Mémoire sur l'évaluation d'inconnues déterminées par un grand nombre d'équations approximatives du premier degré*

Mr. Augustin Cauchy†

Comptes Rendus Hebd. Séances Acad. Sci. 36, 1114-1122.
OC I, 12 (519), 36-46. Read 27 June 1853.

As Mr. Faye has remarked, the new method of interpolation that I have given, in a Memoir lithographed in 1835, is able to be usefully applied to the evaluation of unknowns determined by a great number of approximate equations of the first degree. We enter into this subject in some details.

We consider m unknowns represented by the letters

$$x, y, z, \dots, u, v, w,$$

and we suppose that, n being a very great number, one gives the approximate values

$$k_1, k_2, \dots, k_n$$

of n linear functions of these unknowns, for example of the functions represented by the polynomials

$$a_1x + b_1y + c_1z + \cdots + h_1w, \quad a_2x + b_2y + c_2z + \cdots + h_2w, \dots, \\ a_nx + b_ny + c_nz + \cdots + h_nw.$$

The exact values of these functions will be of the form

$$k_1 - \varepsilon_1, \quad k_2 - \varepsilon_2, \dots, \quad k_n - \varepsilon_n,$$

$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ designating some quantities of which the numerical values will be very small; and one will have rigorously

$$(1) \qquad \begin{cases} a_1x + b_1y + c_1z + \cdots + h_1w = k_1 - \varepsilon_1, \\ a_2x + b_2y + c_2z + \cdots + h_2w = k_2 - \varepsilon_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_nx + b_ny + c_nz + \cdots + h_nw = k_n - \varepsilon_n. \end{cases}$$

*Memoir on the evaluation of unknowns determined by a great number of approximate equations of the first degree.

†Translated by Richard J. Pulskamp, Department of Mathematics & Computer Science, Xavier University, Cincinnati, OH. July 6, 2010

Let now x be that of the unknowns $x, y, z, \ldots, w$ for which the numerical values of the coefficients offer the greatest sum. We designate this greatest sum by Sa_i, the letter i designating any one of the numbers $1, 2, 3, \ldots, n$; and let

$$Sb_i, \quad Sc_i, \ldots, \quad Sh_i$$

be that which Sa_i becomes when one replaces the coefficients

$$a_1, \quad a_2, \ldots, \quad a_n$$

[1115] by the coefficients

$$b_1, b_2, \ldots, b_n; \quad \text{or} \quad c_1, c_2, \ldots, c_n, \ldots; \quad \text{or} \quad h_1, h_2, \ldots, h_n.$$

One will draw from formulas (1)

$$xSa_i + ySb_i + zSc_i + \cdots + wSh_i = Sk_i - S\varepsilon_i. \tag{2}$$

By aid of this last formula, one will be able to eliminate x from equations (1), and by putting, for brevity,

$$\alpha_i = \frac{a_i}{Sa_i}, \tag{3}$$

$$\begin{cases} b_i - \alpha_i Sb_i = \Delta b_i, \quad c_i - \alpha_i Sc_i = \Delta c_i, \ldots, \quad h_i - \alpha_i Sh_i = \Delta h_i, \\ \qquad k_i - \alpha_i Sk_i = \Delta k_i, \quad \varepsilon_i - \alpha_i S\varepsilon_i = \Delta\varepsilon_i, \end{cases} \tag{4}$$

one will obtain, instead of equations (1), the following:

$$\begin{cases} y\Delta b_1 + z\Delta c_1 + \cdots + w\Delta h_1 = \Delta k_1 - \Delta\varepsilon_1, \\ y\Delta b_2 + z\Delta c_2 + \cdots + w\Delta h_2 = \Delta k_2 - \Delta\varepsilon_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y\Delta b_n + z\Delta c_n + \cdots + w\Delta h_n = \Delta k_n - \Delta\varepsilon_n. \end{cases} \tag{5}$$

Let now y be that of the unknowns $y, z, \ldots, w$ for which, in the first members of the equations (5), the sum of the numerical values of the coefficients is the greatest possible. We designate by $S'\Delta b_i$ that greatest sum, and by

$$S'\Delta b_i, \ldots, \quad S'\Delta h_i$$

that which that sum becomes, when one replaces

$$\Delta b_1, \Delta b_2, \ldots, \Delta b_n$$

by

$$\Delta c_1, \Delta c_2, \ldots, \Delta c_n; \ldots; \quad \text{or by} \quad \Delta h_1, \Delta h_2, \ldots, \Delta h_n.$$

