THESIS / THÈSE

MASTER EN SCIENCES MATHÉMATIQUES

Problèmes d'optimisation comportant une infinité de contraintes

Müller, Jean-Pierre

Award date:
1979

Awarding institution:
Universite de Namur

FACULTES UNIVERSITAIRES N.D. DE LA PAIX,

NAMUR

FACULTE DES SCIENCES

Année académique 1978 - 1979

PROBLEMES D'OPTIMISATION

COMPORTANT

UNE INFINITE DE CONTRAINTES

Mémoire présenté pour l'obtention du grade
de Licencié en Sciences
mathématiques

par

Promoteur:
NGUYEN VAN HIEN

Jean-Pierre
MÜLLER

Nous tenons ici à exprimer notre gratitude envers toutes les personnes qui nous ont aidées à la réalisation de ce mémoire.

Nous remercions Monsieur le Professeur NGUYEN VAN HIEN, qui a bien voulu diriger notre travail.

Nous sommes particulièrement reconnaissants envers Monsieur DUCHATEAU dont l'aide et la bonne humeur furent précieuses durant cette année.

CHAPITRE I

INTRODUCTION

Dans ce premier chapitre, nous nous proposons d'introduire une classe particulière de problèmes d'optimisation : les problèmes d'optimisation comportant une infinité de contraintes (1). A cette fin, nous avons porté notre choix sur un problème bien connu de la théorie d'approximation, qui nous servira d'illustration à ce type de problème d'optimisation comportant une infinité de contraintes. La théorie de l'approximation constitue d'ailleurs, comme nous aurons encore l'occasion de le constater, un vaste et riche domaine d'application de la théorie des P.I.C., puisque la plupart des problèmes rencontrés en théorie d'approximation au sens de Chebycheff peuvent se mettre sous la forme d'un P.I.C.

Ensuite, partant de cet exemple introductif, nous énoncerons le P.I.C. sous sa forme générale : celle qui sera employée dans la suite de notre travail. Avant de passer alors à des variantes de cette formulation choisie d'un P.I.C., nous fixerons les objectifs poursuivis dans la suite.

Nous terminerons ce présent chapitre introductif en rappelant quelques définitions et quelques théorèmes essentiels pour les chapitres ultérieurs.

---

(1) Afin de nous faciliter la tâche de rédaction, nous convenons de noter dans la suite "P.I.C." pour "Problème d'optimisation comportant une infinité de contraintes".

I.1. Un problème de la théorie d'approximation comme illustration d'un P.I.C.

Considérons, en effet, le problème d'approximation au sens de Chebycheff avec contraintes d'interpolation. Ce problème s'énonce de la manière suivante :

Soit T, un sous-ensemble compact d'un espace métrique complet.
Soient f et $\phi_1,\ldots,\phi_n$ des fonctions à valeurs réelles définies sur T, continues.
Soit $D = \{ t^1,\ldots,t^q \}$ un ensemble de q points d'interpolation choisis dans l'ensemble T.
Le problème de Chebycheff comportant un nombre fini de contraintes d'interpolation consiste alors à trouver un polynôme généralisé, $\sum_{i=1}^{n} x_i^\star \phi_i$, qui approxime la fonction f pour la norme L^∞ et tel que :

(a) $\sum_{i=1}^{n} x_i^\star \phi_i(t) = f(t)$, quelque soit t dans D,

c'est à dire, le polynôme généralisé approximant prend les mêmes valeurs que la fonction objectif aux points d'interpolation.

(b) $x_1^\star,\ldots,x_n^\star$ sont tels que :

$$\| f - \sum_{i=1}^{n} x_i^\star \phi_i \|_\infty = \inf_{\{x_i\}} \| f - \sum_{i=1}^{n} x_i \phi_i \|_\infty$$

ou, par définition de la norme L^∞

$$\sup_{t\varepsilon T} \left| f(t) - \sum_{i=1}^{n} x_i^\star \phi_i(t) \right| = \inf_{\{x_i\}} \sup_{t\varepsilon T} \left| f(t) - \sum_{i=1}^{n} x_i \phi_i(t) \right| ,$$

c'est à dire, l'approximant réalise l'infimum des "distances" des polynômes généralisés à la fonction objectif f.

Les x_i, pour i=1,...,n, sont appelés : coéfficients des polynômes généralisés. On cherche donc les coéfficients $x_1^\star,\ldots,x_n^\star$ tels que les conditions (a) et (b) soient vérifiées.

Nous pouvons ramener ce problème à un problème d'optimisation comportant une infinité de contraintes.

Posons $m = \sup\limits_{t\varepsilon T} \left| f(t) - \sum\limits_{i=1}^{n} x_i \phi_i(t) \right|$, ce qui nous est permis,

puisque les fonctions $f,\phi_1,\ldots,$et ϕ_n continues sur un sous-ensemble compact T, y atteignent leurs bornes. Nous pouvons alors récrire le problème d'approximation au sens de Chebycheff comportant des contraintes d'interpolation sous la forme:

$$\min_{x_1,\ldots,x_n,m} \quad m$$

sous les contraintes (1) :

$$\begin{cases} (A) \quad -f(t) + \sum\limits_{i=1}^{n} x_i \phi_i(t) - m \leqslant 0 \\ (B) \quad f(t) - \sum\limits_{i=1}^{n} x_i \phi_i(t) - m \leqslant 0 \end{cases} \text{quelque soit t dans T}$$

$$(C) \quad f(t^j) - \sum_{i=1}^{n} x_i \phi_i(t^j) = 0 \;;\; \text{pour } j=1,\ldots,q$$

---

(1) Nous nous permettons dans la suite d'abrêger l'expression "sous les contraintes" par "s.c.".

On remarque que le problème d'approximation au sens de Chebycheff ainsi décrit comporte un nombre infini de contraintes, puisque le paramètre t est uniquement astreint à se trouver dans le sous-ensemble compact T.

Une deuxième transformation de ce problème conduira aux notations, qui vont suggérer la forme du P.I.C. général tel qu'il sera utilisé dans la suite.

