REVUE FRANÇAISE D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE. SÉRIE ROUGE

CLAUDE BREZINSKI

Brève communication. Résultats sur les procédés de sommation et l'ε-algorithme

Revue française d'informatique et de recherche opérationnelle. Série rouge, tome 4, nº R3 (1970), p. 147-153

<http://www.numdam.org/item?id=M2AN_1970__4_3_147_0>

R.I.R.O
4e année, R-3, 1970, p. 147-153)

RESULTATS SUR LES PROCEDES DE SOMMATION ET L'ε-ALGORITHME

par Claude BREZINSKI (*)

Résumé. — *Après avoir posé les problèmes de l'accélération de la convergence des suites à l'aide des procédés de sommation et ceux d'extrapolation, on montre que tout procédé total de sommation est un cas particulier d'extrapolation par des sommes d'exponentielles donc de l'ε-algorithme. On démontre la continuité et la différentiabilité de l'ε-algorithme et on donne une valeur optimale du paramètre a du procédé de Baranger.*

Le problème que nous abordons ici est celui de l'accélération de la convergence de suites qui tendent vers une limite finie. On peut le formuler de la façon générale suivante : soit E un espace vectoriel topologique et E' son dual. Soit $\{ x'_n \}$ une suite d'éléments de E' qui converge faiblement vers x'_∞. Le problème est celui de l'estimation de $\langle x, x'_\infty \rangle$ à partir de certains $\langle x, x'_n \rangle$ $\langle .,. \rangle$ désigne le produit scalaire dans la dualité entre E et E'). Ce problème a été abordé de deux façons différentes : par les procédés de sommation comme ceux d'Euler ou de Césaro et par les méthodes d'extrapolation comme celles de Richardson, Romberg ou le Δ^2 d'Aitken. Après avoir formulé ces deux points de vue nous allons montrer qu'en fait les procédés totaux de sommation sont des cas particuliers d'extrapolation par des sommes d'exponentielles et, par conséquent, sont moins puissants que le Δ^2 d'Aitken généralisé.

LE PROBLEME DE L'EXTRAPOLATION

Soit E un espace vectoriel topologique et E' son dual. Il s'agit d'étudier la nature des $y \in E$ pour lesquels on peut calculer $\langle y, y'_\infty \rangle$ à partir de certains $\langle y, y'_n \rangle$ où $\{ y'_n \}$ est une suite de E' qui converge faiblement vers y'_∞.

(*) Attaché aux Services Techniques des Armées.

Soit $\varphi : \mathbf{R} \times D \to \mathbf{R}$, $D \subseteq \mathbf{R}^k$ telle que $\forall\, a \in D : \lim\limits_{x \to \infty} \varphi(x, a) = 0$. Supposons que φ vérifie la propriété suivante : $\forall\, \alpha = (\alpha_1, ..., \alpha_k) \in \mathbf{R}^k$ et assez voisin de $0, \exists\, n_0 : \forall\, x_{k+1} > x_k > ... > x_1 \geqslant x_{n_0}$, le système :

$$\alpha_1 = \varphi(x_1, a) - \varphi(x_2, a)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_k = \varphi(x_k, a) - \varphi(x_{k+1}, a)$$

admette une et une seule solution a. Soit $V_\varphi \{ y'_n \}$ la variété de E des y tels que $\langle y, y'_\infty - y'_n \rangle = \varphi(x_n, a)\ \forall\, n \in \mathbf{N}$ où les x_n forment une suite monotone croissante. Si $y \in V_\varphi \{ y'_n \}$ et si n_1 est un rang tel que $| \langle y, y'_n \rangle - \langle y, y'_{n-1} \rangle |$ soit assez petit $\forall\, n \geqslant n_1$ alors dès que $n \geqslant n_0, n_1$ on aura :

$$\langle y, y'_{n+1} \rangle - \langle y, y'_n \rangle = \varphi(x_n, a) - \varphi(x_{n+1}, a)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\langle y, y'_{n+k} \rangle - \langle y, y'_{n+k-1} \rangle = \varphi(x_{n+k-1}, a) - \varphi(x_{n+k}, a)$$

d'où un et un seul vecteur a et, puisque $y \in V_\varphi \{ y'_n \}$:

$$\langle y, y'_\infty \rangle = \langle y, y'_p \rangle + \varphi(x_p, a) \qquad \forall\, p \in \mathbf{N}$$

