TD 2 : Convexité

lionel.rieg@ensiie.fr

1 Minimisation d'une fonction sans contrainte

Exercice 1

Soit une fonction quadratique $q(x) = \frac{1}{2}x.Qx + b.x + c$ avec Q symétrique de taille $n \times n$ et $x \in \mathbb{R}^n$.

1. Montrer qu'une condition nécessaire pour que q admette un minimum local est $Q \geqslant 0$ et $b \in \operatorname{Im} Q$, *i.e.* il existe b' tel que $b = Q\,b'$.
2. Montrer que sous ces conditions, le minimum est global.
3. Application : Soit $q(x) = x_1^2 + x_2^2 + 2x_3^2 - 4x_1x_2$ définie sur $\mathbb{R}^3$.
   (a) Calculer b et Q.
   (b) q admet-elle des minima locaux ou globaux ?

Exercice 2

Chercher les minima de la fonction f suivante, définie sur $\mathbb{R}^3$:

$$f(x, y, z) = e^{x-y} + e^{y-x} + z^2$$

2 Convexité

Exercice 3

Soit la fonction f définie sur $\mathbb{R}^3$ par

$$f(x, y, z) = e^{x^2+y^2+z^2}$$

1. Montrer que f est strictement convexe par deux méthodes différentes.
2. Déterminer le minimum de f.

Exercice 4

Soient f_1 une application affine de $\mathbb{R}$ dans $\mathbb{R}$, f_2 une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et g une fonction convexe de $\mathbb{R}^n$ dans $\mathbb{R}$.

1. Montrer que $g \circ f_2$ est convexe.
2. Sous quelles conditions $f_1 \circ g$ est-elle convexe ?

Exercice 5

Soit S un convexe de $\mathbb{R}^n$, g une fonction concave sur S et h une fonction convexe décroissante sur $g(S)$.

1. Montrer que $h \circ g$ est convexe. Que faut-il supposer pour que $h \circ g$ soit strictement convexe ?
2. Application : Si g est de plus strictement positive sur S, montrer que $f(x) = \frac{1}{g(x)}$ est convexe.

Exercice 6

Soit C un convexe de $\mathbb{R}^n$ et $f : C \to \mathbb{R}$. On dit que f est *quasi-convexe* lorsque

$$\forall x, y \in C, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) \leqslant \max(f(x), f(y))$$

Si de plus l'inégalité est stricte lorsque $x \neq y$ et $0 < \lambda < 1$, on dit que f est *fortement quasi-convexe*.

1. Soit x^* un minimum local strict de f quasi-convexe sur C, montrer que x^* est un minimum global strict de f sur C.
2. Soit x^* un minimum local de f fortement quasi-convexe sur C, montrer que x^* est un minimum global strict de f sur C.

Exercice 7

On considère le programme géomérique suivant :

$$\min_{x>0, y>0} g \quad \text{avec} \quad g(x,y) = x + 2\frac{y}{x^2} + \frac{2}{y} \tag{PG}$$

1. Calculer le hessien de g au point de coordonnées $(x, y) = (1, 2)$. En déduire que g n'est pas convexe.
2. Écrire le dual géométrique (DPG) de (PG).
3. Résoudre (DPG).
4. En déduire la solution de (PG).
5. Vérifier que la solution trouvée est bien un point critique de g.

Solutions

▶ Exercice 1

1. Les conditions nécessaires vues en cours sont :
   – premier ordre : $\nabla q(x^*) = 0 \iff Q\,x^* + b = 0 \iff b \in \operatorname{Im} Q$
   – second ordre : $\mathcal{H}(q) \geqslant 0$, ici $\mathcal{H}(q) = Q$
   Si on a oublié la condition du second ordre, on peut la retrouver ainsi à l'aide de la formule de Taylor :

$$q(x^* + t\delta) = q(x^*) + \underbrace{\nabla q(x^*)}_{=0}\,t\delta + \frac{1}{2}t^2\,\delta.Q\delta$$

   Pour que x^* soit minimum local, il faut donc que $\delta.Q\delta \geqslant 0$. Ceci étant nécessaire pour tout δ, en particulier pour les vecteurs propres de Q, on en déduit que toute les valeurs propres de Q doivent être positives ou nulles, *i.e.* Q doit être semi-défini positif.

