE INSTITUTE OF TECHNOLOGY

This paper is concerned with extrapolation of an infinite sequence $y_1, y_2, \ldots$, of real (or complex) numbers. This is accomplished by fitting the sequence (y_n) by a function $p(n)$ taken from a space of functions termed exponomials. The criteria for the fit is given by discounted least squares. This means that $p(n)$ is that exponomial which minimizes the 'error' expression

$$E = \sum_1^{\infty} \theta^n |y_n - p(n)|^2 .$$

Here θ is a positive constant termed the discount factor. Then the extrapolated value of the sequences at the point x is defined to be $p(x)$.

In the previous paper an exponomial was defined as an exponential polynomial of the form

$$p(x) = \sum_1^m d_j \beta_j^x$$

where the β_j are fixed complex numbers assumed distinct and nonzero. The coefficients d_j are arbitrary complex numbers so an m-dimensional vector space results. In this paper the definition of exponomial is extended so as to permit polynomial terms of the form $\beta_j^x, x\beta_j^x, \ldots, x^{k-1}\beta_j^x$. It is seen, therefore, that the space of exponomials can be defined to be the solution set of a certain linear differential equation,

$$\frac{d^m p}{dx^m} + c_{m-1}\frac{d^{m-1} p}{dx^{m-1}} + \cdots + c_0 p = 0$$

*Prepared under Research Grant DA-ARO-D-31-124-G680, Army Research Office (Durham).

where the coefficients e_i are complex constants. This space of exponomials will be denoted as XP. An important special case is defined by the equation $d^m p/dx^m = 0$. Then XP is the space of polynomials of degree less than m.

There are three reason why the extrapolation procedure just described can treat a large class of problems in applied mathematics. The first reason is that the discounted least squares criteria is suited to problems of mechanics and economics for which the progressive discount of the past seems natural. Another reason is that the space XP is invariant under an arbitrary translation of the x axis. This invariance property makes exponomials attractive functions for approximating time series. The third reason for the utility of this extraipolation method is that there is an underlying algebraic structure which is both interesting and significant.

The bases $£_i$ are arbitrary. The selection of the $£_i$ and the discount factor 9 should take into account first the genesis of the data. Also, account must be taken of the genesis of data error and the smoothing property of the extrapolation. These questions are not treated in this paper.

A central problem of this paper is the one step extrapolation of the sequence $\{ y_n \}$ to obtain an extrapolated value at n = 0. The extrapolated value is denoted as y_0^* and is defined as y£ = p(0)• When the minimization is carried out, it results that y_0^* is given by a linear functional

$$y_0^* = \sum_{-1} Q_n y_n .$$

Here the coefficients Q_n do not depend on the sequence

$y_1, y_2, \ldots$ It is then natural to write

$$y_k^* = \sum_1^{\infty} Q_n y_{n+k}$$

and define y_k^* as the predicted value of y_k based only on the previous values $y_{k+1}, y_{k+2}, \ldots$

Theorem 3b to follow gives a simple generating function to evaluate the coefficients Q_n. Thereby the problem of one-step extrapolation is essentially solved. Nevertheless, the above formula is not satisfactory from a computational point of view. This is so because it is an infinite series and so an infinite memory is needed. This situation is remedied by Theorem 2 which provides the following short memory formula,

$$y_k^* = \sum_1^m T_n y_{n+k} + \sum_1^m \phi_n y_{n+k}^*$$

Here the T_n and the ϕ_n are constants which do not depend on the sequence y_n. Thus, this identity gives the extrapolation y^* as a linear combination of the m previous values and the m previous extrapolations. Consequently, the short memory formula is readily adaptible for computer evaluation.

Also of importance are extrapolation formulas for other linear functionals of the sequence $y_1, y_2, \ldots$. For example $p(1/2), p'(0), p''(-2), \int_0^2 p(x)dx$ are linear functionals. We term such functions extrapolators. Let $w(x)$ be a vector whose m components are linearly independent extrapolators. Then Theorem 7 gives the following extrapolation formula.

$$w(x) = Aw(x+1) + by_{x+1}.$$

where A is constant matrix and b is a constant vector. This is termed a very short memory formula because the extrapolation is based only on one previous value of $\vec{w}$ and one previous value of y.

