
ANNALES DE MATHÉMATIQUES PURES ET APPLIQUÉES.

GERGONNE

**Analise. Application de la méthode des moindres carrés
à l'interpolation des suites**

Annales de Mathématiques pures et appliquées, tome 6 (1815-1816), p. 242-252

http://www.numdam.org/item?id=AMPA_1815-1816__6__242_0

© Annales de Mathématiques pures et appliquées, 1815-1816, tous droits réservés.

L'accès aux archives de la revue « Annales de Mathématiques pures et appliquées » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ANALISE.

Application de la méthode des moindres quarrés à l'interpolation des suites ;

Par M. GERGONNE.

LORSQU'UNE fonction d'une seule variable est donnée , on peut toujours déterminer rigoureusement et directement les valeurs , tant de la fonction que de ses divers coefficients différentiels , répondant à une valeur donnée de la variable indépendante ; tout comme , lorsqu'une ligne courbe est donnée , on peut toujours , pour l'une quelconque de ses abscisses , construire l'ordonnée , la tangente , le cercle osculateur , etc.

Mais , de même qu'au lieu de donner une courbe , on peut donner seulement un certain nombre de ses points , on peut aussi , au lieu de donner une fonction d'une variable , donner seulement les valeurs que prend cette fonction pour un certain nombre de valeurs de la variable indépendante , et demander ensuite d'assigner les valeurs , tant de cette fonction que de ses divers coefficients différentiels , pour une autre valeur quelconque de cette variable ; tout comme on pourrait demander quelles sont , pour une abscisse donnée , l'ordonnée , la tangente , le cercle osculateur , etc. , d'une courbe dont on connaîtrait seulement un certain nombre de points. C'est en cela que consiste le problème de *l'interpolation des suites*.

Ce problème se réduit évidemment à remonter des valeurs données à celle de la fonction à laquelle elles appartiennent , ou des points donnés au tracé de la courbe sur laquelle on les suppose situés :

or, par là même il est indéterminé ; car, par des points donnés, non consécutifs, même en nombre infini, on peut toujours faire passer une infinité de courbes différentes (*).

Ces courbes pourront fort bien, dans certaines parties de leur cours, différer les unes des autres d'une manière notable ; et la même différence devra se faire remarquer aussi dans les ordonnées, tangentes, cercles osculateurs, etc., qui répondront à une même abscisse. On conçoit pourtant que, si les points donnés sont assez voisins les uns des autres, les courbes qui les comprendront ne pourront différer notablement, dans l'intervalle embrassé par ces points, du moins si aucune d'elles n'a dans cet intervalle une asymptote parallèle à l'axe des ordonnées ; on conçoit même que ces points pourront toujours être supposés assez multipliés, et, en même temps, assez voisins les uns des autres, pour que les différences entre ces courbes deviennent pour ainsi dire insensibles. Les ordonnées qui répondront à une même abscisse, comprise dans les limites de ces points, seront donc sensiblement égales ; mais la différence entre les tangentes pourra être plus sensible, celle entre les cercles osculateurs encore d'avantage, et ainsi de suite.

Concluons de là que, si des fonctions de formes diverses prennent les mêmes valeurs, pour certaines valeurs déterminées, et voisines les unes des autres, de la variable indépendante, sans devenir infinies pour aucune valeur comprise entre celles-là ; ces fonctions prendront des valeurs peu différentes, pour d'autres valeurs de cette variable, comprises dans les limites qu'embrassent les premières ; mais il n'en sera plus de même des coefficients différentiels successifs qui, d'une fonction à l'autre, pourront différer de plus en plus, à mesure que l'ordre en sera plus élevé.

On pourra donc, sans erreur sensible, adopter indistinctement et arbitrairement l'une des fonctions pour la fonction cherchée ; tout

(*) On peut consulter sur ce sujet une dissertation qui se trouve à la page 252 du V.^e volume de ce recueil.

comme , lorsque plusieurs courbes qui passent par les mêmes points ne présentent entre elles que des différences insensibles , on peut en regarder une quelconque comme étant réellement celle dont ces points font partie.

La courbe et la fonction pouvant ainsi être choisies d'une infinité de manières différentes , il sera convenable de s'arrêter aux plus simples , c'est-à-dire , à la courbe parabolique et à la fonction rationnelle et entière qu'elle représente graphiquement. Ce choix sera d'autant mieux fondé qu'il est connu que toute fonction qui ne devient infinie pour aucune valeur finie de la variable dont elle dépend , est toujours développable en série procédant suivant les puissances ascendantes de cette variable.

