Optimisation sans contrainte

I/ Condition :

Rappel :

1) Sur quelle(s) variable(s) nous allons agir ?

2) Quelle est la fonction de coût ?

3) Quelle(s) sont/est les/la contrainte(s) ?

Hypothèse : On considère f deux fois différentiable

Théorème condition nécessaire d'optimalité d'ordre 1

Soit x^*, un minimum local $\Rightarrow \nabla f(x^*) = 0$

$\cancel{\Leftarrow}$ ⚠

Sketch of proof x^* minimum local $\Rightarrow \nabla^2 f(x^*) \geq 0$

(gradient de la matrice Hessienne définit positif)

Condition d'optimalité pour les problèmes quadratiques (exemple) :

$$\nabla f = Qx + g = 0 \Leftrightarrow x^* = -Q^{-1} \cdot g$$

Exemple 1 (diapo 29 :) $\min\limits_{\underline{x}} f(\underline{x}) = x_1^2 + x_2^2$

Existence des points critiques ? $\nabla f(x) = \begin{bmatrix} 2x_1 \\ -2x_2 \end{bmatrix}$

$\Rightarrow$ 1 seul point critique (candidat à l'optimalité) $x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Calcul de la matrice Hessienne : $\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ (sans utiliser le résultat quadratique)

$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ au point critique

$\rightarrow$ les valeurs propres sont $\lambda_1 = 2$ et $\lambda_2 = -2$

$\nabla^2 f(x^*)$ n'est ni définie positive, ni définie négative

$\rightarrow$ x^* est un point selle

Exemple 2 (diapo 29) : $\min_{\underline{x}} 3(x_2 - x_1^2)^2 + (1 - x_1)^2$

$3(x_2^2 + x_1^4 - 2x_2 x_1^2) + (1 + x_1^2 - 2x_1)$

$$\nabla f(x) = \begin{bmatrix} 12x_1^3 - 12x_2 x_1 + 2x_1 - 2 \\ 6x_2 - 6x_1^2 \end{bmatrix} \begin{matrix} (1) \\ (2) \end{matrix}$$

(2) $\Leftrightarrow$ $\underline{x_2 = x_1^2}$

(1) $\Leftrightarrow$ $12x_1^3 - 12x_1^2 + 2x_1 - 2 = 0 \Rightarrow \underline{x_1 = 0}$

Le candidat à l'optimalité (pt critique) est $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Matrice Hessienne :

$$\nabla^2 f(x) = \begin{bmatrix} 36x_1^2 - 12x_2 + 2 & -12x_1 \\ -12x_1 & 6 \end{bmatrix}$$

Au point $\nabla^2 f(x^*) = \begin{bmatrix} 26 & -12 \\ -12 & 6 \end{bmatrix}$

Il reste à calculer les valeurs propres $\lambda_1 > 0$ et $\lambda_2 > 0$

En conclusion : $x^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ est un minimum local

A faire à la maison : Ex 3 et 4

Exemple 3 : $f(x_1; x_2) = -x_1^4 - x_2^4$

$$\nabla f(x) = \begin{bmatrix} -4x_1^3 \\ -4x_2^3 \end{bmatrix}$$

$\nabla f(x^*) = 0 \Leftrightarrow x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ est le candidat à l'optimalité

$$\nabla^2 f(x) = \begin{bmatrix} -12x_1^2 & 0 \\ 0 & -12x_2^2 \end{bmatrix}$$

Calculée au point x^* on a $\nabla^2 f(x^*) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$\Rightarrow$ on ne peut pas conclure

⚠ Ne pas confondre condition nécessaire et condition suffisante

Démonstration :

Idée : On choisit un point de départ $X^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Et on se déplace avec un pas et une direction $X_1 = X^* + \alpha d$ $\quad d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$