En effet, posons: $x = (x_1, \ldots, x_n, m)$,

$$F(x) = m \ ,$$

$$G_1(x,t) = -f(t) + \sum_{i=1}^{n} x_i \ \phi_i(t) - m \ ;$$

$$G_2(x,t) = f(t) - \sum_{i=1}^{n} x_i \ \phi_i(t) - m \quad \text{et}$$

$$H(x,s) = f(s) - \sum_{i=1}^{n} x_i \ \phi_i(s) \ .$$

Alors, on obtient le problème :

$$\begin{array}{lll} \underset{x}{\text{minimiser}}\ F(x) & & \\ \text{s.c. : } G_1(x,t) \leqslant 0 & & \\ \phantom{\text{s.c. : }} G_2(x,t) \leqslant 0 & & \text{pour tout } t \text{ dans } T \\ \phantom{\text{s.c. : }} H(x,s) = 0 & & \text{pour tout } s \text{ dans } D \end{array}$$

Remarquons que le problème d'approximation au sens de Chebycheff comportant des contraintes d'interpolation ainsi formulé est bien un problème d'optimisation et qu'en plus ce problème englobe une infinité de contraintes, le paramètre t n'étant astreint qu'à appartenir à un ensemble compact quelconque T.

Dans les pages qui suivent, nous décrivons d'autres exemples de problèmes d'approximation qui peuvent se récrire sous la forme de problèmes d'optimisation comportant une infinité de contraintes.

I.2. Le P.I.C. et les objectifs poursuivis de ce mémoire

Motivé par l'existence de problèmes d'approximation pouvant se mettre sous la forme de problèmes d'optimisation comportant une infinité de contraintes, nous nous proposons, en nous inspirant de K.R. GEHNER [1], d'étudier le P.I.C. décrit ci-dessous dans le cadre de la programmation non linéaire.

$$
\text{(PT)} \quad
\boxed{
\begin{array}{lll}
\underset{x \in \mathbb{R}^n}{\text{minimiser}}\ F(x) & & \\
\quad \text{s.c. :} & G_i(x,t) \leqslant 0 & ;\ \text{quelque soit } t \varepsilon T_i,\ \text{pour } i=1,\dots,l \\
 & H_j(x,s) \leqslant 0 & ;\ \text{quelque soit } s \varepsilon S_j,\ \text{pour } j=1,\dots,m \\
 & x \ \varepsilon \ X^\circ &
\end{array}
}
$$

où X° est un ensemble ouvert de $\mathbb{R}^n$

T_i est un sous-ensemble compact d'un espace mètrique complet, pour $i=1,\dots,l$,

S_j est un ensemble quelconque, pour $j=1,\dots,m$,

$F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \rightsquigarrow F(x)$ et $G_i: \mathbb{R}^n x\, T_i \longrightarrow \mathbb{R}$, $(x,t) \rightsquigarrow G_i(x,t)$ sont des fonctions admettant des dérivées partielles continues par rapport à la première variable x, pour tout t dans T_i; en outre les fonctions G_i sont continues, quelque soit x dans X°, pour $i=1,\dots,l$ et

$H_j: \mathbb{R}^n x\, S_j \longrightarrow \mathbb{R}^n$, $(x,s) \rightsquigarrow H_j(x,s)$ est une fonction linéaire en x quelque soit s dans S_j, et H_j est continue en s, quelque soit x dans X°.

Remarquons que le problème d'optimisation ainsi décrit comporte généralement un nombre quelconque - éventuellement infini - de contraintes. En feuilletant la littérature concernant les problèmes d'optimisation comportant une infinité de contraintes, nous constatons que les premiers résultats concernant les P.I.C. ont été établis par FRITZ JOHN [2]. Ce dernier présenta des conditions nécessaires d'optimalité pour des problèmes d'optimisation possédant une infinité de contraintes continûment différentiables du type inégalité, et une preuve que ces conditions nécessaires seraient suffisantes sous des hypothèses supplémentaires assez restrictives.

Ensuite ce furent KUHN et TUCKER qui développèrent des conditions nécessaires et suffisantes d'optimalité pour le problème de Fritz John, dans le cas où celui-ci ne présente plus qu'un nombre fini de contraintes d'inégalité. KUHN et TUCKER furent les premiers [3] à considérer des qualifications de contraintes pour des P.I.C. et à expliquer leurs liens avec les conditions d'optimalité pour les P.I.C.

Les principaux objectifs que nous nous sommes proposés dans ce mémoire, sont de :

- développer un théorème d'existence et de caractère pour une solution d'un P.I.C.,
- appliquer ce critère à un problème de l'approximation au sens de Chebycheff et
- présenter deux algorithmes pour résoudre les P.I.C. .

En ce qui concerne le théorème de caractérisation d'une solution d'un P.I.C., le principal résultat est une généralisation du théorème classique de l'enveloppe convexe de FRITZ JOHN [2] et des conditions de KUHN et TUCKER [3]. Suivant en cela K.R. GEHNER [1], nous allons ajouter un ensemble infini de contraintes d'égalité au problème (PT). De plus, nous présenterons des qualifications de contraintes adéquates pour cette extension.

Quant aux algorithmes que nous développerons, nous verrons que le premier consiste en une généralisation du premier algorithme de REMES et le second, en une du deuxième algorithme de REMES. Les deux algorithmes engendrent en effet, une suite, dont chaque élément est une solution d'un sous-problème de minimisation comportant un nombre fini de contraintes du P.I.C. considéré, tel que la "suite" des sous-problèmes de minimisation tendent en un certain sens vers le P.I.C. traité. La différence entre les deux algorithmes est que pour le premier, tout sous-problème de minimisation fini, engendre une contrainte de plus que le sous-problème le précédant immédiatement; tandis que pour le second, nous n'avons pas cette propriété de croissance du nombre de contraintes pour les sous-problèmes successifs.

I. 3. Quelques variantes du P.I.C. sous la forme (PT)

I.3.1. Le problème de Fritz John

Par la suppression des contraintes du type égalité du P.I.C. (PT), en gardant les mêmes hypothèses de travail, nous obtenons le problème de Fritz John :

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimiser}}\ F(x) & \\ \quad \text{s.c. :} & G_i(x,t) \leqslant 0 \text{ ; quelque soit } t \text{ dans } T_i, \text{ pour } i=1,\ldots,l \\ & x \in X^\circ \end{array}$$

Dès lors, notre problème (PT) ne se distingue du problème de Fritz John que par l'apport d'un ensemble infini de contraintes du type égalité.

Remarquons que, bien que nous puissions ramener le problème (PT) sous la forme du problème de Fritz John - forme plus générale que celle du problème (PT) - en remplaçant les contraintes $H_j(x,s) = 0$ par $H_j(x,s) \leqslant 0$ et $- H_j(x,s) \leqslant 0$ pour $j = 1,\ldots,m$; il est difficile de montrer que les conditions nécessaires d'optimalité de Fritz John pour le problème résultant de ce remplacement des contraintes d'égalité du problème (PT), sont aussi suffisantes sous des qualifications de contraintes raisonnables.