Si $y \notin V_\varphi \{ y'_n \}$ la solution a dépendra de $x_n, ..., x_{n+k}$. Il est donc important de savoir si $\exists z$:

$$\langle z, y'_p \rangle = \langle y, y'_p \rangle + \varphi(x_p, a) \qquad \forall\, p \in \mathbf{N}$$

et si $\langle z, y'_p \rangle$ converge vers $\langle y, y'_\infty \rangle$ et cela plus vite que $\langle y, y'_p \rangle$, c'est-à-dire si :

$$\lim_{p \to \infty} \frac{\langle z, y'_p - y'_\infty \rangle}{\langle y, y'_p - y'_\infty \rangle} = 0.$$

Considérons maintenant la fonction φ suivante :

$$\varphi(n, a) = \sum_{i=1}^{p} A_i(n)\, e^{a_i n} + \sum_{i=p+1}^{q} [B_i(n) \cos b_i n + C_i(n) \sin b_i n]\, e^{a_i n} + \sum_{i=1}^{m} r_i \delta_{in}$$

$n = 1, 2, ...$ où A_i, B_i et C_i sont des polynômes en n tels que si d_i est égal au degré de A_i plus un pour $i = 1, ..., p$ et au plus grand des degrés de B_i et de C_i plus un pour $i = p + 1, ..., q$ on ait :

$$m + \sum_{i=1}^{p} d_i + 2 \sum_{i=p+1}^{q} d_i = k$$

Le vecteur a est formé de toutes les inconnues a_i, b_i, r_i ainsi que des coefficients des polynômes A_i, B_i et C_i. On sait que [2] si $y \in V_\varphi \{ y'_n \}$, $\exists\, c_1, ..., c_k$ tels que :

$$\varphi(n + k, a) + c_1\, \varphi(n + k - 1, a) + \ldots\ldots + c_k\, \varphi(n, a) = 0$$

Si l'on pose $\langle y, y'_n \rangle = S_n$ on a :

$$\langle y, y'_\infty \rangle = \frac{H_{k+1}^{(n)}(S_n)}{H_k^{(n)}(\Delta^2 S_n)} \qquad (1)$$

où $H_k^{(n)}(S_n)$ est le déterminant de Hankel correspondant à $\{ S_n \}$ et défini par :

$$H_k^{(n)}(S_n) = \begin{vmatrix} S_n & \Delta S_n \ldots & \Delta^{k-1} S_n \\ \ldots & \ldots & \ldots \\ \Delta^{k-1} S_n & \Delta^k S_n \ldots & \Delta^{2k-2} S_n \end{vmatrix} \qquad \text{avec} \qquad \Delta S_n = S_{n+1} - S_n$$

si $y \notin V_\varphi \{ y'_n \}$ alors (1) ne sera plus égal à $\langle y, y'_\infty \rangle$ mais à une quantité $e_k(S_n)$. Cette méthode est le procédé Δ^2 d'Aitken généralisé. On voit que le calcul de $e_k(S_n)$ à l'aide de (1) nécessite l'évaluation de déterminants dont l'ordre peut être élevé. On évite ce calcul en utilisant l'ε-algorithme de Wynn [3] :

$$\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{1}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}} \qquad \text{avec} \qquad \varepsilon_{-1}^{(n)} = 0 \text{ et } \varepsilon_0^{(n)} = S_n$$

On montre que $\varepsilon_{2k}^{(n)} = e_k(S_n)$.

Donnons maintenant une interprétation de l'ε-algorithme différente de celle que nous venons de voir. Soit H l'espace des suites infinies convergentes de nombres réels telles que $H_k^{(n)}(\Delta^2 S_n) \neq 0 \; \forall \; k, n$. Soit F_k l'application affine de H dans S espace des suites, définie par :

$$F_k : \{ u_n \} \in H \to \left\{ \sum_{i=0}^{k} \alpha_i u_{n+i} - c \right\} \in S$$

On a

$$\mathrm{Ker}\,(F_k) = \left\{ u = \{ u_n \} \;\middle|\; u_n = \frac{c}{\sum_{i=0}^{k} \alpha_i} + \varphi(n, a) \right\}$$

où $\varphi(n, a)$ est la fonction correspondant à l'ε-algorithme.

