Accélération de la convergence et algorithme proximal

Hervé Le Ferrand *

1995

1 Introduction

Soit C un ensemble convexe fermé, borné non vide d'un espace de Hilbert H dont la norme est notée $\| \ \|$. Dans [5], Martinet a considéré sur de tels ensembles des applications **pseudo-contractantes** :

Définition 1.1 *L'application U de C dans C est appelée* **pseudo-contractante** *si on a*

$$\|U(x) - U(y)\|^2 \leq \|x - y\|^2 - \|(I - U)(x) - (I - U)(y)\|^2, \tag{1}$$

pour tout couple (x, y) de vecteurs de H.

On construit alors la suite des itérés :

$$x_0 \in C, \ x_{n+1} = U(x_n). \tag{2}$$

Martinet a démontré le résultat fondamental :

Théorème 1.1 *Si U est pseudo-contractante de C dans C, la suite $(x_n)_{n \geq 0}$ converge faiblement vers un point fixe de U.*

En dimension finie, la convergence est bien entendu forte. Un exemple important d'application pseudo-contractante est l'application **proximale** selon la terminologie de Moreau. Cette application est associée à une application f convexe, fermée, propre. Plus précisèment, on a deux définitions équivalentes :

Définition 1.2 *L'application $prox_{\lambda f}$ $(\lambda > 0)$ associe à tout élément x de H l'unique vecteur y tel que*

$$f(y) + \frac{1}{2\lambda}\|y - x\|^2 = \min_{u \in H}\left(f(u) + \frac{1}{2\lambda}\|u - x\|^2\right). \tag{3}$$

[1]Université de Bourgogne, Département de Mathématiques
Laboratoire "Analyse Appliquée et Optimisation",
BP 138, 21004 Dijon cedex, France

Définition 1.3 *L'application $prox_{\lambda f}$ ($\lambda > 0$) est définie par*

$$prox_{\lambda f}(x) = (I + \lambda \partial f)^{-1}(x) \tag{4}$$

où ∂f est le sous différentiel de f.

Toujours dans [5], Martinet déduit du théorème précèdent la convergence (faible si l'espace est de dimension infinie) de la suite générée par

$$x_{n+1} = prox_{\lambda f}(x_n),\ x_0 \in H. \tag{5}$$

Beaucoup d'autres auteurs ont étudié cet algorithme et de nouvelles versions en ont été données (version approchée, méthode de l'inverse partielle etc...). Pour avoir une synthèse de tous ces résultats, on pourra consulter l'article de B. Lemaire [4]. On travailles à présent dans l'espace $H = I\!R^N$.

Concernant la vitesse de convergence, il y a des résultat pour des formes approchées de l'algorithme proximal. Souvetn on a au mieux une convergence linéaire (en norme) ou dans certains cas très précis, convergence finie (voir [4]). Dans l'optique d'accélérer la convergence de telles suites , tout ceci n'est pas très encourageant. En effet, si il existe de nombreuses méthodes d'accélération de convergence [1], celles-ci s'appliquent de façon précise à des classes de suites bien déterminées. Ainsi, par exemple, si en dimension 1 les suites à convergence linéaire avec un rapport strictement inférieur à 1 sont facilement accélérables, en dimension supérieure à 1 cette information n'est pas suffisante pour dire si telle ou telle méthode d'accélération peut être utilisée avec succès. Il faudrait par exemple une information du type convergence A-linéaire (i.e on a $\lim_{n\to\infty} \frac{(x_{n+1}-x)-A(x_n-x)}{\|x_n-x\|} = 0$ où A est une matrice). En résumé, il y a deux problèmes majeurs pour l'accélération de convergence :

- en dimension 1, par exemple, si il y a convergence linéaire avec un rapport égal à 1 ;
- en dimension $N > 1$, un manque d'information sur la façon dont la suite converge.

Dans la seconde partie de cette article nous considérerons des suites scalaires générées par des applications pseudo-contractantes. Dans une troisième partie, nous nous placerons en dimension $N > 1$ et regarderons une suite issue de la méthode de l'inverse partiel.

2 Le cas unidimensionnel

2.1 Application pseudo-contractante sur un intervalle fermé borné

Soit I un intervalle fermé borné de $I\!R$ et $f : I \longrightarrow I$ une application pseudo-contractante i.e :

$$(\star)\ \forall (x,y) \in I \times I\ |f(x) - f(y)|^2 \leq |x-y|^2 - |(x - f(x)) - (y - f(y))|^2. \tag{6}$$

On a facilement :

Proposition 2.1 ($\star$) *équivaut à :*

$$\forall (x,y) \in I \times I,\ x \neq y,\ 0 \leq \frac{f(x) - f(y)}{x - y} \leq 1. \tag{7}$$

Preuve
Si $x \neq y$, on a :

$$\left|\frac{f(x)-f(y)}{x-y}\right|^2 \leq 1 - \frac{(x-y-(f(x)-f(y))^2}{(x-y)^2} \tag{8}$$

soit

$$\left(\frac{f(x)-f(y)}{x-y}\right)^2 \leq 1 - \left(1 - \frac{f(x)-f(y)}{x-y}\right)^2. \tag{9}$$

Or $a^2 \leq 1-(1-a)^2$ équivaut à $a^2 - a \leq 0$, i.e $a \in [0,1]$. On a le résultat cherché .

