2 mix

X-591-73-143

X- 70 419

DATA COMPRESSION PRESERVING STATISTICAL INDEPENDENCE

(NASA-TM-X-70419) DATA COMPRESSION PRESERVING STATISTICAL INDEPENDENCE (NASA) 33 p HC $3.75 CSCL 12A

N73-28606

Unclas

G3/19 10742

GEORG E. MORDUCH
WILLIAM M. RICE

MAY 1973

GSFC — GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

DATA COMPRESSION PRESERVING STATISTICAL INDEPENDENCE

Georg E. Morduch
Old Dominion System, Inc.
Gaithersburg, Maryland

and

William M. Rice
Systems Analysis and Requirements Branch
Geodynamics Program Division

PRECEDING PAGE BLANK NOT FILMED

DATA COMPRESSION PRESERVING STATISTICAL INDEPENDENCE

Georg E. Morduch and William M. Rice

ABSTRACT

The purpose of this study was to determine the optimum points of evaluation of data compressed by means of polynomial smoothing. It is shown that a set y of m statistically independent observations $Y(t_1)$, $Y(t_2)$, . . . $Y(t_m)$ of a quantity $X(t)$, which can be described by a $(n - 1)$th degree polynomial in time, may be represented by a set Z of n statistically independent "compressed observations" $Z(\tau_1)$, $Z(\tau_2)$, . . . $Z(\tau_n)$, such that the "compressed" set Z has the same information content as the observed set Y. The times τ_1, τ_2, . . . τ_n are the zeroes of an nth degree polynomial P_n, to whose definition and properties the bulk of this report is devoted. The polynomials P_n are defined as functions of the observation times $t_1, t_2, \ldots t_m$, and it is interesting to note that if the observation times are continuously distributed the polynomials P_n degenerate to Legendre polynomials.

The proposed data compression scheme is a little more complex than those usually employed, but has the advantage of preserving all the information content of the original observations.

PREFACE

It had been our original intention to publish as a separate report the sections in "The Application of Data Compression to Orbit Determination, Old Dominion Systems Inc., January 15, 1973, Contract NAS5-11902" dealing specifically with data compression. However, after some discussions we realized that the data compression that was described in the above mentioned report did not necessarily yield a set of statistically independent "compressed" measurements. Although this in itself does not invalidate the method, there are obvious advantages if the compressed points are statistically independent. In developing the data compression technique described in this report we therefore made it a basic requirement that the "compressed" points should be statistically independent.

PRECEDING PAGE BLANK NOT FILMED

CONTENTS

DATA COMPRESSION PRESERVING STATISTICAL INDEPENDENCE

1.0 Introduction and Statement of the Problem

1.1 Introduction

Data compression and data smoothing is used where large volumes of data must be processed. Furthermore, it is necessary to weight the data properly in order to reflect the confidence in each measurement. Tracking data that is to be used for orbit determination and analysis is an example of where compression and smoothing is required in order to achieve cost effectiveness and at the same time obtain "correct" results.

The "correctness" of results is highly dependent upon the compression and smoothing technique employed. Careful application of smoothing and compression which will yield a single point representation of time varying data may be very accurate but it may also be very destructive of information content.

All tracking data, to continue with our example, undergoes some transformation in programs known as pre-processors, where, in general, calibrations are applied and smoothing and compression is performed. Operational preprocessors make use of least squares polynomial smoothing. Data compression is accomplished by evaluating this polynomial at one or more points. The uncertainties or weights associated with these smoothed data points are seldom obtained, yet they are required in order to properly use the data. Furthermore, one must be careful where the least squares polynomial is evaluated lest serious error be introduced (i.e. the end points are notoriously poor fitting).

Although, as already mentioned, the main application of data compression is in the area of data preprocessing, the technique may also be applied within the main processing program. In this case the data being compressed is not the observational data, but the observational discrepancies, i.e. the differences between observed and predicted measurement values. (See "The Application of Data Compression to Orbit Determination" with particular emphasis on the processing of Minitrack observations, Contract NAS5-11902, Old Dominion Systems Inc., January 15, 1973.)

In this report we shall describe a data compression technique, in which the compressed data points carry the same amount of information as the original data points. In order that this technique may be applied the following two criteria must be met:

1. The "noise" at each data point must be random gaussian and the "noise" at any two points must be uncorrelated. (Although the second criterion rarely applied for real data it is almost invariable assumed to hold true in practical orbit determination.)

2. The "noise free" data must fit some polynomial of degree $\nu - 1$, say, in time.

In order that no information be lost the number of compressed data points must be ν. Statistical independence will, however, still be preserved if the number of compressed points is less than ν. (In order not to introduce any errors into the compressed points it is, of course, necessary that the fitted polynomial be of degree $\nu - 1$ as before.)

In essence, the technique developed will yield a set of compressed data points, which also satisfy the above two criteria, with the addional obvious requirement that the polynomial of the second criterion be the same as for the original set of points.

In Section 1.2 below we define the problem in mathematical terms. In Section 2 we obtain the solution of the problem in terms of a set of orthogonal polynomials P_n, which are defined in Section 3. In this section we also derive all properties of the polynomials required in Section 2. In Section 4 we show that "underfit" (i.e. the effect of assuming that the polynomial of criterion 2 above, is one degree lower than it really is) by itself does not introduce any error in the compressed data points. Finally, in Section 5 we establish that the zeroes of any two of our polynomials, whose degree differ by one, are interleaved. This result adds to our understanding of the behavior of these polynomials. It is also an interesting general result since for example Legendre Polynomials are but particular cases of these polynomials.

1.2 Statement of the Problem

Given a set of data represented by the vector Y, can we find another vector Z, which will give an equally good representation of the data.

It is assumed that Y is an N dimensional vector of the form

$$Y^T = (y_0, y_1, \ldots, y_{N-1}) \tag{1.1}$$

where each y_i can be represented as a polynomial in t_i of degree $\nu - 1$ plus some error or noise e_i.