The extrapolation of time series based on a discounted least squares criteria has previously been treated by Duffin and Schmidt [1], Duffin and Whidden [2], and Morrison [3]. (Various other authors have proposed similar extrapolation formulae but their work is not based on discounted least squares.) The present paper gives a more general treatment of problems posed by references [1], [2], and [3]. In particular the theorems of this paper are aimed at evaluating and interrelating the constants, p_i, ©, Q_n, T_n, $§_n$ * A, and $\vec{b}$ by algebraic formulae.

To begin the proof, let G(z) be a polynomial of degree m in the variable z and let G(0) = 1. Then

$$G(z) = \sum_{j=0}^{m} g_j z^j .$$

Here the g_j are arbitrary complex constants except that $g_o = 1$ and $g_m \neq 0$. Of central concern are functions p(x) of the real variable x satisfying the difference equation

$$o = g_0 p(x) + g_1 p(x + 1) + \cdots + g_m p(x + m) .$$

Let X denote the translation operation defined by the relation Xf(x) = f(x + 1). Then the difference equation can be written in operation form as

$$0 = G(X)p(x) = \prod_{j=1}^{m} (1-\lambda_j^{-1} X)p(x)$$

where the λ_j are the roots of G(z) = 0. First suppose that x is restricted to the integers. If none of the roots are repeated it is seen that there are m linearly independent solutions having the form $p(x) = \beta_j^x$. The general solution is a linear combination of such solutions. If the root β_j is repeated k times there are k linearly independent solutions: $\beta_j^x, x\beta_j^x, \dots, x^{k-1}\beta_j^x$. The general solution is again a linear combination of such solutions and defines an m-dimensional vector space. This is the space of exponomials and it is denoted as XP.

For some problems of extrapolation it is required to continue exponomials to non-integral values. One way to do this is to let $\beta = e^{\alpha}$ and then define $\beta^x = e^{\alpha x} e^{2\pi i L x}$ where L is an integer. A 'natural' choice for L is

$$-\pi \le \text{imag}\, \alpha + 2\pi L \le \pi .$$

This choice minimizes oscillation. The ambiguity here stems from the fact that multiplying a solution of the equation by an arbitrary function of period 1 gives another solution.

Since an exponomial satisfies the difference equation it is seen that prescribing the value of the exponomial at the integers $1,2,\dots,m$ determines the value of the exponomial for all other positive integers. Thus the space XP is m-dimensional even if x is restricted to the integers $1,2,\dots,m$.

The approximation scheme employed in this paper results from embedding the finite dimensional space XP in an infinite Hilbert space of 'discounted squares'. This Hilbert space is denoted as DS. The elements of DS are infinite sequences of complex numbers $y_1, y_2, \dots$, such that

$$\sum_1^{\infty} \delta^n |y_n|^2 < \infty .$$

Here δ is a positive constant termed the <u>discount factor</u>. The discount factor is required to satisfy the inequality

$$0 | £_j . | < 1; j = 1,2,...,m.$$

It is readily shown from this inequality that exponomials are in DS. Thus the space XP is a finite dimensional subspace of DS. Let v and y denote two elements of DS then the Hermitian bilinear form is defined as

$$[v,y] = \sum_{1}^{\infty} \theta^n \bar{v}_n$$

where $\bar{v}_n$ denotes the complex conjugate of v_n. The norm of a element y of DS is defined as

$$||y|| = [y,y]^{1/2}.$$

It is instructive in what follows to regard the sequence $\{y_n\}$ as a time series of observations'. Thus y_x can be regarded as the value of the observation at time -x. To extrapolate the sequence $\{y_n\}$ for values of n not a positive integer we first approximate y by an element p of the subspace XP. By this is meant that p is chosen to minimize the expression

$$E = ||y - p||^2 = \sum_{1}^{\infty} \theta^n |y_n - p(n)|^2.$$

As in the references [1], [2], and [3] this is termed approximation by ~~discounted least squares~~.

The following lemma is aimed at determining p given a sequence $\{y_n\}$.

~~Lemma~~ 1. ~~Let~~ p ~~be the exponomial best approximating~~ y ~~in the Hilbert space~~. ~~Then~~

$$[r,y] = [r,p]$$

~~for every exponomial~~ r.

