Optimisation stochastique

Contents

List of Figures

1 Motivation et quelques rappels

1.1 Un problème d'optimisation "fréquent"

$$\min_{x\in\mathbb{R}^n} h(x) = \min_{x\in\mathbb{R}^n} f(x) + g(x) \tag{1}$$

avec :

- $f : \mathbb{R}^n \to \mathbb{R}$ lisse, à savoir à gradient lipschitzien :
  $\exists L > 0$ tel que $\forall x, y \in \mathbb{R}^n, ||\nabla f(x) - \nabla f(y)|| \leq L||x - y||$
- $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ convexe, potentiellement non-lisse.

Exemple :

i) $g = 0$: problème d'optimisation lisse non-convexe.

ii) $g(x) = \lambda||x||_1$ avec $\lambda > 0$ régularisation parcimonieuse.

iii) Reformulation d'un problème d'optimisation avec contraintes :
$\min\limits_{x\in\mathcal{C}} f(x)$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe non-vide.

1.2 Exemples d'applications

Exemple 1 : Moindres carrés régularisés

On dispose d'un modèle linéaire :

$$\forall x \in \mathbb{R}^n, f(x, \beta) = x^T\beta \qquad \beta \in \mathbb{R}^n \text{ paramètre du modèle}$$

On dispose d'observations $(x_i, y_i) \in (\mathbb{R}^n \times \mathbb{R})^p$ permettant d'estimer β

D'où le problème d'optimisation suivant :

$$\min_{\beta\in\mathbb{R}^n} \frac{1}{2}\sum_{i=1}^{p}(f(x_i, \beta) - y_i)^2 \quad \Leftrightarrow \quad \min_{\beta\in\mathbb{R}^n} \frac{1}{2}||X\beta - y||_2^2 \quad \text{avec } X = \begin{pmatrix} x_1^T \\ \vdots \\ x_p^T \end{pmatrix} \in \mathbb{M}_{p,n}(\mathbb{R}) \tag{2}$$

Texte manquant

i) $g(\beta) = \frac{\lambda}{2}||\beta||_2^2$: régularisation de Tikhonov.

ii) $g(\beta) = \lambda||\beta||_1$: régularisation parcimonieuse (LASSO).

iii) $g(\beta) = \frac{\lambda}{2}||\beta||_\beta^2 =$

Texte manquant

Exemple 2 : SVM (Séparateurs à Vaste Marge)

On dispose de données $(x_i)_{i\in\{1,\dots,p\}} \in \mathbb{R}^p$ labelisés $(y_i)_{i\in\{1,\dots,p\}} \in \{-1, 1\}^p$
On cherche à construire un hyperplan séparant les données (x_i) selons leurs labels (y_i)

Dans un premier temps, on suppose qu'il existe un tel hyperplan de vecteurs normal $\beta \in \mathbb{R}^n$, passant par l'origine.

Figure 1: SVM

Question : Quel hyperplan choisir ?
$\Rightarrow$ Incertitude : nombre de données, répartition dans $\mathbb{R}^n$, etc...
$\Rightarrow$ Maximiser la distance maximale entre les données et l'hyperplan.

Condition de séparabilité :

$$\forall i \in \{1, \ldots, p\}, y_i(x_i^T \beta) \geq 0$$

D'où le problème d'optimisation suivant :

$$\max_{(\beta, M) \in (\mathbb{R}^n \times \mathbb{R}^+)} M \quad \text{s.c} \quad \forall i \in \{1, \ldots, p\}, \quad y_i \frac{(x_i^T \beta)}{||\beta||_2} \geq M \tag{3}$$

<u>Remarque :</u> $d(z, \{\beta^T x = 0\}) = \frac{|\beta^T z|}{||\beta||_2}$

En pratique, la condition de séparabilité n'est pas vérifiée pour tout $i \in \{1, \ldots, p\}$
$\Rightarrow$ Pénalisation des contraintes non satisfaites.
On pose $\forall t \in \mathbb{R}, t^+ = \max(t, 0)$

