THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Une méthode hybride, pénalisation-quasi Newton globalement et superlinéairement convergente

Soblet, Michèle

Award date:
1977

Awarding institution:
Universite de Namur

Une Méthode Hybride
Pénalisation - Quasi Newton
Globalement et Superlinéairement
Convergente

Je remercie Monsieur
V. Hien Nguyen
pour son aide et ses
conseils.

Michele Sablet

Sablet
Michele

CHAPITRE I : INTRODUCTION.

a : But du mémoire

Dans ce mémoire, nous aimerions résoudre le problème suivant:

$$\begin{array}{ll} \text{minimiser} & f(x) \\ \text{sous les contraintes} & g_j(x) \leq 0 \qquad j=1..m \end{array}$$

Les fonctions f et g_j $(j=1...m)$: $R^n \to R$

L'optimisation distingue principalement deux cas; celui où les fonctions f et g_j $(j=1..m)$ sont linéaires, on parle alors de problème de programmation linéaire, et celui où les fonctions f et g_j $(j=1...m)$ ne sont pas linéaires, on parle alors de problème de programmation non linéaire. C'est dans ce cas plus particulier que nous nous placerons.

La solution d'un programme linéaire n'est pas toujours évidente, c'est pourquoi la programmation non linéaire est amenée à construire des algorithmes pour déterminer cette solution.

Le principe des algorithmes est le suivant: à partir d'un point initial, on détermine, par un procédé propre à l'algorithme, une suite de points convergeant vers la solution

On distingue principalement deux types d'algo-rithmes. Les algorithmes du type "global" engendrent la solution quel que soit le point initial. Les algorithmes du type "local" engendrent la solution pourvu que le point initial soit assez proche de cette solution

Les algorithmes locaux convergent plus vite que les autres, mais la difficulté est de déterminer leur point initial. L'avantage des algorithmes globaux est le choix arbitraire du point initial.

L'idéal serait de profiter des avantages des deux types d'algorithmes. Pour cela, il faudrait démarrer d'un point quelconque, engendrer les itérés par un procédé propre aux algorithmes globaux, "mais une fois assez proche de la solution" déterminer les itérés par un procédé propre aux algorithmes locaux.

La difficulté se trouve dans le terme "une fois assez proche de la solution", Que signifie t-il ?

Les maîtres de l'optimisation se sont penchés sur ce problème parce qu'ils savent qu'il est une des clés de l'optimisation. Phillips avait présenté en 1974 une première approche du problème, il s'était surtout interessé à des résultats numériques.
Les références exactes sont:

Phillips, D.A. "A preliminary Investigation of Functions Optimization by a Combination of Methods" Computer Journal, Vol 17 pp 75-79

Le premier à avoir formalisé mathématiquement le terme "Une fois assez proche de la solution" est Chung

Chung, S. "Globally and superlinearly convergent Algorithms for non linear Programming" University of Wisconsin - Madison, Ph.D., 1975. Computer Science.

C'est la thèse de Chung que nous avons suivi dans l'élaboration du mémoire.

B : Lignes directrices de la thèse de Chung.

Chapitre 2

On y présente une classe d'algorithmes de pénalisation exacte convergent globalement. La fonction de pénalisation utilisée est :

$$P(v, x) = v f(x) + \sum_{j=1}^{m} \max\{0, g_j(x)\} . \quad v>0$$

A chaque itération, on utilise un programme linéaire pour déterminer une direction de descente pour la fonction de pénalisation. Une procédure d'Armijo est utilisée pour déterminer la longueur du pas.

Chapitre 3

On étudie une classe d'algorithmes du type Quasi Newton convergent localement et superlinéairement. Un premier type d'algorithmes réduit le programme non linéaire en une suite d'équations linéaires tandis que le deuxième type d'algorithme réduit le programme non linéaire en une suite de programmes quadratiques.

Chapitre 4

Ce chapitre est le plus intéressant, Chung y présente une méthode pour joindre un algorithme Quasi Newton à un algorithme de Pénalisation exacte. Si certaines conditions de régularité sont vérifiées, Chung assure que les itérés peuvent être déterminé par une méthode Quasi Newton. Il appuie ses conclusions par des tests numériques

C : Notations et définitions

1) Nous notirons le programme non linéaire :

minimiser $f(x)$

sous les contraintes $g_j(x) \leq 0 \quad j = 1 \dots m$

P.N.L (1).

2) Un point (x, u) de R^{n+m} sera un point de Kuhn Tucker du P.N.L (1) si et seulement si :

1) $\nabla f(x) + \nabla g(x) \cdot u = 0$

2) $g(x) \leq 0$

3) $u \geq 0$

4) $u^T g(x) = 0$

On appelle ces quatre conditions, conditions de Kuhn Tucker. Ces conditions sont en fait des conditions nécessaires d'optimalité

: Mangasarian O.L., " Nonlinear Programming", Mc Graw-Hill, New York, 1969.

Il existe aussi des conditions suffisantes d'optimalité. Ce sont les conditions de Kuhn Tucker auxquelles on ajoute les conditions du second ordre.

Un couple $(\bar{x},\bar{v})$ vérifie les conditions du second ordre, si et seulement si :

1) $\forall y \in \mathbb{R}^n$, $y \neq 0$ tel que : $y^T \nabla g_j(\bar{x}) = 0 \quad \bar{v}_j > 0 \quad g_j(\bar{x}) = 0$
ou $y^T \nabla g_j(\bar{x}) \leq 0 \quad \bar{v}_j = 0 \quad g_j(\bar{x}) = 0$

on a : $y^T(\nabla^2 f(\bar{x}) + \sum_{j=1}^{m} \bar{v}_j \nabla^2 g_j(\bar{x}))\, y > 0$.

2) $\{\nabla g_j(\bar{x}) \ , \ g_j(\bar{x}) = 0\}$ sont linéairement indépendant

3) $\bar{v}_j > 0$ si $g_j(\bar{x}) = 0$.

3) Tous les vecteurs sont des vecteurs colonnes. Un vecteur ligne se représentera en utilisant un indice supérieur : T
Parfois un vecteur colonne de $\mathbb{R}^{n+m}$ est écrit : (x,u) à la place de $\begin{pmatrix} x \\ u \end{pmatrix}$.

x_j : désigne la $j^{ème}$ composante du vecteur x
soit $J \subset \{1 \dots n\}$ et $x \in \mathbb{R}^n$.
x_J est le sous vecteur de x dont les indices appartiennent à J.

4) Nous appelerons g la fonction $: \mathbb{R}^n \rightarrow \mathbb{R}^m$ telle que pour chaque x de $\mathbb{R}^n$, $g(x) = (g_1(x) \; .... \; g_m(x))$

$\nabla g(x)$ est la matrice dont l'élement de la i^{eme} ligne et de la j^{eme} colonne est : $\dfrac{d g_j}{d x_i}$

5) pour chaque x de $\mathbb{R}^n$ on définit :

$$I(x) = \{ j : g_j(x) = 0 \}$$
$$I^-(x) = \{ j : g_j(x) < 0 \}$$
$$I^+(x) = \{ j : g_j(x) > 0 \} .$$
$$I^{-0}(x) = \{ j : g_j(x) \leq 0 \}$$
$$I^{+0}(x) = \{ j : g_j(x) \geq 0 \}$$

6) soient $x, y \in \mathbb{R}^n$, on dit que $x \leq y$ si et seulement si : $x_i \leq y_i$ $\quad i = 1 ... n$.

on dit que $x < y$ si $x \leq y$ et si $x \neq y$.

7) Un vecteur x de $\mathbb{R}^n$ est admissible si et seulement si $g(x) \leq 0$.

8) pour chaque x de $\mathbb{R}^n$, on note $T(x)$ l'ensemble $\{w \in \mathbb{R}^m \ , \ w \geq 0 \ , \ w_j \leq 1 \text{ pour } j \in I^+(x)\}$

9) Dans le mémoire, il y a des démonstrations longues et compliquées, nous donnerons alors une idée de la démonstration avant de l'entamer. On peut avoir une idée assez précise du travail en ne lisant que les textes soulignés verticalement. Le reste étant surtout des démonstrations techniques.

CHAPITRE II : ALGORITHMES DE PENALISATION EXACTE GLOBALEMENT CONVERGENTS

A Introduction

Dans ce chapitre nous allons donner une classe d'algorithmes de pénalisation exacte globa=
=lement convergents pour trouver la solution du P.N.L. :

$$\left|\begin{array}{l} \underset{x}{\text{minimiser}}\ f(x) \qquad (1) \\ \text{sous les contraintes } g(x) \leq 0. \end{array}\right.$$

où $f : \mathbb{R}^n \to \mathbb{R}$ et $g = (g_1, \dots g_m)$
$g_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1 \dots m$.

L'idée de base des méthodes de pénalisation exacte est d'incorporer les contraintes dans la fonction à minimiser de telle sorte que la so=
=lution du problème sans contrainte qui en résulte, fournisse la solution du P.N.L.(1) avec contraintes.

Cette idée, d'incorporer les contraintes dans la fonction à minimiser, n'est pas nouvelle. En 1968, A.V. Fiacco et G.P. McCormick avaient déjà proposé une méthode de pénalisation.

La fonction de pénalisation qu'ils avaient introduite était :

$$\left\| \begin{array}{l} \mathcal{P} : \mathbb{R}_+^* \times \mathbb{R}^n \longrightarrow \mathbb{R} \\ (\alpha, x) \rightsquigarrow f(x) + \frac{\alpha}{2} \sum_{j=1}^{m} (g_j(x))_+^2 \end{array} \right.$$

où $(g_j(x))_+ = \max\{0, g_j(x)\}$

α devient infiniment grand

Et leur problème de minimisation sans contrainte était donc :

$$\left\| \; \underset{x}{\text{minimiser}} : f(x) + \frac{\alpha}{2} \sum_{j=1}^{m} (g_j(x))_+^2 . \right.$$

La principale difficulté de cette méthode vient de ce que α devient infiniment grand.

Zangwill [1967], Pietrzykowski, [1969] et Howe [1973], ont proposé une autre fonction de pénalisation :

$$\left\| \begin{array}{l} \mathcal{P} : \mathbb{R}_+^* \times \mathbb{R}^n \longrightarrow \mathbb{R} \\ (v, x) \rightsquigarrow v f(x) + \sum_{j=1}^{m} g_j(x)_+ . \end{array} \right.$$

où v devient petit, mais ne s'annule pas.

C'est cette fonction de pénalisation exacte que nous allons utiliser. Le théorème suivant explique la terminologie : "Pénalisation exacte".

<u>Théorème 1</u> : Soit x^* une solution globale ou locale du P.N.L (1)

Soient f et g continument differentiable dans un voisinage de x^*.

Supposons que les contraintes verifient la "qualification" suivante :

Le systeme $\begin{cases} \nabla g_j(\bar{x})^T y \leq 0 & g_j(\bar{x}) = 0,\ j = 1..m \\ \nabla f(\bar{x})^T y \leq 0 \end{cases}$

n'a pas de solution $y \in \mathbb{R}^n$ non nulle en tout point $\bar{x}$, minimum local du P.N.L (1)

Alors :

$$\exists\, V_{x^*},\ \exists\, \bar{r} > 0,\ \forall r < \bar{r},\ P(r, x^*) \leq P(r, x) \quad \forall x \in V_{x^*}$$

<u>Commentaires</u>

Ce théorème indique donc que lorsqu'on a trouvé une solution du P.N.L (1) on a aussi trouvé une solution au problème de pénalisation.

Pour démontrer ce théorème, on choisit d'abord le voisinage V_{x^*} en fonction des contraintes non actives en x^*. Ayant ce voisinage, on démontre par l'absurde qu'il répond à la thèse.

Nier la thèse engendrera une suite de points non admissibles. On démontrera que cette suite converge vers un point $\bar{x}$ admissible qui est en plus minimum local du P.N.L (1).

On terminera en démontrant que ce point $\bar{x}$ ne vérifie pas la qualification des contraintes

démonstration :

première partie : choix de V_{x^*}

soit U_{x^*} le voisinage sur lequel x^* minimise $f(x)$

soit W_{x^*} un voisinage sur lequel les contraintes inactives en x^* restent inactives en tous points de W_{x^*}, c'est à dire :

$$W_{x^*} \subset \left\{ x \in \mathbb{R}^n, \ g_j(x) < 0, \ g_j(x^*) < 0, \ j = 1..m \right\}$$

Ce voisinage W_{x^*} existe parce que la fonction g est continue

soit $V_{x^*} = U_{x^*} \cap W_{x^*}$.

deuxième partie : Vx^* répond à la thèse

pas 1 : exprimons que la deuxième partie est absurde

donc : $\forall \bar{v}, \exists v < \bar{v}, \exists x \in Vx^*, \quad P(v,x) < P(v,x^*) \quad (a)$

Considerons une suite $(\bar{v}^i)_{i \in N} = \left(\frac{1}{i}\right)_{i \in N}$

(a) devient :

$$\forall i, \exists v^i < \frac{1}{i}, \exists x^i \in Vx^*, \; P(v^i, x^i) < P(v^i, x^*) \; (b)$$

remarquons que la suite $(x^i)_{i \in N}$ n'est pas admissible

En effet si $\exists i \in N$, x^i est admissible,

$$P(v^i, x^i) = v^i f(x^i) \geq v^i f(x^*) = P(v^i, x^*).$$

La suite $(x^i)_{i \in N}$ étant infinie et m étant fini nous pouvons affirmer que :

$$\exists j \in \{1 \dots m\} \text{ et } L_1 \subset N \text{ tels que}$$
$$i \in L_1 \text{ entraine } g_j(x^i) > 0.$$

$\{x_i\}_{i \in L_1}$ est une sous suite de $(x_i)_{i \in N}$

nous continuerons cependant à la noter $(x_i)_{i \in N}$ Pour simplifier la suite de la démonstration, supposons que l'indice j est unique. La démonstration dans le cas où j n'est pas unique est une généralisation de celle que nous allons présenter.

pas 2 : la suite (x^i) converge vers un point admissible.

Compte tenu des remarques précédentes l'inégalité (b) peut s'exprimer sous la forme :

$$\forall i, \quad v^i(f(x^i) - f(x^*)) + g_j(x^i) < 0 \quad . \qquad (c)$$

Dans l'inégalité (c) faisons tendre i vers l'infini : nous obtenons alors :

$$\lim_{i \to \infty} v^i(f(x^i) - f(x^*)) + \lim_{i \to \infty} g_j(x^i) \leq 0 \qquad (d)$$

la fonction f est continue et Vx^* est borné le fait que $\forall i$, $x^i \in Vx^*$ entraine donc :

$$(f(x^i) - f(x^*))_{i \in N} \text{ est bornée.}$$

donc,
$$\lim_{i \to \infty} v^i(f(x^i) - f(x^*)) = 0 \qquad (e)$$

$$\text{car } 0 \leq \lim_{i \to \infty} v^i \leq \lim_{i \to \infty} \bar{v}^i = 0 \qquad (f).$$

l'égalité (e) portée dans l'inégalité (d) nous permet d'écrire :

$$\lim_{i \to \infty} g_j(x^i) \leq 0 .$$

Mais nous savons que $\forall i$, $g_j(x^i) \geq 0$
donc,
$$\lim_{i \to \infty} g_j(x^i) \geq 0 .$$

En conclusion $\lim_{i \to \infty} g_j(x^i) = 0$

la suite (x^i) ne converge pas necessairement mais elle est bornée parce que tous ces points appartiennent a V_{x^*} Donc cette suite a au moins un point d'accumulation $\bar{x}$ et il existe une sous suite de (x^i) qui converge vers $\bar{x}$ Dans ce cas, nous ne considererons plus (x^i) mais la sous suite de (x^i) qui converge vers $\bar{x}$. Par abus d'ecriture, nous noterons encore cette sous suite (x^i)

La fonction g_j est continue, nous pouvons donc dire que

$$g_j(\bar{x}) = 0 .$$

De plus, $\forall k \in \{1, \dots m\}$, $k \neq j$, on a

$$\forall i, \quad g_k(x^i) \leq 0 .$$

Donc : $\lim\limits_{i \to \infty} g_k(x^i) \leq 0$

et : $g_k(\bar{x}) \leq 0$

En résumé : $\forall k \in \{1 \dots m\}$. on a

$$g_k(\bar{x}) \leq 0 \text{ et donc } g(\bar{x}) \leq 0 .$$

<u>pas 3 : $\bar{x}$ est un minimum local du P.N.L (1) sur V_{x^*}</u>

L'inégalité (c) et le fait que $g(x^i) > 0 \quad \forall i$ nous permettent de dire que :

$$\forall i, \quad f(x^i) \leq f(x^*)$$

La continuité de f entraine alors que

$$\lim_{i \to \infty} f(x^i) = f(\bar{x}) \leq f(x^*).$$

D'autre part, nous savons que $\forall i, \; x^i \in V_{x^*}$
Donc :

$$\bar{x} = \lim_{i \to \infty} x^i \in \overline{V_{x^*}}.$$

Donc : $\quad f(\bar{x}) \geq f(x^*)$

En résumé nous pouvons affirmer que

$$f(\bar{x}) = f(x^*) \quad , \quad \bar{x} \in V_{x^*}.$$

Parce que x^* est un minimum local du P.N.L (1) sur V_{x^*}, nous pouvons dire que $\bar{x}$ est aussi un minimum local du P.N.L (1) sur V_{x^*}.

pas 4 : démontrons que $\nabla g_j(\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} = 0$

l'inégalité :

$$\forall i, \; v^i(f(x^i) - f(x^*)) + g_j(x^i) < 0 \qquad (c)$$

s'applique au vecteur x^* ; le vecteur $\bar{x}$ appartient à V_{x^*}, la continuité de f et de g nous permet d'affirmer que l'inégalité (c) reste vraie en remplaçant x^* par $\bar{x}$ pour autant que le voisinage V_{x^*} soit assez petit. Nous obtenons donc :

$$\forall i, \; v^i(f(x^i) - f(\bar{x})) + g_j(x^i) < 0$$

En développant cette dernière inégalité en série de Taylor, nous obtenons :

$$v^i(f(x^i) - f(\bar{x})) - \nabla g_j(\bar{x})^T(x^i - \bar{x}) + O(|x^i - \bar{x}|) < 0$$

donc

$$v^i \frac{(f(x^i) - f(\bar{x}))}{|x^i - \bar{x}|} - \nabla g_j(\bar{x})^T \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} + \frac{O(|x^i - \bar{x}|)}{|x^i - \bar{x}|} < 0 \quad (l)$$

Nous savons que f est k. Lipschitz continue. donc le rapport $\frac{f(x^i) - f(\bar{x})}{|x^i - \bar{x}|}$ est borné par k $\forall i$

Dans l'inégalité (l) faisons tendre i vers l'infini nous obtenons :

$$\lim_{i \to \infty} \frac{v^i(f(x^i) - f(\bar{x}))}{|x^i - \bar{x}|} + \lim_{i \to \infty} \nabla g_j(\bar{x})^T \frac{(x^i - \bar{x})}{|x^i - \bar{x}|}$$

$$+ \lim_{i \to \infty} \frac{O(|x^i - \bar{x}|)}{|x^i - \bar{x}|} \leq 0 . \qquad (l)$$

Nous savons que $\lim_{i \to \infty} v^i = 0$.

Nous savons aussi que $\lim_{i \to \infty} x^i = \bar{x}$.

l'inégalité (l) devient donc :

$$- \nabla g_j (\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} \leq 0$$

c'est à dire : $\nabla g_j (\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} \geq 0$

D'autre part, $\forall i$, $g_j (x^i) > 0$

Ceci entraine que $- \nabla g_j (\bar{x})^T \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} + \frac{O(|x^i - \bar{x}|)}{|x^i - \bar{x}|} > 0 \; (p)$

Dans (p), faisons tendre i vers l'infini, nous obtenons :

$$- \nabla g_j (\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} \geq 0$$

c'est à dire : $\nabla g_j (\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} \leq 0$

En résumé : $\nabla g_j (\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} = 0$

pas 5 : $\bar{x}$ ne vérifie pas la qualification des contraintes.

Compte tenu des remarques du pas 1, nous pouvons dire que :

α) $\forall j \,;\, (\forall i,\ g_j(x^i) > 0)$ entraine :

$$g_j(\bar{x}) = 0 \quad \text{et} \quad \nabla g_j(\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} = 0$$

Ceci est une possibilité pour les contraintes actives en $\bar{x}$, il faut envisager deux autres cas : $\forall i,\ g_j(x^i) = 0$ et $g_j(\bar{x}) = 0$

$\forall i,\ g_j(x^i) < 0$. et $g_j(\bar{x}) = 0$

Notons que les autres cas – où la suite $(g_j(x^i))$ n'est pas de signes constant – peuvent se ramener à un des deux cas ci-dessus en considérant une sous suite de (x^i).

supposons donc $\forall i,\ g_j(x^i) = 0$. :

par un argument similaire à celui développé au pas précédent, nous pouvons dire :

β) $\forall j : (\forall i,\ g_j(x^i) = 0)$ entraine :

$$g_j(\bar{x}) = 0 \quad \text{et} \quad \nabla g_j(\bar{x})^T \lim_{i \to \infty} \frac{(x^i - \bar{x})}{|x^i - \bar{x}|} = 0$$

Envisageons maintenant le cas où :

$$g_j(\bar{x})=0 \quad \text{et} \quad \forall i,\ g_j(x^i)<0.$$

Nous avons donc : $g_j(x^i)-g_j(\bar{x})<0$.

En procédant comme au pas précédent nous pourrions demontrer que

$\gamma)$: $\forall j : (\forall i,\ g_j(x^i)<0)$ et $g_j(\bar{x})=0$

entraine : $\nabla g_j(\bar{x})^T \lim\limits_{i\to\infty} \dfrac{(x^i-\bar{x})}{|x^i-\bar{x}|} \leq 0$.

Nous savons, que $\forall i,\ f(x^i)<f(\bar{x})$

et donc :

$\delta)$: $\nabla f(\bar{x})^T \lim\limits_{i\to\infty} \dfrac{(x^i-\bar{x})}{|x^i-\bar{x}|} \leq 0$.

Nous pouvons donc dire que le systeme :

$$\begin{cases} \nabla g_j(\bar{x})^T y \leq 0 & j=1..m \text{ et } g_j(\bar{x})=0 \\ \nabla f(\bar{x})^T y \leq 0 \end{cases}$$

admet la solution $y=\lim\limits_{i\to\infty} \dfrac{(x^i-\bar{x})}{|x^i-\bar{x}|}$

Le vecteur y étant non nul, nous pouvon dire que $\bar{x}$ ne verifie pas la qualification des contraintes.

#.

Commentaires :

Ce théorème n'est pas fondamental pour nous, parce que notre but est de trouver une solution du P.N.L. (1) :

$$\min_{x} \ f(x)$$
$$\text{sous les contraintes : } g(x) \leq 0$$

Lorsqu'on a trouver une solution, peu nous importe qu'elle vérifie d'autres propriétés. Nous préférerions determiner une condition suffisante d'optimalité, c'est l'objet du théorème suivant:

Théorème 2 : Si il existe $\bar{x}$ appartenant a Vx^* tel que $g(\bar{x}) \leq 0$

Alors:

la reciproque du théorème 1 est vrai

Commentaires

Ce théorème affirme donc que lorsqu'on a minimisé la fonction de pénalisation exacte $P(v, x)$ pour $v < \bar{v}$, on a aussi trouvé une solution pour le P.N.L (1) C'est de ce théo=
=rème que l'on peut justifier le terme "penalisation exacte".

La démonstration s'effectue en deux temps
Le premier pas consiste à montrer que x^* minimise effectivement f pour tous les points admissibles du P.N.L (1) situés dans le voisinage V_{x^*} de x^*
Dans un deuxième temps, on démontre que le point x^* est effectivement admissible pour le P.N.L (1).
La démonstration s'achève alors naturellement puisque x^* point admissible minimise f pour tous les points admissibles d'un certain voisinage de x^*
Le deuxième pas de la démonstration se fait par l'ab=
=surde, c'est à dire que l'on suppose que le point x^* n'est pas admissible. On arrive alors à une contradiction relative au paramètre $\bar{v}$

<u>démonstration</u>

Nous savons que $\forall\, v < \bar{v}$, $P(v, x^*) \leq P(v, x)$ $\forall\, x \in V_{x^*}$.

<u>pas 1</u> : $\forall\, x^0$ admissible de V_{x^*}, $f(x^*) \leq f(x^0)$

Nous savons par l'hypothèse qu'il y a au moins un x^0 admissible.
Soit x^0 admissible quelconque de V_{x^*}

$$P(v, x^0) = v f(x^0)$$

$$v f(x^0) = P(v, x^0) \geq P(v, x^*) = \underbrace{v f(x^*) + \sum_{j=1}^{m} g_j(x^*)_+}_{\geq\, v f(x^*)}$$

donc $v f(x^0) \geq v f(x^*)$ et $f(x^0) \geq f(x^*)$.

<u>pas 3</u> : <u>x^* est admissible</u>

supposons que cela soit faux

donc $\sum_{j=1}^{m} g_j(x^*)_+ = a > 0$

$\forall \bar{x}$ admissible de Vx^*, nous avons

$$v f(\bar{x}) \geq v f(x^*) + a \qquad (1)$$

$\bar{x}$ étant admissible, permet de dire que

$$v f(\bar{x}) \geq v f(x^*)$$

c'est à dire : $v f(\bar{x}) = v f(x^*) + v\varepsilon \qquad (2)$

portons (2) dans (1) nous obtenons :

$$v f(x^*) + v\varepsilon \geq v f(x^*) + a$$

et donc $v\varepsilon \geq a \qquad (3)$

Nous avons que $\varepsilon > 0$, parce que si $\varepsilon \leq 0$

(3) donne alors :

$$0 \geq v\varepsilon \geq a > 0 .$$

Dans l'inégalité (3) nous pouvons donc diviser les deux membres par ε. Nous obtenons :

$$v \geq a / \varepsilon \qquad (4)$$

Les hypothèses du théorème nous assurent que v prend n'importe quelles valeurs entre 0 et $\bar{v}$
En appliquant l'hypothèse : " x^* non admessible" nous arrivons à la conclusion que v doit être supe=
=rieur à une certaine valeur $a|\bar{\varepsilon}>0$.
Donc la valeur $v=\frac{1}{2}\min\{\bar{v}, a|\bar{\varepsilon}\}$ est admise par les hypothèses du théoreme mais elle n'est pas admise par l'hypothèse " x^* pas admessible"

#

B : Points stationnaires et Fonction d'optimalité

Dans cette section, nous allons developper quelques shèmes théoriques qui seront utiles pour établir la convergence de l'algorithme. Ces shemes permetteront des raccourcis.

En particulier, nous allons définir un certain type de fonction d'optimalité et un certain type de points stationnaires de $\mathcal{T}(0,x)$

Le coté interessant est que la fonction d'optimalité s'annule pour les points stationnaires et pour les points de Kuhn Tucker du P.N.L (1).

Propriété 1 : $\mathcal{T}(0,x)$ a un minimum global qui est zero et qui est atteint en tout point admissible du P.N.L (1)

Théorème 1 : Soit $\hat{x}$ un minimum non contraint de $P(0,\hat{x})$

Alors :

$\hat{x}$ est une solution locale de :

$$(2) \quad \text{minimiser} \sum_{j \in I^+(\hat{x})} g_j(x)$$

$$\text{sous les contraintes } g_j(x) \leq 0 \ , j \in I^{0-}(\hat{x})$$

Si en plus $g_{I(\hat{x})}$ satisfait une qualification des contraintes en $\hat{x}$

Alors :

$$\exists\ \hat{w} \in \mathbb{R}^m \text{ tel que } \nabla g(\hat{x})\,\hat{w} = 0 .$$

$$\text{et } \hat{w}_j \begin{cases} = 1 & j \in I^+(\hat{x}) \\ \geq 0 & j \in I(\hat{x}) \\ = 0 & j \in I^-(\hat{x}) \end{cases}$$

rappelons que

$$I(\hat{x}) = \{ j \in \bar{m},\ g_j(\hat{x}) = 0 \}$$

$$I^+(\hat{x}) = \{ j \in \bar{m},\ g_j(\hat{x}) > 0 \} \quad I^{-0}(\hat{x}) = I^-(\hat{x}) \cup I(\hat{x})$$

$$I^-(\hat{x}) = \{ j \in \bar{m},\ g_j(\hat{x}) < 0 \} \quad I^{+0}(\hat{x}) = I^+(\hat{x}) \cup I(\hat{x})$$

Commentaires

Ce théorème est interessant parce qu'il nous permettra d'introduire une notion de points stationnaires. Nous verrons en effet qu'un point verifiant la deuxième partie de la thèse – par point on entend ici un couple $(\hat{x}, \hat{w})$ de $\mathbb{R}^{n+m}$ – est un point stationnaire

Nous sommes obligés d'introduire ces points parce que l'algorithme que nous allons présenter dans ce chapitre convergera idéalement vers un point de Kuhn Tucker ; mais il pourra converger vers un point qui n'est pas un point de Kuhn Tucker. On demontrera alors qu'il s'agit d'un point stationnaire

Démonstration

Si $\hat{x}$ est un minimum non contraint de $P(0,x)$ on a que :

$$\exists\ B(\hat{x},\hat{a}) \text{ telle que } \forall x \in B(\hat{x},\hat{a})$$

on a que :

$$\sum_{j \in I^+(\hat{x})} g_j(\hat{x}) = P(0,\hat{x})$$

$$\leq P(0,x)$$

$$= \sum_{j \in I^+(\hat{x})} g_j(x) \quad \text{si } g_j(x) \leq 0 \text{ pour } j \in I^{-0}(\hat{x})$$

La deuxieme partie de la thèse est équivalente à dire que $(\hat{x}, \hat{w}_{I^{-0}(\hat{x})})$ est un point de Kuhn Tucker du probleme (2)
Ceci est vrai puisque $\hat{x}$ est optimal pour le problème (2) et qu'une qualification des contraintes est satisfaite en $\hat{x}$

#

définition : $(\hat{x}, \hat{w})$ de R^{n+m} est un <u>point station-naire</u> de $P(0, x)$ si et seulement si

$\nabla g(\hat{x})\,\hat{w} = 0$

et $\hat{w}_j = 1$ si $j \in I^+(\hat{x})$

≥ 0 si $j \in I(\hat{x})$

$= 0$ si $j \in I^-(\hat{x})$

<u>Commentaire</u>

Remarquons qu'il n'y a pas de rapport direct entre les points stationnaires de $P(0, x)$ et les points de kuhn - tucker du P.N.L (1)

Un point de kuhn Tucker est admissible pour le P.N.L (1), un point stationnaire ne l'est pas toujours.

