Accélération de convergence pour les suites vectorielles de point fixe

Isabelle Ramière

Auteur : Isabelle RAMIÈRE, ingénieur-chercheur en méthodes numériques, Département d'Études des Combustibles, Cadarache.

Surtitre : Analyse Numérique.

Titre : Accélération de convergence pour les suites vectorielles de point fixe.

Résumé

Les problèmes de type point fixe sont fréquemment rencontrés en simulation numérique. La résolution par itérations de point fixe a l'avantage majeur d'être simple et facilement implémentable mais la convergence, quand elle est obtenue, est souvent seulement linéaire et lente. Les méthodes d'accélération de convergence (aussi appelées méthodes d'extrapolation) permettent alors d'améliorer la vitesse de convergence et parfois aussi d'éviter la divergence.

Texte

1 Suite vectorielle de point fixe

Les problèmes de point fixe s'écrivent sous la forme générale

$$X = F(X), \quad \text{où } X \text{ est un vecteur de réels de dimension } d \ (X \in \mathbb{R}^d) \tag{1}$$

On les rencontre dans un nombre varié de problèmes numériques (couplage multiphysique, décomposition de domaine,...) mais également lors de la résolution de problèmes non-linéaires (structure électronique, transfert de chaleur, mécanique non-linéaire, ...). En effet, il y a une dualité naturelle entre les problèmes de point fixe et les problèmes non-linéaires puisque tout problème non-linéaire de la forme $G(X) = 0$ est équivalent à un problème de point fixe de type (1) (un choix trivial, qui n'est souvent pas le meilleur, est de prendre $F(X) = G(X) + X$).

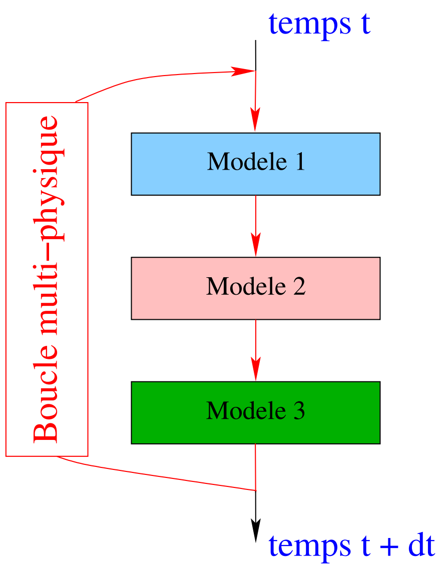

FIGURE 1 – Exemple de problème de point fixe arrivant lors de la résolution d'un couplage multi-physique

À titre d'exemple, le couplage multi-physique de type point fixe entre les modèles 1, 2 et 3, illustré sur la figure 1, s'écrit :

$$\begin{cases} X_1 & = & F_1(X_1, X_2, X_3) \\ X_2 & = & F_2(X_1, X_2, X_3) \\ X_3 & = & F_3(X_1, X_2, X_3) \end{cases}$$

où F_i représente le solveur du modèle i et X_i la (ou les) variable(s) d'intérêt de ce modèle. De façon générale, $X_i \in \mathbb{R}^{d_i}$ (avec $\sum_i d_i = d$). On retrouve bien la forme générale d'un problème de point fixe $X = F(X)$ en posant $X = (X_1, X_2, X_3)$ et $F = (F_1, F_2, F_3)$.

La méthode de résolution naturelle d'un problème de point fixe est de construire une suite de point fixe générée par les itérations suivantes (parfois appelées itérations de Picard) à partir d'un X^0

$$X^{n+1} = F(X^n), \quad \text{où } n \text{ désigne le numéro d'itération} \tag{2}$$

jusqu'à convergence de la suite, détectée numériquement lorsque $\|F(X^n) - X^n\| \leq \epsilon$ avec ϵ un réel positif petit et $\|.\|$ une norme donnée.
La construction des itérations de point fixe a l'avantage majeur d'être simple et facilement mise en

oeuvre informatiquement. En particulier, elle ne nécessite pas la construction de la dérivée de F (ou Jacobienne de F dans le cas vectoriel), ce qui la rend très attrayante vis-à-vis des méthodes de type Newton très connues pour la résolution de problèmes non-linéaires (mais souvent très coûteuses et difficiles à mettre en oeuvre de façon générale). En contre-partie, la convergence de la suite de point fixe vers la solution du problème (1) n'est a priori pas garantie (car elle s'appuie sur des propriétés de F autour du point solution, a priori inconnu !) et souvent linéaire (erreur qui décroît de façon linéaire) avec une vitesse de convergence lente. C'est pourquoi les numériciens se sont intéressés à développer des méthodes d'accélération de convergence, méthodes qui consistent à générer une nouvelle suite qui converge plus vite vers la solution que la suite initiale. Certaines méthodes d'accélération permettent également d'augmenter l'ordre de convergence (super-linéaire, quadratique,...) et même d'atteindre la convergence quand la suite de point fixe diverge.