One will draw from equations (5)

$$yS'\Delta b_i + zS'\Delta c_i + \cdots + wS'\Delta h_i = S'\Delta k_i - S'\Delta\varepsilon_i. \tag{6}$$

By aid of that last formula, one will be able to eliminate y from equations (5), and by putting, for brevity,

$$\beta_i = \frac{\Delta b_i}{S'\Delta b_i}, \tag{7}$$

$$\begin{cases} \Delta c_i - \beta_i S'\Delta c_i = \Delta^2 c_i, \dots, \quad \Delta h_i - \beta_i S'\Delta h_i = \Delta^2 h_i, \\ \Delta k_i - \beta_i S'\Delta k_i = \Delta^2 k_i, \quad \Delta\varepsilon_i - \beta_i S'\Delta\varepsilon_i = \Delta^2 \varepsilon_i, \end{cases} \tag{8}$$

[1116] one will find

$$\begin{cases} z\Delta^2 c_1 + \dots + w\Delta^2 h_1 = \Delta^2 k_1 - \Delta^2 \varepsilon_1, \\ z\Delta^2 c_2 + \dots + w\Delta^2 h_2 = \Delta^2 k_2 - \Delta^2 \varepsilon_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ z\Delta^2 c_n + \dots + w\Delta^2 h_n = \Delta^2 k_n - \Delta^2 \varepsilon_n. \end{cases} \tag{9}$$

By continuing in the same manner, one will obtain definitely, in the place of equation (1), a system of equations of the form

$$\begin{cases} w\Delta^{m-1} h_1 = \Delta^{m-1} k_1 - \Delta^{m-1} \varepsilon_1, \\ w\Delta^{m-1} h_2 = \Delta^{m-1} k_2 - \Delta^{m-1} \varepsilon_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ w\Delta^{m-1} h_n = \Delta^{m-1} k_n - \Delta^{m-1} \varepsilon_n; \end{cases} \tag{10}$$

next, by designating by $S^{(m-1)}\Delta^{m-1}h_i$, the sum of the numerical values of $\Delta^{m-1}h_1$, $\Delta^{m-1}h_2$,..., $\Delta^{m-1}h_n$, and by

$$S^{(m-1)}\Delta^{m-1}k_i, \text{ or by } S^{(m-1)}\Delta^{m-1}\varepsilon_i,$$

that which $S^{(m-1)}\Delta^{m-1}h_i$ becomes when one replaces $h_1, h_2, \dots, h_n$ by

$$k_1, k_2, \dots, k_n \quad \text{or} \quad \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n,$$

one will draw from formulas (10)

$$wS^{(m-1)}\Delta^{m-1}h_i = S^{(m-1)}\Delta^{m-1}k_i - S^{(m-1)}\Delta^{m-1}\varepsilon_i. \tag{11}$$

Finally, by eliminating w from equations (10) by aid of formula (11), and putting, for brevity,

$$\eta_i = \frac{\Delta^{m-1}h_i}{S^{(m-1)}\Delta^{m-1}h_i}, \tag{12}$$

$$\begin{aligned} \Delta^m k_i &= \Delta^{m-1}k_i - \eta_i S^{(m-1)}\Delta^{m-1}k_i, \\ \Delta^m \varepsilon_i &= \Delta^{m-1}\varepsilon_i - \eta_i S^{(m-1)}\Delta^{m-1}\varepsilon_i, \end{aligned} \tag{13}$$

one will find

$$0 = \Delta^m k_1 - \Delta^m \varepsilon_1, \dots \tag{14}$$

consequently,

$$\Delta^m \varepsilon_1 = \Delta^m k_1, \quad \Delta^m \varepsilon_2 = \Delta^m k_2, \; \dots, \quad \Delta^m \varepsilon_n = \Delta^m k_n. \tag{15}$$

These last equations determine completely the values of $\Delta^m \varepsilon_1$, $\Delta^m \varepsilon_2$, …, $\Delta^m \varepsilon_n$, that is to say the diverse values of $\Delta^m \varepsilon_i$. If, for brevity, one puts

$$\theta_i = \Delta^m k_i, \tag{16}$$

[1117] one will have generally, by virtue of formulas (15),

$$\Delta^m \varepsilon_i = \theta_i. \tag{17}$$

If, besides, one puts

$$\lambda = S\varepsilon_i, \quad \mu = S'\Delta\varepsilon_i, \dots, \quad \varsigma = S^{(m-1)}\Delta^{m-1}\varepsilon_i, \tag{18}$$

one will draw from formulas (4), (8), …, (13), (17)

$$\varepsilon_i = \alpha_i\lambda + \beta_i\mu + \gamma_i\nu + \cdots + \eta_i\varsigma + \theta_i. \tag{19}$$