Un point stationnaire verifie $\nabla g(\hat{x})\hat{w} = 0$ un point de kuhn Tucker ne le verifie que si $\nabla f(\hat{x}) = 0$.

Les points stationnaires ne sont introduits que pour caractériser une certaine possibilité de terminaison de l'algorithme

propriété 2 : si $g_1 \cdot\cdot g_m$ sont convexes.
si $(\hat{x}, \hat{w})$ est un point stationnaire de $\mathcal{P}(0, x)$
si $g_{I(\hat{x})}$ vérifie une qualification des contraintes en $\hat{x}$
Alors :
$\hat{x}$ est admissible pour le P.N.L (1)

Commentaires

L'algorithme que nous allons proposer se terminera, comme nous le verrons plus tard, en un point de Kuhn Tucker ou en un point stationnaire non admissible. Si les hypothèses de cette propriété sont vérifiées, il n'y aura pas de points stationnaires non admissibles et l'algorithme se terminera donc en un point de Kuhn Tucker.

Démonstration

supposons que $\hat{x}$ ne soit pas admissible pour le P.N.L (1)

donc $\exists j_0 \in \{1 \cdot\cdot m\}$ tel que $g_{j_0}(\hat{x}) > 0$

Le problème (2) étant convexe, $\hat{x}$ est un minimum global de ce problème

soit x^* un point admissible du P.N.L (1)

nous avons alors:

$$0 < \sum_{j \in I^+(\hat{x})} g_j(\hat{x}) \leq \sum_{j \in I^+(\hat{x})} g_j(x^*) = 0 .$$

$\#$

<u>définition</u> :

la fonction d'optimalité φ_N

soit $\|\cdot\|_N$ une norme quelconque de R^n

on définit la fonction $\varphi_N(v, x, w) =$

$\| v \nabla f(x) + \nabla g(x) w \|_N - g(x)^T w + \sum_{j=1}^{m} g_j(x)^+$

$\forall v \geq 0$, $\forall x \in R^n$, $\forall w \in T(x)$

rappelons que $T(x) = \{ w \in R^m,\ w \geq 0,\ w_j \leq 1,\ j \in I^+(x) \}$.

<u>Théorème 2</u> :

1) $\varphi_N(v, x, w) \geq 0 \quad \forall v \geq 0,\ \forall x \in R^n,\ \forall w \in T(x)$

2) si $v > 0$, si x est admissible et si $w \in T(x)$

alors :

$\varphi_N(v, x, w) = 0 \iff (x, v^{-1} w)$ est un point de Kuhn Tucker du P.N.L (1)

3) si $v = 0$, si x n'est pas admissible et si $w \in T(x)$

alors :

$\varphi_N(v, x, w) = 0 \iff (x, w)$ est un point stationnaire de $P(0, x)$

<u>Commentaires</u> :

On remarque dès lors l'intérêt de la fonction d'optimalité, elle caractérise d'une certaine manière les points de Kuhn Tucker du P.N.L (1) et les points stationnaires de $P(0, x)$. Dans l'algorithme, nous allons construire une certaine suite et nous montrerons que la fonction d'optimalité appliquée à cette suite converge vers zero.
On établira alors le lien avec ce théorème pour dégager une certaine convergence de l'algorithme considéré

Démonstration

$\varphi_N(v, x, w)$ peut s'écrire sous la forme :

$$\underbrace{\| v \nabla f(x) + \nabla g(x)\, w \|_N}_{(a)} + \underbrace{\sum_{j \in I^-(x)} (-w_j\, g_j(x))}_{(b)} + \underbrace{\sum_{j \in I^+(x)} (1 - w_j)\, g_j(x)}_{(c)}$$

(a) : est un terme non négatif

(b) $\quad j \in I^-(x)$ donc : $w_j \geq 0$ donc : $-w_j\, g_j(x) \geq 0$
$\quad g_j(x) < 0$

(c) $\quad j \in I^+(x)$ donc : $1 - w_j \geq 0$ donc : $(1 - w_j)\, g_j(x) \geq 0$
$\quad g_j(x) > 0$

$\varphi_N(v, x, w)$ peut donc s'écrire comme une somme de termes non négatifs ; la première partie de la thèse est donc démontrée.

supposons $v \neq 0$ et $\varphi_N(v, x, w) = 0$. et x admissible
l'expression de $\varphi_N(v, x, w)$ ci dessus permettent de dire que : $v \nabla f(x) + \nabla g(x)\, w = 0$.

$$w_j\, g_j(x) = 0 \qquad \forall j \in I^-(x)$$

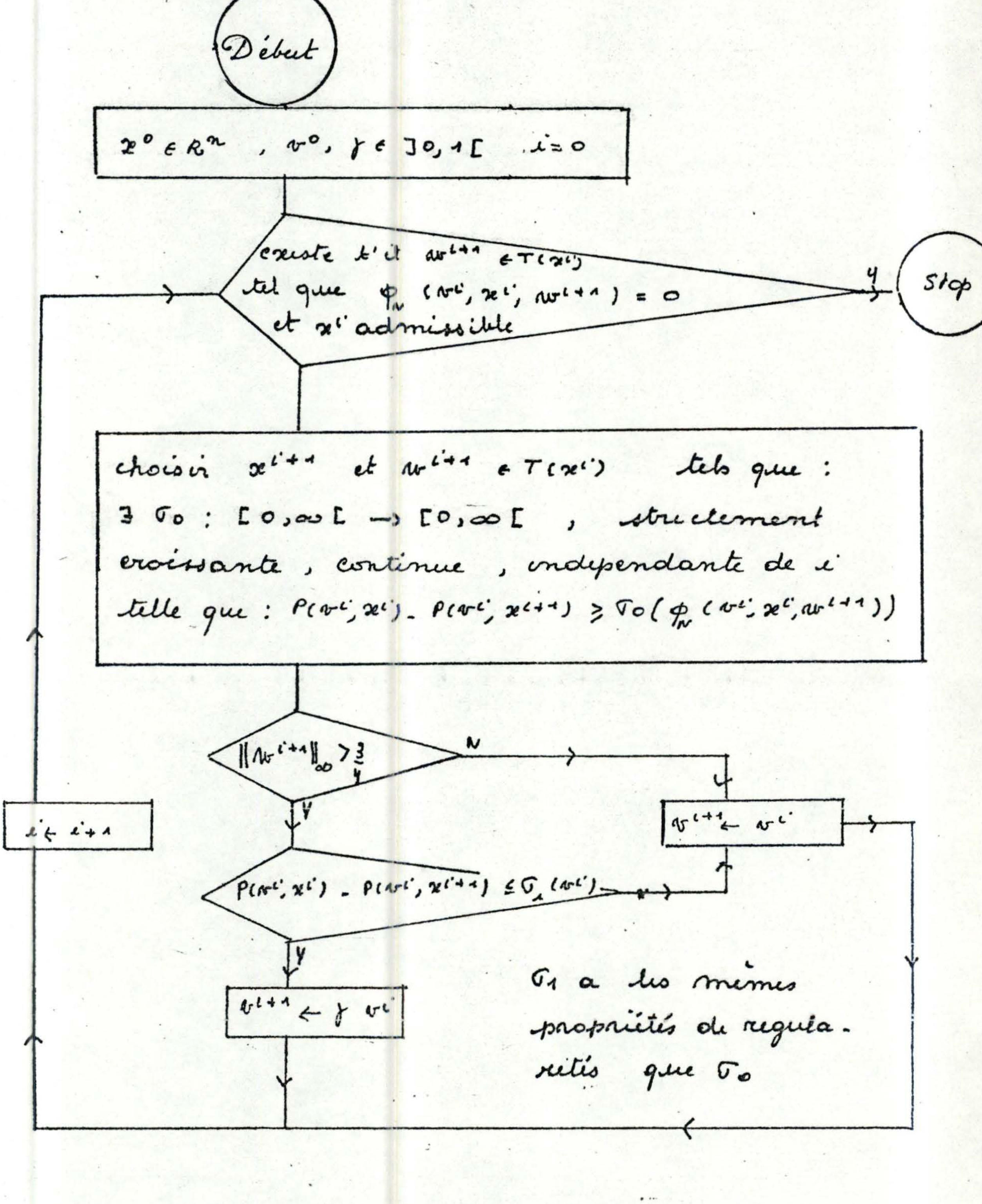
ces deux égalités, le fait que $w \in T(x)$ et que x soit admissible permettent de dire que $(x, w v^{-1})$ est un point de Kuhn Tucker du P.N.L (1)

si nous supposons maintenant que $v \neq 0$ x admissible et $(x, w v^{-1})$ point de kuhn Tucker du P.N.L (1), les conditions de kuhn Tucker permettent de dire que $\varphi_N(v, x, w) = 0$.
La démonstration de la deuxième partie de la thèse est ainsi terminée.

Pour démontrer la troisième partie de la thèse on applique le même raisonnement que celui développé pour la deuxième partie avec les hy= =pothèses dont on dispose

#

C Algorithme de pénalisation exacte globalement convergent



Commentaires

L'algorithme engendre essentiellement trois quantités

- La suite (x^i) qui convergera vers un point de Kuhn Tucker du P.N.L (1) ou vers un point stationnaire de $P(0,x)$. La partie 3 assure une direction de descente pour la fonction de pénalisation et donc aussi pour f.

- La suite (w^i) qui engendrera en fait les multiplicateurs de Kuhn Tucker ou les "multiplicateurs des points stationnaires".

- La suite (v^i) qui convergera par construction (4) vers un nombre $\bar{v}$. Si $\bar{v}>0$ nous obtiendrons ce que nous cherchons c'est à dire un point de Kuhn Tucker du P.N.L (1). Si au contraire nous obtenons $\bar{v}=0$, nous obtiendrons des points stationnaires non admissibles de $P(0,x)$. A vrai dire, cela ne nous intéressera pas beaucoup.

D : Le théorème de convergence

Théorème 1 : si la suite (x^i) est finie, elle s'arrête en un point de Kuhn Tucker du P.N.L (1)

si la suite (x^i) est infinie mais bornée si les gradients des contraintes actives sont linéairement indépendants en tout point d'accumulation de (x^i) noté $\bar{x}$

Alors :

Il existe $\bar{u} \in R^m$ tel que :

soit :

1) $(\bar{x}, \bar{u})$ satisfait les conditions de Kuhn Tucker pour le P.N.L (1)

dans ce cas $\lim_{i \to \infty} v^i = \bar{v} > 0$.

soit :

2) $(\bar{x}, \bar{u})$ est un point stationnaire non admissible de $P(0, x)$

dans ce cas $\lim_{i \to \infty} v^i = \bar{v} = 0$

Commentaires :

Ce théorème donne un résultat déjà annoncé
L'algorithme converge vers un point de Kuhn Tucker dans le cas $\bar{v} > 0$. Sinon il converge vers un point stationnaire non admissible.

Commentaires

Si $g_1 \dots g_m$ sont convexes nous pouvons dire que chacun des points d'accumulation de (x^i) fournira un point de Kuhn Tucker pour le P.N.L (1). En effet la propriete de la section B nous permet de dire qu'aucun point d'accumulation ne fournira un point stationnaire non admissible puisque chaque point d'accumulation de (x^i) verifie une qualification des contraintes.

Avant d'entamer la demonstration du théorème nous allons utiliser une serie de lemmes qui serviront pour démontrer le théorème

---

Lemme 1 : Les suites $(\phi_N(v^i, x^i, w^{i+1}))_{i=0}^{\infty}$

et

$$(P(v^i, x^i))_{i=0}^{\infty}$$

convergent et $\lim\limits_{i \to \infty} \phi_N(v^i, x^i, w^{i+1}) = 0$.

Commentaires

Si on savait que $\hat{x}$ est admissible et $\bar{v} > 0$ [respectivement $\hat{x}$ non admissible et $\bar{v} = 0$], ce lemme suffirait pour conclure que $(\hat{x}, \hat{W} \cdot \bar{v}^{-1})$ est un point de Kuhn Tucker ($\hat{W}$ étant la limite de la suite W^{i+1}) [respectivement, $(\hat{x}, \hat{W})$ est un point stationnaire de $P(0, x)$]

Le lemme démontré, il suffirait de voir que (w^{i+1}) converge et d'étudier l'admissibilité de $\hat{x}$ pour conclure. C'est ce que nous ferons après le lemme 1

Pour démontrer ce lemme, nous allons construire une serie :

$$\left\{ \sum_{i=0}^{j} \left(P(v^i; x^i) - P(v^i; x^{i+1}) \right) \right\}_{j=0}^{\infty}$$

Grâce à la méthode de construction de l'algorithme on peut démontrer que cette seri est croissante et les hypothèses du théorème nous permettrons d'affirmer qu'elle est bornée et donc convergente. Donc le terme général de cette seri converge vers zéro La méthode de construction nous permettra encore de conclure : $\varphi_N(v^i; x^i; w^{i+1}) = 0$

<u>Démonstration</u>

première partie : $\lim\limits_{i \to \infty} \varphi_N(v^i, x^i, w^{i+1}) = 0$

---

soit $\{ i_k = k = 0, 1 \dots \} = \{ i \in N, v^i = \gamma v^{i-1} \}$

Cet ensemble en fait est fini ou infini

<u>pas 1 :</u> $\underline{P(v^{i_k}, x^{i_k}) - P(v^{i_{k-1}}, x^{i_k}) = (\gamma - 1)\, v^{i_{k-1}} f(x^{i_k})}$

$$P(v^{i_k}, x^{i_k}) - P(v^{i_{k-1}}, x^{i_k}) =$$

$$= v^{i_k} f(x^{i_k}) + \sum_{j=1}^{m} g_j(x^{i_k})_+ - v^{i_{k-1}} f(x^{i_k}) - \sum_{j=1}^{m} g_j(x^{i_k})_+$$

$$= (\gamma - 1)\, v^{i_{k-1}} f(x^{i_k}) \qquad \text{parce que } v^{i_k} = \gamma\, v^{i_{k-1}}$$

<u>pas 2 :</u> $\underline{\forall j \in \mathbb{N}_0 \,,\ \sum_{i=0}^{j} (P(v^i, x^i) - P(v^i, x^{i+1}))}$

$$\le P(v^0, x^0) - P(v^j, x^{j+1}) + \frac{\gamma}{1-\gamma}\, v^0 \sup(|f(x^i)|)_{i=0}^{\infty}$$

$$\sum_{i=0}^{j} (P(v^i, x^i) - P(v^i, x^{i+1}))$$

$$= P(v^0, x^0) - P(v^j, x^{j+1}) + \sum_{1 \le i_k \le j} (\gamma - 1)\, v^{i_{k-1}} f(x^{i_k})$$

Ceci vient du pas 1 et du fait que les termes d'indices differents de i_k se suppriment 2 à 2 car $i \neq i_k$ entraine $v^{i-1} = v^i$

et donc $P(v^{i-1}, x^i) = P(v^i, x^i)$

$$v^{i_{k-1}} = \gamma\, v^{i_{(k-2)}} = \gamma^2\, v^{i_{(k-2)}} = \gamma^k v^0$$

donc :

$$\sum_{i=0}^{j} \left(P(v^i, x^i) - P(v^i, x^{i+1}) \right)$$

$$= P(v^0, x^0) - P(v^j, x^{j+1}) + \sum_{1 \leq k \leq j} (\gamma - 1)\, \gamma^k\, v^0\, f(x^k)$$

$$\leq P(v^0, x^0) + P(v^j, x^{j+1}) + \frac{\gamma}{1-\gamma}\, v^0 \sup \{ f(x^i) \}_{i=0}^{\infty}$$

remarquons que $\sup \{ f(x^i) \}_{i=0}^{\infty} < \infty$ parce que (x^i) est borné et que f est continue

pas 3 : la suite $\left\{ \sum_{i=0}^{j} \left(P(v^i, x^i) - P(v^i, x^{i+1}) \right) \right\}_{j=0}^{\infty}$ est croissante

Considérons l'élément j et $j+1$
pour obtenir l'élément $j+1$ de la suite, on ajoute à l'élément j le terme :

$$P(v^{j+1}, x^{j+1}) - P(v^{j+1}, x^{j+2})$$

par la construction de l'algorithme on peut dire

$$P(v^{j+1}, x^{j+1}) - P(v^{j+1}, x^{j+2}) \geq \sigma_0 \left(\phi_N (v^{j+1}, x^{j+1}, w^{j+2}) \right)$$

≥ 0 par la deuxieme partie de l'algorithme.

pas 4 : la suite $\left\{ \sum_{i=0}^{j} (P(v^i; x^i) - P(v^i; x^{i+1})) \right\}_{j=0}^{\infty}$ est bornée

la suite (x^i) est bornée, f et g sont continus

donc:

$$\left\{ f(x^i) \right\}_{i=0}^{\infty} \quad \text{et} \quad \left\{ P(v^j, x^{j+1}) \right\}_{j=0}^{\infty}$$

sont bornées

Et l'expression du pas 2 permettent d'affirmer que:

$$\left\{ \sum_{i=0}^{j} (P(v^i; x^i) - P(v^i; x^{i+1})) \right\}_{j=1}^{\infty} \text{ est bornée}$$

pas 5 : conclusion

Une suite croissante bornée converge et donc, son terme général – puisqu'ici cette suite est en fait une série – converge vers zéro

c'est à dire: $\lim_{i \to \infty} P(v^i, x^i) - P(v^i, x^{i+1}) = 0$.

Mais $P(v^i, x^i) - P(v^i; x^{i+1}) \geq \sigma_0 (\varphi_N (v^i, x^i, w^{i+1}) \geq 0$

Donc en passant à la limite

$$0 = \lim_{i \to \infty} (P(v^i, x^i) - P(v^i, x^{i+1})) \geq \lim_{i \to \infty} \sigma_0 (\varphi_N (v^i, x^i, w^{i+1}) \geq 0$$

donc $\lim_{i \to \infty} \sigma_0 (\varphi_N (v^i, x^i, w^{i+1}) = 0$

$$= \sigma_0 \left(\lim_{i \to \infty} (\varphi_N (v^i, x^i, w^{i+1})) \right) = 0$$

$\Rightarrow \lim_{i \to \infty} (\varphi_N (v^i, x^i, w^{i+1})) = 0$ par les propriétés de σ_0

deuxième partie : $(P(v^i, x^i))_{i=0}^{\infty}$ converge

Dans la démonstration du pas 2 on trouvait :

$$P(v^J, x^{J+1}) = -\sum_{i=0}^{J} (P(v^i, x^i) - P(v^i; x^{i+1})) + P(v^0, x^0)$$

$$+ \sum_{1 \le i_k \le J} (\gamma - 1)\, \gamma^k\, v^0 f(x^{i_k})$$

$f(x^{i_k})$ est borné et $\sum_k (\gamma^k)$ est une serie absolument convergente

$\left(\sum_{i=0}^{J} (P(v^i, x^i) - P(v^i, x^{i+1}))\right)_{J=0}^{\infty}$ est une serie convergente.

Ceci nous permet de dire que le membre de droite de l'egalité ci-dessus converge, on peut donc dire que le membre de gauche converge

donc la suite $(P(v^J, x^{J+1}))_{J=0}^{\infty}$ converge

La premiere partie affirme que

$$\lim_{i \to \infty} (P(v^i, x^i) - P(v^i, x^{i+1})) = 0$$

Donc $\lim_{i \to \infty} (P(v^i, x^i))$ existe aussi et

cette limite egale $\lim_{i \to \infty} P(v^i; x^{i+1})$

#

Lemme 2 :

Il existe une boule ouverte $B(\hat{x}, \hat{a})$ telle que

si $L(\hat{x}) = \{ i,\ x^i \in B(\hat{x}, \hat{a}) \}$

alors :

$$\left\{ W^{l+1}_{I^+(\hat{x})} \right\}_{i \in L(\hat{x})} \text{ converge vers } \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\left\{ W^{l+1}_{I^-(\hat{x})} \right\}_{i \in L(\hat{x})} \text{ converge vers } \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\left\{ W^{l+1}_{I(\hat{x})} \right\}_{i \in L(\hat{x})} \text{ est bornée}$$

rappelons que
$$I^+(\hat{x}) = \{ i \in \overline{m} ,\ g_i(\hat{x}) > 0 \}.$$
$$I^-(\hat{x}) = \{ i \in \overline{m} ,\ g_i(\hat{x}) < 0 \}$$
$$I(\hat{x}) = \{ i \in \overline{m} ,\ g_i(\hat{x}) = 0 \}.$$

Quand il n'y aura pas de confusions possibles nous noterons :
$$I^+ = I^+(\hat{x})$$
$$I^- = I^-(\hat{x})$$
$$I = I(\hat{x})$$

Commentaires

Ce lemme nous fournit une indication sur le comportement de la suite $\{W^i\}$; en fait cette suite fournira les " multiplicateurs stationnaires " dans le cas $\bar{v} = 0$ et les multiplicateurs de Kuhn Tucker dans le cas $\bar{v} > 0$. Ce lemme sera un outil précieux pour démontrer l'admissibilité éventuelle du point d'accumulation $\hat{x}$

En effet, si nous avons que : $\forall j \in \bar{m}$, $\hat{w}_j$ n'est pas égal à 1, alors $\hat{x}$ sera admissible

Il est à remarquer que $\hat{x}$ est un point d'accumulation de (x^i), donc $L(\hat{x})$ est infini

Pour démontrer ce théoreme, on va choisir la boule $B(\hat{x}, \hat{a})$ en fonction des contraintes. Ensuite, on minorera la fonction d'optimalité $\varphi_x(v^i, x^i, w^{i+1})$ par une autre fonction où n'interviennent plus x^i mais $\hat{x}$. Cette quantité minorante est en fait une somme de termes non négatifs qui devront s'annuler grâce au lemme 1. Le choix de $B(\hat{x}, \hat{a})$ et l'introduction d'une fonction $A(x)$ nous permettront de conclure.

<u>démonstration</u>

<u>pas 1</u> : choix de la boule $B(\hat{x}, \hat{a})$

soit $A(x) = \min\limits_{y} \{ \| \nabla g_I(x)^T y \| ,\ \|y\| = 1 ,\ y \geq 0 \}$

par les hypothèses du théorème : $\forall y$, $\| \nabla g_I(\hat{x})^T y \| \neq 0$.

L'ensemble $\{ y ,\ \|y\| = 1 ,\ y \geq 0 \}$ est compact

et donc $A(\hat{x}) > 0$

On choisit $B(\hat{x}, \hat{\alpha})$ telle que

$\forall x$ appartenant à $B(\hat{x}, \hat{\alpha})$, $\quad A(x) > \frac{1}{2} A(\hat{x})$

$$g_{I^-}(x) < \frac{1}{2} g_{I^-}(\hat{x})$$

$$g_{I^+}(x) > \frac{1}{2} g_{I^+}(\hat{x})$$

Le choix de cette boule est possible, parce que les fonctions A, g_{I^-}, g_{I^+} sont des fonctions continues de x

pas 2 : si $i \in L(\hat{x})$ alors $\varphi_N(v^i, x^i, w^{i+1}) \geq$

$\| v^i \nabla f(x^i) + \nabla g(x^i) w^{i+1} \|_N - \sum_{j \in I^-} (w^{i+1})_j g_j(x^i) + \sum_{j \in I^+} (1 - w_j^{i+1}) g_j(x^i)$

En effet :

$$\varphi_N(v^i, x^i, w^{i+1}) = \| v^i \nabla f(x^i) + \nabla g(x^i) w^{i+1} \|_N$$

$$- \sum_{I^+} g_j(x^i) w_j^{i+1} - \sum_{I^-} g_j(x^i) w_j^{i+1} - \sum_{I} g_j(x^i) w_j^{i+1}$$

$$+ \sum_{I^+} g_j(x^i)_+ + \sum_{I^-} g_j(x^i)_+ + \sum_{I} g_j(x^i)_+ .$$

le cinquième terme :

Le choix de $B(\hat{x}, \hat{\alpha})$ permet de dire que si $x^i \in B(\hat{x}, \hat{\alpha})$ alors $g_j(x^i) > 0 \quad \forall\, j \in I^+$

donc :

$$\sum_{I^+} g_j(x^i)_+ = \sum_{I^+} g_j(x^i)$$

le sixième terme :

Le choix de $B(\hat{x}, \hat{\alpha})$ permet encore de dire que si $x^i \in B(\hat{x}, \hat{\alpha})$ alors $g_j(x^i) < 0 \quad \forall j \in I^-$

donc

$$\sum_{I^-} g_j(x^i)_+ = 0 .$$

le quatrième et le septième terme :

si $j \in I$ et $g_j(x^i)_+ \neq 0$

c'est que $g_j(x^i) > 0$ et donc $j \in I^+(x^i)$

donc $w_j^{l+1} \leq 1$ car $w^{l+1} \in T(x^i)$

donc : $(1 - w_j^{l+1})\, g_j(x^i) \geq 0$.

si $j \in I$ et $g_j(x^i)_+ = 0$.

c'est que $g_j(x^i) \leq 0$

donc $- w_j^{l+1}\, g_j(x^i) \geq 0$.

Compte tenu de ce qui précède, en faisant les mises en evidence necessaires, on arrive à la conclusion du pas 2.

pas 3 : propriétés de $L(\bar{x})$

par le pas 2, $\phi_N(v^i, x^i, w^{l+1}) \geq$

$$\| v^i \nabla f(x^i) + \nabla g(x^i)\, w^{l+1} \|_N - \sum_{j \in I^-} w_j^{l+1} g_j(x^i) + \sum_{j \in I^+} (1 - w_j^{l+1})\, g_j(x^i)$$

En appliquant les propriétés de $B(\hat{v}, \hat{\alpha})$, si $i \in L(\bar{x})$ nous pourrons majorer cette dernière quantité par :

$$\underbrace{\| v^i \nabla f(x^i) + \nabla g(x^i)\, w^{l+1} \|_N}\ \underbrace{- \frac{1}{2} \sum_{j \in I^-} w_j^{l+1} g_j(\hat{x})}\ \underbrace{+ \frac{1}{2} \sum_{j \in I^+} (1 - w_j^{l+1})\, g_j(\bar{x})}$$

chacun des trois termes de cette somme sont non négatifs pas 4 conclusion

Le lemme 1 nous permet de dire que

$$\lim_{i\to\infty} \varphi_N(v^i, x^i, w^{i+1}) = 0 .$$

Nous pourons donc dire que

$$\lim_{\substack{i\to\infty \\ i\in L(\hat{x})}} \| v^i \nabla f(x^i) + \nabla g(x^i) w^{i+1} \|_N = 0$$

$$\lim_{\substack{i\to\infty \\ i\in L(\hat{x})}} \sum_{j\in I^-} (w^{i+1})_j \, g_j(\hat{x}) = 0$$

$$\lim_{\substack{i\to\infty \\ i\in L(\hat{x})}} \sum_{j\in I^+} (1 - w_j^{i+1}) \, g_j(\hat{x}) = 0 .$$

- Etudions les consequences de $\lim\limits_{i\in L(\hat{x})} \sum\limits_{j\in I^-} (w^{i+1})_j \, g_j(\hat{x}) = 0$

Cela entraine d'abord que

$$\forall j \in I^-, \lim_{i\in L(\hat{x})} w_j^{i+1} \, g_j(\hat{x}) = 0$$

donc $\forall j\in I^- \cdot g_j(\hat{x}) \lim\limits_{i\in L(\hat{x})} w_j^{i+1} = 0$

comme $g_j(\hat{x}) \neq 0 \quad \forall j \in I^-$,

Nous pourons affirmer que

$$\forall j \in I^-, \quad \lim_{i\in L(\hat{x})} w_j^{i+1} = 0$$

ou encore $\lim\limits_{i\in L(\hat{x})} w_{I^-}^{i+1} = 0 .$

- Etudions maintenant le fait que :

$$\lim_{i \in L(\hat{x})} \left\{ \sum_{j \in I^+} (1 - w^{l+1})_j \, g_j(\hat{x}) \right\} = 0$$

donc $\forall j \in I^+$, $\lim\limits_{i \in L(\hat{x})} (1 - w^{l+1})_j \, g_j(\hat{x}) = 0$

et donc $\forall j \in I^+$ $g_j(\hat{x}) \lim\limits_{i \in L(\hat{x})} (1 - w^{l+1})_j = 0$

Comme $\forall j \in I^+$, $g_j(\hat{x}) > 0$, nous pouvons dire que

$$\lim_{i \in L(\hat{x})} (1 - w^{l+1})_j = 0 \quad \forall j \in I^+$$

ou encore $\lim\limits_{i \in L(\hat{x})} \left\{ w^{l+1}_{I^+} \right\} = 1$

- Il reste un dernier cas à étudier :

$$\lim_{i \in L(\hat{x})} \| \nabla g(x^i) \, w^{l+1} + v^i \, \nabla f(x^i) \| = 0 .$$

(x^i) est bornée, donc $(v^i \nabla f(x^i))$ est bornée parce que f, ∇f sont des fonctions continues de x. Nous pouvons donc conclure que $\| \nabla g(x^i) \, w^{l+1} \|$ est bornée.