Soit G_k le sous-espace de H des suites finies de $2k + 1$ termes et soit E_k l'application qui à $x = (S_n, \ldots, S_{n+2k}) \in G_k$ fait correspondre

$$\varepsilon_{2k}^{(n)} = H_{k+1}^{(n)}(S_n)/H_k^{(n)}(\Delta^2 S_n) \in \mathbf{R} \text{ (ou } \mathbf{C}).$$

Étant donné $x \in G_k$ il est facile de montrer que l'ε-algorithme revient à chercher les coefficients $\alpha_0, \ldots, \alpha_k$ de l'équation aux différences vérifiée par la suite unique $\{ u_n \} \in \mathrm{Ker}\,(F_k)$ telle que $u_i = S_i$ pour $i = n, \ldots, n + 2k$ puis à calculer $S = c \Big/ \sum_{i=0}^{k} \alpha_i = \varepsilon_{2k}^{(n)}$. En effectuant cette démonstration on remarque que S est indépendant de c et que la condition $H_k^{(n)}(\Delta^2 S_n) \neq 0$ est identique à $\sum_{i=0}^{k} \alpha_i \neq 0$.

D'autre part on sait que $H_{k+1}^{(n)}(S_n)$ et $H_k^{(n)}(\Delta^2 S_n)$ sont des applications continues et différentiables par rapport à $x = (S_n, ..., S_{n+2k})$, donc, puisque $x \in G_k$, l'application E_k est continue et différentiable dans G_k. D'où le résultat fondamental suivant :

Théorème 1 : l'ε-algorithme est continu et différentiable et l'on a

$$\frac{\partial \varepsilon_{2k}^{(n)}}{\partial S_j} = \frac{\partial}{\partial S_j} \cdot \frac{H_{k+1}^{(n)}(S_n)}{H_k^{(n)}(\Delta^2 S_n)} \qquad \text{pour} \qquad j = n, ... n + 2k$$

LES PROCEDES DE SOMMATION

Soient E et F deux espaces de Banach, E' et F' leurs duals topologiques, $\{ e'_n \}$ et $\{ f'_n \}$ des suites de E' et F' qui convergent faiblement vers e'_∞ et f'_∞ et $\{ e_n \}$ un système total de E. Soit T un opérateur linéaire continu de E dans F et T' son dual. Soit enfin $\{ x'_n \}$ la suite d'éléments de E' définie par $x'_n = T' f'_n \, \forall \, n$. On appelle procédé de sommation le procédé qui consiste à remplacer $x \in E$ par $Tx \in F$. On veut évidemment que le procédé de sommation soit régulier c'est-à-dire que $\langle Tx, f'_\infty \rangle = \langle x, e'_\infty \rangle = \langle x, x'_\infty \rangle$. Le problème des conditions à imposer à T pour que ce soit un procédé régulier de sommation peut être remplacé par celui des conditions à imposer à $\{ x'_n \}$ pour que ce soit une suite faiblement convergente et pour que $\langle x, x'_\infty \rangle = \langle x, e'_\infty \rangle$. La réponse à la première question est fournie par le théorème de Banach-Steinhaus tandis que la réponse à la seconde question est donnée par le principe de prolongement des identités.

Théorème 2 : Une condition nécessaire et suffisante pour que $\{ x'_n \}$ converge faiblement est que :

$$\|x'_n\| < M \qquad \forall \, n$$

$$\langle e_k, x'_n \rangle \qquad \text{converge} \; \forall \, k$$

Théorème 3 : Supposons que $\{ x'_n \}$ soit faiblement convergente. Si $\langle x, x'_\infty \rangle = \langle x, e'_\infty \rangle \, \forall \, x \in D$ partout dense dans E alors

$$\langle x, x'_\infty \rangle = \langle x, e'_\infty \rangle \, \forall \, x \in E.$$

Nous allons maintenant particulariser à un espace de suites et retrouver les conditions de Tœplitz pour que le procédé soit régulier. Soit c l'espace des suites convergentes de nombres réels. Muni de la norme du Sup c'est un espace de Banach. Son dual topologique est l^1. Le sous-espace D engendré par le système $e_0 = (1, 1, ...)$, $e_1 = (1, 0, 0, ...)$, $e_2 = (0, 1, 0, ...)$, etc..., est partout dense dans c. Prenons $E = F = c$ et $e'_n = f'_n =$ forme linéaire continue qui à une suite de c associe son $n^{\text{ième}}$ terme. L'opérateur T est défini par une matrice infinie $A = (a_{ij})$ et T' par $A' = (a_{ji})$ transposée de A.

