MAT1112 - Optimisation avec ou sans contrainte

Notes de cours supplémentaires

D'APRÈS M. ANEL

L'*optimisation* est la recherche de maxima ou de minima (locaux) d'une fonction f. L'étude se fait toujours en deux étapes :

1. on cherche les points critiques de la fonction f qui sont les candidats pour être des maxima ou des minima,
2. pour chaque point critique, on utilise un test sur les dérivées secondes de f pour savoir s'il donne un maximum ou un minimum (mais ce test n'est pas toujours concluant).

1 Optimisation sans contraintes

1.1 Recette pour les fonctions à une variable

Soit $f : D \subset \mathbb{R} \longrightarrow \mathbb{R}$ une fonction deux fois différentiable dont la dérivée seconde est continue, on calcule ses maxima et ses minima locaux de la manière suivante.

1. on trouve les points critiques : ce sont les solutions à l'équation $f'(x) = 0$
2. pour chaque point critique x_0, on calcule $f''(x_0)$
   (a) si $f''(x_0) > 0$ alors f atteint en x_0 un minimum local,
   (b) si $f''(x_0) < 0$ alors f atteint en x_0 un maximum local,
   (c) si $f''(x_0) = 0$ on ne peut rien dire.

1.2 Recette pour les fonctions à plusieurs variables

Soit $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction deux fois différentiable dont les dérivés secondes sont continues, on calcule ses maxima et ses minima locaux de la manière suivante.

1. On trouve les points critiques : ce sont les solutions $(x_1, \ldots, x_n)$ à l'équation

$$\nabla f|_{(x_1,\ldots,x_n)} = 0 \iff \begin{cases} \left.\frac{\partial f}{\partial x_1}\right|_{(x_1,\ldots,x_n)} = 0 \\ \ldots \\ \left.\frac{\partial f}{\partial x_n}\right|_{(x_1,\ldots,x_n)} = 0 \end{cases}$$

2. La suite de la méthode ne marche que pour une fonction à deux variables $f(x, y)$.
   Pour chaque point critique (x_0, y_0) de f, on calcule la matrice hessienne de f en (x_0, y_0) :

$$H(f)|_{(x_0,y_0)} = \begin{pmatrix} \left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} & \left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0,y_0)} \\ \left.\frac{\partial^2 f}{\partial y \partial x}\right|_{(x_0,y_0)} & \left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0,y_0)} \end{pmatrix}$$

(ce doit être une matrice symétrique). Puis on calcule le "discriminant" de la matrice (c'est en fait l'opposé du déterminant) :

$$\Delta = \left(\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(x_0,y_0)} \right)^2 - \left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} \left.\frac{\partial^2 f}{\partial y^2}\right|_{(x_0,y_0)}$$

(Cette formule est un peu complexe à lire, si on note

$$\begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

la matrice hessienne, la formule pour Δ est $\Delta = B^2 - AC$.)

3. Le critère est alors le suivant :
   (a) si $\Delta < 0$ et $\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} < 0$ alors f admet en (x_0, y_0) un maximum local
   (b) si $\Delta < 0$ et $\left.\frac{\partial^2 f}{\partial x^2}\right|_{(x_0,y_0)} > 0$ alors f admet en (x_0, y_0) un minimum local
   (c) si $\Delta > 0$ alors $(x_0, y_0, f(x_0, y_0))$ est un point selle du graphe de f (on dit aussi un point col) : en particulier, ce n'est ni un maximum, ni un minimum.
   (d) dans tous les autres cas, on ne peut pas conclure.

2 Optimisation sous contraintes

Une *contrainte* sur les points de $\mathbb{R}^n$ est une équation du type $g(x_1, \ldots, x_n) = 0$ pour une certaine fonction $g : \mathbb{R}^n \to \mathbb{R}$.

On cherche maintenant à maximiser ou minimiser une fonction $f(x_1, \ldots, x_n)$ sous la contrainte $g(x_1, \ldots, x_n) = 0$. On suppose que f et g sont des fonctions deux fois différentiables dont les dérivées partielles secondes sont continues. On résout le problème de la manière suivante.

1. On définit la fonction de Lagrange

$$L(x_1, \ldots, x_n, \lambda) = f(x_1, \ldots, x_n) + \lambda g(x_1, \ldots, x_n).$$

2. On résout le sytème

$$\nabla L|_{(x_1,\ldots,x_n,\lambda)} = 0 \iff \begin{cases} \left.\frac{\partial f}{\partial x_1}\right|_{(x_1,\ldots,x_n)} = -\lambda \left.\frac{\partial g}{\partial x_1}\right|_{(x_1,\ldots,x_n)} \\ \ldots \\ \left.\frac{\partial f}{\partial x_n}\right|_{(x_1,\ldots,x_n)} = -\lambda \left.\frac{\partial g}{\partial x_n}\right|_{(x_1,\ldots,x_n)} \\ g(x_1, \ldots, x_n) = 0 \end{cases}$$

C'est un système de $n+1$ équations à $n+1$ inconnues, il admet en général plusieurs solutions.
Si $(x_1, \ldots, x_n, \lambda)$ est une solution, le vecteur $(x_1, \ldots, x_n)$ est appelé un *point critique de* $f(x_1, \ldots, x_n)$ *sous la contrainte* $g(x_1, \ldots, x_n) = 0$.

3. La suite de la méthode ne marche que pour les fonctions à deux variables $f(x,y)$ et $g(x,y)$ (et est hors programme). Pour chaque point critique (x_0, y_0, λ_0) de $L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$, on calcule la matrice hessienne de L en (x_0, y_0, λ_0) :

$$H(L)|_{(x_0,y_0,\lambda_0)} = \begin{pmatrix} \left.\frac{\partial^2 L}{\partial x^2}\right|_{(x_0,y_0,\lambda_0)} & \left.\frac{\partial^2 L}{\partial x \partial y}\right|_{(x_0,y_0,\lambda_0)} & \left.\frac{\partial^2 L}{\partial x \partial \lambda}\right|_{(x_0,y_0,\lambda_0)} \\ \left.\frac{\partial^2 L}{\partial y \partial x}\right|_{(x_0,y_0,\lambda_0)} & \left.\frac{\partial^2 L}{\partial y^2}\right|_{(x_0,y_0,\lambda_0)} & \left.\frac{\partial^2 L}{\partial y \partial \lambda}\right|_{(x_0,y_0,\lambda_0)} \\ \left.\frac{\partial^2 L}{\partial \lambda \partial x}\right|_{(x_0,y_0,\lambda_0)} & \left.\frac{\partial^2 L}{\partial \lambda \partial y}\right|_{(x_0,y_0,\lambda_0)} & \left.\frac{\partial^2 L}{\partial \lambda^2}\right|_{(x_0,y_0,\lambda_0)} \end{pmatrix}$$

(rappel : ce doit être une matrice symétrique). Puis on calcule le déterminant $\det(H(L))$ de cette matrice. (Le déterminant d'une matrice 3×3

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

est donné par la formule

$$\begin{aligned} \det(M) &= a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix} \\ &= a(ei - hf) - b(di - gf) + c(dh - ge).) \end{aligned}$$

4. Le critère est alors le suivant :
   (a) si $\det(H(L)) > 0$ alors, en (x_0, y_0), f atteint un maximum local sous la contrainte $g(x,y) = 0$,
   (b) si $\det(H(L)) < 0$ alors, en (x_0, y_0), f atteint un minimum local sous la contrainte $g(x,y) = 0$,
   (c) si $\det(H(L)) = 0$ alors on ne peut rien dire.