JOURNAL OF APPROXIMATION THEORY 28, 299–308 (1980)

Approximation of the Unit Step Function by a Linear Combination of Exponential Functions

J. SULLIVAN

Center for Environmental Assessment Services, Environmental Data and Information Service, National Oceanic and Atmospheric Administration, Washington, D.C. 20235

L. CRONE

National Environmental Satellite Service, National Oceanic and Atmospheric Administration, Washington, D.C. 20233

AND

J. JALICKEE

Center for Environmental Assessment Services, Environmental Data and Information Service, National Oceanic and Atmospheric Administration, Washington, D.C. 20235

Communicated by Oved Shisha

Received August 30, 1978

We approximate the unit step function, which equals 1 if $t \in [0, T]$ and equals 0 if $t > T$, by functions of the form $\sum_{n=1}^{N} A_n^{(N)} e^{-\lambda_n t/T}$, where each λ_n is a given positive constant. We find the coefficients $A_n^{(N)}$ by minimizing the integrated square of the difference between the unit step function and the approximating function. We first solve the specialized case where each $\lambda_n = n$. The resulting sum can be shown to converge in the mean to the unit step function as $N \to \infty$. The general case is then solved and some interesting properties of the numbers $A_n^{(N)}$ are noted.

INTRODUCTION

The problem addressed in this paper concerns the fitting of the unit step function by a series of exponential functions. It resembles Widder's [1] treatment of the passage from power series to Laplace transforms via Dirichlet series. The unit step function f we consider is defined by:

$$\begin{aligned} f(t) &= 1, \qquad 0 \leqslant t \leqslant T\,(T > 0); \\ f(t) &= 0, \qquad t > T. \end{aligned} \tag{1}$$

0021-9045/80/040299-10$02.00/0

Copyright © 1980 by Academic Press, Inc.
All rights of reproduction in any form reserved.

The functions to be used to approximate the unit step function are of the form

$$F_N(t) = \sum_{n=1}^{N} A_n^{(N)} e^{-\lambda_n t/T}. \tag{2}$$

N is a given integer ≥ 1, and also λ_n are fixed numbers: $0 < \lambda_1 < \lambda_2 < \ldots,$ λ_N. The coefficients $A_n^{(N)}$ are found by requiring that they minimize the integral

$$\int_0^\infty [f(t) - F_N(t)]^2\, dt. \tag{3}$$

In the first part of the paper we treat the simple case where each $\lambda_n = n$. This leads to shifted Legendre polynomials, which are moderately familiar functions. More importantly, it suggests relations and techniques useful in solving the more general case. Next we digress and present some lesser-known properties of these polynomials which arose from our work.

In the second part of the paper, the general case of λ_n is treated. Finally, we show that, as $N \to \infty$ in the case $\lambda_n = n$, $F_N(t)$ approaches $f(t)$ in the mean, i.e., the integral (3) $\to 0$.

Specialized Case

We turn to the simple case $\lambda_n = n$. Carrying out the minimization procedure for (3) leads to the set of simultaneous equations

$$\sum_{n=1}^{N} A_n^{(N)}/(n - m) = [1 - (1/e)^m]/m, \qquad m = 1, 2, \ldots, N. \tag{4}$$

We solved this set in two steps. First, we ignored $1/e$ in the numerator of the right-hand side of (4). We denote the coefficient on the left of (4) by $a_n^{(N)}$ and obtain

$$\sum_{n=1}^{N} a_n^{(N)}/(n - m) = 1/m. \tag{5}$$

With the advantage of hindsight, Eqs. (5) can be easily solved by utilizing a property of the shifted or asymmetric Legendre polynomials. These are a set of orthogonal polynomials over [0, 1] with the weight function 1. There appears to be no standard symbol for representing them, but we feel it is least confusing to denote them $P_N(x)$, the symbol usually reserved for the symmetric Legendre polynomials. The function $P_N(x)$ is of the form

$$P_N(x) = \sum_{n=0}^{N} b_n^{(N)} x^n. \tag{6}$$

with $b_0^{(N)} = 1$ for all N. The polynomials have the property that

$$\int_0^1 P_N(x)\, x^{m-1}\, dx = 0, \qquad m = 1, 2, \ldots, N. \tag{7}$$

Substituting the expression for $P_N(x)$ from (6) into (7) results in

$$\sum_{n=0}^{N} b_n^{(N)}/(n+m) = 0.$$

This is (5), with $a_n^{(N)} = -b_n^{(N)}$ for $n > 0$. Thus $a_n^{(N)}$ can be taken as the negatives of the shifted Legendre polynomials coefficients. Their values are [2]:

$$a_n^{(N)} = (-1)^{n-1} \frac{(N+n)!}{(N-n)!\, n!\, n!} = (-1)^{n+1} \binom{N}{n}\binom{N+n}{n}. \tag{8}$$

