Agrégation Interne de Mathématiques
Thierry Champion

Convexité et théorème du minimax

1 Fonctions convexes de la variable réelle : quelques propriétés

Dans cette partie on redémontre des résultats plus ou moins classiques sur les fonctions convexes définies sur un intervalle ouvert non vide I de $\mathbb{R}$. On rappelle qu'une telle fonction vérifie la propriété :

$$\forall x, y \in I, \quad \forall t \in [0,1], \qquad f\left((1-t)x + ty\right) \leq (1-t)f(x) + tf(y). \tag{1}$$

Dans toute cette partie, f est une fonction convexe sur l'intervalle I.

1. Pour $x \in I$, on définit la fonction g_x sur $I \setminus \{x\}$ par $g_x(y) = \frac{f(y)-f(x)}{y-x}$.

   (a) Démontrer que f est convexe sur I si et seulement si pour tout $x \in I$ la fonction g_x est croissante sur $]-\infty, x[\cap I$ et $I \cap]x, +\infty[$.

   (b) En déduire que f est convexe sur I si et seulement si pour tous $a < b < c$ dans I on a

   $$\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(a)}{c-a} \leq \frac{f(c)-f(b)}{c-b}.$$

2. Montrer que f admet des dérivées à droite et à gauche sur I, et que celles-ci sont croissantes. On les notera respectivement $f^{'+}$ et $f^{'-}$. Montrer de plus que

   $$\forall x, y \in I, \qquad f(x) + f^{'-}(x)(y-x) \leq f(y) \tag{2}$$

   et que f est continue sur I.

3. Réciproquement, montrer que si une fonction h dérivable sur I a une dérivée croissante alors on a

   $$\forall x, y \in I, \qquad h(x) + h^{'}(x)(y-x) \leq h(y)$$

   et en déduire que h est convexe.

4. L'inégalité de Jensen comme conséquence de (2).

   (a) Version discrète : soit $a_1, \ldots, a_n$ dans I et $t_1, \ldots, t_n$ positifs tels que $\sum_{i=1}^{n} t_i = 1$, montrer que

   $$f\left(\sum_{i=1}^{n} t_i a_i\right) \leq \sum_{i=1}^{n} t_i f(a_i).$$

(b) Version continue (en lien avec les probabilités) : soit φ une densité de probabilité sur I et $r : I \mapsto I$ une fonction continue telle que $r\varphi$ et $(f \circ r)\varphi$ sont intégrables sur I. Montrer que $\displaystyle\int_I r(x)\varphi(x)dx \in I$ et en déduire que

$$f\left(\int_I r(x)\varphi(x)dx\right) \leq \int_I f(r(x))\,\varphi(x)dx.$$

5. On définit l'épigraphe de f, noté $epi(f)$ par

$$epi(f) = \{(x,r) : x \in I, f(x) \leq r\}.$$

   (a) Soit h une fonction définie sur I. Démontrer que $epi(h)$ est convexe (dans $\mathbb{R}^2$) si et seulement si h est convexe.

   (b) Soit $x \in I$, démontrer que le vecteur $(f'^{-}(x), -1)$ est un *vecteur normal* à $epi(f)$ au point $(x, f(x))$, c'est-à-dire

$$\forall \xi \in epi(f), \qquad \langle (f'^{-}(x), -1), \xi - (x, f(x))\rangle \leq 0$$

   où $\langle \cdot, \cdot \rangle$ est le produit scalaire usuel dans $\mathbb{R}^2$.

   (c) Pour $I = \mathbb{R}$ et $x = 0$, illustrer graphiquement le résultat précédent avec les fonctions exponentielle et valeur absolue.

   (d) Soit $x \in I$ et $\alpha \in \mathbb{R}$. Montrer que $(\alpha, -1)$ un vecteur normal à $epi(f)$ au point $(x, f(x))$ si et seulement si $f'^{-}(x) \leq \alpha \leq f'^{+}(x)$.
   Remarque : l'ensemble des réels α tel que $(\alpha, -1)$ est un vecteur normal à $epi(f)$ au point $(x, f(x))$ est appelé sous-gradient de f en x, c'est une notion de dérivée généralisée pour les fonctions convexes.

2 Fonctions convexes sur $\mathbb{R}^d$

On généralise ici quelques-unes des propriétés précédentes. Dans cette partie, U est un ouvert convexe de $\mathbb{R}^d$, et $f : U \to \mathbb{R}$ est une fonction convexe sur U, c'est-à-dire que

$$\forall x, y \in U, \qquad \text{la fonction } t \mapsto f(x + t(y-x)) \text{ est convexe.}$$

1. Montrer que

$$\forall x, y \in U, \quad \forall t \in [0,1], \qquad f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

   Remarque : la réciproque est vraie mais fastidieuse.

2. On suppose désormais et jusqu'à la fin de cette partie que f est différentiable sur U, et on note ∇f son gradient. Démontrer que

$$\forall x, y \in U, \qquad f(x) + \langle \nabla f(x), y - x\rangle \leq f(y)$$

   où $\langle \cdot, \cdot \rangle$ est le produit scalaire usuel dans $\mathbb{R}^d$.

3. Démontrer que

$$\forall x, y \in U, \qquad \langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0 \tag{3}$$

4. Soit $x \in U$. On suppose que f est deux fois différentiable en x, démontrer que la Hessienne $D^2 f(x)$ vérifie
$$\forall y \in U, \qquad \langle D^2 f(x)(y-x), y-x \rangle \ \geq \ 0$$
En déduire que $D^2 f(x)$ est une matrice symétrique positive.