Nous savons que $\lim\limits_{i \in L(\hat{x})} w^{l+1}_{I^+} = 1$

$$\lim_{i \in L(\hat{x})} w^{l+1}_{I^-} = 0 .$$

Comme $(\nabla g(x^i))$ est bornée (car ∇g est continue) nous pouvons conclure que $\left\{ \| \nabla g_I(x^i) \, w^{l+1}_I \| \right\}_{i \in L(\hat{x})}$ est bornée par définition de $A(x^i)$,

$$\| \nabla g_I (x^i) w_I^{l+1} \| \geq A(x^i) \| w_I^{l+1} \| \qquad i \in L(\hat{x})$$

$$\geq \frac{1}{2} A(\hat{x}) \| w_I^{l+1} \| \qquad i \in L(\hat{x})$$

Nous pouvons donc conclure que la suite

$$\left\{ A(\hat{x}) \| w_I^{l+1} \| \right\}_{i \in L(\hat{x})} \text{ est bornée}$$

Comme $A(\hat{x}) > 0$, nous pouvons affirmer que

$$\left\{ w_I^{l+1} \right\}_{i \in L(\hat{x})} \text{ est bornée.} \qquad \#$$

Démonstration du Théorème

Commentaire

Pour démontrer le théorème on résume d'abord ce qui est à notre disposition. En d'autres mots on peut déterminer une sous suite d'indices appartenant à $L_1 \subseteq L(\hat{x}) \subseteq \mathbb{N}$ telle que

$$\lim_{i \in L_1} \phi_N (v^i, x^i, w^{i+1}) = 0 = \phi(\bar{v}, \hat{x}, \hat{w})$$

$$\lim_{i \in L_1} (w^{i+1}) = \hat{w} \geq 0$$

$$\hat{w}_j = 1 \qquad j \in I^+(\hat{x})$$

$$\hat{w}_j = 0 \qquad j \in I^-(\hat{x})$$

Nous envisagerons tout d'abord le cas $\bar{v} > 0$. Le théorème des fonctions d'optimalité nous dit que $(\hat{x}, \hat{w}\,\bar{v}^{-1})$ est un point de Kuhn Tucker si et seulement si $\hat{x}$ est admissible. Pour cela, il suffirait de voir que $\|\hat{w}\|_\infty < 1$ et donc par le lemme 2, $\hat{x}$ serait admissible.

Pour démontrer cela, nous utiliserons la quatrième partie de l'algorithme, celle où on ajuste le paramètre v^i. On remarquera que ce paramètre ne peut être ajusté qu'un nombre fini de fois et donc la condition " $\|w^{i+1}\|_\infty > \frac{3}{4}$ " ne sera vrai qu'un nombre fini de fois, la deuxième condition étant toujours vérifiée. Nous aurons donc que $\|\hat{w}\|_\infty < \frac{3}{4}$

Commentaires $\bar{v}=0$

Toujours en appliquant le théorème de la section B sur les fonctions d'optimalité, il suffirait de prouver que $\bar{x}$ n'est pas admissible. Cela se démontre par l'absurde.

On introduit alors un autre point d'accumulation de x^i que l'on note $\tilde{x}$, il est le point de convergence de la sous suite dont les indices sont ceux où l'on ajuste le paramètre v^i. A partir de $\hat{x}$ on avait construit une sous suite de (x^i) qui convergeait vers $\hat{x}$ et telle que (v^{i+1}) convergeait vers $\hat{v}$. On va suivre ici le même chemin, c'est à dire construire une sous suite (x^{i_k}) convergeant vers $\tilde{x}$ telle que (v^{i_k+1}) convergera vers $\tilde{v}$. On démontrera alors que $\tilde{x}$ est admissible et que $\tilde{v}>0$ et cela en utilisant les lemmes 1 et 2. On parviendra ensuite à contredire l'indépendance des gradients en $\tilde{x}$

démonstration

pas 1 : conclusions des lemmes précédents

Le lemme 2 assure que la suite $\{w^{i+1}\}_{i\in L(\hat{x})}$ est bornée

On peut donc en extraire une sous suite convergente :

$$\{w^{i+1}\}_{i\in L_1 \subseteq L(\hat{x})} \text{ converge vers } \hat{w}\in R^m$$

$$\{x^i\}_{i\in L_1} \text{ converge vers } \hat{x}$$

par le lemme 2 : $\hat{w}\geq 0$ (limite de "nombres" non negatifs)

$$\lim_{i\in L_1} w_j^{i+1} = 1 \quad \forall j\in I^+(\hat{x})$$

$$\text{donc } \hat{w}_j = 1 \quad \forall j\in I^+(\hat{x})$$

Nous pouvons donc dire que $\hat{w}\in T(\hat{x})$

Le lemme 1 affirme que $\lim_{i\to\infty} \varphi_N(v^i, x^i, w^{i+1}) = 0$

donc $\lim_{i\in L_1} \varphi_N(v^i, x^i, w^{i+1}) = 0$

et par continuité de φ_N nous affirmons :

$$\varphi_N(\bar{v}, \hat{x}, \hat{w}) = 0$$

<u>pas 2:</u> <u>supposons $\bar{v} > 0$</u>

donc $\lim_{i \to \infty} v^i > 0$ et $\{i, v^i = \gamma v^{i-1}\}$ est fini.

Nous allons essayer de demontrer que $\hat{x}$ est admissible.

dire que $\{i, v^i = \gamma v^{i-1}\}$ est equivalent à dire :

$$\{i, P(v^i, x^i) - P(v^i, x^{i+1}) \leq \sigma_1(v^i)\} \text{ est fini.}$$

ou

$$\{i, \|w^{i+1}\|_\infty > \frac{3}{4}\} \text{ est fini.}$$

Le lemme 1 pas 5 disait que.

$$\lim_{i \to \infty} \{P(v^i, x^i) - P(v^i, x^{i+1})\} = 0$$

Nous savons aussi, par les propriétés de σ_1 que :

$$\lim_{i \to \infty} \sigma_1(v^i) = \sigma_1(\bar{v}) > 0.$$

donc $\exists i_0$, $i > i_0 \Rightarrow P(v^i, x^i) - P(v^i, x^{i+1}) \leq \sigma_1(v^i)$

donc $\{i, P(v^i, x^i) - P(v^i, x^{i+1}) \leq \sigma_1(v^i)\}$ est infini.

On en conclut que :

$$\{i, \|w^{i+1}\|_\infty > \frac{3}{4}\} \text{ est fini}$$

et donc $\exists \hat{i}$, $i > \hat{i} \Rightarrow \|w^{i+1}\|_\infty \leq \frac{3}{4}$

nous pourrons donc affirmer que $\|\hat{w}\|_\infty \leq \frac{3}{4}$

donc $\forall j, \ \hat{w}_j \leq \frac{3}{4}$

et donc $\forall j, \ \hat{w}_j \neq 1$. et $I^+(\hat{x}) = \phi$

$\hat{x}$ est donc admissible et le theorème des fonction d'optimalité permettent d'affirmer que $(\hat{x}, \hat{w} . \bar{v}^{-1})$ est un point de kuhn Tucker pour le P.N.L (1) (puisque $\bar{v} > 0$, $\hat{x}$ admis-sible, $\varphi_N(\bar{v}, \hat{x}, \hat{w}) = 0$.

pas 3 : supposons $\bar{v} = 0$

nous allons essayer de demontrer que $\hat{x}$ est non admissible, $\bar{v} = 0$ et $\varphi_N(\bar{v}, \hat{x}, \hat{w}) = 0$ nous permettront de dire que $(\hat{x}, \hat{w})$ est un point stationnaire de $P(0, x)$

Supposons que $\hat{x}$ est admissible et montrons qu'il existe un point d'accumulation de (x^i) qui ne verifie pas l'independance lineaire des gradients des contraintes actives.

a) $\lim\limits_{i \to \infty} P(v^i, x^i) = 0$.

par le lemme 1, $\lim\limits_{i \to \infty} P(v^i, x^i)$ existe.

$\hat{x}$ est un point d'accumulation de (x^i)
P est continue en x et (x^i) est bornée
donc :

$$\lim_{i \to \infty} P(v^i, x^i) = \cdots P(0, \hat{x}) = 0$$

car $\hat{x}$ est admissible.

b) choix de $\tilde{x}$ point d'accumulation de (x^i)

soit $L_2 = \{ i ,\ v^{i+1} = \gamma v^i \}$

$\{x^i\}_{i \in L_2}$ est une suite bornée, on peut donc en extraire une sous suite convergente :

$$\{x^i\}_{i \in L_3 \subset L_2} \text{ converge vers } \tilde{x} \in \mathbb{R}^n$$

Le lemme 2 affirme que $\{w^{i+1}\}_{i \in L_3}$ est bornée

et donc il existe un sous ensemble $L_4 \subset L_3$ tel que :

$$\{w^{i+1}\}_{i \in L_4 \subset L_3} \text{ converge vers } \tilde{w} \in \mathbb{R}^m$$

$$\{x^i\}_{i \in L_4 \subset L_3} \text{ converge vers } \tilde{x} \in \mathbb{R}^n$$

c) recherche d'une contradiction

$\forall i \in L_4$, $v^{i+1} = \gamma v^i$, donc $\forall i \in L_4$, $\|w^{i+1}\|_\infty > \frac{3}{4}$

donc $\exists j_0 \in \bar{m}$, et un ensemble infini $L_5 \subset L_4$ tels que

$$w^{i+1}_{j_0} \geq \frac{3}{4} \qquad \forall i \in L_5$$

donc : $\tilde{w}_{j_0} = \lim\limits_{i \in L_5} w^{i+1}_{j_0} = \lim\limits_{i \in L_4} w^{i+1}_{j_0} \geq \frac{3}{4}$

ceci entraine que

$$\tilde{w} > 0$$

$(\hat{x}, \tilde{w})$ ont les mêmes propriétés que $(\hat{x}, \hat{w})$

donc $\phi_N(0, \hat{x}, \tilde{w}) = 0$. et $\nabla g(\hat{x})\tilde{w} = 0$

par le lemme 2 : $\forall j \in I^-(\hat{x})$, $\tilde{w}_j = 0$

d'autre part : $P(0, \hat{x}) = \lim\limits_{i \to \infty} P(v^i, x^i) = 0$

ce qui entraine que $\sum\limits_{j=0}^{m} g_j(\hat{x})_+ = 0$

et donc $I^+(\hat{x}) = \phi$

en conclusion : $\left.\begin{array}{l} \tilde{w}_{I^-(\hat{x})} = 0 \\ I^+(\hat{x}) = \phi \\ \tilde{w} > 0 \\ \nabla g(\hat{x})\tilde{w} = 0 \end{array}\right\}$ donc : $\tilde{w}_{I(\hat{x})} > 0$

et $\nabla g_{I(\hat{x})} \tilde{w}_{I(\hat{x})} = 0$

Ceci contredit l'hypothèse d'independance linéaire des gradients des contraintes actives en $\hat{x}$

$\hat{x}$ est non admissible et le theorème des fonctions d'optimalités permettent d'affirmer que $(\hat{x}, \hat{w})$ est un point stationnaire non admissible de $P(0, x)$.

#

C. Determination de (x^{i+1}, w^{i+1})

L'algorithme presenté dans ce chapitre dit de choisir le couple (x^{i+1}, w^{i+1}) tel que:

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq \sigma_0 (\varphi_N (x^i, v^i, w^{i+1}))$$

Il ne précise cependant pas la manière de déterminer (x^{i+1}, w^{i+1}) c'est ce que nous allons faire maintenant

<u>La méthode est la suivante:</u>

$$x^{i+1} = x^i + \lambda^i t^i$$

où t^i est la solution d'un programme convexe à contraintes lineaires

où λ^i est determiné par une procedure d'Armijo

w^{i+1} sera lui aussi déterminé par l'intermediaire du programme convexe à contraintes lineaires. Nous determinerons un problème lineaire qui sera equi=valent au premier. w^{i+1} sera alors le vecteur des multiplicateurs de Kuhn Tucker associés à certaines contraintes détermination de (x^{i+1}, w^{i+1})

soit t^i solution du problème :

$$\min_{t \in \mathbb{R}^n} \; v^i \nabla f(x^i)^T t + \sum_{j \in I^+(x^i)} \left((g_j(x^i) + \nabla g_j(x^i)^T t)_+ - g_j(x^i)_+ \right)$$

sous les contraintes :

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq 0 \qquad j = 1 .. m \text{ et } g_j(x^i) \leq 0$$

$$- d^i \leq t_j \leq d^i \qquad j = 1 .. n \, , \; d^i \in [d_L, d_U] \, , \; d_L, d_U > 0 .$$

le problème est équivalent au programme linéaire :

$$\min_{t, r_j} \; v^i \nabla f(x^i)^T t + \sum_{j \in I^+(x^i)} (r_j - g_j(x^i))$$

sous les contraintes :

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq 0 \qquad j = 1 .. m \text{ et } g_j(x^i) \leq 0 .$$

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq r_j \qquad j = 1 .. m \text{ et } g_j(x^i) > 0$$

$$r_j \geq 0$$

$$- d^i \leq t_j \leq d^i \qquad j = 1 .. n \, , \; d^i \in [d_L, d_U] .$$

soit θ^i la valeur de la fonction à minimiser dans le programme linéaire.

soit w_j^{i+1} le multiplicateur de Kuhn Tucker associé à la contrainte :

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq 0 \qquad \text{si } g_j(x^i) \leq 0$$

$$\nabla g_j(x^i)^T t + g_j(x^i) > 0 \qquad \text{si } g_j(x^i) > 0$$

Soit $\lambda^i = \max\ (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16} \ldots\ldots)$

satisfaisant à :

$$P(v^i, x^i) - P(v^i, x^i + \lambda^i t^i) \geq -\frac{1}{3} (\lambda^i)^2 \theta^i$$

On appelle une telle procedure, une procedure d'ARMIJO

Soit $x^{i+1} = x^i + \lambda^i t^i$

Commentaires.

Nous venons de determiner une suite (x^i, w^i), nous devons voir qu'elle verifie bien la condition imposée par l'algorithme proposé, c'est à dire :

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq \sigma_0 (\varphi_w (v^i, x^i, w^{i+1}))$$

C'est la seule propriété à verifier, c'est en effet le seul critère dont on dispose pour engendrer la suite (x^i, w^i)

<u>définition</u> :

soit $F : \mathbb{R}^n \to \mathbb{R}$

$\Delta F : \mathbb{R}^{2n} \to \mathbb{R}$ est une <u>fonction tangente</u> de F si et seulement si :

1) $\forall (x,t) \in \mathbb{R}^n \times \mathbb{R}^n$, ΔF est convexe en t $\forall t$, $\forall x$ fixé.

2) $F(x+t) - F(x) = \Delta F(x,t) + \beta(x,t) \|t\|$

avec $\beta(x,0) = 0$ et $\lim_{t \to 0} \beta(x,t) = 0$

<u>Commentaires</u>

Pour des raisons de simplicité nous noterons :

$$\Delta P(v, x, t) = v \nabla f(x)^T t + \sum_{j=1}^{m} ((g_j(x) + \nabla g_j(x)^T t)_+ - g_j(x)_+)$$

On remarque que $\Delta P(v, x, t)$ n'est pas sans rapport avec la fonction à minimiser dans le programme convexe à contraintes linéaires proposé pour déterminer t^i. Nous verrons aussi que la fonction à minimiser et $\Delta P(v, x, t)$ coïncident à l'optimum.

La notation "$\Delta P(v, x, t)$" va se justifier dans le premier lemme que nous allons étudier. Nous allons voir que $\Delta P(v, x, t)$ est une fonction tangente de $P(v, x)$.

<u>Lemme 1</u> : $\Delta P(v, x, t)$ est une fonction convexe de t $\forall v > 0$, $\forall x \in \mathbb{R}^n$ fixés.
De plus, il existe $k_{21} > 0$ tel que
$\forall x \in \mathbb{R}^n$, $\forall t \in \mathbb{R}^n$, $\forall v > 0$, on ait :
$|P(v, x+t) - P(v, x) - \Delta_2 P(v, x, t)| \leq K_{21} \|t\|^2$

<u>Commentaires</u>

Pour demontrer ce lemme, nous allons proceder en deux étapes

- demontrer la convexité.
- demontrer la formule d'Approximation
Nous utiliserons pour cela, le fait que les fonctions ∇f, $\nabla g_1 \dots \nabla g_m$ sont k Lipschitz continues et aussi la proprieté :

$$\forall a, b \in \mathbb{R} \quad , \quad |b - a| \geq |b_+ - a_+| \qquad (1)$$

<u>démonstration</u>

première partie : $\Delta P(v, x, t)$ est convexe en t

Il faut remarquer qu'une combinaison non négative de fonctions convexes est encore une fonction convexe il suffirait donc de demontrer que :

$(g_j(x) + \nabla g_j(x)^T t)_+$ est une fonction convexe de t puisque les autres termes intervenant dans $\Delta P(v, x, t)$ sont des fonctions convexes de t.

soient donc $\bar{t}$ et $\hat{t}$ 2 points de $\mathbb{R}^n$ et λ entre 0 et 1

$$g_j(x) + \nabla g_j(x)^T(\lambda\hat{t} + (1-\lambda)\bar{t})$$

$$= \lambda(g_j(x) + \nabla g_j(x)^T\hat{t}) + (1-\lambda)(g_j(x) + \nabla g_j(x)\bar{t})$$

ce qui entraine que

$$(g_j(x) + \nabla g_j(x)^T(\lambda\hat{t} + (1-\lambda)\bar{t}))_+$$
$$\leq$$
$$\lambda(g_j(x) + \nabla g_j(x)^T\hat{t})_+ + (1-\lambda)(g_j(x) + \nabla g_j(x)\bar{t})_+ .$$

Ceci entraine la convexité de $\Delta P(v,x,t)$ en' t

deuxième partie : la formule d'approximation

Il faut donc demontrer que :

$$|P(v,x+t) - P(v,x) - \Delta_e P(v,x,t)| \leq K_{21} \|t\|^2$$

rappelons que 1) $P(v,x) = v f(x) + \sum_{j\in I^+(x)} g_j(x)$

$$= v f(x) + \sum_{j=1}^{m} g_j(x)_+ .$$

2) ∇f, $\nabla g_1 \dots \nabla g_m$ sont L. Lipchitz continues .
et donc :

$$|f(x+t) - f(x) - \nabla f(x)^T t| \leq \frac{1}{2} K \|t\|^2. \qquad (2)$$
$$|g(x+t) - g(x) - \nabla g(x)^T t| \leq \frac{1}{2} K \|t\|^2. \qquad (3)$$

$j = 1 \dots m$

Nous allons maintenant examiner

$P(v, x+t) - P(v,x) - \Delta P(v,x,t)$, cette différence fait intervenir les fonctions f et g

$$|P(v, x+t) - P(v,x) - \Delta P(v,x,t)| =$$

$$\Big| v f(x+t) + \sum_{j=1}^{m} g_j(x+t)_+ - v f(x) - \sum_{j=1}^{m} g_j(x)_+$$

$$- v \nabla f(x)^T t - \sum_{j=1}^{m} \left((g_j(x) + \nabla g_j(x)^T t)_+ - g_j(x)_+ \right) \Big|$$

$$\leq |v f(x+t) - v f(x) - v \nabla f(x)^T t|$$

$$+ \sum_{j=1}^{m} | g_j(x+t)_+ - g_j(x)_+ - (g_j(x) + \nabla g_j(x)^T t)_+ + g_j(x)_+ |$$

$$= |v f(x+t) - v f(x) - v \nabla f(x)^T t|$$

$$+ \sum_{j=1}^{m} | g_j(x+t)_+ - (g_j(x) + \nabla g_j(x)^T t)_+ |$$

$$\leq |v f(x+t) - v f(x) - v \nabla f(x)^T t|$$

$$+ \sum_{j=1}^{m} | g_j(x+t) - g_j(x) - \nabla g_j(x)^T t | \qquad \text{par (1)}$$

$$\leq \frac{1}{2} (m+1) k \|t\|^2 \qquad \text{par } (2-3)$$

et donc $\quad |P(v, x+t) - P(v,x) - \Delta P(v,x,t)|$

$$\leq \frac{1}{2} (m+1) k \|t\|^2 .$$

#

Commentaire

Nous venons de déterminer une fonction tangente de $P(v, x)$. Ceci est intéressant parce que cette fonction tangente est à mettre en rapport avec la fonction à minimiser dans le programme convexe à contraintes linéaires utilisé pour déterminer t^i

Lemme 2 : si t^i est la solution du programme convexe à contraintes linéaire utilisé pour déterminer t^i
si θ^i est la valeur de la fonction à minimiser à l'optimum
alors :

$$\Delta P(v^i, x^i, t^i) = \theta^i \leq 0.$$

démonstration

première partie : $\theta^i = \Delta P(v^i, x^i, t^i)$

$$\theta^i = v^i \nabla f(x^i)^T t^i + \sum_{j \in I^+(x^i)} \left((g_j(x^i) + \nabla g_j(x^i)^T t)_+ - g_j(x^i)_+ \right)$$

avec : $\nabla g_j(x^i)^T t^i + g_j(x^i) \leq 0 \quad j = 1 \dots m$ et $g_j(x^i) \leq 0$

$-d^i \leq t_j \leq d^i \quad d^i \in [d_L, d_U]$.

donc θ^i :

$$v^i \nabla f(x^i)^T t + \sum_{j \in I^+(x^i)} ((g_j(x^i) + \nabla g_j(x^i)^T t)_+ - g_j(x^i)_+)$$

$$+ \sum_{j \in I^{-0}(x^i)} ((g_j(x^i) + \nabla g_j(x^i)^T t)_+ - g_j(x^i)_+)$$

$$= v^i \nabla f(x^i)^T t + \sum_{j=1}^{m} ((g_j(x^i) + \nabla g_j(x^i)^T t)_+ - g_j(x^i)_+)$$

$$= \Delta P(v^i, x^i, t^i)$$

deuxieme partie $\theta^i \leq 0$

$t = 0$ verifie les contraintes qui se transforment en : $g_j(x^i) \leq 0$ si $g_j(x^i) \leq 0$ $j = 1..m$

$- d^i \leq 0 \leq d^i$

la fonction a minimiser se transforme en

$$0 + \sum_{j \in I^+(x^i)} ((g_j(x^i)_+ + 0)_+ - g_j(x^i)_+) = 0.$$

puisque t^i est l'optimum nous concluons :

$$v^i \nabla f(x^i)^T t^i + \sum_{j \in I^+(x^i)} ((g_j(x^i) + \nabla g_j(x^i)^T t^i)_+ - g_j(x^i)_+)$$

$$= \theta^i \leq 0$$

#

Commentaires

Nous allons maintenant étudier un lemme nous renseignant sur le comportement de la suite (t^i)
En fait, nous allons voir que la suite (t^i) est bornée.

Lemme 3 : Il existe une constante $k_{22} > 0$ telle que $\|t^i\| \leq k_{22} \quad \forall i \in N$

démonstration

La contrainte $\quad -d_i \leq t^i \leq d^i$

entraîne : $\|t^i\| \leq \sqrt{n}\, d^i$

$\|t^i\| \leq \sqrt{n}\, d_M$.

#

Commentaires

Nous allons maintenant essayer de nous rapprocher de notre but, c'est à dire :

$$P(v^i; x^i) - P(v^i, x^{i+1}) \geq \sigma_0 (\varphi_N (v^i; x^i, w^{i+1}))$$

Pour cela, nous allons procéder en deux étapes
d'abord démontrer que:

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq \sigma_0 (-\theta^i)$$

ensuite démontrer que

$$-\theta^i \geq \varphi_N (v^i, x^i, w^{i+1}).$$

Lemme 4 : Il existe une fonction continue strictement croissante σ_0 telle que

$\sigma_0 : [0,\infty[\longrightarrow [0,\infty[$

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq \sigma_0(-\theta^i) \quad \forall i$$

indépendante de i

Commentaires

Dans la démonstration, nous traiterons d'abord le cas $\theta^i = 0$. et nous verrons qu'il est trivial. Ensuite nous verrons que la procédure d'Armijo a un sens, c'est à dire qu'il existe au moins une valeur λ vérifiant la condition d'Armijo, pour cela, nous utiliserons principalement le lemme 1.

Remarquons que :

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq -\frac{1}{2} (\lambda^i)^2 \theta^i$$

ne nous suffit pas pour conclure parce que la fonction $\sigma_0(-\theta)$ serait telle que

$$\sigma_0(-\theta^i) = \frac{-1}{2} (\lambda^i)^2 \theta^i ,$$

elle dépendrait donc de "i"

En utilisant le lemme 1 et la condition d'Armijo, nous parviendrons à minorer λ^i par une valeur ne dépendant que de θ^i, la démonstration s'achève alors naturellement

démonstration

pas 1 : soit $\theta^i = 0$

si $\theta^i = 0$, alors $t^i = 0$

En effet le problème convexe à contraintes linéaires utilisé pour déterminer t^i est equivalent comme nous l'avons vu à un programme linéaire. si $\theta^i = 0$, la solution de ce problème linéaire sera zero.

soit $\lambda^i = 1$ et regardons si nous verifions la condition d'Armijo.

$$
\begin{aligned}
P(v^i, x^i) - P(v^i, x^{i+1}) &= P(v^i, x^i) - P(v^i, x^i + 0 \cdot 1) \\
&= 0 \\
&= \sigma_0(0) \\
&= \sigma_0(-\theta^i)
\end{aligned}
$$

Donc si $\theta^i = 0$, la procedure d'Armijo a un sens et le lemme 4 est verifié quelle que soit la fonction σ_0 pourvu que $\sigma_0(0) = 0$.

Ce cas est trivial et peu intéressant parce qu'il ne nous permet pas d'avancer. Nous supposerons donc $\theta^i < 0$.

pas 2 : $\exists \lambda$ tel que : $P(v^i, x^i) - P(v^i, x^i + \lambda t^i) \geq \frac{-1}{2} \lambda^2 \theta^i$

le lemme 1 assure que : $\forall v > 0, \forall x \in R^n, \forall t \in R^n$

$$|P(v, x+t) - P(v, x) - \Delta P(v, x, t)| \leq K_{21} \|t\|^2$$

donc:

$$P(v, x+t) - P(v, x) - \Delta P(v, x, t) \leq K_{21} \|t\|^2$$

donc

$$P(v^i, x^i + \lambda t^i) - P(v^i, x^i) - \Delta P(v^i, x^i, \lambda t^i) \leq K_{21} \lambda^2 \|t^i\|^2$$

et par la convexité de $\Delta P(v, x, t)$ en t :

$$P(v^i, x^i + \lambda t^i) - P(v^i, x^i) - \lambda \Delta P(v^i, x^i, \lambda^i) \leq K_{21} \lambda^2 \|t^i\|^2$$

et donc

$$-P(v^i, x^i + \lambda t^i) + P(v^i, x^i) \geq -\lambda \Delta P(v^i, x^i, t^i) - K_{21} \lambda^2 \|t^i\|^2 \quad (2)$$

soit λ tel que

$$0 < \lambda < \frac{-\Delta P(v^i, x^i, t^i)}{K_{21} \|t^i\|^2 - \frac{1}{2} \theta^i} \quad (1)$$

on peut determiner un λ verifiant (1)

parce que : $-\Delta P(v^i, x^i, t^i) = -\theta^i > 0$.

$K_{21} \|t^i\|^2 - \frac{1}{2} \theta^i > 0$.

(1) entraine que :

$$\lambda K_{21} \|t^i\|^2 - \frac{1}{2} \lambda \theta^i < -\Delta P(v^i, x^i, t^i)$$

et en multipliant les 2 membres par λ.

$$-\lambda \Delta P(v^i, x^i, t^i) - \lambda^2 K_{21} \|t^i\|^2 \geq \frac{-1}{2} \lambda^2 \theta^i$$

portant cette inégalité dans (2), nous obtenons :

$$-P(v^i, x^i + \lambda t^i) + P(v^i, x^i) \geq \frac{-1}{2} \lambda^2 \theta^i$$

Le pas 2 assure l'existence d'une valeur λ répondant à la condition d'Armijo

<u>pas 3 :</u> <u>soit $\lambda^i \leq \frac{1}{2}$, alors $\lambda^i \geq \frac{-\theta^i}{4K_2 + K_{22}^2}$</u>

$\lambda^i \leq \frac{1}{2}$, l'inégalité proposée par Armijo n'est donc pas vérifiée pour $\lambda = 2\lambda^i$

et donc :

$$P(v^i, x^i) - P(v^i, x^i + 2\lambda^i t^i) < -\frac{1}{2}(2\lambda^i)^2 \theta^i$$

$$\begin{aligned} P(v^i, x^i) - P(v^i, x^i + 2\lambda^i t^i) &< -2(\lambda^i)^2 \theta^i \\ &= -2(\lambda^i)^2 \Delta . P(v^i; x^i; t^i) \\ &\leq -\lambda^i \Delta P(v^i; x^i, 2\lambda^i t^i) \end{aligned}$$

et donc :

$$P(v^i; x^i + 2\lambda^i t^i) - P(v^i, x^i) \geq \lambda^i \Delta P(v^i; x^i; 2\lambda^i t^i)$$

Le pas 2 (2) nous permet de dire :

$$P(v^i, x^i) - P(v^i, x^i + \lambda t^i) \geq -\lambda \Delta P(v^i, x^i; t^i) - K_{21} \lambda^2 \|t^i\|^2$$

et donc

$$P(v^i, x^i) - P(v^i, x^i + 2\lambda^i t^i) \geq -\Delta P(v^i, x^i, 2\lambda^i t^i) - 4K_{21} \lambda^{i2} \|t^i\|^2$$

ceci entraine :

$$\begin{aligned} 4K_{21} (\lambda^i)^2 \|t^i\|^2 \geq & -\Delta P(v^i, x^i, 2\lambda^i t^i) \\ & - (P(v^i; x^i) - P(v^i; x^i + 2\lambda^i t^i)) \end{aligned}$$

Cette dernière affirmation, et le fait que

$$P(v^i, x^i + 2\lambda^i t^i) - P(v^i; x^i) \geq \lambda^i \Delta P(v^i, x^i, 2\lambda^i t^i)$$

nous permettent de dire :

$$4K_{21} (\lambda^i)^2 \|t^i\|^2 \geq -(1-\lambda^i) \Delta P(v^i, x^i, 2\lambda^i t^i)$$

et donc

$$-2\lambda^i (1-\lambda^i) \Delta P(v^i, x^i, t^i) \leq 4K_{21} (\lambda^i)^2 \|t^i\|^2 .$$

comme $\lambda^i \leq \frac{1}{2}$, $-\lambda^i \geqslant -\frac{1}{2}$

et donc $-2\lambda^i \geqslant -1$

$1-\lambda^i \geqslant \frac{1}{2}$

$-2\lambda^i(1-\lambda^i) \geqslant -\frac{1}{2}$

de plus nous avons : $(\lambda^i)^2 \leq \frac{1}{2}\lambda^i$

et donc : $4K_{21}(\lambda^i)^2 \leq 2K_{21}\lambda^i$

Nous obtenons donc :

$-\frac{1}{2}\Delta P(v^i; x^i, t^i) \leq 2K_{21}\lambda^i \|t^i\|^2$

donc $\lambda^i \geqslant \dfrac{-\Delta P(v^i; x^i, t^i)}{4K_{21}\|t^i\|^2} \geqslant \dfrac{-\Delta P(v^i; x^i, t^i)}{4K_{21}K_{22}^2} = \dfrac{-\theta^i}{4K_{21}K_{22}^2}$

ces dernieres affirmation se justifient par les lemmes 2, 3

et donc si $\lambda^i < 1$, $\lambda^i \geqslant \dfrac{-\theta^i}{4K_{21}K_{22}^2}$

<u>pas 4 : conclusion</u>

Quel que soit λ^i ($\lambda^i < 1$ ou $\lambda^i = 1$), nous obtenons grace au pas 3 :

$$\lambda^i \geqslant \min\left\{1, \frac{-\theta^i}{4K_{21}K_{22}^2}\right\}$$

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geqslant -\frac{1}{2}(\lambda^i)^2\theta^i .$$

$$\geqslant -\frac{1}{2}\left\{\min\left\{1, \frac{-\theta^i}{4K_{21}K_{22}^2}\right\}\right\}^2 \theta^i$$

nous obtenons que : $P(v^i; x^i) - P(v^i; x^{i+1}) \geqslant \sigma_0(-\theta^i)$

avec $\sigma_0(-\theta) = +\frac{1}{2}(-\theta)\left\{\min\left\{1, \dfrac{-\theta}{4K_{21}K_{22}^2}\right\}\right\}^2$

On peut voir que σ_0 est continue et strictement croissante sur $[0,\infty[$.