Proof. This is merely a reformulation of the basic theorem concerning Hilbert space which states that if p is the best approximation to y for p constrained to a subspace then p is the orthogonal projection of y into the subspace.

By choosing a basis for the subspace XP the orthogonality relation stated by this lemma leads to a system of m linear equations which could be used to determine p as a unique linear combination of the basis elements.

Then p is determined uniquely at the positive integer points. The convention introduced above permits an exponomial, given at the positive integer points, to be determined for all real x. Thus it is possible to define the extrapolation of the sequence $y_1, y_2, \ldots$ at the point x to be p(x).

Lemma 2. *Let* r *be a given exponomial and let* k *be a given number*. *Then there is a unique sequence of numbers* $c_1, c_2, \ldots, c_m$ *such that the formula*

$$[r,v] = c_1v(k + 1) + c_2v(k + 2) + \cdots + c_mv(k + m)$$

holds for every exponomial v.

Conversely given a number k *and a sequence of numbers* $c_1, c_2, \ldots, c_m$ *there is a unique exponomial* r *for which the above formula holds*.

Proof. As is well known the linear functionals f(v) in an m-dimensional space such as XP have the form f(v) = [r,v]. Moreover, these linear functionals themselves form an m-dimensional space. Clearly the expression on the right side of the formula of the lemma is a linear functional for any choice of constants $c_1, c_2, \ldots, c_m$. Suppose that for some choice $c_1v(k + 1) + \ldots c_mv(k + m)$ vanishes for all v in XP. However, an exponomial can be prescribed arbitrarily at a sequence of m points obtained by successive unit translations to the right. Hence all the constants

c_1^* must be zero. This shows that $v(k + 1), v(k + 2), \ldots, v(k + m)$ are m independent functionals. There can be no more than m independent functionals and consequently the left and right side of the formula represent the same space of linear functionals. The proof of the lemma follows from this observation.

Theorem 1. Let r be a given exponomial and let k be a given number, then there is a unique sequence of numbers $c_1, c_2, \ldots, c_m$ such that the formula

$$[r,y] = c_1 p(k + 1) + c_2 p(k + 2) + \cdots + c_m p(k + m)$$

holds for every y of the Hilbert space DS provided that p is the orthogonal projection of y into the exponomial subspace XP.

Conversely given a sequence of numbers $c_1, c_2, \ldots, c_m$ there is a unique exponomial r satisfying this formula.

Proof. By Lemma 1 we have $[r,y] = [r,p]$. Then apply Lemma 2 with $v = p$ and the proof is seen to be complete.

A general problem of extrapolation is to extrapolate the sequence $y_1, y_2, \ldots$ to obtain a value for y at a point x not a positive integer. This extrapolated value is denoted by the symbol $\text{ext}(y_x)$ and is defined as $\text{ext}(y_x) = p(x)$. The following corollary of Theorem 1 gives a formula for computing $\text{ext}(y_x)$.

Corollary 1. Let $p(x)$ be the approximating exponomial to the sequence $y_1, y_2, \ldots$ Then there is a kernel function $q(x,n)$ such that

$$p(x) = \sum_{n=1}^{\infty} e^{n} q(x,n) y_n = [g,y] .$$

For n fixed $q(x,n)$ is an exponomial in x. For x fixed

$\bar{q}(x,n)$ *is an exponomial in* n. *The kernel function may depend on* θ *but not on* y.

Proof. In Theorem 1 take $k = x - 1$, $c_1 = 1$, $c_2 = 0, \ldots, c_m = 0$. This proves the formula of Corollary 1 with $\bar{q} = r$. To show that the kernel function is an exponomial in x take $y_n = 0$ except for $n = n_o$. This is seen to complete the proof.

A significant special case of the formula of Corollary 1 is $x = 0$. This extrapolates the sequence $y_1, y_2, \cdots$ one unit to the left to obtain $\mathrm{ext}(y_o)$. This one step extrapolation is sufficiently important to warrant a special notation and we write $y_o^* = p(0)$. The general extrapolation formula becomes

$$y_o^* = \sum_{n=1}^{\infty} \theta^n q(n) y_n .$$

Here $q(n) = q(0,n)$ and $q(n)$ is also termed a kernel.

