Leçon 219 : Extremums : existence, caractérisation, recherche. Exemples et Application

Introduction

Définition 1 Soit X un ensemble et $f: X \to \mathbb{R}$. f admet un minimum (respectivement maximum) global en $a \in X$ si

$$\forall x \in X,\ f(x) \geq f(a) \quad (\text{resp } f(x) \leq f(a)).$$

- Un extremum est un maximum ou un minimum.
- Un extremum en a est strict si $\forall x \in X \setminus \{a\},\ f(x) \neq f(a)$

Définition 2 Soit $(X, \mathcal{T})$ un espace topologique, $f: X \to \mathbb{R}$ admet un extremum local en $a \in X$ s'il existe V un voisinage de a tel que $f_{/V}$ admet un extremum en a.

I Conditions d'existence globale

1) Par compacité

Proposition 3 : Soit X compact et $f: X \to \mathbb{R}$ continue. Alors f admet un maximum et un minimum global

[Pom] p78 Corollaire 4 : Soit X métrique compact et $f: X \to X$ contractante. Alors f admet un unique point fixe

p295 Proposition 5 : Soit E un espace vectoriel normé (evn) de dimension finie et $f: E \to \mathbb{R}$ continue telle que $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

Alors f admet un minimum global

Proposition 6 Soit $x \in E$ et F un ensemble fermé non vide de E. Alors $\exists y \in F$ tel que $d(x, F) = \|x-y\|_E$

Exemples :

(i) Polynômes de meilleure approximation.

Soit $n \in \mathbb{N}^*$ et $f: [0,1] \to \mathbb{R}$ continue.

$\exists P \in \mathbb{R}_n[X]$ tel que $\|f - P\|_\infty = d(f, \mathbb{R}_n[X])$

(ii) Un triangle équilatéral est d'aire maximale parmi les triangles inscrits dans un cercle donné.

2) En utilisant la convexité [Pom] p 98

Proposition 7 : tout extremum local (resp. local strict) d'une fonction convexe est global (resp global strict)

Proposition 8 Soit $\mathcal{C} \subset \mathbb{R}^n$ compact convexe et $f: \mathcal{C} \to \mathbb{R}$, une fonction strictement convexe. Alors les maximums locaux de f sont inclus dans l'ensemble des points extrémaux de f

Application (Optimisation linéaire : Algorithme du simplexe)

- Problème : soit $c: \mathbb{R}^n \to \mathbb{R}$ une forme linéaire et $A \in \mathcal{M}_n(\mathbb{R})$, $b \in \mathbb{R}^n$. On cherche à minimiser $c(x)$ en respectant les contraintes : $(Ax)_i \geq b_i$, $\forall i \in \mathbb{R}^n$.
- Idée de l'algorithme : à partir d'un sommet du polyèdre délimité par les contraintes, progresser le long d'une arête où le coût décroît jusqu'à un autre sommet, et répéter.
- L'algorithme termine et est de complexité exponentielle dans le pire des cas

II Applications dans les espaces de Hilbert

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert. On a alors l'ensemble des théorèmes suivants

Théorème 9 (de projection) Soit $\mathcal{C} \subset H$ un convexe fermé non vide, et $a \in H$. Alors il existe un unique $v \in \mathcal{C}$ tel que $\|a - v\| = d(a, H)$. v est caractérisé par

$$\begin{cases} v \in \mathcal{C} \\ \forall v \in \mathcal{C},\ (a - v, v - v) \leq 0 \end{cases}$$

[BR] p79

on note $v = P_{\mathcal{C}}(a)$

Théorème 10 (de représentation de Riesz) [BR] p 81

Soit $\varphi \in H'$, alors il existe un unique $u \in H$ tel que $\forall v \in H, \; \varphi(v) = \langle f, v \rangle$.

Théorème 11 (de Stampacchia) p 84

Soit a une forme bilinéaire coercive sur H et $K \subset H$ convexe fermé non vide, et $\varphi \in H'$.

Il existe un unique $u \in K$ tel que $\forall v \in K$, $a(u, v-u) \geq \varphi(v-u)$

Si a est symétrique, alors u est caractérisé par

$$\begin{cases} u \in K \\ \frac{1}{2} a(u,u) - \varphi(u) = \min_{v \in K} \left(\frac{1}{2} a(v,v) - \varphi(v) \right) \end{cases}$$

Application : problème de l'obstacle [KIN] p 40

Soit $\Omega \subset \mathbb{R}^n$, convexe borné non vide, de bords réguliers et soit $\psi \in H^1_0(\Omega)$ la condition d'obstacle. On cherche à minimiser

$S(v) = \int_\Omega \sqrt{1 + \|\nabla v\|^2} \, dx$ la surface de la fonction.