Digression

In practice, we first obtained the coefficients in the form (8) and did not recognize their relation to the shifted Legendre polynomials. In the process of discovering this, we obtained a procedure for calculating Legendre polynomials that is a hybrid of the generating function approach and a Rodrigues-type formula. We found that the Nth shifted Legendre polynomial is

$$P_N(x) = \frac{1}{N!} \frac{\partial^N}{\partial y^N} [(y+x)^N/(y+1)^{N+1}]_{y=0}. \tag{9}$$

This can be shown most clearly by using Leibnitz's rule for the Nth derivative of the product of two functions. Applied to (9), we have

$$\frac{1}{N!} \frac{\partial^N}{\partial y^N} [(y+x)^N (y+1)^{-(N+1)}]$$
$$= \sum_{n=0}^{N} (-1)^n \binom{N}{n}\binom{N+n}{n} (y+x)^n (y+1)^{-(N+1+n)}. \tag{10}$$

Setting $y = 0$, we regain (6). An interesting by-product of this approach is that it can also be shown that the expression

$$\frac{2^{N+1}}{N!} \frac{\partial^N}{\partial y^N} [(y+x)^N/(y+1)^{N+1}]_{y=1}$$

gives the more familiar *symmetric* Legendre polynomial.

Return to the Specialized Case

We now solve the system (4) without ignoring any term. Upon computation of a few lower-order coefficients $A_n^{(N)}$, we note that they appear to be polynomials in $(1/e)$ times $a_n^{(N)}$ of (8). For example,

$$A_1^{(2)} = 6[1 - 3(1/e) + 2(1/e)^2]$$
$$= a_1^{(2)}[1 - 3(1/e) + 2(1/e)^2].$$

After various conjectures were tested, we found that

$$A_n^{(N)} = a_n^{(N)} \left[n e^n \int_0^{1/e} y^{n-1} P_N(y)\, dy \right], \qquad n = 1, \ldots, N. \tag{11}$$

We have not yet found a short proof that the right-hand side of (11) satisfies (4). However, our present proof contains an interesting intermediate result. The proof is as follows.

Using the fact that

$$P_N(y) = 1 - \sum_{n=1}^{N} a_n^{(N)} y^n,$$

the integration in (11) is performed and the result is substituted into (4). The term free of $(1/e)$ gives (5) back again, and so it remains to show that

$$\sum_{k=1}^{N} a_k^{(N)} (1/e)^k \left[\sum_{n=1}^{N} a_n^{(N)} n/((n+m)(n+k)) \right] = (1/e)^m/m. \tag{12}$$

for $m = 1, 2, \ldots, N$. The term in square brackets in (12) resembles the Kronecker delta function, equalling 0 when $k \neq m$. To see this, consider

$$- \int_0^1 y^m \frac{d}{dy} \left[\int_0^1 P_N(yz)\, z^{k-1}\, dz \right] dy, \tag{13}$$

and carry out the integrations systematically. This quantity is seen to equal the square-bracketed term in (12). Integrating in (13) by parts and using the property (7), we are led to posit that

$$\int_0^1 \int_0^1 y^{m-1} z^{k-1} P_N(yz)\, dy\, dz = 0, \qquad k \neq m. \tag{14}$$

for $1 \leqslant m \leqslant N$, $1 \leqslant k \leqslant N$. Relation (14) is a two-dimensional analogue of (7). It can be proved by transforming to new variables:

$$u = yz,$$
$$v = z.$$

and then integrating. When $m = k$, integration in the new variables shows that the polynomials defined in (11) will satisfy (4) if

$$\int_0^1 u^{m-1} P_N(u) \log u \, du = 1/(m^2 a_m^{(N)}). \tag{15}$$

The validity of (15) can be established directly by inductive reasoning involving the Rodrigues formula for $P_N(u)$. It is simpler to regard it as a particular case of the more general proof given in the next section.

