4ème année IE

ENSSEA

Ex : INPS

CH4 : Optimisation non linéaire sans contrainte
La méthode de Newton-Raphson

Introduction :

Les méthodes d'optimisation sans contraintes, IR^n, consistent à rechercher un point x^* point stationnaire ($\nabla f(x) = 0 \Leftrightarrow (\partial f(x) / \partial x_i = 0 \; i = 1,\dots,n)$ par des procédures itératives, engendrant une suite de points $x_0, x_1, \dots, x_n$ convergeant vers un optimum local de f.

A chaque étape k, x_{k+1} est défini par : $x_{k+1} = x_k + \lambda_k d_k$ où

1) d_k est une direction de déplacement qui peut être :

_ Soit le gradient de f en x_k, $d_k = -\nabla f(x_k)$

_ Soit calculée à partir du gradient $\nabla f(x_k)$

_ Soit choisie de façon plus ou moins arbitraire, à condition que ce soit une direction de descente ie $(\nabla f(x_k))^t d_k < 0$

2) λ_k est un coefficient qui représente la valeur ou le pas de déplacement.

Parmi ces méthodes, on va étudier la méthode de Newton-Raphson, dans IR et IR^n

I) Méthode de Newton-Raphson, dans IR

Soit f : IR → IR fonction bicontinûment différentiable sur IR ; on veut minimiser f, par cette méthode.

A chaque étape k, on doit évaluer $f(x_k)$, $f'(x_k)$ et $f''(x_k)$. Dans le cas où $f'(x_k) \neq 0$, le point x_{k+1} est calculé de la manière suivante :

Si x_k est le point obtenu, à l'étape k, le développement Taylor de f, au voisinage de x_k est égal : $f(x) = f(x_k) + (x - x_k)f'(x_k) + 1/2!\,(x - x_k)^2 f''(x_k) + \theta(x - x_k)^3$ avec $\lim_{x \to x_k} \theta(x - x_k)^3 = 0$

On remplace f par son approximation quadratique :

$$g(x) = f(x_k) + (x - x_k)f'(x_k) + 1/2!\,(x - x_k)^2 f''(x_k)$$

Remarque1 :

Si $f''(x_k) > 0$, la fonction g est strictement convexe, en effet :

$$g'(x) = f'(x_k) + (x - x_k)f''(x_k)$$

$g''(x) = f''(x_k) > 0 \quad \forall x$ d'où g est strictement convexe et g admet un minimum unique x_{k+1} tels que $g'(x_{k+1}) = 0$, or $g'(x) = f'(x_k) + (x - x_k)f''(x_k) \Rightarrow$

$g'(x_{k+1}) = f'(x_k) + (x - x_k)f''(x_k) = 0 \Rightarrow (x_{k+1} - x_k)f''(x_k) = -f'(x_k) \Rightarrow x_{k+1} - x_k = -f'(x_k)/f''(x_k)$

$f''(x_k) \neq 0 \Rightarrow \mathbf{x_{k+1} = x_k - f'(x_k)/f''(x_k)}$ **Formule de Newton-Raphson**

Remarque2 :

1) Le point x_{k+1} est choisi comme l'intersection de la droite $Z = g'(x) = f'(x_k) + (x - x_k)f''(x_k)$ avec l'axe 0x.

2) Cette procédure a une propriété de convergence finie lorsqu'elle est appliquée à une fonction quadratique de la forme : $g(x) = ux^2 + vx + w$ où $u > 0$

$$g'(x) = 2ux + v$$
$$g''(x) = 2u$$

En effet, pour un point de départ x_0 quelconque on a :

$$g'(x_0) = 2ux + v$$
$$g''(x_0) = 2u > 0$$

et $x_1 = x_0 - g'(x_0)/g''(x_0) = x_0 - (2ux_0 + v)/2u = x_0 - x_0 - v/2u = -v/2u$

$g'(x_1) = 2u(-v/2u) + v = -v + v = 0$; on a $g''(x) > 0$, $\forall x$; g est alors strictement convexe

donc $(-v/2u, g(-v/2u))$ Min G de f

Algorithme de Newton-Raphson

Pas 0 : Soit x_0 point de départ, ε>0

Si $f'(x_0) = 0$ ie $|f'(x_0)| < \varepsilon$

Si $f''(x_0) > 0$ alors $(x_0, f(x_0))$ minloc de f

Sinon, aller en 1

Pas1 : A l'étape k, calculer

$$\mathbf{X_{k+1} = x_k - f'(x_k)/f''(x_k)}$$

Faire k = k+1, aller au pas 0

Remarque3 :

Lorsque le point de départ des itérations est suffisamment proche du minimum, la convergence vers ce minimum est plus rapide, si par contre, il est trop éloigné du min risque de divergence est plus grand.

II) Méthode de Newton-Raphson, dans R^n

Soit f : R^n → R fonction deux fois continûment différentiable sur R^n ; on veut minimi cette méthode.

A chaque étape k, on doit évaluer $f(x_k)$, le gradient $\nabla f(x_k)$ et la matrice hessienne $\nabla^2 f$ dans le cas où $\nabla f(x_k) \neq 0$, le point x_{k+1} est défini de la manière suivante :

Le développement de Taylor de f au voisinage de x_k est égal :

$f(x) = f(x_k) + (\nabla f(x_k))^t (x - x_k) + 1/2\,(x - x_k)^t \nabla^2 f(x_k)(x - x_k) + \| x - x_k \|^2 \theta(x - x_k)$

avec $\lim_{x \to x_k} \theta(x - x_k) = 0$

On remplace f par son approximation quadratique :

$$q(x) = f(x_k) + (\nabla f(x_k))^t (x - x_k) + 1/2 (x - x_k)^t \nabla^2 f(x_k)(x - x_k)$$

Si $\nabla^2 f(x_k)$ est une matrice définie positive, la fonction q est alors strictement convexe minimum unique x_{k+1} défini par : $\nabla q(x_{k+1}) = 0$

Or $\nabla q(x) = (\nabla f(x_k))^t + (x - x_k)^t \nabla^2 f(x_k)$ comme $\nabla q(x_{k+1}) = 0 \Rightarrow (\nabla f(x_k))^t + (x_{k+1} - x_k)^t \nabla$

$\Rightarrow \nabla f(x_k) + \nabla^2 f(x_k)(x_{k+1} - x_k) = 0 \Rightarrow$ d'où $\nabla^2 f(x_k)(x_{k+1} - x_k) = -\nabla f(x_k)$

$\Rightarrow x_{k+1} - x_k = (\nabla^2 f(x_k))^{-1} \nabla f(x_k) \quad \Rightarrow \mathbf{x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k)}$

Remarque :

$d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$ direction de déplacement qui est de descente car $(\nabla f(x_k))^t d_k <$

$(\nabla f(x_k))^t d_k = -(\nabla f(x_k))^t (\nabla^2 f(x_k))^{-1} \nabla f(x_k)$

Algorithme de Newton-Raphson

Pas 0 : Soit x_0 point de départ, ε>0

Si $\nabla f(x_0) = 0$ ie $\|\nabla f(x_0)\| < \varepsilon$

Si $\nabla^2 f(x_0)$ semi définie positive alors $(x_0, f(x_0))$ minloc de f

Stop

Sinon, aller en (1)

Pas 1 : A l'étape k, calculer

$$\mathbf{X_{k+1} = x_k - (\nabla^2 f(x_k))^{-1} \nabla f(x_k)}$$

Faire k = k+1 et aller en (0)