Dans le cas de petites variations, la linéarisation donne un nouvelle fonction à minimiser : $\frac{1}{2} a(v,v) = \frac{1}{2} \int_\Omega \|\nabla v\|^2 dx \sim S(v) - \lambda(\Omega)$

Le théorème de Stampacchia donne existence et unicité de solution au problème ∎

Corollaire 12 (Théorème de Lax Milgram) [BR] p 84

Soit a une application bilinéaire continue coercive et $\varphi \in H'$, alors il existe un unique $u \in H$ tel que $\forall v \in H$, $a(u,v) = \varphi(v)$.

Si a est symétrique, alors u est caractérisé par

$$\begin{cases} u \in H \\ \frac{1}{2} a(u,u) - \varphi(u) = \min_{v \in H} \frac{1}{2} a(v,v) - \varphi(v) \end{cases}$$

Application : Résolution d'EDP. Soit Ω ouvert convexe borné régulier. Soit $A \in (L^\infty(\Omega))^{n^2}$, uniformément elliptique ($\forall \xi \in \mathbb{R}^n$, $\langle A\xi, \xi \rangle \geq c \|\xi\|^2$ presque partout

Alors l'équation $-\nabla \cdot A \nabla u + \varepsilon u = f$ admet une unique solution $\forall f \in L^2(\Omega)$, $\forall \varepsilon > 0$ ∎

Cas particulier : Algorithme en dimension finie [DVPT]

Soit $A \in S_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$. Il existe un algorithme pour approcher le minimum de $f : \mathbb{R}^n \to \mathbb{R}$, $x \longmapsto \frac{1}{2}\langle Ax, x \rangle + \langle b, x \rangle$

(admis : Inégalité de Kantorovitch)

III Différentiabilité et extremums

1) Conditions d'ordre 1 [POM]

Définition 13 Soit E un evn, Ω un ouvert de E et $f : \Omega \to \mathbb{R}$ différentiable.

$a \in \Omega$ est un point critique de f si $df(a) = 0$.

Proposition 14 Tout extremum local d'une application différentiable est un point critique p 297

Contre exemple : Si Ω n'est pas un ouvert, $[0,1] \longrightarrow \mathbb{R}$, $x \longmapsto x$ est un contre exemple

Exemples : (i) U ouvert borné de $\mathbb{R}^n$, $f : \overline{U} \to \mathbb{R}$ continue, nulle sur ∂U et différentiable sur U. Alors f admet un point critique p 298

(ii) Si f est $\mathcal{C}^1$, convexe, ses minimums sont ses points critiques p 300

2) Conditions d'ordre 2

On travaille en dimension finie

Théorème 15 Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $\mathcal{C}^2$ sur Ω, et $a \in \Omega$.

Si f admet un minimum local en a, alors a est un point critique et $d^2 f(a)$ est symétrique positive

Si a est un point critique et $d^2 f(a)$ est symétrique définie positive, alors a est un minimum local strict de f.

3) Discussion en dimension 2

[POM] 299

Soit $f : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$, C^2.

On note $p = \frac{\partial f}{\partial x}$, $q = \frac{\partial f}{\partial y}$, $r = \frac{\partial^2 f}{\partial x^2}$, $s = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, $t = \frac{\partial^2 f}{\partial y^2}$

Alors a est un point critique $\iff p = q = 0$
et pour a un tel point, on a la disjonction de cas suivante :

- $rt < s^2 \Rightarrow a$ n'est pas un extremum
- $rt = s^2$ — on ne peut pas conclure
- $rt > s^2$ a est un $\left|\begin{array}{l}\text{minimum}\\ \text{maximum}\end{array}\right.$ local strict si $\left|\begin{array}{l} r > 0 \\ r < 0\end{array}\right.$

III Optimisation sous contrainte

Théorème 16 (extrema liés) [GOU], 317

Soit $(f, g_1, \dots g_r) \in \mathcal{C}^1(\Omega, \mathbb{R})^{r+1}$ où $r \in \mathbb{N}$ et $\Omega \subset \mathbb{R}^n$