Thus

$$y_i = \sum_{\mu=0}^{\nu-1} b_\mu t_i^\mu + e_i \quad \text{for} \quad i = 0, 1, \ldots N-1 \tag{1.2}$$

where

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} \tag{1.3}$$

and where

$$\left.\begin{array}{ll} E(e_i) = 0 & \\ E(e_i\, e_j) = 0 \quad \text{if} \quad i \neq j \\ E(e_i\, e_i) = w_i^{-1}, & \end{array}\right\} \text{for} \quad i, j = 0, 1, \ldots N-1 \qquad \begin{array}{r} (1.4) \\ (1.5) \\ (1.6) \end{array}$$

where E denotes "the expected value of".

It is further assumed that

$$w_i > 0 \quad \text{for} \quad i = 0, 1, \ldots N-1 \tag{1.7}$$

and that

$$\nu \leq N \tag{1.8}$$

Equation (1.2) may equivalently be written in the form,

$$y_i = \sum_{\mu=0}^{\nu-1} a_\mu P_\mu(t_i) + e_i, \quad \text{for} \quad i = 0, 1, \ldots N-1 \tag{1.9}$$

where $P_\mu(t)$ is some polynomial in t of degree μ.

Equation (1.9) may be expressed in matrix form as,

$$Y = Fa + e, \tag{1.10}$$

where

$$a^T = (a_0, a_1, \ldots, a_{\nu-1}) \tag{1.11}$$

$$e^T = (e_0, e_1, \dots, e_{N-1}) \tag{1.12}$$

and

$$F = \begin{bmatrix} P_0(t_0) & P_1(t_0) & P_2(t_0) & \dots & P_{\nu-1}(t_0) \\ P_0(t_1) & P_1(t_1) & P_2(t_1) & \dots & P_{\nu-1}(t_1) \\ \vdots & \vdots & \vdots & & \vdots \\ P_0(t_{N-1}) & P_1(t_{N-1}) & P_2(t_{N-1}) & \dots & P_{\nu-1}(t_{N-1}) \end{bmatrix} \tag{1.13}$$

Equation (1.4) may be written as

$$E(e) = 0 \tag{1.14}$$

and equations (1.5) and (1.6) as,

$$E(ee^T) = W^{-1}. \tag{1.15}$$

where W is the diagonal matrix,

$$W = \begin{bmatrix} w_0 & 0 & 0 & \dots & 0 \\ 0 & w_1 & 0 & \dots & 0 \\ 0 & 0 & w_2 & \ddots & \vdots \\ 0 & 0 & 0 & & w_{N-1} \end{bmatrix} \tag{1.16}$$

The problem may now be stated as: Can we find a vector Z such that

$$Z = Ga + \epsilon, \tag{1.17}$$

where

$$Z^T = (z_0, z_1, \dots z_{\nu-1}), \tag{1.18}$$

$$\epsilon^T = (\epsilon_0, \epsilon_1, \dots \epsilon_{\nu-1}). \tag{1.19}$$

and

$$G = \begin{bmatrix} P_0(\tau_0) & P_1(\tau_0) & P_2(\tau_0) & \cdots & P_{\nu-1}(\tau_0) \\ P_0(\tau_1) & P_1(\tau_1) & P_2(\tau_1) & \cdots & P_{\nu-1}(\tau_0) \\ \vdots & & & & \\ P_0(\tau_{\nu-1}) & P_1(\tau_{\nu-1}) & P_2(\tau_{\nu-1}) & \cdots & P_{\nu-1}(\tau_{\nu-1}) \end{bmatrix} \tag{1.20}$$

and where

$$E(\epsilon) = 0 \tag{1.21}$$

$$E(\epsilon\, \epsilon^T) = V^{-1} \tag{1.22}$$

for some diagonal matrix,

$$V = \begin{bmatrix} v_0 & 0 & \cdots & 0 \\ 0 & v_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & v_{\nu-1} \end{bmatrix} \tag{1.23}$$

In other words can we find quantities

$$\tau_0,\ \tau_1, \ldots, \tau_{\nu-1}, \quad v_0,\ v_1, \ldots v_{\nu-1}$$

such that the above is true.

2.0 Solution of the Problem

In order to estimate the vector we proceed in the usual manner by obtaining a weighted least squares solution, i.e. we minimize the quantity

$$(Y - F\bar{a})^T\, W(Y - F\bar{a}),$$

where $\bar{a}$ denotes our estimate of a.

Letting the partial derivative of the above quantity with respect to a vanish we deduce that

$$F^T W Y = (F^T W F)\,\bar{a}.$$

It follows from the above and equation (1.10) that

$$\bar{a} = a + (F^TWF)^{-1} F^TWe \tag{2.1}$$

Hence if we write,

$$Z = G\bar{a},$$

it follows from the above and equation (1.17) that

$$\epsilon = G(F^TWF)^{-1} F^TWe \tag{2.2}$$

Since $E(e) = 0$ it follows that

$$E(\epsilon) = 0, \tag{2.3}$$

i.e. equation (1.21) is satisfied.

From equations (2.2) and (1.15) we deduce that

$$\begin{aligned} E(\epsilon\epsilon^T) &= G(F^TWF)^{-1} F^TWE(ee^T)WF(F^TWF)^{-1}G^T \\ &= G(F^TWF)^{-1} F^TWW^{-1}WF(F^TWF)^{-1}G^T \\ &= G(F^TWF)^{-1} G^T \end{aligned} \tag{2.4}$$

It will be shown in the next section that there exists a set of polynomials $P_0(t)$, $P_1(t), \ldots P_N(t)$ with the property that

$$\sum_{K=0}^{N-1} P_i(t_K)\, P_j(t_K)\, w_K = \delta_{ij}, \quad \text{for} \quad i, j - 0, 1, \ldots N-1 \tag{2.5}$$

where δ_{ij} is the Kronecker delta satisfying $\delta_{ij} = 0$ if $i \neq j$, and $\delta_{ij} = 1$ if $i = j$. These polynomials further have the property that their roots are all real and distinct. If the roots of $P_\nu(t) = 0$ are arranged in ascending order and are denoted by $\tau_0, \tau_1, \ldots \tau_{\nu-1}$, then

$$t_0 \leq \tau_0 < \tau_1 <, \ldots < \tau_{\nu-1} \leq t_{N-1} \tag{2.6}$$

Equation (2.5) may be expressed in matrix form as

$$F^TWF = I, \tag{2.7}$$

where I is the identity matrix. The polynomials also have the property that

$$GG^T = V^{-1}, \tag{2.8}$$

where V is a diagonal matrix. G being defined by equation (1.20) with τ_0, $\tau_1, \ldots, \tau_{\nu-1}$ being the roots of $P_\nu(t) = 0$.

