NOTES ON NUMERICAL ANALYSIS II

INTERPOLATION AND CURVE FITTING BY SECTIONALLY LINEAR FUNCTIONS

Hans Schwerdtfeger

(received August 5, 1959)

Functions employed for interpolation and curve fitting are for the most part polynomials with numerical coefficients. Indeed these are functions whose values for numerically given arguments can be computed directly without resorting to non-algebraic designs. It is known, however, that there are cases where polynomial interpolation does not yield an adequate approximation to a given function (cf. [4], p. 34).

Another class of functions whose values can be computed without the aid of tables or advanced methods is the class of the sectionally linear functions. Their use for the purpose of interpolation and curve fitting is further suggested by the well known fact that every function F(x), continuous over a closed interval [a, b], can be approximated uniformly by a sequence of sectionally linear functions over a distinguished sequence of partitions of the interval [a, b]. It seems that this approximation by sectionally linear functions which is based on a linear interpolation over the part intervals of the partitions, has not been studied so far for other than theoretical purposes; Lebesgue's proof of Weierstrass' first approximation theorem makes use of the process (cf. [5], p. 3). This interpolation will be considered here with explicit expressions for the approximating sectionally linear functions. In the same way also curve fitting by the method of least squares will be studied. Moreover an orthogonal system of sectionally linear functions over a given partition is introduced in order to simplify the computation of the coefficients of a certain expansion. The entirely algebraic character of the discussion makes the results immediately applicable to numerical work.

Can. Math. Bull. vol. 3, no. 1, Jan. 1960

1. Sectionally linear functions over a partition. Let $a = x_o < x_1 < \dots < x_{n-1} < x_n = b$ be a partition $\mathcal{P}_n$ of the interval $[a,b]$, i.e. $a \leq x \leq b$. A real function $f(x)$ is said to be sectionally linear over the partition $\mathcal{P}_n$ if its graph is a polygon curve having the points $(x_\nu, f(x_\nu))$ ($\nu = 0, 1, \dots, n$) as vertices. Clearly every such sectionally linear function is continuous over $[a,b]$. If $f(x)$ and $g(x)$ are two sectionally linear functions over $\mathcal{P}_n$, then also $\alpha f(x) + \beta g(x)$ is sectionally linear over $\mathcal{P}_n$, where α, β are real constants. All sectionally linear functions therefore form a linear space.

This space has the dimension $n + 1$.

Proof. Let

$$\varphi(t) = \tfrac{1}{2}(|t| + t) = \begin{cases} 0 & \text{if} \quad t \leq 0 \\ t & \text{if} \quad t > 0 \end{cases}.$$

Put

$$\varphi_o(x) \equiv 1, \quad \varphi_\nu(x) = \varphi(x - x_{\nu-1}) = \begin{cases} 0 & \text{if } x \leq x_{\nu-1} \\ x - x_{\nu-1} & \text{if } x \quad x_{\nu-1} \end{cases} (\nu = 1, \dots, n)$$

The functions $\varphi_\nu(x)$ are linearly independent. Indeed suppose that with real coefficients $\lambda_o, \lambda_1, \dots, \lambda_n$ the relation

$$\lambda_o + \lambda_1 \varphi_1(x) + \dots + \lambda_n \varphi_n(x) = 0$$

is satisfied for all x in $[a,b]$. For $x = x_o$ we have $\varphi_\nu(x_o) = 0$ ($\nu = 1, 2, \dots, n$) and therefore $\lambda_o = 0$. Put $x = x_1$; then $\varphi_1(x_1) = x_1 - x_o > 0$, but $\varphi_2(x_1) = 0, \dots, \varphi_n(x_1) = 0$; hence $\lambda_1 = 0$, etc. Thus the dimension of our space is at least $n + 1$ and if a function $f(x)$ can be written in the form

$$(1) \qquad f(x) = c_o + c_1 \varphi_1(x) + \dots + c_n \varphi_n(x)$$

then the coefficients $c_o, c_1, \dots, c_n$ are uniquely defined.

Conversely every sectionally linear function $f(x)$ can be written in the form (1) which is equivalent with

$$f(x_o) = c_o$$

$$f(x) = c_o - c_1 x_o + c_1 x \qquad x_o < x \leq x_1$$

$$f(x) = c_o - c_1 x_o - c_2 x_1 + (c_1 + c_2)x \qquad x_1 < x \leq x_2$$

. . . .

From here it is evident that the c_o, c_1, $c_2, \ldots$ can be adapted to every prescribed polygon curve over $\mathcal{P}_n$.