Ainsi $f(X_1) = -\alpha^4 d_1^4 - \alpha^4 d_2^4 < 0$ $\quad ; X_1 = \begin{bmatrix} \alpha_1 d_1 \\ \alpha_2 d_2 \end{bmatrix}$

$\forall \alpha ; \forall d \quad f(X_1) < 0$

Exemple 4 : $f(x_1 ; x_2) = \frac{1}{2} x_1^2 + x_1 \cos(x_2)$ $\quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\nabla f(x) = \begin{bmatrix} 1 x_1 + \cos(x_2) \\ -x_1 \sin(x_2) \end{bmatrix}$$

$$\nabla f(X^*) = 0 \Leftrightarrow \begin{cases} 1 x_1 + \cos(x_2) = 0 \\ -x_1 \sin(x_2) = 0 \end{cases} \Leftrightarrow x_2 = 0 + k\pi ; \; k \in \mathbb{Z}$$

$$\nabla f(X^*) = 0 \Leftrightarrow \begin{cases} x_1 = -\cos(x_2) \\ -x_1 \sin(x_2) = 0 \quad (2) \end{cases} \quad (1) \Leftrightarrow x_1 = 0 \Leftrightarrow \cos(x_2) = 0 \; (1)$$

(1) $\Leftrightarrow x_2 = \pm \frac{\pi}{2} \; [k\pi] \Leftrightarrow X^*_{(1)} = \begin{bmatrix} 0 \\ \pm \frac{\pi}{2} \; [k\pi] \end{bmatrix}$

(2) $\sin(x_2) = 0 \Rightarrow x_2 = 0 \; [k\pi] \Rightarrow \begin{cases} x_1 = -1 \text{ si } x_2 = 0 \\ x_1 = 1 \text{ si } x_2 = \pi \end{cases}$

Au final, $X^*_{(2)} = \begin{bmatrix} \pm 1 \\ 0 \; [k\pi] \end{bmatrix}$

$\Rightarrow$ il faut considérer la matrice Hessienne pour chaque famille de points critiques : $X^*_{(1)}$ et $X^*_{(2)}$

Matrice Hessienne dans le cas général :

$$\nabla^2 f(x) = \begin{bmatrix} 1 & -\sin(x_2) \\ -\sin(x_2) & -x_1 \cos(x_2) \end{bmatrix}$$

Considérons la famille $X^*_{(2)} = \begin{bmatrix} \pm 2 \\ 0\ [k\pi] \end{bmatrix}$

Au point $X^*_{(2)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$, on a $\nabla^2 f\left(X^*_{(2)}\right) = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$

dans ce cas, $\lambda_1 > 0$ et $\lambda_2 > 0$ $\Rightarrow$ $X^*_{(2)}$ correspond à des minimum locaux

Au point $X^*_{(2)} = \begin{pmatrix} 4 \\ \pi \end{pmatrix}$, on a $\nabla f(X^*_{(2)}) = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$

dans ce cas ~~là à nouveau~~, $\lambda_1 > 0$ et $\lambda_2 > 0$ $\Rightarrow$ $X^*_{(2)}$ correspond à des minimum locaux.

Considérons la famille $X^*_{(1)} = \begin{bmatrix} 0 \\ \pm \frac{\pi}{2}\ [k\pi] \end{bmatrix}$

Au point $X^*_{(1)} = \begin{pmatrix} 0 \\ \frac{\pi}{2} \end{pmatrix}$, $\nabla f\left(X^*_{(1)}\right) = \begin{bmatrix} 4 & -2 \\ -2 & 0 \end{bmatrix}$

dans ce cas $\lambda_1 > 0$ et $\lambda_2 < 0$ $\Rightarrow$ point selle

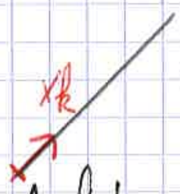
Au point $X^*_{(1)} = \begin{pmatrix} 0 \\ -\frac{\pi}{2} \end{pmatrix}$, $\nabla f\left(X^*_{(1)}\right) = \begin{bmatrix} 4 & +2 \\ +2 & 0 \end{bmatrix}$

dans ce cas $\lambda_1 > 0$ et $\lambda_2 < 0$ $\Rightarrow$ point selle

La famille $X^*_{(1)}$ représente une infinité de points selle.

