

Université de Marseille
Licence de Mathématiques, 3ème année, analyse numérique et optimisation
Méthodes de Newton “dégradées”

Quelques précisions sur les possibilités de “dégrader” la méthode de Newton de manière à simplifier les calculs à chaque itération, sans trop perdre en vitesse convergence. . .

Rappel : la méthode de Newton pour calculer \bar{x} solution de $g(x) = 0$ s’écrit (pour $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$) :

Initialisation $x^{(0)} \in \mathbb{R}^n$.

Itérations $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)}), k \geq 0$.

Première idée : Ne pas calculer toute la jacobienne de g

C’est la méthode qui est décrite dans les notes de cours sous le nom “Newton incomplet”.

Pour mieux comprendre son intérêt, on la détaille ici dans l’exemple du système que vous avez à résoudre pour le projet (à rendre à la fin des cours).

Le système que vous avez à résoudre pour le projet s’écrit $Au + R(u) = b$ (autrement dit la fonction g est $g(u) = Au - R(u) - b$ pour $u \in \mathbb{R}^n$). La matrice A est appartient à $M_n(\mathbb{R})$. Elle est issue de la discrétisation de $-u''$, elle n’a que 3 diagonales non nulles. Par contre la fonction R est une fonction non linéaire de \mathbb{R}^n dans \mathbb{R}^n . Sa i -ème composante, notée R_i , dépend de toutes les composantes de u mais dépend de u_j surtout pour j “proche” de i (regarder la forme de R . . .).

Si on applique à ce système la méthode de Newton, la jacobienne de g sera à chaque itération une matrice “pleine”, c’est-à-dire que tous les termes de $J_g(u^{(k)})$ seront probablement non nuls.

Une idée intéressante consiste à ne mettre dans la Jacobienne de g que les termes $\partial_j R_i(u^{(k)})$ pour j “proche” de i .

Par exemple, en ne considérant dans la Jacobienne de g que les termes $\partial_j R_i(u^{(k)})$ pour $j = i + 1$ et $i - 1$ (c’est-à-dire en remplaçant $\partial_j R_i(u^{(k)})$ par 0 si $j \notin \{i - 1, i + 1\}$), la matrice $J_g(u^{(k)})$ a alors la même structure que A , c’est-à-dire qu’elle n’a que 3 diagonales non nulles. Ceci simplifie grandement à chaque itération de Newton la résolution du système linéaire. Bien sûr, l’inconvénient est que l’on perd la convergence quadratique.

Deuxième idée : Ne pas calculer la jacobienne de g

Il s’agit ici, dans les notes de cours, de la méthode de la sécante pour $n = 1$ et sa généralisation en méthode de “quasi-Newton” pour le cas $n > 1$.

L’idée consiste ici à ne jamais calculer la matrice jacobienne de g mais à l’approcher de mieux de mieux à chaque itération.

Pour $n = 1$ et $k > 1$, on remplace $g'(x^{(k)})$ par $\frac{g(x^{(k)}) - g(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}$ (qui est une approximation de $g'(x^{(k)})$). L’intérêt est de ne pas calculer g' (lorsque ce calcul est difficile). L’inconvénient est que l’on perd aussi ici la convergence quadratique (la convergence est d’ordre $\frac{1+\sqrt{5}}{2}$).

La même idée pour $n > 1$ suggère de remplacer $J_g(x^{(k)})$ par une matrice $B^{(k)}$ telle que (égalité (2.29) des notes de cours)

$$B^{(k)}(x^{(k)} - x^{(k-1)}) = g(x^{(k)}) - g(x^{(k-1)}).$$

Mais cette relation est insuffisante pour déterminer $B^{(k)}$. Pour compléter on demande à $B^{(k)}$ de rester “proche” de $B^{(k-1)}$. Un exemple classique est la méthode de Broyden (voir les notes de cours).

En pratique, ce choix de $B^{(k)}$ suggéré dans la méthode de Broyden ne semble pas très efficace. Mais pour les problèmes de minimisation de fonctions f de \mathbb{R}^n dans \mathbb{R} (minimiser f est lié à chercher un point où le gradient de f s’annule, donc à chercher $x \in \mathbb{R}^n$ tel que $g(x) = 0$ avec $g = \nabla f$), cette idée est à l’origine de méthodes très performantes (que nous verrons dans la suite du cours).