2. Soit x quelconque dans $\mathbb{R}^n$. On pose $\delta = x - x^*$. La formule de Taylor donne :

$$q(x^* + \delta) = q(x^*) + \underbrace{\nabla q(x^*)}_{=0}\,t\delta + \frac{1}{2}\,\underbrace{\delta.Q\delta}_{\geqslant 0}$$

   Ainsi, $q(x) = q(x^* + \delta) \geqslant q(x^*)$.

3. (a) les termes du premier ordre donnent b et ceux du second ordre donnent Q. On a donc :

$$b = 0 \quad \text{et} \quad Q = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

   (b) Puisque $b = 0$, $\nabla q(x^*) = 0 \iff Qx + b = 0 \iff x = 0$. Les mineurs principaux de Q sont $\Delta_1 = 2$, $\Delta_2 = -8$ et $\Delta_3 = -32$. Si non pose $u = (1, 0, 0)$ et $v = (1, 1, 0)$, on a $u.Qu = 2 > 0$ et $v.Qv = -4 < 0$ donc il s'agit d'un point selle.

▶ Exercice 2

Calculons le gradient de f :

$$\nabla f(x, y, z) = \begin{pmatrix} e^{x-y} - e^{y-x} \\ e^{y-x} - e^{x-y} \\ 2z \end{pmatrix}$$

Il est nul ssi $x = y$ et $z = 0$. On a alors $f(\lambda, \lambda, 0) = 2$.
Étudions le hessien :

$$\mathcal{H}(f)(x, y, z) = \begin{pmatrix} e^{x-y} + e^{y-x} & -e^{x-y} - e^{y-x} & 0 \\ -e^{x-y} - e^{y-x} & e^{x-y} + e^{y-x} & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{d'où} \quad \mathcal{H}(f)(\lambda, \lambda, 0) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Calculant les mineurs principaux, on obtient : $\Delta_1 = 2$, $\Delta_2 = 0$ et $\Delta_3 = 0$. $\mathcal{H}(f)(\lambda, \lambda, 0)$ est donc semi-défini positif mais non défini positif. Il s'agit donc de minima locaux mais potentiellement non stricts. En effet, ils sont situés sur une même droite donc ne peuvent être stricts.

De plus, le hessien est globalement semi-défini positif (avec les mêmes valeurs pour les mineurs principaux), donc les minima sont en fait globaux.

▶ Exercice 3

1. On pose $g(x) = e^x$ et $h(x) = x^2 + y^2 + z^2$. On a donc $f = g \circ h$ et on va utiliser le théorème de composition des fonctions convexes. Le hessien de h est :

$$\mathcal{H}(h)(x) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

   Il est strictement positif donc h est strictement convexe. La fonction g est strictement croissante et strictement convexe (strictement convexe n'est pas nécessaire, convexe suffit). Par le théorème de composition des fonctions convexes, $f = g \circ h$ est strictement convexe.
   Cette méthode est nettement plus simple que de calculer le hessien de f.

$$\mathcal{H}(f)(x,y,z) = \begin{pmatrix} (4x^2+2)f(x,y,z) & 4xyf(x,y,z) & 4xzf(x,y,z) \\ 4xyf(x,y,z) & (4y^2+2)f(x,y,z) & 4yzf(x,y,z) \\ 4xzf(x,y,z) & 4yzf(x,y,z) & (4z^2+2)f(x,y,z) \end{pmatrix}$$

2. On calcule le gradient de f :

$$\nabla f(x) = \begin{pmatrix} 2xf(x,y,z) \\ 2yf(x,y,z) \\ 2zf(x,y,z) \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} f(x,y,z)$$

   Puisque f est strictement positive, ∇f s'annule en $(0,0,0)$ et on a $f(0,0,0) = 1$.

