Nouvelles annales de mathématiques

Housel

Méthode de M. Cauchy pour modifier la méthode de Newton dans la résolution des équations numériques

Nouvelles annales de mathématiques 1re série, tome 15 (1856), p. 244-256

<http://www.numdam.org/item?id=NAM_1856_1_15__244_0>

MÉTHODE DE M. CAUCHY

Pour modifier la méthode de Newton dans la résolution des équations numériques ;

PAR M. HOUSEL,
Professeur.

1. On sait que la méthode de Newton est quelquefois en défaut, c'est-à-dire que la correction qu'elle fournit, ajoutée à la première valeur approchée de la racine, donne quelquefois une somme plus éloignée de cette racine que les limites qui la comprennent. Aussi divers mathématiciens, tels que Fourier, ont cherché avec plus ou moins de succès à la perfectionner en écartant ces causes d'erreur. M. Cauchy a résolu complétement ce problème; mais, quoique son travail soit publié depuis vingt ans, quoique M. l'abbé Moigno ait essayé plusieurs fois, et notamment dans le tome X des *Nouvelles Annales,* de diriger sur ce sujet l'attention du monde savant, ce procédé n'est pas connu comme il devrait l'être, et nous croyons utile de le rappeler, en y joignant quelques observations et surtout la manière de resserrer une racine entre deux valeurs, l'une supérieure, l'autre inférieure.

2. Soit

$$y = f(x)$$

une fonction quelconque algébrique ou transcendante ; on demande la plus petite racine α de l'équation

$$f(x) = 0$$

comprise entre deux valeurs de x, a et A.

Nous remarquerons que cette question, étant résolue d'une manière générale, dispensera de la séparation des racines : en effet, dès qu'on aura trouvé avec une approximation suffisante la plus petite racine α comprise entre a et A, en ajoutant à cette racine α une quantité suffisamment petite, on obtiendra sans peine une valeur a' telle, que a et a' ne contiennent que la seule racine α. Donc on pourra trouver entre a' et A une seconde racine, et ainsi de suite, s'il y a lieu.

On peut toujours supposer α positif, sauf à changer le signe de x si cela est nécessaire ; dès lors A sera positif : mais on peut aussi supposer $a \geqq 0$, car rien n'empêche de partir de la valeur $a = 0$. Enfin, observons que l'on peut encore supposer $f(a) > 0$: en effet, considérons la courbe dont l'équation est

$$y = f(x);$$

si $f(a)$ est négatif, il suffira de changer le sens des ordonnées, ce qui revient à changer le signe de tous les termes de $f(x)$.

3. Cela posé, toute la méthode est fondée sur le théorème suivant qui se trouve dans quelques Algèbres, mais que nous croyons devoir démontrer ici, parce qu'il n'est pas très-répandu.

Soient R *la plus petite et* S *la plus grande valeur que reçoit la fonction dérivée* $f'(x)$ *lorsque* x *croît par degrés insensibles depuis* a *jusqu'à* A ; *la valeur du rapport*

$$F(x) = \frac{f(x) - f(a)}{x - a}$$

sera toujours comprise entre R *et* S, *si* x *est compris entre* a *et* A.

En effet, $f'(x)$ étant la limite du rapport de l'accroisment de $f(x)$ à celui de x, on peut toujours trouver un nombre h assez petit pour que l'on ait

$$\frac{f(x+h)-f(x)}{h} > f'(x) - \varepsilon$$

et

$$\frac{f(x+h)-f(x)}{h} < f'(x) + \varepsilon,$$

ε étant une quantité aussi petite que l'on voudra. Donc, à plus forte raison,

$$\frac{f(x+h)-f(x)}{h} > \mathrm{R} - \varepsilon$$

et

$$\frac{f(x+h)-f(x)}{h} < \mathrm{S} + \varepsilon.$$

D'après cela, concevons que l'on interpose entre les limites a et x des valeurs croissantes de x, savoir $x_1, x_2, \ldots, x_n$ assez rapprochées pour que les conditions précédentes soient toujours remplies quand on prendra l'une de ces valeurs pour x et la suivante pour $x+h$. Les fractions

$$\frac{f(x_1)-f(a)}{x_1-a}, \quad \frac{f(x_2)-f(x_1)}{x_2-x_1}, \ldots, \quad \frac{f(x)-f(x_n)}{x-x_n}$$

seront toutes plus grandes que $\mathrm{R}-\varepsilon$ et toutes plus petites que $\mathrm{S}+\varepsilon$.