Le procédé auquel nous venons d'être conduit est aussi celui qu'on suit communément ; on suppose que l'ordonnée de la courbe cherchée est une fonction complète , rationnelle et entière de l'abscisse , dans laquelle on admet autant de termes qu'il y a de systèmes de valeur donnés ; les coefficients de ces termes sont inconnus , et on les détermine en exprimant que la courbe passe par les points donnés. Ces coefficients une fois déterminés , rien n'est plus facile ensuite que d'assigner l'ordonnée et les coefficients différentiels qui répondent à une abscisse quelconque ; mais on ne peut compter sur les valeurs que la formule leur assignera qu'autant qu'on n'en fera l'application qu'à une abscisse comprise entre celles des points donnés , et même ne se rapprochant pas trop de la plus grande ni de la plus petite.

Cette méthode qui , en particulier , a été employée par M. Laplace , dans son mémoire sur la *Recherche des orbites des comètes* (*), renferme une source d'erreur , dans la supposition , tout à fait gratuite , d'une courbe du genre parabolique. Néanmoins , si l'on pouvait compter en toute rigueur sur les valeurs données de la fonction , et si ces valeurs étaient très-multipliées et très-voisines , ce que

(*) Voyez les *Mémoires de l'académie des sciences* de Paris , pour 1780.

nous avons dit ci-dessus, montre assez que l'erreur résultant de cette supposition ne serait jamais bien considérable.

Mais il n'en va pas ordinairement ainsi ; les valeurs discontinues de la fonction, sur lesquelles on s'appuie pour construire la formule, sont communément déduites d'expériences ou d'observations susceptibles d'une exactitude assez bornée ; et il arrive alors, comme M. Legendre l'a fort bien observé (*), que les erreurs qui les affectent peuvent avoir d'autant plus d'influence sur la formule finale et sur les résultats qu'on en déduit, que ces valeurs sont en plus grand nombre.

Concevons, en effet, qu'on ait tracé une courbe quelconque, et qu'on lui ait mené plusieurs ordonnées peu distantes les unes des autres ; si l'on vient à faire subir à ces ordonnées des altérations, très-légères d'ailleurs, tantôt en plus et tantôt en moins, et qu'ensuite on tente de faire passer une courbe continue par les extrémités de ces ordonnées ainsi altérées, on s'apercevra aisément que, si les altérations qu'elles ont subi n'ont qu'une faible influence sur la grandeur des ordonnées intermédiaires, il n'en est plus ainsi à l'égard de la direction de la tangente qui souvent pour une même abscisse aura pu subir un changement très-notable ; la différence pourra être plus sensible encore à l'égard de la grandeur du cercle osculateur.

Ces aperçus graphiques peuvent facilement être confirmés par le calcul. Supposons, en effet, un nombre impair d'ordonnées données, toutes équidistantes, et dont la distance commune soit prise pour unité. Soient o l'abscisse et b l'ordonnée du milieu ; $1, 2, 3, \dots$ les abscisses et b', b'', b''', \dots les ordonnées qui les suivent ; $-1, -2, -3, \dots$ les abscisses et $b_I, b_{II}, b_{III}, \dots$ les ordonnées qui les précèdent ; et cherchons les coefficients différentiels qui repondent à l'ordonnée du milieu ; nous trouverons, pour le cas de trois ordonnées seulement,

(*) Voyez ses *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, 1806, (pag. iv)

$$\frac{dy}{dx} = \frac{b' - b_1}{2}, \quad \frac{d^2y}{dx^2} = (b' + b_1) - 2b;$$

pour le cas de cinq ordonnées

$$\frac{dy}{dx} = \frac{8(b' - b_1) - (b'' - b_{II})}{12}, \quad \frac{d^2y}{dx^2} = -\frac{30b - 16(b' + b_1) + (b'' + b_{II})}{12};$$

pour le cas de sept coordonnées

$$\frac{dy}{dx} = \frac{45(b' - b_1) - 9(b'' - b_{II}) + (b''' - b_{III})}{60},$$

$$\frac{d^2y}{dx^2} = -\frac{490b - 270(b' + b_1) + 27(b'' + b_{II}) - 2(b''' + b_{III})}{180};$$

et ainsi de suite.

