THE MEANING OF THE PERIODICITY FOR THE APROXIMATION

Ing. Omer SALIH, Doctoral Degree Programme (2)
Dept. of Telecommunications, FEEC, BUT
E-mail: saliho@feec.vutbr.cz

Supervised by: Dr. Vít Novotný, Dr. Josef Zapletal

ABSTRACT

It is known that the approximation of the periodic functions with the aid of the interpolation polynomials outperforms fairy stable extrapolation values. We furnish concrete example of such functions in the introduction. In the next part we build one of the theories for the determination of one ore more hidden eminent period.

1 INTRODUCTION

We often meet with the following problem in technical practice. A function $f(x)$ which is rather complied is given on interval I and it is necessary to recompense it by a more simple function whose values can be simply enumerated and which is "near enough" to $f(x)$ on interval I. We can suppose for such a function for example polynomial

$$P_n(x) = a_0 + a_1 x + \ldots + a_n x^n \tag{1}$$

or trigonometrically polynomial

$$T_n(x) = a_0/2 + (a_1 \cos x + b_1 \sin x) + \ldots + (a_n \cos nx + b_n \sin nx). \tag{2}$$

We suppose that the function $f(x)$ is defined on the interval $[a, b]$ and $G_n(x)$ is either algebraically or trigonometrically polynomial. We suppose choice of $n + 1$ different points in the interval $[a, b]$:

$$x_0, x_1, x_2, \ldots, x_n. \tag{3}$$

It is to find a function $G_n(x)$ satisfying the following properties: $f(x_i) = G_n(x_i)$ for $i = 0, 1, 2, \ldots, n$. Function $G_n(x)$ is called interpolating function and the points x_i bundles of interpolation. We put $x_0 = a$ and $x_0 = b$. The process of replacement of the function $f(x)$ by function $G_n(x)$ outside of the interval $[a, b]$ is called extrapolation. The most applied interpolating polynomials are: Lagrange, Newton, Gauss, Stirling, Bessel and other polynomials. The function $f(x)$ has a significant influence on the condition

$$|f(x) - G_n(x)| < \varepsilon \tag{4}$$

for x lying outside of interval $[a, b]$ from both the sides. We select for the example of interpolation two continuous functions satisfying the Weierstrass conditions for interpolation namely $f_1(x)$ = cos x and $f_2(x) = |x-6|$. We chose a 16 points system of bundles (3) of the interval [4, 7]

4, 4.2, 4.4, 4.6, 4.8, 5, 5.2, 5.4, 5.6, 5.8. 6, 6.2, 6.4, 6.6, 6.8, 7

We calculate Newton interpolation formula $G_{15}(x)$ and create the table of absolute values of differences of functions

$$G_{15}(x) - \cos x \qquad \text{and} \qquad G_{15}(x) - |x-6|$$

at some points lying on the left side of the point 4 and it's strait right neighbourhood.

x	$G_{15}(x)$ - cos x	$G_{15}(x)$ - abs (x - 6)
0.47	0.098	$1.17 . 10^{10}$
1.11	0.027	$3.24 . 10^{9}$
2.10	0.00082	$9.81 .10^{7}$
3.03	$4.68 . 10^{-6}$	$5.61 .10^{5}$
3.97	$3.64 . 10^{-11}$	4.99
4.14	$5.55 . 10^{-12}$	$2.48 . 10^{-8}$

Likewise precise extrapolations as for given function cosine can be obtained also for other periodical functions or functions constructed as the quotient of polynomials optionally powers, extractions or logarithms of these quotients if they exist in the domain of real numbers. Previous results imply the following conclusion. When we know about a (discrete) set of values of a function $f(x)$ which is in the investigated interval periodical or arose as an elementary function of the quotient of two polynomials then it is possible to suppose their extrapolations for arbitrary small number ε as very good in the sense of (4). As introduced, one class of functions, which can be interpolated well, is the class of periodical functions. This fact incites following question. Does there exist any possibility to verify by a mathematical method the existence of periodicity or at least of a hidden periodicity in the sequence of given values? Calculate Fourier coefficients for the approximation by polynomial (2). Suppose that the studied functions are time functions independent on t, their values at the points $t_1, t_2, \ldots, t_n$ generate time series. It is well known that time series can be described by a system of sine and cosine waves with various amplitudes and frequencies in full generality in the form of trigonometrical polynomial.