2 Accélération de convergence

L'accélération de la convergence de suites est un domaine à part entière de l'analyse numérique. La littérature sur ce sujet est vaste, on peut citer comme point d'entrée le livre de Brezinski et Zaglia [1]. Les méthodes d'accélération de convergence consistent de manière générale à définir les itérés de la suite accélérée comme une fonction des derniers itérés de la suite initiale :

$$Y^n = \psi(X^n, X^{n-1}, X^{n-2}, ...) \tag{3}$$

telle que $\lim_{n\to\infty} \frac{Y^n - X}{X^n - X} = 0$, où X est la limite de la suite $(X^n)_n$. Cette dernière équation permet de constater que la suite $(Y^n)_n$ converge plus vite vers la solution que la suite initiale $(X^n)_n$ (puisque le numérateur tend plus vite vers 0 que le dénominateur). En pratique, on utilise l'accélération de convergence de manière dynamique, c'est-à-dire que le nouvel itéré sera directement l'itéré accéléré $X^{n+1} = Y^n$.
Pour les suites scalaires ($d = 1$), la méthode d'accélération de suites la plus connue et la plus performante reste la méthode du Δ^2 (« delta carré ») d'Aitken [2]. Son adaptation en dynamique pour les suites de point fixe est la méthode de Steffensen [3] qui converge quadratiquement et permet une convergence même lorsque la suite de point fixe diverge. Cependant cette méthode nécessite deux évaluations successives de la fonction F par itération ce qui en réduit son efficacité. D'autres méthodes très connues avec un seul appel à F sont donc souvent utilisées, comme la méthode de la sécante et la méthode de relaxation (voir par exemple [4]).
Cependant, dans les cas pratiques d'utilisation, ce sont des suites vectorielles qui sont générées. On trouve aussi beaucoup de méthodes d'accélération de suites vectorielles dans la littérature. La plupart d'entre elles sont des extensions de méthodes d'accélération de suites scalaires. D'autres méthodes sont, au contraire, directement construites pour des suites vectorielles. Elles sont dans ces cas-là généralement basées sur un processus de minimisation.
Le but de la formulation proposée dans [5] par Département d'Etudes des Combustibles de la Direction de l'Energie Nucléaire du CEA, est d'introduire un cadre commun pour construire de nouvelles méthodes d'accélération de suites vectorielles mais également pour retrouver sous un même formalisme les méthodes d'accélération vectorielles les plus populaires et efficaces. D'un point de vue pratique, ce formalisme commun permet de tester facilement différentes méthodes d'accélération.

Le principe de la méthode est de générer une nouvelle suite $(Y^n)_n$ telle que

$$\begin{aligned} Y^n = X^n - \sum_{i=1}^{M} \lambda_i^n Z_i^n, \quad & \text{avec } 1 \leq M \leq d \text{ un entier,} \\ & \text{et } \forall i \in [\![1, ..., M]\!] : \ \lambda_i^n \in \mathbb{R} \text{ , } (Z_i^n) \in \mathbb{R}^d \text{ tel que } \lim_{n\to\infty} Z_i^n = 0. \end{aligned} \tag{4}$$

On voit immédiatement que lorsque les λ_i^n ne dépendent pas des itérations (donc de n), la suite $(Y^n)_n$ converge vers la même limite que la suite $(X^n)_n$. En pratique, comme les λ_i^n varient avec n, cette convergence vers la même limite ne peut être garantie pour toutes les suites $(Y^n)_n$ générées (la limite

des $(\lambda_i^n)_n$ ne doit pas tendre vers l'infini).
A partir de l'expression (4), les λ_i^n sont choisis de sorte à minimiser $\|Y^{n+1} - Y^n\|$ et ainsi garantir l'accélération de convergence vis-à-vis de la suite $(X^n)_n$. Pour la norme Euclidienne, l'expression des λ_i^n peut être obtenue en utilisant la solution de l'équation normale résultant d'une minimisation par moindres carrés.