An integral representation for this specialized, consecutive integer case is obtained by substituting in

$$F_N(t) = \sum_{n=1}^{N} A_n^{(N)} (e^{-t/T})^n$$

the value of $A_n^{(N)}$ from (11).

Interchanging summation and integration, we get

$$F_N(t) = -\int_0^{1/e} P_N(y) \frac{d}{dy} [P_N(eye^{-t/T})] \, dy.$$

With the variable change $w \equiv ey$, we finally obtain

$$F_N(t) = -\int_0^1 P_N(w/e) \frac{d}{dw} [P_N(we^{-t/T})] \, dw. \tag{16}$$

Using the integral representation (16) and the inequality [3]

$$P_N(w) \leqslant [w(1-w)]^{-1/2}/(2\pi N)^{1/2}, \tag{17}$$

we can examine the behaviour of $F_N(t)$ at certain points as $N \to \infty$. Note that $P_N(0) = 1$, $P_N(1) = (-1)^N$. The inequality in (17) shows that, for every fixed w in the interval (0, 1),

$$P_N(w) \to 0 \qquad \text{as} \quad N \to \infty.$$

If $t/T = 0$, integration of (16) gives

$$1 + (-1)^{N+1} P_N(1/e),$$

which converges to 1 as $N \to \infty$. If $t = T$, integration of (16) results in

$$F_N(t = T) = [1 - P_N{}^2(1/e)]/2.$$

As $N \to \infty$, $F_N(t = T) \to \frac{1}{2}$ at the point of discontinuity where the unit step function changes from 1 to 0.

General Case

For the general case, the set of equations that define the coefficients $A_n^{(N)}$ is

$$\sum_{n=1}^{N} A_n^{(N)}/(\lambda_n + \lambda_m) = [1 - (1/e)^{\lambda_m}]/\lambda_m, \qquad m = 1, 2, \dots, N \tag{18}$$

The $A_n^{(N)}$ are uniquely determined since the determinant of the coefficient matrix is non-zero as long as the λ's are distinct. This is a special case of a general theorem due to Cauchy [5]. Taking our cue from the technique that worked for the specialized case, we ignored the dependence on $1/e$ in the right-hand side of (18) and solved for the coefficients. In Appendix A we prove that the coefficients, denoted once again by $a_n^{(N)}$, are

$$a_n^{(N)} = 2 \prod_{k \neq n}^{N} (\lambda_k + \lambda_n)/(\lambda_k - \lambda_n), \qquad n = 1, 2, \dots, N. \tag{19}$$

These coefficients have some interesting properties. If all the λ's are multiplied by a constant scale factor, $\lambda_n = s\lambda_n$, $s \neq 0$, the coefficients remain unchanged. If the λ's are transformed into their reciprocals, $\lambda_n = 1/\lambda_n$, the new coefficients are the original ones multiplied by $(-1)^{N-1}$. A set of polynomials which in some respects can be regarded as a generalization of the shifted Legendre polynomials is defined by

$$R_n(x) = 1 - \sum_{n=1}^{N} a_n^{(N)} x^{\lambda_n}. \tag{20}$$

As for the shifted Legendre polynomials, $R_N(0) = 1$. Also, it is shown in Appendix A that

$$\sum_{n=1}^{N-1} a_n^{(N-1)} - \sum_{n=1}^{N} a_n^{(N)} = 1. \tag{21}$$

Since $a^{(1)} = 2$, (21) implies that $R_N(1) = (-1)^N$, the same as for the shifted Legendre polynomials.

