Sur l'interpolation des valeurs fournies par les observations*

Pafnuty Chebyshev

Bulletin physico-mathématique de de l'Académie Impériale de St.-Pétersbourg, T. XVI , p. 353–358.
(Read 18 (30) March 1858)

If the number of interpolated values surpasses the one of the terms that one conserves in their expression, interpolation is able to be executed by diverse methods. But these methods, in each particular case, are far from being equally good; they differ among themselves either by the prolixity more or less great of the calculations, or by the magnitude of the mean error to fear, as long as there is concern of interpolation of the values furnished by the observations, and consequently affected of errors. As one is not able to gain beyond a certain limit, under one of these relations without loss under the other, it is impossible to give a method of interpolation which is in general preferable to all the others; for, according to the case, one holds more either to the simplification of the calculations, or to the precision of the results. It is thus that the choice of the method of interpolation depends on the number of the values to interpolate.If this number is small enough, the data of interpolation offer only quite little resources in order to attenuate the influence of their errors on that of the result sought, and then it is important to draw from it all the part possible in order to diminish the mean error to fear, that which one is able to do only by aid of *the method of least squares*. In the contrary case, the considerable number of data that one has to its disposition, dispenses us of recourse to *the method of least squares* which requires some calculations too long. In this case, for the simplification of the numerical operations, one is able also to sacrifice a part more or less considerable of that which the given values offer in order to estimate the sought result. In the Memoir *sur les fractions continues*, presented to the Academy in 1854, we have treated of the interpolation according to *the method of least squares*, and we are arrived to a series which gives directly the results of one such interpolation, indispensable, as we just saw, if the number of their values to interpolate is rather small. In the present Memoir we will show how, according to our methods, one arrives to other formulas of interpolation which are able to replace with advantage that of which we just spoke, as much as its application, because with the great number of interpolated values, on the one hand, ceases to be important, and on the other, becomes not very practical.

*Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. November 3, 2011

We do not treat the different particular cases which are able to present interpolation according to the number, more or less great, of the interpolated values; we limit ourselves to consider the one which is the limit of all the others, where the number of interpolated values is infinite. Although, in reality, this number is never infinite, the formulas that one finds under this assumption are able to be however of a useful application; for they present the limit towards which the results of interpolation converge very rapidly, in measure as this number increases, and it will not be difficult to see, in each particular case, by what degree of approximation these formulas are susceptible according to the given values.

Thus, among other formulas, let us arrive to this here:

$$
\begin{aligned}
f(X) \quad = \quad & \frac{1}{2a}\int_{-a}^{a} f(x)\,dx + \left[\int_{0}^{a} f(x)\,dx - \int_{-a}^{0} f(x)\,dx\right]\frac{X}{x^2} \\
+ \quad & \left[\int_{\frac{a}{2}}^{a} f(x)\,dx - \int_{-\frac{a}{2}}^{\frac{a}{2}} f(x)\,dx + \int_{-a}^{-\frac{a}{2}} f(x)\,dx\right]\frac{2X^2 - \frac{2}{3}a^2}{a^3} \\
+ \quad & \left[\int_{\frac{a}{\sqrt{2}}}^{a} f(x)\,dx - \int_{0}^{\frac{a}{\sqrt{2}}} f(x)\,dx + \int_{-\frac{a}{\sqrt{2}}}^{0} f(x)\,dx - \int_{-a}^{-\frac{a}{\sqrt{2}}} f(x)\,dx\right]\frac{4X^3 - 2a^2X}{a^4} \\
+ \quad & \left[\int_{\frac{\sqrt{5}+1}{4}a}^{a} f(x)\,dx - \int_{\frac{\sqrt{5}-1}{4}a}^{\frac{\sqrt{5}+1}{4}a} f(x)\,dx + \int_{-\frac{\sqrt{5}-1}{4}a}^{\frac{\sqrt{5}-1}{4}a} f(x)\,dx - \right. \\
& \qquad \left. \int_{-\frac{\sqrt{5}+1}{4}a}^{-\frac{\sqrt{5}-1}{4}a} f(x)\,dx + \int_{-a}^{-\frac{\sqrt{5}+1}{4}a} f(x)\,dx\right]\frac{8X^4 - 6a^2X^2 + \frac{2}{5}a^4}{a^5} \\
+ \quad & \left[\int_{\frac{\sqrt{3}}{2}a}^{a} f(x)\,dx - \int_{\frac{a}{2}}^{\frac{\sqrt{3}}{2}a} f(x)\,dx + \int_{0}^{\frac{a}{2}} f(x)\,dx - \int_{-\frac{a}{2}}^{0} f(x)\,dx \right. \\
& \qquad \left. \int_{-\frac{\sqrt{3}}{2}a}^{-\frac{a}{2}} f(x)\,dx - \int_{-a}^{-\frac{\sqrt{3}}{2}a} f(x)\,dx\right]\frac{16X^5 - 16a^2X^3 + \frac{8}{3}a^4X}{a^6} \\
+ \quad & \text{etc.}
\end{aligned}
$$

Although this formula contains some integrals, in order to evaluate its terms with a sufficient approximation and beyond from that which the errors of the data they themselves comprise, one has need ordinarily only of a very limited number of values of $f(x)$ between $x = -a$ and $x = +a$. but as long as one has a sufficient number of values of $f(x)$, this formula is able to be advantageously employed for interpolation; for here, on the one hand, the numerical operations, in regard to the complication of the problem, are rather short, and on the other, the influence of the errors of the interpolated values over those of the sought result is notably attenuated.