▶ Exercice 4

1. On utilise la définition : pour tout $x, y \in \operatorname{dom} f$ et tout $\lambda \in [0,1]$,

$$\begin{aligned} (g \circ f_2)(\lambda x + (1-\lambda)y) &= g(f_2(\lambda x + (1-\lambda)y)) \\ &= g(\lambda f_2(x) + (1-\lambda) f_2(y)) && f \text{ linéaire} \\ &\leqslant \lambda g(f_2(x)) + (1-\lambda) g(f_2(y)) && g \text{ convexe} \end{aligned}$$

2. On a $f_1(x) = ax + b$, avec $(a,b) \in \mathbb{R}^2$. Par la formule de Taylor pour g en 0 (de $\mathbb{R}^n$), on a

$$g(y) = g(0) + \nabla g(0) y + y.\mathcal{H}(g)(0) y + \|y\|^2 \varepsilon(y)$$

   avec $\lim_{y \to 0} \varepsilon(y) = 0$. On en déduit

$$\begin{aligned} (f_1 \circ g)(y) &= a\left(g(0) + \nabla g(0) y + y.\mathcal{H}(g)(0) y + \|y\|^2 \varepsilon(y)\right) + b \\ &= (f_1 \circ g)(0) + a \nabla g(0) y + ay.\mathcal{H}(g)(0)) y + a\|y\|^2 \varepsilon(y) \end{aligned}$$

   Puisque $\lim_{y \to 0}(a\varepsilon(y)) = 0$, on peut identifier avec la formule de Taylor pour $f_1 \circ g$: $\nabla(f_1 \circ g) = a \nabla g$ et $\mathcal{H}(f_2 \circ g) = a\mathcal{H}(g)$. Puisque $\mathcal{H}(g)$ est semi-défini positif par hypothèse sur g, $a\mathcal{H}(g)$ est semi-défini positif ssi $a \geqslant 0$.

   Autre méthode : Puisque g est convexe, on a pour tout $(x,y) \in S^2$ et tout $\lambda \in [0,1]$, $g(\lambda x + (1-\lambda)y) \leqslant \lambda g(x) + (1-\lambda) g(y)$. Supposons f_1 croissante, *i.e.* $a \geqslant 0$. On en déduit :

$$\begin{aligned} (f_1 \circ g)(\lambda x + (1-\lambda)y) &\leqslant f_1(\lambda g(x) + (1-\lambda) g(y)) && f_1 \text{ croissante} \\ &= a(\lambda g(x) + (1-\lambda) g(y)) + b \\ &= \lambda(ag(x) + b) + (1-\lambda)(ag(y) + b) \\ &= \lambda(f_1 \circ g)(x) + (1-\lambda)(f_1 \circ g)(x) \end{aligned}$$

et $f_1 \circ g$ est convexe.

Si au contraire, f_1 est décroissante, *i.e.* $a \leqslant 0$, alors la première inégalité est dans l'autre sens et $f_1 \circ g$ est concave.

▶ Exercice 5

1. par concavité de g, on a pour tout $(x, y) \in S^2$ et tout $\lambda \in [0,1]$, $g(\lambda x + (1-\lambda)y) \geqslant \lambda g(x) + (1-\lambda)g(y)$. On en déduit :

$$\begin{aligned} h(g(\lambda x + (1-\lambda)y)) &\leqslant h(\lambda g(x) + (1-\lambda)g(y)) && h \text{ décroissante} \\ &\leqslant \lambda h(g(x)) + (1-\lambda)h(g(y)) && h \text{ convexe} \end{aligned}$$

donc $h \circ g$ est convexe.

Si g est strictement concave et h strictement décroissantes, les inégalités précédentes deviennent strictes et $h \circ g$ est strictement convexe.

2. On pose $h(x) = \frac{1}{x}$ pour $x > 0$. On a alors $h'(x) = -\frac{1}{x^2} < 0$ (donc h décroissante) et $h''(x) = \frac{2}{x^3} > 0$ (donc h convexe). On conclut par la première question : $f = h \circ g$ est convexe.

