Minima locaux, fonctions marginales et hyperplans séparants dans l'optimisation discrète

Christer O. Kiselman [a],

[a] *Département de Mathématiques, Université d'Uppsala, boîte postale 480, SE-751 06 Uppsala, Suède*

Reçu le 28 juin 2007 ; accepté après révision le 29 octobre 2007

Présenté par Jean-Michel Bony

1. Trois difficultés en convexité digitale

Regardons trois propriétés fondamentales dans la théorie de la convexité en variables réelles : un minimum local d'une fonction convexe est global ; la fonction marginale d'une fonction convexe est convexe ; et deux ensembles convexes disjoints peuvent toujours être séparés par un hyperplan. Comment les adapater aux espaces discrets n'est pas évident. Nous proposons ici une nouvelle classe de fonctions qui étend considérablement la classe des fonctions $L^\natural$-convexes.

2. Les fonctions fortement convexes

Définition 2.1 — *Une fonction* $f\colon \mathbf{Z}^2 \to \mathbf{R}$ *sera dite* ***fortement convexe*** *si elle est la restriction à* $\mathbf{Z}^2$ *d'une fonction convexe dans* $\mathbf{R}^2$ *qui est affine sur tout segment* $[a, a + (1, 0)]$ *ainsi que sur tout segment* $[a, a + (0, 1)]$, $a \in \mathbf{Z}^2$.

Définissons deux opérateurs de différence opérant sur les fonctions définies sur $\mathbf{Z}^2$, l'ensemble des points dans $\mathbf{R}^2$ à coordonnées entieres :

$$(D_1f)(x, y) = f(x + 1, y) - f(x, y) \text{ et } (D_2f)(x, y) = f(x, y + 1) - f(x, y), \qquad (x, y) \in \mathbf{Z}^2.$$

La quantité

$$(D_2D_1f)(x, y) = f(x + 1, y + 1) - f(x + 1, y) - f(x, y + 1) + f(x, y), \qquad (x, y) \in \mathbf{Z}^2,$$

que j'appellerai *le module de* f, va jouer un grand rôle dans la suite.

Étant donnée une fonction $f\colon \{0, 1\}^2 \to \mathbf{R}$ on peut toujours la prolonger en une fonction convexe F définie dans le carré $[0, 1]^2$. On peut prendre par exemple $F = \mathbf{cvx}(f)$, l'enveloppe convexe de f ; c'est le prolongement convexe le plus grand.

Si $D_2D_1f(0, 0) = 0$ (f est alors dite *modulaire*), $\mathbf{cvx}(f)$ est affine dans le carré $[0, 1]^2$.

Si $D_2D_1f(0,0) < 0$ (f est dite *sousmodulaire*), $\mathbf{cvx}(f)$ est affine dans le triangle à sommets $(0,0)$, $(1,0)$, $(1,1)$ ainsi que dans celui à sommets $(0,0)$, $(1,1)$, $(0,1)$. La diagonale $[(0,0),(1,1)]$ est une ligne de brisure.

Enfin si $D_2D_1f(0,0) > 0$ (f est dite *supermodulaire*), $\mathbf{cvx}(f)$ est affine dans les triangles à sommets $(0,0)$, $(1,0)$, $(0,1)$ et $(1,0)$, $(1,1)$, $(0,1)$. La diagonale $[(0,1),(1,0)]$ est une ligne de brisure.

Définition 2.2 — *Étant donnée une fonction* $f\colon \mathbf{Z}^2 \to \mathbf{R}$, *on définit son* ***prolongement canonique****, noté* $\mathbf{can}(f)\colon \mathbf{R}^2 \to \mathbf{R}$, *de la façon suivante : dans chaque carré* $a + [0,1]^2$ *où* $a \in \mathbf{Z}^2$, $\mathbf{can}(f)$ *vaut* $\mathbf{cvx}\left(f|_{a+\{0,1\}^2}\right)$.