Ceci conduit à formuler un autre problème d'optimisation :

$$\max_{(\beta, M) \in (\mathbb{R}^n \times \mathbb{R}^+)} M - \lambda \sum_{i=1}^{p} (1 - y_i \frac{(x_i^T \beta)}{||\beta||_2 M})^+ \tag{4}$$

avec : $||\beta||_2 = \frac{1}{M}$ et en reformulant pour obtenir un problème de minimisation :

$$\min_{\beta \in \mathbb{R}^n} ||\beta||_2^2 + \lambda \sum_{i=1}^{p} (1 - y_i(x_i^T \beta))^+ \tag{5}$$

avec : $\sum_{i=1}^p (1 - y_i(x_i^T \beta))^+$ non-lisse (non différentiable)

1.3 Rappels de convexité

<u>Définition :</u>
Soit $\mathcal{C} \subset \mathbb{R}^n$. On dit que $\mathcal{C}$ est convexe si $\forall x, y \in \mathcal{C}$, $\forall \lambda \in [0, 1]$, $\lambda x + (1 - \lambda) y \in \mathcal{C}$

<u>Exemples :</u>

i) partie affine : $\{x_0 + s \text{ avec } s \in S\}$ avec S un sous-espace vectoriel de $\mathbb{R}^n$ et $x_0 \in \mathbb{R}^n$.

ii) hyperplan : $\{x \in \mathbb{R}^n \text{ tel que } \alpha^T x = \beta\}$

iii) demi-espace : $\{x \in \mathbb{R}^n \text{ tel que } \alpha^T x \leq \beta\}$

iv) polyèdre : $\{x \in \mathbb{R}^n \text{ tel que } Ax \leq b\}$ avec $A \in \mathbb{M}_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$.

v) ellipoïde : $\{x \in \mathbb{R}^n \text{ tel que } x^T C x \leq 1\}$ avec $C \in \mathbb{S}_n(\mathbb{R})$ (matrice symétrique semi-définie positive).

Propriétés : Opérations préservant la convexité

i) intersection

ii) somme

iii) multiplication par un scalaire

iv) produit cartésien

v) image réciproque par une application linéaire

vi) image directe par une application linéaire

vii) projection : $\{x_1 \text{ tel que } (x_1, x_2) \in \mathcal{C}\}$ avec $\mathcal{C}$ convexe.

Définition : Fonction convexe sur $\mathcal{C}$ convexe

$f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.

- f est convexe sur $\mathcal{C}$ si $\forall x, y \in \mathcal{C},\ \forall \lambda \in [0, 1],\ f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$
- f est strictement convexe sur $\mathcal{C}$ si $\forall x, y \in \mathcal{C},\ \forall \lambda \in]0, 1[,\ f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$

Propriétés : CNS (Condition Nécessaire et Suffisante) de convexité dans le cas dérivable.

Soit $f : \Omega \to \mathbb{R}$ avec Ω ouvert de $\mathbb{R}^n$ et $\mathcal{C} \subset \Omega$ convexe.

1. Si f est dérivable sur Ω, alors f est convexe sur $\mathcal{C}$ convexe si et seulement si
   $\forall x, y \in \mathcal{C},\ f(y) \geq f'(x)(y - x) + f(x)$
   $\Leftrightarrow \forall x, y \in \mathcal{C},\ f(y) \geq f(x) + \nabla f(x)^T(y - x)$
2. Si f est deux fois dérivable sur Ω, alors f est convexe sur $\mathcal{C}$ convexe si et seulement si
   $\forall x \in \mathcal{C},\ f''(x)(y - x, y - x) \geq 0$
   $\Leftrightarrow \forall x \in \mathcal{C}, (y - x)^T \nabla^2 f(x)(y - x) \geq 0$