Soit $\Gamma = \{ x \in \Omega \mid g_1(x) = g_2(x) = \dots = g_r(x) = 0 \}$

Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$ et la famille $(dg_1(a), \dots dg_r(a))$ est libre, alors $\exists (\lambda_1, \dots \lambda_r) \in \mathbb{R}$ tels que

$$df(a) = \sum_{i=1}^{r} \lambda_i \, dg_i(a)$$

Application Inégalité arithmético géométrique

Soit $f : x \in \mathbb{R}_+^n \longmapsto \prod_{i=1}^{n} x_i$ à maximiser sur $\{ \|x\|_1 = c \}$

On pose $g : x \longmapsto \sum_{i=1}^{n} x_i - c$. $dg(x) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ donc $\exists \lambda \in \mathbb{R}$

tel que $df(x) = \left(\prod_{i \neq j} x_i\right)_{1 \leq j \leq n} = \begin{pmatrix} \lambda \\ \vdots \\ \lambda \end{pmatrix} \Rightarrow \forall i \in [\![1, n]\!],\ x_i = \lambda^{\frac{1}{n-1}}$

On montre l'inégalité au point maximum, et on en déduit l'inégalité sur $\mathbb{R}_+^n$: $\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^{n} x_i$ ∎

Le théorème des extrema liés admet une généralisation avec des contraintes inégalitaires :

Théorème 17 (condition nécessaire de Fritz John) [DVPT]

Soit $(f, g_1, \dots g_m, h_1, \dots h_p) \in \mathcal{C}^1(\Omega, \mathbb{R})^{m+p+1}$ et

$$\Gamma = \{ x \in \Omega,\ g_i(x) \leq 0,\ h_j(x) = 0 \}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq p}}$$

Soit a un minimum local de f sur Γ. Alors

$\exists \lambda_0 \in \{0, 1\}$, $\exists (\gamma_i)_{1 \leq i \leq m} \in \mathbb{R}_+^m$. $\exists (\lambda_j)_{1 \leq j \leq p} \in \mathbb{R}^p$ tels que

(i) $\nabla\left(\lambda_0 f + \sum_{i=1}^{m} \gamma_i g_i + \sum_{j=1}^{p} \lambda_j h_j\right)(a) = 0$

(ii) $(\lambda_0, \gamma, \lambda) \neq 0$

(iii) $\sum_{i=1}^{m} \gamma_i g_i = 0$

Et dans le cas de contraintes convexes, on a le théorème suivant.

Théorème 18 Conditions de Karush Kuhn Tucker

Soient $J, g_1, \dots g_m$ des fonctions convexes de $\mathbb{R}^N$, on note $K = \{ v \in X,\ g_i(v) \leq 0,\ 1 \leq i \leq m \}$

On définit la fonction Lagrangien par

$$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \longrightarrow \mathbb{R}$$
$$x, \gamma, \lambda_0 \longmapsto \lambda_0 J(x) + \sum_{i=1}^{m} \gamma_i g_i(x)$$

Soit $v \in K$. Alors

1) Si v est minimum global de J sur K, $\exists (\lambda_0, \gamma)$ tels que
   a) v est minimum global de $\mathcal{L}(\cdot, \gamma, \lambda_0)$
   b) $\lambda_0 \in \{0, 1\}$, $\gamma_i \geq 0$ $1 \leq i \leq m$.
   c) $\gamma_i g_i = 0$, $1 \leq i \leq m$

2) Si $\lambda_0 \neq 0$, les trois conditions a) b) et c) sont suffisantes

3) S'il existe $x \in X \mid g_i(x) < 0$, $1 \leq i \leq m$, alors on peut prendre $\lambda_0 = 1$.