Explicit values of the c_ν can be expressed by means of the successive divided differences of the function f(x). We apply Newton's interpolation formula to the function f(x) over the partition $\mathcal{P}_n$; making use of Steffensen's notation (cf. [4], § 4) we thus obtain the following expression for f(x).

$$f(x) = f(x_o) + f(x_o,x_1)(x - x_o) + f(x_o,x_1,x_2)(x - x_o)(x - x_1)$$
$$+ \ldots + f(x_o,x_1,\ldots,x_{n-1})(x - x_o)(x - x_1)\ldots(x - x_{n-2})$$
$$+ f(x_o,x_1,\ldots,x_{n-1},x)(x - x_o)(x - x_1)\ldots(x - x_{n-2})(x - x_{n-1}).$$

Hence follows that

$$f(x_1) = f(x_o) + f(x_o,x_1)(x_1 - x_o)$$
$$f(x_2) = f(x_o) + f(x_o,x_1)(x_2 - x_o) + f(x_o,x_1,x_2)(x_2 - x_o)(x_2 - x_1)$$
$$f(x_3) = f(x_o) + f(x_o,x_1)(x_3 - x_o) + f(x_o,x_1,x_2)(x_3 - x_o)(x_3 - x_1)$$
$$+ f(x_o,x_1,x_2,x_3)(x_3 - x_o)(x_3 - x_1)(x_3 - x_2)$$

. . . .

On the other hand we have by (1)

$$f(x_o) = c_o$$
$$f(x_1) = c_o + c_1(x_1 - x_o)$$
$$f(x_2) = c_o + c_1(x_2 - x_o) + c_2(x_2 - x_1)$$
$$f(x_3) = c_o + c_1(x_3 - x_o) + c_2(x_3 - x_1) + c_3(x_3 - x_2)$$

. . . . ,

and since the c_ν are uniquely defined it follows that

$$c_o = f(x_o)$$
$$c_1 = f(x_o,x_1)$$
$$c_2 = f(x_o,x_1,x_2)(x_2 - x_o)$$
$$c_3 = f(x_o,x_1,x_2,x_3)(x_3 - x_o)(x_3 - x_1)$$

. . . .

https://doi.org/10.4153/CMB-1960-009-4 Published online by Cambridge University Press

$$c_n = f(x_0, x_1, \ldots, x_{n-1}, x_n)(x_n - x_0)(x_n - x_1)\ldots(x_n - x_{n-2}) .$$

These c_ν define the function $f(x)$, and its properties are reflected in certain properties of the sequence $c_0, c_1, \ldots, c_n$. Thus we find $f(x)$ to be monotonically increasing if and only if

$$c_1 > 0,\ c_1 + c_2 > 0,\ c_1 + c_2 + c_3 > 0 ,\ \ldots,$$

monotonically non-decreasing if in some (or all) of these relations the sign $>$ may be replaced by $=$. The function $f(x)$ will be strictly convex (from below) if the sequence c_1, $c_1 + c_2$, $c_1 + c_2 + c_3$, . . . is monotonically increasing. A convex function $f(x)$ will have its minimum within the interval $[x_{\mu-1}, x_\mu]$ if $c_1 + \ldots + c_{\mu-1} < 0$, $c_1 + \ldots + c_{\mu-1} + c_\mu = 0$, $c_1 + \ldots + c_{\mu+1} > 0$; the minimum will be at $x_{\mu-1}$ if

$$c_1 + \ldots + c_{\mu-1} < 0,\ c_1 + \ldots + c_{\mu-1} + c_\mu > 0 .$$

The equation $f(x) = 0$ can be solved by a trial and error method. Suppose for example that $f(x)$ is increasing and therefore the sums $c_1 + \ldots + c_\nu > 0$ ($\nu = 1, 2, \ldots$). The root of the equation will lie in the left-open interval $(x_{\mu-1}, x_\mu]$ if

$$x_{\mu-1} < - \frac{c_0 - c_1 x_0 - \ldots - c_\mu x_{\mu-1}}{c_1 + c_2 + \ldots + c_\mu} \leq x_\mu .$$

The middle term represents the value of the root.

This is in outline the elementary analysis of the sectionally linear functions.

In the case of a partition $\mathcal{P}_n$ with equidistant points

$$x_\nu = a + \nu(b - a)/n \qquad (\nu = 0, 1, \ldots, n)$$

one has in the usual notation

$$c_0 = f(a), \qquad c_\nu = (1/\Delta a)\ \Delta^\nu f(a) \qquad (\Delta a = (b - a)/n).$$

2. Sectionally linear functions, orthogonal over a partition. In most practical cases the direct computation of the coefficients c_ν will be tedious. Therefore we shall establish for $f(x)$ another expansion the coefficients of which can be computed more easily.

The inner product of two functions f(x), g(x) over the partition $\mathcal{P}_n$ is defined as

$$(f,g)_{\mathcal{P}_n} = (f,g) = \sum_{\nu=0}^{n} f(x_\nu) g(x_\nu) .$$

It has the usual formal properties of an inner product. It is symmetric, and linear and homogeneous in each of its factors f(x) and g(x). The square of the norm of f(x) over $\mathcal{P}_n$ is given by $(f,f) \geqslant 0$; in the case of a sectionally linear function f(x) one has $(f,f) = 0$ if and only if $f(x) \equiv 0$ over $[a,b]$.

The two functions f(x) and g(x) are said to be <u>orthogonal over the partition</u> $\mathcal{P}_n$ if $(f,g) = 0$ (cf. [1], p.80; [3], p.283).