#

Commentaires

Nous allons maintenant etablir le rapport entre $(-\theta^i)$ et $\varphi_N(v^i, x^i, w^{i+1})$

Lemme 5 :

Il existe $w^{i+1} \in T(x^i)$ tel que :

$$-\theta^i \geq \varphi_N(v^i, x^i, w^{i+1})$$

avec $\|\cdot\|_N = d_i \|\cdot\|_1$.

Commentaires

Pour demontrer ce théorème nous allons utiliser le programme linéaire equivalent au programme convexe à contraints lineairs utilisé pour determiner t^i. Nous utiliserons surtout le programme dual de ce programme lineaire. La programmation lineaire nous a appris que les variables duales optimales sont les multiplicateurs de Kuhn Tucker du programme primal à l'optimum. Plus precisement nous verrons que w_j^{i+1} est le multiplicateur associé à :

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq 0 \qquad j = 1 \ldots m \text{ et } g_j(x^i) \leq 0$$

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq r_j \qquad j = 1 \ldots m \text{ et } g_j(x^i) > 0$$

$$r_j \geq 0$$

Ce resultat avait été annoncé précédemment

Remarquons que

$-\theta^i \geq \varphi_N(v^i, x^i, w^{i+1})$ entraine par le lemme 4

que : $P(v^i, x^i) - P(v^i; x^{i+1}) \geq \sigma_0 (\varphi_N(v^i, x^i, w^{i+1}))$

Ce qu'on voulait demontrer sera donc démontré.

Démonstration

Le problème :

$$\underset{t}{\text{minimiser}} \quad v^i \nabla f(x^i)^T t + \sum_{j \in I^+(x^i)} \left((g_j(x^i) + \nabla g_j(x^i)^T t)_+ - g_j(x^i)_+ \right)$$

$$\text{sous les contraintes : } \nabla g_j(x^i)^T t + g_j(x^i) \leq 0 \qquad j = 1 .. m \;,\; g_j(x^i) \leq 0 .$$

$$- d^i \leq t_j \leq d^i \qquad j = 1 .. n \;,\; d^i \in [d_L , d_U]$$

est equivalent au problème linéaire :

$$\underset{t, r}{\text{minimiser}} \quad v^i \nabla f(x^i)^T t + \sum_{j \in I^+(x^i)} (r_j - g_j(x^i))$$

$$\text{sous les contraintes : } \nabla g_j(x^i)^T t + g_j(x^i) \leq 0 \quad j = 1 .. m \,,\; g_j(x^i) \leq 0$$

$$\nabla g_j(x^i)^T t + g_j(x^i) \leq r_j \quad j = 1 \ldots m \,,\; g_j(x^i) > 0$$

$$r_j \geqslant 0$$

$$- d^i \leq t_j \leq d^i \qquad j = 1 .. n \,,\; d^i \in [d_L , d_U] .$$

le dual de ce problème est

$$\underset{w, p, z}{\text{maximiser}} \quad w^T g(x^i) - d^i \sum_{j=1}^{n} p_j - d^i \sum_{j=1}^{n} z_j - \sum_{j \in I^+(x^i)} g_j(x^i)$$

$$\text{sous les contraintes : } w^T \nabla g(x^i) + p_j - z_j = - v^i \nabla f(x^i)$$

$$- w_j \geqslant -1 \qquad j \in I^+(x^i)$$

$$w, p, z \geqslant 0 .$$

soient $(w^{i+1}, p^{i+1}, z^{i+1})$ solution de ce problème. Nous savons que la valeur de la fonction à maximiser dans le dual et la valeur de la fonction à minimiser dans le primal coincident à leur optimum respectif donc

$$\theta^i = (w^{i+1})^T g(x^i) - d^i \sum_{j=1}^{n} p_j^{i+1} - d^i \sum_{j=1}^{n} z_j^{i+1} - \sum_{j \in I^+(x^i)} g_j(x^i)$$

et : $(w^{i+1})^T \nabla g(x^i) + p^{i+1} - z^{i+1} = -v^i \nabla f(x^i)$

$w_j^{i+1} \leq 1 \qquad j \in I^+(x^i)$

$w^{i+1},\ p^{i+1},\ z^{i+1} \geq 0.$

donc : $-\theta^i = -(w^{i+1})^T g(x^i) + d^i \sum_{j=1}^{n} p_j^{i+1} + d^i \sum_{j=1}^{n} z_j^{i+1} + \sum_{j \in I^+(x^i)} g_j(x^i)$

$$= -(w^{i+1})^T g(x^i) + d^i \sum_{j=1}^{n} | p_j^{i+1} + z_j^{i+1} | + \sum_{j \in I^+(x^i)} g_j(x^i)$$

(car $p^{i+1},\ z^{i+1} \geq 0$)

$$\geq -(w^{i+1})^T g(x^i) + d^i \sum_{j=1}^{n} | p_j^{i+1} - z_j^{i+1} | + \sum_{j \in I^+(x^i)} g_j(x^i)$$

$$= -(w^{i+1})^T g(x^i) + d^i \| p^{i+1} - z^{i+1} \|_1 + \sum_{j \in I^+(x^i)} g_j(x^i)$$

$$= -(w^{i+1})^T g(x^i) + d^i \| v^i \nabla f(x^i) + (w^{i+1})^T \nabla g(x^i) \|_1 + \sum_{j \in I^+(x^i)} g_j(x^i)$$

$$\geq \varphi_N (v^i, x^i, w^{i+1}) \qquad \text{si} \quad \| \cdot \|_N = d_i \| \cdot \|_1$$

et donc $\quad -\theta^i \geq \varphi_N (v^i, x^i, w^{i+1})$.

#

CHAPITRE III : ALGORITHMES DE CONVERGENCE LOCALE.

A Introduction

Dans le deuxième chapitre, nous avons considéré un algorithme convergent globalement vers un point de Kuhn Tucker du PNL(1) ou vers un point stationnaire de $P(0,x)$; globalement en ce sens que : " quel que soit le point d'où l'on demarre, on a toujours la convergence de la méthode.

Dans ce chapitre, nous nous interesserons à des algorithmes locaux, c'est à dire qu'ils auront certaines propriétés de convergence pourvu que le point d'où l'on demarre soit assez proche du point $(\bar{x},\bar{u})$ où l'on veut arriver.

Nous allons principalement étudier deux types d'Algorithmes

premier type : Le premier type d'algorithme reduit le P.N.L(1) en une suite d'équations linéaires

deuxième type : Le deuxième type d'algorithme reduit le P.N.L(1) en une suite de programmes quadratiques.

Les algorithms presentés ici sont du Type Quasi Newton. Une bonne approche des methods est donnée par : J.E. Dennis, Jr. et Jorge J. Moré "Quasi - Newton methods, Motivation and Theory", SIAM Review 19 (1977), 46-89.

B : Algorithme Quasi Newton réduisant un P.N.L en une suite d'équations linéaires

1) Préliminaires

Avertissement

A partir de maintenant, nous considérerons que le P.N.L (1) possède un point $(\bar{x}, \bar{u})$ de Kuhn Tucker qui vérifie les conditions de non singularité de Jacobi, c'est à dire :

1) $\forall y \in R^n$, y non nul tel que $y^T \nabla g_j(\bar{x}) = 0$, $j \in I(\bar{x})$, $\bar{u}_j > 0$
$y^T \nabla g_j(\bar{x}) \leq 0$ $j \in I(\bar{x})$, $\bar{u}_j = 0$

on a :

$$y^T (\nabla^2 f(\bar{x}) + \sum_{j=1}^{m} \bar{v}_j \nabla^2 g_j(\bar{x})) y > 0$$

2) $\{\nabla g_j(\bar{x}) ; j \in I(\bar{x})\}$ sont linéairement indépendants

3) $\bar{v}_j > 0$ si $j \in I(\bar{x})$

Ces conditions sont aussi appelées : conditions du second ordre.

remarques

Un point de Kuhn Tucker du P.N.L (1) vérifie : $\sum_{j=1}^{m} \bar{u}_j g_j(\bar{x}) = 0$, $I^+(\bar{x}) = \phi$

donc $\forall j \in I^-(\bar{x})$, $\bar{v}_j = 0$.

La troisième condition de Jacobi est que :

si $j \in I(\bar{x})$, alors $\bar{v}_j > 0$.

Un point de Kuhn Tucker vérifiant les conditions de Kuhn Tucker vérifie donc : $j \in I(\bar{x}) \iff \bar{v}_j > 0$.

Les points de Kuhn Tucker que nous allons considérer maintenant vérifieront les conditions du second ordre. Les algorithmes que nous allons construire convergeront vers des points de ce type.

Dans ce chapitre, nous supposerons aussi que les fonctions $\nabla^2 f$, $\nabla^2 g_1, \dots \nabla^2 g_m$ sont k_2 lipchitz continues sur une boule ouverte $B(\bar{x}, \delta_0)$

Nous noterons aussi

$$I^+(\bar{x}) = I^+$$

$$I(\bar{x}) = I$$

$$I^-(\bar{x}) = I^-$$

Un point vérifiant les conditions de Kuhn Tucker et les conditions du second ordre, vérifie :

$$\left\| \begin{array}{l} \nabla f(\bar{x}) + \bar{U}_I^T \nabla g_I(\bar{x}) = 0 \\ g_I(\bar{x}) = 0 \end{array} \right.$$

<u>définition</u> : soit J une partie de $\{1 \dots m\}$

on définit une fonction :

$$\underline{H_J} : \mathbb{R}^{n+\#J} \longrightarrow \mathbb{R}^{n+\#J}$$

$$(x, U_J) \rightsquigarrow \begin{bmatrix} \nabla_x L(x, U_J) \\ g_J(x) \end{bmatrix}$$

avec : $L(x, U_J) = f(x) + U_J^T \; g_J(x)$

$\nabla_x L(x, U_J) = \nabla f(x) + \nabla g_J(x) \cdot U_J$

$\#J$ est le nombre d'éléments de l'ensemble J

Commentaires

La quantité $L(x, u_J) = f(x) + u_J^T g_J(x)$ n'est pas le lagrangien au sens habituel. Dans le lagrangien au sens habituel, on considère :

$$f(x) + u^T g(x)$$

c'est à dire qu'on tient compte de toutes les contraintes g_j et de tous les multiplicateurs u_j

Nous ne nous interessons qu'aux contraintes correspondant aux indices de J. Nous verrons plus tard que cet ensemble J sera en fait I, ensemble des indices des contraintes actives en $\bar{x}$.

Par abus de langage, nous continuerons d'appeler $L(x, u_J)$ le lagrangien.

soit $\nabla H_J(x, u_J) = \begin{bmatrix} \nabla^2_{xx} L(x, u_J) & \nabla g_J(x) \\ \nabla g_J(x)^T & 0 \end{bmatrix}$

propriété 1 : $\nabla H_I(\bar{x}, \bar{u}_I)$ est régulière.

<u>démonstration</u>

supposons que $\nabla H_I(\bar{x}, \bar{v}_I)$ soit singulière

c'est à dire :

$$\exists (y, p) \in \mathbb{R}^{n+\# I}, \quad (y,p) \neq (0,0)$$

tels que $\nabla H_I(\bar{x}, \bar{v}_I)\begin{pmatrix} y \\ p \end{pmatrix} = 0$

donc :

$$\begin{bmatrix} \nabla^2_{xx} L(\bar{x}, \bar{v}_I) & \nabla g_I(\bar{x}) \\ \nabla g_I(\bar{x})^T & 0 \end{bmatrix} \begin{bmatrix} y \\ p \end{bmatrix} = 0$$

donc

$$\nabla^2_{xx} L(\bar{x}, \bar{v}_I)\, y + \nabla g_I(\bar{x})\, p = 0$$

$$\nabla g_I(x)^T y = 0.$$

donc

$$y^T(\nabla^2_{xx} L(\bar{x}, \bar{v}_I)\, y + \nabla g_I(\bar{x})\, p) = 0$$

$$y^T \nabla^2_{xx} L(\bar{x}, \bar{v}_I)\, y + y^T \nabla g_I(\bar{x})\, p = 0.$$

comme $\nabla g_I(x)^T y = 0$, $y^T \nabla g_I(\bar{x})\, p = 0$

donc $y^T \nabla^2_{xx} L(\bar{x}, \bar{v}_I)\, y = 0$

et donc $y = 0$ par la première condition du second ordre.

et donc $\nabla g_I(\bar{x})\, p = 0$.

ce qui entraine $p = 0$ par la deuxième condition du second ordre

Nous avons donc $(y, p) = (0, 0)$

et $\nabla H_I(\bar{x}, \bar{v}_I)$ est regulière.

$\ddagger$

Commentaires

Cette propriété est fondamentale, parce que nous verrons dans la suite, que pour construire un certain algorithme et déterminer les points qu'il engendre, nous devrons inverser la matrice $\nabla H_I(\bar{x}, \bar{u}_I)$ ou alors une matrice approchant $\nabla H_I(\bar{x}, \bar{u}_I)$.

Nous allons maintenant introduire une autre quantité, $E(x, u)$; cette quantité englobe tous les indices de contraintes vérifiant une certaine propriété.

définition : $$E(x,u) = \left\{ j : u_j \, \|\nabla g_j(x)\|^2 + g_j(x) \, \|\nabla f(x)\| > 0 \right\}.$$

Théorème 1 : Si (x, u) est "assez proche" de $(\bar{x}, \bar{u})$ alors:

$$E(x, u) = I.$$

Commentaires :

Ce théorème est intéressant parce qu'il permet, une fois "assez proche" de $(\bar{x}, \bar{u})$ de déterminer les contraintes actives en $(\bar{x}, \bar{u})$

Attention, ce théorème ne signifie pas que les contraintes actives sont restées inchangées proches de $(\bar{x}, \bar{u})$. Le théorème dit simplement qu'une certaine propriété est invariante proche de $(\bar{x}, \bar{u})$

<u>démonstration</u>

les fonctions $f, \nabla f, g, \nabla g, \|\cdot\|$ étant continues

$\exists\, \varepsilon, \quad \forall (x,u) \in B((\bar{x},\bar{u}), \varepsilon)$

$$\bar{u}_j \|\nabla g_j(\bar{x})\|^2 + g_j(\bar{x}) \|\nabla f(\bar{x})\| > 0$$
$$\iff$$
$$u_j \|\nabla g_j(x)\|^2 + g_j(x) \|\nabla f(x)\| > 0$$

soit $j \in I$; donc $\bar{u}_j > 0$

et $\bar{u}_j \|\nabla g_j(\bar{x})\|^2 > 0$ sinon $\|\nabla g_j(\bar{x})\| = 0$ et ceci est contraire à l'independance linéaire des gradients en $\bar{x}$

donc $\bar{u}_j \|\nabla g_j(\bar{x})\|^2 + g_j(\bar{x}) \|\nabla f(\bar{x})\| > 0$

et donc $u_j \|\nabla g_j(x)\|^2 + g_j(x) \|\nabla f(x)\| > 0$

$\forall (x,u) \in B((\bar{x},\bar{u}), \varepsilon)$

et donc $j \in E(x,u)$

soit $j \in E(x,u) \quad \forall (x,u) \in B((\bar{x},\bar{u}), \varepsilon)$

alors $u_j \|\nabla g_j(x)\|^2 + g_j(x) \|\nabla f(x)\| > 0$

et donc

$$\bar{u}_j \|\nabla g_j(\bar{x})\|^2 + g_j(\bar{x}) \|\nabla f(\bar{x})\| > 0$$

supposons $j \notin I$, alors $\bar{u}_j = 0$

donc $g_j(\bar{x}) \|\nabla f(\bar{x})\| > 0$

mais $\left.\begin{array}{l} g_j(\bar{x}) \leq 0 \\ \|\nabla f(\bar{x})\| \geq 0 \end{array}\right\}$ donc $g_j(\bar{x}) \|\nabla f(\bar{x})\| \leq 0$

nous avons donc une contradiction

nous pouvons donc dire que $j \in I$.

#

2 L'algorithme proprement dit

Commentaires :

Considérons le système d'équations linéaires en (x^{i+1}, u^{i+1}) défini par

$$\nabla H_I(x^i, u^i_I) \begin{bmatrix} x^{i+1} - x^i \\ u^{i+1}_I - u^i_I \end{bmatrix} + H_I(x^i, u^i_I) = 0$$

Ce système fait appel à la matrice $\nabla_{xx} L(x^i, u^i_I)$. Cette matrice, nous ne la connaissons généralement pas parce que nous ne connaissons pas $\nabla_{xx} L(x^i, u^i_I)$. Le système n'est donc pas toujours "traitable tel quel". C'est pourquoi on introduit un autre système d'équations linéaires :

$$LS(x, u, G, J) = 0$$

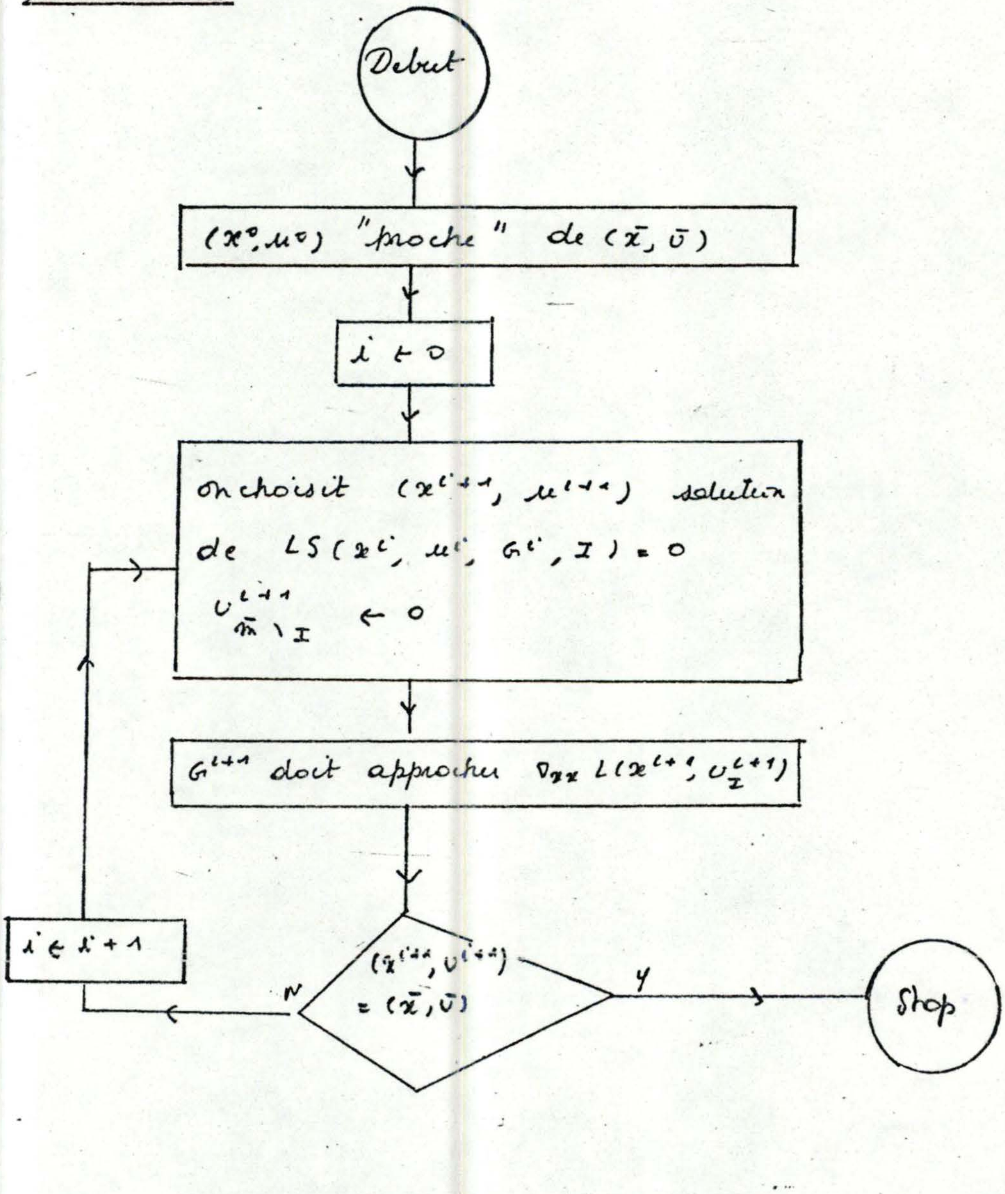
$$\Leftrightarrow$$

$$\begin{bmatrix} G & \nabla g_J(x) \\ \nabla g_J(x)^T & 0 \end{bmatrix} \begin{bmatrix} y - x \\ z - u_J \end{bmatrix} + \begin{bmatrix} \nabla_x L(x, u_J) \\ g_J(x) \end{bmatrix} = 0$$

$x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $J \subset \overline{m}$, $G \in \mathbb{R}^{n \times n}$

La matrice $DH_J(x, u_J)$ a donc été transformée on a remplacé la partie $D_{xx} L(x, u_J)$ par une matrice "quelconque" G
Pour le moment, cette matrice est quelconque mais on verra que le but de la théorie sera de résoudre $LS(x^i, u^i, G^i, J) = 0$ où G^i approximera $D_{xx} L(x^i, u^i_I)$,

Algorithme



Description de l'algorithme

L'algorithme consiste donc en une suite d'équations linéaires à résoudre :

$$G^i (x^{i+1} - x^i) + \nabla g_I (x^i) (u_I^{i+1} - u_I^i) + \nabla f(x^i) + U_I^T \nabla g_I (x^i) = 0$$

$$\nabla g_I (x^i)^T (x^{i+1} - x^i) + g_I (x^i) = 0$$

Notons que $U^{i+1}_{m \setminus I} = 0$.

Nous allons maintenant établir un théorème de convergence LOCALE.

Cette convergence sera superlinéaire à condition que le point (x^0, u^0) soit assez proche du point de Kuhn Tucker $(\bar{x}, \bar{U})$.

La notion de convergence superlinéaire utilisée ici est celle d'Orthega et Rheinboldt.

Les références exactes sont :

" Iteratives Solution of Nonlinear Equations in Several Variables.
J.M Orthega and W.C. Rheinboldt .
Academic Press New York and London 1970
Chapitre 9 .

3. Théorème de convergence

Théorème 1 :

$\forall \delta_1$ appartenant à $]0,1[$

Il existe $\delta_2(\delta_1)$, $\delta_3(\delta_1)$, k_{32}, $k_{31} > 0$

telles que si $\forall i$ on a :

1) $\| (x^i, u^i) - (\bar{x}, \bar{u}) \| \leq \delta_2$

2) a) $\| G^i - \nabla_{xx} L(x^i, \bar{u}) \| \leq \frac{1}{2} \beta_I$

avec $\beta_I = \| \nabla H_I (\bar{x}, \bar{u}_I)^{-1} \|$

ou alors

b) $\begin{bmatrix} G^i & \nabla g_I(x^i) \\ \nabla g_I(x^i)^T & 0 \end{bmatrix}$ est regulière

3) $\dfrac{\| (G^i - \nabla_{xx} L(x^i, \bar{u}))(x^{i+1} - x^i) \|}{\| x^{i+1} - x^i \|} = \Delta^i \leq \delta_3$

alors on a

4) $E(x^i, u^i) = I$

(x^{i+1}, u^{i+1}) défini par l'algorithme existe et

$\| (x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u}) \| \leq \delta_1 \| x^i - \bar{x} \|$

5) si en plus $G^i = \nabla_{xx}(x^i, u^i)$ alors

$\| (x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u}) \| \leq k_{31} \| x^i - \bar{x} \| \| (x^i, u^i) - (\bar{x}, \bar{u}) \|$

6) soit S le sous espace engendré par

$\{ \nabla g_j(\bar{x}) \ : \ j \in I \}$

$\forall y \in R^n$, soit y_s la projection de y sur S

alors : $\| x_s^{i+1} - \bar{x}_s \| \leq k_{32} \| x^i - \bar{x} \|^2$.

Commentaire : explications du théorème

La première hypothèse indique simplement que l'on doit partir d'un point assez proche de $(\bar{x}, \bar{u})$

La deuxième hypothèse indique que la matrice

$$\begin{bmatrix} G^i & \nabla g_I(x^i) \\ \nabla g_I(x^i)^T & 0 \end{bmatrix}$$

est une matrice régulière.

ou du moins qu'elle possède certaines propriétés qui permettent d'affirmer qu'elle est régulière. On pourra donc calculer son inverse, de manière à déterminer (x^{i+1}, u^{i+1}).

La troisième hypothèse nous indique que le comportement de G^i et celui de $L(x^i, \bar{u})$ sont proches en ce qui concerne $(x^{i+1} - x^i)$

La première affirmation : $F(x^i, u^i) = I$ a déjà été prouvée. Le fait de dire que (x^{i+1}, u^{i+1}) existe vient de la deuxième hypothèse. La propriété de convergence est en fait une propriété de convergence superlinéaire puisqu'elle est vérifiée $\forall\ \delta_1 \in]0,1[$.

Dans le cas idéal où G^i est exactement $\nabla_{xx} L(x^i, u^i)$ on nous donne une indication sur la vitesse de convergence : $\|(x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u})\| \leq K_{31} \|x^i - \bar{x}\| \, \|(x^i, u^i) - (\bar{x}, \bar{u})\|$

La troisième affirmation s'intéresse au sous espace vectoriel engendré par les gradients des contraintes actives en $\bar{x}$. Ce sous espace est noté S.

Les deux dernières propriétés seront des outils de calcul.