[BR] H. Brezis — Analyse fonctionnelle

[GOU] Gourdon — Cours d'Analyse

[KIN] Kinderlehrer, Stampacchia — Introduction to variational inequalities and their applications

[POM] Pommelet — Agrégation de mathématiques, cours d'Analyse

Algorithme du gradient à pas optimal

[Hiriart-Urruty, optimisation et analyse convexe]

Soit $A \in S_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$

On cherche à minimiser $f : x \longmapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle$ sur $\mathbb{R}^n$

Conditions de minimalité :

Soit $\bar{x}$ un point minimal, alors $\nabla f(x) = 0$

Or
$$\begin{aligned} f(x+h) &= \frac{1}{2}\langle A(x+h), x+h\rangle + \langle b, x+h\rangle \\ &= \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + \frac{1}{2}\left(\langle Ax, h\rangle + \langle Ah, x\rangle\right) + \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \\ &= f(x) + \langle Ax, h\rangle + \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \quad \text{par symétrie de } A \\ &= f(x) + \langle Ax+b, h\rangle + o(\|h\|) \end{aligned}$$

donc $\boxed{\nabla f(x) = Ax + b}$

$\Rightarrow \boxed{\bar{x} = -A^{-1}b}$

Minimiser f revient à inverser une matrice, de manière peut-être plus efficace

D'autre part,

$$\bar{f} := \min_{x \in \mathbb{R}^n} f(x) = f(\bar{x}) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle$$
$$= \frac{1}{2}\langle Ax, x\rangle + \langle -Ax, x\rangle = \boxed{-\frac{1}{2}\langle Ax, x\rangle}$$
$$= -\frac{1}{2}\langle A(-A^{-1}b), -A^{-1}b\rangle = \boxed{-\frac{1}{2}\langle A^{-1}b, b\rangle}$$

Algorithme

On choisit $x_0 \in \mathbb{R}^n$, et on définit la suite $(x_k)_{k\geq 0}$ par la relation de récurrence

$$\forall k \geq 0, \quad \begin{cases} d_k = -\nabla f(x_k) = -Ax_k - b \\ x_{k+1} = x_k - t_k d_k \end{cases}$$

, où t_k minimise la fonction $t \longmapsto f(x_k + t d_k)$

Preuve de la convergence de l'algorithme

Dans toute la preuve, on suppose $d_k \neq 0$. Dans le cas contraire, $Ax_k = -b$ donc l'algorithme a convergé en temps fini.

• Calcul de t_k

$$g(t) := f(x_k + t d_k) = f(x_k) + \langle \underbrace{Ax_k + b}_{-d_k}, t d_k \rangle + \frac{1}{2} \langle A t d_k, t d_k \rangle$$

$$= f(x_k) + t \left(- \|d_k\|^2 \right) + \frac{1}{2} t^2 \langle A d_k, d_k \rangle$$

$g'(t) = -\|d_k\|^2 + t \langle A d_k, d_k \rangle$ est nul pour $\boxed{t = t_k = \dfrac{\|d_k\|^2}{\langle A d_k, d_k \rangle}}$

• Calcul de l'erreur sur $f(x_k) - \bar{f}$

$$f(x_{k+1}) = f(x_k + t_k d_k) = f(x_k) + \frac{-\|d_k\|^4}{\langle A d_k, d_k \rangle} + \frac{1}{2} \frac{\|d_k\|^4}{\langle A d_k, d_k \rangle^2} \langle A d_k, d_k \rangle$$

$$= f(x_k) - \frac{1}{2} \frac{\|d_k\|^4}{\langle A d_k, d_k \rangle}$$

$$f(x_{k+1}) - \bar{f} = \left(f(x_k) - \bar{f} \right) - \frac{1}{2} \frac{\|d_k\|^4}{\langle A d_k, d_k \rangle}$$

$$= \left(f(x_k) - \bar{f} \right) \left(1 - \frac{1}{2 \left(f(x_k) - \bar{f} \right)} \frac{\|d_k\|^4}{\langle A d_k, d_k \rangle} \right) \qquad (*)$$

or $\langle A^{-1} d_k, d_k \rangle = \langle A^{-1} (Ax_k + b), (Ax_k + b) \rangle$

$$= \langle x_k, Ax_k \rangle + \langle x_k, b \rangle + \langle A^{-1} b, Ax_k \rangle + \langle A^{-1} b, b \rangle$$

$$= 2 \left(\frac{1}{2} \langle Ax_k, x_k \rangle + \langle x_k, b \rangle - \bar{f} \right) \quad \text{par symétrie de } A$$

$$= 2 \left(f(x_k) - \bar{f} \right)$$

On utilise ce résultat dans $(*)$

$$f(x_{k+1}) - \bar{f} = \left(f(x_k) - \bar{f} \right) \left(1 - \frac{\|d_k\|^4}{\langle A^{-1} d_k, d_k \rangle \langle A d_k, d_k \rangle} \right)$$

On applique maintenant l'inégalité de Kantorovitch :

$$\langle A d_k, d_k\rangle \langle A^{-1} d_k, d_k\rangle \leqslant \frac{\|d_k\|^4}{4}\left(\sqrt{c(A)} + \sqrt{\frac{1}{c(A)}}\right)^2$$

où $c(A)$ est le conditionnement de A, le rapport de la plus grande valeur propre sur la plus petite.