En effet, il est nécessaire de développer un nouveau théorème de caractérisation pour le problème de Fritz John avec des contraintes linéaires du type égalité, plutôt que de scinder les égalités en deux inégalités et d'appliquer ensuite les conditions nécessaires d'optimalité de FRITZ JOHN [2] au problème qui en résulterait.

Supposons, pour simplifier, que nous devions résoudre le P.I.C. suivant :

$$\underset{x \in \mathbb{R}^n}{\text{minimiser}}\ F(x)$$

$$\text{s.c. :}\quad H(x,s) = 0,\ \text{quelque soit } s \text{ dans } S$$

$$x \in X^\circ$$

où X° est un ensemble ouvert de $\mathbb{R}^n$,

$F : \begin{array}{l}\mathbb{R}^n \longrightarrow \mathbb{R} \\ x \rightsquigarrow F(x)\end{array}$ est une fonction admettant des dérivées partielles continues par rapport à x,

$H : \begin{array}{l}\mathbb{R}^n \times S \longrightarrow \mathbb{R} \\ (x,s) \rightsquigarrow H(x,s)\end{array}$ est une fonction linéaire en x, quelque soit s dans S et continue en s quelque soit x dans X°.

Récrivons ce problème sous la forme d'un problème de Fritz John, en scindant les contraintes d'égalité et en employant la relation ci-dessous, valide puisque les contraintes sont linéaires en x, quelque soit $s \in S$.

$$H(x,s) = \nabla_x H(x,s).x + b(s)$$

Dès lors, on a :

$$\underset{x \in \mathbb{R}^n}{\text{minimiser}}\ F(x)$$

$$\text{s.c. :}\quad \nabla_x H(x,s).x + b(s) \leqslant 0$$

$$-\nabla_x H(x,s).x - b(s) \leqslant 0$$

$$x \in X^\circ$$

Appliquons le théorème des conditions nécessaires de FRITZ JOHN [2] à ce problème avec une infinité de contraintes d'inégalité.

On sait qu'il existe des nombres $\lambda_0 \geqslant 0$, $\lambda_1 > 0$ et $\lambda_2 > 0$ tel que

$$\lambda_0\ F(\overline{x}) + \lambda_1\ \nabla_x H(\overline{x},s^1) - \lambda_2\ \nabla_x H(\overline{x},s^2) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

si $H(\overline{x},s^1) = H(\overline{x},s^2) = 0$ et si $\overline{x}$ est une solution du problème, où en outre s^1 et s^2 sont dans l'ensemble quelconque S. Si maintenant nous sommes dans le cas où $\lambda_0 = 0$ et $\lambda_1 = \lambda_2 = 1/2$, alors cette dernière relation s'écrit :

$$1/2\ \nabla_x H(x,s^1) - 1/2\ \nabla_x H(x,s^2) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Il apparaît dès lors, qu'il n'existe aucun autre chemin adéquat pour éviter ce genre de résultat ne fournissant aucun renseignement utile, que de reformer des critères correspondant à des P.I.C. comportant une infinité de contraintes du type égalité [4].

I.3.2. Exemple d'un problème de Fritz John

Le problème d'approximation de Chebycheff sans contraintes décrit ci-dessous constitue un exemple d'illustration d'un problème de Fritz John.

Soient $T = [-1,+1]$ un sous-ensemble compact de la droite réelle

$$f : \begin{array}{ccc} T & \longrightarrow & \mathbb{R} \\ t & & f(t) \end{array}$$ une fonction continue et $\{ \phi_1, \ldots, \phi_n \}$ un ensemble fini de fonctions continues à valeurs réelles, définies sur T.

Le but de l'approximation au sens de Chebycheff est d'approcher, pour la norme L^{∞}, la fonction objectif f par des polynômes généralisés de la forme

$\sum_{i=1}^{n} x_i \phi_i$.

Le problème d'approximation au sens de Chebycheff est dès lors celui de la recherche des coëfficients $x_1^{\star},\ldots,x_n^{\star}$ d'un polynôme généralisé tel que :

$\max_{t\varepsilon T} | f(t) - \sum_{i=1}^{n} x_i^{\star} \phi_i(t)| = \min_{\{x_i\}} \max_{t\varepsilon T} | f(t) - \sum_{i=1}^{n} x_i \phi_i(t) |$ et en posant

$m = \max_{t\varepsilon T} | f(t) - \sum_{i=1}^{n} x_i \phi_i(t) |$, le problème peut se récrire sous la forme d'un P.I.C. de Fritz John :

$$\min_{x_1,\ldots,x_n,m} \quad m$$

$$\text{s.c.} \quad : \quad \left.\begin{array}{l} f(t) - \sum_{i=1}^{n} x_i \phi_i(t) - m \leqslant 0 \\ -f(t) - \sum_{i=1}^{n} x_i \phi_i(t) - m \leqslant 0 \end{array}\right\} \quad \text{quelque soit } t\varepsilon T$$

Remarquons que, pour ce problème de Chebycheff, les contraintes et la fonction objectif sont linéaires en m et x, les variables du problème, et donc à fortiori convexes. Une notation analogue à celle employée par le problème d'approximation au sens de Chebycheff comportant des contraintes d'interpolation (1) montre que le problème d'approximation au sens de Chebycheff sans contraintes est un problème de Fritz John.

---

(1) Ibid page 4

I.3.3. Une autre variante du P.I.C.

Une deuxième variante du problème initial est obtenu à partir du problème de Fritz John. Posons, à cette fin dans le problème de Fritz John : $G = \underset{i=1,\ldots,l}{\text{maximum}} \; G_i$. Nous obtenons alors le P.I.C. :

$$\underset{x \in \mathbb{R}^n}{\text{minimiser}} \; F(x)$$

$$\text{s.c. :} \quad G(x,t) \leqslant 0, \text{quelque soit } t \in T$$
$$x \in X^\circ$$

L'exemple suivant est un P.I.C. de ce type :

$$\underset{x \in \mathbb{R}}{\text{minimiser}} \; -x$$

$$\text{s.c. : } x.t \leqslant t^2 \text{ , quelque soit } t \in [0,1]$$

Pour opérer le passage au problème général il suffit de poser :

$$F(x) = -x \,,$$
$$G(x,t) = x.t - t^2 \,,$$
$$T = [0,1] \text{ et}$$
$$X^\circ = \mathbb{R} \,.$$

I.4. Quelques rappels

Avant de passer au second chapitre, rappelons quelques définitions et propriétés [5] , [6] , dont nous nous servirons dans la suite.