Propriété :

i) f convexe sur $\mathcal{C}$ convexe $\Rightarrow \alpha f$ convexe sur $\mathcal{C}$ convexe pour $\alpha > 0$

ii) combinaisons linéaires à coefficients positifs de fonctions convexes sont convexes

iii) f convexe sur $\mathcal{C}$ convexe :
Soit $A \in \mathbb{M}_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$.
alors $\mathcal{C}' = \{x \in \mathbb{R}^n \text{ tel que } Ax + B \in \mathcal{C}\}$ est convexe et $x \mapsto f(Ax + b)$ est convexe sur $\mathcal{C}'$.

iv) $(f_i)_{i \in \{1,\ldots,m\}}$ convexes sur $(\mathcal{C}_i)_{\{i \in \{1,\ldots,m\}\}}$, alors $\max\limits_{i \in \{1,\ldots,m\}} f_i$ convexe sur $\bigcap_{i=1}^n \mathcal{C}_i$.

v) $g : \mathbb{R}^n \to \mathbb{R}$ convexe sur $\mathcal{C} \subset \mathbb{R}^n$
$h : \mathbb{R} \to \mathbb{R}$ croissante et convexe sur $\mathcal{C}'$ tel que $g(\mathcal{C}) \subset \mathcal{C}'$
alors $h \circ g$ est convexe sur $\mathcal{C}$.

vi) $g : \mathbb{R}^n \to \mathbb{R}^p$ avec $\forall i \in \{1, \ldots, p\}$, g_i convexe sur $\mathbb{R}^n$
$h : \mathbb{R}^p \to \mathbb{R}$ croissante et convexe vis-à-vis de chacun de ses arguments.
alors $f : \underset{x \mapsto h \circ g(x) = h(g_1(x), \ldots, g_p(x))}{\mathbb{R}^n \Rightarrow \mathbb{R}}$ est convexe sur $\mathbb{R}^n$.

1.4 Régularité des fonctions convexes

Définition : Epigraphe de f
Soit $f : \mathcal{C} \to \mathbb{R}$
On appelle *épigraphe* de f, noté $\varepsilon(f)$, l'ensemble suivant :

$$\varepsilon(f) = \{(x, w) \in \mathcal{C} \times \mathbb{R} \text{ tel que } f(x) \leq w\}$$

Propriété :
$f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.
f est convexe sur $\mathcal{C}$ si et seulement si $\varepsilon(f)$ est convexe.

Preuve :

i) Supposons f convexe sur $\mathcal{C}$ convexe
$\forall x, y \in \mathcal{C}$, $\forall \lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$
Soient $(x_1, w_1) \in \varepsilon(f)$ et $(x_2, w_2) \in \varepsilon(f)$ et $\lambda \in [0,1]$
$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2) \leq \lambda w_1 + (1-\lambda) w_2$
$\Rightarrow \lambda(x_1, w_1) + (1-\lambda)(x_2, w_2) \in \varepsilon(f)$

ii) Supposons $\varepsilon(f)$ convexe
Soit $x, y \in \mathcal{C}$ et $\lambda \in [0,1]$
$(x, f(x)), (y, f(y)) \in \varepsilon(f)$
$\Rightarrow \lambda(x, f(x)) + (1-\lambda)(y, f(y)) \in \varepsilon(f)$
$\Rightarrow (\lambda x + (1-\lambda)y, \lambda f(x) + (1-\lambda) f(y)) \in \varepsilon(f)$
$\Rightarrow f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$

Propriété : Inégalité de Jensen
Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.
Soit $x_1, \ldots, x_p \in \mathcal{C}$ et $\lambda_1, \ldots, \lambda_p \in \mathbb{R}^+$ tel que $\sum_{i=1}^p \lambda_i = 1$.
Alors $f(\sum_{i=1}^p \lambda_i x_i) \leq \sum_{i=1}^p \lambda_i f(x_i)$

Propriété :
Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.
Soit $x_0 \in \mathring{\mathcal{C}}$
alors f est continue en x_0.