Le prolongement canonique est donc défini dans $\mathbf{R}^2$ et convexe dans chacun des carrés $a+[0,1]^2$, $a \in \mathbf{Z}^2$, mais il n'est pas forcément convexe dans $\mathbf{R}^2$ tout entier, car il n'y a aucune garantie que les brisures sur les lignes verticales ou horizontales aillent dans le bon sens.

Lemme 2.3 — *L'opération* $f \mapsto \mathbf{can}(f)$ *est croissante et superadditive dans le sens que*

$$\mathbf{can}(f+g) \geqslant \mathbf{can}(f) + \mathbf{can}(g).$$

Pour tout point $a \in \mathbf{Z}^2$ *on a*

$$\mathbf{can}(f+g) = \mathbf{can}(f) + \mathbf{can}(g)$$

dans le carré $a + [0,1]^2$ *si et seulement si* $[D_2D_1f(a)] \cdot [D_2D_1g(a)] \geqslant 0$.

Théorème 2.4 — *Étant donnée une fonction* $f\colon \mathbf{Z}^2 \to \mathbf{R}$ *les propriétés suivantes sont équivalentes :*

A. f *est fortement convexe ;*

B. Le prolongement canonique de f *est convexe dans* $\mathbf{R}^2$ *;*

C. Pour tout point $(x,y) \in \mathbf{Z}^2$ *et tous nombres* $s,t \in \{-1,0,1\}$ *on a*

$$D_1f(x+1,y+t) - D_1f(x,y) \geqslant 0 \text{ et } D_2f(x+s,y+1) - D_2f(x,y) \geqslant 0.$$

D. Pour tous points x,y *avec* $\|x-y\|_1 = 1$ *les ensembles* $\partial f(x)$ *et* $\partial f(y)$ *de leurs sous-gradients en* x *et* y *ont un point commun.*

On peut se demander si les modules $D_2D_1u(x,y)$ peuvent être préscrits librement. C'est le cas :

Théorème 2.5 — *Étant donnée une fonction* $f\colon \mathbf{Z}^2 \to \mathbf{R}$ *quelconque il existe* $u\colon \mathbf{Z}^2 \to \mathbf{R}$ *solution de* $D_2D_1u = f$. *On peut préscrire* u *sur les axes* $x=0$ *et* $y=0$ *de façon arbitraire ; la solution est alors unique.*

Est-ce qu'il est possible de trouver une solution fortement convexe à toute équation $D_2D_1u = f$? La réponse est non :

Théorème 2.6 — *Soit* $f\colon \mathbf{Z}^2 \to \mathbf{R}$ *une fonction donnée et considérons les propriétés suivantes :*

A. Pour tous $a,b \in \mathbf{Z}$ *on a*

$$\sup_{r,s\in\mathbf{Z}} \left| \sum_{y=r}^{s} f(a,y) \right| < +\infty \text{ et } \sup_{r,s\in\mathbf{Z}} \left| \sum_{x=r}^{s} f(x,b) \right| < +\infty;$$

B. Pour tous $a,b \in \mathbf{Z}$ *et tous* $s,t \in \{-1,0,1\}$ *on a, en notant* I_k *un iverse à droite de* D_k,

$$\inf_{y\in\mathbf{Z}} \big((I_2f)(a+1,y+t) - (I_2f)(a,y)\big) > -\infty \text{ et } \inf_{x\in\mathbf{Z}} \big((I_1f)(x+s,b+1) - (I_1f)(x,b)\big) > -\infty;$$

C. Il existe une solution fortement convexe de $D_2D_1u = f$ *;*

D. Pour tous $a,b \in \mathbf{Z}$ *on a*

$$\inf_{y\in\mathbf{Z}} \big(I_2D_1f(a,y)\big) > -\infty \text{ et } \inf_{x\in\mathbf{Z}} \big(I_1D_2f(x,b)\big) > -\infty;$$

E. Il existe une solution séparément convexe de $D_2D_1u = f$.

Alors $A \Rightarrow B \Leftrightarrow C \Rightarrow D \Leftrightarrow E$.

En particulier ce résultat montre que les signes des quantités $D_2D_1u(x,y)$, $(x,y)\in\mathbf{Z}^2$, avec u fortement convexe peuvent être choisis indépendamment l'un de l'autre sans aucune contrainte.