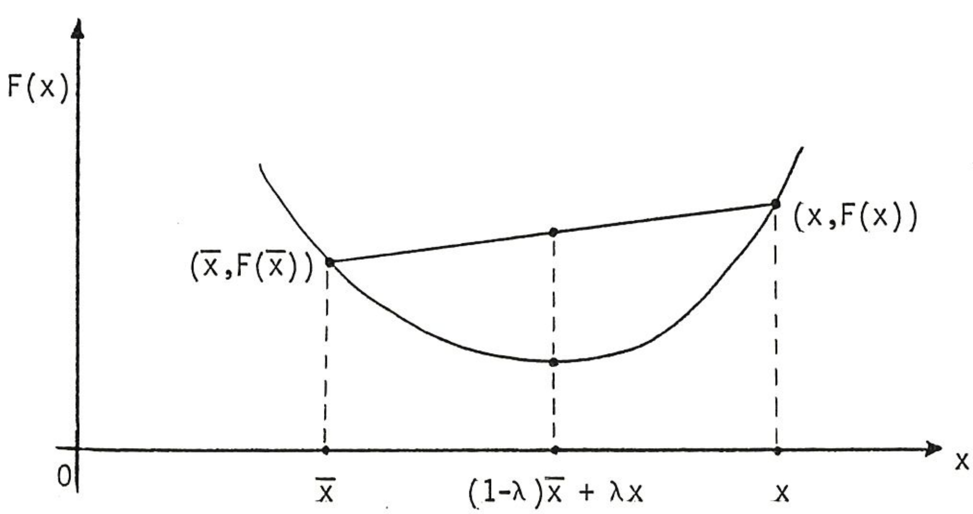
1. Fonction convexe

- Une fonction F à valeurs réelles et définie sur un ensemble $\Gamma \subset \mathbb{R}^n$ est dite *convexe* en $\bar{x} \in \Gamma$ si :

$$\left.\begin{array}{l} \forall x \in \Gamma \\ \forall \lambda \in \mathbb{R},\ 0 \leqslant \lambda \leqslant 1 \\ \text{et } (1-\lambda)\bar{x} + \lambda x \in \Gamma \end{array}\right\} \text{ on a } \quad F[(1-\lambda)\bar{x} + \lambda x] \leqslant (1-\lambda)\, F(\bar{x}) + \lambda\, F(x)$$

- F est convexe sur Γ, si F est convexe en x, quelque soit $x \in \Gamma$.

- Une interprétation géométrique est donnée, par le fait que si on prend un point arbitraire appartenant à Γ, soit x, alors le graphe de la fonction F est toujours en dessous de la corde passant par les deux points $(\bar{x}, F(\bar{x}))$ et $(x,F(x))$.



2. Fonction quasi-convexe

- Une fonction F, définie sur un ensemble $\Gamma \subset \mathbb{R}^n$ est dite *quasi-convexe* en $\bar{x} \in \Gamma$ (par rapport à Γ) si et seulement si

$$\left.\begin{array}{l} \forall x \in \Gamma \\ F(x) \leqslant F(\bar{x}) \\ \forall \lambda \in \mathbb{R},\ 0 \leqslant \lambda \leqslant 1 \\ (1-\lambda)\bar{x} + \lambda x \in \Gamma \end{array}\right\} \text{ on a } \quad F[(1-\lambda)\bar{x} + \lambda x] \leqslant F(\bar{x})$$

en d'autres termes, quelque soit x appartenant à Γ tel que $F(x) \leqslant F(\bar{x})$, la fonction F n'atteint pas de valeurs supérieures à $F(\bar{x})$ sur tout point de l'intersection du segment $[\bar{x},x]$ et de Γ.

- F est quasi-convexe sur Γ, si F est quasi-convexe en x,quelque soit x appartenant à Γ.

Remarquons qu'il n'est pas nécessaire d'imposer que l'ensemble Γ soit convexe (comme ce n'était pas le cas pour la définition d'une fonction convexe).

- Si l'ensemble Γ est convexe, alors la définition de la quasi-convexité s'énonce :

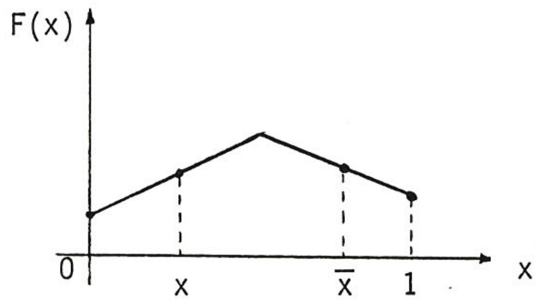
Une fonction F, à valeurs réelles et définie sur un ensemble convexe $\Gamma \subset \mathbb{R}^n$ est quasi-convexe sur Γ si et seulement si

$$\left.\begin{array}{l} \forall x^1, x^2 \in \Gamma \\ F(x^2) \leqslant F(x^1) \\ \forall \lambda,\ 0 \leqslant \lambda \leqslant 1 \end{array}\right\} \text{ on a } \quad F[(1-\lambda)x^1 + \lambda x^2] \leqslant F(x^1)$$

En outre, remarquons que, si nous remplaçons dans la définition de la quasi-convexité les inégalités par des inégalités strictes, nous obtenons la notion de quasi-convexité stricte pour une fonction F.

