Optimisation convexe – Méthodes itératives

Méthode de Newton

Bashar Dudin

June 23, 2020

EPITA

Une méthode de second ordre

Les descentes de gradient qu'on a pu voir jusqu'à présent sont des méthodes de descentes dites de premier ordre ; elles minimisent (sur la boule unité pour une norme donnée) l'approximation au premier ordre de la fonction objectif.

Une méthode de second ordre

Les descentes de gradient qu'on a pu voir jusqu'à présent sont des méthodes de descentes dites de premier ordre ; elles minimisent (sur la boule unité pour une norme donnée) l'approximation au premier ordre de la fonction objectif.

La méthode de Newton est une méthode de second ordre ; on cherche à mininiser l'approximation de second ordre (en un point donnée) de la fonction objectif.

Une méthode de second ordre

Soit f une fonction objectif convexe et x un point du domaine de f. Le DL de f au second ordre en x s'écrit

$$f(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T H_f(x) v + \|v\|^2 \varepsilon(v)$$

Une méthode de second ordre

Soit f une fonction objectif convexe et x un point du domaine de f. Le DL de f au second ordre en x s'écrit

$$f(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T H_f(x) v + \|v\|^2 \varepsilon(v)$$

On choisit d'approcher $f(x+v)$ par l'expression de second ordre

$$f(x) + \nabla f(x)^T v + \frac{1}{2} v^T H_f(x) v$$

C'est une fonction convexe en v qu'on sait minimiser. On obtient ici un minimisant donné par

$$\Delta x_N = -\left(H_f(x)\right)^{-1} \nabla f(x).$$

Méthode de Newton – Condition d'arrêt

Traditionnellement, la condition d'arrêt de la méthode de Newton est décrite par le fait que le carré du coefficient de décroissance suivant tombe sous un certain seuil de tolérance.

$$\lambda(x) = \left(\nabla f(x)^T H_f(x)^{-1} \nabla f(x)\right)^{1/2}.$$

Méthode de Newton – Condition d'arrêt

Traditionnellement, la condition d'arrêt de la méthode de Newton est décrite par le fait que le carré du coefficient de décroissance suivant tombe sous un certain seuil de tolérance.

$$\lambda(x) = \left(\nabla f(x)^T H_f(x)^{-1} \nabla f(x)\right)^{1/2}.$$

Celui-ci estime l'écart entre le minimisant de l'approximation du second ordre de f en un point et la valeur de f en ce même point.

Algorithme de Newton

Algorithm 1 Méthode de Newton

Input: f : a function, x_0 : an initial point in the domain of f, ε : tolerance.

Output: x^* : an optimal solution of (P) if bounded from below

function NEWTON_METHOD(f, x_0, ε)
    $x \leftarrow x_0$
    $\Delta x_N \leftarrow -\left(H_f(x)\right)^{-1} \nabla f(x)$
    $\lambda^2(x) = -\nabla f(x)^T \Delta x_N$
    while $\frac{\lambda^2(x)}{2} > \varepsilon$ **do**
        $\Delta x_N \leftarrow -\left(H_f(x)\right)^{-1} \nabla f(x)$
        $\lambda^2(x) = -\nabla f(x)^T \Delta x_N$
        compute step $t > 0$ of descent
        $x \leftarrow x + t\Delta x_N$
    end while
    return x
end function

Méthode de Newton – Quelques remarques

- Si T est une transformation affine satisfaisant $y = Tx$, les itérés de la méthode de Newton dans ces coordonées satisfont $x + \Delta x_N = T(y + \Delta y_N)$. Dans ce sens une transformation affine n'affecte pas les résultats de convergence de la méthode.

Méthode de Newton – Quelques remarques

- Si T est une transformation affine satisfaisant $y = Tx$, les itérés de la méthode de Newton dans ces coordonées satisfont $x + \Delta x_N = T(y + \Delta y_N)$. Dans ce sens une transformation affine n'affecte pas les résultats de convergence de la méthode.
- La méthode de Newton a 2 phases de convergence : une relativement lente (linéaire) lors de l'approche du point optimal puis une très rapide (quadratique) une fois suffisamment près de celui-ci.

Méthode de Newton – Quelques remarques

- Si T est une transformation affine satisfaisant $y = Tx$, les itérés de la méthode de Newton dans ces coordonées satisfont $x + \Delta x_N = T(y + \Delta y_N)$. Dans ce sens une transformation affine n'affecte pas les résultats de convergence de la méthode.
- La méthode de Newton a 2 phases de convergence : une relativement lente (linéaire) lors de l'approche du point optimal puis une très rapide (quadratique) une fois suffisamment près de celui-ci.
- Le principal ***bottle neck*** de la méthode de Newton est celui relatif au calcul des inverses des hessiennes aux itérés de la méthode, les méthodes de quasi-Newton cherchent à optimiser les méthodes de Newton en approchant les inverses des hessiennes en ces points.

C'est tout pour la méthode de Newton!
Il vous revient d'enrichir votre compréhension de cette démarche. Un mot clé : *quasi-Newton*.