$$f(t) = A_0 + \sum_{j=1}^{H} a_j \sin \omega_j t + \sum_{j=1}^{H} b_j \sin \omega_j t \qquad (5)$$

where $H = \frac{n}{2}$ for even n and $H = \frac{n-1}{2}$ for odd n, $\omega_j = \frac{2\pi j}{n}$ for $j = 1, \ldots, H$ is the j^{th}-frequency and A_0 , a_j , b_j are coefficients, which will be estimated later. We introduce the notion of Fourier's period $\tau_j = \frac{2\pi}{\omega_j} = \frac{n}{j}$, which has the dimension of the time t which is different from the physical definition of frequency, where the quotient $\frac{\omega}{2\pi}$ is supposed as the number of cycles per time period. Formally we suppose that the first m (where $m < H$) of these Fourier period τ_j are significant. The other $H - m$ are supposed as insignificant and we do not work with them hereafter. The equation (5) transforms into the form

$$f(t) = A_0 + \sum_{j=1}^{m} a_j \sin \omega_j t + \sum_{j=1}^{m} b_j \sin \omega_j t \qquad (6)$$

where the estimation of A_0 is $\overline{A_0} = \frac{1}{n}\sum_{i=1}^{n} f(t_i)$ and the estimations of a_j are

$\overline{a_j} = \frac{2}{n}\sum_{i=1}^{n} f(t_i)\sin \omega_j t_i$ and similarly b_j are estimated by

$\overline{b_j} = \frac{2}{n}\sum_{i=1}^{n} f(t_i)\cos \omega_j t_i$, where $j = 1, 2, \ldots m$.

For the establishment of existence of significant periodical component in time period we are in need of determination of the variance of estimated values $\overline{f(t_i)}$. The following formula can be proved:

$$\text{var } \overline{f(t_i)} = \frac{1}{2}\sum_{j=1}^{m}\left(\overline{a_j}^2 + \overline{b_j}^2\right) , \qquad (7)$$

An important role in determining significant hidden periods has so called index of determination I^2 which is defined by the formula

$$I^2 = \frac{\text{var}\,\overline{f(t_i)}}{\text{var}\,f(t_i)} \qquad (8)$$

For identification of periodicity besides visual methods, objective methods of the analysis of time periods are often used. Especially auto-correlation functions, coherent functions but also the analysis of periodogram. We will deal with the last method. We understand by periodogram the set of all the values of theoretical variances given by (7) which contains variances no only for significant frequencies but for all frequencies. It is

$$\omega_j = \frac{2\pi j}{n} \text{ , kde } j = 1, 2, \ldots, H .$$

We put a concrete problem. At first we construct for them the periodogram and using the statistical Fisher test we determine statistical significant periods. We study the following data:

t_i	1	2	3	4	5	6	7	8	9	10	11	12
$f(t_i)$	228	200	188	158	97	104	83	93	110	117	168	181
t_i	13	14	15	16	17	18	19	20	21	22	23	24
$f(t_i)$	199	229	248	222	225	166	161	99	80	61	90	96

2 WE CALCULATE PERIODOGRAM

Order	Period in time un.	Periodogram $I(\omega_j)$	Order	Period in time un.	Periodogram $I(\omega_j)$
2	12	2008.057 (v_1)	11	2.18	38.315 (v_7)
1	24	652.982 (v_2)	5	4.80	29.056 (v_8)

3	0	294.968 (v_3)	10	2.40	21.858 (v_9)
12	2	79.170 (v_4)	7	3.43	18.566 (v_{10})
9	2.67	71.793 (v_5)	8	3	8.173 (v_{11})
4	6	68.150 (v_6)	6	4	2.740 ($v_{12}=v_H$)

We do Fisher test now. The values of periodogram of the time series $f(t_i)$ intended for the frequencies ω_j are chosen as the testing statistic. The null hypothesis H_o is tested:
The values of periodogram form the sequence of independent random variables having a normal distribution with the mean value equal to zero ii

$$H_o : \mathrm{var} f(t_i) = 0 \text{ against the alternative } H_1 : \mathrm{var} f(t_i) \neq 0 .$$

In other words: We test the hypothesis H_o that the investigated time period does not contain a period against the alternative that it is contrary. We use the Fisher test of signification. Its order is following: We take the first value (v_1) from the ordered table of periodogram and calculate V_j for $j = 1, 2, ... , H,$ according to the following formula

$$V_j = \frac{v_j}{\sum_{i=j}^{H} v_i} \tag{9}$$

If all values v_i, $i = 1, 2, . . . , H$ will be approximately equal all values V_i will be near to the number $\frac{1}{H}$. If vice versa the values of v_i will acquire very big values for some indexes after that the values of V_i will be near to the unity. The big values of V_i will create the critical domain of the hypothesis H_0 and the length of period will be $\frac{n}{j}$. The self test has then a simple form:

$$W = \max_j V_j , \quad j = 1, 2, . . . , H \tag{10}$$

Hypothesis disapproves when $W > g_f(\alpha)$, where $g_f(\alpha)$ is a critical value of the Fisher test at the level of importance α. As soon as the first statistical significant periodical component of certain frequency ω_j is determined, the significance of the next component determined by the second biggest value of the periodogram is tested and so one till the hypothesis H_o is accept.