The following formulae are readily established.

$$(f, \varphi_\mu) = \sum_{\nu=\mu}^{n} f(x_\nu)(x_\nu - x_{\mu-1})$$

and in particular

$$(\varphi_0, \varphi_0) = n+1, \quad (\varphi_\lambda, \varphi_\mu) = \sum_{\nu=\lambda}^{n} (x_\nu - x_{\lambda-1})(x_\nu - x_{\mu-1})$$

if $\lambda > \mu$.

Since the functions $\varphi_0(x), \ldots, \varphi_n(x)$ are linearly independent one can determine an orthogonal system $\psi_0(x), \ldots, \psi_n(x)$ spanning the same space so that

$$(2) \qquad (\psi_\mu, \psi_\nu) = 0 \quad \text{if } \mu \neq \nu .$$

The orthogonal functions $\psi_\nu(x)$ may be determined either by the well-known Schmidt orthogonalization process or, as it will be done here, by the determinant formulae.

Notice first that if

$$\tilde{f}(x) = \xi_0 + \xi_1 \varphi_1(x) + \ldots + \xi_n \varphi_n(x)$$

is a sectionally linear function with the variable coefficients ξ_ν, then its norm square

$$(f,f) = \sum_{\nu=0}^{n} f(x_\nu)^2 = \sum_{i=0}^{n} \sum_{j=0}^{n} (\varphi_i, \varphi_j) \xi_i \xi_j$$

is a positive definite quadratic form. Similarly one can see that the matrices

https://doi.org/10.4153/CMB-1960-009-4 Published online by Cambridge University Press

$$Q_o = \left((\varphi_o, \varphi_o)\right), \quad Q_1 = \begin{pmatrix} (\varphi_o, \varphi_o) & (\varphi_o, \varphi_1) \\ (\varphi_1, \varphi_o) & (\varphi_1, \varphi_1) \end{pmatrix}, \ldots,$$

$$Q_m = \begin{pmatrix} (\varphi_o, \varphi_o) & \ldots & (\varphi_o, \varphi_m) \\ \vdots & & \vdots \\ (\varphi_m, \varphi_o) & \ldots & (\varphi_m, \varphi_m) \end{pmatrix}, \ldots$$

are positive definite and therefore their determinants

$$\delta_m = |Q_m| > 0 .$$

Now we put

$$\psi_o(x) = \varphi_o(x) = 1, \quad \psi_1(x) = \delta_o^{-1} \begin{vmatrix} (\varphi_o, \varphi_o) & (\varphi_o, \varphi_1) \\ \varphi_o(x) & \varphi_1(x) \end{vmatrix}, \ldots,$$

$$\psi_m(x) = \delta_{m-1}^{-1} \begin{vmatrix} (\varphi_o, \varphi_o) & (\varphi_o, \varphi_1) & \ldots & (\varphi_o, \varphi_m) \\ (\varphi_1, \varphi_o) & (\varphi_1, \varphi_1) & \ldots & (\varphi_1, \varphi_m) \\ \vdots & & & \vdots \\ (\varphi_{m-1}, \varphi_o) & (\varphi_{m-1}, \varphi_1) & \ldots & (\varphi_{m-1}, \varphi_m) \\ \varphi_o(x) & \varphi_1(x) & & \varphi_m(x) \end{vmatrix}, \ldots$$

Referring to elementary properties of determinants it is readily seen that

$$(3) \qquad (\psi_m, \varphi_\mu) = 0 \; (m = 1, \ldots, n; \; \mu = 0, 1, \ldots, m-1) .$$

On the other hand we observe that the functions $\psi_\nu(x)$, defined by the determinant formulae can be represented in the form

$$\psi_o(x) = 1$$

$$\psi_1(x) = \alpha_o^{(1)} + \varphi_1(x)$$

$$\psi_2(x) = \alpha_o^{(2)} + \alpha_1^{(2)} \varphi_1(x) + \varphi_2(x)$$

$$\cdots\cdots\cdots\cdots$$

$$\psi_n(x) = \alpha_o^{(n)} + \alpha_1^{(n)} \varphi_1(x) + \alpha_2^{(n)} \varphi_2(x) + \ldots + \alpha_{n-1}^{(n)} \varphi_{n-1}(x) + \varphi_n(x)$$

where the $\alpha_\mu^{(\nu)}$ are numerical coefficients, uniquely defined for a given partition $\mathcal{P}_n$. The orthogonality relations (2) now follow

https://doi.org/10.4153/CMB-1960-009-4 Published online by Cambridge University Press

immediately from (3). The systems (2) and (3) are indeed equivalent.

In the same way one can calculate the norm square

(4) $\sigma_m = (\psi_m, \psi_m) = (\psi_m, \varphi_m) = \delta_m / (\delta_{m-1})$ $(m = 1, \ldots, n)$.