Preuve :
Soit $x_0 \in \mathring{\mathcal{C}}$
Soit Δ un simplexe inclus dans $\mathring{\mathcal{C}}$ et contenant x_0.
Notons $(s_i)_{i \in \{1,\ldots,n+1\}}$ les sommets de Δ.
$\forall x \in \Delta$, $\exists! (\lambda_i)_{i \in \{1,\ldots,n+1\}} \in \mathbb{R}^{n+1}$ tel que $\lambda_i \geq 0$ et $\sum_{i=1}^{n+1} \lambda_i = 1$ et $x = \sum_{i=1}^{n+1} \lambda_i s_i$
(coordonnées barycentriques de x vis-à-vis de Δ).

d'où

$$\begin{aligned} f(x) &\leq \sum_{i=1}^{n+1} \lambda_i f(s_i) \text{ par convexité de } f \text{ et inégalité de Jensen.} \\ &\leq \max_{i \in \{1,\ldots,n+1\}} f(s_i) \text{ car } \sum_{i=1}^{n+1} \lambda_i = 1 \end{aligned}$$

Donc f est majorée sur Δ.
En particulier, $\forall \delta > 0$ tel que $B(x_0, \delta) \subset \Delta$, f est majorée sur $B(x_0, \delta)$.

Fixons un tel δ et notons M un majorant de f sur $B(x_0, \delta)$.

Texte manquant

$\Rightarrow f(\delta) \leq 2f(x_0) - M$

Bilan : $2f(x_0) - M \leq f(z) \leq M, \forall z \in B(x_0, \delta)$
$\Rightarrow f$ est bornée sur $B(x_0, \delta)$

Soit $K > 0$ tel que $\forall z \in B(x_0, \delta)$, $|f(z)| \leq K$

On montre que f est lipschitzienne sur $B(x_0, \frac{\delta}{2})$
Soit $x, y \in B(x_0, \frac{\delta}{2})$, $x \neq y$

On pose : $\begin{cases} x' = x - \frac{\delta}{2}\frac{y-x}{||y-x||} \\ y' = y + \frac{\delta}{2}\frac{y-x}{||y-x||} \end{cases}$

Alors $x' \in B(x_0, \delta)$: $||x' - x_0|| = ||x - \frac{\delta}{2}\frac{y-x}{||y-x||} - x_0|| \leq ||x - x_0|| + \frac{\delta}{2} < \delta$
De même, $y' \in B(x_0, \delta)$

D'où $|f(x')| \leq K$ et $|f(y')| \leq K$

De plus, $x = \frac{2||y-x||}{2||y-x||+\delta}x' + \frac{\delta}{2||y-x||+\delta}y = \lambda x' + (1-\lambda)y'$ avec $\lambda = \frac{2||y-x||}{2||y-x||+\delta} \in]0, 1[$

Par convexité de f sur $\mathcal{C}$: $f(x) \leq \lambda f(x') + (1-\lambda)f(y)$
$\Rightarrow f(x) - f(y) \leq \lambda(f(x') - f(y)) \leq 2K\lambda \leq \frac{4K}{2||y-x||+\delta}||y-x||$
$\Rightarrow |f(x) - f(y)| \leq \frac{4K}{\delta}||y-x||$

De même, $y = \lambda y' + (1-\lambda)x$ et de la même manière, on montre que $|f(y) - f(x)| \leq \frac{4K}{\delta}||y-x||$
$\Rightarrow |f(x) - f(y)| \leq \frac{4K}{\delta}||y-x||$

Bilan : f est lipschitzienne sur $B(x_0, \frac{\delta}{2})$
En particulier, f est continue en x_0

2 Sous-différentiell d'une fonction

2.1 Sous-gradient et sous différentiel

Définition :
Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.
Soient $x \in \mathcal{C}$ et $g \in \mathbb{R}^n$
g est appelé *sous-gradient* de f en x si :

$$\forall y \in \mathcal{C}, f(y) \geq g^T(y-x) + f(x)$$

On appelle *sous-différentiel* de f en x, noté $\partial f(x)$, l'ensemble des sous-gradients de f en x :

Exemple :
$f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$
Soit $x \in \mathbb{R}$