It thus follows from equations (2.4), (2.7) and (2.8) that

$$E(\epsilon\epsilon^T) = V^{-1}, \tag{2.9}$$

i.e. equation (1.22) is satisfied.

In the next section we shall demonstrate that the polynomials $P_i(t)$ with the desired properties do indeed exist.

3.0 Definition of Orthogonal Polynomials $P_n(t)$

3.1 Definition of Polynomials $\overline{P}_n(t)$

Let us define the polynomial $\overline{P}_n(t)$* of degree n in t by the determinant

$$\overline{P}_n(t) = \begin{vmatrix} H_0 & H_1 & H_2 & \cdots & H_n \\ H_1 & H_2 & H_3 & \cdots & H_{n+1} \\ \vdots & & & & \vdots \\ H_{n-1} & H_n & H_{n+1} & \cdots & H_{2n-1} \\ 1 & t & t^2 & \cdots & t^n \end{vmatrix} \tag{3.1}$$

where

$$H_i = \sum_{K=0}^{N-1} t_K^i w_K \tag{3.2}$$

* It is important to note that the polynomial defined is $\overline{P}_n(t)$ and not $P_n(t)$. The relationship between these two polynomials is given in section 3.2.2 equation 3.24.

If we denote the cofactor of t^i in the determinant for $\overline{P}_n(t)$ by $C_{n,i}$ then equation (3.1) may be written in the form

$$\overline{P}_n(t) = \sum_{i=0}^{n} C_{n,i}\, t^i \tag{3.4}$$

3.2 Some Properties of the Polynomials $\overline{P}_n(t)$

3.2.1 Proof that $C_{n,n} > 0$

We shall first establish that for $0 \leq n \leq N$,

$$C_{n,n} > 0$$

By definition, $C_{n,n}$ is given by the determinant

$$C_{n,n} = \begin{vmatrix} H_0 & H_1 & \cdots & H_{n-1} \\ H_1 & H_2 & \cdots & H_n \\ \vdots & & & \\ H_{n-1} & H_{n-2} & \cdots & H_{2n-2} \end{vmatrix} \tag{3.5}$$

Let us define the $n \times n$ matrix A, as the matrix whose components A_{ij} are given by

$$A_{ij} = \sum_{K=0}^{N-1} t_K^i w_K t_K^j \tag{3.6}$$

A comparison of definitions (3.2) and (3.6) shows that

$$H_{i+j} = A_{ij} \tag{3.7}$$

It hence follows from equation (3.5) that

$$C_{n,n} = |A| \tag{3.8}$$

To show that $|A| > 0$ it is sufficient to show that the symmetric matrix A is positive definite, i.e. that for any real non-zero vector x,

$$x^T A x > 0$$

If we define the $(n \times N)$ matrix E, as the matrix whose components E_{iK} are given by

$$E_{iK} = t_K^i \tag{3.9}$$

then by definition

$$A = EWE^T \tag{3.10}$$

We thus have to demonstrate that for any non-zero x

$$x^T EWE^T x > 0,$$

i.e. that $(x^T E)\ W\ (x^T E)^T > 0$.

Since W is a diagonal matrix whose diagonal elements are greater than zero the above is certainly true provided that $x^T E \neq 0$ for any real non-zero vector x.

In other words, we have to show that the r rows of the $n \times N$ matrix E are linearly independent, i.e. we have to show that following vectors are linearly independent,

$$\begin{array}{c}
(1,\ 1,\ \ldots\ 1) \\
(t_0,\ t_1,\ \ldots\ t_{N-1}) \\
(t_0^2,\ t_1^2,\ \ldots\ t_{N-1}^2) \\
\vdots \\
(t_0^{n-1},\ t_1^{n-1},\ \ldots\ t_{N-1}^{n-1})
\end{array}$$

Since $n \leq N$ it is clearly sufficient to show that the vectors

$$\begin{array}{c}
(1,\ 1,\ \ldots\ 1) \\
(t_0,\ t_1,\ \ldots\ t_{n-1}) \\
\vdots \\
(t_0^{n-1},\ t_1^{n-1},\ \ldots\ t_{n-1}^{n-1})
\end{array}$$

are linearly independent.

But the determinant of the above vectors is given by (see Appendix A)

$$(t_1 - t_0)(t_2 - t_0)(t_{n-1} - t_0)(t_2 - t_1) \cdots (t_{n-1} - t_1) \cdots (t_{n-1} - t_{n-2})$$

Since by assumption all the t's are different none of the terms appearing in the product can vanish. Therefore the determinant cannot vanish and thus the rows of the matrix E are linearly independent from which it follows that

$$C_{n,n} > 0 \quad \text{for} \quad n = 0, 1, \dots N \tag{3.11}$$

3.2.2 Orthogonal Properties and Definition of $P_n(t)$

Let $d(t_K)$ denote the vector,

$$d(t_K) = (1, t_K, t_K^2, \dots, t_K^n) \quad \text{for} \quad K = 0, 1, \dots N - 1 \tag{3.12}$$

and f_s the vector

$$f_s = (H_s, H_{s+1}, \dots H_{s+n}) \quad \text{for} \quad s = 0, 1, \dots n - 1 \tag{3.13}$$

We deduce from equations (3.2), (3.12) and (3.13) that

$$f_s = \sum_{K=0}^{N-1} t_K^s w_K d(t_K) \quad \text{for} \quad s = 0, 1, \dots n - 1 \tag{3.14}$$