For any given partition of the interval $[a, b]$ one can thus set up a system of $n + 1$ orthogonal functions $\psi_o(x), \ldots, \psi_n(x)$, sectionally linear over $\mathfrak{P}_n$, and every sectionally linear function $f(x)$ over $\mathfrak{P}_n$ can be represented by its "Fourier expansion"

$$f(x) = a_o + a_1 \psi_1(x) + \ldots + a_n \psi_n(x)$$

whose coefficients are given by

$$a_\mu = (1/\sigma_\mu)(f, \psi_\mu) = (1/\sigma_\mu) \sum_{\nu=o}^{n} f(x_\nu) \psi_\mu(x_\nu).$$

They satisfy a "Parseval equation"

$$(f, f) = \sum_{\mu=o}^{n} \sigma_\mu a_\mu^2$$

which may be used here as a check relation. For practical applications it is thus necessary and sufficient to have a complete table of the orthogonal sectionally linear functions for a partition.

3. Minimum property of the orthogonal functions. Let

$$f_m(x) = \xi_o + \xi_1 \varphi_1(x) + \ldots + \xi_{m-1} \varphi_{m-1}(x) + \varphi_m(x)$$

be a sectionally linear function with the variable coefficients $\xi_o, \xi_1, \ldots, \xi_{m-1}$ (if necessary put $\xi_m = 1$). For a fixed index m, i.e. one of the numbers $1, 2, \ldots, n$, the function $f_m(x)$ is to be determined such that its norm square

$$(f_m, f_m) = \sum_{i=o}^{m-1} \sum_{j=o}^{m-1} (\varphi_i, \varphi_j) \xi_i \xi_j + 2 \sum_{i=o}^{m-1} (\varphi_i, \varphi_m) \xi_i + (\varphi_m, \varphi_m)$$

has the smallest possible value.

Using matrix notations with the prime indicating transposition we have

$$(f_m, f_m) = \xi' Q_{m-1} \xi + 2q'\xi + q_m \quad (\xi' = (\xi_o, \xi_1, \ldots, \xi_{m-1}))$$

where q' denotes the row vector $((\varphi_o, \varphi_m), \ldots, (\varphi_{m-1}, \varphi_m))$ and $q_m = (\varphi_m, \varphi_m)$. The matrix Q_{m-1} was seen to be positive definite. Consequently the expression (f_m, f_m) has a minimum. In fact, if we introduce the new variable column η by the substitution

$$\xi = \eta + \beta$$

we obtain

$$(f_m, f_m) = \eta' \, Q_{m-1} \eta + 2(\beta' Q_{m-1} + q')\eta + \beta' Q_{m-1}\beta + 2q'\beta + q_m .$$

Now the column β can be chosen such that the linear terms in η vanish:

$$Q_{m-1}\beta + q = 0.$$

This is a system of linear equations with the coefficient matrix Q_{m-1}; it has a unique solution β such that

$$(f_m, f_m) = \eta' \, Q_{m-1} \eta + \beta' Q_{m-1} \beta + 2q'\beta + q_m$$

and the minimum of (f_m, f_m) is assumed for $\eta = 0$, i.e. for $\xi = \beta$. Thus the minimum function $f_m(x)$ has to satisfy the condition

$$Q_{m-1} \xi + q = 0$$

which by returning to the original notations is seen to be equivalent with

$$\sum_{j=o}^{m-1} (\varphi_\lambda, \varphi_j) \xi_j + (\varphi_\lambda, \varphi_m) = (\varphi_\lambda, f_m) = 0 \quad (\lambda = 0, 1, \ldots, m-1).$$

By comparison with (3) one concludes that $f_m(x) = \psi_m(x)$. Thus the orthogonal functions have the required minimum property.

4. Partition with equidistant points. Instead of $[a, b]$ we take now the unit interval $[0, 1]$. By $\mathcal{P}_n$ we denote its partition in n equal parts, i.e.

$$x_\nu = \nu/n \qquad (\nu = 0, 1, \ldots, n) .$$

The basis functions are then

$$\varphi_\mu(x) = \begin{cases} 0 & \text{if } x \leq (\mu - 1)/n \\ x - (\mu - 1)/n & \text{if } x > (\mu - 1)/n \end{cases}$$

In this case the orthogonal functions $\psi_\nu(x)$ can readily be computed for every fixed n. (It might be pointed out that in our notation we are constantly omitting an index n; thus instead of $\varphi_\mu(x)$, $\psi_\nu(x)$ we really should write $\varphi_\mu^{(n)}(x)$ and $\psi_\nu^{(n)}(x)$ and similarly in other symbols. This omission is however irrelevant as long as n is supposed to be fixed throughout the discussion.)

