PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 41 (55), 1987, pp. 143–148

AN APPROXIMATION OF TABULATED FUNCTION

Bora Jovanović

Abstract. Using a recurrent approach a method for approximate analytical representation of tabulated function as the sum of linear function, harmonics and pairs of exponential functions is derived. The method is demonstrated by two examples.

1. Introduction. The analytical description of observed data is of great importance in many sciences and is often used.

If the function followed by the data is unknown then polynomial and Fourier approximation is commonly applied in practice. Although, these procedures are quite good for data intervals, both of them manifest certain shortcomings in forecasting the observed phenomenon. In particular, this is true for the method first mentioned because it neglects the periodicities which may exist within data set.

On the other side, the use of Fourier analysis makes the precise separation of periodical components possible only if both the sample considered does not contain some nonperiodical trend and fundamental frequency is well chosen. These difficulties can be avoided if the form of the trend function is previously determined, but then the action becomes rather roundabout and certain statistical considerations are required. Moreover, periodicities exceeding the data interval cannot be detected.

One should be reminded that several authors considered the hidden periodicities in a more direct way using algebraic approach. Let us mention only some of them: Lagrange, Kühnen, Bruns, Dale and Plummer ([4], [5], [3], [1],[2], [7])[1]. Their methods have never been widely used since in the past there were no appropriate computing devices. Afterwards, these methods have been overshadowed by new spectral analysis techniques based on discrete Fourier transform and the theory of random functions.

In this paper a variant of these "old" algebraic methods will be presented in a simpler and more general way. The method gives the resolution of the arbitrary

---

AMS Subject Classification (1980): Primary 65P15, 41A99.

[1]A more completed list of references is given in [8].

tabulated function into a linear function, harmonics and exponential functions proceeding from a recurrence relation.

2. Preliminaries. Let

$$\{F(x_i)\} \qquad i = \overline{1, N} \tag{1}$$

be a sample containing N elements which represent the values of a function $F(x)$ with equidistant argument step $\varepsilon = x_i - x_{i-1}$.

Let us determine the analytical form of the best approximation of $F(x)$ along the interval $[x_1, x_N]$, assuming that it may be expressed in the form[2]

$$F(x) = \sum_{i=1}^{n} f_i(x) \tag{2}$$

where functions $f_i(x)$ are defined by

$$f_i(x) = k_i f_i(x-\varepsilon) - f_i(x-2\varepsilon) \tag{3}$$

Here n denotes their total number and k_i are real positive constants.

Obviously, the condition $n \leq N/3$ must be satisfied as each of relations in (3) is determined by three values: three initial values or two initial values and constant k_i.

Since the solution of the functional equation (3) may have three forms for $k > 0$ (subscript i is omitted)

$$\begin{array}{lll} (1) & f(x) = C_1 \sin \omega x + C_2 \cos \omega x, & k \in [0, 2) \\ (2) & f(x) = C_1 x + C_2, & k = 2 \\ (3) & f(x) = C_1 \lambda_1^x + C_2 \lambda_2^x, & k \in (2, +\infty) \end{array} \tag{4}$$

where C_1 and C_2 are arbitrary constants and

$$\omega = \frac{1}{\varepsilon} \arccos \frac{k}{2}, \quad \lambda_{1,2} = \left(\frac{k \pm \sqrt{k^2 - 4}}{2} \right)^{1/\varepsilon} \tag{5}$$

the task just formulated is reduced to the estimation of parameters k, C_1, C_2 for each function $f(x)$.

3. Determination of $F(x)$. The function $F(x)$ can be derived by simple addition of $f_i(x)$ in (2) for $n = 2, 3, \ldots$. Then it is easy to prove by mathematical induction the validity of the expression

$$\sum_{i=0}^{n} \binom{n}{i} F(x - 2i\varepsilon) = \sum_{i=1}^{n} -P_i \binom{n-i}{j} F(x - (i+2j)\varepsilon) \tag{6}$$

where coefficients P_i are defined by

[2]For the sake of simplicity the function $F(x)$ and its approximation are denoted identically.

$$P_1 = -\sum_{i=1}^{n} k_i, \quad \sum_{\substack{i,j \\ 1\le i<l\le n}} k_i k_j, \;\; P_3 = -\sum_{\substack{i,j \\ 1\le i<l\le n}} k_i k_j k_l, \ldots, P_n = (-1)^n \prod_{i=1}^{n} k_i \tag{7}$$

The formula (6) gives $F(x)$ represented in the form

$$F(x) = \sum_{i=1}^{2n} Q_i F(x - i\varepsilon) \tag{8}$$

which is the recursive equivalent of the analytical function $F(x)$. Hence, Q_i are the coefficients directly depending on P_i.

