CH5 : La méthode de Steepest Ascent

Introduction :

On va voir, dans ce chapitre, la méthode de steepest ascent qui est une méthode d'optimisation sans contraintes, dans R^n

La méthode de steepest ascent :

C'est une méthode itérative de recherche, par le gradient, d'un x^* qui maximise une fonction f de la manière suivante :

Soit $f : R^n \to R$ une fonction deux fois continûment différentiable sur R^n, il s'agit de déterminer un point x^* de R^n tels que : $\forall x \in R^n \; f(x) \leq f(x^*)$

On choisit un point de départ x_0.

On définit le gradient $\nabla f(x_k)$, au $k^{ème}$ point x_k.

L'idée principale : c'est de déterminer le sentier optimal (p), sur lequel $\partial f/\partial p$ est maximisée, à un point donné ; ce qui revient à dire :

si $\nabla f(x_k) \neq \underline{0}$, alors on cherche le point x_{k+1} défini par :

$$\mathbf{x_{k+1} = x_k + r_k \nabla f(x_k)}, \; k \in N$$

où r_k est le paramètre qui mesure le « pas » (steep size) à franchir, à chaque fois.

Le r_k est déterminé tels que les résultats de x_{k+1} est une amélioration de f si la fonction h définie par : $h : R \to R$

$$r \to h(r) = f(x_k + r_k \nabla f(x_k)) = f(x_{k+1})$$

r_k mesure le pas de déplacement

$h(r) = f(x_{k+1})$

on détermine r qui maximise h.

<u>d'où</u> : $h'(r) = 0$, $h''(r) < 0 \to r \to x_{k+1}$

et où r_k maximise h :

$$h'(r_k) = \partial h(r_k)/\partial r_k = 0$$
$$h''(r_k) = \partial^2 h(r_k)/(\partial r_k)^2 \leq 0$$

Remarque :

On cherche d'abord le r_k qui maximise h et ensuite, on calcule x_{k+1} ; la procédure sera terminée lorsque les deux points successifs x_k et x_{k+1} sont approximativement égaux c'est-à-dire $\| x_{k+1} - x_k \| = \| x_k + r_k \nabla f(x_k) - x_k \| = \| r_k \nabla f(x_k) \| \leq \varepsilon, \; \forall \varepsilon > 0$ ou $r_k \nabla f(x_k) \approx \underline{0}$

c-à-d x_k est un point stationnaire approximatif.

Algorithme de steepest ascent

Pas 0 : x_0 point initial, $\varepsilon > 0$, $k=0$

pt stationnaire ← Si $\nabla f(x_0) = \underline{0}$ (x_0 point stationnaire, voir sa nature par la matrice Hessienne)

Sinon, aller en (1)

Pas 1 : A l'étape k

On calcule $x_{k+1} = x_k + r_k \nabla f(x_k)$

$h(r) = f(x_{k+1})$

on cherche r qui maximise h :

$\partial h(r_k)/\partial r_k = 0$

et $\qquad \to r \to x_{k+1}$

$\partial^2 h(r_k)/(\partial r_k)^2 \leq 0$

Vais tester ← si $\nabla f(x_{k+1}) = \underline{0}$ alors $(x_{k+1}, f(x_{k+1}))$ Max de f

sinon aller en (2)

Pas 2 : si $\| x_{k+1} - x_k \| = \| r_k \nabla f(x_k) \| \leq \varepsilon$, alors $(x_k, f(x_k))$ Max de f

sinon, faire $k = k+1$ et on revient à (1)