If the sequence $y_1, y_2, \ldots$ is in the Hilbert space DS then the sequences $y_{x+1}, y_{x+2}, \cdots$ is also in DS for any positive or negative integer x. This follows from the relation

$$\sum_{1}^{\infty} \theta^n |y_{x+n}|^2 = \theta^{-k} \sum_{x}^{\infty} \theta^n |y_n|^2 < \infty .$$

As a natural extension of the previous notation let y_x^* denote the extrapolation of the sequence $y_{x+1}, y_{x+2}, \cdots$ to the point x. The extrapolation formula given by discounted least squares is

$$y_x^* = \sum_{1}^{\infty} \theta^n q^*(n) y_{n+x} .$$

This formula involves all values of the sequence $y_{x+1}, y_{x+2}, \cdots$ and so is an infinite series. For this reason a formula of this nature may be termed a *long memory formula*. It is now to be shown that y_k^* is also given by a finite recursion relation and hence by a *short memory formula*.

Theorem 2. Let yf be defined by the formula

$$y_k^* = \sum_{n=1}^{\infty} \theta^n q(n) y_{n+k}.$$

Then y£* the one step extrapolation, is given by the recursion formula

$$y_k^* = -\sum_{n=1}^{m} g_n y_{n+k} - \sum_{n=1}^{m} f_n \theta^n \delta_{n+k}.$$

Here, for convenience of notation, $f_n = \bar{g}_{n-m}/\bar{g}_m$. Also $\delta_n = y_n^* - y_n$ and is termed the discrepancy.

Proof. Let $q' = 0$ for $n \leq 0$ and $q' = q$ for $n > 0$ so

$$y_x^* = \sum_{-\infty} \theta^n q'(n) y_{n+x},$$

$$\theta^x y_x^* = \sum_{-\infty}^{\infty} \theta^n q'(n - x) y_n.$$

Since $\bar{q}(x)$ is a polynomial $G(X)\bar{q}(x) = 0$. Thus $\sum_{0}^{m} \bar{g}_j q(x + j) = 0$ or $\sum_{0}^{m} \bar{g}_{m-j} q(x + m - j) = 0$. A polynomial F(z) related to the polynomial G(z) plays an important role in what follows. It is defined as

$$F(z) = \sum_{j=0}^{m} f_j z^j = \sum_{j=0}^{m} \bar{g}_{j-m} z^j/\bar{g}_m.$$

It is seen that $F(X)q(n - x) = 0$ for all n. Consequently $F(X)q'(n - x) = 0$ for $n > x + m$ or $n \leq x$ so

$$F(X)\theta^x y_x^* = \sum_{n=-\infty}^{\infty} \Big[\sum_{j=0}^{m} f_j q'(n - x - j)\Big]\theta^n y_n$$

$$= \sum_{n=x+1}^{x+m} s_n y_n$$

where the s_n are certain absolute constants. To evaluate these constants,, first set x = 0 in the last relation and obtain

(*) $$\sum_{j=0}^{m} f_j \theta^j y_j^* = \sum_{n=1}^{m} s_n y_n.$$

Next let $y_k = r(k)$ an exponomial. Since the extrapolation of exponomials is error-free, $y_k^* = r(k)$ also. Note that $f_0 = 1$ so after substituting $y_k = r(k)$ the relation (*) can be written as

$$r(0) = \sum_{j=1}^{m} (s_j - f_j \theta^j) r(j).$$

But since $G(X)r(x) = 0$ it is also true that

$$r(0) = -\sum_{j=1}^{m} g_j r(j).$$

Subtracting these two equations for $r(0)$ gives

$$0 = \sum_{j=1}^{m} (s_j - f_j \theta^j + g_j) r(j).$$

But an exponomial can be defined arbitrarily on the integers 1,2,...,m. Thus $s_j = f_j \theta^j - g_j$ and relation (*) becomes

$$y_0^* = \sum_{j=1}^{m} (f_j \theta^j - g_j) y_j - \sum_{j=1}^{m} f_j \theta^j y_j^*.$$

This is seen to be equivalent to the short memory formula stated in Theorem 2.

Corollary 2. The kernel function $q(x)$ satisfies the constant coefficient difference equation $G(X)\bar{q}(x) = 0$, an equation of the mth order. But $q(x)$ does not satisfy any such equation of lower order.