The values of P_i; may be calculated from the system of linear equations

$$-\sum_{j=1}^{n} A_{i,n-j+1} P_j = A_{i,n+1}, \quad i = \overline{1, N-2n} \tag{9}$$

which includes the sample (1) and where

$$A_{ij} = \sum_{l=1}^{j-1} \binom{j-1}{l} F(x_{n+i+j-1-2l}), \quad \begin{cases} i = \overline{1, N-2n} \\ j = 1\overline{1, n+1} \end{cases}$$

The system (9) has the unique solution only if both $n = N/3$ and its determinant is not equal to zero. In the case $n < N/3$, the best coefficients P_j may be obtained, for example, by least squares method.

Since the relations (7) are Viete rules, constants k_i can be computed by solving the algebraic equation

$$\sum_{i=1}^{n+1} b_i k^{n-i+1} = 0 \tag{10}$$

where $b_1 = 1$ and $b_i = P_{i-1}$, $i = \overline{2, n+1}$.

Each positive k defines one function in (4) with constants C_1 and C_2. If there exist m such functions, then $2m$ unknown constants are obtained by solving the linear system

$$F(x_i) = \sum_{j=1}^{m} f_j(x_i), \quad i = \overline{1, N} \tag{11}$$

which means that initial task is fulfilled.

4. Examples. *Example* 1. Separate the component functions (4) in the sample made by the function

$$F(x) = 2x + 5 - 3\cos 7x - e^x$$

for $x = 0, 0.1, 0.2, \ldots, 3$.

Solution. The matrix of the system (9) may be created by the algorithm presented in Figure 1, where C and D are auxiliary arrays for binomial coefficients computation, and F are the values of function with $F(i) = F(x_i)$. Solving the system by the least squares technique and applying Bairstow method to find the roots of the equation (10), depending on given number n, the following results are obtained:

n	i	$b(i)$	$k=\mathrm{re}+\mathrm{im}\sqrt{-1}$ re	im	parameters
2	1	1.00000000	2.0011872	0.0000000	$\to \lambda_1 = 1.4113,\quad \lambda_2 = 0.7085$
	2	-3.52837370	1.5271865	0.0000000	$\to \omega = 7.0194$
	3	3.05618600			
3	1	1.00000000	1.9998854	0.0000000	$\to \omega_1 = 0.1070$
	2	-5.54056400	1.5296839	0.0000000	$\to \omega_2 = 7.0000$
	3	10.15713800	2.0109948	0.0000000	$\to \lambda_1 = 2.8522,\quad \lambda_2 = 0.3506$
	4	-6.15202000			
4	1	1.00000000	2.0098360	0.0000000	$\to \lambda_1 = 2.6949,\quad \lambda_2 = 0.3711$
	2	-3.11615270	-2.4234163	0.0000000	
	3	-3.27105460	2.0000494	0.0000000	$\to$ linear function
	4	18.45748800	1.5296836	0.0000000	$\to \omega = 7.0000$
	5	-14.90153400			
9	1	1.00000000	2.0100142	0.0000000	$\to \lambda_1 = 2.7191,\quad \lambda_2 = 0.3676$
	2	-2.03534460	1.9999959	0.0000000	$\to$ linear function
	3	3.94588460	-1.3052951	0.6385531	
	4	-4.85274540	-1.3052951	-0.6385531	
	5	12.12078900	-3.0148948	1.1822640	
	6	-50.89238600	-3.0148948	-1.1822640	
	7	-107.03534460	0.2528301	1.7630826	
	8	-107.46158000	0.2528301	-1.7630826	
	9	-0.28770447	1.5296844	0.0000000	$\to \omega = 7.0000$
	7	-431.99773000			

Tab. 1

Obviously, for greater n the better resolution is achieved.

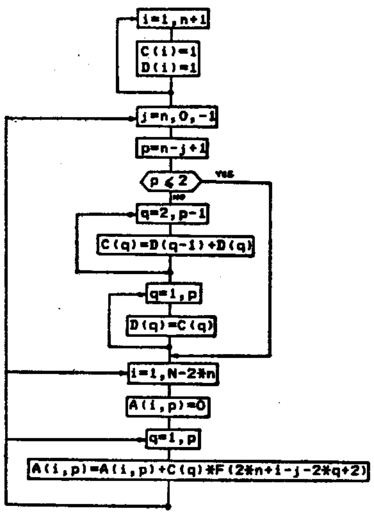

Fig. 1. The algorithm for computing the matrix $A(i,j)$

Example 2. Approximate the function $F(x) = \sqrt{\ln x + \sin x}$ along the interval $[1, 31]$, assuming that it is tabulated with $\varepsilon - 1$.