II. Méthode

α variable optimal (diapo 41)



$f(x_k - \alpha_k \nabla f(x_k))$

Exemple: Calcul du pas optimal qui minimise la fonction $f(x) = \frac{1}{2}x_1^2 + \frac{9}{2}x_2^2$

A l'instant k, $f(x_k) = \frac{1}{2}x_{1k}^2 + \frac{9}{2}x_{2k}^2$

Le gradient à l'instant k vaut $\nabla f(x_k) = \begin{bmatrix} x_{1k} \\ 9x_{2k} \end{bmatrix}$ et $\nabla f(x_{k+1}) = \begin{bmatrix} x_{1k} - \alpha_k x_{1k} \\ x_{2k} - \alpha_k 9 x_{2k} \end{bmatrix}$

après un pas.

$\underbrace{f(x_k - \alpha_k \nabla f(x_k))}_{x_{k+1}} = \frac{1}{2}\left[x_{1k}(1-\alpha_k)\right]^2 + \frac{9}{2}\left[x_{2k}(1-9\alpha_k)\right]^2$

Le pas optimal vérifie $\dfrac{\partial f(x_{k+1})}{d\alpha_k} = 0$

$$\boxed{\alpha_k = \frac{x_{1k}^2 + 9x_{2k}^2}{x_{1k}^2 + 9_2 x_{2k}^2}}$$

Le pas optimal est

⚠ Si on ne trouve pas de solution analytique comme c'est le cas ici, il faut faire une boucle d'optimisation pour trouver une valeur optimale au sein de la boucle du problème d'optimisation càd. $x_{k+1} = x_k + \alpha_k \underbrace{\begin{bmatrix} -x_{1k} \\ -9x_{2k} \end{bmatrix}}_{-\nabla f(x)}$

Exemple : Condition sur la longueur de pas (diapo 49)

$f(x) = \frac{1}{2} x_1^2 + \frac{1}{4} x_2^2 - \frac{1}{2} x_2^2$

$\nabla f(x) = \begin{bmatrix} x_1 \\ x_2^3 - x_2 \end{bmatrix}$

Partons du point $X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Ainsi, le point suivant est $X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$= \begin{bmatrix} 1-\alpha \\ 0 \end{bmatrix}$

Puis, le point $X_2 = \begin{bmatrix} 1-\alpha \\ 0 \end{bmatrix} = \begin{bmatrix} (1-\alpha)^2 \\ 0 \end{bmatrix}$

La récurrence est $X_n = \begin{bmatrix} (1-\alpha)^n \\ 0 \end{bmatrix}$ ⚠ L'optimium est $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ mais on ne connait pas sa nature

Vérification par la méthode analytique : $\nabla f(x) = 0 \iff \begin{cases} x_1 = 0 \\ x_2(x_2^2 - 1) = 0 \end{cases}$

On a deux familles de points critiques :

$X^*_{(1)} = \begin{bmatrix} x_1 = 0 \\ x_2 = 0 \end{bmatrix}$ et $X^*_{(2)} = \begin{bmatrix} 0 \\ \pm 1 \end{bmatrix}$

Dans le cas général, la matrice Hessienne est $\nabla^2 f(x) = \begin{bmatrix} 1 & 0 \\ 0 & 3x_2^2 - 1 \end{bmatrix}$

Pour la famille $X^*_{(1)}$, on a $\nabla^2 f(X^*_{(1)}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \Rightarrow$ il s'agit d'un point selle

Pour la famille $X^*_{(2)}$, on a $\nabla^2 f(X^*_{(2)}) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \Rightarrow$ minimum local