Or, comme toutes ces fractions ont des dénominateurs positifs, si l'on divise la somme de leurs numérateurs par la somme de leurs dénominateurs, on obtiendra une nouvelle fraction qui sera elle-même comprise entre $\mathrm{R}+\varepsilon$ et $\mathrm{S}+\varepsilon$; car si l'on additionne les inégalités

$$f(x_1)-f(a) > (x_1-a)(\mathrm{R}-\varepsilon),$$
$$f(x_2)-f(x_1) > (x_2-x_1)(\mathrm{R}-\varepsilon), \ldots,$$
$$f(x)-f(x_n) > (x-x_n)(\mathrm{R}-\varepsilon),$$

on aura, en réduisant,

$$f(x)-f(a)>(x-a)(\mathrm{R}-\varepsilon)$$

ou

$$\frac{f(x)-f(a)}{x-a}>\mathrm{R}-\varepsilon.$$

De même

$$\frac{f(x)-f(a)}{x-a}<\mathrm{S}+\varepsilon;$$

donc à la limite, pour $\varepsilon=0$, on trouvera le théorème énoncé.

4. Ainsi, x étant compris entre a et A, F (x) sera une des valeurs que prendra $f'(x)$ pour x compris entre a et A; or, puisque $f(a)>0$, α étant la plus petite racine ou, ce qui revient au même, α étant la première racine de $f(x)=0$ depuis a jusqu'à A, on voit que $f(x)$ diminue quelque part entre a et α, et, par conséquent, $f'(x)$ prend des valeurs négatives dans ce même intervalle.

5. Cela posé, l'équation

$$y=f(x)$$

de la courbe que nous considérons pouvant être mise sous la forme

$$y=f(a)+(x-a)\,\mathrm{F}(x),$$

comparons-la avec l'équation

$$y_1=f(a)+(x-a)\,m$$

d'une droite qui partira évidemment du même point que la courbe, puisque $x-a$ donnera

$$y=y_1=f(a),$$

et cherchons quel doit être le coefficient m pour que la droite rencontre l'axe des x avant le point de la courbe qui correspond à $x = \alpha$. Il suffit pour cela que, dans l'intervalle de a jusqu'à α, on ait toujours $y_1 < y$, car si cette condition est remplie, l'ordonnée de la courbe sera encore positive quand celle de la droite sera annulée. Il est clair qu'on y parviendra en donnant à m une valeur inférieure à toutes celles que peut avoir F (x) dans l'intervalle indiqué. Ainsi tout se réduit à prendre le minimum de F (x), c'est-à-dire celui de $f'(x)$ depuis $x = a$ jusqu'à $x = \alpha$, ou plutôt depuis a jusqu'à A, intervalle qui comprend le premier. Comme nous avons vu que $f'(x)$ devenait négatif quelque part entre a et α, on conçoit que ce minimum sera un maximum numérique des valeurs négatives de $f'(x)$.

