Optimisation non linéaire — Grille de lecture

Michel Bierlaire

Ce document a pour but de guider les étudiants pour l'utilisation du livre dans le cadre de l'auto-apprentissage, afin d'éviter de devoir tout lire. Les parties du livre à apprendre pour ce thème sont :

- Chapitre 10 (pp. 235–244),
- Section 11.1 (pp. 245–251),
- Section 11.3 (pp. 263–276),
- Section 11.4 (pp. 277–277),
- Section 11.5 (pp. 277-281),
- Section 11.6 (pp. 281–283).

Les chapitres 5 (conditions d'optimalité) et 7 (résolution de systèmes d'équations) sont des révisions d'analyse et d'algèbre linéaire, et donc censés être connus. Il est conseillé de les lire pour reviser les concepts, et se familiariser avec les notations.

Les étapes pour développer un algorithme pour l'optimisation non linéaire sans contrainte sont les suivantes.

1. Les conditions d'optimalité nous disent que, à un optimum local, le gradient est égal à zéro:
$$\nabla f(x) = 0.$$
Il s'agit d'un système de n équations non linéaires à n inconnues. Dès lors, la méthode de Newton pour résoudre ce type de système semble adéquate. Le chapitre 10 applique cette méthode dans le cadre de l'optimisation, et discute ses limitations. De plus, une interprétation géométrique est proposée (section 10.2). En résumé, quand elle fonctionne bien, elle fonctionne rapidement, Mais il est possible qu'elle ne fonctionne pas.

2. On investigue alors une autre idée : les méthodes de descente. Afin de trouver un minimum, l'idée est de suivre exclusivement des direction de descente (c-à-d telles que la dérivées directionnelle de la fonction objective $\nabla f(x)^\top d$ est négative, et d'identifier un itéré "meilleur" que le précédent. L'idée naturelle qui consiste à suivre la direction de la plus forte pente s'avère être inefficace. Un changement de variables, appelé "préconditionnement" permet d'accélérer les choses. Ceci est décrit à la section 11.1.

3. Une fois que la direction est choisie, il faut décider jusqu'où on la suit. L'idée qui consiste à prendre le point le plus bas dans cette direction (décrite à la section 11.2, qui n'est pas matière ce ce cours) est trop coûteuse en temps calcul. Nous analysons alors une méthode dite inexacte (section 11.3), qui est basée sur deux conditions : le nouvel itéré doit garantir une diminution suffisante de la fonction objectif, et un progrès suffisant de l'algorithme (interdit de faire des pas trop petits). Tout pas qui vérifie ces deux conditions peut alors être considéré.

4. La méthode dite de la plus forte de pente est décrite à la section 11.4. Cette méthode, notoirement lente, est décrite ici pour pouvoir la comparer avec d'autres.

5. Lorsque le changement de variable permettant de préconditionner la méthode de la plus forte pente est choisi d'une manière appropriée, on constate que, dans les bonnes conditions, cette méthode est équivalente avec la méthode de Newton. La Section 11.5 explique comment les différents ingrédients permettent de bénéficier de la vitesse de la méthode de Newton dans les bonnes circonstances, et de modifier celle-ci dans les autres circonstances afin qu'elle converge vers une solution du problème d'optimisation.

6. Enfin, la section 11.6 applique ces méthodes pour résoudre un problème à deux variables, afin d'illustrer leur comportement.