Solution. In order to determine the structure of the function the procedure shown in the previous example should be repeated. The results or $n = 2, 3, 4$ are given in Tab 2.

n	i	$b(i)$	$k = \text{re} + \text{im}\sqrt{-1}$		parameters
			re	im	
2	1	1.00000000			$\to \omega_1 = 0.0594$
	2	-2.98843180	1.9964689	0.0000000	$\to \omega_2 = 1.0518$
	3	1.98042300	0.9919629	0.0000000	
3	1	1.00000000	1.0493319	0.0000000	$\to \omega_1 = 1.0185$
	2	-2.37858270	-0.6675544	0.0000000	
	3	0.06184917	2.0109948	0.0000000	$\to \omega_2 = 0.0565$
	4	1.39873430			
4	1	1.00000000	1.0557793	0.0000000	$\to \omega_1 = 1.0147$
	2	-1.76715210	-1.2915686	0.0000000	
	3	-1.82524640	1.9976164	0.0000000	$\to \omega_2 = 0.0488$
	4	2.73374210	0.0053250	0.0000000	$\to \omega_3 = 1.5681$
	5	-0.01450509			

Tab. 2

Since all extracted components are harmonics, the approximate function searched for is

$$\overline{F}(x) = \sum_{i=1}^{m} C_{1i} \sin \omega_i x + C_{2i} \cos \omega_i x$$

where $m = 2$ in the first two cases, and $m = 3$ in the last one.

Tab. 3 contains the coefficients C_{1i} and C_{2i} obtained by solving the system (11) using the least squares method. Also, mean square errors of the coefficients $(\Delta C_{1i}, \Delta C_{2i})$ and conditional equations (σ_0) are given. The last one is obtained by the formula

$$\sigma_0 = \left((N - 2m)^{-1} \sum_{i=1}^{N} \delta_i^2 \right)^{1/2}$$

where δ_i are the residuals of conditional equations.

n	i	C_{1i}	C_{2i}	ΔC_{1i}	ΔC_{2i}	σ_0
2	1	1.745	0.606	0.043	0.051	
	2	0.211	-0.191	0.039	0.040	0.156
3	1	0.292	-0.065	0.030	0.030	
	2	1.727	0.644	0.034	0.039	0.119
4	1	0.290	-0.068	0.023	0.023	
	2	1.652	0.762	0.030	0.030	
	3	-0.018	-0.013	0.022	0.023	0.088

Tab. 3

All calculations in the examples presented were carried out with nine significant digits.

5. Conclusions. The method has been tested on various examples. Several basic features have been noted.

– The separation precision of individual components depends considerably on their total number and on the number of significant digits used in calculations. The smaller the number of individual components and the greater the number of significant digits, the better is the result.

– The efficiency of the method is fully achieved if it is applied to the determination of hidden periodicities by

$$T = \frac{2\pi\varepsilon}{\arccos k/2}$$

because they may belong to a broad interval $[4\varepsilon, +\infty)$. In addition to that, the possible existence of linear or exponential trend will not change their values. Of course, if the real periodicity is smaller then the more accurate result will be attained.

– The better approximation is achieved if the larger number n of components is assumed. The choice of n depends on the particular requirements, the dimension of the sample N and the capacity of the computer used.

– If considered data have a great dispersion of random errors then it is advisable to apply a smoothing or a filtering method in order to increase the accuracy and to decrease the number of components separated.

Finally, it should be noted that the idea used here, could also be applied to other resolutions of a tabulated function. For example, if the relation (3) is

$$f_i(x) = k_i t f_i(x - \varepsilon)$$

then the described procedure is reduced to the well known Prony method.

REFERENCES

[1] H. Bruns, *Über die Analyse periodischer Vorgänge*, Astronomische Nachrichten **188** (1911), 55–60.

[2] J.B. Dale, *The resolution of a compound function into simple periodic functions*, Monthly Notices Roy. Astron Soc. **74** (1914), 628–648.

[3] F. Kühnen, *Methode zur Aufsuchung periodischer Erscheinungen in Reihen äquidistanter Beobachtungen*, Astronomische Nachrichten **182** (1909), 1–12.

[4] J. Lagrange, *Sur les interpolations, Oeuvres* VII, Gauthier-Villars, Paris, 1877,

[5] J.L. Lagrange, *Recherches sur la maniere de former des tables des planetes d'apres les seules observations, Oeuvres* VI, Gauthier-Villars, Paris, 1873., pp. 507–627.

[6] G. Milovanović, *Numerička analiza*, Naučna knjiga, Beograd, 1985., pp. 413–416.

[7] H.C. Plummer, *Note on professor Dale's method of periodic analysis*, Monthly Notices Roy. Astron. Soc. **76** (1915), 60–63.

[8] М.Г. Серебренников, А.А. Первозванский, *Выявление скрытых периодичностей*, Наука, Москва, 1965. 137–170.

[9] Б.М. Щиголев, *Матемачеcкая обработка наблюдений*, Наука, Москва, 1969 239–277.

Astronomska opservatorija
Volgina 7
11050 Beograd
Jugoslavija

(Received 28 08 1986)