Murota (2003:178) étudie une classe de fonctions nommée fonctions $L^\natural$-convexes (≪ L-naturellement-convexes ≫), définie à l'aide des fonctions sousmodulaires, définition qui remonte jusqu'à la notion des fonctions fortement sousadditives de Choquet (1955:132). Leur relation avec les fonctions convexes des variables réelles est donnée par le prolongement de Lovász (1983). Notons que ces fonctions $L^\natural$-convexes constituent une sous-classe assez restreinte des fonctions fortement convexes. On peut les caracteriser facilement à l'aide de ses segments de brisure :

Théorème 2.7 — *Une fonction $u\colon\mathbf{Z}^2\to\mathbf{R}$ est $L^\natural$-convexe si et seulement si elle est fortement convexe et son prolongement canonique n'a aucun segment de brisure de la forme $[a+(0,1),a+(1,0)]$, $a\in\mathbf{Z}^2$.*

La condition sur les segments de brisure équivaut à $D_2D_1u\leqslant 0$ en tout point, à comparaître au fait que les signes de cette quantité peuvent être choisis librement en chaque point pour les fonctions fortement convexes.

3. Résultats

Théorème 3.1 — *Soit $f\colon\mathbf{Z}^2\to\mathbf{R}$ une fonction fortement convexe et bornée inférieurement. Alors sa function marginale, $h(x)=\inf_{y\in\mathbf{Z}}f(x,y)$, $x\in\mathbf{Z}$, est convexe, c'est-à-dire restriction d'une fonction convexe sur $\mathbf{R}$.*

La fonction $f(x)=|x-2my|$, $(x,y)\in\mathbf{Z}^2$, où m est un entier aussi grand que l'on veut, est restriction d'une fonction convexe sur $\mathbf{R}^2$, mais sa fonction marginale n'est convexe dans aucun sens.

Démonstration. — Un résultat classique sur les fonctions convexes dit que

$$H(x)=\inf_{y\in\mathbf{R}}F(x,y),\qquad x\in\mathbf{R},$$

est convexe si $F\colon\mathbf{R}^2\to\mathbf{R}$ et convexe. Sous les hypothèses faites on obtient en posant $F=\mathbf{can}(f)$

$$H(x)=\inf_{y\in\mathbf{Z}}F(x,y)=\inf_{y\in\mathbf{Z}}f(x,y)=h(x),\qquad x\in\mathbf{Z}.$$

Pour un minimum local la question clef est de trouver un voisinage qui peut servir. Pour obtenir un résultat optimal nous définirons un voisinage qui dépend de la fonction étudiée. (Il ne s'agit pas de voisinage topologique, bien sûr.)

Définition 3.2 — *Étant donnée une fonction $f\colon\mathbf{Z}^2\to\mathbf{R}$ et un point $a\in\mathbf{Z}^2$ on dit qu'un point $b\in\mathbf{Z}^2$ est **f-voisin de** a si $\|b-a\|_\infty=1$ et si le segment $[a,b]$ est un segment de brisure de $\mathbf{can}(f)$. On notera $V_f(a)$ l'ensemble de tous les f-voisins de a.*

Un point peut donc avoir $0,2,3,\dots,8$ voisins, tous éléments de $S(a)=\{x\in\mathbf{Z}^2;\|x-a\|_\infty=1\}$.

Lemme 3.3 — *Si une fonction fortement convexe f est $\geqslant f(a)$ dans $V_f(a)$, alors $\mathbf{can}(f)\geqslant f(a)$ dans l'enveloppe convexe de $V_f(a)\cup\{a\}$.*

Démonstration. — Si $x\in\mathbf{cvx}\big(V_f(a)\cup\{a\}\big)$, alors il existe au plus trois points dans $V_f(a)\cup\{a\}$ dont l'enveloppe convexe contient x et tels que $\mathbf{can}(f)$ soit affine dans l'enveloppe convexe de ces points. Donc $\mathbf{can}(f)\geqslant f(a)$ aussi dans cette enveloppe convexe et par conséquent dans $\mathbf{cvx}\big(V_f(a)\cup\{a\}\big)$ tout entier. □