Ainsi, $f(x_{k+1}) - \bar{f} \leqslant \left(f(x_k) - \bar{f}\right)\left(1 - \frac{4}{\left(\sqrt{c(A)} + \frac{1}{\sqrt{c(A)}}\right)^2}\right)$

$$\leqslant \left(f(x_k) - \bar{f}\right)\left(1 - \frac{4\, c(A)}{(c(A)+1)^2}\right)$$

$$\leqslant \left(f(x_k) - \bar{f}\right)\left(\frac{c(A)-1}{c(A)+1}\right)^2$$

Donc par récurrence,

$$\boxed{f(x_k) - \bar{f} \leqslant \left(f(x_0) - \bar{f}\right)\left(\frac{c(A)-1}{c(A)+1}\right)^{2k}}$$

- Calcul de l'erreur sur $\|x_k - \bar{x}\|$

On va travailler sur $\langle A(x_k - \bar{x}), x_k - \bar{x}\rangle$.
En effet, avec $\lambda > 0$ la plus petite valeur propre de A,
on a $\|x_k - \bar{x}\|^2 \leqslant \frac{1}{\lambda}\langle A(x_k - \bar{x}), x_k - \bar{x}\rangle$

$$\leqslant \frac{1}{\lambda}\left(\langle A x_k, x_k\rangle - \langle A x_k, \bar{x}\rangle - \langle A\bar{x}, x_k\rangle + \langle A\bar{x}, \bar{x}\rangle\right)$$

$$\leqslant \frac{1}{\lambda}\left(\langle A x_k, x_k\rangle - 2\langle x_k, A\bar{x}\rangle - 2 f(\bar{x})\right)$$

$$\leqslant \frac{1}{\lambda}\left(\langle A x_k, x_k\rangle + 2\langle x_k, b\rangle - 2 f(\bar{x})\right)$$

$$\leqslant \frac{2}{\lambda}\left(f(x_k) - \bar{f}\right)$$

$$\Longrightarrow \boxed{\|x_k - \bar{x}\| \leqslant \sqrt{\frac{2}{\lambda}\left(f(x_0) - \bar{f}\right)}\left(\frac{c(A)-1}{c(A)+1}\right)^k}$$

Condition de Fritz John

[Allaire, Analyse numérique et optimisation]

Soit $(f, g_1 \dots g_m, h_1, \dots h_p) \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})^{m+p+1}$

$$K = \{x \in \mathbb{R}^n, \ \forall i \in [\![1, m]\!], \ \forall j \in [\![1, p]\!], \ g_i(x) \leqslant 0, \ h_j(x) = 0\}$$

Soit $\bar{x} \in K$. Alors si $\bar{x}$ est un minimum local de f sur K,

on a $\exists \lambda_0 \in \{0, 1\}$, $\gamma_i \geqslant 0$ $1 \leqslant i \leqslant m$, $\lambda_j \in \mathbb{R}$ $1 \leqslant j \leqslant p$ tels que

(i) $\nabla\left(\lambda_0 f + \sum_{i=1}^{m} \gamma_i g_i + \sum_{j=1}^{p} \lambda_j h_j\right) = 0$

(ii) $(\lambda_0, \gamma, \lambda) \neq 0$

(iii) $\sum_{i=1}^{m} \gamma_i g_i = 0$

On fait la preuve dans le cas $m = p = 1$

Soit $\bar{x}$ minimum local de f, appartenant à K.