Remarque 2.1 *Si f est en fait strictement croissante sur I, elle réalise une bijection de I sur $f(I)$. Posons $g = f^{-1} - I$ (g est continue sur $f(I)$). La fonction g est croissante sur $f(I)$, en effet :*

$$\frac{g(x)-g(y)}{x-y} = \frac{f^{-1}(x)-x-f^{-1}(y)+y}{x-y} = \frac{f^{-1}(x)-f^{-1}(y)}{x-y} - 1. \tag{10}$$

Or $\frac{f^{-1}(x)-f^{-1}(y)}{x-y} \geq 1$, d'où le résultat. En conclusion, f est de la forme $(I+g)^{-1}$ avec g croissante et continue sur $f(I)$.

2.2 Accélération de convergence

La suite générée par $x_{n+1} = f(x_n)$ où f est pseudo-contractante vérifie donc :

i) $(x_n)_{n\geq 0}$ converge

ii) $\forall n,\ 0 \leq \frac{x_{n+1}-x}{x_n-x} \leq 1$ où $x = \lim_{n\to\infty} x_n$.

D'après les travaux de Delahaye [2], il n'existe pas de transformation accélérant l'ensemble des suites (y_n) vérifiant i) et ii). En général, on ne peut donc pas accélérer les suites récurrentes générées par une application pseudo-contractante. On peut cependant faire les remarques suivantes :

Remarque 2.2 *Si on pose $r_n = \frac{x_{n+1}-x}{x_n-x}$, il existe une sous-suite $(r_{\varphi(n)})$ convergente, i.e :*

$$\lim_{n\to\infty} \frac{x_{\varphi(n)}-x}{x_{\varphi(n)}-x} = \rho \in [0,1]. \tag{11}$$

Comme on a :

$$0 \leq \frac{x_{\varphi(n+1)}-x}{x_{\varphi(n)}-x} = \frac{x_{\varphi(n+1)}-x}{x_{\varphi(n+1)-1}-x} \times \frac{x_{\varphi(n+1)-1}-x}{x_{\varphi(n+1)-2}-x} \times \cdots \times \frac{x_{\varphi(n)+1}-x}{x_{\varphi(n)}-x} \tag{12}$$

il vient

$$0 \leq \frac{x_{\varphi(n+1)}-x}{x_{\varphi(n)}-x} \leq \frac{x_{\varphi(n)+1}-x}{x_{\varphi(n)}-x}. \tag{13}$$

Si $\rho \in [0,1[$, il existe $\rho < \rho' < 1$ tel que pour n assez grand

$$\frac{x_{\varphi(n)+1}-x}{x_{\varphi(n)}-x} \leq \rho'. \tag{14}$$

On peut améliorer la convergence de la suite $(x_{\varphi(n)})$ (voir "contractive sequence transformations" dans [1] p 201-210).

Remarque 2.3 *Même dans le cas de la convergence linéaire, l'accélération de convergence peut être difficile à réaliser. C'est le cas si le rapport est égal à 1. On dit alors que la suite est logarithmique. Les résultats d'accélération connus portent sur des sous-classes bien précises (voir [7][8]).*

Remarque 2.4 *Si l'on connait de nombreuses méthodes d'accélération de convergence pour certaines classes de suites, dans la pratique quand on considère une suite on ne sait pas à priori dans quelle classe elle se situe ! Pour cette raison, Delahaye a proposé un procédé de sélection automatique ([2] p 255).*

3 Accélération en dimension $N > 1$

3.1 Généralités

Comme nous l'avons signalé dans l'introduction, l'accélération de suites vectorielles est un problème très délicat. Par exemple, contrairement au cas scalaire, si la convergence linéaire pour une suite vectorielle est une information intéressante en soi, elle est peu utile pour décider si telle ou telle méthode d'accélération peut être efficace. Il faut des informations plus précises, disons de type algèbrique (comme par exemple le comportement A-linéaire).
Une autre difficulté est qu'un procédé d'accélération vectoriel uniformise, dans le sens que ce sont les composantes les plus lentes qui sont accélérées. Or on peut trés bien souhaiter des résultats sur les composantes les plus rapides. Nous allons aborder cette situation dans le paragraphe suivant.

