980 - PREMIER SEMESTRE

CHASAP-CHDBAN ISSN 0151-0509

013

# COMPTES RENDUS
# DES SÉANCES DE
# L'ACADÉMIE DES SCIENCES

TOME 290 SÉRIES A-B N° 17,18 5 MAI 1980

## Série A
## SCIENCES MATHÉMATIQUES

## Série B
## SCIENCES PHYSIQUES

PARIS

gauthier-villars

ANALYSE NUMÉRIQUE. — *Sur la différentiabilité de la fonction d'appui du sous-différentiel approché.* Note (*) de **Claude Lemaréchal** et **Evgueni Nurminskii**, présentée par Jacques-Louis Lions.

On étudie la différentiabilité de la fonction d'appui du ε-sous-différentiel d'une fonction convexe. On donne l'expression de ses dérivées directionnelles.

*Differentiability of the support function of the ε-subdifferential of a convex function, is studied; a formula is given, which expresses its directional derivatives.*

Soit $f$ une fonction convexe, $f^*$ sa fonction conjuguée, $\partial_\varepsilon f$ son sous-différentiel à ε près. Nous supposons $f$ localement lipschitzienne et coercive (i. e. $f(x)/|x| \to +\infty$ lorsque $|x| \to +\infty$). Rappelons que

$$\partial_\varepsilon f(x) = \{ g / f(y) \geqq f(x) + (g, y-x) - \varepsilon, \forall y \}.$$

Si $\varepsilon = 0$, la fonction d'appui de $\partial_\varepsilon f(x)$ n'est pas continue par rapport à $x$. En revanche, si $\varepsilon > 0$, on sait que cette fonction d'appui est localement lipschitzienne ([1], [5]). Notre but est alors d'étudier ses dérivées directionnelles.

Dans toute la suite, $\varepsilon > 0$ et $p \in \mathrm{H}$ sont fixés, où H est un espace de Hilbert réel. Pour alléger l'écriture, nous adoptons donc des notations indépendantes de ε et $p$ :

(1) $$v(x) = \max \{ (g, p) / g \in \partial_\varepsilon f(x) \}$$

est la fonction d'appui du ε-sous-différentiel, et

(2) $$q(t) = \frac{f(x+tp) - f(x) + \varepsilon}{t}$$

est le quotient différentiel perturbé (défini pour $t > 0$). On sait que

(3) $$v(x) = \inf \{ q(t) / t > 0 \}.$$

On pose

(4) $$\mathrm{G}(x) = \{ g \in \partial_\varepsilon f(x) / (g, p) = v(x) \},$$

qui est un ensemble convexe fermé borné, et

(5) $$\mathrm{T}(x) = \{ t > 0 / q(t) = v(x) \}.$$

La coercivité de $f$ implique que T est non vide et demeure borné au voisinage de $x$. L'hypothèse de Lipschitz implique que 0 n'est pas adhérent à $\mathrm{T}(x)$; en fait, si M est une constante de Lipschitz locale pour $f$, on peut montrer que $\varepsilon / 2\mathrm{M}|p|$ minore tout $t$ de T.

Les multi-applications $\mathrm{G}(x)$ et $\mathrm{T}(x)$ sont semi-continues supérieurement de H fort dans H faible et $\mathrm{R}^+$ respectivement (*voir* par exemple [2], [3]).