Proof. Suppose $q(x)$ satisfied the equation $G'(X)\bar{q}(x) = 0$ of order $m' < m$. Then the proof of Theorem 2 could be carried out with G' replacing G. This would lead to a relation of the form

$$r(0) = \sum_{j=1}^{m'} (s'_j - f'_j \theta^j) r(j)$$

for every exponomial $r(j)$. But an exponomial can be defined arbitrarily at m successive points. Thus take $r(0) = 1$ but $r(j) = 0$ for $j = 1, \ldots, m-1$. This contradicts the assumed relation and the proof is complete.

So far the kernel function $q(n)$ of the long formula has only been defined implicitly. On the other hand the coefficients of the short formula are given explicitly in Theorem 2. However, since the short memory formula and the long memory formula are essentially equivalent it is possible to use the short formula to give an explicit procedure for evaluating $q(n)$. Different ways of doing this are given in Theorem 3a and Theorem 3b to follow.

<u>Theorem</u> 3a. <u>The kernel function</u> $q(n)$ <u>satisfies the recursion formulae</u>:

$$q(n) = f_n - \theta^n g_n - \sum_{j=1}^{n-1} f_j q(n-j), \quad 1 \le n \le m$$

$$q(n) = -\sum_{j=1}^{m} f_j q(n-j), \quad n > m.$$

<u>Thus</u>

$$q(1) = f_1 - \theta g_1$$

$$q(2) = f_2 - \theta^2 g_2 - f_1^2 + \theta f_1 g_1$$

$$q(3) = f_3 - \theta^3 g_3 - 2 f_1 f_2 + f_1^3 - \theta f_1^2 g_1 + \theta^2 f_1 g_2 + \theta f_2 g_1 .$$

Proof. Let $y_j = 0$ for all j except that $y_n = 1$. Then it follows directly from the long formula that $y_0^* = \theta^n q(n)$. It is also seen that: $y_j^* = \theta^{n-j} q(n-j)$ if $j \le n-1$, $y_j^* = 0$ if $j \ge n$. The short formula may be written in the form

$$y_0^* = \sum_{j=1}^{m} (\theta^j f_j - g_j) y_j - \sum_{j=1}^{m} \theta^j f_j y_j^*$$

Substituting the above special values of y_j and y_j^* in this formula leads directly to the formulae of Theorem 3a.

HUNT LIBRARY
CARNEGIE-MELLON UNIVERSITY

Theorem 3b. The kernel function q(n) satisfies the generating identity

$$\frac{F(\theta z) - G(z)}{F(\theta z)} = \sum_{1}^{\infty} \theta^n q(n) z^n$$

provided |z| is small.

Proof. If |z| is small the sequence $1, z, z^2, \ldots$ is in the Hilbert space DS. Thus, substituting $y_n = z^n$ in the long formula gives

$$y_0^* = \sum_{1}^{\infty} \theta^n q(n) z^n = z y_0^*$$

Then substituting $y_n = z^n$ and $y_k^* = z^k y_0^*$ in the short formula gives

$$y_0^* = \sum_{j=1}^{m} (\theta^j f_j - g_j) z^j - \sum_{j=1}^{m} \theta^j f_j z^j y_0^*,$$

But $f_0 = g_0 = 1$ so

$$y_0^* \sum_{j=0}^{m} f_j \theta^j z^j = \sum_{j=0}^{m} f_j \theta^j z^j - \sum_{j=0}^{m} g_j z^j$$

and the proof is complete.

We now turn from one-step extrapolation to multi-step extrapolation. Thus the two-step extrapolation of the sequence $y_1, y_2, \ldots$ is given by $\text{ext}(y_{-1}) = p(-1)$ etc.

Theorem 4a. Multistep extrapolators of the sequence $y_1, y_2, \ldots$ are given by the long memory formulae:

$$p(0) = \sum_{1}^{\infty} \theta^n q_n y_n,$$

$$p(-1) = \sum_{1}^{\infty} \theta^{n+1} (q_{n+1} + q_1 q_n) y_n,$$

$$p(-2) = \sum_{1}^{\infty} \theta^{n+2} (q_{n+2} + q_1 q_{n+1} + q_2 q_n + q_1^2 q_n) y_n,$$

$$p(-3) = \sum_1^{\infty} \theta^{n+3}[q_{n+3} + q_1 q_{n+2} + (q_1^2 + q_2)q_{n+1} + (q_1^3 + 2q_1 q_2 + q_3)q_n]y_n.$$

Here q_n is a condensed notation for the kernel $q(n)$.