Le résultat fondamental à retenir est la convergence superlinéaire de (x^i, u^i) vers $(\bar{x}, \bar{u})$

Commentaires : principe de demonstration

La démonstration de ce theorème est longue et technique, nous allons d'abord presenter le sheme de la démonstration, nous presenterons ensuite la démonstration proprement dite, elle ne presente pas beaucoup d'interêts au point de vue de l'optimisation elle est interessante au point de vue des techniques qu'elle presente.

premiere affirmation

Nous ferons intervenir la suite (y^i, p^i), cette suite est en fait celle qu'on obtiendrait si $G^i = \nabla_{xx} L(x^i, \bar{U})$.
On montrera d'abord que

$$\| (y^{i+1}, p^{i+1}) - (\bar{x}, \bar{U}_I) \| \leq K_{33} \| x^i - \bar{x} \|^2 \qquad (a)$$

et ensuite que (x^{i+1}, u^{i+1}) existe effectivement
On calculera alors la différence entre ce que l'on obtient c'est à dire (x^{i+1}, u^{i+1}) et ce qu'on aurait voulu avoir c'est à dire (y^{i+1}, p^{i+1}) et on montrera que

$$\| (x^{i+1}, u_I^{i+1}) - (y^{i+1}, p^{i+1}) \|$$
$$\leq 4 \beta_I \Delta^i (K_{33} \| x^i - \bar{x} \|^2 + \| x^i - \bar{x} \|) \qquad (b)$$

En combinant (a) et (b), nous obtiendrons la thèse

deuxième affirmation

Nous nous plaçons dans le cas idéal, c'est à dire $G^i = \nabla_{xx} L(x^i, U^i)$; nous essayerons alors de majorer Δ^i par : $\sqrt{m}\ k_{34} \| U^i - \bar{U} \|$. Il faut remarquer que G^i et $\nabla_{xx} L(x^i, \bar{U})$ ne diffèrent que par des termes en u
On remplacera alors Δ^i par son majorant dans une des inégalités obtenues précédemment

Pour démontrer la dernière affirmation, nous allons distinguer 2 cas :

1) $x_s^{i+1} = y_s^{i+1}$

nous verrons que ce cas est trivial

2) $x_s^{i+1} \neq y_s^{i+1}$

Pour chaque $j \in I$, on définit A_j comme étant l'angle entre $\nabla g_j(\bar{x})$ et $(x_s^{i+1} - y_s^{i+1})$ ou encore l'angle entre $\nabla g_j(\bar{x})$ et $\dfrac{(x_s^{i+1} - y_s^{i+1})}{\| x_s^{i+1} - y_s^{i+1} \|}$

De là, nous exprimerons le produit scalaire entre $\nabla g_j(\bar{x})$ et $(x_s^{i+1} - y_s^{i+1})$

Nous arriverons à la conclusion que :

$$\forall j \in I, \ \| \nabla g_j(\bar{x}) \| \, \| x_s^{i+1} - y_s^{i+1} \| \, | \cos A_j | \leq k \, \| x^i - \bar{x} \| \, \| x^{i+1} - y^{i+1} \|$$

et grace à un lemme assurant que $\cos A_{j^*} > 0$ (j^* tel que $\cos A_{j^*} = \max_{j \in I} \cos A_j$) nous arriverons à la conclusion que

$$\| x_s^{i+1} - y_s^{i+1} \| \leq \frac{4 \beta_I \, \delta_3 (k_{33} \delta_2 + 1) \, k \, \| x^i - \bar{x} \|^2}{k_{35} \min \| \nabla g_j(\bar{x}) \|}.$$

en posant $k_{32} = k_{33} + \dfrac{4 \beta_I \, \delta_3 (k_{33} \delta_2 + 1) \, k}{k_{35} \min \| \nabla g_j(\bar{x}) \|}$ nous parviendrons à la thèse

Démonstration

première partie : préliminaires

pas 1 : notations

$$\begin{bmatrix} G^i & \nabla g_I(x^i) \\ \nabla g_I(x^i)^T & 0 \end{bmatrix} = F_I^i$$

pas 2 : $\| \nabla H_I(x^i, u_I^i)^{-1} \| \leq \frac{3}{2} \beta_I$

Ceci est évident parce que $\nabla H_I(\bar{x}, \bar{u}_I)$ est régulière et que toutes les fonctions intervenant dans $(\nabla H_I(x^i, u_I^i))^{-1}$ sont des fonctions continues

pas 3) $\| (y^{i+1}, p^{i+1}) - (\bar{x}, \bar{u}_I) \| \leq K_{33} \| x^i - \bar{x} \|^2$

où (y^{i+1}, p^{i+1}) est défini par :

$$\nabla H_I(x^i, \bar{u}_I) \begin{bmatrix} y^{i+1} - x^i \\ p^{i+1} - \bar{u}_I \end{bmatrix} + H_I(x^i, \bar{u}_I) = 0$$

Nous savons que $H_I(\bar{x}, \bar{u}_I) = 0$

Un développement de Taylor de $H_I(x^i, \bar{u}_I)$ donne :

$$\nabla H_I(x', \bar{u}_I) \begin{bmatrix} x^i - \bar{x} \\ \bar{u}_I - \bar{u}_I \end{bmatrix} + H_I(x', \bar{u}_I) = 0$$

avec x' point intermédiaire entre $\bar{x}$ et x^i

donc

$$\begin{bmatrix} x^{i} - \bar{x} \\ \bar{v}_I - \bar{v}_I \end{bmatrix} = - \nabla H_I (x', \bar{v}_I)^{-1} \; H_I (x^{i}, \bar{v}_I)$$

la matrice $\nabla H_I (x', \bar{v}_I)$ est une matrice inversible par un raisonnement semblable à celui effectué au pas 2.

nous avons aussi que

$$\begin{bmatrix} y^{i+1} - x^{i} \\ p^{i+1} - \bar{v}_I \end{bmatrix} = - \nabla H_I (x^{i}, \bar{v}_I)^{-1} \; H_I (x^{i}, \bar{v}_I)$$

et donc

$$\begin{bmatrix} y^{i+1} - \bar{x} \\ p^{i+1} - \bar{v}_I \end{bmatrix} = \left(\nabla H_I (x', \bar{v}_I)^{-1} - \nabla H_I (x^{i}, \bar{v}_I)^{-1} \right) \cdot \left(H_I (x^{i}, \bar{v}_I) - \underset{0}{\underbrace{H_I (\bar{x}, \bar{v}_I)}} \right)$$

donc :

$$\| (y^{i+1}, p^{i+1}) - (\bar{x}, \bar{u}_I) \| \leq \| \nabla H_I (x^{i}, \bar{v}_I)^{-1} - \nabla H_I (x', \bar{v}_I)^{-1} \| \cdot \| H_I (x^{i}, \bar{v}_I) - H_I (\bar{x}, \bar{v}_I) \|$$

Les fonctions intervenant dans $\nabla H_I (x, u_J)$ et $H_I (x, u_J)$ sont toutes k Lipschitz continues on a donc :

$$\begin{aligned} \| (y^{i+1}, p^{i+1}) - (\bar{x}, \bar{v}_I) \| & \leq k a \| x^{i} - x' \| \; k b \| x^{i} - \bar{x} \| \\ & \leq 2 \, k a \| x^{i} - \bar{x} \| \; k b \| x^{i} - \bar{x} \| \\ & \leq k_{33} \| x^{i} - \bar{x} \|^{2} . \end{aligned}$$

deuxième partie : démonstration de 4

pas 1 (x^{i+1}, u^{i+1}) existe :

On considère le système $LS(x^i, u^i, G^i, I) = 0$

c'est à dire :
$$\begin{cases} G^i(x^{i+1} - x^i) + \nabla g_I(x^i)(u_I^{i+1} - u_I^i) \\ \quad + \nabla f(x^i) + \nabla g_I(x^i)\, u_I^i = 0 \\ \nabla g_I(x^i)(x^{i+1} - x^i) + g_I(x^i) = 0 \end{cases}$$

La contribution de u_I^i étant nulle, on peut remplacer ce système par :
$$\begin{cases} G^i(x^{i+1} - x^i) + \nabla g_I(x^i)(u_I^{i+1} - \bar{u}_I) \\ \quad + \nabla f(x^i) + \nabla g_I(x^i)\, \bar{u}_I = 0 \\ \nabla g_I(x^i)(x^{i+1} - x^i) + g_I(x^i) = 0 \end{cases}$$

c'est à dire : $$F_I^i \begin{bmatrix} x^{i+1} - x^i \\ u_I^{i+1} - \bar{u}_I \end{bmatrix} + H_I(x^i, \bar{u}_I) = 0 \qquad (c)$$

soit $\delta H_I^i = F_I^i - \nabla H_I(x^i, \bar{u}_I)$

donc $\| \delta H_I^i \| = \| G^i - \nabla_{xx} L(x^i, \bar{u}_I) \|$

(2.a) entraine que $\| G^i - \nabla_{xx} L(x^i, \bar{u}_I) \| \leq \dfrac{1}{2\beta_I}$

comme $\| \nabla H_I(x^i, u_I^i)^{-1} \| \leq \frac{3}{2} \beta_I$
$\forall (x^i, u^i) \in B((\bar{x}, \bar{u}), \delta_2)$

on a que $\| \nabla H_I(x^i, \bar{u}_I)^{-1} \| \leq \frac{3}{2} \beta_I$

et donc $\| \delta H_I^i \| \leq \dfrac{3}{4} \| \nabla H_I(x^i, \bar{u}_I)^{-1} \|^{-1}$.

si nous parvenons à démontrer que F_I^i est non singu-
-lière, nous aurons par (e) que (x^{i+1}, u_I^{i+1}) existe.
Les composantes $v_{\overline{m}/I}^{i+1}$ existent toujours par construction
de l'algorithme.

Supposons que F_I^i soit singulière
alors, il existe x, $\|x\| = 1$ et $F_I^i\, x = 0$.

soit donc $F_I^i\, x = 0$

donc $(\delta H_I^i + \nabla H_I(x^i, \bar{v}_I))\, x = 0$

et $-\delta H_I^i\, x = \nabla H_I(x^i, \bar{v}_I)\, x = 0$

donc $-(\nabla H_I(x^i, \bar{v}_I)^{-1}\, \delta H_I^i)\, x = x$

donc $\|-\nabla H_I(x^i, \bar{v}_I)^{-1}\, \delta H_I^i\, x\| = \|x\| = 1$

donc $|-1|\, \|\nabla H_I(x^i, \bar{v}_I)^{-1}\|\, \|\delta H_I^i\|\, \|x\| \geq 1$

c'est à dire :

$$\|\nabla H_I(x^i, \bar{v}_I)^{-1}\|\ \|\delta H_I^i\| \geq 1$$

mais $\|\delta H_I^i\| \leq \frac{3}{4} \|\nabla H_I(x^i, \bar{u}_I)^{-1}\|^{-1}$

donc :

$$\|\nabla H_I(x^i, \bar{v}_I)^{-1}\| . \frac{3}{4} \|\nabla H_I(x^i, \bar{v}_I)^{-1}\|^{-1} \geq 1$$

et donc : $\frac{3}{4} \geq 1$

En supposant F_I^i singulière, nous aboutissons
donc à une contradiction manifeste,
F_I^i est donc régulière

pas 2 : $\| (x^{i+1}, u_I^{i+1}) - (y^{i+1}, p^{i+1}) \| \leq 2\beta_I \Delta^i (\|x^{i+1} - y^{i+1}\| + \|y^{i+1} - x^i\|$

rappelons que

d) : $\| (y^{i+1}, p^{i+1}) - (\bar{x}, \bar{u}_I) \| \leq K_{33} \|x^i - \bar{x}\|^2$

e) $DH_I(x^i, \bar{v}_I) \begin{bmatrix} y^{i+1} - x^i \\ p^{i+1} - \bar{v}_I \end{bmatrix} + H_I(x^i, \bar{v}_I) = 0$

et donc :

$$H_I(x^i, \bar{v}_I) = - DH_I(x^i, \bar{v}_I) \begin{bmatrix} y^{i+1} - x^i \\ p^{i+1} - \bar{v}_I \end{bmatrix}$$

f) $F_I^i \begin{bmatrix} x^{i+1} - x^i \\ v_I^{i+1} - \bar{v}_I \end{bmatrix} + H_I(x^i, \bar{v}_I) = 0$

g) $\delta H_I^i = F_I^i - DH_I(x^i, \bar{v}_I)$.

on déduit que :

$$DH_I(x^i, \bar{v}_I) \begin{bmatrix} x^{i+1} - y^{i+1} \\ v_I^{i+1} - p^{i+1} \end{bmatrix} + \delta H_I^i \begin{bmatrix} x^{i+1} - x^i \\ u_I^{i+1} - \bar{v}_I \end{bmatrix} = 0$$

donc : $\| (x^{i+1}, v_I^{i+1}) - (y^{i+1}, p^{i+1}) \|$

$\leq \| DH_I(x^i, \bar{v}_I)^{-1} \| \, \| (\delta H_I^i)' \, (x^{i+1} - x^i) \|$

$= \| DH_I(x^i, \bar{v}_I)^{-1} \| \, \underbrace{\| (G^i - D_{xx} L(x^i, \bar{v}_I)) (x^{i+1} - x^i \|}_{= \Delta^i \|x^{i+1} - x^i\|}$

$\leq 2\beta_I \Delta^i \|x^{i+1} - x^i\|$

parce que $\|DH_I(x^i, \bar{v}_I)^{-1}\| \leq 2 \|DH_I(\bar{x}, \bar{v}_I)^{-1}\| = 2\beta_I$

$\leq 2\beta_I \Delta^i (\|x^{i+1} - y^{i+1}\| + \|y^{i+1} - x^i\|$

<u>pas 3</u> : $\|(x^{i+1}, u_I^{i+1}) - (y^{i+1}, p^{i+1})\| \leq 4\beta_I \Delta^i (K_{33}\|x^i - \bar{x}\|^2 + \|x^i - \bar{x}\|)$

on a que :

$$\|(x^{i+1}, v_I^{i+1}) - (y^{i+1}, p^{i+1})\| \geq \|x^{i+1} - y^{i+1}\|$$

donc :

$$-2\beta_I \Delta^i \|(x^{i+1}, u_I^{i+1}) - (y^{i+1}, p^{i+1})\| \leq -2\beta_I \Delta^i \|x^{i+1} - y^{i+1}\|$$

additionnons membre à membre cette inégalité à celle obtenue par le pas 2, nous obtenons :

$$(1 - 2\beta_I \Delta^i)\|(x^{i+1}, u_I^{i+1}) - (y^{i+1}, p^{i+1})\| \leq 2\Delta^i \beta_I \|y^{i+1} - x^i\|$$

soit $\delta_3 < \dfrac{1}{4\beta_I}$ donc $\Delta^i < \delta_3 < \dfrac{1}{4\beta_I}$

$$(1 - 2\Delta^i \beta_I) \geq 1 - 2\beta_I \cdot \frac{1}{4\beta_I} = \frac{1}{2}$$

donc :

$$\frac{1}{2}\|(x^{i+1}, v_I^{i+1}) - (y^{i+1}, p^{i+1})\| \leq 2\beta_I \Delta^i \|y^{i+1} - x^i\|$$

$$\begin{aligned}
\|(x^{i+1}, v_I^{i+1}) - (y^{i+1}, p^{i+1})\| &\leq 4\beta_I \Delta^i \|y^{i+1} - x^i\| \\
&\leq 4\beta_I \Delta^i (\|y^{i+1} - \bar{x}\| + \|x^i - \bar{x}\|) \\
&\leq 4\beta_I \Delta^i (\|(y^{i+1}, p^{i+1}), (\bar{x}, \bar{v}_I)\| + \|x^i - \bar{x}\|) \\
&\leq 4\beta_I \Delta^i (K_{33}\|x^i - \bar{x}\|^2 + \|x^i - \bar{x}\|).
\end{aligned}$$

<u>pas 4</u> : <u>conclusion</u>

$$\begin{aligned}
\|(x^{i+1}, v_I^{i+1}) - (\bar{x}, \bar{v}_I)\| &\leq \|(y^{i+1}, p^{i+1}) - (\bar{x}, \bar{v}_I)\| \\
&\quad + \|(x^{i+1}, v_I^{i+1}) - (y^{i+1}, p^{i+1})\| \\
&\leq K_{33}\|x^i - \bar{x}\|^2 + 4\beta_I \Delta^i (K_{33}\|x^i - \bar{x}\|^2 + \|x^i - \bar{x}\|) \\
&\leq [K_{33}\delta_2 + 4\beta_I \Delta^i (K_{33}\delta_2 + 1)]\|x^i - \bar{x}\|
\end{aligned}$$

parce que $\|(x^i, v^i) - (\bar{x}, \bar{v}_I)\| < \delta_2$.

$$\leq \delta_1 \|x^i - \bar{x}\|$$

en diminuant δ_2, δ_3 si nécessaire

troisième partie : démonstration de 5

pas 1 : $\Delta^i \leq \sqrt{m}\ K_{34}\ \| v^i - \bar{v} \|$

si $G^i = \nabla_{xx} L(x^i, v^i)$

alors :

$$\Delta^i \leq \| G^i - \nabla_{xx} L(x^i, \bar{v}) \|$$
$$= \| \nabla_{xx} L(x^i, v^i) - \nabla_{xx} L(x^i, \bar{v}) \|$$
$$\leq \sum_{j=1}^{m} \| (v_j^i - \bar{v}_j) \nabla^2 g_j (x^i) \|$$

soit k_{34} tel que $\forall j \in \bar{m}$,

$$\| \nabla^2 g_j (x^i) \| \leq k_{34} \qquad \forall x \in B(\bar{x}, \delta_2).$$

k_{34} existe parce que $\nabla^2 g_j(x)$ $j = 1..m$ sont lipschitz continues sur une boule centrée en $\bar{x}$

nous obtenons donc :

$$\Delta^i \leq \sum_{j=1}^{m} | u_j^i - \bar{v}_j |\ k_{34}$$

$$= k_{34}\ \| u^i - \bar{u} \|_1$$

$$\leq \sqrt{m}\ k_{34}\ \| u^i - \bar{v} \|$$

<u>pas 2:</u> $\|(x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u})\| \leq K_{31} \|x^i - \bar{x}\| \, \|(x^i, u^i) - (\bar{x}, \bar{u})\|$.

Le pas 3 de la deuxième partie permet de dire :

$$\|(x^{i+1}, u_I^{i+1}) - (\bar{x}, \bar{u}_I)\|$$
$$\leq K_{33} \|x^i - \bar{x}\|^2 + 4\beta_I \Delta^i (K_{33} \|x^i - \bar{x}\|^2 + \|x^i - \bar{x}\|)$$

Nous savons aussi que :

$$\|(x^{i+1}, v_I^{i+1}) - (\bar{x}, \bar{v}_I)\| = \|(x^{i+1}, v^{i+1}) - (\bar{x}, \bar{v})\|$$

Et donc, en remplaçant Δ^i par son majorant, nous obtenons :

$$\|(x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u})\|$$
$$\leq K_{33} \|x^i - \bar{x}\|^2 + 4\beta_I \sqrt{m}\, K_{34} \|v^i - \bar{v}\| (K_{33} \|x^i - \bar{x}\|^2 + \|x^i - \bar{x}\|)$$

$$= \left(K_{33} \|x^i - \bar{x}\| + 4\beta_I \sqrt{m}\, K_{34} (K_{33} \delta_2 + 1) \|v^i - \bar{v}\|\right) \|x^i - \bar{x}\|$$

$$\leq \left(K_{33} + 4\beta_I \sqrt{m}\, K_{34} (K_{33} \delta_2 + 1)\right) \|(x^i, u^i) - (\bar{x}, \bar{u})\| \, \|x^i - \bar{x}\|$$

soit $k_{31} = k_{33} + 4\beta_I \sqrt{m}\, K_{34} (K_{33} \delta_2 + 1)$

donc $\|(x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u})\|$
$$\leq k_{31} \|(x^i, u^i) - (\bar{x}, \bar{u})\| \, \|x^i - \bar{x}\|.$$

quatrième partie : démonstration de 6

pas 1 : Lemme

⌜ $\forall x \in S$, soit $e(x) = \max \{ |\cos \theta_x^j| : j \in I \}$

θ_x^j ~~est~~ l'angle entre x et $\nabla g_j(\bar{x})$

alors, $k_{35} = \min \{ e(x) , x \in S , \|x\| = 1 \} > 0$. ⌟

Cela vient du fait que $e(x)$ est une fonction continue sur un ensemble compact $\{ x \in S , \|x\| = 1 \}$. et que $e(x) > 0$

en effet $e(x) > 0$ $\begin{cases} : \text{par définition de } S \\ : \text{par l'indépendance linéaire des} \end{cases}$

gradients des contraints actives en $\bar{x}$.

pas 2 : $\nabla g_I(x^i)(y^{i+1} - x^{i+1}) = 0$.

nous savons par la définition de x^{i+1} et y^{i+1} que :

$$\nabla g_I(x^i)(y^{i+1} - x^i) + g_I(x^i) = 0$$
$$\nabla g_I(x^i)(x^{i+1} - x^i) + g_I(x^i) = 0$$

donc $\nabla g_I(x^i)(y^{i+1} - x^{i+1}) = 0$

pas 3 : $\forall j \in I$, $|\nabla g_j(\bar{x})(x_s^{i+1} - y_s^{i+1})| \leq k \|x^i - \bar{x}\| \, \|x^{i+1} - y^{i+1}\|$ $j \in I$

$$|\nabla g_j(\bar{x})(x_s^{i+1} - y_s^{i+1})|$$
$$= |\nabla g_j(\bar{x})(x^{i+1} - y^{i+1})|$$
$$= |(\nabla g_j(x^i) - \nabla g_j(\bar{x}))(x^{i+1} - y^{i+1})|$$

grace au pas 2 .

$$\leq k \|x^i - \bar{x}\| \, \|x^{i+1} - y^{i+1}\|$$

parce que les fonctions $\nabla g_j(\cdot)$ $j = 1 \dots m$ sont k. lipschitz continues.

pas 4 : supposons $x_s^{i+1} = y_s^{i+1}$

alors :

$$\begin{aligned} \| x_s^{i+1} - \bar{x}_s \| &= \| y_s^{i+1} - \bar{x}_s \| \\ &\leq \| y^{i+1} - \bar{x} \| \\ &\leq \| (y^{i+1}, p^{i+1}) - (\bar{x}, \bar{u}_I) \| \\ &\leq K_{33} \| x^i - \bar{x} \|^2 . \end{aligned}$$

Ce cas est donc trivial, nous envisagerons donc $x^{i+1} \neq y^{i+1}$.

pas 5 : définition de A_j

soit A_j l'angle entre $\nabla g_j(\bar{x})$ et $\dfrac{(x_s^{i+1} - y_s^{i+1})}{\| x_s^{i+1} - y_s^{i+1} \|}$

$$\| \nabla g_j(\bar{x}) \| \; \| x_s^{i+1} - y_s^{i+1} \| \; | \cos A_j |$$

$$= | \nabla g_j(\bar{x}) (x_s^{i+1} - y_s^{i+1}) | \qquad \forall j \in I.$$

$$\leq k \| x^i - \bar{x} \| \; \| x^{i+1} - y^{i+1} \| \qquad \text{par le pas 3}.$$

soit j^* tel que $\cos A_{j^*} = \max\limits_{j \in I} \{ \cos A_j \}$.

<u>pas 6 : conclusion</u>

Le pas 5 entraine que

$$\| x_s^{i+1} - y_s^{i+1} \| \leq \frac{k \|x^i - \bar{x}\| \, \|x^{i+1} - y^{i+1}\|}{\min_{j \in I} \|\nabla g_j(\bar{x})\| \, |\cos A_j^*|}$$

par les étapes précédentes :

$$\begin{aligned} \|x^{i+1} - y^{i+1}\| &\leq \|(x^{i+1}, u^{i+1}) - (y^{i+1}, p^{i+1})\| \\ &\leq 4\beta_I \Delta^i (k_{33} \|x^i - \bar{x}\|^2 + \|x^i - \bar{x}\|) \\ &\leq 4\beta_I \delta_3 (k_{33} \delta_2 + 1) \|x^i - \bar{x}\| . \end{aligned}$$

donc : $$\| x_s^{i+1} - y_s^{i+1} \| \leq \frac{k \, 4\beta_I \delta_3 (k_{33} \delta_2 + 1) \|x^i - \bar{x}\|^2}{\min_{j \in I} \|\nabla g_j(\bar{x})\| \, |\cos A_j^*|}$$

$$\leq \frac{k \, 4\beta_I \delta_3 (k_{33} \delta_2 + 1) \|x^i - \bar{x}\|^2}{k_{35} \cdot \min_{j \in I} \|\nabla g_j(\bar{x})\|}$$

$\downarrow$ existe par le pas 1 .

soit $$k_{32} = \frac{4 k \beta_I \delta_3 (k_{33} \delta_2 + 1)}{k_{35} \min_{j \in I} \|\nabla g_j(\bar{x})\|} + k_{33} .$$

alors $$\begin{aligned} \| x_s^{i+1} - \bar{x}_s \| &\leq \| y_s^{i+1} - \bar{x}_s \| + \| x_s^{i+1} - y_s^{i+1} \| \\ &\leq \| y^{i+1} - \bar{x} \| + \| x_s^{i+1} - y_s^{i+1} \| \\ &\leq k_{32} \|x^i - \bar{x}\|^2 . \end{aligned}$$

$\#$

C: Algorithmes réduisant un P.N.L en une suite de programmes quadratiques

1) préliminaires

Dans cette section nous allons développer une classe d'Algorithmes réduisant le P.N.L en une suite de programmes quadratiques. Nous proposerons un premier algorithme et nous montrerons que sous certaines hypothèses, les points qu'il engendre sont les mêmes que ceux engendrés par l'algorithme de la section B. Nous proposerons ensuite un deuxième Algorithme qui est une modification du premier mais qui finalement engendrera les mêmes points

définition : Considérons le programme quadratique $Q(x, G)$:

$$\underset{y \in \mathbb{R}^n}{\text{minimiser}} \quad \nabla f(x)^T (y - x) + \frac{1}{2}(y-x)\, G\,(y-x)$$

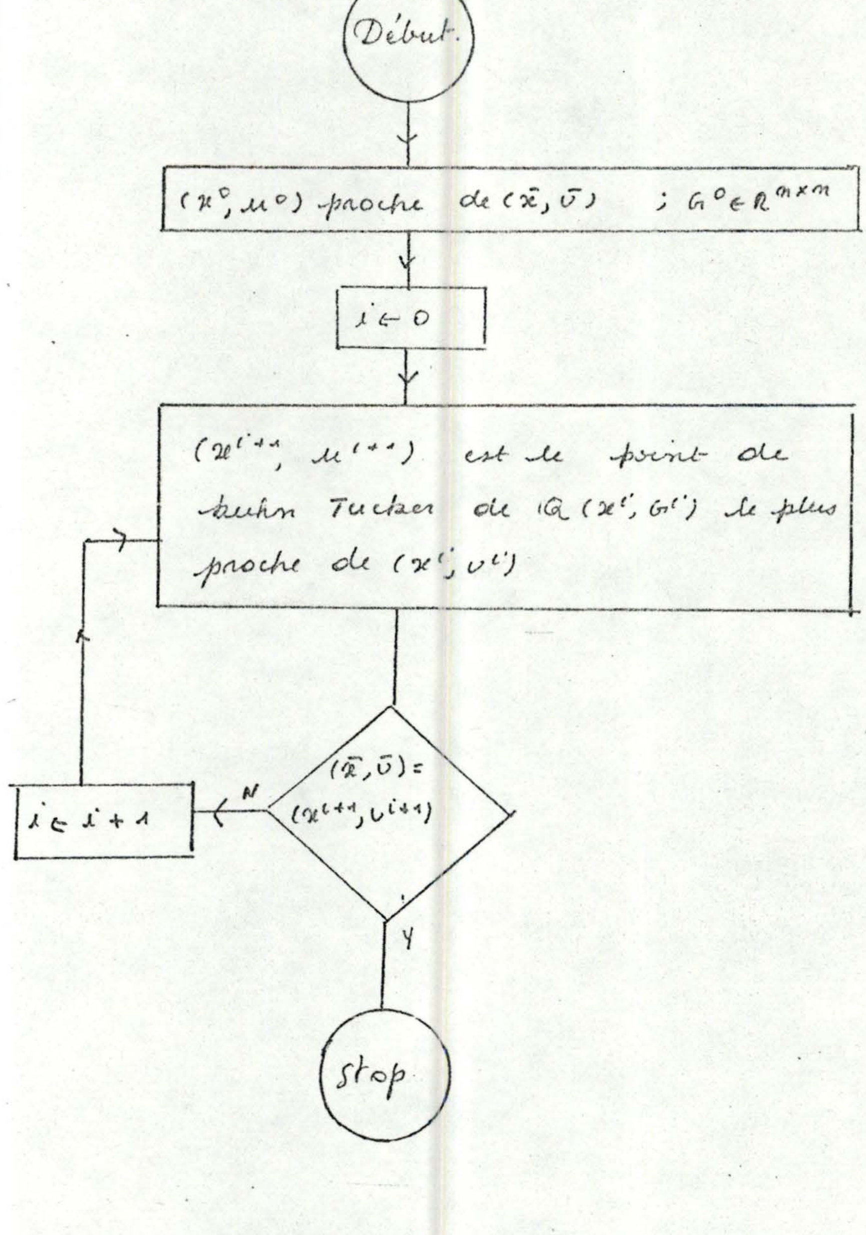
sous les contraintes : $g(x) + \nabla g(x)^T (y - x) \leq 0.$

Commentaires

$Q(x, G)$ est donc un problème quadratique en y parce que la fonction à minimiser et les contraintes de $Q(x, G)$ sont quadratiques en y.

Ce qu'on va faire, c'est déterminer des points de Kuhn Tucker du problème $Q(x, G)$. Plus précisement on va construire un algorithme à partir de ce problème quadratique.

2: description du premier algorithme



Commentaires

(x^{i+1}, u^{i+1}) est donc le point de Kuhn Tucker le plus proche de (x^i, v^i) en terme de norme du problème $Q(x^i, G^i)$.

(x^{i+1}, u^{i+1}) vérifie donc :

$$\begin{cases} g(x^i) + \nabla g(x^i)^T (x^{i+1} - x^i) \leq 0 . \\ (u^{i+1})^T \left(g(x^i) + \nabla g(x^i)^T (x^{i+1} - x^i) \right) = 0 \\ \nabla f(x^i) + G^i (x^{i+1} - x^i) + \nabla g(x^i) u^{i+1} = 0 . \\ u^{i+1} \geq 0 \end{cases}$$

Comme dans l'algorithme de la section B, nous prendrons G^i approximant le Hessien du lagrangien en $(x^i; v_I^i)$, c'est à dire :

$$G^i \simeq \nabla_{xx} L(x^i, v_I^i) \quad , \quad \text{si } I = E(x^i, v^i) .$$

On montrera que le point de Kuhn Tucker déterminé par cet algorithme, vérifie aussi les conditions imposées par l'algorithme de la section B.

C'est à dire :

$$\begin{cases} \nabla g_I (x^i) + \nabla g_I (x^i)^T (x^{i+1} - x^i) = 0 \\ \nabla f(x^i) + G^i (x^{i+1} - x^i) + \nabla g_I (x^i) v_I^i \\ \quad + \nabla g_I (x^i) (v_I^{i+1} - v_I^i) = 0 . \end{cases}$$

Cette propriété sera vérifiée pour autant que la matrice G^i vérifie certaines conditions principalement : être assez près de $\nabla_{xx} L(x^i, \bar{v}_I)$ et donc aussi de $\nabla_{xx} L(x^i, v_I^i)$.

Et donc les propriétés que nous avons étudiées a propos de l'algorithme de la section B resteront verifiées pour ce nouvel algorithme.
Et donc, on voit que le procédé qui consiste à re-duire un P.N.L en une suite d'équations linéaires est equivalent au procédé qui consiste à reduire le P.N.L en une suite de programmes quadratiques.