▶ Exercice 6

1. Vu que x^* est un minimum local strict de f, il existe $\varepsilon > 0$ tel que pour tout h, si $0 < \|h\| < \varepsilon$, alors, $f(x^*) < f(x^* + h)$. Prenons $y \in C$ quelconque. En posant $\lambda = \frac{\varepsilon}{2\|y - x^*\|}$, on a

$$\begin{aligned} f(x^*) &< f(x^* + \lambda(y - x^*)) && x^* \text{ minimum local strict} \\ &= f(\lambda y + (1-\lambda)x^*) \leqslant \max(f(y), f(x^*)) && f \text{ quasi-convexe} \end{aligned}$$

Si $f(y) \leqslant f(x^*)$, on obtient $f(x^*) < f(x^*)$, ce qui est absurde. Donc $f(x^*) < f(y)$, *i.e.* $f(x^*)$ est un minimum global strict pour f.

2. Pour λ assez petit, on a

$$\begin{aligned} f(x^*) &\leqslant f(x^* + \lambda(y - x^*)) && x^* \text{ minimum local} \\ &= f(\lambda y + (1-\lambda)x^*) < \max(f(y), f(x^*)) && f \text{ strictement quasi-convexe} \end{aligned}$$

On conclut de même.

▶ Exercice 7

1. On a

$$\nabla g = \begin{pmatrix} 1 - \frac{4y}{x^3} \\ \frac{2}{x^2} - \frac{2}{y^2} \end{pmatrix} \quad \text{puis} \quad \mathcal{H}(g) = \begin{pmatrix} \frac{12y}{x^4} & \frac{-4}{x^3} \\ \frac{-4}{x^3} & \frac{4}{y^3} \end{pmatrix} \quad \text{donc} \quad \mathcal{H}(g)(1,2) = \begin{pmatrix} 24 & -4 \\ -4 & \frac{1}{2} \end{pmatrix}$$

On a $\Delta_n = -4$ et la dimension $n = 2$ est paire donc $\mathcal{H}(g)(1,2)$ est indéfini. En particulier, g n'est pas convexe.

2. Rappel : le dual géométrique du problème de minimisation d'un posinôme $g = \sum_{i=1}^m u_i(x)$ où $u_i(x) = c_i \prod_{j=1}^n x_j^{\alpha_{ij}}$ avec $c_i > 0$ et α_{ij} réels quelconques est :

$$\max \nu(\delta) = \prod_{i=1}^m \left(\frac{c_i}{\delta_i}\right)^{\delta_i} \quad \text{sous les contraintes} \quad \begin{cases} \delta_i > 0 & 1 \leqslant i \leqslant m \\ \sum_{i=1}^m \delta_i = 1 & \\ \sum_{i=1}^m \alpha_{ij}\delta_i = 0 & 1 \leqslant j \leqslant n \end{cases} \tag{DPG}$$

Le théorème de dualité donne alors $\nu(\delta^*) = g(x^*)$ et $u_i(x^*) = \delta_i^* g(x^*)$.

Ici, cela donne : (DPG) $\max \nu(\delta)$ avec $\nu(\delta) = \left(\frac{1}{\delta_1}\right)^{\delta_1} \left(\frac{2}{\delta_2}\right)^{\delta_2} \left(\frac{2}{\delta_3}\right)^{\delta_3}$
sous les contraintes $\delta_i > 0$, $\delta_1 + \delta_2 + \delta_3 = 1$, $\delta_1 - 2\delta_2 = 0$, $\delta_2 - \delta_3 = 0$.

3. Les contraintes du dual donnent $\delta_1 = \frac{1}{2}$, $\delta_2 = \delta_3 = \frac{1}{4}$ et $\max \nu(\delta) = (2)^{\frac{1}{2}}(2^3)^{\frac{1}{4}}(2^3)^{\frac{1}{4}} = 2^{\frac{1}{2}+\frac{3}{4}+\frac{3}{4}} = 2^2 = 4$.
4. On a donc $g(x^*) = \nu(\delta^*) = 4$. Les relations $u_i(x^*) = \delta_i \cdot g(x^*)$ donnent $x = \frac{1}{2} \cdot 4$, $\frac{2y}{x^2} = \frac{1}{4} \cdot 4$ et $\frac{2}{y} = \frac{1}{4} \cdot 4$. On en tire $x^* = 2$ et $y^* = 2$.
5. Remplaçant x^* et y^* dans ∇g, on a bien $\nabla g(2,2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.