Remarque : cette dernière propriété justifie que la propriété (3) *soit appelée "monotonie du gradient de f".*

3 Théorème du minimax de Von Neumann

Pour $d \geq 1$ et $x, y \in \mathbb{R}^d$ on note $x \leq y$ la propriété : $x_i \leq y_i$ pour tout $i \in \{1, \ldots, d\}$.
On note Δ_d le simplexe de $\mathbb{R}^d$ défini par
$$\Delta_d = \left\{ x \in \mathbb{R}^d \ : \ x \geq 0, \ \sum x_i = 1 \right\}.$$

Dans ce qui suit, n et k sont deux entiers non nuls fixés et A est une matrice de $M_{n,k}(\mathbb{R})$. Par convention, un vecteur sera toujours représenté par une matrice colonne.

On veut démontrer l'égalité suivante due à J. Von Neumann :
$$\max_{x \in \Delta_n} \min_{y \in \Delta_k} {}^t x A y \ = \ \min_{y \in \Delta_k} \max_{x \in \Delta_n} {}^t x A y. \tag{4}$$

On notera $A_1, \ldots, A_k$ les vecteurs colonnes de A.

1. Démontrer que Δ_d est un convexe compact de $\mathbb{R}^d$.

2. Démontrer que le nombre réel $\max_{x \in \Delta_n} \min_{y \in \Delta_k} {}^t x A y$ est bien défini (question difficile, on peut admettre ce résultat).
On notera
$$V_-(A) = \max_{x \in \Delta_n} \min_{y \in \Delta_k} {}^t x A y$$
et
$$V_+(A) = \min_{y \in \Delta_k} \max_{x \in \Delta_n} {}^t x A y.$$

3. Soit X et Y deux ensembles non vides et $f : X \times Y \to \mathbb{R}$. En remarquant que pour tout x, y on a $f(x,y) \leq \sup\limits_{x \in X} f(x,y)$, démontrer que
$$\sup_{x \in X} \inf_{y \in Y} f(x,y) \ \leq \ \inf_{y \in Y} \sup_{x \in X} f(x,y).$$
En déduire que $V_-(A) \leq V_+(A)$.

4. *Projection sur un convexe.* Soit $z \in \mathbb{R}^d$ et C un convexe fermé non vide de $\mathbb{R}^d$. On appelle projeté de z sur C l'unique élément $p \in C$ tel que
$$\|z - p\|^2 = \min\{\|z - c\|^2 \ : \ c \in C\}.$$
   (a) Justifier que le minimum ci-dessus est atteint et prouver que si ce minimum est atteint en p alors
$$\forall c \in C, \qquad \langle z - p, c - p \rangle \leq 0 \tag{5}$$
   En déduire que p est bien défini.

(b) *Version faible du Théorème de séparation de Hahn-Banach.* On pose $\xi = p - z$, prouver qu'il existe $\alpha \in \mathbb{R}$ tel que

$$\langle z, \xi \rangle \ < \ \alpha \ < \ \inf\{\langle c, \xi \rangle : c \in C\}.$$

5. *Théorème de Carathéodory.* Soit Z une partie non vide de $\mathbb{R}^d$, on note $conv(Z)$ son enveloppe convexe définie par

$$conv(Z) = \left\{ \sum_{i=1}^{n} t_i z_i \ : \ n \geq 1, \ z_1, \ldots, z_n \in Z, \ t \in \Delta_n \right\}.$$

(a) Soit $\xi \in conv(Z)$, montrer que si $\xi = \displaystyle\sum_{i=1}^{n} t_i z_i$ avec $n \geq d+2$ et $z_1, \ldots, z_n \in Z, \ t \in \Delta_n$, alors il existe $z'_1, \ldots, z'_{n-1} \in Z, \ t \in \Delta_{n-1}$ tels que $\xi = \displaystyle\sum_{i=1}^{n-1} t'_i z'_i$.

Remarque : ceci indique que dans la définition ci-dessus on peut imposer $1 \leq n \leq d+1$, *ce qui constitue le théorème de Carathéodory.*

(b) En déduire que l'enveloppe convexe d'un compact Z est compacte.

6. Dans cette question on veut démontrer l'alternative : soit $V_-(A) > 0$, soit $V_+(A) \leq 0$.

(a) On note $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. Démontrer que $e_i \in \Delta_n$ pour tout i.

(b) Démontrer l'équivalence : $V_+(A) \leq 0 \quad \Leftrightarrow \quad \exists y \in \Delta_k, Ay \leq 0$.

(c) Démontrer l'équivalence :

$$\exists y \in \Delta_k, \ Ay \leq 0 \quad \Leftrightarrow \quad 0 \in conv\left(\{A_1, \ldots, A_k, e_1, \ldots, e_n\}\right).$$

(d) On suppose que $V_+(A) > 0$. En appliquant la version faible du Théorème de séparation de Banach, prouver que $V_-(A) > 0$.

7. Pour tout $\beta \in \mathbb{R}$, on note $A + \beta = (A_{i,j} + \beta)_{i,j}$. Calculer $V_-(A + \beta)$ en fonction de $V_-(A)$ et en déduire l'alternative : soit $V_-(A) > \beta$, soit $V_+(A) \leq \beta$.

8. Prouver que $V_-(A) = V_+(A)$.

L'égalité (4)*, appelée théorème du minimax, est un résultat de la théorie des jeux et a de nombreuses démonstrations et généralisations.*