Or, supposons que, toutes les autres ordonnées étant d'ailleurs exactes, l'ordonnée b' seule soit en erreur d'une quantité β , et désignons par $E. \frac{dy}{dx}$, $E. \frac{d^2y}{dx^2}$ les erreurs qui en résulteront sur les coefficients différentiels; il est aisé de voir qu'on aura, dans le cas de trois ordonnées,

$$E. \frac{dy}{dx} = \frac{1}{2}\beta, \quad E. \frac{d^2y}{dx^2} = \frac{1}{2}\beta;$$

dans le cas de cinq coordonnées

$$E. \frac{dy}{dx} = \frac{1}{3}\beta, \quad E. \frac{d^2y}{dx^2} = \frac{4}{3}\beta;$$

dans le cas de sept coordonnées

$$E. \frac{dy}{dx} = \frac{3}{4}\beta, \quad E. \frac{d^2y}{dx^2} = \frac{6}{4}\beta;$$

de sorte que les erreurs sur le coefficient différentiel du premier ordre croissent comme les nombres $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{2}{5}$, ... et tendent ainsi sans cesse à devenir égales à l'erreur même commise sur l'ordonnée b' ;

et que l'erreur commise sur le coefficient différentiel du second ordre est constamment double de celle-là.

M. Legendre a donc été fondé à dire qu'en multipliant les données on s'exposait à faire croître aussi les erreurs dans la même proportion. Il est pourtant juste de remarquer que c'est en supposant qu'il n'y a qu'une seule ordonnée fautive, ce qui exclut toute possibilité de compensation d'erreurs; et en supposant de plus que l'ordonnée fautive est précisément celle dont la valeur, exacte ou non, exerce l'influence la plus notable sur nos deux coefficients différentiels.

Quoi qu'il en soit, cette source d'erreur paraît n'avoir point échappé à l'attention de M. Laplace. Voici, en effet, comment il s'exprime (*Mécanique céleste*, tom. I, pag. 201): « Ces expressions sont » d'autant plus précises, qu'il y a plus d'observations, et que les » intervalles qui les séparent sont plus petits; on pourrait donc » employer toutes les observations voisines de l'époque choisie, si » elles étaient exactes; mais les erreurs dont elles sont toujours » susceptibles conduiraient à un résultat fautif; ainsi, *pour diminuer » l'influence de ces erreurs, il faut augmenter l'intervalle des obser- » vations extrêmes, à mesure que l'on emploie plus d'observations.* »

Il serait peut-être plus exact de dire qu'il faut employer des observations de plus en plus distantes entre elles, à mesure qu'on en emploie un plus grand nombre; et nous allons voir, en effet, qu'avec cette attention, on peut, à volonté, atténuer les erreurs. Soit a l'intervalle, supposé constant, qui sépare les valeurs consécutives de x ; intervalle que, ci-dessus, nous avons pris pour unité. Nos résultats deviendront alors

$$\text{Pour 3 observations, } E. \frac{dy}{dx} = \frac{1}{2} \frac{\beta}{a}, \quad E. \frac{d^2y}{dx^2} = \frac{1}{2} \frac{\beta}{a^2};$$

$$\text{Pour 5 observations, } E. \frac{dy}{dx} = \frac{1}{3} \frac{\beta}{a}, \quad E. \frac{d^2y}{dx^2} = \frac{2}{3} \frac{\beta}{a^2};$$

$$\text{Pour 7 observations, } E. \frac{dy}{dx} = \frac{1}{4} \frac{\beta}{a}, \quad E. \frac{d^2y}{dx^2} = \frac{3}{4} \frac{\beta}{a^2}.$$

Pourvu donc que nous prenions pour a des nombres qui croissent plus rapidement que ceux de la suite $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$ nos erreurs iront continuellement en décroissant, à mesure que nous aurons recours à un plus grand nombre d'observations. Supposons, par exemple, que nous fassions croître les valeurs de a suivant les nombres de la suite naturelle ; et prenons pour unité la valeur de cette quantité qui répond au cas de trois observations, nous aurons alors

Pour 3 observations , $E. \frac{1}{2}\beta$, $E. \frac{1}{3}\beta$;

Pour 5 observations , $E. \frac{1}{3}\beta$, $E. \frac{1}{4}\beta$;

Pour 7 observations , $E. \frac{1}{4}\beta$, $E. \frac{1}{5}\beta$;

d'où l'on voit qu'alors les erreurs sur les coefficients différentiels du premier ordre décroîtront comme les inverses des nombres naturels, et que celles qui affecteront les coefficients différentiels du second ordre décroîtront suivant la progression, plus rapide encore, des inverses des nombres triangulaires. La méthode de M. Laplace est donc, du moins de ce côté, tout à fait à l'abri du reproche.