It follows from equations (3.1) and (3.13) that

$$\overline{P}_n(t) = \begin{vmatrix} f_0 \\ f_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ f_{n-1} \\ d(t) \end{vmatrix}$$

Since each of the first n rows of the above determinant is a linear combination (given by equation (3.14)) of the N vectors $d(t_0)$, $d(t_1)$, . . . $d(t_{N-1})$ it follows that they cannot be linearly independent if n exceeds N, i.e.

$$\overline{P}_n(t) \equiv 0 \quad \text{for} \quad n > N \tag{3.16}$$

Similarly

$$\overline{P}_N(t_K) = 0 \quad \text{for} \quad K = 0,\ 1, \ldots N \tag{3.17}$$

Equations (3.4) and (3.17) imply that

$$\overline{P}_N(t) = c_{n,n}(t - t_0)(t - t_1)\ldots(t - t_{N-1}) \tag{3.18}$$

It follows from equations (3.14) and (3.15) that

$$\sum_{K=0}^{N-1} t_K^s w_k \overline{P}_n(t_K) = \begin{vmatrix} f_0 \\ f_1 \\ \vdots \\ f_{n-1} \\ f_s \end{vmatrix} \tag{3.19}$$

Hence

$$\sum_{K=0}^{N-1} t_K^s w_K \overline{P}_n(t_K) = 0 \quad \text{if} \quad 0 \le s < n \tag{3.20}$$

Also since

$$\begin{vmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{vmatrix} = C_{n+1,\, n+1}$$

since it is the cofactor of t^{n+1} in the expansion of $\overline{P}_{n+1}(t)$ it follows from equation (3.19) that

$$\sum_{K=0}^{N-1} t_K^n w_K \overline{P}_n(t_K) = C_{n+1,\, n+1} \tag{3.21}$$

It follows from equation (3.4) and (3.20) that

$$\sum_{K=0}^{n} \bar{P}_s(t_K)\, w_K \bar{P}_n(t_K) = 0 \quad \text{if} \quad s \neq n \tag{3.22}$$

[The inequality sign follows since by symmetry equation (3.22) must hold both for $s < n$ and $n < s$].

From equation (3.4), (3.20) and (3.21) we deduce that

$$\sum_{K=0}^{N-1} \bar{P}_n(t_K)\, w_K \bar{P}_n(t_K) = C_{n,n} C_{n+1,\, n+1} \tag{3.23}$$

Since by equation (3.11) $C_{n,n}$ and $C_{n+1,\, n+1}$ are positive, it follows from the above that we may define a polynomial $P_n(t)$ by

$$P_n(t) = \bar{P}_n(t)/\sqrt{C_{n,n} C_{n+1,\, n+1}} \quad \text{for} \quad n = 0,\ 1, \ldots\ N-1 \tag{3.24}$$

We may now deduce from equations (3.22), (3.23) and (3.24) that

$$\sum_{K=0}^{N-1} P_i(t_K)\, P_j(t_K)\, w_K = \delta_{ij}, \quad \text{for} \quad i,\ j = 0,\ 1, \ldots\ N-1 \tag{3.25}$$

[We have thus shown that the polynomials $P_i(t)$ satisfy equation (2.5) – which is the same as (3.25)].

3.2.3 Proof That the Zeroes of $P_n(t)$ are Real and Distinct

We shall next show that the zeroes of a polynomial $P_n(t)$ are real, distinct and lie in the range $t_0 \le t \le t_{N-1}$. The proof given here follows that given by Hildebrand [1] for orthogonal polynomials. Since any polynomial $g_i(t)$ of degree i may be written as a linear combination of polynomials $P_0, P_1, \ldots P_i$ it follows from equation (3.25) that

$$\sum_{K=0}^{K-1} g_i(t_K)\, P_j(t_K)\, w_K = 0 \quad \text{for} \quad i < j \tag{3.26}$$

Assume that $P_j(t)$ has complex zeroes. Since they appear in complex conjugate pairs there must be at least two of them x and $\bar{x}$ say. Then let $g_i(t) = P_j(t)/(t - x)(t - \bar{x})$ which is a polynomial of degree j - 2. It then follows from equation (3.26) that

$$\sum_{K=0}^{N-1} w_K P_j^2(t_K)/(t_K - x)(t_K - \bar{x}) = 0$$

but this is clearly impossible since all the quantities in the sum are positive. (Remember that $(t_K - x)(t_K - \bar{x}) = t_K^2 + |x|^2$. A similar argument leads to the conclusion that all zeroes are distinct [if two zeroes are coincident and equal to x, say we choose $g_i(t) = P_j(t)/(t - x)^2$].

If a zero is less than t_0 or greater than t_{N-1} and equals x say, then we choose $g_i(t) = P_i(t)/(t - x)$ and apply the same argument to show that equation (3.26) cannot possibly be satisfied.

We thus conclude that all roots are real, distinct and lie in the range to $t_0 \leq t \leq t_{N-1}$.

3.2.4 A Recurrence Relation Between the Polynomials $\bar{P}_i(t)$

(The derivation given here is based on that given by Hildebrand [2].) According to equation (3.4) we have

$$\bar{P}_{n+1}(t) = \sum_{i=0}^{n+1} C_{n+1,\,i}\, t^i \quad \text{and} \quad \bar{P}_n(t) = \sum_{i=0}^{n} C_{n,\,i}\, t^i$$

Hence,

$$C_{n,n}\bar{P}_{n+1}(t) - C_{n+1,\,n+1}\, t\bar{P}_n(t) = \sum_{i=1}^{n} (C_{n,n}C_{n+1,\,i} - C_{n+1,\,n+1}C_{n,\,i-1})\, t^i + C_{n,n}C_{n+1,0}$$