3 WE CALCULATE FOR GIVEN DATA WITH RESPECT TO (9)

$$V_1 = \frac{2008.057}{3293.828} = 0.6096$$

Using linear interpolation we determine the value of $g_f(0.01)$ for H = 12 from the table **V** in [6] which is equal to 0.48443.
For W = 0.6096 > 0.488443 the hypothesis H_o is disapproved and the contrary hypothesis for the existence of significant period of the length 12 in given series is accepted. Analysis of the time series continues by the test for the second biggest value of peridogram. We obtain in this case

$$V_2 = \frac{652.982}{1285.771} = 0.507852$$

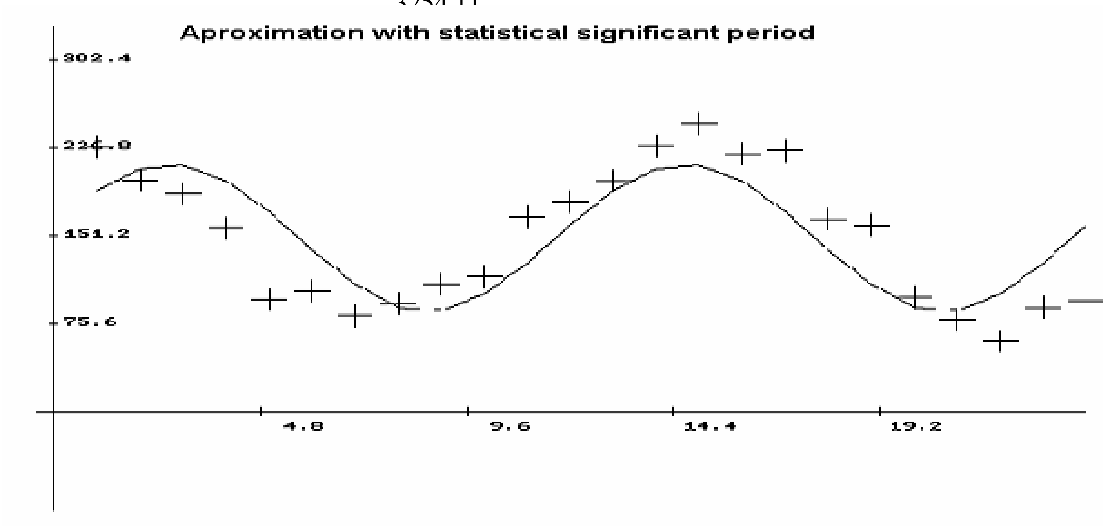
Using linear interpolation we determine the value of $g_f(0.01)$ for H = 11 from the table **V** in [6] which is equal to 0.51022. Because W = 0.507852 < 0.51022, the next period of the length 6 is not significant.

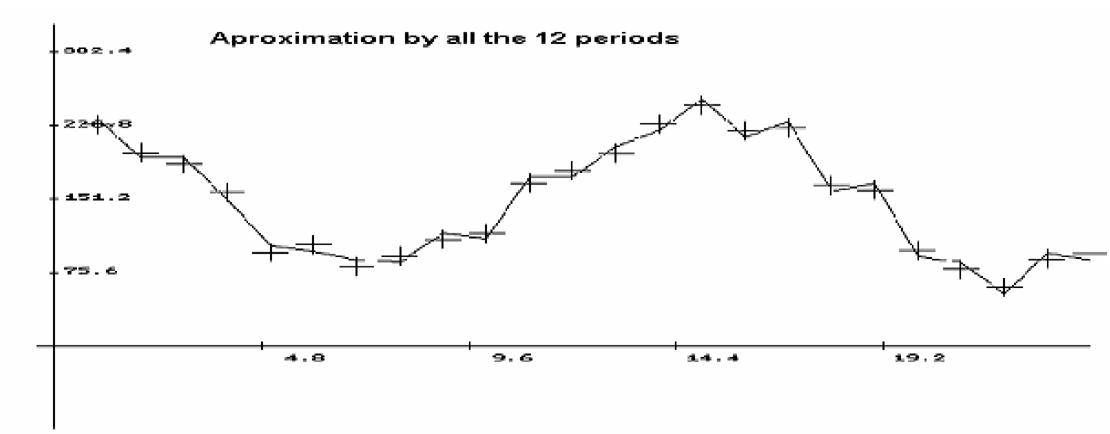
4 WE CALCULATED M = 1 AND WE HAVE

$$\text{var } \overline{f(t_i)} = \frac{1}{2}\left(67.474^2 + 10.633^2\right) = 4016.178$$

and for index of determination I^2 the following formula holds:

$$I^2 = \frac{2008.06}{3254.11} = 0.61709 \qquad \text{tj. } 61.71\% \ .$$





REFERENCES

[1] Anděl, J. : Matematická statistika. Praha, Matfyzpress 1993.

[2] Cramér, H.: Mathematical Methods of Statistics. Princeton, Princeton University Press 1956.

[3] David, H. : Order Statistics. New York, J. Wiley 1966.