Lors de l'application dynamique de ce procédé d'accélération de convergence aux suites de point fixe, nous disposons de deux suites : la suite des $(X^n)_n$ et celle des $(F(X^n))_n$. Il en découle deux classes de méthodes d'accélération vectorielles : les méthodes de suites croisées ou les méthodes de suites alternées. Les méthodes de suites croisées se basent sur la suite $(F(X^n))_n$ et utilisent la suite $(X^n)_n$ uniquement pour définir les résidus Z^n (typiquement les Z^n s'expriment en fonction des résidus de point fixe $F(X^n) - X^n$). Pour les méthodes de suites alternées, l'itéré Y^n s'appuie uniquement sur la suite $(X^n)_n$ et l'itéré Y^{n+1} uniquement sur la suite $(F(X^n))_n$ (même dans la définition des résidus). Afin de s'assurer de la convergence des suites générées vers la solution du problème de point fixe initial, le critère de convergence doit toujours porter sur le résidu du système de point fixe $\|F(X^n) - X^n\|$.
Avec ces deux classes de méthodes, on retrouve la majorité des méthodes d'accélération vectorielles connues (pour $M = 1$: méthodes de sécante multi-dimensionnelle [6], Irons & Tuck [7] (généralisation de Steffensen),..., pour $M \geq 1$: méthode d'Anderson [8], interface quasi-Newton [9],...) et on peut en construire une multitude d'autres. On a pu montrer que pour la résolution d'un problème non-linéaire, ce type de méthodes d'accélération vectorielles de point fixe à convergence linéaire concurrence la méthode de Newton (à convergence quadratique) en termes d'itérations sans avoir à évaluer la matrice Jacobienne (voir une illustration en figure 2). Elles constituent donc une stratégie de résolution très intéressante.

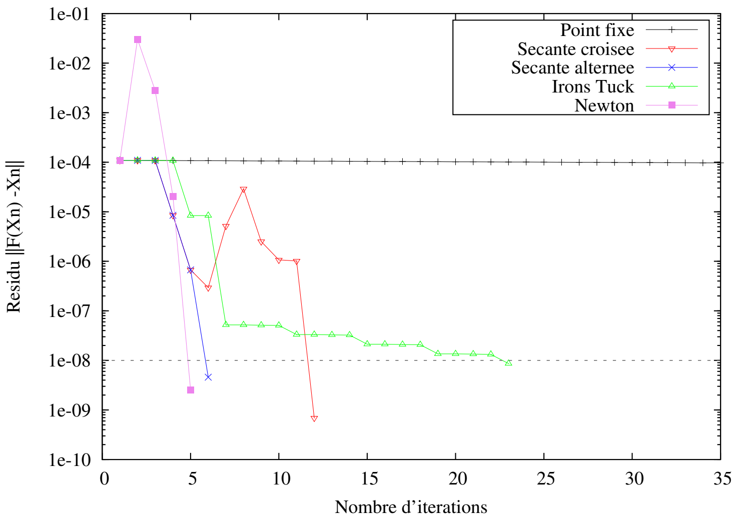

FIGURE 2 – Exemple de résolution d'un problème non-linéaire par différentes méthodes : Newton, Point fixe standard, Point fixe accéléré avec le formalisme proposé (Sécante croisée, Sécante alternée, Irons et Tuck). Seuil de convergence $\epsilon = 10^{-8}$.

On voit sur cet exemple la convergence très lente de la méthode de point fixe standard (autour de 1000 itérations). Les méthodes d'accélération de convergence de point fixe améliorent bien la vitesse de convergence. Les méthodes de type suites croisées semblent avoir un comportement moins monotone (toujours décroissant) que celles de type suites alternées. La méthode de Newton converge en 5 itérations (avec des premières itérations chaotiques) ce qui est seulement 1 itération de moins que la méthode de sécante alternée.

Différentes méthodes d'accélération tirées de ce formalisme général ont été depuis utilisées avec succès dans la littérature pour des applications diverses telles que le couplage multi-physique ou multi-champs, la résolution mécanique non-linéaire par transformée de Fourier rapide, la réduction de modèle pour des procédés de soudage, la distribution de population,...

Références

[1] C. Brezinski and M. Redivo Zaglia. *Extrapolation Methods : Theory and Practice.* North Holland, 1991.

[2] A. C. Aitken. On bernouilli's numerical solution of algebraic equations. *Proc. Roy. Soc. Edinburgh*, 46 :289–305, 1926.

[3] Y. Nievergelt. Aitken's and Steffensen's accelerations in several variables. *Numerische Mathematik*, 59(1) :295–310, 1991.

[4] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*. Springer, 2007.

[5] I. Ramière and Th. Helfer. Iterative residual-based vector methods to accelerate fixed point iterations. *Computers & Mathematics with Applications*, 70(9) :2210 – 2226, 2015.

[6] H. Fang and Y. Saad. Two classes of multisecant methods for nonlinear acceleration. *Numerical Linear Algebra with Applications*, 16 :197–221, 2009.

[7] B. M. Irons and R.C. Tuck. A version of the Aitken accelerator for computer iteration. *International journal of numerical methods in engineering*, 1 :275–277, 1969.

[8] D. G. Anderson. Iterative procedures for nonlinear integral equations. *Journal of the Association for Computing Machinery*, 12(4) :547–560, 1965.

[9] U. Küttler and W. A. Wall. Fixed-point fluid-structure interaction solvers with dynamic relaxation. *Computational Mechanics*, 43 :61–72, 2008.