3. Théorème d'identification

Théorème 1 : Il existe $\delta_4 > 0$ tel que
si : 1) $\| (x^0, u^0) - (\bar{x}, \bar{u}) \| \leq \delta_4$

2) $\| G^i - \nabla_{xx} L(x^i, \bar{u} \| \leq \frac{1}{10\beta}$.

où $\beta = \frac{3}{2} \| \nabla H(\bar{x}, \bar{u})^{-1} \|$

$$H(\bar{x}, \bar{u}) = \begin{bmatrix} \nabla f(x) + \nabla g(x) u \\ u_1 g_1(x) \\ \vdots \\ u_m g_m(x). \end{bmatrix}$$

Alors : en partant du même point (x^0, u^0) cet algorithme et celui de la section precedente engendre les mêmes points (x^i, u^i) et ont donc les mêmes propriétés de convergence.

démonstration

Sous les hypothèses du théorème, Garcia et Mangasarian ont démontré que

1) $g_j(x^i) + \nabla g_j(x^i)^T (x^{i+1} - x^i) \leq 0$
et l'inégalité devient une égalité si et seulement si $j \in I$

2) $u_j^{i+1} \geq 0$ et l'inégalité est stricte si et seulement si $j \in I$.

les références exactes sont :
" Superlinearly Convergent Quasi-Newton Algorithms for Nonlinearly Constrained Optimization Problems
by U.M. Garcia Palomares and O.L. Mangasarian
Computer Sciences Technical Report # 195' March 1974
Théorème 3.1 page 6.7.8.

Nous avons donc : $g_j(x^i) + \nabla g_j(x^i)^T (x^{i+1} - x^i) = 0 \quad \forall j \in I$

Nous avons aussi : $\nabla f(x^i) + G^i (x^{i+1} - x^i) + \nabla g(x^i) U_I^{i+1} = 0$

finalement : $g_j(x^i) + \nabla g_j(x^i)^T (x^{i+1} - x^i) = 0 \quad \forall j \in I$

$$\nabla f(x^i) + G^i (x^{i+1} - x^i) + \nabla g(x^i) (u_I^{i+1} - u_I^i) + \nabla g(x^i) u_I^i = 0.$$

Nous remplissons donc bien les conditions de l'algorithme de la section B.

#

4 introduction d'un nouvel algorithme

Nous venons de présenter un algorithme basé sur la recherche de points de Kuhn Tucker. Nous avons vu l'équivalence entre cet algorithme et celui présenté dans la section B. Nous allons maintenant introduire un nouvel algorithme, basé lui aussi sur la recherche de points de Kuhn Tucker.

Considérons le problème :

$$\left|\begin{array}{ll} \underset{x,r}{\text{minimiser}} & v f(x) + e^T r \\ \text{sous les contraintes :} & g(x) \leq r \\ & r \geq 0 \end{array}\right. \qquad e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

remarques :

Ce problème est le même (c'est à dire équivalent) que :

$$\underset{x \in R^n}{\text{minimiser}} \; v f(x) + \sum_{j=1}^{m} g_j(x)_+$$

le premier problème atteint en effet son minimum en r pour la valeur $r_j = g_j(\overset{*}{x})_+$.

notations

Le problème que nous venons d'introduire, nous le noterons P.N.L.M c'est à dire : programme non linéaire modifié

Justification de l'introduction du P.N.L.M

Théorème 2 : Soit $(\bar{x}, \bar{u})$ le point de Kuhn Tucker du P.N.L (1) vérifiant les conditions du second ordre.

soit : $\bar{r} = 0$, $\bar{w} = \bar{u}.v$, $\bar{s} = e - \bar{w}$

$(0 < v < \|\bar{U}\|_{\infty}^{-1})$

Alors :

Les conditions de Kuhn Tucker et les conditions du second ordre sont vérifiées pour le P.N.L.M au point $(\bar{x}, \bar{r}, \bar{w}, \bar{s})$ où $\bar{s}$ est le multiplicateur associé à $r \geq 0$

Commentaires :

Ce théorème est la justification de l'introduction du programme non linéaire modifié. Nous voyons en effet qu'à partir des points de Kuhn Tucker du P.N.L nous pouvons déterminer des points de Kuhn Tucker du P.N.L.M La réciproque est bien entendu vrai

<u>démonstration</u>

première partie : $(\bar{x}, \bar{\pi}, \bar{w}, s)$ est un point de kuhn Tucker du P.N.L.M

<u>pas 1</u> : <u>la première condition de kuhn Tucker</u> :

appliqué au P.N.L.M cela donne :

$$v \nabla f(\bar{x}) + \nabla g(\bar{x}) \bar{w} = 0 \qquad (1)$$

$$e - \bar{w} - \bar{s} = 0 \qquad (2)$$

(1) : $v (\nabla f(\bar{x}) + \nabla g(\bar{x}) \bar{u}) = v \cdot 0 = 0$

(2) : vrai par hypothèse.

<u>pas 2</u> : <u>la deuxième condition de kuhn Tucker</u>.

c'est à dire ici : $g(\bar{x}) \leq \bar{\pi}$ (1)

$\bar{\pi} \geq 0$. (2).

(2) : vrai car $\bar{\pi} = 0$.

(1) : $g(\bar{x}) \leq 0$ car $\bar{x}$ est un point admissible du P.N.L

<u>pas 3</u> : <u>la troisième condition de kuhn Tucker</u>

appliqué au P.N.L.M cela devient :

$$\bar{w} \geq 0. \qquad (1)$$

$$\bar{s} \geq 0. \qquad (2)$$

(1) : $\bar{w} = \bar{u} \cdot v \geq 0$.

(2) : $\bar{s} = e - \bar{w}$

$$\bar{s}_j = 1 - \bar{w}_j = 1 - v \bar{u}_j$$

$$> 1 - \bar{u}_j \| \bar{u} \|_\infty^{-1}$$

$$> 1 - \bar{u}_j \bar{u}_j^{-1}$$

$$> 0$$

donc $\bar{s} > 0$.

pas 4 : La quatrième condition de Kuhn Tucker :

appliqué au P.N.L.M :

$$\bar{w}^T (g(\bar{x}) - \bar{r}) = 0 \qquad (1)$$

$$- s^T (\bar{r}) = 0 \qquad (2)$$

(1) $\bar{r} = 0$ donc il faut que $\bar{w}^T (g(\bar{x})) = 0$

$$\bar{w}^T (g(\bar{x})) = v \; \underbrace{\bar{U}^T g(\bar{x})}_{0} = 0$$

(2) $\bar{r} = 0$ donc $- s^T (0) = 0$.

deuxième partie : $(\bar{x}, \bar{r}, \bar{w}, \bar{s})$ vérifie les conditions du second ordre pour le P.N.L.M.

---

pas 1 : la première condition

a) soit e_j le vecteur de $\mathbb{R}^m = (\delta_j^i)$

b) Quelles sont les contraintes actives en $(\bar{x}, \bar{r})$:

1) : $g_I(x) - r_I \leq 0$.

2) : $r_j \leq 0 \qquad j = 1 \ldots m$

c) soit (t, q) orthogonal aux gradients des contraintes actives en $(\bar{x}, \bar{r})$

pour les contraintes du type 1 :

$$(t, q) \begin{bmatrix} \nabla g_j(\bar{x}) \\ - e_j \end{bmatrix} = 0 \qquad j \in I.$$

pour les contraintes du type 2 :

$$(t, q) \begin{bmatrix} 0 \\ - e_j \end{bmatrix} = 0 \qquad j = 1 \ldots m$$

d) propriétés de (t, q)

: $\nabla g_I(\bar{x})^T t = 0$

$q = 0$.

e) en conclusion

nous savons que $\bar{s} > 0$

les hypothèses affirment que $\bar{v}_I > 0$.

d) entraîne que $\nabla g_I(\bar{x})^T t = 0$

et donc, par la première condition du second ordre appliquée au P.N.L :

$$t\,[\nabla^2 f(\bar{x}) + \sum_{j=1}^{m} \bar{v}_j \nabla^2 g_j(\bar{x})]\, t > 0$$

donc : $v\, t\,[\nabla^2 f(\bar{x}) + \sum_{j=1}^{m} \bar{v}_j \nabla^2 g_j(\bar{x})]\, t > 0$

: $t\,[\nabla^2 v f(\bar{x}) + \sum_{j=1}^{m} v \bar{v}_j \nabla^2 g_j(\bar{x})]\, t > 0$

$$t\,[\nabla^2 v f(\bar{x}) + \sum_{j=1}^{m} v\bar{v}_j \nabla^2 g_j(\bar{x})]\, t > 0.$$

de plus $q = 0$

Et donc :

$$[t, q] \begin{bmatrix} v\nabla^2 f(\bar{x}) + \sum_{j=1}^{m} v\bar{v}_j \nabla^2 g_j(\bar{x}) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} t \\ q \end{bmatrix} = 0$$

$\downarrow$

ceci représente le Hessien du lagrangien pour le P.N.L.M.

pas 2 : la deuxième condition

Les gradients des contraintes actives en $(\bar{x}, \bar{\pi})$ sont :

$$[\ \nabla g_j(\bar{x}),\ -e_j\ ] \qquad j \in I.$$

$$[\ 0,\ -e_j\ ] \qquad j \in \bar{m}.$$

Ces vecteurs sont linéairement indépendants par définition des e_j et par l'indépendance de la famille $\{\nabla g_I(\bar{x}),\ j \in I\}$

pas 3 : la troisième condition

Nous savons déjà que $\bar{s} > 0$.

Nous savons aussi que $\bar{v}_j > 0$, $j \in I$.

donc $\bar{v}_j\, v > 0$ et $\bar{w}_j > 0$ $\quad j \in I$.

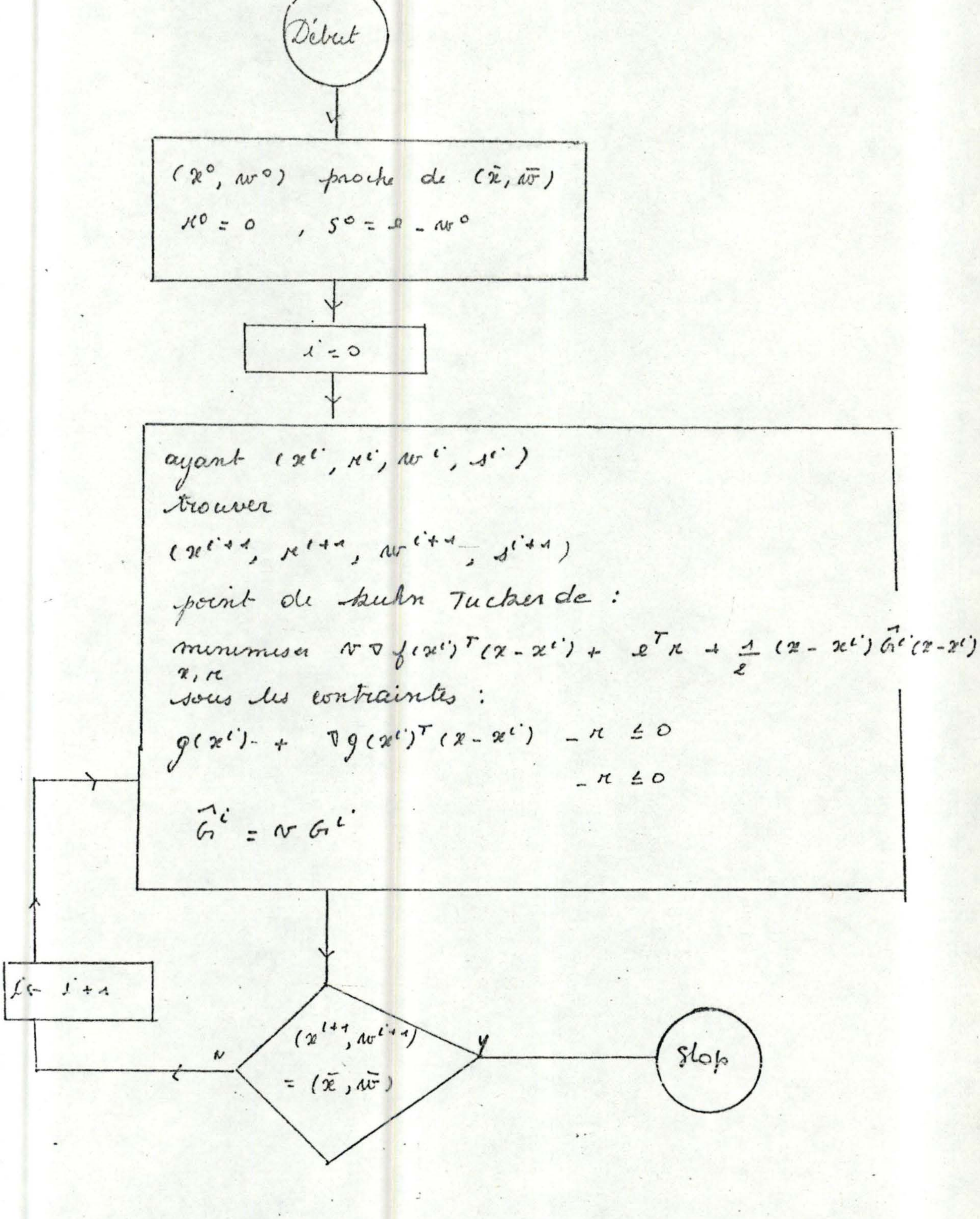
#

remarques

Remarquons que la démonstration s'est effectuée en deux étapes
1) démontrer que $(\bar{x}, \bar{\pi}, \bar{w}, \bar{s})$ est un point de Kuhn Tucker du P.N.L.M
2) démontrer que les conditions du second ordre sont vérifiées.

Pour démontrer la première partie, on a utilisé qu'une hypothèse : $(\bar{x}, \bar{u})$ est un point de Kuhn Tucker du P.N.L (1). On a donc aussi l'équivalence des points de Kuhn Tucker simples

5 description de l'algorithme



Commentaires :

Le fait d'avoir introduit un nouveau programme non linéaire à partir du P.N.L (1) nous a permis d'introduire un nouvel algorithme. Il est basé sur le fait suivant :

Au problème quadratique :

(1) $\underset{y \in \mathbb{R}^n}{\text{minimiser}} \quad \nabla f(x)^T (y-x) + \frac{1}{2} (y-x) G (y-x)$

sous les contraintes : $g(x) + \nabla g(x)^T (y-x) \leq 0$

On applique la transformation :

(2) : $\underset{y \in \mathbb{R}^n, \, r \in \mathbb{R}^m}{\text{minimiser}} \quad v \, \nabla f(x)^T (y-x) + \frac{1}{2} (y-x) \, v \, G (y-x) + e^T r$

sous les contraints : $g(x) + \nabla g(x)^T (y-x) - r \leq 0$

$-r \leq 0$.

Et c'est à ce problème qu'on applique la recherche des points de Kuhn Tucker. Par le theorème precedent nous savons que le point de Kuhn Tucker engendré par le problème (2) fournira un point de Kuhn Tucker du problème 1. Les deux algorithms presentés dans la section 6. sont donc equivalents. Le dernier algorithme minimise en fait $P(v, x)$ tandis que l'avant dernier cherche la solution du P.N.L (1) et ces deux algorithmes sont equivalents, ceci est sans doute à mettre en rapport l'equivalence du P.N.L (1) et le fait de minimiser $P(v, x)$.

CHAPITRE IV : ALGORITHME GLOBAL DE CONVERGENCE SUPERLINEAIRE

A Introduction

Dans ce dernier chapitre, nous allons presenter un algorithme de convergence globale et superlineaire.

Le principe est de demarrer avec l'algorithme global presenté au chapitre 2 et une fois "assez proche" de $(\bar{x},\bar{u})$, - point de R^{n+m} verifiant les conditions de Kuhn Tucker et les conditions du deuxième ordre pour le P.N.L (1), - on calcule l'itération par une des methodes presentées au chapitre 3. Mais si les iterés engendrés par cette methode, ne fournissent pas une position de descente pour la fonction de penalisation $P(v,x)$ on calcule l'iteré par la methode du chapitre 2.

Le théoreme de convergence montrera que si $(\bar{x},\bar{u})$ est un point d'accumulation de la suite engendrée par l'algorithme, alors l'algorithme converge superlinéairement vers $(\bar{x},\bar{u})$

L'algorithme fait intervenir une suite de matrice G^i; cette suite doit approcher la suite $\{\nabla_{xx} L(x^i,u^i)\}$. Un procedé d'ajustement des matrices G^i sera presentée dans la derniere partie du chapitre 4.

B : description de l'algorithme

Pour faciliter les notations, nous appelerons algorithme 2 l'algorithme présenté au deuxième chapitre et nous appelerons algorithme 3 le premier algorithme présenté au chapitre 3. L'algorithme 4 sera lui l'algorithme du quatrième chapitre.

Le principe de l'algorithme 4 est simple : joindre l'algorithme 3 à l'algorithme 2 de manière à obtenir un algorithme 4 qui converge superlinéairement vers $(\bar{x},\bar{u})$.

Notons que nous utilisons l'algorithme 3 réduisant le P.N.L (1) en une suite d'équations linéaires. Nous pourrions aussi utiliser l'algorithme réduisant le P.N.L (1) en une suite de programmes quadratiques

Initialisation de l'algorithme :

On démarre avec un vecteur $x^0 \in \mathbb{R}^n$ quelconque et un vecteur $u^0 \in \mathbb{R}^m$ quelconque. (x^i, u^i) sera le couple d'itérés appartenant à R^{n+m} qui convergera vers $(\bar{x},\bar{u})$

La matrice G^0 est elle aussi quelconque, les matrices G^i appartenant à $R^{n\times n}$ seront déterminées par une procédure d'ajustement que nous étudierons plus tard et qui sera telle que

$$\lim_{i\to\infty} G^i = \nabla_{xx} L(\bar{x},\bar{u}) .$$

La constante c appartenant à $]0,1]$ est un paramètre généralisant la fonction σ_0 imposée par l'algorithme 2.
La constante $\gamma \in]0,1[$ est le paramètre qui ajuste v^i, variable auxilliaire de la fonction de pénalisation $P(v,x)$
Le paramètre $v^0 \in]0,1]$

première phase

La première phase effectue une itération par l'algorithme 2 de manière à se rapprocher de $(\bar{x},\bar{v})$.
La procedure de fonction nous renverra parfois à la première phase lorsque la deuxième phase ne sera pas utilisable. L'algorithme 2 determine un couple (x^{i+1}, w^{i+1}), nous poserons $v^{i+1} = w^{i+1}.(v^i)^{-1}$.

deuxième phase

La deuxième phase, pour autant que la procedure de fonction le permette, determinera (x^{i+1}, u^{i+1}) par l'intermédiaire de l'algorithme 3.

(x^{i+1}, u_J^{i+1}) sera la solution de

$$\begin{bmatrix} G^i & \nabla g_J(x^i) \\ \nabla g_J(x^i)^T & 0 \end{bmatrix} \begin{bmatrix} x^{i+1} - x^i \\ u_J^{i+1} - u_J^i \end{bmatrix} + H_J(x^i, v_J^i) = 0$$

Notons cette solution $(y^i, \tilde{z}_J^i)$., $\tilde{z}^i_{\overline{m}\setminus J} = 0 = u^{i+1}_{\overline{m}\setminus J}$

ou alors (x^{i+1}, u_J^{i+1}) sera la solution de

$$\begin{bmatrix} G^i & \nabla g_J(x^i) \\ \nabla g_J(x^i)^T & 0 \end{bmatrix} \begin{bmatrix} x^{i+1} - y^i \\ u_J^{i+1} - \tilde{z}_J^i \end{bmatrix} + H_J(y^i, \tilde{z}_J^i) = 0$$

nous noterons cette solution (t^i, s_J^i), $u_{\overline{m} \setminus J}^{i+1} = s_{\overline{m} \setminus J}^i = 0$
Il est à noter que les deux systèmes sont compatibles ou incompatibles ensembles.

En lignes générales, on prendra (x^{i+1}, v^{i+1}) solution du premier système s'il en résulte une direction de descente non négligeable pour la fonction de pénalisation sinon on choisira $(x^{i+1}, \bar{v}^{i+1})$ solution du deuxième système s'il en résulte une direction de descente non négligeable pour la fonction de pénalisation. Si ce n'est pas le cas, on détermine (x^{i+1}, u^{i+1}) par la première phase. A l'intérieur de la deuxième phase, se trouve le procédé de jonction

procédé de jonction

Le procédé de jonction permet de passer de la phase 1 à la phase 2.
Pour appliquer la phase 2, il faut qu'elle ait un sens, c'est à dire que le système $LS(x^i, v^i, G^i, J) = 0$ soit compatible, si ce n'est pas le cas, le procédé de jonction nous renvoie à la première phase.

L'algorithme 3 détermine le couple (x^{i+1}, u^{i+1}), mais nous aimerions aussi déterminer un vecteur v^{i+1} comme dans l'algorithme 2, de manière à garder une certaine analogie avec cet algorithme. Nous ne voulons en effet pas perdre les propriétés que nous connaissons sur l'algorithme 2.

Comme dans l'algorithme 2, nous voulons que :

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq \sigma_0 (\phi_N (v^i; x^i; v^{i+1}))$$

et cela pour assurer une direction de descente à la fonction de pénalisation

Nous nous assurons d'abord que les vecteurs $\tilde{z}^i$ et s^i définis sont positifs ou nuls, si ce n'est pas le cas, nous sommes "trop loin" de $\bar{u}$ et nous déterminons (x^{i+1}, u^{i+1}) par la première phase.

Nous définissons ensuite deux vecteurs, z^i et $\overset{*}{z}^i$.

$$z^i_j = \begin{cases} 1 & \text{si } v^i \tilde{z}^i_j > 1 \\ v^i \tilde{z}^i_j & \text{sinon} \end{cases} \qquad (1)$$

$$\overset{*}{z}^i_j = \begin{cases} 1 & \text{si } v^i s^i_j > 1 \\ v^i s^i_j & \text{sinon.} \end{cases} \qquad (1)$$

Et v^{i+1} prendra la valeur z^i ou $\overset{*}{z}^i$ suivant que $(x^{i+1}, u^{i+1}) = (y^i, \tilde{z}^i)$ ou (t^i, s^i)

On prend la précaution (1, pour être sûr que (v^{i+1}) appartienne à $T(x^i)$

Comme nous avons défini $u^{i+1} = w^{i+1}(v^i)^{-1}$ pour la phase 1, nous aimerions que $w^{i+1} = u^{i+1} v^i$ nous verrons que cela sera vrai pour i assez grand

En résumé la procédure de fonction renvoie à la phase si une des éventualités suivantes n'est pas satisfaites :
1) $LS(x^i, u^i, v^i; J) = 0$ a une solution unique
2) $\tilde{z}^i \geq 0$
3) $s^i \geq 0$
4) $P(v^i, x^i) - P(v^i, t^i) \geq c \min \left\{ \phi_N(v^i, x^i, \overset{*}{z}^i) , (\phi_N(v^i, x^i, \overset{*}{z}^i))^3 \right\}$.

La procedure de fonction ne permettra pas de prendre pour (x^{i+1}, u^{i+1}) la valeur $(y^i, \tilde{z}^i)$ si

$$P(v^i, x^i) - P(v^i, y^i) \geq c \min \left\{ \phi_N(v^i, x^i, z^i), (\phi_N(v^i, x^i, z^i))^3 \right\}$$

elle nous proposera alors de prendre pour (x^{i+1}, u^{i+1}) la valeur (t^i, s^i) pour autant que les conditions 1) 2) 3) 4) soient vérifiées
La procédure de fonction assure donc une direction de descente non négligeable pour la fonction de pénalisation

A propos du paramètre v^i :

Nous ne parlons pas explicitement de v^{i+1} dans la phase 1 et la phase 2 parce que les deux phases le déterminent de la même manière, celle proposée au chapitre 2.

Dans la suite, nous dirons que (x^{i+1}, u^{i+1}) est déterminé par la phase 1, s'il est déterminé par l'algorithme 2 et par la phase 2, s'il est déterminé par l'algorithme 3

Propriété 1 : | L'algorithme 4 est un cas particulier de l'algorithme 2

démonstration :

si (x^{i+1}, u^{i+1}) est déterminé par la première phase, le deuxième chapitre assure que :

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq \min \left\{ \frac{1}{2} \varphi_v(v^i, x^i, w^{i+1}), \frac{1}{8} \frac{\varphi_v(v^i, x^i, w^{i+1})^3}{k_{21} k_{22}^2} \right\}$$

où les constantes k_{12}, k_{22} sont définies au chapitre 2.

si (x^{i+1}, u^{i+1}) est déterminé par la deuxième phase, l'algorithme 4 assure que :

$$P(v^i, x^i) - P(v^i, x^{i+1}) \geq c \min \left\{ \varphi_v(v^i, x^i, w^{i+1}), \varphi_v(v^i, x^i, w^{i+1})^3 \right\}$$

#

Conséquence :

Le théorème de convergence étudié au chapitre 2 reste valable pour l'algorithme 4 et si $(\bar{x}, \bar{v})$ est un point d'accumulation de (x^i, v^i) - avec $(\bar{x}, \bar{v})$ point de Kuhn Tucker pour le P.N.L (1) - nous aurons que :

$$\bar{v} = \lim_{i \to \infty} v^i > 0$$

propriété 2 : Il existe i_1 appartenant à $\mathbb{N}$ tel que $\forall i > i_g$, on a : $v^{i+1} = u^{i+1} v^i$ pour autant que $\lim_{i \to \infty} v^i > 0$.

démonstration

si (x^{i+1}, u^{i+1}) est déterminé par la phase 1 alors : $v^{i+1} = u^{i+1} v^i$ quel que soit i

supposons que (x^{i+1}, u^{i+1}) soit déterminé par la phase 2 le fait que $\bar{v} > 0$ entraîne que :

$$\exists i_g, \quad i > i_g \quad \Rightarrow \quad v^i = v^{i+1} = \bar{v}$$

et donc par un raisonnement semblable à celui présenté dans le théorème de convergence de l'algorithme 2 nous pouvons affirmer sans perdre de généralités :

$$\exists i_g, \quad i > i_g \quad \Rightarrow \quad \|W^{i+1}\|_\infty < \frac{3}{4}. \qquad (a)$$

rappelons que

$$W_j^{i+1} = \begin{cases} 1 & \text{si } v^i u_j^{i+1} > 1 \\ v^i u_j^{i+1} & \text{sinon.} \end{cases} \qquad (b)$$

les relations (a) et (b) permettent de dire que

$$\exists i_g, \quad i > i_g \quad \Rightarrow \quad v^{i+1} = u^{i+1} v^i. \qquad \#$$

Conséquence :

$$\exists i_g, \quad i > i_g \quad \Rightarrow \quad \|v^{i+1}\|_\infty = \|u^{i+1}\|_\infty \bar{v}$$

et donc $\frac{3}{4} > \|u^{i+1}\|_\infty \bar{v}$

et donc $\frac{3}{4} > \|\bar{u}\|_\infty \bar{v}$

et $\bar{v} < \frac{3}{4} \|\bar{u}\|_\infty^{-1}$.

C Théorème de convergence

<u>Théorème</u> :

Si la suite (x^i) est finie, elle se termine en un point de Kuhn Tucker du P.N.L (1)

supposons que :

1) la suite (x^i) est ∞ et $(\bar{x})$ est un point d'accumulation de (x^i), $\bar{x}$ est tel qu'il existe $\bar{u} \in \mathbb{R}^m$ tel que $(\bar{x}, \bar{u})$ vérifie les conditions de Kuhn Tucker et les conditions du second ordre pour le P.N.L (1)

2) $\exists\, i_0$, $i \geq i_0$ entraîne que :

$\| G^i - \nabla_{xx} L(x^i, \bar{u}) \| \leq (2\beta_x)^{-1}$

$\| G^i - \nabla_{xx} L(x^i, u^i) \| \leq (10\beta)^{-1}$

3) soit $L = \{ i,\ LS(x^i, u^i, G^i, J) = 0.$ a une solution unique $\}$

et $$\lim_{i \in L} \frac{\| (G^i - \nabla_{xx} L(x^i, \bar{u}))(y^i - x^i) \|}{\| y^i - x^i \|} = 0$$

et $$\lim_{i \in L} \frac{\| (G^i - \nabla_{xx} L(x^i, \bar{u}))(t^i - y^i) \|}{\| t^i - y^i \|} = 0$$

Alors on a que :

$$\lim_{i \to \infty} \frac{\| (x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u}) \|}{\| (x^i, u^i) - (\bar{x}, \bar{u}) \|} = 0$$

Commentaires

première hypothèse :

On suppose donc que la suite (x^i, v^i) a comme point d'accumulation un couple $(\bar{x}, \bar{v})$ vérifiant les conditions de Kuhn Tucker et les conditions du second ordre pour le P.N.L (1). On peut tirer cette conclusion de la première hypothèse grâce à un raisonnement similaire à celui effectué dans le théorème de convergence du chapitre 2.

A priori rien ne nous permet de justifier cette hypothèse parce que nous savons que les points d'accumulation de l'algorithme 2 – et donc aussi de l'algorithme 4 sont soit des points de Kuhn Tucker soit des points stationnaires de $P(0, x)$

deuxième hypothèse

La première assertion est nécessaire pour appliquer le théorème fondamental du chapitre 3. en effet pour (x^i, u^i) proche de $(\bar{x}, \bar{v})$, la première assertion est une condition suffisante pour assurer la consistance du système $LS(x^i, u^i, \sigma^i, J) = 0$

La deuxième assertion est utile lorsqu'on utilise pour la phase 2, le principe qui réduit le P.N.L (1) en une suite de programmes quadratiques

Troisième hypothèse.

La troisième hypothèse indique que le comportement de G^i et $\nabla_{x^2} L(x^i, \bar{u})$ sont les mêmes vis à vis de $(y^i - x^i)$ et $(t^i - y^i)$., propriétés nécessaires comme nous le verrons, pour appliquer un lemme technique. version du théorème fondamental du chapitre précédent. On prend la précaution d'isoler les éléments pour lesquels $LS(x^i, u^i, b^i, J) = 0$ n'est pas consistant, en effet, que vaut alors y^i et t^i.