For all n it is found that the orthogonal sectionally linear functions over $\mathcal{P}_n$ are given by

$$\psi_0(x) = 1\ ,\quad \psi_1(x) = -\tfrac{1}{2} + x$$

$$\psi_2(x) = \frac{n-1}{(n+1)(n+2)} - \frac{(n-1)(n+4)}{(n+1)(n+2)}\,x + \varphi_2(x)$$

$$\psi_3(x) = \frac{(n-2)(n-1)}{n(n+1)}\,x - 2\,\frac{n-2}{n}\,\varphi_2(x) + \varphi_3(x)\ ,$$

and generally for $k = 0, 1, 2, \ldots, n-3$:

$$(5)\quad \psi_{n-k}(x) = \frac{(k+1)(k+2)}{(k+3)(k+4)}\,\varphi_{n-k-2}(x) - 2\,\frac{k+1}{k+3}\,\varphi_{n-k-1}(x) + \varphi_{n-k}(x)\ .$$

In particular we find

$$\psi_{n-2}(x) = (2/5)\varphi_{n-4}(x) - (6/5)\varphi_{n-3}(x) + \varphi_{n-2}(x) \qquad (\text{if } n \geqslant 5)$$

$$\psi_{n-1}(x) = (3/10)\varphi_{n-3}(x) - \varphi_{n-2}(x) + \varphi_{n-1}(x) \qquad (\text{if } n \geqslant 4)$$

$$\psi_n(x) = (1/6)\,\varphi_{n-2}(x) - (2/3)\varphi_{n-1}(x) + \varphi_n(x)\ (n \geqslant 3).$$

We notice that the coefficients of these expansions do not depend on n.

The first two or three functions $\psi_\nu(x)$ are readily obtained from the general formula in §2. For the rest it has been observed in cases of small values of n that each $\psi_{n-k}(x)$ can be written as linear combination of only three consecutive basis functions: $\varphi_{n-k-2}(x)$, $\varphi_{n-k-1}(x)$, $\varphi_{n-k}(x)$ with coefficients depending on k only. Consequently it was conjectured - and verified - that this is so for every n.

Thus let us write

$$\psi_{n-k}(x) = \alpha\,\varphi_{n-k-2}(x) + \beta\,\varphi_{n-k-1}(x) + \varphi_{n-k}(x)\ .$$

We express according to (3) that $\psi_{n-k}(x)$ is orthogonal to $\varphi_{n-k-2}(x)$ and $\varphi_{n-k-1}(x)$. This gives us two linear equations for the unknown coefficients α and β :

$$(\varphi_{n-k-2}, \psi_{n-k}) = (\varphi_{n-k-2}, \varphi_{n-k-2})\alpha + (\varphi_{n-k-2}, \varphi_{n-k-1})\beta$$
$$+ (\varphi_{n-k-2}, \varphi_{n-k}) = 0$$

$$(\varphi_{n-k-1}, \psi_{n-k}) = (\varphi_{n-k-1}, \varphi_{n-k-2})\alpha + (\varphi_{n-k-1}, \varphi_{n-k-1})\beta$$
$$+ (\varphi_{n-k-1}, \varphi_{n-k}) = 0$$

We evaluate the inner products and obtain the equations in the form

$$(k+3)(k+4)(2k+7)\alpha + (k+2)(k+3)(2k+8)\beta = -(k+1)(k+2)(2k+9)$$
$$(k+2)(k+3)(2k+8)\alpha + (k+2)(k+3)(2k+5)\beta = -(k+1)(k+2)(2k+6)$$

whence

$$\alpha = \frac{(k+1)(k+2)}{(k+3)(k+4)}, \qquad \beta = -2\frac{k+1}{k+3}.$$

To prove the formula (5) we have to show that with these coefficients all the orthogonality relations

$$(\varphi_\nu, \psi_{n-k}) = 0 \qquad (\nu = 0, 1, \ldots, n-k-1)$$

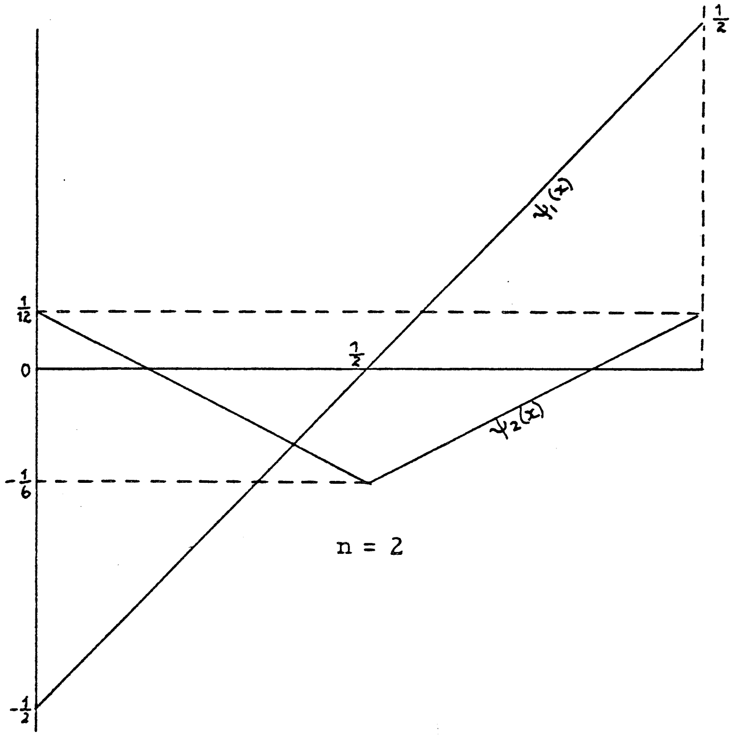
are satisfied. Indeed

$$(\varphi_\nu, \psi_{n-k}) = \frac{(k+1)(k+2)}{(k+3)(k+4)}(\varphi_\nu, \varphi_{n-k-2}) - 2\frac{k+1}{k+3}(\varphi_\nu, \varphi_{n-k-1}) + (\varphi_\nu, \varphi_{n-k})$$