Proof. Let $Y_n = y_n$ for $n > 0$ and let $Y_n = p(n)$ for $n \leq 0$. Let $E_0 = \sum_0^{\infty} \theta^n |Y_n - P(n)|^2$ where $P(n)$ is an exponomial. Thus

$$E_0 = |p(0) - P(0)|^2 + \sum_1^{\infty} \theta^n |y_n - P(n)|^2.$$

Clearly E_0 is minimized when $P(n) = p(n)$. Hence applying the one-step extrapolation formula to the sequence $Y_0, Y_1, Y_2, \ldots$ gives

$$p(-1) = \sum_1^{\infty} \theta^n q_n Y_{n-1}$$

$$p(-1) = \theta q_1 p(0) + \sum_2^{\infty} \theta^n q_n y_{n-1}.$$

But

$$p(0) = \sum_1^{\infty} \theta^n q_n y_n$$

so

$$p(-1) = \sum_1^{\infty} \theta^{n+1} q_1 q_n y_n + \sum_1^{\infty} \theta^{n+1} q_{n+1} y_n.$$

This proves the formula for the two-step extrapolator.

Next consider $E_{-1} = \sum_{-1}^{\infty} \theta^n |Y_n - P(n)|^2$ and following similar reasoning to that given above shows that

$$p(-2) = \sum_1^{\infty} \theta^n q_n Y_{n-2}$$

$$p(-2) = \theta q_1 p(-1) + \theta^2 q_2 p(0) + \sum_3^{\infty} \theta^n q_n y_{n-2}.$$

Substituting the series expression just derived for $p(-1)$ is seen to prove the formula given for the three-step extrapolation. Further formulae are derived analogously and the proof is complete.

Theorem 4b. Multistep extrapolators of the sequence $y_1, y_2, \ldots$ are given by the short memory formulae;

$$-p(0) = \sum_1^m g_n y_n + \sum_1^m \theta^n f_n \delta_n$$

$$-p(-1) = \sum_1^m (g_{n+1} - g_1 g_n) y_n + \sum_1^m \theta^n (\theta f_{n+1} - g_1 f_n) \delta_n ,$$

$$-p(-2) = \sum_1^m (g_{n+2} - g_1 g_{n+1} - g_2 g_n + g_1^2 g_n) y_n ,$$

$$\sum_1^m \theta^n (\theta^2 f_{n+2} - \theta g_1 f_{n+1} - g_2 f_n + g_1^2 f_n) \delta_n .$$

where $g_n = 0$ and $f_n = 0$ if $n > m$.

Proof: Apply the short memory formula to the sequence $Y_0, Y_1, \ldots$ which was introduced in the proof of Theorem 4a. But $Y^*_{-1} = p(-1)$

so

$$* \quad \sum_1^m \ldots_{n} \, {}_{n-1} \quad \sum_1^m \theta^n f_n \Delta_{n-1} .$$

Also $Y_0 = p(0)$, $\Delta_0 = 0$ and so

$$-p(-1) = g_1 p(0) + \sum_2^m g_n y_{n-1} + \sum_2^m \theta^n f_n \delta_{n-1}$$

$$-p(-1) = -g_1 \left(\sum_1^m g_n y_n + \sum_1^m \theta^n f_n \delta_n \right) + \sum_1^{m-1} g_{n+1} y_n +$$

$$\sum_1^{m-1} \theta^{n+1} f_{n+1} \delta_n .$$

This seen to prove the stated formulae for $p(-1)$. Further formulae are derived analogously.

Theorem 5. The relation

$$F(\theta z) \sim \sum^{\infty} e\, t(n) z^n$$

<u>gives a one to one linear correspondence between polynomials of the form</u> \ (z) = $\sum_{1}^{m} a_j z^{j}$ <u>and conjugate exponomials</u> t(n) .