Conclusion:

La suite (x^i, u^i) converge superlinéairement vers $(\bar{x}, \bar{u})$

principe de démonstration

Le principe est de parvenir à l'aide de lemmes techniques à ceci :

si on parvient à déterminer un couple (x^i, u^i) par la phase 1 ou 2 et que ce couple est "assez proche" de $(\bar{x}, \bar{u})$ - fait justifié par les hypothèses d'accumulation. - alors, toutes les itérations suivantes seront déterminées par la phase 2.

Lemme 1 : $\forall\ 0 < v < \|\bar u\|_\infty^{-1}$.

Il existe $\varepsilon, K_S, K_Q, K_{SQ} > 0$

tels que, $\forall x \in B(\bar x, \varepsilon)$

(a) : $P(v, x) - P(v, \bar x) \geq K_{QS}\ \|x - \bar x\|^2$

(b) : $P(v, x) - P(v, \bar x) \leq K_Q\ \|x - \bar x\|^2 + K_S\ \|x_S - \bar x_S\|$

Commentaires

La propriété 2 nous permet d'appliquer ce lemme au paramètre $\bar v = \lim_{i \to \infty} v^i > 0$

Le théorème présente deux assertions, nous les démontrerons chacunes séparement

On choisit tout d'abord la valeur ε de telle manière que dans la boule $B(\bar x, \varepsilon)$ les contraints strictement négatives en $\bar x$ le restent.

On introduit ensuite W qui est la sphère unité de l'espace tangent aux gradients des contraints actives en $\bar x$

on définit Q_y : $y^T \nabla_{xx} L(\bar x, \bar u)\ y \quad \forall y \in W$

on choisit alors $K_{SQ} < \frac{1}{2} \overline{K_{SQ}} = \frac{1}{2} \min_{y \in W} Q_y$

La démonstration se poursuit en utilisant l'hypothèse de l'absurde. On suppose que K_{SQ} ne répond pas à la thèse, cela nous amène à introduire une suite $- x^i = \bar x + \delta_i\, l^i$, $\|l^i\| = 1$ — niant la thèse.

ce qui signifie en gros :

$$K_{SQ}\ (\delta_i)^2 > P(v; x^i) - P(v, \bar x)$$

On développe ce résultat en série de Taylor du premier ordre, et on arrive à la conclusion que $\bar l = \lim_{i \to \infty} l^i \in W$.

On développe de nouveau l'inégalité en série de Taylor du deuxième ordre. pour obtenir cette fois une contradiction : $KSQ \geq \frac{1}{2}\overline{KSQ} > KSQ$.
La deuxième partie de la démonstration est plus simple, on utilise le fait que $f, g_1 \cdots g_m$, $\nabla f, \cdots \nabla g_1 \cdots \nabla g_m$, $\nabla^2 f \cdots \nabla^2 g_m$ sont k Lipschitz continues

démonstration

première partie : démonstration de la première assertion.

pas 1 : choix de ε.
on choisit ε tel que :
$$\left.\begin{array}{l} x \in B(\bar{x}, \varepsilon) \\ g_j(\bar{x}) < 0 \end{array}\right\} \text{ entraine que } g_j(x) < 0$$

pas 2: choix de KSQ
soit
$$W = \left\{ y \in \mathbb{R}^n,\ \|y\| = 1,\ \begin{array}{ll} \nabla g_j(\bar{x})^T y = 0 & \text{si } \bar{u}_j > 0 \\ \nabla g_j(\bar{x})^T y \leq 0 & \text{si } j \in I,\ \bar{u}_j = 0 \end{array} \right\}$$

W est compact et les conditions du second ordre sont satisfaites en $(\bar{x}, \bar{u})$ donc :
$$a_y = y^T \nabla_{xx} L(\bar{x}, \bar{u})\, y > 0$$
et
$$\overline{KSQ} = \min_{y \in W} a_y > 0$$

soit KSQ tel que $0 < KSQ < \frac{1}{2}\overline{KSQ}$.

pas 3 : contredisons (a), qu'obtient-on :

en contredisant (a) on a :

$\forall \varepsilon, \exists x \in B(\bar{x}, \varepsilon)$ tel que $P(v, \bar{x}) - P(v, x) < KSQ \|x - \bar{x}\|^2$

c'est à dire : $\exists (x^i)_{i=0}^{\infty}$ qui converge vers $\bar{x}$

telle que : $P(v, x^i) - P(v, \bar{x}) < K_{SQ} \|x^i - \bar{x}\|^2$

soit $x^i = \bar{x} + \delta_i \cdot l^i$ avec $\lim\limits_{i \to \infty} \delta_i = 0$

$\|l^i\| = 1 \ \forall i$

la suite (l^i) étant bornée, il existe une sous suite de (l^i) qui converge.

c'est à dire :

Il existe $L_0 \subset \mathbb{N}$, $\lim\limits_{i \in L_0} l^i = \bar{l}$.

on a donc : $KSQ(\delta^i)^2 > P(v, x^i) - P(v, \bar{x})$.

pas 4 : developpons le pas 3 en serie de Taylor.

nous savons que $g_j(x^i) = \nabla g_j(\bar{x} + \delta'_{ij} l^i)^T \delta_i \cdot l^i$

$0 < \delta'_{ij} < 1$

et donc : $\left(g_j(x^i)\right)_+ = \left[\nabla g_j(\bar{x} + \delta'_{ij} l^i)^T \delta_i \cdot l^i\right]_+$

nous avons donc :

$$KSQ(\delta^i)^2 > P(v, x^i) - P(v, \bar{x})$$
$$= v \nabla f(\bar{x} + \delta'_i l^i)^T \delta_i \cdot l^i$$
$$+ \sum_{j \in I} (\nabla g_j(\bar{x} + \delta'_{ij} l^i)^T \delta_i \cdot l^i)_+ .$$

$\checkmark$ $(\bar{x})$ est admissible, donc seules les contraintes actives en $\bar{x}$ peuvent devenir strictement positives.

$0 < \delta_i' , \delta_{ij}' < \delta_i$ par le théorème de la moyenne

divisons l'inégalité obtenue par δ^i et faisons tendre i vers l'infini, nous obtenons :

$$0 \geq v \nabla f(\bar{x})^T \bar{l} + \sum_{j \in I} (\nabla g_j(\bar{x})^T \bar{l})_+ .$$

$$= - v \sum_{j \in I} \bar{u}_j \nabla g_j(\bar{x})^T \bar{l} + \sum_{j \in I} (\nabla g_j(\bar{x})^T \bar{l})_+ .$$

parce que $(\bar{x}, \bar{u})$ vérifie les conditions de Kuhn Tucker pour le P.N.L (1)

$$= \sum_{j \in I} \Big[(1 - v \bar{u}_j)(\nabla g_j(\bar{x})^T \bar{l})_+ + v \bar{u}_j ((\nabla g_j(\bar{x})^T \bar{l})_+ - \nabla g_j(\bar{x})^T \bar{l}) \Big] \geq 0$$

parce que $1 - v \bar{u}_j > 0 \quad (v < \|\bar{u}\|_\infty^{-1})$

$(\nabla g_j(\bar{x})^T \bar{l})_+ \geq \nabla g_j(\bar{x})^T \bar{l}$

<u>pas 5</u> : <u>conséquences du pas 4 pour $\bar{l} = \lim_{i \in L_0} l^i$</u>

$$\left.\begin{array}{ll} 1) : & (1 - v \bar{u}_j)(\nabla g_j(\bar{x})^T \bar{l})_+ = 0 \\ 2) & v \bar{u}_j ((\nabla g_j(\bar{x})^T \bar{l})_+ - \nabla g_j(\bar{x})^T \bar{l}) = 0 \end{array}\right\} j \in I .$$

1) entraîne que $(\nabla g_j(\bar{x})^T \bar{l})_+ = 0$

parce que $(1 - v \bar{u}_j) > 0$

donc $\nabla g_j(\bar{x})^T \bar{l} \leq 0$,

2) nous savons que $\forall j \in I$, $\bar{v}_j > 0$

donc : $(\nabla g_j(\bar{x})^T \ell)_+ - \nabla g_j(\bar{x})^T \ell = 0$

et donc $\nabla g_j(\bar{x})^T \ell = 0$ car $(\nabla g_j(\bar{x})^T \ell)_+ = 0$

nous avons donc que $\bar{\ell} \in W$

<u>pas 6:</u> <u>développons de nouveau le pas 3 en série</u>.

$$k s_a (\delta_i)^2$$

$$> v f(x^i) - v f(\bar{x}) + \sum_{j \in I} g_j(x^i)_+$$

$$> v f(x^i) - v f(\bar{x}) + \sum_{j \in I} \underbrace{v \bar{v}_j}_{< 1} g_j(x^i)_+$$

$$= v \Big\{ \delta_i \nabla f(\bar{x})^T \ell^i + \frac{1}{2} (\delta_i)^2 (\ell^i)^T \nabla^2 f(\bar{x} + \hat{\delta}_i \ell^i) \ell^i$$

$$+ \sum_{j \in I} \Big[\delta_i \bar{v}_j \nabla g_j(\bar{x})^T \ell^i$$

$$+ \frac{1}{2} (\delta_i)^2 (\ell^i)^T \nabla^2 g_j(\bar{x} + \hat{\delta}_{ij} \ell^i) \ell^i \cdot \bar{v}_j \Big] \Big\}.$$

avec $0 < \hat{\delta}_i, \hat{\delta}_{ij} < \delta_i$

$$= v \Big\{ \delta_i \big(\nabla f(\bar{x}) + \sum_{j \in I} \bar{v}_j \nabla g_j(\bar{x})\big)^T \ell^i$$

$$+ \frac{1}{2} (\delta_i)^2 \Big[(\ell^i)^T \Big(\nabla^2 f(\bar{x} + \hat{\delta}_i \ell^i)$$

$$+ \sum_{j \in I} \nabla^2 g_j(\bar{x} + \hat{\delta}_{ij} \ell^i) \bar{v}_j \Big) \Big] \ell^i \Big\}$$

pas 7 : recherche de la contradiction :

Dans la dernière relation obtenue, on divise les deux membres par $(\delta^i)^2$ et on fait tendre i vers l'infini on obtient :

$$KSQ \geq \frac{1}{2} \bar{\ell}^T (\nabla^2 f(\bar{x}) + \sum_{j \in I} \bar{u}_j \nabla^2 g_j(\bar{x})) \bar{\ell})$$

$$\geq \frac{1}{2} K\overline{SQ} \qquad \text{car } \bar{\ell} \in W$$

hors $KSQ < \frac{1}{2} K\overline{SQ}$

deuxième partie : démonstration de la deuxième assertion

soit ε petit tel que

$\nabla f, \nabla g_1 \ldots \nabla g_m, \nabla^2 f, \nabla^2 g_1 \ldots \nabla^2 g_m$ soient lipschitz continus dans $B(\bar{x}, \varepsilon)$

soit Q le complément orthogonal de S

donc, $\forall y \in \mathbb{R}^n$, on a : $y_Q^T \nabla g_I(\bar{x}) = 0$.

on a aussi que $\nabla f(\bar{x}) + \nabla g_I(\bar{x}) \bar{v}_I = 0$

donc : $y_Q^T \nabla f(\bar{x}) = 0$

$$
\begin{aligned}
. \; & P(v,x) - P(v,\bar{x}) \\
&= v \nabla f(\bar{x})^T (x-\bar{x}) + \frac{v}{2}(x-\bar{x})^T \nabla^2 f(\bar{x})(x-\bar{x}) \\
&\quad + \sum_{j\in I} \left[\nabla g_j(\bar{x})^T (x-\bar{x}) + \frac{1}{2}(x-\bar{x})^T \nabla^2 g_j(\bar{x})(x-\bar{x})\right]_+ \\
&\quad + \beta(x,\bar{x}) \|x-\bar{x}\|^2 \qquad \text{avec } \lim_{x\to\bar{x}} \beta(x,\bar{x}) = 0
\end{aligned}
$$

$$
\begin{aligned}
&= v \nabla f(\bar{x})^T (x_S-\bar{x}_S) + \frac{v}{2}(x-\bar{x})^T \nabla^2 f(\bar{x})(x-\bar{x}) \\
&\quad + \sum_{j\in I} \left[\nabla g_j(\bar{x})^T (x_S-\bar{x}_S) + \frac{1}{2}(x-\bar{x})^T \nabla^2 g_j(\bar{x})(x-\bar{x})\right]_+ \\
&\quad + \beta(x,\bar{x}) \|x-\bar{x}\|^2 .
\end{aligned}
$$

$$
\begin{aligned}
&\leq v \nabla f(\bar{x})^T (x_S-\bar{x}_S) + \sum_{j\in I} \left[\nabla g_j(\bar{x})^T (x_S-\bar{x}_S)\right]_+ \\
&\quad + \frac{v}{2}(x-\bar{x})^T \nabla^2 f(\bar{x})(x-\bar{x}) \\
&\quad + \sum_{j\in I} \frac{1}{2}\left[(x-\bar{x})^T \nabla^2 g_j(\bar{x})(x-\bar{x})\right]_+ + \beta(x,\bar{x})\|x-\bar{x}\|^2 .
\end{aligned}
$$

$$
\leq K_S \|x_S - \bar{x}_S\| + K_Q \|x-\bar{x}\|^2 .
$$

#

Lemme 2. : $\forall \delta \in \,]0,1[$, Il existe $\delta_{11}(\delta), \delta_{12}(\delta) > 0$ tels que :

soient : $L_1 = \{ i \in L,\ i \geq i_0 \}$

$$\Delta_1^i = \frac{\| (G^i - \nabla_{xx} L(x^i,\bar{u}))(y^i - x^i) \|}{\| y^i - x^i \|}$$

$$\Delta_2^i = \frac{\| (G^i - \nabla_{xx} L(x^i,\bar{u}))(y^i - t^i) \|}{\| t^i - y^i \|}$$

si $i \in L_1$

$(x^i, u^i) \in B[(\bar{x},\bar{u}), \delta_{11}]$

$\Delta_1^i, \Delta_2^i \leq \delta_{12}$

Alors on a :

$\| (t^i, s^i) - (\bar{x},\bar{u}) \| \leq \delta \| x^i - \bar{x} \|$

$\| (t_s^i - \bar{x}_s) \| \leq \delta \| x^i - \bar{x} \|^2.$

La démonstration se base sur le théorème fondamental du chapitre précédent.

Il nous dit en particulier : $\forall \delta_{13} \in \,]0,1[$

Il existe $\delta_{11}, \delta_{12}, K_{41}, K_{42} > 0$ tels que :

$\forall (x',u'), (x'',u'') \in B((\bar{x},\bar{u}), 2\delta_{11})$ on a :

1) $E(x', v') = I$

2) $\| \nabla H_I (x', v')^{-1} \| \leq \frac{3}{2} \beta_I$

3) $\| \nabla H_I (x', v') - \nabla H_I (x'', v'') \| \leq K_{41} \| (x', v') - (x'', v'') \|$

si $\| (x^i, u^i) - (\bar{x}, \bar{v}) \| \leq \delta_{11}$ on a :

4) $\| (y^i, \tilde{z}^i) - (\bar{x}, \bar{v}) \| \leq \delta_{13} \| x^i - \bar{x} \|$

$$\| (\tilde{q}^i, \tilde{v}^i) - (\bar{x}, \bar{v}_I) \| \leq K_{42} \| (y^i, \tilde{z}^i_I) - (\bar{x}, \bar{v}_I) \|^2$$

avec : $\nabla H_I (y^i, \tilde{z}^i_I) \begin{bmatrix} \tilde{q}^i - y^i \\ \tilde{v}^i - \tilde{z}^i_I \end{bmatrix} + H_I (y^i, \tilde{z}^i_I) = 0$

En partant de ces valeurs δ_{11}, δ_{12}, on parvient à démontrer que :

$$\| (t^i, s^i) - (\bar{x}, \bar{v}) \| \leq \delta_{15} \| x^i - \bar{x} \|$$

$$\| t^i_s - \bar{x}_s \| \leq \delta_{17} \| x^i - \bar{x} \|^2$$

où δ_{17}, δ_{15} dépendent explicitement de δ_{12}, δ_{11}, δ_{13}. En diminuant δ_{13}, δ_{12}, δ_{11} on obtient δ_{15}, δ_{17} plus petits que δ, on a donc bien que δ_{12}, δ_{11} dépendent de δ ; δ_{13} peut bien entendu être pris arbitrairement petit.

démonstration du théorème

pas 1 : application des lemmes précédents

lemme 1 : $\exists\ \varepsilon,\ K_S,\ K_Q,\ K_{SQ} > 0$

tels que : $\forall x \in B(\bar{x}, \varepsilon)$

$$P(\bar{v}, x) - P(\bar{v}, \bar{x}) \leq K_S \|x_S - \bar{x}_S\| + K_Q \|x - \bar{x}\|^2$$

$$P(\bar{v}, x) - P(\bar{v}, \bar{x}) \geq K_{SQ} \|x - \bar{x}\|^2 .$$

lemme 2 : $\forall\ \bar{\delta}$, si $L_1 = \{ i \in L,\ i \geq i_{10} \}$

$$\left.\begin{array}{l} \exists\ \delta_{11}(\bar{\delta}),\ \delta_{12}(\bar{\delta}) \text{ tels que} \\ \text{si } \forall i \in L_1,\ \Delta_1^i, \Delta_2^i \leq \delta_{12} \\ \text{si } (x^i, u^i) \in B((\bar{x}, \bar{u}), \delta_{11}) \end{array}\right\} \text{(a)}$$

Alors : $\|(t^i, s^i) - (\bar{x}, \bar{u}_z)\| \leq \bar{\delta} \|x^i - \bar{x}\|$

$\|t_s^i - \bar{x}_s\| \leq \bar{\delta} \|x^i - \bar{x}\|^2$

$s^i \geq 0$

théorème du chapitre 3 : $\forall\ \bar{\delta}$, si i vérifie (a)

Alors : $\|(y^i, \tilde{z}^i) - (\bar{x}, \bar{u})\| \leq \bar{\delta} \|x^i - \bar{x}\|$

: $\tilde{z}^i \geq 0$

$E(x^i, u^i) = I$.

Ces résultats techniques sont les résultats de travaux précédents, ils sont indispensables pour arriver à la thèse. En effet si nous parvenons à montrer qu'à partir d'un certain moment (x^{i+1}, u^{i+1}) est soit $(y^i, \tilde{z}^i)$ soit (t^i, s^i), ces résultats techniques permettront de démontrer la convergence superlinéaire.

pas 2 : détermination de $\bar{\delta}$

pour demontrer qu'une suite (a^i) converge vers zero on demontre que :

$$\forall \bar{\delta} , \exists i_{\bar{\delta}} , i \geq i_{\bar{\delta}} \Rightarrow |a^i| \leq \bar{\delta} \qquad (b)$$

il faut donc choisir $\bar{\delta}$ arbitraire et verifier la relation (b)

choisissons $\bar{\delta} > 0$ tel que

$$\frac{1}{2} K_{Sa} - K_a \bar{\delta}^2 - K_S \bar{\delta} \geq 0 \qquad \text{et } \delta > 1 \qquad (c)$$

Cette determination n'enleve rien à la généralité du paramètre $\bar{\delta}$

en effet : soit $\bar{\delta} = 0$ la relation (c) devient :

$$\frac{1}{2} K_{Sa} \geq 0$$

Or fait $\frac{1}{2} K_{Sa} > 0$, on peut donc déterminer un intervalle $[0, d[$, tel que $\forall \delta \in [0, d[$. la relation (c) est satisfaite pas 3 : majoration de ϕ_N

nous savons par le chapitre 2 que

$$\lim_{i \to \infty} \phi_N (v^i, x^i, w^{i+1}) = 0$$

soit $\bar{w} = \bar{u}\,\bar{v}$

alors : $\phi_N (\bar{v}, \bar{x}, \bar{w}) = 0$

Nous savons aussi que $\nabla f, \nabla g_1 \cdots \nabla g_m$ sont k lipchitz continues sur $B(\bar{x}, \varepsilon)$
la fonction ϕ_N est elle aussi lipchitz continue sur $B((\bar{x}, \bar{w}), \varepsilon)$ parce qu'elle s'exprime sous la forme :

$$\phi_N (\bar{v}, x, w) = \| \bar{v} \nabla f(x) + \nabla g(x) w \|_N - g(x)^T w + \sum_{j=1}^{m} g_j(x)_+ .$$

on peut donc ecrire : $\exists\, K_\phi > 0$ tel que :

$$\phi_N (\bar{v}, x, w) \leq K_\phi \left(\| x - \bar{x} \| + \| w - \bar{w} \| \right)$$

pas 4 : définition de $i_{\bar{\delta}}$

rappelons que $L_1 = \{ i \in L \, , \, i \geq i_{10} \}$

a) : les hypothèses du théorème affirment que :

$$\lim_{i \in L} \Delta_1^i = \lim_{i \in L} \Delta_2^i = 0$$

donc : $\forall \delta_{12} \, , \, \exists i_{11} \, , \, i \geq i_{11} \, , \, i \in L \Rightarrow \Delta_1^i \leq \delta_{12} \; (\bar{\delta})$

$$\Delta_2^i \leq \delta_{12} \; (\bar{\delta})$$

b) les hypothèses du théorème affirment que $(\bar{x}, \bar{v})$ est un point d'accumulation de (x^i, v^i)

donc : $\exists M \subset \mathbb{N} \, , \, \lim_{i \in M} (x^i, v^i) = (\bar{x}, \bar{v})$

donc $\exists i_{12} \, , \, i \geq i_{12} \, , \, i \in M \Rightarrow$

$$\| (x^i, v^i) - (\bar{x}, \bar{v}) \| \leq \min \left\{ \delta_{11} \, , \, \frac{2 c K_\varphi^3 (1 + \bar{v} \bar{\delta})^3}{K_{S\Omega}} \right\}$$

c) $\exists i_9 \, , \, i \geq i_9 \Rightarrow v^L = \bar{v}$.

. soit $i \geq i_{10}$

alors $\| G^i - \nabla_{xx} L(x^i, \bar{v}) \| \leq (2\beta_2)^{-1}$

si $i \in M$, $i \geq \max \{ i_{10} \, , \, i_{12} \}$ alors par le théorème fondamental du chapitre 3, on a que le système $LS(x^i, v^i, G^i \; I) = 0$ a une solution unique et donc $i \in L$

notons que si (x^i, v^i) est "assez proche" de $(\bar{x}, \bar{v})$

$J = E(x^i, v^i) = I$

posons $i_{\bar{\delta}} = \max \{ i_9 \, , \, i_{10} \, , \, i_{11} \, , \, i_{12} \}$

pas 5 : si $i > i_{\bar\delta}$, alors $(x^i v^i)$ est déterminé par la phase 2

soit $i = i_{\bar\delta}$

alors : $P(\bar v, x^i) - P(\bar v, t^i)$

$= P(\bar v, x^i) - P(\bar v, \bar x) - (P(\bar v, t^i) - P(\bar v, \bar x))$

$\geq K_{SQ} \|x^i - \bar x\|^2 - k_S \|t^i_s - \bar x_s\| - k_Q \|t^i - \bar x\|^2$

par le lemme 1

remarquons que δ_{11} peut être pris assez petit de manière à ce que $t^i, s^i \in B(\bar x, \varepsilon)$.

$\geq K_{SQ} \|x^i - \bar x\|^2 - k_S \bar\delta \|x^i - \bar x\|^2 - k_Q \bar\delta^2 \|x^i - \bar x\|^2$

par le lemme 2

remarquons que $i_{\bar\delta}$ répond à toutes les hypothèses du lemme 2

$\geq \frac{1}{2} K_{SQ} \|x^i - \bar x\|^2$

parce que $\frac{1}{2} K_{SQ} - k_Q \bar\delta^2 - K_S \bar\delta \geq 0$.

Nous venons de minorer $P(\bar v, x^i) - P(\bar v, t^i)$ par $\frac{1}{2} K_{SQ} \|x^i - \bar x\|^2$.

Nous allons maintenant majorer $c \min \{ \phi_N(v^i, x^i, \overset{*}{\Sigma}{}^i), \phi_N(v^i, x^i \overset{*}{\Sigma}{}^i)^3 \}$

Nous calculerons ensuite la différence :

$P(\bar v, x^i) - P(\bar v, t^i) - c \min \{ \phi_N(v^i, x^i, \overset{*}{\Sigma}{}^i), \phi_N(v^i, x^i, \overset{*}{\Sigma}{}^i)^3 \}$

Nous minorerons alors cette quantité par zéro et nous arriverons à la conclusion que (x^{i+1}, u^{i+1}) est déterminé par la phase 2 nous aurons alors que

1) $LS(x^i, u^i, G^i, I) = 0$ est consistant

2) $s^i \geq 0$

3) $\tilde{z}^i \geq 0$

4) $P(\bar{v}, x^i) - P(\bar{v}, t^i) \geq c \min \{ \varphi_N(\bar{v}, x^i, \overset{*}{z}^i), \varphi_N(\bar{v}, x^i, \overset{*}{z}^i)^3 \}$

La procedure de fonction ne nous renverra donc pas à la phase 1.

$$c\,(\varphi_N(\bar{v}, x^i, \overset{*}{z}^i)^3)$$

$$\leq c\,K_\varphi^3 \left(\|x^i - \bar{x}\| + \|\bar{w} - \overset{*}{z}^i\| \right)^3$$

$$= c\,K_\varphi^3 \left(\|x^i - \bar{x}\| + \bar{v}\,\|\bar{u} - s^i\| \right)^3$$

par la propriété 2

$$\leq c\,K_\varphi^3 \left(\|x^i - \bar{x}\| + \bar{v}\,\bar{\delta}\,\|x^i - \bar{x}\| \right)^3$$

par le lemme 2

$$\leq c\,K_\varphi^3\,(1 + \bar{v}\,\bar{\delta})^3\,\|x^i - \bar{x}\|^3$$

$$P(\bar{v}, x^i) - P(\bar{v}, t^i) - c \min \{ \varphi_N(\bar{v}, x^i, \overset{*}{z}^i), \varphi_N(\bar{v}, x^i, \overset{*}{z}^i)^3 \}$$

$$\geq P(\bar{v}, x^i) - P(\bar{v}, t^i) - c\,\varphi_N(\bar{v}, x^i, \overset{*}{z}^i)^3$$

$$\geq \frac{1}{2} K_{SQ}\,\|x^i - \bar{x}\|^2 - c\,K_\varphi^3\,(1 + \bar{v}\,\bar{\delta})^3\,\|x^i - \bar{x}\|^3$$

≥ 0 parce que $\|x^i - \bar{x}\| \leq \min \left\{ \delta_{11} , \dfrac{2\,c\,K_\varphi^3\,(1 + \bar{v}\,\bar{\delta})^3}{K_{SQ}} \right\}$

Nous voyons donc que pour $i = i_{\bar{\delta}}$; (x^{i+1}, u^{i+1}) est déterminé par la phase 2, nous avons donc

$(x^{i+1}, u^{i+1}) = (t^i, s^i)$

ou alors

$(x^{i+1}, u^{i+1}) = (y^i, \tilde{z}^i)$

Le pas 1 et les propriétés de $i = i_{\bar{\delta}}$ nous permettent d'affirmer que :

$$\| (x^{i+1}, u^{i+1}) - (\bar{x}, \bar{u}) \| \leq \bar{\delta} \, \| (x^i, u^i) - (\bar{x}, \bar{v}) \|$$

$i+1$ a donc les mêmes propriétés que $i_{\bar{\delta}}$

Donc en répétant les arguments précédents, nous arrivons à la conclusion que :

$\forall k \in \mathbb{N}$, $(x^{i_{\bar{\delta}}+k}, u^{i_{\bar{\delta}}+k})$ est déterminé par la phase 2

donc, $\forall i \geq i_{\bar{\delta}}$, $\dfrac{\| (x^{i+1}, u^{i+1}) - (\bar{x}, \bar{v}) \|}{\| (x^i, v^i) - (\bar{x}, \bar{v}) \|} < \bar{\delta}$

#

D : Formules d'ajustement

La méthode présentée par l'algorithme utilise une matrice G^i ; cette matrice devrait approcher le Hessien du Lagrangien au point (x^i, v_I^i) ; nous verifierions alors les hypothèses du théorème de convergence superlinéaire

Il est donc necessaire de determiner des formules d'ajustement de la matrice G^i de manière à verifier les conditions :

$$\lim_{i \to \infty} G^i = \nabla_{xx} L(\bar{x}, \bar{v}_I)$$

ce qui est equivalent à :

$$\lim_{i \to \infty} G^i = \lim_{i \to \infty} \nabla_{xx} L(x^i, u_I^i)$$

si la suite (x^i, v^i) converge vers $(\bar{x}, \bar{v})$.

Le principe est le suivant :
On part d'une matrice G^0 appartenant à $\mathbb{R}^{n \times n}$,
les formules d'ajustement determinent $G^1, G^2 \dots G^n$
On reprend le procédé d'ajustement avec G^n comme matrice de départ et on determine $G^{n+1} \dots G^{2n}$
De manière générale, a partir d'une matrice $G^{m.n}$ $(m, n \in \mathbb{N})$ on determine les matrices $G^{m.n+1}, \dots G^{(m+1)n}$ grâce aux formules d'ajustement.

Dans ce chapitre, nous allons utiliser la norme 1
Il est entendu qu'une convergence établie en norme 1 est équivalente à une convergence établie en norme 2

1 : construction de "n" matrices

Pour une question de facilité, nous considérerons que nous demarrons avec la matrice G^0 et que nous voulons construire $G^1, G^2 \dots G^n$. Le procédé pour construire $G^{n+1}, G^{n+2}, \dots G^{2n}$ a partir de G^n est evidemment le même.

<u>definitions</u> :

soient pour $k = 0, \dots n-1$

$\tilde{s}^k = x^{k+1} - x^k$, c'est à dire le vecteur différence entre deux itérés successifs

$e^0 = \tilde{s}^0 / \| \tilde{s}^0 \|$, c'est à dire le vecteur différence entre x^0 et x^1 normalisé.