To solve for the case where the dependence on $1/e$ in the right-hand side of (18) is not ignored, we use one more important property of the polynomials $R_N(x)$. Using (18), the analogue of (7) for $R_N(x)$ is

$$\int_0^1 R_N(x)\, x^{\lambda_m - 1}\, dx = 0, \qquad m = 1, 2, \dots, N. \tag{22}$$

as can be shown. Substituting (20) into (22) gives

$$1/\lambda_m - \prod_{n=1}^{N} a_n^{(N)}/(\lambda_n + \lambda_m) = 0, \tag{23}$$

which is the form (18) takes when the dependence on $1/e$ is ignored. Using the "quasi-orthogonal" relation (22), we can now prove that the set of equations (18) has the solution

$$A_n^{(N)} = a_n^{(N)} \left[\lambda_n e^{\lambda_n} \int_0^{1/e} y^{\lambda_n - 1} R_N(y)\, dy\right], \qquad n = 1, \ldots, N. \tag{24}$$

When (24) is substituted into (18), we obtain

$$\sum_{k=1}^{N} (1/e)^{\lambda_k} a_k^{(N)} \lambda_m \left[\sum_{n=1}^{N} a_n^{(N)} \lambda_n / ((\lambda_n - \lambda_m)(\lambda_n - \lambda_k))\right] = (1/e)^{\lambda_m}. \tag{25}$$

This closely resembles (12) (slightly rearranged) with the proviso that n now becomes λ_n, etc. That the bracketed term in (15) is 0 if $k \neq m$ can be proven by the same method that resulted in (13) and (14). We now prove that, when $k = m$, the term multiplying $(1/e)^{\lambda_m}$ on the left-hand side of (25) is equal to 1. From (23) we find that

$$\sum_{n=1}^{N} a_n^{(N)} \lambda_m / (\lambda_n - \lambda_m) = 1,$$

and further, that

$$\sum_{k=1}^{N} a_k^{(N)} \lambda_n \left[\sum_{n=1}^{N} a_n^{(N)} \lambda_m / (\lambda_n + \lambda_m)\right] \Big/ (\lambda_k - \lambda_n) = 1.$$

Rearranging, we obtain,

$$\sum_{k=1}^{N} a_k^{(N)} \lambda_m \left[\sum_{n=1}^{N} a_n^{(N)} \lambda_n / ((\lambda_n - \lambda_m)(\lambda_n + \lambda_k))\right] = 1. \tag{26}$$

The bracketed term in (26) is the same as the bracketed term in (25) and equals 0 when $k \neq m$. The only term contributing in (26) occurs when $k = m$ and must equal 1. This term is precisely the one that multiplies $(1/e)^{\lambda_m}$ in (25), which completes the proof. The validity of (15) follows from this proof.

Convergence in the Mean

For the specialized case where $\lambda_n = n$, we now show that the integral in (3) $\to 0$ as $N \to \infty$. Using (2) and letting the new variable $w \equiv e^{-t/T}$, minimizing (3) is equivalent to minimizing the integral

$$T \int_0^1 \left[f(-T \log w) - w \sum_{n=0}^{N-1} A_{n+1}^{(N)} w^n\right]^2 w^{-1}\, dw. \tag{3a}$$

We now assume that the sum of powers of w in (3a) is also the sum of a set of orthogonal polynomials,

$$\sum_{n=0}^{N-1} A_{n-1}^{(N)} w^n = \sum_{n=0}^{N-1} B_n^{(N)} G_n(w). \tag{27}$$

The function $G_n(w)$ is the nth polynomial and $B_n^{(N)}$ is the coefficient associated with it. When (27) is substituted into the integral in (3a) and the squaring of the bracketed term is carried out, we find that orthogonality requires that

$$\int_0^1 w G_n(w)\, G_{n'}(w)\, dw = 0 \qquad \text{for } n' \neq n. \tag{28}$$

The set of orthogonal polynomials which satisfy (28) is a particular set of hypergeometric polynomials of Jacobi [3]. They are denoted in the literature $G_n(2, 2, w)$ and satisfy the differential equation

$$w(1-w)\, G_n''(2, 2, w) + (2-3w)\, G_n'(2, 2, w) + n(n+2)\, G_n(2, 2, w) = 0.$$

It is known that these classical orthogonal polynomials form a complete set and thus the series developement in terms of them converges in the mean [4] to any piecewise continuous function in the interval $[0, 1]$. If the Jacobi polynomials are normalized so that the constant term equals 1, the connection between them and the shifted Legendre polynomials is

$$2nwG_{n-1}(2, 2, w) = P_{n-1}(w) - P_n(w).$$

Further examination shows that the same type of proof can be repeated for the case where $\lambda_n = sn$, s being a positive number. Therefore, as long as the λ_n are evenly spaced, $F_N(t)$ will converge in the mean to the unit step function.