In order to be assured of it we note that all the difficulty of interpolation, according to this formula, is reduced to the evaluation of the integrals

$$\int_{-a}^{a} f(x)\,dx, \quad \int_{0}^{a} f(x)\,dx, \quad \int_{-a}^{0} f(x)\,dx, \quad \text{etc.}$$

according to the known values of $f(x)$. Now, though the number of different arithmetic operations that this requires increases to infinity with the one of the interpolated values of $f(x)$, these two numbers are only of the same order of magnitude, while in the *method of least squares* the first is of a superior order relative to the second. On the other hand, the composition of this formula shows that the mean error of the result, coming from those of the interpolated values, is in general of the same order of magnitude as unity divided by the square root of their number, as this holds in the *method of least squares*.

As for the determination of the integrals

$$\int_{-a}^{a} f(x)\,dx, \quad \int_{0}^{a} f(x)\,dx, \quad \int_{-a}^{0} f(x)\,dx, \quad \text{etc.}$$

which enter into our formula, they are able to be evaluated according to the known values of $f(x)$, with an approximation more or less great. But if these values are brought rather closer, one will be able often, in their approximate evaluation, to be content with this very simple formula:

$$\begin{aligned}\int_{h}^{H} f(x)\,dx &= \frac{1}{2}\left[(x_\lambda + x_{\lambda+1} - 2h)f(x_\lambda) + (x_{\lambda+2} - x_\lambda)f(x_{\lambda+1}) + (x_{\lambda+3} - x_{\lambda+1})f(x_{\lambda+2})\right.\\ &\quad \left. + \cdots + (x_\mu - x_{\mu-2})f(_{\mu-1}) + (2H - x_\mu - x_{\mu-1})f(x_\mu)\right].\end{aligned}$$

where

$$f(x_1), f(x_2), \ldots f(x_\lambda), f(x_{\lambda+1}), \ldots f(x_{\mu-2}), f(x_{\mu-1}), f(x_\mu)$$

being the known values of $f(x)$, and

$$x_\lambda, \quad x_{\lambda+1}, \ldots \quad x_{\mu-1}, x_\mu$$

those of x, comprehended between $x = h$ and $x = H$.—The error of this expression of the integral $\int_h^H f(x)\,dx$, as it is easy to understand, will be always inferior to

$$\left(A + \frac{(H-h)B}{24}\right)\Delta^2,$$

where A, B designate the greatest values of $f'(x)$, $f''(x)$ between $x = h$ and $x = H$, and Δ the greatest of the differences

$$x_\lambda - h, x_{\lambda+1} - x_\lambda, \ldots x_\mu - x_{\mu-1}, H - x_m u.$$

Moreover, one will be able to find the integrals

$$\int_{-a}^{a} f(x)\,dx, \quad \int_{0}^{a} f(x)\,dx, \quad \int_{-a}^{0} f(x)\,dx, \quad \text{etc.}$$

very nearly without calculation, if one has a graphical representation of the function $f(x)$, constructed according to its known values; for then, in order to evaluate all these integrals, one will have only to determine the areas of the curve

$$y = f(x),$$

between $x = -a$ and $x = +a$, between $x = 0$ and $x = a$, etc., this which will be done very easily by aid of the *planimeter*.

We note further that by making in this formula

$$\int_0^x f(x)\,dx = F(x),$$

one finds

$$\begin{aligned}
F'(X) &= \frac{F(a) - F(-a)}{2a} + \frac{F(a) - 2F(0) + F(-a)}{a^2} X \\
&\quad + \frac{F(a) - 2F\left(\frac{a}{2}\right) + 2F\left(-\frac{a}{2}\right) - F(-a)}{a^3}(2X^2 - \frac{2}{3}a^2) \\
&+ \frac{F(a) - 2F\left(\frac{a}{\sqrt{2}}\right) + 2F(0) - 2F\left(-\frac{a}{\sqrt{2}}\right) + F(-a)}{a^4}(4X^3 - 2a^2 X) \\
&+ \frac{F(a) - 2F\left(\frac{\sqrt{5}+1}{4}a\right) + 2F\left(\frac{\sqrt{5}-1}{4}a\right) - 2F\left(-\frac{\sqrt{5}-1}{4}a\right) + 2F\left(-\frac{\sqrt{5}+1}{4}a\right) - F(-a)}{a^5} \times \\
&\qquad (8X^4 - 6a^2X^2 + \frac{2}{5}a^4) \\
&+ \frac{F(a) - 2F\left(\frac{\sqrt{3}}{2}a\right) + 2F\left(\frac{a}{2}\right) - 2F(0) + 2F\left(-\frac{a}{2}\right) - 2F\left(-\frac{\sqrt{3}}{2}a\right) + F(-a)}{a^6} \times \\
&\qquad (16X^5 - 16a^2X^3 + \frac{8}{3}a^4 X) \\
&+ \text{ etc.,}
\end{aligned}$$

a formula which is able to be advantageously employed for the determination of the first derivative $F'(x)$, according to the given values of $F(x)$, if however these values are rather close among themselves, so that one is able to evaluate according to them, with a sufficient approximation, all the values of $F(x)$ which figure in the formula.

One understands easily the advantage of this formula over that which one finds according to the calculation *of the finite differences*, by noting that here the divisors are comparatively greater, and consequently the errors of the known values of $F(x)$ have less influence over that of $F'(X)$ that one seeks, that which is very important in many cases.