Voici maintenant ce qu'il y a de simple et d'ingénieux à la fois dans la manière dont M. Cauchy détermine cette quantité m. Il met à part les termes positifs et les termes négatifs de $f(x)$ et pose

$$f(x) = \lambda(x) - \mu(x),$$

ce qui donnera aussi

$$f'(x) = \lambda'(x) - \mu'(x),$$

de sorte que $f'(x)$ sera décomposé de la même manière en un groupe positif et un groupe négatif. Or, x croissant constamment depuis a jusqu'à A, il est évident que l'on aura le minimum de $f'(x)$ en remplaçant x par a dans $\lambda'(x)$ et par A dans $\mu'(x)$; donc enfin

$$m = \lambda'(a) - \mu'(A).$$

m ainsi déterminé donne naturellement $y_1 < y$. Donc, si l'ordonnée $y_1 = 0$, la valeur correspondante de l'abscisse x sera comprise entre a et la racine α, et sera une valeur plus rapprochée de α; cette valeur de x, pour laquelle

$$f(a) + m(x - a) = 0,$$

donnera

$$x - a = -\frac{f(a)}{m}$$

et

$$x = a - \frac{f(a)}{m}.$$

On voit maintenant comment on pourra approcher indéfiniment de α. Soit

$$a_1 = a - \frac{f(a)}{m}$$

cette première valeur de x; nous aurons ensuite

$$a_2 = a_1 - \frac{f(a_1)}{m_1},$$

en posant

$$m_1 = \lambda'(a_1) - \mu'(A).$$

On aura de même

$$a_3 = a_2 - \frac{f(a_2)}{m_2},$$

et ainsi de suite. On trouvera d'une manière analogue la plus grande des racines comprises entre a et A, en revenant de A à a.

6. Cette méthode présente le cachet de rigueur et de généralité qui caractérise les travaux de M. Cauchy, mais elle est moins avantageuse que celle de Newton, quand celle-ci est applicable. En effet, au lieu de prendre

$$x = a - \frac{f(a)}{\lambda'(a) - \mu'(A)},$$

Newton pose

$$x = a - \frac{f(a)}{\lambda'(a) - \mu'(a)},$$

puisque

$$f'(a) = \lambda'(a) - \mu'(a).$$

Or, de ces deux quantités, m et $f'(a)$, toutes deux négatives, du moins quand la correction de Newton est applicable, la plus grande numériquement est m, parce que la portion négative $\mu'(\mathrm{A})$ est plus grande que dans $f'(a)$. Donc, dans la méthode de Newton, le dénominateur de la correction étant plus petit, la correction sera plus considérable, et, par suite, plus avantageuse.

D'après cela, il est important de pouvoir reconnaître dans quelles circonstances la méthode de Newton est applicable. M. Cauchy est encore parvenu à résoudre ce problème, ainsi qu'on va le voir.

Si $f'(x)$ croît d'une manière continue de a jusqu'à A, les valeurs négatives de cette quantité décroissant numériquement, il est clair que $f'(a)$ sera le minimum cherché, c'est-à-dire le maximum des valeurs négatives que nous avons appelé m; on aura donc

$$a_1 = a - \frac{f(a)}{f'(a)}.$$

C'est la correction de Newton.

Si, au contraire, $f'(x)$ va toujours en décroissant depuis a jusqu'à A, les valeurs négatives de cette quantité croissant numériquement, on pourra évidemment prendre $f'(\mathrm{A})$ pour valeur de m, ce qui donnera

$$a_1 = a - \frac{f(a)}{f'(\mathrm{A})}.$$

Cette correction a encore été indiquée par Newton, cependant on néglige de la faire connaître dans les ouvrages élémentaires, sans doute parce qu'elle n'a pas, comme la correction $-\frac{f(a)}{f'(a)}$, une signification géométrique relative à la tangente en un point de la courbe. Mais elle n'en est pas moins utile, et l'on doit remarquer qu'on trouve bien plus souvent qu'on ne le pense l'occasion d'employer la méthode de Newton, si l'on appelle ainsi l'ensemble des deux corrections que nous venons d'examiner.