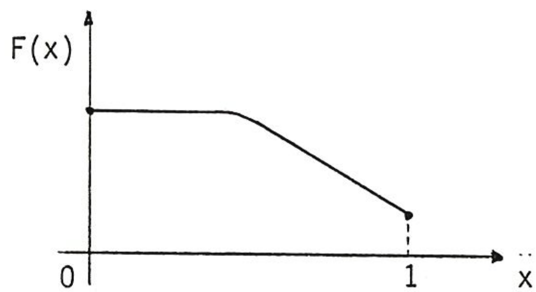
Exemples

a)



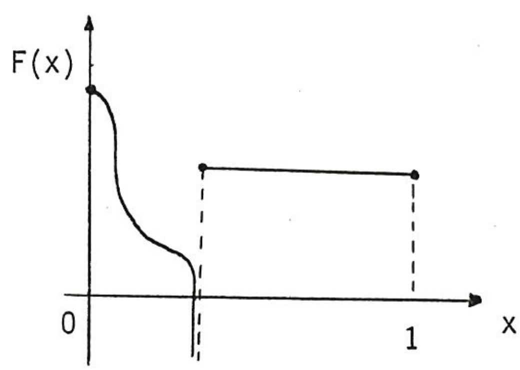
Fonction non quasi-convexe sur [0,1]

b)



Fonction quasi-convexe, qui n'est pas convexe, mais concave sur [0,1]

c)



Fonction quasi-convexe, non convexe et non concave sur [0,1]

Dans le cas différentiable, une autre interprétation géométrique de la quasi-convexité d'une fonction F au point $\overline{x} \varepsilon \Gamma \subset \mathbb{R}^n$ est possible. Dans ce cas, F est quasi-convexe en $\overline{x}$ si, quelque soit $x \varepsilon \Gamma$ on a : $F(x) \leqslant F(\overline{x})$ entraîne que $\nabla_x F(\overline{x}).(x-\overline{x}) \leqslant 0$, ce qui nous permet de formuler l'interprétation suivante :

Si au point considéré $\overline{x}$, la dérivée directionnelle de la fonction F est positive, dans la direction pointant vers un autre point de l'ensemble Γ, alors la valeur de la fonction F dépasse la valeur initiale de la fonction F lorsqu'on se déplace sur cette direction en partant de $\overline{x}$ (lorsque la fonction est définie sur les points intermédiaires).

3. Fonction pseudo-convexe

- On dit que la fonction F à valeur réelle et définie sur un ouvert de $\mathbb{R}^n$ contenant l'ensemble Γ est *pseudo-convexe* en $\overline{x} \in \Gamma$ (par rapport à Γ) si F est différentiable en $\overline{x}$ et si

$$\left.\begin{array}{l} \forall x \in \Gamma \\ \nabla_x F(\overline{x}).(x-\overline{x}) \geqslant 0 \end{array}\right\} \text{ on a } F(x) \geqslant F(\overline{x})$$

- F est pseudo-convexe sur Γ, si F est pseudo-convexe en x, quelque soit $x \in \Gamma$.

- Rappelons deux propriétés utiles [5].
  Soit F une fonction à valeurs réelles et définie sur un ensemble ouvert convexe Γ de $\mathbb{R}^n$. Soit $\overline{x} \in \Gamma$ et supposons que F soit différentiable en $\overline{x}$. Alors

(i) $\quad F(\overline{x}) = \underset{x\in\Gamma}{\text{minimum}}\, F(x) \qquad \Rightarrow \qquad \nabla_x F(\overline{x}) = 0$

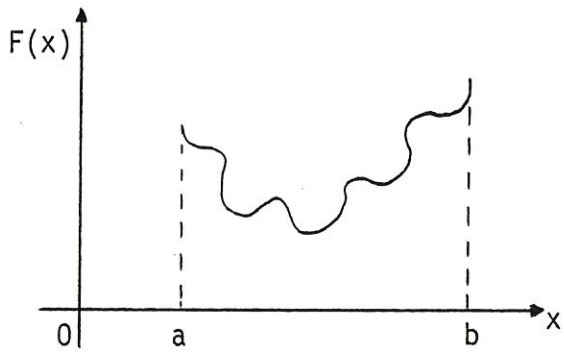
(ii) si de plus, F est pseudo-convexe en $\overline{x}$, alors

$$\nabla_x F(\overline{x}) = 0 \qquad \Rightarrow \qquad F(\overline{x}) = \underset{x\in\Gamma}{\text{minimum}}\, F(x).$$

- Une autre définition de la pseudo-convexité est possible et montre le lien avec la quasi-convexité. Une fonction à valeurs réelles F, différentiable en $\overline{x} \in \Gamma$ et définie sur Γ est dite pseudo-convexe en $\overline{x} \in \Gamma$, si F est quasi-convexe en $\overline{x}$ et quelque soit $x\in\Gamma$:

$$F(x) < F(\overline{x}) \qquad \Rightarrow \qquad \nabla_x F(\overline{x}).(x-\overline{x}) < 0$$

Exemple



Fonction pseudo-convexe, non convexe et non concave sur [a,b]

Une interprétation géométrique de la pseudo-convexité analogue à celle donnée pour la quasi-convexité et basée sur la deuxième définition (équivalente à la première, [5]) peut être formulée à présent : la dérivée directionnelle non négative en un point de la fonction pseudo-convexe pointe vers un point où la fonction admet une valeur supérieure ou égale à celle où l'on a considéré la dérivée directionnelle.

4. Fonction différentiable, fortement convexe

Une fonction F à valeurs réelles et définie sur $\mathbb{R}^n$ est dite *différentiable, fortement convexe*, si quelque soit x^1 et x^2 dans $\mathbb{R}^n$ et pour γ dans $\mathbb{R}$, strictement positif, on a

$$F(x^1) - F(x^2) \geqslant \nabla_x F(x^2) \cdot (x^1-x^2) + \gamma . \| x^1 - x^2 \|^2$$

5. Propriétés

Nous résumons les relations existantes entre des fonctions différentiables qui sont quasi-convexes, strictement quasi-convexes, pseudo-convexes et convexes, toutes étant définies sur un ensemble ouvert convexe Γ de $\mathbb{R}^n$ et à valeurs réelles dans le schéma suivant [5] :

F est convexe

$\Downarrow$

F est pseudo-convexe

$\Downarrow$

F est strictement quasi-convexe

$\Downarrow$

F est quasi-convexe

Ceci nous montre immédiatement que la classe des fonctions quasi-convexes est la plus large, tandis que celle des fonctions convexes est la plus petite parmi les classes de fonctions considérées.

6. Fonction Lipschitz-continue

On dit que la fonction $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ est *Lipschitz-continue* (de rapport k), si il existe un réel k, strictement positif tel que quelque soit x_1 et x_2 dans $\mathbb{R}^n$ on ait :

$$F(x_1) - F(x_2) \leqslant k.\| x_1 - x_2 \| .$$

I.5 Théorèmes préliminaires

Voyons maintenant deux théorèmes qui nous seront utiles dans le chapitre suivant.

I.5.1. Le théorème de Carathéodory

Soit $\Gamma \subset \mathbb{R}^n$.