LEMME 1. — *$\forall g \in \mathrm{G}(x)$, $\forall t \in \mathrm{T}(x)$, on a l'inclusion*

(6) $$g \in \partial f(x+tp).$$

*Démonstration.* — Définissons $v(x)$ par

$$\begin{cases} \max(g, p), \\ f(x)+f^*(g)-(g, x) \leqq \varepsilon. \end{cases} \tag{7}$$

Il s'agit d'un problème convexe, avec une contrainte scalaire. Nous considérons alors le lagrangien spécial.

$$L(g, t)=(g, p)-\frac{1}{t}[f(x)+f^*(g)-(x, g)-\varepsilon], \tag{8}$$

défini pour $t>0$ et nous appliquons la dualité ([4] § 8.6). La fonction duale est obtenue lorsque $g$ minimise L, auquel cas

$$0 \in p-\frac{1}{t}[\partial f^*(g)-x],$$

ce qui est équivalent à

$$g \in \partial f(x+tp) \tag{9}$$

Il reste à démontrer que la fonction duale est minimisée sur $T(x)$. Pour cela nous écrivons (9) sous la forme

$$f^*(g)+f(x+tp)=(g, x+tp)$$

obtenant ainsi une valeur pour $f^*(g)$ qui, portée dans le lagrangien (8), montre que la fonction duale n'est autre que $q(t)$ donnée par (2).

C.Q.F.D.

On notera que $q(t)$, en tant que fonction duale, est convexe en $1/t$. $T(x)$ est un intervalle (compact).

LEMME 2. — *Soient x et y donnés. Alors* $\forall t_x \in T(x)$, $\forall g_y \in G(y)$, $\forall \gamma_x \in \partial f(x)$ :

$$v(y)-v(x) \leqq \frac{1}{t_x}(g_y-\gamma_x, y-x). \tag{10}$$

*Démonstration.* — D'après (3), $\forall t_y \in T(y)$, $\forall \gamma_x \in \partial f(x)$ :

$$t_y v(y) \leqq f(y+t_y p)-f(x)-(\gamma_x, y-x)+\varepsilon. \tag{11}$$

D'autre part, d'après le lemme 1 : $\forall t_x \in T(x)$, $\forall g_y \in G(y)$ :

$$t_x v(x) \geqq f(y+t_y p)+(g_y, x+t_x p-y-t_y p)-f(x)+\varepsilon,$$

ce qui peut s'écrire :

$$t_x v(x) \geqq f(y+t_y p)-(g_y, y-x)+(t_x-t_y) v(y)-f(x)+\varepsilon. \tag{12}$$

L'inégalité requise s'obtient alors en soustrayant (12) de (11).

C.Q.F.D.

Nous pouvons maintenant prouver l'existence de la dérivée directionnelle de $v$ et donner son expression :

THÉORÈME. — *Soient $x$ et $d$ donnés dans* H. *Alors* $[v(x+sd)-v(x)]/s$ *a une limite* $v'(x, d)$ *lorsque* $s \to 0^+$, *et*

$$v'(x, d) = \min_{\substack{t \in T(x) \\ \gamma \in \partial f(x)}} \max_{g \in G(x)} \left( \frac{g-\gamma}{t}, d \right). \tag{13}$$

*Démonstration*. — Posons $y = x + s d$ dans (10), divisons par $s > 0$ et passons à la limite en utilisant la semi-continuité supérieure de G : $\exists g^* \in G(x)$ tel que

$$\limsup \frac{v(y)-v(x)}{s} \leqq \frac{1}{t}(g^* - \gamma, d), \qquad \forall t \in T(x), \quad \forall \gamma \in \partial f(x). \tag{14}$$

On en déduit

$$\limsup \frac{v(y)-v(x)}{s} \leqq \min_{\substack{t \in T(x) \\ \gamma \in \partial f(x)}} \frac{1}{t}(g^* - \gamma, d).$$

Appliquons à nouveau le lemme 2 en intervertissant les rôles de $x$ et $y$ :

$$\frac{1}{t_y}(g_x - \gamma_y, y-x) \leqq v(y) - v(x), \qquad \forall t_y \in T(y), \quad \forall g_x \in G(x), \quad \forall \gamma_y \in \partial f(y).$$