Pour reconnaître si $f'(x)$ varie d'une manière continue, observons quel sera le signe de

$$f''(x) = \lambda''(x) - \mu''(x),$$

qui est la dérivée de $f'(x)$. Si l'on fait tour à tour dans la somme des termes positifs $x = a$ et $x = A$, et dans la somme des termes négatifs $x = A$ et $x = a$, on obtiendra deux différences $\lambda''(a) - \mu''(A)$ et $\lambda''(A) - \mu''(a)$, dont la première est évidemment inférieure, la seconde évidemment supérieure à toutes les valeurs de $f''(x)$ dans l'intervalle de $x = a$ à $x = A$. Donc si ces deux différences sont de même signe, il en sera de même, à plus forte raison, pour toutes les valeurs de $f''(x)$ comprises dans cet intervalle; ainsi le caractère cherché pour la continuité sera

$$\frac{\lambda''(a) - \mu''(A)}{\lambda''(A) - \mu''(a)} > 0,$$

puisque les deux termes de cette fraction devront avoir le même signe.

Cette condition étant remplie, si les deux signes des termes de la dernière fraction sont positifs, $f'(x)$ croît toujours dans l'intervalle indiqué : donc la correction est $-\frac{f(a)}{f'(a)}$; si le deux signes sont négatifs, $f'(x)$ décroît constamment: donc la correction sera $-\frac{f(a)}{f'(A)}$.

Nous allons voir ce qu'il faut faire quand les deux signes sont opposés.

7. Jusqu'à présent nous avons supposé que les corrections étaient toujours additives, ce qui ne permettait d'approcher de la racine que d'un seul côté. Nous allons

chercher maintenant à resserrer cette racine entre deux valeurs, l'une inférieure et toujours croissante, comme précédemment, l'autre supérieure et toujours décroissante. Seulement, il faut admettre pour cela que cette racine soit séparée, et nous avons vu que les recherches précédentes permettent d'y parvenir.

Nous supposerons donc que α est la seule racine comprise entre a et A; alors, puisque $f(a) > 0$, on a nécessairement $f(A) < 0$.

On s'approchera toujours de a vers α en posant

$$a_1 = a - \frac{f(a)}{m}$$

et

$$m = \lambda'(a) - \mu'(A).$$

En modifiant légèrement les raisonnements qui ont conduit à cette correction, on trouvera pour approcher de A vers α, la valeur

$$A_1 = A - \frac{f(A)}{m}$$

que l'on peut d'ailleurs vérifier en remarquant que la correction $-\frac{f(A)}{m}$ est négative comme cela doit être puisque $f(A)$ et m sont tous deux négatifs; de plus

$$m = \lambda'(a) - \mu'(A)$$

étant maximum numérique des valeurs négatives de $f'(x)$, la correction ne peut être trop considérable.

Cette méthode réussit dans tous les cas possibles; mais pour que l'on soit forcé de l'appliquer, c'est-à-dire pour que la méthode de Newton ne puisse être employée, il faut que les limites $x = a$ et $x = A$ de la racine α comprennent un changement dans la marche des valeurs de $f'(x)$ qui croîtrait après avoir décru ou réciproquement;

en d'autres termes, il faut que la courbe représentée par l'équation

$$y = f(x)$$

ait une inflexion dans cet intervalle. Dans cette circonstance même, on peut généralement ramener cette recherche au cas ordinaire en prenant pour l'une des limites de α la valeur de x qui correspond à $f''(x) = 0$ et qui donne par conséquent le point d'inflexion.

8. Voyons ce qu'il faut faire pour resserrer la racine en plus et en moins quand la méthode de Newton est applicable, puisque l'emploi en est plus avantageux quand il est légitime.

On reconnaîtra sans peine, d'après ce qui a été dit jusqu'à présent, que si $f'(x)$ croît toujours de a jusqu'à A, on doit prendre

$$a_1 = a - \frac{f(a)}{f'(a)}$$

et

$$A_1 = A - \frac{f(A)}{f'(a)}.$$

Si, au contraire, $f'(x)$ décroît toujours, il faut poser

$$a_1 = a - \frac{f(a)}{f'(A)}$$

et

$$A_1 = A - \frac{f(A)}{f'(A)} \text{ (*)}.$$

Enfin nous terminerons en rappelant que nous avons supposé la racine α positive, ainsi que ses limites a et A : de plus, nous avons aussi supposé $f(a) > 0$, et, par suite,

(*) Fourier indique cette méthode, mais il la soumet à des restrictions qui ne sont pas toutes nécessaires.