Les conditions du théorème 2 s'écrivent :

$$\|x'_n\| = \|T'f'_n\| = \sum_k |a_{nk}| < M \qquad \forall\, n$$

$$\langle e_k, x'_\infty \rangle = \langle Te_k, f'_\infty \rangle = \lim_{n\to\infty} a_{nk} = b_k \qquad \forall\, k > 0$$

$$\langle e_0, x'_\infty \rangle = \lim_{n\to\infty} \sum_k a_{nk} = b_0$$

Si $b_k = 0\ \forall\, k > 0$ et $b_0 = 1$ alors les conditions du théorème 3 sont remplies car $\langle e_k, e'_\infty \rangle = 0\ \forall\, k > 0$ et $\langle e_0, e'_\infty \rangle = 1$; d'où le théorème de Tœplitz :

Théorème 4 : Une condition nécessaire et suffisante pour que la matrice $A = (a_{ij})$ définisse un procédé régulier de sommation est que :

$$\sum_k |a_{nk}| < M \qquad \forall\, n$$

$$\lim_{n\to\infty} a_{nk} = 0 \qquad \forall\, k$$

$$\lim_{n\to\infty} \sum_k a_{nk} = 1$$

Étudions maintenant une classe particulière de procédés de sommation : les procédés totaux. On dit que T est une transformation totale si elle est régulière et si $\langle x, x'_n \rangle = \langle x, x'_\infty \rangle\ \forall\, n$ entraîne $\langle Tx, x'_n \rangle = \langle x, x'_\infty \rangle\ \forall\, n$.

Soit un procédé total de sommation défini par une matrice $A = (a_{ij})$. Nous supposerons que cette matrice est telle que $\forall\, i\, \exists\, k(i) : \forall\, j > k(i)\ a_{ij} = 0$. Soit $\{x_n\}$ une suite de c : $\lim_{n\to\infty} x_n = S$ et soit $\{y_n\} \in c$ la suite obtenue en appliquant le procédé de sommation. On a :

$$y_i = \sum_{j=1}^{k(i)} a_{ij} x_j \qquad \forall\, i$$

Pour i fixé, y_i sera égal à S si l'équation aux différences $S = \sum_{j=1}^{k(i)} a_{ij} x_{j+n}$ est vérifiée. La solution générale de cette équation aux différences est $x_n = S + \varphi(n, a)$ où φ est la fonction correspondant à l'ε-algorithme avec $k = k(i) - 1$ et où le polynôme $\sum_{j=1}^{k(i)} a_{ij} t^{j-1}$ admet les p racines distinctes $\overline{e^{a_i}}$ pour $i = 1, ..., p$, les $2(q-p)$ racines distinctes e^{a_i} $(\cos b_i \pm i \sin b_i)$ pour $i = p+1, ..., q$ et m racines nulles.

Les conséquences de ce résultat sont importantes. On remarque d'abord que les a_i et les b_i de φ sont fixés et spécifiques du procédé de sommation utilisé (il en est de même des coefficients de l'équation aux différences alors que pour l'ε-algorithme ceux-ci sont des inconnues). Les seules inconnues sont les coefficients des polynômes A_i, B_i et C_i, les r_i et S. Ce qui explique pourquoi les procédés dits de sommation employés pour trouver S sont linéaires, tandis que l'ε-algorithme, qui doit également déterminer les a_i et les b_i, est non linéaire. On voit également que pour trouver S il faudra $k(i)$ termes par un procédé de sommation alors que l'ε-algorithme en nécessitera $2k(i) - 1$. Mais la conclusion la plus importante est la suivante : tout procédé total de sommation est un cas particulier de l'ε-algorithme.

Par conséquent tout procédé total de sommation convergera moins bien que l'ε-algorithme. Enfin ceci montre que le procédé de Romberg est un procédé d'extrapolation sans qu'il y ait pour cela besoin de le considérer comme provenant de la méthode d'extrapolation polynomiale de Richardson.