By virtue of formula (19), the value of ε_i depends on the values of the m sums represented by the letters

$$\lambda, \mu, \nu, \dots, \varsigma.$$

The most simple hypothesis that one is able to make on the values of these same sums is to suppose them null, that is to say to take

$$S\varepsilon_i = 0, \quad S'\Delta\varepsilon_i = 0, \; \dots, \quad S^{(m-1)}\Delta^{m-1}\varepsilon_i = 0. \tag{20}$$

Then one has generally

$$\varepsilon_i = \theta_i, \tag{21}$$

and the formulas (2), (6), …, (11) give

$$\left\{\begin{array}{l} xSa_i + ySb_i + \cdots + wSh_i = Sk_i, \\ \quad yS'\Delta b_i + \cdots + wS'\Delta h_i = S'\Delta k_i, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \qquad\quad wS^{(m-1)}\Delta^{m-1}h_i = S^{(m-1)}\Delta^{m-1}k_i \end{array}\right. \tag{22}$$

These last equations are those to which the method of interpolation already cited leads. They furnish, for the unknowns x, y, z, …, w, the values that one is able to easily calculate, by beginning with w. These values, which are only approximate, enjoy remarkable properties indicated in the Memoir on interpolation. If one designates them by x, y, z,

…, w, if, besides, one names $\xi, \eta, \zeta, \ldots, \omega$ the errors that they involve, one will have rigorously

$$(23)\qquad \begin{cases} \mathrm{x}Sa_i + \mathrm{y}Sb_i + \mathrm{z}Sc_i + \cdots + \mathrm{w}wSh_i = Sk_i, \\ \quad \mathrm{y}S'\Delta b_i + \mathrm{z}S'\Delta c_i + \cdots + \mathrm{w}S'\Delta h_i = S'\Delta k_i, \\ \quad\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \quad\quad\quad \mathrm{w}S^{(m-1)}\Delta^{m-1}h_i = S^{(m-1)}\Delta^{m-1}k_i \end{cases}$$

and

$$(24)\qquad x = \mathrm{x} - \xi, \quad y = \mathrm{y} - \eta, \quad z = \mathrm{z} - \zeta, \ldots, \quad w = \mathrm{w} - \omega;$$

[1118] and, from equations (2), (6), …, (11), joined to formulas (18), (23), (24), one will draw[1]

$$(25)\qquad \begin{cases} \xi Sa_i + \eta Sb_i + \zeta Sc_i + \cdots + \omega Sh_i = \lambda, \\ \quad \eta S'\Delta b_i + \zeta S'\Delta c_i + \cdots + \omega S'\Delta h_i = \mu, \\ \quad\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \quad\quad\quad \omega S^{(m-1)}\Delta^{m-1}h_i = \varsigma. \end{cases}$$

It is good to observe that by virtue of the formulas (3) and (4), (7) and (8), etc., one has generally

$$(26)\qquad \begin{cases} S\alpha_i = 1, & S\beta_i = 0, & S\gamma_i = 0, \ldots, & S\theta_i = 0, \\ & S'\beta_i = 1, & S'\gamma_i = 0, \ldots, & S'\theta_i = 0, \\ & & S''\gamma_i = 1, \ldots, & S''\theta_i = 0, \\ \cdots\cdots\cdots & \cdots\cdots\cdots & \cdots\cdots\cdots\cdots & \cdots\cdots\cdots \end{cases}$$

This put, one will draw successively from formula (19)

$$(27)\qquad \begin{cases} S\varepsilon_i = \lambda, \\ S'\varepsilon_i = \lambda S'\alpha_i + \mu, \\ S''\varepsilon_i = \lambda S''\alpha_i + \mu S'\beta_i + \nu, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ S^{(m-1)}\varepsilon_i = \lambda S^{(m-1)}\alpha_i + \mu S^{(m-1)}\beta_i + \cdots + \varsigma, \end{cases}$$

and one will be able from formulas (27), joined to equations (25), draw first the values of the coefficients

$$\lambda, \mu, \nu, \ldots, \varsigma,$$

next those of the errors

$$\xi, \eta, \zeta, \ldots, \omega,$$

in a manner to obtain these diverse values expressed in linear functions of the sums

$$S\varepsilon_i, \quad S'\varepsilon_i, \ldots, \quad S^{(m-1)}\varepsilon_i,$$

[1] *Translator's note*: In formula (25), Cauchy gives y rather than η.