The right-hand side, whose leading term is $(C_{n,n}\, C_{n+1,n} - C_{n+1,\,n+1}\, C_{n,n-1})\, t^n$, is a polynomial of degree n. Hence it may be written as a linear combination of polynomials $\bar{P}_i(t)$ of order n and less. We may, therefore write

$$C_{n,n}[C_{n,n}\bar{P}_{n+1}(t) - C_{n+1,\,n+1}t\bar{P}_n(t)] = (C_{n,n}C_{n+1,n} - C_{n+1,n+1}C_{n,n-1})\bar{P}_n(t) + \sum_{i=0}^{n-1} h_i\bar{P}_i(t) \tag{3.27}$$

where the coefficients h_i are to be determined. Since by equation (3.20)

$$\sum_{K=0}^{N-1} t_K^s w_K \overline{P}_n(t_K) = 0$$

for $s = 0, 1, \ldots n-1$ we deduce from the above that $h_0 = h_1 = \ldots = h_{n-2} = 0$ [we do this by successively performing the summation implied by equation (3.20) for $s = 0, 1, \ldots, n-2$]. Performing the summation with $s = n-1$, we obtain

$$-C_{n,n}C_{n+1,n+1}\sum_{K=0}^{N-1} t_K^n \overline{P}_n(t_K) = h_{n-1}\sum_{K=0}^{N-1} t_K^{n-1}\overline{P}_{n-1}(t_K),$$

whence by equation (3.21) $h_{n-1} = -(C_{n+1,n+1})^2$. Equation (3.27) therefore reduces to

$$(C_{n,n})^2\overline{P}_{n+1}(t) - [C_{n,n}C_{n+1,n+1}t + (C_{n,n}C_{n+1,n} - C_{n+1,n+1}C_{n,n-1})]\,\overline{P}_n(t) + (C_{n+1,n+1})^2\overline{P}_{n-1}(t) = 0 \tag{3.28}$$

Equation (3.28) holds true for all values of n, provided that we define $\overline{P}_{-1}(t)$ to be identically equal to zero. Note that the equation becomes meaningless for $n > N$ since in that case all terms will vanish.

3.3 The Existence of an Equivalent Set of Points

We shall show that there exists a set of ν points (where $\nu \le N$) $\tau_0, \tau_1, \ldots \tau_{\nu-1}$ and associated with these points a set of non-zero positive numbers $v_0, v_1, \ldots v_{\nu-1}$ such that

$$\sum_{\mu=0}^{\nu-1} \tau_\mu^i v_\mu = \sum_{K=0}^{N-1} t_K^i w_K \quad \text{for} \quad i = 0, 1, \ldots \nu - 1 \tag{3.29}$$

If equation (3.29) were true then it would follow from equations (3.1) and (3.2) that we could construct our polynomials $\overline{P}_i(t)$ (for $i = 0, 1, \ldots \nu$) over the set $(\tau_0, \tau_1, \ldots \tau_\nu)$ with the associated set of weights $v_0, v_1, \ldots v_\nu$. But then it would follow by analogy with equation (3.17) that

$$\overline{P}_\nu(\tau_\mu) = 0 \quad \text{for} \quad \mu = 0, 1, \ldots \nu - 1 \tag{3.30}$$

In other words $(\tau_0, \tau_1, \ldots \tau_{\nu-1})$ are the zeroes of the polynomial $\overline{P}_\nu(t)$. In Section 3.2.3 we established that the zeroes of a polynomial $\overline{P}_i(t)$ are all real and distinct and lie in the range $t_0 \leq t \leq t_{N-1}$. If we denote the smallest zero by τ_0, the next one by τ_1, etc. it follows that,

$$t_0 \leq \tau_0 < \tau_1 < \tau_2 \cdots < \tau_{\nu-1} \leq t_{N-1} \tag{3.31}$$

Next we are going to show that if the τ's are the zeroes of $P_\nu(t)$ then only ν of the 2ν equations (3.29) are independent. In other words we shall show that if

$$\sum_{\mu=0}^{\nu-1} \tau_\mu^i v_\mu = \sum_{K=0}^{\nu-1} t_K^i w_K \quad \text{for} \quad i = 0, 1, \ldots \nu - 1 \tag{3.31}$$

then the remainder of equations (3.29) will follow. The proof is by induction. Let us assume that

$$\sum_{\mu=0}^{\nu-1} \tau_\mu^i v_\mu = \sum_{K=0}^{N-1} t_K^i w_K \quad \text{for} \quad i = 0, 1, \ldots s - 1 \tag{3.32}$$

where $\nu \leq s < 2\nu$. The righthand side of equation (3.32) by definition (3.2) equals H_i. The last ν of the above equations may then be written as

$$\sum_{\mu=0}^{\nu-1} \tau_\mu^j v_\mu \tau_\mu^m = H_{j+m}, \quad \text{where} \quad j = s - \nu, \quad \text{and} \quad m = 0, 1, \ldots \nu - 1 \tag{3.33}$$

Multiplying both sides of the above equation by $C_{\nu,m}$ and summing over m we deduce that

$$\sum_{\mu=0}^{\nu-1} \tau_\mu^j v_\mu \sum_{m=0}^{\nu-1} C_{\nu,m} \tau_\mu^m = \sum_{m=0}^{\nu-1} C_{\nu,m} H_{j+m} \tag{3.34}$$

Since by equation (3.4)

$$\overline{P}_\nu(\tau_\mu) = \sum_{m=0}^{\nu} C_{\nu,m} \tau_\mu^m$$

and since τ_μ is a zero of $\overline{P}_\nu(t)$ it follows that

$$\sum_{m=0}^{\nu-1} C_{\nu,m} \tau_\mu^m = - C_{\nu,\nu} \tau_\mu^\nu \tag{3.35}$$

Since $C_{\nu,m}$ is the cofactor of t^m, the (m + 1) th element in the (ν + 1)th row in the determinant defining $P_\nu(t)$, and since H_{j+m} is the (m + 1)th element in the (j + 1)th row it follows from the theory of determinants that (provided that $j < \nu$)

$$\sum_{m=0}^{\nu} C_{\nu,m} H_{j+m} = 0,$$

and hence that

$$\sum_{m=0}^{\nu-1} C_{\nu,m} H_{j+m} = - C_{\nu,\nu} H_{j+\nu} \tag{3.36}$$

It follows from equations (3.34), (3.35) and (3.36) that

$$\sum_{\mu=0}^{\nu-1} \tau_\mu^j v_\mu \tau_\mu^\nu = H_{j+\nu} \quad \text{for} \quad j < \nu \tag{3.37}$$

We have thus shown that equation (3.33) is valid for $m = \nu$. It follows by induction that equations (3.29) may be deduced from equations (3.30) and (3.31). By an argument identical to the one used at the end of para. 3.2.1 above it follows that the $\nu \times \nu$ matrix of coefficients t_μ^i has a non-zero determinant so that the matrix is non-singular. Hence we can use the set of ν linear equations (3.31) to solve for $v_0, v_1, \ldots, v_{\nu-1}$. We have thus demonstrated the existence of a set of points ($\tau_0, \tau_1, \ldots \tau_{\nu-1}$) and the j associated numbers $v_0, v_1, \ldots v_{\nu-1}$) which satisfy equation (3.29).