3.2 La méthode de l'inverse partiel et le problème de Fermat-Weber

Dans [3], on trouve un algorithme fondé sur la méthode de l'inverse partiel de Spingarn (voir [4] pour une synthèse et des références) pour résoudre un problème de Fermat-Weber. Sans entrer dans les détails, on travaille dans un espace vectoriel euclidien H ($\mathbb{R}^n$ dupliqué plusieurs fois) qui est somme directe de deux sous-espaces orthogonaux A et B. L'algorithme produit deux suites $(x_k) \in A^{\mathbb{N}}$ et $(p_k) \in B^{\mathbb{N}}$ telles que :

a) (x_k) converge vers x solution du problème primal de Fermat-Weber

b) (p_k) converge vers p solution du problème dual de Fermat-Weber

c) $\|(x_{k+1} + p_{k+1}) - (x + p)\| \leq \|(x_k + p_k) - (x + p)\|$.

On a donc une information sur la vitesse de convergence de la suite somme $u_k = x_k + p_k$. Supposons qu'il existe une transformation T transformant (u_k) en une suite (U_k) telle que :

i) $\lim_{k\to\infty} U_k = u = x + p$

ii) $\lim_{k\to\infty} \frac{\|U_k - u\|}{\|u_k - u\|} = 0$.

Notons P_A (respectivement P_B) la projection sur A parallèlement à B (respectivement la projection sur B parallèlement à A). On a $\lim_{k\to\infty} P_A(U_k) = x$ et $\lim_{k\to\infty} P_B(U_k) = p$.
A-t-on

$$\lim_{k\to\infty} \frac{\|P_A(U_k) - x\|}{\|x_k - x\|} = 0 \ ? \tag{15}$$

Comme on s'y attend, la réponse va dépendre de la comparaison des convergences des deux suites. De manière élémentaire, de $\lim_{k\to\infty} \frac{U_k - u}{\|u_k - u\|} = 0$, il vient :

$$\lim_{k\to\infty} P_A\left(\frac{U_k - u}{\|u_k - u\|}\right) = 0. \tag{16}$$

De plus $\|u_k - u\| = \left(\|x_k - x\|^2 + \|p_k - p\|^2\right)^{\frac{1}{2}}$, d'où

$$P_A\left(\frac{U_k - u}{\|u_k - u\|}\right) = \frac{P_A(U_k) - x}{\|x_k - x\| \left(1 + \frac{\|p_k - p\|^2}{\|x_k - x\|^2}\right)^{\frac{1}{2}}}. \tag{17}$$

Si il existe une constante $M > 0$ telle que $\frac{\|p_k - p\|}{\|x_k - x\|} < M$ à partir d'un certain rang, on a alors

$$\lim_{k\to\infty} \frac{P_A(U_k) - x}{\|x_k - x\|} = 0. \tag{18}$$

Pour conclure, on peut formuler la

Proposition 3.1 *Si il existe $M_1 > 0$ et $M_2 > 0$ tels que pour tout $k > k_0$ on ait*

$$M_1 \leq \frac{\|p_k - p\|}{\|x_k - x\|} \leq M_2$$

alors

$$\lim_{k\to\infty} \frac{P_A(U_k) - x}{\|x_k - x\|} = \lim_{k\to\infty} \frac{P_B(U_k) - p}{\|p_k - p\|} = 0.$$

4 Conclusion

On a vu que d'un point de vue général, il n'est pas possible d'accélérer les suites récurrentes générées par une application pseudo-contractante. Pour des problèmes particuliers, il faut des informations très précises sur la nature de la convergence de la suite. On propose d'utiliser la procédure de sélection automatique de Delahaye. Il reste donc à faire des expérimentations numériques.

5 Références

1. C. Brezinski et M. Redivo-Zaglia, Extrapolation Methods, Theory and Practise, North-Holland, Amsterdam, 1992.

2. J.P. Delahaye, Sequence transformations, Springer Verlag, Berlin, 1988.

3. O. Lefebvre et C. Michelot, A primal-dual algorithm for the Fermat-Weber problem involving mixed gauges, Math. Program. 39 (1987) 319-335.

4. B. Lemaire, The proximal algorithm, proceedings of “New methods of optimization and their industrial use ” (1987, Pau, France) in International Series of Numerical Mathematics, Birkhäuser Verlag.

5. B. Martinet, Détermination approchée d’un point fixe d’une application pseudo-contractante. Cas de l’application *prox.*, C.R. Acad. Sc. Paris, t.274 Série A, p 163-165 (10 Janvier 1972).

6. J.J Moreau, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France, 93, p 273-299, 1965.

7. N. Osada, A convergence acceleration method for some logarithmically convergent sequences. SIAM J. Numer. Anal., 27, p 178-189, 199O.

8. P. Sablonnière, Convergence acceleration of logarithmic fixed point sequences. J. Comput. Appl. Math., 19, p 55-60, 1987.