Appendix A

The set of N equations (18) can be written, for the case where the dependence on $1/e$ is ignored, as

$$\begin{aligned} \frac{a_1^{(N)}}{2\lambda_1} + \frac{a_2^{(N)}}{\lambda_1+\lambda_2} + \cdots + \frac{a_N^{(N)}}{\lambda_1+\lambda_N} &= \frac{1}{\lambda_1} \\ &\vdots \\ \frac{a_1^{(N)}}{\lambda_N+\lambda_1} + \frac{a_2^{(N)}}{\lambda_N+\lambda_2} + \cdots + \frac{a_N^{(N)}}{2\lambda_N} &= \frac{1}{\lambda_N}. \end{aligned} \tag{29}$$

The unknown $a_N^{(N)}$ may be eliminated by dividing the mth equation by $2\lambda_N$, dividing the last equation by $\lambda_m + \lambda_N$, and then subtracting. This is done for $m = 1, 2, ..., N - 1$. After cancellation, a set of $N - 1$ equations is obtained;

$$\frac{b_1^{(N)}}{2\lambda_1} + \frac{b_2^{(N)}}{\lambda_1 + \lambda_2} + \cdots + \frac{b_{N-1}^{(N)}}{\lambda_1 + \lambda_{N-1}} = \frac{1}{\lambda_1}$$
$$\vdots$$
$$\frac{b_1^{(N)}}{\lambda_N + \lambda_1} + \frac{b_2^{(N)}}{\lambda_N + \lambda_2} + \cdots + \frac{b_{N-1}^{(N)}}{\lambda_N + \lambda_{N-1}} = \frac{1}{\lambda_{N-1}},$$

where

$$b_m^{(N)} = \frac{\lambda_N - \lambda_m}{\lambda_N + \lambda_m} a_m^{(N)}.$$

This same process can now be used to eliminate $b_{N-1}^{(N)}$, which leaves $N - 2$ equations to solve. Repeating this procedure until only $a_1^{(N)}$ is left gives

$$(\tfrac{1}{2}\lambda_1) \prod_{m\neq 1}^{N} (\lambda_m - \lambda_1)/(\lambda_m + \lambda_1)\, a_1^{(N)} = 1/\lambda_1,$$

or

$$a_1^{(N)} = 2 \prod_{m\neq 1}^{N} (\lambda_m + \lambda_1)/(\lambda_m - \lambda_1). \tag{30}$$

This same elimination procedure can be used to isolate any of the unknowns as the remaining one. This is how (19) is obtained.

Multiplying the last equation in (29) by $2\lambda_N$ gives

$$\sum_{n=1}^{N-1} a_n^{(N)} 2\lambda_N/(\lambda_N + \lambda_n) + a_N^{(N)} = 2.$$

Using the fact that $2\lambda_N/(\lambda_N + \lambda_n) = 1 + (\lambda_N - \lambda_n)/(\lambda_N + \lambda_n)$, and recognizing from (30) that

$$\frac{(\lambda_N - \lambda_n)}{(\lambda_N + \lambda_n)} a_n^{(N)} = a_n^{(N-1)},$$

we finally obtain

$$\sum_{n=1}^{N-1} a_n^{(N-1)} + \sum_{n=1}^{N} a_n^{(N)} = 2. \tag{31}$$

REFERENCES

1. D. V. WIDDER, "Advanced Calculus," pp. 437 ff, Prentice-Hall, Englewood Cliffs, N.J., 1961.
2. H. K. CROWDER AND S. W. MCCUSKEY, "Topics in Higher Analysis," p. 413, Macmillan, New York; Collier-Macmillan, London, 1964.
3. "Handbook of Mathematical Functions" (M. Abramowitz and I. A. Stegun, Eds.), Chap. 22, Dover, New York, 1965).
4. P. M. MORSE AND H. FESHBACK, "Methods of Mathematical Physics," Part I, pp. 736 ff, McGraw-Hill, New York, 1953.
5. A. RALSTON, "A First Course in Numerical Analysis," p. 263, McGraw-Hill, New York, 1965.