or, that which reverts to the same, in linear functions of the errors

$$\varepsilon_1, \quad \varepsilon_2, \dots, \quad \varepsilon_n.$$

By operating thus, one will arrive to some equations of the form

$$(28) \qquad \begin{cases} \xi = \xi_1\varepsilon_1 + \xi_2\varepsilon_2 + \cdots + \xi_n\varepsilon_n, \\ \eta = \eta\varepsilon_1 + \eta_2\varepsilon_2 + \cdots + \eta_n\varepsilon_n, \\ \\ \omega = \omega_1\varepsilon_1 + \omega_2\varepsilon_2 + \cdots + \omega_n\varepsilon_n, \end{cases}$$

[1119] $\xi_1, \xi_2, \dots, \xi_n; \eta_1, \eta_2, \dots, \eta_n; \omega_1, \omega_2, \dots, \omega_n$ being some quantities of which the values will be given in numbers; and, by aid of these equations, one will be able to form an idea of the degree of precision with which each of the unknowns

$$x, y, z, \dots, w$$

is determined by formulas (21), or, that which reverts to the same, by the equations

$$(29) \qquad x = \mathrm{x}, \quad y = \mathrm{y}, \quad z = \mathrm{z}, \dots, \quad w = \mathrm{w}.$$

In fact, the errors

$$\xi, \eta, \zeta, \dots, \omega,$$

that one will commit in taking x, y, z, ..., w for values of the unknowns $x, y, z, \dots, w$, will be equivalent, by virtue of formulas (18), to some linear functions and determined from the errors

$$\varepsilon_1, \quad \varepsilon_2, \dots, \quad \varepsilon_n;$$

and, hence, the limits that the numerical values of $\xi, \eta, \zeta, \dots, \omega$ will be able to attain will depend on the limits that the numerical values of ε_1, $\varepsilon_2, \dots$, ε_n will be able to attain.

We imagine, in order to fix the ideas, that the quantities $k_1, k_2, \dots, k_n$ are all of the same nature, and that, in the determination of each of them, the error to fear is contained between the limits $-\varepsilon$ and ε. Let, besides, Ξ be the sum of the numerical values of the quantities $\xi_1, \xi_2, \dots, \xi_n$; H the sum of the numerical values of the quantities $\eta_1, \eta_2, \dots, \eta_n; \dots$; Ω the sum of the numerical values of the quantities $\omega_1, \omega_2, \dots, \omega_n$. By virtue of formulas (28), when one will take x, y, z, ..., w for approximate values of the unknowns $x, y, z, \dots, w$, the numerical values of the errors to fear will have for limits the products

$$\Xi\varepsilon, \quad H\varepsilon, \dots, \quad \Omega\varepsilon.$$

Hence, if, beneath the unknowns

$$x, y, z, \dots, w$$

one writes the corresponding numbers

$$\Xi, \quad H, \dots, \quad \Omega,$$

then, to a greater number will correspond an unknown for which the limit of the errors to fear will be more considerable. The respective magnitudes [1120] of the inverse numbers

$$\frac{1}{\Xi}, \quad \frac{1}{H}, \ldots, \quad \frac{1}{\Omega}, \tag{30}$$

will furnish therefore an idea of the precision with which the unknowns

$$x, y, z, \ldots, w$$

will be determined by formulas (29).

One will form a more exact idea yet of this precision, if, instead of supposing the numerical values of the errors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ inferiors to a certain limit ε that they are not able to pass, one considers each of them as being able to attain in rigor any numerical value, but with a probability that decreases very rapidly when this numerical value begins to increase, and if one takes for $\Xi, H, \ldots, \Omega$ some numbers proportional to those which would express then the respective probability of the lowering of the numerical values of the errors $\xi, \eta, \ldots, \omega$, below a common and infinitely small limit. This is that which I myself propose to explicate more in detail in another article, by researching how the numbers $\Xi, H, \ldots, \Omega$ would depend then on the coefficients $\xi_1, \xi_2, \ldots, \xi_n$; $\eta_1, \eta_2, \ldots, \eta_n$; $\omega_1, \omega_2, \ldots, \omega_n$.