Proof. Given such a polynomial $\tau(z)$ then T(z) = H (z)/F(€)z) is a convergent power series if |z| is sufficiently small. Thus write

$$T(z) = \sum_{-\infty}^{\infty} \theta^{n} t(n) z^{n}$$

where t(n) =0 for n £ 0. Since %(z) = F(0z)T(z) we have

$$\tau(z) = \sum_{0}^{m} \sum_{-\infty}^{\infty} f_j \theta^{j+n} z^{j+n} t(n),$$

Let j + n = k so

$$\tau(z) = \sum_{-\infty}^{\infty} \theta^{k} z^{k} \sum_{0}^{m} f_j t(k - j)$$

Since *C(z) is a polynomial of degree not exceeding m it follows from the above relation that

$$\sum_{0}^{m} f_j t(k - j) = 0 \quad \text{for} \quad k > m, \text{ or}$$

$$\sum_{0}^{m} f_j t(x + m - j) = 0 \quad \text{for} \quad x > 0, \text{ so}$$

$$\sum_{0}^{m} f_{m-j} t(x + j) = 0 \quad \text{for} \quad x > 0.$$

Hence $G(X)\bar{t}(x) = 0$ and it follows that $\bar{t}(x)$ is an exponomial for x > 0.

We can conclude from this last result that the m-dimensional vector space p of polynomials of the form 't(z) is mapped linearly into a space S of conjugate exponomials. Moreover, this is a one to one mapping because power series are unique so S has dimension m. However, the space XP of all conjugate exponomials has dimension m so S = XP and the proof is complete.

Lemma 3. Let $1(z) = \sum_{i}^{m} a_i z^i$ be a given polynomial. Then there are polynomials

$$H(z) = \sum_{1}^{m} h_j z^j \quad \text{and} \quad K(z) = \sum_{1}^{m} k_j z^j$$

such that

$$C(z) = -H(z)G(z) + K(z)F(\theta z).$$

Hence the system of 2m equations

$$a_i A_i = -\sum_{1}^{i} g_{i-j} h_j + \sum_{1}^{i} e^{i-j} f_{i-j} k_j, \quad 1 \le i \le 2m$$

can be used to find the coefficients h_j and k_j. Here $A_j = 1$ for $i \le m$ and $A_i = 0$ for $i > m$.

Proof. The roots of G(z) are $\{\beta_i\}$ and the roots of $F(\theta z)$ are $\{\theta/\beta_j\}$. The discount factor θ was choosen so that
ount factor
$|\beta$.

It follows that the roots of G(z) are outside a circle of radius $\theta^{1/2}$ and the roots of F(θz) are inside this circle. Consequently G(z) and F(θz) cannot have a common root so by a basic theorem in the algebra of polynomials

$$C(z)/z = -H_0(z)G(z) + K_0(z)F(\theta z)$$

where $H_0(z)$ and $K_0(z)$ are polynomials of degree less than the degree of G and F. Of course G and F are of degree m so multiplying through by z leads to the relation stated in the lemma.

Theorem 6. Given an arbitrary exponomial $\bar{t}(n)$ let w(x) be the corresponding extrapolator

$$w(x) = \sum_{1}^{\infty} e^n t(n) y_{n+x}.$$

Then a short formula for this extrapolator is

$$w(x) = \sum_{1}^{m} h_j \delta_{j+x} + \sum_{1}^{m} k_j y_{j+x}.$$

Here h_j and k_j are coefficients defined by the polynomial $\Gamma(z)$ which is the image of $t(n)$ according to Theorem 5 and Lemma 3.

Proof. Let $Q(z) = \sum_{1}^{\infty} \theta^n q(n) z^n$ then according to Theorem 3b $1 - Q(z) = G(z)/F(\theta z)$. Let $T(z) = \sum_{1}^{\infty} \theta^n t(n) z^n$. Then according to Theorem 5 we have $T(z) = C(z)/F(\theta z)$. Then Lemma 3 with X replacing z gives

$$T(X) = H(X) [Q(X) - 1] + K(X).$$

Here X is interpreted as the translation operator. Operating on the function y_x with the above identity we see that $[Q(X) - 1] y_x = \delta_x$ and so the short formula follows. This formalism is justified by the absolute convergence of the resulting series when y is in the Hilbert space and so the proof is complete.

A different proof for the existence of the short formula for w(x) results from combining Theorem 1 with k = -m and Theorem 4b.