$f(A) < 0$. Si l'énoncé de la question ne rentrait pas dans ces suppositions, il serait facile de l'y ramener.

9. Il nous reste maintenant à donner quelques exemples.

Soit l'équation

$$x^3 - 2x^2 - x + 1 = 0$$

dont les racines sont comprises l'une entre 0 et -1, l'autre entre 0 et 1, et enfin la troisième entre 2 et 3. Cherchons la seconde racine pour laquelle $a = 0$, $A = 1$, ce qui donne

$$f(a) = 1 \quad \text{et} \quad f(A) = -1.$$

On trouve

$$f'(x) = 3x^2 - 4x - 1$$

et

$$f''(x) = 6x - 4;$$

donc $f''(x)$ prenant des valeurs différentes pour les limites qui comprennent la racine, on ne pourra pas compter sur la méthode de Newton, et il faut prendre avec M. Cauchy

$$m = \lambda'(a) - \mu'(A),$$

ce qui donne

$$m = -5, \quad a_1 = 0,2$$

et

$$A_1 = 0,8.$$

Ensuite

$$m_1 = \lambda'(a_1) - \mu'(A_1) = -4,08,$$

d'où l'on tire

$$a = 0,38 \quad \text{et} \quad A_2 = 0,66;$$

on trouvera de même

$$a_3 = 0,5, \quad A_3 = 0,584.$$

Une fois arrivé à ce point, on pourra employer la méthode de Newton. En effet, le point d'inflexion de la courbe dont l'équation est

$$y = f(x)$$

étant donné par la valeur de x qui rend nulle la dérivée seconde $6x - 4$, c'est-à-dire par la valeur $x = 0,666$, on voit que la courbe ne subit pas d'inflexion depuis $x = 0,5$ jusqu'à $x = 0,584$. Comme $f''(x)$ est négatif pour ces deux valeurs, $f'(x)$ diminue dans cet intervalle; donc on devra poser

$$a_4 = a_3 - \frac{f(a_3)}{f'(A_3)}$$

et

$$A_4 = A_3 - \frac{f(A_3)}{f'(A_3)},$$

ce qui donnera

$$a_4 = 0,554 \quad \text{et} \quad A_4 = 0,555.$$

En continuant ainsi l'on trouve

$$x = 0,55496.$$

10. Considérons encore l'équation transcendante

$$x.2^x = 30 \ (*).$$

Il est clair qu'il n'y a pas de racines négatives, puisque le premier membre serait alors négatif; de plus, il n'y aura qu'une seule racine positive, car le premier membre croît indéfiniment à partir de $x = 0$.

(*) Cette équation est déjà résolue par Pacciolo. Tm.

Posons

$$f(x) = 30 - x.2^x;$$

on reconnaît que la racine est comprise entre $a = 3$ qui donne

$$f(a) = 6,$$

et $A = 4$ pour lequel

$$f(A) = -34.$$

Ensuite

$$f'(x) = -2^x(1 + xl2)$$

et

$$f''(x) = -2^x l2(2 + xl2),$$

en indiquant par $l2$ le logarithme népérien de 2, c'est-à-dire 0,69314718. On peut voir facilement que $f''(x)$ étant toujours négatif, $f'(x)$ décroît constamment; on a donc

$$a_1 = a - \frac{f(a)}{f'(A)}$$

et

$$A_1 = A - \frac{f(A)}{f'(A)},$$

car la méthode de Newton peut s'appliquer ici puisqu'il n'y a pas de changement de courbure dans l'intervalle de $x = 3$ à $x = 4$.

On a

$$f'(34) = -60,3614,$$

ce qui donne

$$a_1 = 3,1 \quad \text{et} \quad A_1 = 3,4.$$

On calculera facilement 2^x par logarithmes, et l'on parviendra bientôt à la valeur $x = 3,22$ qui est suffisamment approchée, car

$$f(3,22) = 0,00001.$$