- si $x < 0$, alors $f(x) = -x$
  Soit $g \in \mathbb{R}$ tel que $\forall y \in \mathbb{R},\ f(y) \geq g(y-x) + f(x)$
  - Soit $y \leq 0,\ f(y) = -y = -y + x - x = -(y-x) + f(x) \geq g(y-x) + f(x)$ avec $g = -1$
  - Soit $y > 0,\ f(y) = y \geq -y + x - x \geq -(y-x) + f(x) \geq g(y-x) + f(x)$ avec $g = -1$

  Donc $\partial f(x) = \{-1\}$
- si $x > 0$, $\partial f(x) = \{1\}$ (même raisonnement)
- si $x = 0$, $\partial f(x) = [-1, 1]$

Propriété :
Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.
Soit $x \in \mathcal{C}$
Alors $\partial f(x)$ est un convexe non-vide.

Preuve :
Supposons $\partial f(x) = \emptyset$
alors $\partial f(x) = \bigcap_{y \in \mathcal{C}} \{g \in \mathbb{R}^n$ tel que $f(y) \geq g^T(y-x) + f(x)\}$

or, $\forall y \in \mathcal{C}$, $\{g \in \mathbb{R}^n$ tel que $f(y) \geq g^T(y-x) + f(x)\}$ est convexe (demi-espace) et fermé (comme image réciproque d'un fermé par une application continue ψ)

$$\begin{aligned} \psi : \mathbb{R}^n &\to \mathbb{R} \\ g &\mapsto f(y) - f(x) - g^T(y-x) \end{aligned}$$

$\Rightarrow \partial f(x)$ est convexe et fermé comme intersections de parties convexes et fermées.

Propriété :
Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.
Soit $x \in \overset{\circ}{\mathcal{C}}$ tel que f est continue en x. alors $\partial f(x)$ est borné.

Preuve :
Soit $x \in \overset{\circ}{\mathcal{C}}$ tel que f est continue en x.
$x \in \overset{\circ}{\mathcal{C}}$: $\exists y_1 > 0$ tel que $B(x, y_1) \subset \mathcal{C}$
f continue en x : $\forall \varepsilon > 0,\ \exists y_2 > 0$ tel que $\forall y \in B(x, y_2),\ |f(y) - f(x)| < \varepsilon$

Posons $\eta = min(y_1, y_2)$

Montrons que $\partial f(x)$ est borné.
Supposons le contraire.
$\forall M > 0,\ \exists g \in \partial f(x)$ tel que $||g||_2 > M$
Soit $M > 0$. Fixons un tel g tel que $||g||_2 > M \Rightarrow g \neq 0$

Soit $y = x + \frac{\eta}{2}\frac{g}{||g||_2}$
d'où $||y - x||_2 = \frac{\eta}{2} < \eta \Rightarrow y \in B(x, \eta) \subset \mathcal{C}$

Par définition de g : $f(y) \geq g^T(y-x) + f(x)$
$\Rightarrow f(y) - f(x) \geq \frac{\eta}{2}||g||_2 > \frac{\eta}{2}M$ avec $M = \frac{2}{\eta}$

$\Rightarrow |f(y) - f(x)| > \varepsilon$

Or, $y \in B(x, \eta) \Rightarrow y \in B(x, y_2) \Rightarrow |f(y) - f(x)| < \varepsilon$

D'où $\varepsilon < |f(y) - f(x)| < \varepsilon$: contradiction.

Donc $\partial f(x)$ est borné.