$\exists \varepsilon > 0 \mid \forall x \in K \cap \overline{B(\bar{x}, \varepsilon)}$, $f(x) \geq f(\bar{x})$

Soit alors $\varphi_k : x \longmapsto f(x) + \|x - \bar{x}\|^2 + k\left(g_+^2(x) + |h(x)|^2\right)$
$\overline{B(\bar{x}, \varepsilon)} \longrightarrow \mathbb{R}$

g_+ est la partie positive de g

Par compacité de $\overline{B(\bar{x}, \varepsilon)}$, φ_k admet un minimum global en x_k

Par compacité également, il existe une suite extraite telle que $x_{k_i} \xrightarrow[i \to +\infty]{} x_\infty$

Montrons le théorème en trois étapes

(1) $\bar{x} = x_\infty$

(2) Condition de minimum sur x_k

(3) Passage à la limite : condition sur $x_\infty = \bar{x}$

(1) $\forall k \geq 1, \quad \underbrace{f(x_k) + \|x_k - \bar{x}\| \leq}_{\text{ajout de termes positifs}} \underbrace{\varphi_k(x_k) \leq}_{\text{minimalité de } x_k} \underbrace{\varphi_k(\bar{x}) = f(\bar{x})}_{\bar{x} \in K.}$

Par passage à la limite, on obtient

$$f(x_\infty) + \|x_\infty - \bar{x}\| \leq f(\bar{x}) \text{ soit}$$

$$\text{si } x_\infty \neq \bar{x}, \quad f(x_\infty) < f(\bar{x}).$$

Montrons que $x_\infty \in K \cap \overline{B(\bar{x}, \varepsilon)}$ pour obtenir une contradiction

$$f(x_k) + \|x_k - \bar{x}\|^2 + k\left(g_+(x_k)^2 + h(x_k)^2\right) \leq \varphi_k(\bar{x}) = f(\bar{x})$$

donc
$$k\left(g_+(x_k)^2 + h(x_k)^2\right) \leq f(\bar{x}) - f(x_k)$$
$$\leq f(\bar{x}) - \min_{x \in \overline{B(\bar{x},\varepsilon)}} f(x)$$

$$g_+(x_k)^2 + h(x_k)^2 \leq \frac{1}{k}\left(f(\bar{x}) - \min_{x \in \overline{B(\bar{x},\varepsilon)}} f(x)\right) \xrightarrow[k \to +\infty]{} 0$$

$\Longrightarrow \quad g_+(x_\infty) = 0$ soit $g(x_\infty) \leq 0$

et $h(x_\infty) = 0$

Donc $x_\infty \in \bar{x}$

On en conclut que $x_\infty = \bar{x}$

(2) $\exists k_0 \mid k \geq k_0 \implies x_k \in B(\bar{x}, \varepsilon)$

Donc la minimalité de x_k pour φ_k se traduit par

$\nabla \varphi_k(x_k) = 0$

$$\Longrightarrow \quad \nabla f(x_k) + 2(x_k - \bar{x}) + 2k\left(g_+(x_k)\nabla g(x_k) + h(x_k)\nabla h(x_k)\right) = 0$$

Soit alors $\gamma_k = 2k\, g_+(x_k) \geq 0$, $\quad \lambda_k = 2k\, h(x_k) \in \mathbb{R}$.

) On exprime ces contraintes et on passe à la limite.

On a deux cas :

(a) $(\gamma_k)_{k\in\mathbb{N}}$ et $(\lambda_k)_{k\in\mathbb{N}}$ sont bornées, on peut donc extraire des sous suites convergentes, de limites γ et λ.

On est dans le cas $\lambda_0 = 1$, et les conditions (i) et (ii) sont vérifiées

Si $g(\bar{x}) < 0$, alors $\exists k' \ (k \geq k' \Rightarrow g_+(x_k) = 0$

On a alors $\gamma = 0$.

On a donc prouvé (iii).

(b) Dans le cas où $(\gamma_k)_{k\in\mathbb{N}}$ et $(\lambda_k)_{k\in\mathbb{N}}$ ne sont pas bornées,

on a une sous suite telle que $|\gamma_k| + |\lambda_k| \longrightarrow +\infty$.

On divise donc la condition au rang k par

$|\gamma_k| + |\lambda_k|$, et avec $\begin{cases} \gamma_k' = \dfrac{\gamma_k}{|\gamma_k| + |\lambda_k|} \\ \lambda_k' = \dfrac{\lambda_k}{|\gamma_k| + |\lambda_k|} \end{cases}$ $k \in \mathbb{N}$,

On refait le raisonnement du (a) pour obtenir γ et λ

tels que $0 = \gamma \nabla g(\bar{x}) + \lambda \nabla h(\bar{x})$. On a (i) et (ii).

Si $g(\bar{x}) < 0$, $\exists k' \mid k \geq k' \Rightarrow g(x_k) = 0$

et alors $\gamma = 0$ On a (iii)

▦