<u>Commentaires</u> :

e^0 est donc un vecteur normé. En fait, on construira une suite $(e^0, \dots e^{n-1})$ de vecteurs orthonormés. On définit ensuite un vecteur y^0 qui est la difference entre deux gradients :

$$y^0 = \nabla_x L(x^0 + s^0, u^1) - \nabla_x L(x^0, u^1)$$

C'est donc l'accroissement du gradient du lagrangien dans la direction s_0.

La première formule d'ajustement est :

$$G^1 = G^0 + \frac{(y^0 - G^0 s^0)\, e^{0T}}{\| s^0 \|}$$

Commentaires

La formule d'ajustement fait intervenir :
G^0 : matrice calculée à l'étape précédente ou connue
y^0 : différence des gradiens du lagrangien dans la direction s^0

Notons le rapport entre y^0 et $G^0(x^1 - x^0)$
dans le cas où : $x^1 \to x^0$
$u^1 \to u^0$
G^0 approxime $\nabla_{xx} L(x^0, u^0)$.

définitions : pour $k = 1, \dots n-1$, on définit

$$\hat{e}^k = \hat{s}^k - \sum_{j=0}^{k-1} ((\hat{s}^k)^T e^j)\, e^j$$

Commentaires

$\hat{e}^k$ est donc défini à partir de $\hat{s}^k$. C'est la composante de $\hat{s}^k$ dans l'espace normal de l'espace engendré par les vecteurs $e^0 \dots e^{k-1}$; $\hat{e}^k$ est donc un vecteur normal à la famille $(e^0 \dots e^{k-1})$
Nous allons maintenant déterminer e^k ; e^k sera un vecteur orthogonal aux vecteurs $(e^0, \dots e^{k-1})$ déjà construits. Pour cela, on introduit une constante β, $\frac{1}{2} < \beta < 1$.

définitions : soit β appartenant à $]\frac{1}{2}, 1[$

si $\frac{\|\tilde{e}^k\|}{\|\tilde{s}^k\|} \geq \beta$ alors $e^k = \frac{\tilde{e}^k}{\|\tilde{e}^k\|}$

$s^k = \tilde{s}^k$

si $\frac{\|\tilde{e}^k\|}{\|\tilde{s}^k\|} < \beta$ alors e^k est un vecteur quelconque normal à $(e^0 \ldots e^{k-1})$

$s^k = \|\tilde{s}^k\| e^k$

Commentaires

Il se peut que $\|\tilde{e}^k\|$ soit très petite, ou même nulle, alors on prend pour e^k un vecteur quelconque orthogonal à $(e^0, \ldots e^{k-1})$.

Notons que quelles que soit le choix de e^k et s^k on a toujours :

$$\|s^k\| = \|\tilde{s}^k\| = \|x^{k+1} - x^k\|$$

tout comme on avait défini y_0, on défini y_k qui mesurera la différence les gradients dans la direction s^k

$$y^k = \nabla_x L(x^k + s^k, u^{k+1}) - \nabla_x L(x^k, u^{k+1})$$

Le procédé d'ajustement définit alors G^{k+1} :

$$G^{k+1} = G^k + \frac{(y^k - G^k s^k)(e^k)^T}{(s^k)^T e^k}$$

Commentaires

Remarquons que $(s^0)^T e^0 = (s^0)^T \dfrac{s_0}{\|s_0\|} = \|s_0\|$ et donc cette formule est une généralisation de la formule introduite pour déterminer G^1.
Si la base $(e^0, e^1 \ldots e^{n-1})$ était la base canonique de $\mathbb{R}^n$, la formule d'ajustement déterminant la matrice G^{k+1} consisterait en un changement de la $(k+1)$ième colonne de la matrice G^k, les autres colonnes restant inchangées.

Nous allons maintenant montrer que cette procedure d'ajustement verifie :

$$\lim_{i \to \infty} \| G^i - \nabla_{xx} L(\bar{x}, \bar{u}) \| = 0$$

$$\lim_{i \to \infty} \| G^i - \nabla_{xx} L(x^i, u^i) \| = 0 .$$

3 Théorème de convergence.

Théorème 1 :

Supposons que :

(1) la suite engendrée par l'algorithme 4 converge vers $(\bar{x}, \bar{u})$

ou alors :

(2) $\bar{x}$ est la solution unique du P.N.L (1)
$\bar{x}$ est un point d'accumulation de (x^i)
et (x^i) est bornée

Alors :

$$\lim_{i \to \infty} \| G^i - \nabla_{xx} L(\bar{x}, \bar{u}) \| = 0$$

et donc (x^i, u^i) converge superlinéairement vers $(\bar{x}, \bar{u})$

Commentaires

Remarquons que si $\lim_{i \to \infty} \| G^i - \nabla_{xx} L(\bar{x}, \bar{u}) \| = 0$
et si $\lim_{i \to \infty} (x^i, u^i) = (\bar{x}, \bar{u})$ alors on a que

$$\lim_{i \to \infty} \frac{\| (G^i - \nabla_{xx} L(x^i, \bar{u}))(y^i - x^i) \|}{\| y^i - x^i \|} = 0$$

$$\lim_{i \to \infty} \frac{\| G^i - \nabla_{xx} L(x^i, u))(t^i - y^i) \|}{\| t^i - y^i \|} = 0$$

et les hypothèses du théorème de convergence superlinéaire sont satisfaites

Il reste donc à démontrer que :

$$\lim_{i \to \infty} \| G^i - \nabla_{xx} L(\bar{x}, \bar{v}) \| = 0$$

Pour cela on va d'abord démontrer que l'hypothèse (2) entraine l'hypothèse (1) nous démontrerons alors le théorème sous l'hypothèse (1) Nous démontrerons d'abord trois lemmes, le deuxième se démontrant grace au premier. La thèse du théorème deviendra ensuite naturelle.

<u>propriété 1 :</u>

si $\bar{x}$ est la solution unique du P.N.L (1)
$\hat{x}$ est un point d'accumulation de (x^i)
(x^i) est bornée
alors :
la suite engendrée par l'algorithme y converge vers $(\bar{x}, \bar{u})$

<u>preuve :</u>

<u>1) demontrons que $\lim_{i \to \infty} x^i = \bar{x}$</u>

supposons que cela soit faux
Comme (x^i) est une suite bornée, il existe un autre point d'accumulation $\hat{x}$ de la suite (x^i).
Nous savons par le deuxième chapitre que le paramètre $\bar{v}$ est strictement positif et donc $\hat{x}$ est admissible.

Nous savons aussi par le chapitre 2 que $\lim_{i\to\infty} P(v^i; x^i)$ existe et que

$$\bar{v}\, f(\hat{x}) = P(\bar{v}, \hat{x}) = \lim_{i\to\infty} P(v^i; x^i) = P(\bar{v}, \bar{x}) = v\, f(\bar{x})$$

$$\text{et donc } f(\hat{x}) = f(\bar{x})$$

Nous contredisons l'hypothèse affirmant l'unicité de la solution du P.N.L (1)

2) démontrons que $\lim_{i\to\infty} u^i = \bar{u}$

supposons l'inverse

nous savons que : $\exists\, i_1$, $i > i_1 \Rightarrow v^i = \bar{v}$

et sans perdre de généralité, nous pouvons dire, par un raisonnement semblable à celui présenté dans le théorème de convergence du chapitre 2, que :

$$\forall i > i_1 \;,\quad \|w^i\|_\infty < \frac{3}{4}$$

La suite (w^i) est donc bornée.

Si (u^i) ne converge pas vers $(\bar{u})$, alors (w^i) ne converge pas non plus vers $\bar{w} = \bar{u}(\bar{v})^{-1}$ par la propriété 2 du chapitre 4.

Il existe donc un deuxième point d'accumulation de la suite (w^i), notons le $\tilde{w}$.

Nous savons que : $\forall i > i_1$, $\|w^i\| < \frac{3}{4}$ donc $w^i \in T(x^{i-1})$

Le chapitre 2 nous permet de dire que

$$\lim_{i\to\infty} \phi_N(\bar{v}, \bar{x}, \tilde{w}) = 0 .$$

$\bar{v} > 0$, $\varphi_v(\bar{v}, \bar{x}, \tilde{w}) = 0$, $\bar{x}$ admissible

ces trois affirmations nous permettent de dire que $(\bar{x}, \frac{\tilde{w}}{\bar{v}})$ est un point de Kuhn Tucker du P.N.L (1)

Nous avons donc que :

$$\bar{v} f(\bar{x}) + \nabla g_I(\bar{x}) \tilde{w}_I = 0.$$

Mais nous savons que

$$\bar{v} f(\bar{x}) + \nabla g_I(\bar{x}) \bar{w}_I = 0$$

donc : $\nabla g_I(\bar{x})(\tilde{w}_I - \bar{w}_I) = 0$

Nous concluons que $\tilde{w}_I = \bar{w}_I$ parce que les gradients des contraintes actives en $\bar{x}$ sont linéairement independants.

Nous avons donc que $\tilde{w} = \bar{w}$

car $\forall j \notin I$, $\tilde{w}_j = \bar{w}_j = 0$

#

<u>définition</u> : $\forall\ k, j$ tels que $0 \leq j, k \leq n-1$ soit

$$a_{kj} = \frac{(s^k)^T e^j}{\|s^k\|}$$

<u>Commentaires</u>

La matrice A ainsi défini $- A \equiv (a_{kj}) -$ est la matrice des coordonnées des vecteurs $\frac{s^k}{\|s^k\|}$ $(k = 0 \ldots n-1)$ par rapport à la base orthonormée $(e^0 \ldots\ldots\ldots e^{n-1})$.

Nous allons voir que les elements diagonaux de cette matrice A sont tous differents de zero, en fait ces elements seront plus grands que β cela nous sera utile pour construire une suite de vecteurs z^k $(k = 0 \ldots n-1)$.

Nous avons que $\forall k,\ a_{kk} \geq \beta$.

si $\frac{\|\hat{e}^k\|}{\|\hat{s}^k\|} \geq \beta$ alors $(\hat{e}^k, z^k) = 1$

$(s^k, e^k) = \|\hat{e}^k\|$

et $a_{kk} = \frac{\|\hat{e}^k\|}{\|\hat{s}^k\|} \geq \beta$

si $\frac{\|\hat{e}^k\|}{\|\hat{s}^k\|} < \beta$ alors $(\hat{e}^k, \hat{s}^k) = 1$

$$\frac{(s^k, e^k)}{\|s^k\|} = \frac{\|\hat{s}^k\|}{\|s^k\|} = 1 > \beta$$

<u>définitions</u> :

$$z^0 = y^0 / \| s^0 \|$$

$$z^k = \left(\frac{y^k}{\| s^k \|} - \sum_{j=0}^{k-1} a_{kj} z^j \right) (a_{kk})^{-1}$$

$k = 1 \dots n-1.$

<u>lemme 1</u> : $\forall\ 0 \le k \le n$ on a $G^k e^j = z^j \quad 0 < j < k$

<u>Commentaires</u>

La famille de vecteurs $(z^j,\ j=0 \dots n-1)$ est donc l'image de la base $(e^j,\ j=0, \dots n-1)$ par l'application linéaire représentée par la matrice G^n

La démonstration se fait par récurrence

le cas $k=0$ est trivial

On suppose donc la propriété vraie pour $k = 1 \dots l$

on doit donc la démontrer pour $k = l+1$

Cela s'effectue en deux étapes

1) $\forall j < l$, $G^{l+1} e^j = z^j$

2) $j = l$, $G^{l+1} e^j = z^j$

Cette deuxième partie est la plus difficile pour la démontrer on essaie de remplacer

$\dfrac{y^l}{(s^l)^T e^l}$, $\dfrac{G^l s^l}{(s^l)^T e^l}$ par des expressions

faisant intervenir z^j $(j = 0 \dots l)$ et $G^l e^j$ $(j = 0 \dots l)$ de manière à utiliser l'hypothèse de récurrence.

démonstration

1) le lemme est vrai pour $k=0$

2) supposons le lemme vrai pour $k = 0 \dots l$
et démontrons le pour $k = l+1$

soit $j < l$

$$G^{l+1} e^j = G^l e^j + \frac{(y^l - G^l s^l)(e^l)^T e^j}{(s^l)^T e^l}$$

$$= G^l e^j \qquad \text{car } (e^l)^T e^j = 0 \text{ si } j<l$$

soit $j = l$

$$G^{l+1} e^l = G^l e^l + \frac{(y^l - G^l s^l)(e^l)^T e^l}{(s^l)^T e^l}$$

$$\frac{y^l}{(s^l)^T e^l} = \frac{1}{a_{ll}} \sum_{j=0}^{l} a_{lj} z^j$$

$$\frac{G^l s^l}{(y^l)^T e^l} = \frac{\|s^l\|\; G^l s^l}{\|s^l\| \; (s^l)^T e^l}$$

$$= \frac{1}{a_{ll}} \; \frac{G^l s^l}{\|s^l\|}$$

$$s^l = \sum_{j=0}^{l} ((s^l)^T e^j)\, e^j$$

car s^l se décompose entièrement dans $(e^0 \ldots e^l)$

$$= \frac{1}{a_{ll}} \; \frac{G^l}{\|s^l\|} \left[\sum_{j=0}^{l} ((s^l)^T e^j)\, e^j \right]$$

$$= \frac{1}{a_{ll}} \; G^l \sum_{j=0}^{l} a_{lj}\, e^j$$

donc $G^{l+1} e^l = G^l e^l + \dfrac{1}{a_{ll}} \displaystyle\sum_{j=0}^{l} a_{lj} \, (z^j - G^l e^j)$

$$= G^l e^l + \frac{1}{a_{ll}} \, a_{ll} \, (z^l - G^l e^l)$$

parce que $z^j = G^l e^j \qquad j = 0 \ldots l-1$

$$= z^l .$$

#

interprétation qualitative du lemme

Une des conséquences du lemme est :

$$G^n e^0 = z^0 \qquad G^{n-1} e^0 = z^0$$
$$G^n e^1 = z^1 \qquad G^{n-1} e^1 = z^1$$
$$\vdots \qquad\qquad \vdots$$
$$G^n e^{n-1} = z^{n-1} \qquad G^{n-1} e^{n-2} = z^{n-2}$$

et donc : $(G^n - G^{n-1}) e^0 = 0$

$$(G^n - G^{n-1}) e^1 = 0$$
$$\vdots$$
$$(G^n - G^{n-1}) e^{n-2} = 0$$

$(e^0, \dots e^{n-2})$ sont linéairement indépendants donc la matrice $(G^{n-1} - G^n)$ est une matrice de rang 1

Le passage de la matrice G^{n-1} à G^n est donc "équivalente" à un changement de lignes ou de colonnes au plus.

Si la base $(e^i \}\ i = 0, \dots n-1)$ était la base canonique le passage de G^{n-1} à G^n consisterait à changer la dernière colonne de G^{n-1}.

De même le passage de G^{n-2} à G^{n-1} "correspond" – par le lemme – à deux changements de lignes ou de colonnes et le passage de G^{n-2} à G^n "correspond" lui aussi à deux changements de lignes ou de colonnes.

$$\underbrace{G^{n-2} \xrightarrow[\text{au plus}]{\text{2 changements}} G^{n-1} \xrightarrow[\text{au plus}]{\text{1 changement}} G^n}_{\text{2 changements au plus}}$$

Donc si on a un changement entre G^{n-1} et G^{n} on a au plus un changement entre G^{n-2} et G^{n-1}. Les formules d'ajustement correspondent donc à un changement de rang 1 à chaque étape.

<u>Lemme 2 :</u>

Il existe ε, $k_{44} > 0$ tels que

si : $(x^i, u^{i+1}), (x^i + s^i, u^{i+1}) \in B((\bar{x}, \bar{u}), \varepsilon)$

pour $i = 0 \ldots n-1$

alors : $\| Z^k - \nabla_{xx} L(\bar{x}, \bar{u}) e^k \| \leq \beta^{-1} (k+1) K_{44} M$

$k = 0 \ldots n-1$

$$M = \text{maximum} \left\{ \left\{ \| (x^i, u^{i+1}) - (\bar{x}, \bar{u}) \|, \ i = 0 \ldots n-1 \right\} \cup \left\{ \| (x^i + s^i, u^{i+1}) - (\bar{x}, \bar{u}) \|, \ i = 0 \ldots n-1 \right\} \right\}$$

<u>Commentaires</u>

L'hypothèse du lemme est :

"$(x^i, u^{i+1}), (x^i + s^i, u^{i+1}) \in B((\bar{x}, \bar{u}), \varepsilon)$"

notons que pour i assez grand, cette hypothèse est vérifiée parce que $\lim\limits_{i \to \infty} (x^i, u^i) = (\bar{x}, \bar{u})$

L'énoncé du lemme mentionne : " pour $i = 0 \ldots n-1$ "

Ceci est à comprendre dans le sens suivant :

$$\begin{array}{llll} i = & 0 \ (\mathrm{mod}\, n) & \ldots & n-1 \ (\mathrm{mod}\, n) \\ = & m \cdot n & & m \cdot n + (n-1) \end{array}$$

et le lemme prend sa valeur pour m grand.

La démonstration se fera par induction sur k. Le cas $k = 0$ se démontrera grâce à un lemme que nous mentionnerons sans le démontrer. Pour $k = l \leq n-1$, nous majorerons la quantité $\| z^l - \nabla_{xx} L(\bar{x}, \bar{u})\, e^l \|$, pour cela nous exprimerons

1) z^l en fonction de y^l et $z^{l-1} \dots z^0$.

2) e^l en fonction de s^l et $e^{l-1} \dots e^0$

de manière à pouvoir utiliser l'hypothèse de récurrence et le lemme auxiliaire

<u>démonstration</u>

1) $k = 0$.

Pour démontrer l'assertion dans ce cas ci nous allons utiliser un lemme auxiliaire, nous l'appelerons le lemme de Han :

lemme de Han : $\exists k_{44}$, $\varepsilon > 0$ tels que si

$$(x^k, u^{k+1}), (x^k + s^k, u^{k+1}) \in B((\bar{x}, \bar{u}), \varepsilon)$$

alors :

$$\| y^k - \nabla_{xx} L(\bar{x}, \bar{u})\, s^k \| \leq k_{44}\, M \| s^k \|$$

$$k = 0, \dots n-1.$$

La référence exacte de ce lemme est :

Han, S. P., " Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems, Ph. D. dissertation, University of Wisconsin, 1974. Lemme 2.3.7.

de ce fait

$$\|Z^0 - \nabla_{xx} L(\bar{x},\bar{v})\, e^0\| = \frac{\|y^0 - \nabla_{xx} L(\bar{x},\bar{v})\, s^0\|}{\|s^0\|}$$

$$\leq K_{44}\ M\ \frac{\|s^0\|}{\|s^0\|}$$

$$\leq \beta^{-1}\ M\ K_{44} \qquad (\beta \leq 1)$$

1) supposons le lemme vrai pour $k = 0 \ldots l-1$

démontrons qu'il est vrai pour l $\qquad l \leq n-1$

$$s^l = \sum_{j=0}^{l} \left[(s^l)^T e^j \right] e^j$$

$$= \|s^l\| \sum_{j=0}^{l} a_{lj}\, e^j \qquad \text{par définition de } a_{lj}$$

$$\frac{s^l}{\|s^l\|} = \sum_{j=0}^{l} a_{lj}\, e^j$$

$$= \sum_{j=0}^{l-1} a_{lj}\, e^j + a_{ll}\, e^l .$$

et donc $\displaystyle a_{ll}\, e^l = \frac{s^l}{\|s^l\|} - \sum_{j=0}^{l-1} a_{lj}\, e^j$

donc $a_{ll}(z^l - \nabla_{xx} L(\bar{x},\bar{u}) e^l)$

$$= \frac{y^l}{\|s^l\|} - \sum_{j=0}^{l-1} a_{lj} z^j - \nabla_{xx} L(\bar{x},\bar{u}) \left(\frac{s^l}{\|s^l\|} - \sum_{j=0}^{l-1} a_{lj} e^j \right)$$

par définition de z^l et ce qui précède

$$= \frac{y^l - \nabla_{xx} L(\bar{x},\bar{u}) s^l}{\|s^l\|} + \sum_{j=0}^{l-1} a_{lj} (z^j - \nabla_{xx} L(\bar{x},\bar{u}) e^j)$$

$|a_{ll}| \, \|z^l - \nabla_{xx} L(\bar{x},\bar{u}) e^l\|$

$$\leq K_{44} M + \sum_{j=0}^{l-1} |a_{lj}| \, (j+1) \beta^{-1} K_{44} M$$

par le lemme de Han et l'hypothèse de récurrence.

et donc $\|z^l - \nabla_{xx} L(\bar{x},\bar{u}) e^l\|$

$$\leq \beta^{-1} K_{44} M + \sum_{j=0}^{l-1} \frac{|a_{lj}|}{|a_{ll}|} (j+1) \beta^{-1} K_{44} M$$

car $|a_{ll}| \geq \beta$.

si nous parvenons à minorer

$\displaystyle\sum_{j=0}^{l-1} \frac{|a_{lj}|}{|a_{ll}|} (j+1)$ par l, nous aurons fini

parce qu'alors :

$$\|z^l - \nabla_{xx} L(\bar{x},\bar{u}) e^l\| \leq \beta^{-1} K_{44} M (l+1)$$

$$\sum_{j=0}^{l-1} \frac{|a_{lj}|}{|a_{ll}|} (j+1) < l \sum_{j=0}^{l-1} \frac{|a_{lj}|}{|a_{ll}|}$$

nous savons que $\dfrac{s^l}{\|s^l\|} = \sum_{j=0}^{l} a_{lj}\, e^j$

donc $\sum_{j=0}^{l} |a_{lj}| = 1$

parce que nous travaillons en norme 1.

nous avons donc : $\sum_{j=0}^{l-1} |a_{lj}| = 1 - |a_{ll}|$

$$\leq 1 - \beta .$$

$$\sum_{j=0}^{l-1} \frac{|a_{lj}|}{|a_{ll}|} \leq \frac{1-\beta}{\beta} < 1 \quad \text{car } \beta > \frac{1}{2}$$

nous avons donc $\sum_{j=0}^{l-1} \frac{|a_{lj}|}{|a_{ll}|} (j+1) < l$.

#

Lemme 3 : si (x^i, u^{i+1}), $(x^i + s^i, u^{i+1}) \in B((\bar{x}, \bar{u}), \varepsilon)$ pour $i = 0, \dots n-1$

alors :

1) $\| G^n - \nabla_{xx} L(\bar{x}, \bar{u}) \| \leq n(n+1) \beta^{-1} K_{44} M$

2) $\| G^k - \nabla_{xx} L(\bar{x}, \bar{u}) \| \leq k(k+1) \beta^{-1} K_{44} M + \| G^0 - \nabla_{11} L(\bar{x}, \bar{u}) \|$

$k = 1 \dots n-1$.

Démonstration

1) demontrons le 1)

soit $t = t_0 e^0 + \dots t_{n-1} e^{n-1}$ vecteur arbitraire de norme 1 de $\mathbb{R}^n$.

$$\| (G^n - \nabla_{xx} L(\bar{x}, \bar{u})) t \|$$

$$\leq \sum_{j=0}^{n-1} |t_j| \, \| \underbrace{G^n e^j}_{zj} - \nabla_{xx} L(\bar{x}, \bar{u}) e^j \|$$

$$\leq \sum_{j=0}^{n-1} |t_j| \, \beta^{-1} (j+1) K_{44} M$$

$$\leq n(n+1) \beta^{-1} K_{44} M$$

donc $\| (G^n - \nabla_{xx} L(\bar{x}, \bar{u})) \| \leq n(n+1) \beta^{-1} K_{44} M$

2) démontrons la deuxième assertion

soit $k = 1 \ldots n-1$.

$$\| (G^k - \nabla_{xx} L(\bar{x}, \bar{u})) t \|$$

$$= \| (G^k - \nabla_{xx} L(\bar{x}, \bar{u})) (\sum_{j=0}^{k-1} b_j e^j) \|$$

$$\leq \sum_{j=0}^{k-1} |b_j| \, \| G^k e^j - \nabla_{xx} L(\bar{x}, \bar{u}) e^j \|$$

$$+ \| (G^k - \nabla_{xx} L(\bar{x}, \bar{u})) (\sum_{j=k}^{n-1} b_j e^j) \|$$

$$\leq k \beta^{-1} (k+1) K_{44} M$$

$$+ \| (G^k - \nabla_{xx} L(\bar{x}, \bar{u})) (\sum_{j=k}^{n-1} b_j e^j) \|$$

$(e^0, \ldots e^{n-1})$ est une base orthonormée

par la définition des G^k, nous avons donc :

$$G^k e^j = G^{k-1} e^j = G^0 e^j \qquad j \geq k.$$

donc $\| (G^k - \nabla_{xx} L(\bar{x}, \bar{u})) (\sum_{j=k}^{n-1} b_j e^j) \|$

$$= \| (G^0 - \nabla_{xx} L(\bar{x}, \bar{u})) (\sum_{j=k}^{n-1} b_j e^j) \|$$

$$\leq \| (G^0 - \nabla_{11} L(\bar{x}, \bar{u}) \|$$

donc $\| (G^k - \nabla_{xx} L(\bar{x}, \bar{u}) \|$

$$\leq \beta^{-1} k (k+1) K_{44} M + \| G^0 - \nabla_{xx} L(\bar{x}, \bar{u}) \| .$$

Démonstration du théorème

Les commentaires du lemme 2 sont valables pour le lemme 3

Nous savons que $i = 0 \ldots n-1$ veut dire :

$$i = 0 \pmod n \ldots \qquad n-1 \pmod n$$

$$i = m.n \qquad\qquad m.n + n - 1$$

et donc lorsque m grandit, M tend vers zero.

et $(x^i ; u^i) \in B((\bar{x}, \bar{u}), \epsilon)$

$$(x^i + s^i , u^{i+1}) \in B((\bar{x}, \bar{u}), \epsilon)$$

forts de ces remarques, nous appliquons le lemme 3 :

$$: \lim_{m \to \infty} G^{m.n} = \nabla_{11} L(\bar{x}, \bar{u}) .$$

et donc "lim" $\| G^0 - \nabla_{11} L(\bar{x}, \bar{u}) \| = 0$

$$(0 = 0 \pmod n) = m.n \pmod n)$$

En appliquant de nouveau le lemme 3, nous disons que :

$$\lim_{m \to \infty} \underset{k = 1, \ldots n-1}{G^{m.n + k}} = \nabla_{xx} L(\bar{x}, \bar{u})$$

En conclusion $\lim_{i \to \infty} G^i = \nabla_{xx} L(\bar{x}, \bar{u})$.

CHAPITRE V : CONCLUSION

Nous avons étudié un Algorithme Hybride Penalisation. Quasi Newton superlinéairement et globalement convergent. Nous n'avons pas programmé l'algorithme, mais cela a été fait par Chung. Il a testé son algorithme sur le deuxième et le troisième problème de Colville. Il a choisi ces deux problèmes, parce qu'ils étaient les seuls parmi les problèmes présentés par Colville dont les contraintes étaient non linéaires et deux fois continument différentiables.

Pour déterminer la variable t^i par l'intermédiaire du programme linéaire, on a utilisé la méthode du simplex. Pour la matrice G^i, on a pris la matrice $\nabla_{xx} L(x^i, v^i)$ parce que pour les deux problèmes traités, le Hessien du Lagrangien était accessible.

Tous les calculs ont été effectués sur le "Système Univac 1110" à l'université du Wisconsin Madison.

Le deuxième problème de Colville comporte 15 variables indépendantes et 20 contraintes. Que l'on parte d'un point admissible ou d'un point non admissible l'algorithme effectue 15 itérations par la phase 1 et 3 itérations par la phase 2

Le troisième problème de Colville a 5 variables indépendantes et 15 contraintes. L'algorithme arrive à la solution après 3 itérations en partant d'un point admissible et après 4 itérations en partant d'un point non admissible

Bibliographie

Chung S. "Globally and Superlinearly Convergent Algorithms for Nonlinearly Programming" Ph.D Computer Science, University of Wisconsin Madison 1975

Colville, A. R., "A Comparative Study of Nonlinear Programming Codes," I.B.M New York Scientific Center, Tech. Rep. 320-2925, 1968.

Dennis, J. E., Jr and Jorge J. Moré, "Quasi Newton Methods, Motivation and Theory. Siam Vol 19 No 1 1977 pp 46-89

Fiacco, A. V. and G. P. McCormick, "Nonlinear Programming: Sequential Unconstrained Minimization Techniques", Wiley, New York 1968.

Garcia-Palomares, U. M. and Mangasarian, O. L. "Superlinear Convergent Quasi Newton Algorithms for Nonlinearly Constrained Optimization Problems" Computer Sciences Technical Report # 195, University of Wisconsin Madison, 1974.

Han, S.P., " Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems", Ph.D. Dissertation, University of Wisconsin, 1974.

Howe, S., " New Conditions for Exactness for a simple Penality Function", Siam J. Control, Vol. 11, No.2, 1973, p.p 378-382

Lessman, R. " Mixed Nonderivative Algorithms for unconstrained optimization " University of Missouri - Rolla, Ph D, 1975 Computer Science

Mangasarian, O.L., " Nonlinear Programming Theory and Computation", Computer Sciences Technical Report # 185, University of Wisconsin, July 1973 pp. 10, 11.

McCormick, G.P., " Penalty Function Versus Nonpenalty Function Methods For Constrained Nonlinear Programming Problems", Mathematical Programming 1 (1971) pp. 217. 238.

Ortega J.M and Rheinboldt W.C : " Iterative Solution of Nonlinear Equations in Several Variables Academic Press, New York and London 1970.

Phillips, D.A " A Preliminary Investigation of Functions Optimization by a combination of Methods " Computer Journal, Vol.17 pp 75-79

Pietrzykowski, T., " An Exact Potential Method for Constrained Maxima ", Siam J. Numer. Anal. Vol. 6, No. 2 (1969), pp. 299-304.

Rall, L.B., " Computational Solution of Nonlinear Operator Equations ", John Wiley, 1969.

Zangwill, W. Z., " Nonlinear Programming via Penalty Functions " Management Science, Vol. 13, No. 5 (1967), pp. 344-358.

table des matières