The short memory formulae given in Theorem 2 and Theorem 6 may be termed m th order formulae because the right sides are expressed as translation operators of order m. Thus the extrapolator of Theorem 6 can be written as

$$w(x) = \left(\sum_{1}^{m} h_j X^j\right) \delta_x + \left(\sum_{1}^{m} k_j X^j\right) y_x.$$

Now an mth order scalar difference equation is equivalent to a first-order vector difference equation. This suggests that the mth order scalar extrapolation formula can be replaced by a first order vector extrapolation formula. In fact such formulae are to be found in the paper by Morrison [3]. for the special case of polynomial extrapolation. Moreover, Morrison's paper indicates that first order vector extrapolators may be advantageous in numerical work because of economy in memory.

The following is a general theorem on first order vector extrapolation formula.

Theorem 7. Let $w_1(x), \ldots, w_m(x)$ denote a set of m independent extrapolators. Then there is a set of constants a_{ij} and b_i such that

$$w_i(x) = \sum_{j=1}^{m} a_{ij} w_j(x + 1) + b_i y_{x+1}; \quad i = 1, \ldots, m.$$

This is termed a very short memory formula.

Proof. Then $w_i(x) = \sum_{n=1}^{\infty} \theta^n r_i(n) y_{n+x}$ and the exponomials $r_1(n), r_2(n), \ldots, r_m(n)$ form a basis for the space X^p. Hence it is possible to find constants a_{ij} such that

$$\theta r_i(n + 1) = \sum_{j=1}^{m} a_{ij} r_j(n); \quad i = 1, \ldots, m$$

because $\theta r_i(n + 1)$ is also an exponomial. These identities are multiplied by $\theta^n y_{n+x+1}$ and summed. This gives

$$\sum_{n=1}^{\infty} \theta^{n+1} r_i(n + x) y_{n+x+1} = \sum_{j=1}^{m} a_{ij} w_j(x + 1)$$

$$\sum_{n=1}^{\infty} \theta^n r_i(n) y_{n+x} = \sum_{j=1}^{m} a_{ij} w_j(x + 1) + \theta r_i(1) y_{x+1}$$

Let $b_i = \theta r_i(1)$ and this is seen to complete the proof of the very short formula.

Corollary 3. The extrapolation of y_n at m successive points is given by a very short memory formula.

Proof. The extrapolation of y_n at m successive points is $p(x + 1)$, $p(x + 2), \ldots, p(x + m)$. By virtue of Theorem 1 these are independent extrapolators and the proof is complete.

Lemma 4. Let x_0 *denote an arbitrary fixed point and let* $v(x)$ *denote a function of the exponomial space* XP. *Then*

$$v(x_0), \ \frac{dv}{dx}\Big|_{x=x_0}, \ \ldots, \ \frac{d^{m-1}v}{dx^{m-1}}\Big|_{x=x_0}$$

are a set of m *independent linear functionals on the space* XP.

Proof. The above expressions are obviously linear functionals. If they were not independent then there would be an identity of the form

$$\sum_{i=0}^{m-1} c_i v^{(i)}(x) = 0 ; \ x = x_0$$

which holds for v in XP. But if $v(x)$ is in XP so also is $v(x + k)$ for every k. Hence the above identity actually holds for all x. It is a differential equation of order $m - 1$ and can have at most $m - 1$ independent solutions. This is a contradiction, since there are m independent functions in XP,

Corollary 4. *The extrapolation of* y *together with its derivatives up to order* $m - 1$ *evaluated at the same point is given by the very short memory formula.*

Proof. The extrapolators are $p(x), p^{(1)}(x), \ldots, p^{(m-1)}(x)$ evaluated at a certain point $x = x_0$. Then the corollary follows from Lemma 4 and Theorem 7.

REFERENCES

1. Duffin, R. J. and Schmidt, Th. W., 'An extrapolator and scrutator', J. Math. Anal, and Appl. 1 (1960), 215-224.
2. Duffin, R. J. and Whidden, Phillips, 'An exponomial extrapolator', J. Math. Anal, and Appl. 3 (1961), 526-536.

3. Morrison, Norman, "Smoothing and extrapolation of time series by means of discrete Laguerre Polynomials", SIAM Jour, (to appear).