Mais, supposons qu'on ait, entre deux limites fixes données, des observations assez nombreuses pour rendre très-petite la différence entre les valeurs consécutives de x . Suivant ce qui vient d'être dit, on devra rejeter un d'autant plus grand nombre de ces observations qu'on en voudra employer davantage dans la recherche de $\frac{dy}{dx}$ et $\frac{d^2y}{dx^2}$.

Or, c'est là un inconvénient assez grave, sur-tout si l'on n'a aucun motif de suspecter plutôt les données que l'on rejette que celles dont on se propose de faire exclusivement usage ; puisqu'on se prive ainsi des compensations d'erreurs sur lesquelles on pourrait compter en les employant toutes.

En réfléchissant sur ce sujet, il m'a paru qu'il était possible de tout concilier, au moyen de la méthode des moindres carrés (*),

(*) On sait que la méthode des moindres carrés repose sur ce principe que la valeur moyenne, la plus probablement voisine de l'exactitude, d'une quantité

et d'arriver par elle à toute la précision qu'il est possible d'espérer dans la recherche qui nous occupe. Voici pour cela de quelle manière je conçois qu'on en doit faire usage.

Soient a, a', a'', \dots des valeurs de x , en nombre quelconque, et soient b, b', b'', \dots les valeurs données et correspondantes de y . Soit posé

$$y = A + Bx + Cx^2 + Dx^3 + \dots ; \quad (1)$$

en prenant dans cette fonction autant de termes seulement qu'on en admettrait si, suivant ce qui vient d'être dit ci-dessus, on ne se proposait d'employer qu'une partie des valeurs correspondantes de x et de y ; il s'agira de déterminer les valeurs des coefficients A, B, C, D, \dots . Si leur nombre était égal à celui des observations, on pourrait leur assigner des valeurs qui rendissent les erreurs tout à fait nulles; mais la chose sera impossible dans le cas actuel, et il faudra se contenter de rendre *minimum* la somme de leurs carrés.

Ces erreurs étant respectivement

$$A + Ba + Ca^2 + Da^3 + \dots - b ;$$

$$A + Ba' + Ca'^2 + Da'^3 + \dots - b' ,$$

$$A + Ba'' + Ca''^2 + Da''^3 + \dots - b'' ,$$

$$\dots \dots \dots ,$$

il faudra faire

dont on a plusieurs valeurs approchées, est celle qui, étant supposée tout à fait exacte, rendrait *minimum* la somme des carrés des erreurs dont les autres seraient alors affectées. Le premier ouvrage imprimé dans lequel il ait été fait mention de cette méthode est le mémoire de M. Legendre, déjà cité dans une précédente note (1806). Dans un ouvrage publié en 1809, M. Gauss a déclaré faire usage d'une semblable méthode depuis 1795; et M. Laplace a démontré postérieurement que cette méthode est rigoureusement conforme à la doctrine des probabilités.

$$\left. \begin{aligned} &(A+Ba +Ca^2+Da^3+\dots-b)^2 \\ &+(A+Ba'+Ca'^2+Da'^3+\dots-b')^2 \\ &+(A+Ba''+Ca''^2+Da''^3+\dots-b'')^2 \\ &+\dots \end{aligned} \right\} = \text{minimum} ;$$

c'est-à-dire, en différentiant par rapport à A, B, C, D, \dots

$$\left. \begin{aligned} &(A+Ba +Ca^2+\dots-b)(dA+a dB+a^2dC+\dots) \\ &+(A+Ba'+Ca'^2+\dots-b')(dA+a' dB+a'^2dC+\dots) \\ &+(A+Ba''+Ca''^2+\dots-b'')(dA+a''dB+a''^2dC+\dots) \\ &+\dots \end{aligned} \right\} = 0 .$$

A cause de l'indépendance entre A, B, C, \dots les multiplicateurs de dA, dB, dC, \dots devront séparément être nuls ; faisant donc en général, pour abrégé,

$$\begin{aligned} \Sigma a^m &= a^m + a'^m + a''^m + \dots , \\ \Sigma a^m b &= a^m b + a'^m b' + a''^m b'' + \dots ; \end{aligned}$$

on aura cette suite d'équations

$$\left. \begin{aligned} \Sigma a^0 . A + \Sigma a^1 . B + \Sigma a^2 . C + \Sigma a^3 . D + \dots &= \Sigma a^0 b , \\ \Sigma a^1 . A + \Sigma a^2 . B + \Sigma a^3 . C + \Sigma a^4 . D + \dots &= \Sigma a^1 b , \\ \Sigma a^2 . A + \Sigma a^3 . B + \Sigma a^4 . C + \Sigma a^5 . D + \dots &= \Sigma a^2 b , \\ \dots & ; \end{aligned} \right\} (1)$$

en nombre précisément égal à celui des coefficients A, B, C, D, \dots qu'il s'agit de déterminer ; et, tandis que les méthodes ordinaires donnent pour y et ses coefficients différentiels des valeurs d'une précision toujours un peu inférieure à celle des données d'après lesquelles on les calcule, on pourra le plus souvent espérer ici de l'emporter en précision sur ces données elles-mêmes.