Equations (3.24) may be written in the equivalent form

$$\sum_{\mu=0}^{\nu-1} q_i(\tau_\mu) v_\mu = \sum_{K=0}^{N-1} q_i(t_K) w_K \quad \text{for} \quad i = 0, 1, \ldots 2\nu - 1 \tag{3.38}$$

where $q_i(t)$ is any ith degree polynomial in t. In particular, let us choose $q_i(t) = P_m(t) P_n(t)$ where $m + n = i$. Then

$$\sum_{\mu=0}^{\nu-1} P_m(\tau_\mu) P_n(\tau_\mu) v_\mu = \sum_{K=0}^{N-1} P_m(t_K) P_n(t_K) w_K \quad \text{for} \quad m, n = 0, 1, \ldots \nu - 1$$

But by equation (3.25) the righthand side of the above equation may be written as δ_{mn} it follows that

$$\sum_{\mu=0}^{\nu-1} P_m(\tau_\mu)\, P_n(\tau_\mu)\, v_\mu = \delta_{mn} \quad \text{for} \quad m, n = 0, 1, \ldots \nu - 1. \tag{3.39}$$

In matrix form this may be expressed as

$$G^T V G = I, \tag{3.40}$$

where I is the $\nu \times \nu$ identity matrix, G is defined by equation (1.20) and V by equation (1.23).

Inverting both sides of equation (3.40) we obtain

$$G^{-1} V^{-1} (G^T)^{-1} = I.$$

Hence

$$G G^T = V^{-1} \tag{3.41}$$

which is the same as equation (2.8). Thus we have derived all results which we set out to obtain.

3.4 Some Additional Properties of the Polynomial P_i (t)

In non-matrix notation equations (3.41) may be written as

$$\sum_{i=0}^{\nu-1} P_i(\tau_m)\, P_i(\tau_n) = v_m^{-1}\, \delta_{mn} \tag{3.42}$$

if m = n equation (3.42) becomes

$$\sum_{i=0}^{\nu-1} [P_i(\tau_m)]^2 = v_m^{-1} \tag{3.43}$$

One consequence of equation (3.43) is that the set of polynomials P_0 (t), P_1 (t), . . . $P_{\nu-1}$(t) cannot have a common zero. Since we may deduce from equation (3.28) that if two polynomials, whose degrees differ by one, have the same zero, then all polynomials have the same zero. But since this is impossible we conclude that two polynomials, whose degree differ by one cannot have the same zero.

In general, we can now say that the zeroes of the polynomial $P_n(t)$ cannot equal the zeroes of $P_{n+1}(t)$ and are bounded above and below by the largest and smallest zero of $P_{n+1}(t)$.

3.5 Relationship to Gaussian Quadrature

If the set of points $t_0, t_1, \ldots t_{N-1}$ is infinite the sum on the righthand side of equation (3.38) gets replaced by an integral. We thus obtain

$$\sum_{\mu=0}^{\nu-1} g_i(\tau_\mu)\, v_\mu = \int_a^b g_i(t)\, w(t)\, dt, \quad \text{for} \quad i = 0, 1, \ldots 2\nu - 1, \tag{3.44}$$

where g_i is any ith degree polynomial in t. Equation (3.44) tells us that we may approximate the above integral by a finite sum.

3.6 Uniqueness of the Orthogonal Polynomials

We shall show that the polynomials $P_i(t)$ are the only ith degree polynomial having the property that,

$$\sum_{K=0}^{N-1} P_i(t_K)\, P_j(t_K)\, w_K = \delta_{ij} \quad \text{for} \quad i, j = 0, 1, \ldots N - 1 \tag{3.45}$$

Suppose

$$\sum_{K=0}^{N} Q_i(t_K)\, Q_j(t_K)\, w_K = \delta_{ij} \tag{3.46}$$

Since $Q_i(t)$ is an ith degree polynomial in t we may express it as a linear combination of $P_0(t)$, $P_1(t)$, . . . $P_i(t)$. Thus there exists a set of numbers h_j such that

$$Q_i(t) = h_0 P_0(t) + h_1 P_1(t) + \ldots h_i P_i(t) \tag{3.47}$$

It follows from equations (3.45) and (3.47) that

$$\sum_{K=0}^{N-1} Q_i(t_K)\, P_j(t_K)\, w_K = h_j, \quad \text{for} \quad j = 0, 1, \ldots\ i \tag{3.48}$$

But equation (3.46) implies that for any polynomial R_i (t),

$$\sum_{K=0}^{N-1} Q_i(t_K)\, R_j(t_K)\, w_K = 0 \quad \text{if} \quad j < i$$

We hence deduce from equation (3.48) that $h_0 = h_1 = \ldots h_{i-1} = 0$, whence equation (3.47) becomes

$$Q_i(t) = h_i P_i(t) \tag{3.49}$$

Substituting the expression for $Q_i(t)$ from equation (3.49) on the lefthand side of equation (3.46) (with j = i) we obtain,

$$\sum_{K=0}^{N-1} h_i^2\, [P_i(t_K)]^2\, w_K = 1 \tag{3.50}$$

We hence conclude from equations (3.50) and (3.45) that

$$h_i^2 = 1 \tag{3.51}$$

Thus $Q_i(t) = \pm P_i(t)$ for $i = 0, \ldots N - 1$. The statement that $P_i(t)$ is unique is, therefore, not absolutely true since the sign is arbitrary.