Si x est une combinaison convexe de points de Γ,

alors x est une combinaison convexe de au plus n + 1 points de Γ; c'est à dire :

$$\text{si} \quad x = \sum_{i=1}^{m} p_i \, x^i \ , \quad x^i \in \Gamma, \quad p_i \in \mathbb{R} \ , \ p_i \geq 0 \ \text{et}$$

$$\sum_{i=1}^{m} p_i = 1 \ \text{avec} \ m > n - 1 \ ;$$

alors x peut s'écrire comme :

$$\sum_{i=1}^{m'} p'_i \, x^i \ , \ x^i \in \Gamma \ , \ p'_i \in \mathbb{R} \ , \ p'_i \geq 0 \quad \text{et}$$

$$\sum_{i=1}^{m'} p'_i = 1 \quad \text{avec} \quad m' \leq n + 1 \ .$$

Nous ne démontrons pas ce théorème qui est un résultat bien connu dans la littérature [5] . Ce théorème nous servira pour simplifier certaines expressions mathématiques obtenues dans les divers théorèmes qui suivront.

I.5.2. Le théorème généralisé de Motzkin (1), [4]

Soient U et V deux ensembles compacts dans $\mathbb{R}^n$
Soit W un ensemble quelconque dans $\mathbb{R}^n$

Alors

(A) $\begin{cases} \text{soit le système} & uz < 0 \quad \forall\, u \in U \\ & vz \leqslant 0 \quad \forall\, v \in V \\ & wz = 0 \quad \forall\, w \in W \\ \text{admet une solution } z \in \mathbb{R}^n, & \end{cases}$

(B) soit pour $u^\circ \in U$, il existe $s \leqslant n$ avec

(i) s vecteurs $u^i \in U, \quad 1 \leqslant i \leqslant s_1$
$v^i \in V, \; s_1 + 1 \leqslant i \leqslant s_2$
$w^i \in W, \; s_2 + 1 \leqslant i \leqslant s$

(ii) s+1 nombres réels λ_i, $0 \leqslant i \leqslant s$ tels que $\lambda_i \geqslant 0$ pour $i=0,\ldots,s_2$ avec soit $\lambda_0 > 0$ soit $s_1 \geqslant 1$ tels que

$$\lambda_0 u^\circ + \sum_{i=1}^{s_1} \lambda_i u^i + \sum_{i=s_1+1}^{s_2} \lambda_i v^i + \sum_{i=s_2+1}^{s} \lambda_i w^i = 0$$

mais non les deux à la fois.

Démonstration:

1) Montrons qu'il est impossible d'avoir en même temps les assertions (A) et (B).
Supposons que (A) soit satisfaite.

---

(1) Ce théorème est aussi appelé: Théorème de Gordon (cfr. Cours de Claude Dierieck: Théorie de l'approximation).

Alors il existe $z^* \in \mathbb{R}^n$ vérifiant le système (A).
Si (B) était aussi vérifié, alors

$$\sum_{i=0}^{s_1} \lambda_i\, u^i z^* + \sum_{i=s_1+1}^{s_2} \lambda_i\, v^i z^* + \sum_{i=s_2+1}^{s} \lambda_i\, w^i z^* = 0$$

or, $u^i z^* < 0$, car $u^i \in U \quad i=0,\dots,s_1$
$v^i z^* \leqslant 0$, car $v^i \in V \quad i=s_1+1,\dots,s_2$
$w^i z^* = 0$, car $w^i \in W \quad i=s_2+1,\dots,s$.

Donc,

$$0 = \underbrace{\sum_{i=0}^{s_1} \lambda_i\, u^i z^*}_{<0} + \underbrace{\sum_{i=s_1+1}^{s_2} \lambda_i\, v^i z^*}_{\leqslant 0} + \underbrace{\sum_{i=s_2+1}^{s} \lambda_i\, w^i z^*}_{=0} < 0$$

ce qui est absurde.

2) Montrons que si l'on suppose que (A) soit non vérifié, alors l'affirmation (B) est forcément correcte.

- Définissons les sous-ensembles de $\mathbb{R}^n$

$$Z_1 = \left\{ z \text{ tel que } \begin{array}{l} uz < 0,\ \forall\, u \in U \\ vz \leqslant 0,\ \forall\, v \in V \end{array} \right\}$$

$$Z_2 = \left\{ z \text{ tel que } wz = 0,\ \forall\, w \in W \right\}$$

Remarquons que : + $Z_2 = W^\perp$ est un sous-espace :
$\{ z \in \mathbb{R}^n \text{ tel que } wz = 0,\ \forall\, w \in W \}$
+ $Z_1 \cap Z_2$ est l'ensemble des solutions de (A) et donc $Z_1 \cap Z_2 = \phi$
+ Z_1 et Z_2 sont convexes.
Montrons le, par exemple, pour Z_2.

<u>H.</u>: $\forall\, \lambda \in [0,1]$
$\forall\, z_1, z_2 \in Z_2$

<u>Th.</u>: $\lambda z_1 + (1-\lambda) z_2 \in Z_2$

En effet: $z_i \varepsilon Z_2 \Rightarrow w z_i = 0, \forall\ w \varepsilon W,\ i=1,2$

et donc $\lambda\ w\ z_1 + (1-\lambda)\ w\ z_2 = 0$.

De même on démontre facilement que

Z_1 est convexe.

- <u>1^{er} cas</u>: $Z_1 \neq \phi$

Puisque $Z_2 \neq \phi$, car l'origine de $\mathbb{R}^n$ y est toujours, nous pouvons appliquer le théorème de la séparation des convexes dans $\mathbb{R}^n$ [5] aux deux ensembles convexes Z_1 et Z_2:

il existe un hyperplan $\{ z \varepsilon \mathbb{R}^n$ tel que $qz = \alpha \}$, $q \neq 0$ qui sépare Z_1 et Z_2; c'est à dire

$$(5.2.1.) \quad \begin{cases} qz \geqslant \alpha & \forall\ z \varepsilon Z_1 \\ qz \leqslant \alpha & \forall\ z \varepsilon Z_2 \end{cases} \qquad \begin{matrix} \text{où } q \varepsilon \mathbb{R}^n, q \neq 0 \\ \alpha \varepsilon \mathbb{R} \end{matrix}$$

D'autre part, puisque $Z_2 = W^{\perp}$ un sous- espace, nous avons que $qz=0$, $\forall\ z \varepsilon Z_2$. Si ce n'était pas le cas, il existerait $\overline{z} \varepsilon Z_2$ tel que $q\overline{z} \neq 0$. Puisque, maintenant $\lambda \overline{z}\ \varepsilon\ Z_2$, $\forall\ \lambda \varepsilon \mathbb{R}$, la relation (5.2.1.) entraîne que: $\lambda(q\overline{z}) \leqslant \alpha \quad \forall\ z \varepsilon Z_2$ pour α,q et $\overline{z}$ fixé, ce qui est absurde.