Le cas le plus simple, et en même temps le plus fréquent, est celui où les valeurs de x sont en progression par différences ; il est alors permis de substituer à cette progression la suite naturelle des

nombres. Soit $2n+1$ le nombre des valeurs connues et correspondantes de x et y , on pourra numérotter 0 la valeur de x qui se trouvera occuper le milieu, de manière que le numérotage soit

$$-n, -(n-1), \dots, -3, -2, -1, 0, +1, +2, +3, \dots, +(n-1), +n$$

si alors on désigne par Σn^m la somme des m^{mes} puissances des nombres de la suite naturelle, on aura

$$\Sigma a^0 = 2n+1, \Sigma a = 0, \Sigma a^2 = 2\Sigma n^2, \Sigma a^3 = 0, \Sigma a^4 = 2\Sigma n^4, \dots;$$

au moyen de quoi les équations (1) deviendront

$$\begin{aligned} (2n+1)A + 2\Sigma n^2.C + \dots &= \Sigma b, & 2\Sigma n^2.B + 2\Sigma n^4.D + \dots &= \Sigma a b, \\ 2\Sigma n^2.A + 2\Sigma n^4.C + \dots &= \Sigma a^2 b, & 2\Sigma n^4.B + 2\Sigma n^6.D + \dots &= \Sigma a^3 b, \\ 2\Sigma n^4.A + 2\Sigma n^6.C + \dots &= \Sigma a^4 b, & 2\Sigma n^6.B + 2\Sigma n^8.D + \dots &= \Sigma a^5 b, \\ \dots & \dots & \dots & \dots \end{aligned}$$

ainsi, outre que les sommes de puissances semblables des nombres naturels sont données par des formules connues et générales, on aura ici l'avantage de pouvoir calculer séparément les coefficients de rangs pairs et ceux de rangs impairs, ce qui simplifiera le travail d'une manière notable.

Dans le cas même où ni les valeurs de x ni celles de y ne marcheraient en progression par différences, on pourrait encore profiter de ces simplifications, en procédant comme il suit : on supposerait que x et y sont toutes deux fonctions d'une troisième variable z , dont les valeurs, tout à fait arbitraires, pourraient être numérotées comme nous l'avons dit ci-dessus à l'égard de x ; on chercherait par notre procédé, les valeurs de $\frac{dx}{dz}, \frac{dy}{dz}, \frac{d^2x}{dz^2}, \frac{d^2y}{dz^2}, \dots$; et on aurait ensuite par les formules connues, relatives au changement de la variable indépendante,

$$\frac{dy}{dx} = \frac{\frac{dy}{dz}}{\frac{dx}{dz}}, \quad \frac{d^2y}{dx^2} = \frac{\frac{dx}{dz} \frac{d^2y}{dz^2} - \frac{dy}{dz} \frac{d^2x}{dz^2}}{\left(\frac{dx}{dz}\right)^2};$$

cette méthode me semblerait préférable à celle qui consiste à interpoler entre les observations, afin de les rendre équidistantes; attendu qu'il peut être dangereux, dans un problème d'une nature aussi délicate, de dénaturer les données avant d'en faire usage.

Il nous paraît que l'introduction des procédés que nous venons d'indiquer, dans la méthode de M. Laplace, pour la détermination des orbites des comètes, ne peut qu'ajouter beaucoup à sa précision, du moins dans le cas où l'on peut disposer d'un grand nombre d'observations; mais cette méthode, comme beaucoup d'autres, aura toujours le grave inconvénient de n'être, au fond, qu'un tâtonnement bien dirigé.

Il resterait ici au surplus un autre problème à résoudre, lequel pourrait être énoncé comme il suit: *On sait que des points donnés, en nombre quelconque, sont à peu près situés sur une courbe parabolique d'un degré déterminé, mais inconnu; et l'on demande de découvrir quel est le plus probablement le degré de cette courbe?* La solution de ce problème leverait complètement l'incertitude du calculateur qui, voulant appliquer la méthode de M. Laplace, se trouve pouvoir disposer d'un grand nombre d'observations.