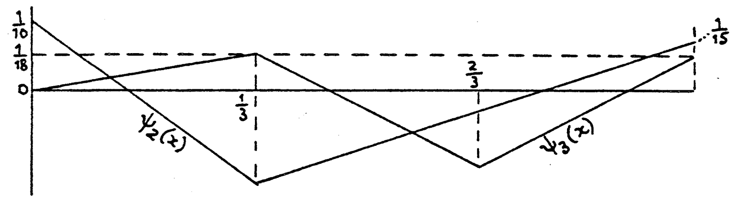
and taking $n^2(k+3)(k+4)$ as common denominator on the right side we obtain as numerator

$$(k+1)(k+2)(n-k-\nu-1) + 2(n-k-\nu) + 3(n-k-\nu+1) + \ldots + (k+3)(n-\nu+1))$$
$$- 2(k+1)(k+4)(n-k-\nu + 2(n-k-\nu+1) + \ldots + (k+2)(n-\nu+1))$$
$$+ (k+3)(k+4)(n-k-\nu+1) + 2(n-k-\nu+2) + \ldots + (k+1)(n-\nu+1))$$
$$= (k+1)(k+2)(\tfrac{1}{2}(k+3)(k+4)(n-k-\nu) + (1/6)(k+1)(k+2)(2k+9) - 1)$$
$$- 2(k+1)(k+4)(\tfrac{1}{2}(k+2)(k+3)(n-k-\nu) + (1/6)(k+1)(k+2)(2k+6))$$
$$+ (k+3)(k+4)(\tfrac{1}{2}(k+1)(k+2)(n-k-\nu) + (1/6)(k+1)(k+2)(2k+3))$$

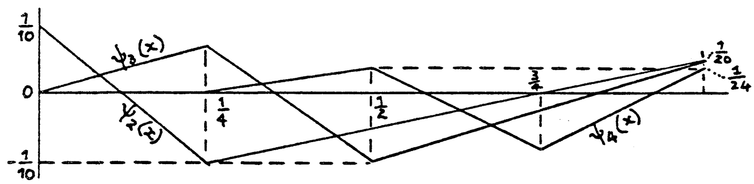
which indeed equals zero. Thus (5) is proved.



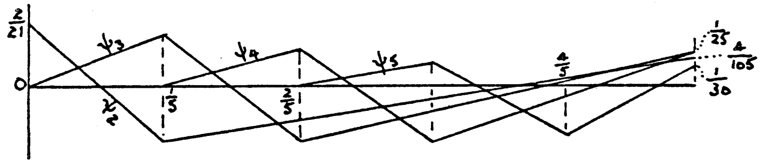
n = 2



n = 3



n = 4



n = 5

https://doi.org/10.4153/CMB-1960-009-4 Published online by Cambridge University Press

For small values of n complete systems of sectionally linear orthogonal functions over $\mathcal{P}_n$ are shown in the following table, their graphs in the figure.

$n = 2: \quad \psi_2(x) = (1/12) - \frac{1}{2}\, x + \varphi_2(x)$

$n = 3: \quad \psi_2(x) = (1/10) - (7/10)\, x + \varphi_2(x)$

$\psi_3(x) = (1/6)\, x - (2/3)\, \varphi_2(x) + \varphi_3(x)$

$n = 4: \quad \psi_2(x) = (1/10) - (4/5)\, x + \varphi_2(x)$

$\psi_3(x) = (3/10)\, x - \varphi_2(x) + \varphi_3(x)$

$\psi_4(x) = (1/6)\varphi_2(x) - (2/3)\varphi_3(x) + \varphi_4(x)$

$n = 5: \quad \psi_2(x) = (2/21) - (6/7)\, x + \varphi_2(x)$

$\psi_3(x) = (2/5)\, x = (6/5)\, \varphi_2(x) + \varphi_3(x)$

$\psi_4(x) = (3/10)\varphi_2(x) - \varphi_3(x) + \varphi_4(x)$

$\psi_5(x) = (1/6)\, \varphi_3(x) - (2/3)\varphi_4(x) + \varphi_5(x)\,.$

For large-scale numerical applications it would of course be necessary to compute orthogonal systems for greater values of n.