Théorème 3.4 — *Soit $f\colon\mathbf{Z}^2\to\mathbf{R}$ une fonction fortement convexe. Supposons que l'enveloppe convexe de l'ensemble $\{x\in S(a);f(x)\geqslant f(a)\}$ contient a dans son intérieur pris dans la topologie usuelle de $\mathbf{R}^2$. Alors $f(a)$ est le minimum global de f, c'est-à-dire que $f(a)=\inf_{y\in\mathbf{Z}^2}f(y)$. On peut distinguer trois cas :*

A. Si l'enveloppe convexe de $V_f(a)$ contient a dans son intérieur, il suffit de supposer que $f(b) \geqslant f(a)$ pour tout $b \in V_f(a)$.

B. Si a a des f-voisins mais si leur enveloppe convexe ne contient pas a dans son intérieur, alors $\mathbf{cvx}\big(V_f(a)\big)$ contient quand même a, et il suffit de supposer que l'on a $f(x), f(y) \geqslant f(a)$ pour deux points $x, y \in S(a)$ (non pas nécessairement distincts) tels que $\mathbf{cvx}\big(V_f(a) \cup \{x\} \cup \{y\}\big)$ contienne a dans son intérieur.

C. Enfin si a n'a pas de f-voisins, il suffit de supposer qu'il existe deux points $x, y \in S(a)$ satisfaisant à $x - a \neq \pm(y - a)$ et tels que $f(\pm x), f(\pm y) \geqslant f(a)$.

Dans ce théorème l'ensemble $S(a)$ peut toujours servir de voisinage. Or pour chaque sous-ensemble propre B de $S(a)$ il existe une fonction fortement convexe non bornée inférieurement qui satisfait à $\inf_{b \in B} f(b) \geqslant f(a)$. Par exemple $f(x, y) = \sup(x, y)$ est $\geqslant 0$ sur $S \smallsetminus \{(-1, -1)\}$ et $g(x, y) = \sup(x + y, x - y)$ est $\geqslant 0$ sur $S \smallsetminus \{(-1, 0)\}$, alors que $V_f(0) = \{(1, 1), (-1, -1)\}$ et $V_g(0) = \{(1, 0), (-1, 0)\}$ (cas B).

Aussi pour le troisième problème, celui des hyperplans séparants, les fonctions fortement convexes nous offrent une solution convenable :

Théorème 3.5 — *Soient $f, g \colon \mathbf{Z}^2 \to \mathbf{R}$ deux fonctions fortement convexes. Définissons deux ensembles A et B dans $\mathbf{Z}^3$ par*

$$A = \big\{(x, y, z) \in \mathbf{Z}^3; z \geqslant f(x, y)\big\} \text{ et } B = \big\{(x, y, z) \in \mathbf{Z}^3; z \leqslant -g(x, y)\big\}.$$

Alors il existe un plan dans $\mathbf{R}^3$ séparant A et B si et seulement si $\mathbf{can}(f) + \mathbf{can}(g) \geqslant 0$.

Théorème 3.6 — *Soient f, g, A et B comme dans le théorème précédant, et supposons que leur modules sont du même signe partout dans le sens non strict que $[D_2 D_1 f(a)] \cdot [D_2 D_1 g(a)] \geqslant 0$ pour tout point $a \in \mathbf{Z}^2$. Alors il existe un plan séparant les ensembles A et B si et seulement si $f \geqslant -g$.*

Références

[1] Gustave Choquet (1955). Theory of capacities. *Ann. Inst. Fourier,* Grenoble 5 (1953/1954), 131—295.

[2] L. Lovász (1983). Submodular functions and convexity. *Dans :* A. Bachem ; M. Grötschel ; B. Korte (réd.), *Mathematical Programming—The State of the Art.* Berlin : Springer-Verlag, pp. 235—257.

[3] Kazuo Murota (2003). *Discrete Convex Analysis.* Philadelphia : Society for Industrial and Applied Mathematics.