Effectuant sur cette minoration la même opération que pour obtenir (14) : $\exists t^* \in T(x)$ et $\gamma^* \in \partial f(x)$ tels que

$$\frac{1}{t^*}(g - \gamma^*, d) \leqq \liminf \frac{v(y)-v(x)}{s}, \qquad \forall g \in G(x),$$

d'où

$$\max_{g \in G(x)} \frac{1}{t^*}(g - \gamma^*, d) \leqq \liminf \frac{v(y)-v(x)}{s},$$

et enfin

$$\min_{t,\gamma} \max_{g} \frac{1}{t}(g-\gamma, d) \leqq \max_{g} \frac{1}{t^*}(g-\gamma^*, d) \leqq \liminf \frac{v(y)-v(x)}{s}$$

$$\leqq \limsup \frac{v(y)-v(x)}{s} \leqq \min_{t,\gamma} \frac{1}{t}(g^*-\gamma, d) \leqq \max_{g} \min_{t,\gamma} \frac{1}{t}(g-\gamma, d).$$

La démonstration se termine alors en remarquant que, le min max étant supérieur au max min, la chaîne d'inégalités ci-dessus est en fait une chaîne d'égalités.

C.Q.F.D.

*Remarque*. — La fin de cette démonstration montre que la fonction $(1/t)(g-\gamma, d)$ admet un point-selle, si bien que l'ordre des min et max dans (13) n'intervient pas.

*Remarque*. — Si $f$ est dérivable et nulle part linéaire, alors $T(x)$, $G(x)$ et $\partial f(x)$ sont des singletons, et $v$ est différentiable. Toutefois, cette propriété ne se conserve pas lorsque $\varepsilon \to 0$.

*Remarque.* – Si l'on supprime les hypothèses faites sur $f$ (Lipschitz et coercivité), les résultats précédents restent valides à condition que G reste borné, et que T reste doublement borné, au voisinage de $x$.

*Remarque.* – L'une des applications possibles du théorème serait une définition de la dérivée seconde d'une fonction convexe. Une autre application possible serait la construction d'une méthode de type Newton pour minimiser une fonction convexe.

Dire que $x$ minimise $f$ à $\varepsilon$ près, i. e. $f(y) \geqq f(x) - \varepsilon, \forall y$, c'est dire que $0 \in \partial_\varepsilon f(x)$. Adoptant maintenant les notations $v_p(x)$, $G_p(x)$ et $T_p(x)$ pour les définitions (1), (4), (5), $x$ minimise $f$ à $\varepsilon$ près si et seulement si

$$v_p(x) \geqq 0, \qquad \forall p.$$

Une méthode de Newton consisterait alors à remplacer $v_p(x+d)$ par son approximation au premier ordre $v_p(x) + v'_p(x, d)$, et à trouver $d$ de façon que

$$v_p(x) + v'_p(x, d) \geqq 0, \qquad \forall p.$$

La formule (13) montre qu'une telle méthode reviendrait à résoudre le système d'inéquations (linéaires en $d$) :

$$v_p(x) + \min_{\substack{t \in T_p(x) \\ \gamma \in \partial f(x)}} \max_{g \in G_p(x)} \left( \frac{g-\gamma}{t}, d \right) \geqq 0, \qquad \forall p.$$

(*) Remise le 12 mai 1980.

[1] J. B. Hiriart-Urruty, *Lipschitz r-Continuity of the Approximate Subdifferential of a Convex Function.*

[2] W. W. Hogan, *S.I.A.M. Rev.*, 15, 1973, p. 591-603.

[3] C. Lemaréchal, *Extensions diverses des méthodes de gradient et applications* (*Thèse d'État*, Paris).

[4] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.

[5] E. A. Nurminskii, *Nondifferentiable Optimization with ε-Subgradient Methods*, I.I.A.S.A., working paper, WP.78-55.

*I.N.R.I.A., Rocquencourt*, 78150 *Le Chesnay;*

*I.I.A.S.A.*, 2361 *Laxenburg, Autriche*

et *Institute of Cybernetics,*

*Ukrainian Academy of Sciences, Kiev* 207, *U.R.S.S.*