<u>Propriété :</u>
Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe.

i) $\forall x \in \mathring{\mathcal{C}}$, $\partial f(x) \neq \emptyset$ (Et $\partial f(x)$ compact convexe non-vide)

ii) Si f dérivable en $x \in \mathring{\mathcal{C}}$, alors $\partial f(x) = \{\nabla f(x)\}$

<u>Preuve :</u>

i) Soit $\mathcal{C} \subset \mathbb{R}^n$ convexe non-vide.
Soit $x_0 \in \mathcal{C}^c \cup (\overline{\mathcal{C}} \backslash \mathring{\mathcal{C}})$
alors $\exists \alpha \in \mathbb{R}^n \backslash \{0\}$ tel que $\sup\limits_{z \in \mathcal{C}} \alpha^T z \leq \alpha^T x_0$

Texte manquant

Soit $x \in \mathring{\mathcal{C}}$

Texte manquant

ii) On suppose de plus f dérivable en $x \in \mathring{\mathcal{C}}$

- f convexe sur $\mathcal{C}$ convexe et dérivable en $x \in \mathring{\mathcal{C}}$
$\Rightarrow \forall y \in \mathcal{C},\ f(y) \geq \nabla f(x)^T (y - x) + f(x)$
$\Rightarrow \nabla f(x) \in \partial f(x)$
$\Rightarrow \{\nabla f(x)\} \subset \partial f(x)$

- Soit $g \in \partial f(x)$
$\forall y \in \mathcal{C},\ f(y) \geq g^T (y - x) + f(x)$
Or, $x \in \mathring{\mathcal{C}}$: $\exists N \in \mathbb{N}$ tel que $y_N = x + \frac{u}{N} \in \mathcal{C}$ avec $u \in \mathbb{R}^n$

  Fixons un tel N.
  $\forall n \geq N,\ f(y_n) \geq \frac{1}{n} g^T u + f(x)$

  Or, f dérivable en x :
  $f(y_n) = f(x) + \frac{1}{n} \nabla f(x)^T u + \frac{1}{n} ||u||_2 \varepsilon(\frac{1}{n} u)$ avec $\varepsilon(h) \xrightarrow[h \to 0]{} 0$
  $\Rightarrow f(x) + \frac{1}{n} \nabla f(x)^T u + \frac{1}{n} ||u||_2 \varepsilon(\frac{1}{n} u) \geq \frac{1}{n} g^T u + f(x)$
  $\Rightarrow (\nabla f(x) - g)^T u + ||u||_2 \varepsilon(\frac{1}{n} u) \geq 0$

  A la limite : $(\nabla f(x) - g)^T u \geq 0, \forall u \in \mathbb{R}^n$
  $\Rightarrow g = \nabla f(x)$

Donc $\partial f(x) \subset \{\nabla f(x)\}$

Bilan : $\partial f(x) = \{\nabla f(x)\}$

Propriété :

Soit $f : \mathcal{C} \to \mathbb{R}$ avec $\mathcal{C} \subset \mathbb{R}^n$ convexe. Soit $x^* \in \mathcal{C}$
Alors x^* est un minimum global de f sur $\mathcal{C}$ si et seulement si $0 \in \partial f(x^*)$

Preuve :

$0 \in \partial f(x^*) \Leftrightarrow \forall y \in \mathcal{C},\ f(y) \geq 0^T(y - x^*) + f(x^*) = f(x^*) \Leftrightarrow x^*$ est un minimum global de f sur $\mathcal{C}$

2.2 Calculs de sous-gradients

Pour simplifier, on suppose avoir (f_i) convexes sur $\mathbb{R}^n$

Propriété :

i) Soit $(\alpha_1, \alpha_2) \in (\mathbb{R}_+^*)^2$
On pose $f = \alpha_1 f_1 + \alpha_2 f_2$
alors $\partial f(x) = \alpha_1 \partial f_1(x) + \alpha_2 \partial f_2(x)$

ii) Soit $h : x \mapsto f(Ax + b)$ avec $A \in \mathbb{M}_{m,n}(\mathbb{R})$ et $b \in \mathbb{R}^m$
alors $\partial h(x) = A^T \partial f(Ax + b)$

iii) Soit $f : x \mapsto \max\limits_{i \in \{1,\dots,m\}} f_i(x)$
Soit $I_0 = \{i \in \{1, \dots, m\}$ tel que $f_i(x) = f(x)\}$
alors $\forall gin \partial f_{I_0}(x),\ g \in \partial f(x)$

iv) Soit $f : x \mapsto \sup\limits_{a \in A} f_a(x)$
Soit