Donc $q \varepsilon Z_2^{\perp} = W^{\perp\perp}$. Or, l'algèbre nous enseigne que $W^{\perp\perp}$ est isomorphe à l'espace engendré par W. Ceci permet de réexprimer q:

$$q = \sum_{i=1}^{l} \mu_i\ w^i \qquad \text{où } w^i\ \varepsilon\ W,\ \ 1 \leqslant\ i \leqslant l \text{ et } l \leqslant n\ .$$

La relation (5.2.1.), nous montre alors que le système

$$(5.2.2.)\begin{cases} uz < 0 & \forall\ u\varepsilon U \\ vz \leqslant 0 & \forall\ v\varepsilon V \\ \left(\sum_{i=1}^{l} \mu_i\ w^i\right) z & < 0 \end{cases}$$

n'admet pas de solution $z\varepsilon\,\mathbb{R}^n$.

- $\underline{2^e \text{ cas}}$: $Z_1 = \phi$

Alors le système $(5.2.3.)\begin{cases} uz < 0 & \forall\ u\varepsilon U \\ vz \leqslant 0 & \forall\ v\varepsilon V \end{cases}$

n'admet pas de solution.

- Vu que (5.2.2) et(5.2.3.) sont de la même forme regroupons ces résultats.

Définissons $R = \left\{ \begin{array}{l} U \quad , \text{ si } Z_1 = \phi \\ U \cup \sum_{i=1}^{l} \mu_i\ w^i \quad = \quad U \cup \{q\} , \text{ si } Z_1 \neq \phi \end{array} \right\}$.

Alors le système $(5.2.4.)\begin{cases} qz < 0 & q\varepsilon\,R \\ vz \leqslant 0 & v\varepsilon V \end{cases}$

n'admet pas de solution $z\varepsilon\,\mathbb{R}^n$.

- Puisque R et V sont compact, pour $u°\varepsilon U$ choisi,

$$(5.2.5.) \quad Z = \left\{ \begin{array}{l} z \text{ tel que } \sum_{i=1}^{I} \beta_i\ q^i + \sum_{j=1}^{J} \delta_j\ v^j,\ q^i{=}u° \text{ pour un } i \\ \qquad\qquad \sum_{i=1}^{I} \beta_i = 1,\ \beta_i \geqslant 0,\ \delta_j \geqslant 0,\ I,J \text{ arbitraires} \\ \qquad\qquad q^i\varepsilon R,\ v^j\varepsilon V \end{array} \right\}$$

est fermé et convexe.

En effet, pour voir que Z est fermé récrivons (5.2.5.) :

$$Z = \left\{ \begin{array}{l} z \text{ tel que} \sum_{i=1}^{I} \beta_i \, q^i + \sum_{j=1}^{J} \dfrac{\delta_j \, v^j}{\sum_{j=1}^{J} \delta_j} \quad , q^i = u^\circ \text{ pour un } i \\ \\ \qquad \sum_{i=1}^{I} \beta_i = 1, \ \beta_i \geqslant 0, \ \delta_j \geqslant 0, \quad I, J \ \text{ arbitraires} \end{array} \right\}$$

ce qui peut encore s'écrire, en employant la notion d'enveloppe convexe:

$$Z = Co(R) + \mathbb{R}^+[\ Co(V)\ ] \quad .$$

Or, $Co(R)$ est l'enveloppe convexe d'un compact et est donc compact. De même pour $Co(V)$ et $\mathbb{R}^+[\ Co(V)\ ]$ reste un fermé. De plus la somme d'un compact et d'un fermé reste fermé et dès lors Z est un fermé.

Supposons par l'absurde $0 \notin Z$.

Alors par le théorème de séparation stricte pour un ensemble convexe fermé et un point n' appartenant pas à celui-ci [5], on a:

(5.2.6.) $\left\{ \begin{array}{l} \text{il existe } d \in \mathbb{R}^n, \ d \neq 0 \\ \text{il existe } \alpha \in \mathbb{R}, \ \alpha < 0 \end{array} \right.$ tel que $dz < \alpha < 0 \quad \forall \ z \in Z$

De (5.2.5.) et (5.2.6.) , on déduit que:

(5.2.7.) $\left\{ \begin{array}{ll} rd < 0, & \forall \ r \in R \\ vd \leqslant 0, & \forall \ v \in V. \end{array} \right.$

En effet: si $r \in R$, on rend $\beta_i = 0, \forall i$ sauf pour le coéfficient de r dans la combinaison de z. Dès lors, $\sum_{i=1}^{I} \beta_i = 1$, si on pose le coéfficient de r égal à 1;

si $v \varepsilon V$, prenons $\delta v + r \varepsilon Z$, $\delta \geqslant 0$ arbitraire;

donc $\delta v + r < 0$

ou $\delta vd < -rd$ où $-rd > 0$;

ou $vd < -rd/\delta \quad \forall \delta \geqslant 0$,

donc $vd \leqslant 0$.

Vu (5.2.7.) , on aurait donc trouvé une solution $d \neq 0$ appartenant à $\mathbb{R}^n$ du système (5.2.4.), ce qui contredit notre hypothèse.

Dès lors, $0 \varepsilon Z$ et donc pour I_0, J_0, β_i° et δ_j° nous avons:

$$(5.2.8.) \quad \begin{cases} \sum_{i=1}^{I_0} \beta_i^\circ q^i + \sum_{j=1}^{J_0} \delta_j^\circ v^j = 0 \\ q^i = u^\circ, \text{ pour un } i,\ 1 \leqslant i \leqslant I_0 \\ \sum_{i=1}^{J_0} \beta_i^\circ = 1 ,\ \beta_i^\circ \geqslant 0,\ \delta_j^\circ \geqslant 0. \end{cases}$$

En remplaçant q^i par $\sum_{i=1}^{l} \alpha_i w^i$, dans (5.2.8.), s'il en existent, nous obtenons une combinaison linéaire des vecteurs du système (A) .

En employant le fait que n+1 ou plus de vecteurs sont linéairement dépendant dans $\mathbb{R}^n$, on se ramène finalement au résultat (B).