The fact that the orthogonal polynomials are unique implies that if the set of points $t_0, t_1, \ldots t_{N-1}$ are evenly distributed in the range (-1, 1) and if W = 1/N, then our polynomials $P_n(t)$ will degenerate to normalized Legendre polynomials as $N \to \infty$.

4. The Effect of Using a "Fitting" Polynomial, which is One Degree Too Low

In this section we shall investigate the effect of the components of the data vector being of the form,

$$y_i = \sum_{\mu=0}^{\nu} a_\mu P_\mu(t_i) + e_i, \tag{4.1}$$

rather than by equation (1.9). Corresponding to equation (1.10) we then have,

$$Y = Fa + M + e, \tag{4.2}$$

where the (i + 1)th component of the N-vector M is given by

$$M_i = a_\nu P_\nu(t_i) \tag{4.3}$$

Corresponding to equation (2.1) we now have,

$$\overline{a} = a + (F^T wF)^{-1} F^T w(M + e) \tag{4.4}$$

Instead of equation (1.17) we have,

$$Z = Ga + S + \epsilon \tag{4.5}$$

where the (i + 1)th component of the ν-vector S is given by

$$\begin{aligned} S_i &= a_\nu \overline{P}_\nu(\tau_i) \\ &= 0, \quad \text{inaccordance with equation (3.30)} \end{aligned}$$

Therefore,

$$Z = Ga + \epsilon \tag{4.6}$$

Corresponding to equation (2.2) we then have

$$\epsilon = G(F^T WF)^{-1} F^T W(M + e) \tag{4.7}$$

The (i + 1)th component of the ν-vector $F^T w M$ is given by,

$$\begin{aligned} (F^T WM)_i &= \sum_{K=0}^{N-1} F_{Ki} w_K M_K \\ &= \sum_{K=0}^{N-1} P_i(t_K) w_K a_\nu P_\nu(t_K) \\ &= a_\nu \delta_{i\nu}, \quad \text{by equation (3.25)} \end{aligned} \tag{4.8}$$

Since by equation (2.7) $F^TW\,F = I$, it follows from equation (4.7) and (4.8) that the (j + 1)th component of ϵ is given by

$$\epsilon_j = (GF^TWe)_j + \sum_{i=0}^{\nu-1} G_{jK}\, a_\nu\, \delta_{i\nu}$$

The sum on the righthand side of the above equation vanishes since i is always less than ν. We therefore conclude that

$$\epsilon = GF^TWe, \tag{4.9}$$

which is the same as equation (2.2) [since $F^TWF = I$ by equation (2.7)].

In other words "underfitting" (i.e. the effect of assuming that the given data Y can be described by a polynomial of degree $\nu - 1$, when in reality a polynomial of degree ν is required) does not by itself introduce an error in the compressed data points. On the other hand, some information loss obviously does result.

5.0 The Zeroes of $\overline{P}_S(t)$ are Separated by the Zeroes of $\overline{P}_{S-1}(t)$

Proof: First we shall show that between each pair of points t_s and t_{s+1} lies one and only one zero of $\overline{P}_{N-1}(t)$. We may deduce from equation (C.5) in Appendix 3 that the sign of $\overline{P}_{N-1}(t_s)$ is positive if the number of terms of the form $(t_s - t_K)$, where $k > s$, is even. In other words $\overline{P}_{N-1}(t_s)$ is positive if $(N - 1) - (S + 1) + 1$ is even, i.e.

$$\overline{P}_{N-1}(t_s) \text{ is positive if } (N - 1 - S) \text{ is even.}$$

It follows that $\overline{P}_{N-1}(t_s)$ and $\overline{P}_{N-1}(t_{s+1})$ must be of opposite sign. Since $\overline{P}_{N-1}(t)$ is a polynomial and hence a continuous function we deduce that $\overline{P}_{N-1}(t)$ must have a zero between t_s and t_{s+1}. It thus follows that $\overline{P}_{N-1}(t)$ must have a zero between t_0 and t_1, t_1 and t_2, . . . , t_{N-2} and t_{N-1}. Since the N - 1 zeroes of $\overline{P}_{N-1}(t)$ are thus accounted for, it follows that $\overline{P}_{N-1}(t)$ has one and only one zero between any pair of points t_0 and t_{s+1}. Since the points $t_0, t_1, \ldots t_{N+1}$ are the zeroes of $\overline{P}_N(t)$ we have shown that the zeroes of $\overline{P}(t)$ are separated by the zeroes of $\overline{P}_{N-1}(t)$. In Section (3.3) we showed that the set of polynomials $\overline{P}_0(t), \overline{P}_1(t), \ldots \overline{P}_{s-1}(t)$ may be constructed on the zeroes of $\overline{P}_s(t)$. Hence the result follows.

6.0 Acknowledgments

The authors wish to acknowledge the very useful discussions with Mr. E. Lefferts and Dr. P. Argentiero which eventually led us to our present scheme of data compression to statistically independent data points.

7.0 References

1. Hildebrand, F. B., Introduction ot Numerical Analysis, McGraw-Hill, 1956, Chapter 7, Section 7.4.

2. Hildebrand, F. B., Introduction to Numerical Analysis, McGraw-Hill, 1956, Chapter 8, Section 8.4.

APPENDIX A

The Determinant D Whose Elements d_{ij} are Given by $d_{ij} = t_j^i$ is Equal to the Product of all Terms of the Form $t_i - t_j$, where $i > j$ (If D is an $N \times N$ determinant then its elements are given by d_{ij} where $i, j = 0, 1, \ldots N - 1$)

Proof: By definition

$$D = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ t_0 & t_1 & t_2 & \cdots & t_{N-1} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ t_0^{N-1} & t_1^{N-1} & t_2^{N-1} & \cdots & t_{N-1}^{N-1} \end{vmatrix} \qquad \text{(A.1)}$$

Since the determinant vanishes if any two columns are identical it will vanish if any of the t_i's are equal. Hence D must be of the form,

$$D = KP \qquad \text{(A.2)}$$

where

$$P = \prod_{i=1}^{N-1} \prod_{j=0}^{i-1} (t_i - t_j) \qquad \text{(A.3)}$$

and where K is some polynomial in the t_i's. From equation (A.1) we see that D is a polynomial of degree $(1 + 2 + \ldots N - 1)$ in the t_i's. But the degree of the polynomial P is the same, since each term in the product is of degree 1 and the number of terms is $(1 + 2 + \ldots + N - 1)$. It then follows from equation (4.2) that K must be of degree zero, i.e. K is a constant.