For the practical computation of the "Fourier coefficients" of a given function we also need the norm squares $\sigma_m = (\psi_m, \psi_m)$. With regard to (4) and the preceding formulae we find for any n :

$\sigma_0 = n + 1\,, \qquad \sigma_1 = (n+1)(n+2)/12n\,,$

$\sigma_2 = n-1/n(n+1)(n+2)\,, \quad \sigma_3 = (n-1)(n-2)/n^3(n+1)\,;$

and for any $k \leqslant n - 3$:

$$\sigma_{n-k} = \frac{\alpha}{n^2} = \frac{1}{n^2}\,\frac{(k+1)(k+2)}{(k+3)(k+4)}\,.$$

In particular for n = 5 one has

$\sigma_0 = 6,\ \sigma_1 = 7/10,\ \sigma_2 = 2/105,\ \sigma_3 = 2/125,\ \sigma_4 = 3/250,\ \sigma_5 = 1/150\,.$

https://doi.org/10.4153/CMB-1960-009-4 Published online by Cambridge University Press

As an application we shall obtain the sectionally linear interpolation of the function x^2 over the partition $\mathcal{P}_5$ of $[0, 1]$. Evidently

$$f(0) = 0, \; f(1/5) = 1/25, \; f(2/5) = 4/25, \; f(3/5) = 9/25, \; f(4/5) = 16/25,$$

$$f(1) = 1$$

and

$$(f, \varphi_0) = 11/5, \; (f, \varphi_1) = 9/5, \; (f, \varphi_2) = 34/25, \; (f, \varphi_3) = 116/125,$$

$$(f, \varphi_4) = 66/125, \; (f, \varphi_5) = 1/5 .$$

Therefore

$$(f, \psi_0) = 11/5, \; (f, \psi_1) = 7/10, \; (f, \psi_2) = 2/75, \; (f, \psi_3) = 2/125,$$

$$(f, \psi_4) = 1/125, \; (f, \psi_5) = 1/375$$

and

$$a_0 = 11/30, \; a_1 = 1, \; a_2 = 7/5, \; a_3 = 1, \; a_4 = 2/3, \; a_5 = 2/5 .$$

From the "Fourier expansion"

$$f(x) = \sum_{\nu=0}^{5} a_\nu \psi_\nu(x)$$

one derives the interpolating function in the form

$$f(x) = (1/5)\varphi_1(x) + (2/5)\varphi_2(x) + (2/5)\varphi_3(x) + (2/5)\varphi_4(x) + (2/5)\varphi_5(x),$$

which indeed satisfies the given conditions.

5. Curve fitting by least squares. The problem of curve fitting (or data fitting) is really a generalized interpolation problem. Like the ordinary interpolation problem it is usually solved by polynomials. Given a function F(x) tabulated over the partition

$$\bar{\mathcal{P}}_m : \quad a = \bar{x}_0 < \bar{x}_1 < \ldots < \bar{x}_{m-1} < \bar{x}_m = b$$

of the interval $[a, b]$, it is required to find a polynomial p(x) of degree not greater than $n < m$ such that the values of p(x) fit best to the values of F(x) at $\bar{\mathcal{P}}_m$ "in the sense of the method of least squares", i.e. such that

$$\sum_{\mu=0}^{m} (F(\bar{x}_\mu) - p(\bar{x}_\mu))^2$$

becomes as small as possible. If n = m the minimum of this

https://doi.org/10.4153/CMB-1960-009-4 Published online by Cambridge University Press

expression will be zero; it will be obtained if one uses for p(x) the ordinary interpolation polynomial of degree $\leq n$ as it is given by Lagrange's formula or Newton's formula without the remainder term; if $n < m$ the minimum will in general be positive.

A treatment of the data fitting problem is given for instance by Nielsen [3], Chap. VIII; a more elegant method has been proposed by Forsythe [1] and Herzberger [2]. We shall adapt this method for data fitting by means of sectionally linear functions instead of polynomials.

Again we assume that the number m + 1 of the data

$$y_o = F(\bar{x}_o) , \ldots , y_m = F(\bar{x}_m)$$

is larger than n + 1 where n is the "degree" of the sectionally linear function f(x) over $\mathcal{P}_n$, by which the data are to be fitted best in the sense of the method of least squares. We also suppose that the partition $\bar{\mathcal{P}}_m$ contains the partition $\mathcal{P}_n$, i.e. each x_ν is an $\bar{x}_\mu$. (The reason for the inequality $n < m$, preventing strict interpolation, could be that orthogonal systems are not available for sufficiently high values of n.)

Let

$$\begin{aligned} f(x) &= c_o + c_1 \ \varphi_1(x) + \ldots + c_n \ \varphi_n(x) \\ &= a_o + a_1 \ \psi_1(x) + \ldots + a_n \ \psi_n(x) \quad . \end{aligned}$$

The coefficients c_ν or a_ν have to be found such that

$$\Delta = \sum_{\mu=o}^{m} (f(x_\mu) - y_\mu)^2 = \text{Min.}$$

It will be sufficient to establish the procedure for the c_ν as it will be formally the same for the a_ν .