Comme illustration de ceci nous allons montrer que le procédé de Baranger [1] est un cas particulier du Δ^2 d'Aitken habituel. Baranger a étudié les p_i et les A_i ($i = 0, \ldots, n$) qui rendent la quantité $\sum_{i=0}^{\infty} u_i - \sum_{i=0}^{n} A_i u_{p_i}$ optimale en un certain sens pour un sous-ensemble déterminé de l'ensemble des séries absolument convergentes. Soit H_a l'espace des suites :

$$H_a = \left\{ u = \{ u_n \} \,\middle|\, \sum_0^{\infty} |u_n| < \infty \quad \text{et} \quad \sum_0^{\infty} \frac{(\Delta u_n)^2}{a^n} < \infty \right\}$$

avec $0 < a < 1$.

Baranger a montré que pour $p_i = i$ le meilleur choix des A_i était $A_0 = A_1 = \ldots = A_{n-1} = 1$ et $A_n = \dfrac{1}{1 - a}$.

Posons $V_n = \sum_{i=0}^{n} A_i u_i$ et $S_n = \sum_{i=0}^{n} u_i$. On vérifie immédiatement que :

$$V_n = A_n S_n + (1 - A_n) S_{n-1} \tag{2}$$

et que, par conséquent, cette transformation est totale. On a donc $\lim_{n \to \infty} V_n = \sum_{i=0}^{\infty} u_i$. Supposons maintenant que $0 < a_1 < a_2 < 1$. On a $H_{a_1} \subset H_{a_2}$.

En effet $\dfrac{1}{a_1^n} > \dfrac{1}{a_2^n}$ et $\sum_0^{\infty} \dfrac{(\Delta u_n)^2}{a_2^n} < \sum_0 \dfrac{(\Delta u_n)^2}{a_1^n}$. Donc si $u = \{ u_n \} \in H_{a_1}$ alors $u \in H_{a_2}$ car la série est absolument convergente. D'autre part si $u \in H_{a_1}$ et si $\forall\, \varepsilon > 0$ $u \notin H_{a_1 - \varepsilon}$ alors $u \in H_a$ $\forall\, a \in [a_1, 1[$.

Peut-être existe-t-il une valeur de a optimale dans $[a_1, 1[$?

C'est cette valeur que nous allons maintenant caractériser. Nous dirons que $\{ V_n \}$ converge plus vite que $\{ S_n \}$ si $\lim\limits_{n\to\infty} \Delta V_n/\Delta S_n = 0$ d'où, en remplaçant V_n par sa valeur et A_n par $\dfrac{1}{1-a}$: $\lim\limits_{n\to\infty} \dfrac{\Delta S_{n-1}}{\Delta S_n} = \dfrac{1}{a}$.

C'est la condition nécessaire et suffisante pour que $\{ V_n \}$ converge plus vite que $\{ S_n \}$, ce qui nous donne comme valeur optimale pour a :

$$a_{\text{opt}} = \lim_{n\to\infty} \frac{\Delta S_n}{\Delta S_{n-1}} = \lim_{n\to\infty} \frac{u_{n+1}}{u_n}$$

On voit que l'étude de (2) en tant que procédé de sommation n'impose aucune contrainte sur a. Si $S_n = S + \alpha\lambda^n$ on trouve $a_{\text{opt}} = \lambda$ et en prenant $A_n = \dfrac{1}{1-\lambda}$ on obtient $V_n = S$. On peut se rendre compte encore mieux que cette méthode est un cas particulier du Δ^2 d'Aitken en prenant, au lieu de $a = a_{\text{opt}}$, la valeur $a = \dfrac{\Delta S_n}{\Delta S_{n-1}}$. On trouve alors $V_n = \dfrac{S_{n-1} \cdot S_{n+1} - S_n^2}{\Delta^2 S_{n-1}} = \varepsilon_2^{(n-1)}$ qui n'est autre que le procédé Δ^2 habituel.

BIBLIOGRAPHIE

[1] J. BARANGER, *Une formule optimale pour le calcul de la somme d'une série*. Colloque d'Analyse numérique du CNRS, Super Besse, juin, 1970.

[2] C. BREZINSKI et M. CROUZEIX, *Remarques sur le procédé Δ^2 d'Aitken*, CRAS, série A, t. 270, pp. 896-898, 6 avril 1970.

[3] P. WYNN, *On a device for computing the $e_m(S_n)$ transformation*, MTAC, vol. 10, n° 54, 1956, pp. 91-96.