Before terminating this article, we will remark that of the values of $x, y, z, \ldots, w$ furnished by the new method of interpolation, one is able easily to deduce those that the method known as *least squares* would provide. One will arrive to it, in fact, by operating as follows.

We designate by $\sum \varepsilon_i^2$ the sum of the squares of the errors

$$\varepsilon_1, \quad \varepsilon_2, \ldots, \quad \varepsilon_n;$$

In order that this sum become a *minimum*, as the method of least squares requires, it will suffice to attribute to the quantities

$$\lambda, \mu, \nu, \ldots, \varsigma,$$

contained in the second member of formula (19), some values which verify the linear equations

$$\begin{aligned} &\Sigma\alpha_i(\alpha_i\lambda + \beta_i\mu + \gamma_i\nu + \cdots + \eta_i\zeta + \theta_i) = 0, \\ &\Sigma\beta_i(\alpha_i\lambda + \beta_i\mu + \gamma_i\nu + \cdots + \eta_i\zeta + \theta_i) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\Sigma\eta_i(\alpha_i\lambda + \beta_i\mu + \gamma_i\nu + \cdots + \eta_i\zeta + \theta_i) = 0. \end{aligned} \tag{31}$$

Besides, the diverse values of θ_i being generally very small, one [1121] will be able to say as much of the values of $\lambda, \mu, \nu, \ldots, \varsigma$, and, in calculating them, one will be able to express each of them by aid of a very small number of significant digits. This circumstance will permit to resolve easily equations (31). The resolution being effected, the values of the unknowns $x, y, z, \ldots, w$, will be furnished by the equations

$$x = \mathrm{x} - \xi, \quad y = \mathrm{y} - \eta, \quad z = \mathrm{z} - \zeta, \ldots, \quad w = \mathrm{w} - \omega; \tag{24}$$

the corrections $\xi, \eta, \zeta, \ldots, \omega$ being themselves determined by the system of equations

$$(32) \qquad \begin{cases} \xi Sa_i + \eta Sb_i + \zeta Sc_i + \cdots + \omega Sh_i = \lambda, \\ \quad \eta S'\Delta b_i + \zeta S'\Delta c_i + \cdots + \omega S'\Delta h_i = \mu, \\ \qquad \zeta S''\Delta^2 c_i + \cdots + \omega S''\Delta^2 h_i = \nu, \\ \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad \omega S^{(m-1)}\Delta^{m-1} h_i = \varsigma. \end{cases}$$

By virtue of equations (31) and (32), the corrections $\xi, \eta, \zeta, \ldots, \omega$ will offer some numerical values which will be in general sensibly inferior to those of the quantities $\theta_1, \theta_2, \ldots, \theta_n$. The reason for it is that the coefficients of λ in the first of the equations (31), of μ in the second, etc., of ς in the last, that is to say the sums

$$\Sigma\alpha_i^2, \quad \Sigma\beta_i^2, \ldots, \quad \Sigma\eta_i^2,$$

will be composed of terms which will be all positive, while the other coefficients and the sums

$$\Sigma\alpha_i\theta_i, \quad \Sigma\beta_i\theta_i, \ldots, \quad \Sigma\eta_i\theta_i,$$

will be composed of terms which will be in general positive ones, the others negative. Therefore, seeing that the numerical values of the quantities

$$\theta_1, \quad \theta_2, \ldots, \quad \theta_n$$

will be generally very small, one will be able to say as much a fortiori of the numerical values of the quantities

$$\lambda, \mu, \nu, \ldots, \varsigma,$$

and of the quantities

$$\xi, \eta, \zeta, \ldots, \omega,$$

which will be deduced successively from the first by aid of equations (31) and (32). One must not be surprised therefore to see the results which [1122] the new method of interpolation furnishes to coincide in general very nearly with those to which one is led by the method of least squares.

We remark again that one could apply to equations (32) the method of resolution employed for equations (1). This application will be so much more easy, as the numerical values of the quantities

$$\theta_1, \quad \theta_2, \ldots, \quad \theta_n$$

will be smaller. In fact, when these numerical values, and for stronger reason those of $\lambda, \mu, \nu, \ldots, \varsigma$, will be very near to zero, one will be able ordinarily, in the calculation of these last, to stop after the determination of a small number of decimal digits, for example one or two significant digits.