We introduce the vectors (columns)

$$c = \begin{pmatrix} c_o \\ c_1 \\ \vdots \\ c_n \end{pmatrix}, \quad y = \begin{pmatrix} y_o \\ y_1 \\ \vdots \\ y_m \end{pmatrix}, \quad \delta = \begin{pmatrix} f(\bar{x}_o) - y_o \\ f(\bar{x}_1) - y_1 \\ \vdots \\ f(\bar{x}_m) - y_m \end{pmatrix}$$

and the rectangular $(m + 1) \times (n + 1)$ - matrix

$$\Phi = \begin{pmatrix} 1 & \varphi_1(\bar{x}_0) & \varphi_2(\bar{x}_0) & \dots & \varphi_n(\bar{x}_0) \\ 1 & \varphi_1(\bar{x}_1) & \varphi_2(\bar{x}_1) & \dots & \varphi_n(\bar{x}_1) \\ 1 & \varphi_1(\bar{x}_2) & \varphi_2(\bar{x}_2) & \dots & \varphi_n(\bar{x}_2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \varphi_1(\bar{x}_m) & \varphi_2(\bar{x}_m) & \dots & \varphi_n(\bar{x}_m) \end{pmatrix}.$$

Then

$$\delta = \Phi c - y$$

and

$$\Delta = \delta' \delta = c' \Phi' \Phi\, c - 2c' \Phi' y + y' y$$

is the expression to be minimized. Putting

$$G = \Phi' \Phi \quad , \qquad z = \Phi' y$$

one has

$$\Delta = c'Gc - c'z - z'c + y'y \ .$$

Now we observe that the matrix Φ has the rank $n + 1$. It contains as submatrix the regular $(n+1) \times (n+1)$ - matrix

$$(6) \qquad \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & \varphi_1(x_1) & 0 & \dots & 0 \\ 1 & \varphi_1(x_2) & \varphi_2(x_2) & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \varphi_1(x_n) & \varphi_2(x_n) & \dots & \varphi_n(x_n) \end{pmatrix}$$

because $\mathcal{P}_n \subset \mathcal{P}_m$. Thus the Gram matrix G of Φ is positive definite, and G^{-1} exists, and we can write

$$\begin{aligned} \Delta &= c'Gc - c'GG^{-1} z - z'G^{-1} Gc + z'G^{-1} z + y'y - z'G^{-1} z \\ &= (c - G^{-1} z)'G(c - G^{-1} z) + y'y - z'G^{-1} z \ . \end{aligned}$$

It follows that Δ has a minimum, viz. $y'y - z'G^{-1} z$, which actually will be assumed if

$$c - G^{-1} z = 0 \ ;$$

hence c has to be the unique solution of the "normal system"

$$Gc = \Phi' y \ ,$$

or in explicit form

$$(m+1)c_0 + \sum_{\mu=0}^{m} \varphi_1(\bar{x}_\mu)c_1 + \dots + \sum_{\mu=0}^{m} \varphi_n(\bar{x}_\mu)c_n = \sum_{\mu=0}^{m} y_\mu$$

$$\sum_{\mu=0}^{m} \varphi_1(\bar{x}_\mu)c_0 + \sum_{\mu=0}^{m} \varphi_1(\bar{x}_\mu)^2 c_1 + \dots + \sum_{\mu=0}^{m} \varphi_1(\bar{x}_\mu)\varphi_n(\bar{x}_\mu)c_n = \sum_{\mu=0}^{m} \varphi_1(\bar{x}_\mu)y_\mu$$

$$\vdots$$

$$\sum_{\mu=0}^{m} \varphi_n(\bar{x}_\mu)c_0 + \sum_{\mu=1}^{m} \varphi_n(\bar{x}_\mu)\varphi_1(\bar{x}_\mu)c_1 + \dots + \sum_{\mu=0}^{m} \varphi_n(\bar{x}_\mu)^2 c_n = \sum_{\mu=0}^{m} \varphi_n(\bar{x}_\mu)y_\mu \ .$$

If m = n the matrix Φ coincides with the matrix (6) and we have $\Delta = 0$; the data fitting problem then is the interpolation problem.

6. Final remark. We have studied the theory and some applications of the sectionally linear functions over a finite partition of a finite interval only. The theory can be extended to functions over an infinite partition of a finite or infinite interval. This will be the object of a forthcoming paper.

Acknowledgment. This paper was written while the author was receiving a Research Grant from the National Research Council of Canada.

REFERENCES

1. G.E. Forsythe, Generation and use of orthogonal polynomials for data-fitting with a digital computer, J. Soc. Indust. Appl. Math. 5 (1957) 74-88.

2. M. Herzberger, The normal equations of the method of least squares and their solution, Quarterly Appl. Math. 7 (1949) 217-223.

3. K.L. Nielsen, Methods in Numerical Analysis, (New York, 1956).

4. J.F. Steffensen, Interpolation, (Baltimore, 1927).

5. C. de la Vallée Poussin, Leçons sur l'approximation des fonctions d'une variable réelle, (Paris, 1919).

McGill University