To show that $K = 1$ we look at the coefficient of $t_1 \, t_2^2 \, t_3^3 \ldots t_{N-1}^{N-1}$ both in D and in P. It is easily seen that there is only one term of the form $t_1 \, t_2^2 \, t_3^3 \ldots t_{N-1}^{N-1}$ in equation (A.1). This is the product of the diagonal elements. The desired coefficient in D thus equals 1.

Since,

$$t_1 t_2^2 t_3^3 \cdots t_{N-1}^{N-1} = \prod_{i=1}^{N-1} \prod_{j=0}^{i-1} t_i$$

it follows from equation (A.3) that the desired coefficient in P also equals 1. Hence K must equal 1. Therefore D = P, i.e.

$$D = \prod_{i=1}^{N-1} \prod_{j=0}^{i-1} (t_i - t_j) \tag{A.4}$$

APPENDIX B

The Expression of the Polynomial $\overline{P}_N(t)$ as a Product

In accordance with equation (3.1), $\overline{P}_N(t)$ is given by

$$\overline{P}_N(t) = \begin{vmatrix} H_0 & H_1 & H_2 & \cdots & H_N \\ H_1 & H_2 & H_3 & \cdots & H_{N+1} \\ \vdots & & & & \\ H_{N-1} & H_N & H_{N+1} & \cdots & H_{2N-1} \\ 1 & t & t^2 & \cdots & t^N \end{vmatrix} \tag{B.1}$$

where

$$H_i = \sum_{K=0}^{N-1} t_K^i w_K \tag{B.2}$$

Let us define the $N \times N$ matrices A, B and W by,

$$A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ t_0 & t_1 & t_2 & \cdots & t_{N-1} & t \\ \vdots & \vdots & \vdots & & \vdots & \\ t_0^{N-1} & t_1^{N-1} & t_2^{N-1} & \cdots & t_{N-1}^{N-1} & t^{N-1} \\ T_0^N & t_1^N & t_2^N & \cdots & t_{N-1}^N & t^N \end{bmatrix} B = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 0 \\ t_0 & t_1 & t_2 & \cdots & t_{N-1} & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ t^{N-1} & t_1^{N-1} & t_2^{N-1} & \cdots & t_{N-1}^{N-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{B.3}$$

and

$$W = \begin{bmatrix} w_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & w_1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & w_{N-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{B.4}$$

It is easily verified that

$$\bar{P}_N(t) = |B^T W A|$$

$$= |B|\,|W|\,|A| \qquad \text{(B.5)}$$

since the determinant of the product of square matrices equals the product of their determinants.

Since W is a diagonal matrix it is easily seen that

$$|W| = \prod_{i=0}^{N-1} w_i \qquad \text{(B.6)}$$

It follows from equation (B.3) and the result of Appendix 1 that

$$|A| = \left\{ \prod_{i=1}^{N-1} \prod_{j=0}^{i-1} (t_i - t_j) \right\} \prod_{K=0}^{N-1} (t - t_K) \qquad \text{(B.7)}$$

and

$$|B| = \prod_{i=1}^{N-1} \prod_{j=0}^{i-1} (t_i - t_j) \qquad \text{(B.8)}$$

Hence we deduce from equations (B.5), (B.6), (B.7) and (B.8) that,

$$\bar{P}_N(t) = \left\{ \prod_{i=1}^{N-1} \prod_{j=0}^{i-1} (t_i - t_j) \right\}^2 \prod_{K=0}^{N-1} w_K(t - t_K) \qquad \text{(B.9)}$$

APPENDIX C

The Expression of $\overline{P}_{N-1}(t_s)$ as a Product (for s = 0, 1, . . . N - 1)

By definition (3.1),

$$\overline{P}_{N-1}(t_j) = \begin{vmatrix} H_0 & H_1 & H_2 \cdots & H_{N-1} \\ H_1 & H_2 & H_3 \cdots & H_N \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{N-2} & H_{N-1} & H_N \cdots & H_{2N-3} \\ 1 & t_s & t_s^2 \cdots & t_s^{N-1} \end{vmatrix}, \tag{C.1}$$

where

$$H_i = \sum_{K=0}^{N-1} t_K^i w_K \tag{C.2}$$

If we now subtract $(w_j \, t_j^i)$ times the last row from the (i + 1)th row the value of the determinant will remain unchanged. We then obtain,

$$\overline{P}_{N-1}(t_j) = \begin{vmatrix} H_0' & H_1' & H_2' \cdots & H_{N-1}' \\ H_1' & H_2' & H_3' \cdots & H_N' \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ H_{N-2}' & H_{N-1}' & H_N' \cdots & H_{2N-3}' \\ 1 & t_s & t_s^2 \cdots & t_s^{N-1} \end{vmatrix} \tag{C.3}$$

where

$$H_i' = \sum_{\substack{K=0 \\ K \neq s}}^{N-1} t_K^i w_K \tag{C.4}$$

Comparing equations (C.3), (C.4) with equations (B.1), (B.2) and (B.4) in Appendix B we deduce that

$$\overline{P}_{N-1}(t_s) = \left\{ \prod_{\substack{i=1 \\ i \neq s}}^{N-1} \prod_{\substack{j=0 \\ j \neq s}}^{i-1} (t_i - t_j) \right\}^2 \prod_{\substack{K=0 \\ K \neq s}}^{N-1} w_K (t_s - t_K) \qquad \text{(C.5)}$$