CQFD.

CHAPITRE II

CONDITIONS NECESSAIRES ET SUFFISANTES D' OPTIMALITE

Nous allons formuler dans un premier temps des conditions nécessaires d'optimalite pour le P.I.C. sous la forme ci-dessous, vérifiant les mêmes hypothèses qu'au premier chapitre.

$$
(PT) \quad \begin{array}{ll}
\underset{x \in \mathbb{R}^n}{\text{minimiser}}\ F(x) & \\
\text{s.c. :} & G_i(x,t) \leqslant 0, \text{ quelque soit } t \in T_i;\ i=1,\dots,l \\
 & H_j(x,s) = 0, \text{ quelque soit } s \in S_j,\ j=1,\dots,m \\
 & x \in X^\circ
\end{array}
$$

Ensuite, dans un deuxième paragraphe nous postulerons des qualifications de contraintes, conditions qui vont nous permettre d'obtenir des conditions suffisantes d'optimalité pour caractériser la solution d'un P.I.C. sous la forme (PT).

Finalement, nous appliquerons ces critères à un problème simple de la théorie d'approximation (traduit sous la forme d'un P.I.C.).

II.1. Conditions nécessaires d'optimalité

Avant d'aborder le théorème exprimant des conditions nécessaires d'optimalité pour un P.I.C. (PT), nous nous proposons de mettre en évidence un lemme établissant un résultat à propos de la linéarisation de la fonction objectif et des contraintes autour d'un minimum local du problème d'optimisation (PT).

Lemme 1: Soit $\overline{x}$ est un minimum local du problème (PT)
Définissons les ensembles

$$D_i = \{ t \in T_i \text{ tel que } G_i(\overline{x},t) = 0 \};\ i=1,\dots,l$$

D_i, sous-ensemble de T_i, est donc constitué par l'ensemble des points t appartenant à T_i et qui saturent la i-ème contrainte d'inégalité de (PT)

Alors le système

$$(LS)\begin{cases} \nabla_x F(\overline{x})z < 0 & \\ \nabla_x G_i(\overline{x},t)z < 0 & \text{quelque soit } t \in D_i,\ i=1,\dots,l \\ \nabla_x H_j(\overline{x},s)z = 0 & \text{quelque soit } s \in S_j,\ j=1,\dots,m \end{cases}$$

n'admet pas de solution z dans $\mathbb{R}^n$.

Démonstration : Le lemme sera démontré par l'absurde. En effet, en supposant que le système (LS) admette une solution dans $\mathbb{R}^n$, nous construirons à partir de cette solution un nouvel élément de X°, vérifiant toutes les contraintes et tel que la valeur de la fonction objectif en cet élément soit strictement inférieure à la valeur de la fonction objectif en $\overline{x}$. $\overline{x}$ ne serait dès lors pas solution du problème (PT), ce qui contredit les hypothèses du lemme.

Supposons donc, qu'il existe une solution z^* appartenant à $\mathbb{R}^m$ du système (LS). Quelque soit ε strictement positif, définissons les deux ensembles suivants :

$D_i(\varepsilon) = \{ \overline{t} \in T_i$ tel que il existe $t \in D_i$: $d_i(\overline{t},t) \leqslant \varepsilon \}$ i=1,...,l (1)

$X°(\varepsilon) = \{ x \in X°$ tel que $\| x - \overline{x} \|_{\mathbb{R}^m} \leqslant \varepsilon \}$, c'est donc la boule fermée centrée au minimum local de (PT) et de rayon $\varepsilon > 0$.

Nous affirmons qu'on peut trouver alors deux nombres strictement positifs δ et ε tels que l'on ait :

$$(2.1.1.)\quad \forall\, x \in X°(\varepsilon) \quad \begin{cases} \nabla_x F(x) z^* < -\delta \\ \nabla_x G_i(x,t) z^* < -\delta \quad \forall\, t \in D_i(\varepsilon), i=1,\dots,l \end{cases}$$

En effet, supposons que (2.1.1.) ne soit pas vérifié quelque soient ε et δ strictement positifs. Comme, par définition du problème, F, respectivement G_i, et leurs dérivées partielles sont continues en x, respectivement en x, quelque soit t appartenant à T_i, i=1,...,l; F et toutes les G_i, i=1,...,l, et leurs dérivées partielles atteignent leurs bornes comme fonctions continues sur des compacts. Donc la négation de (2.1.1.), c.à.d. soit $\nabla_x F(x)$, soit $\nabla_x G_i$ non bornées et par le caractère compact de $X°(\varepsilon)$ et $D_i(\varepsilon)$, on déduit qu'il faut soit $\nabla_x F(\overline{x}) z^* \geqslant 0$, soit $\nabla_x G_i(\overline{x},\overline{t}) z^* \geqslant 0$, i=1,...,l, ce qui contredit notre hypothèse que z^* est solution du système (LS).

---

(1) Nous nous sommes proposés, de noter $d_i(.,.)$, la métrique sur T_i ainsi que $\| . - . \|_{\mathbb{R}^m}$, la métrique de $\mathbb{R}^m$, définie à partir de sa norme.

Sachant que (2.1.1.) est satisfait pour un ε et un δ strictement positifs, montrons que l'on peut trouver un nombre strictement positif μ tel qu'on ait :

(2.1.2.) $\quad G_i(\overline{x},t) < -\mu \qquad \forall\ t \in T_i \setminus D_i(\varepsilon),\ i=1,\dots,l.$

ce qui signifie que quelque soit $t \in T_i \setminus D_i(\varepsilon)$, la ième contrainte d'inégalité du problème (PT) n'est pas saturée au minimum local considéré $\overline{x}$, $i=1,\dots,l$.

En effet, supposons qu'il existe un indice i_o avec $1 \leqslant i_o \leqslant l$ tel que la relation (2.1.2.) ne soit pas vérifiée. Alors, il existe une suite dont les éléments t^k appartiennent à $T_{i_o} \setminus D_{i_o}(\varepsilon)$ pour $k = 1,2,\dots$ Par la compacité de T_{i_o}, il existe alors une sous-suite convergente que nous continuons à noter t^k, pour alléger l'écriture, telle que t^k tende vers $\overline{t}$ pour k tendant vers l'infini, où $\overline{t}$ appartient à T_{i_o}. Ce qui prouve que $G_{i_o}(\overline{x},\overline{t}) = 0$ par la continuité de G_{i_o} en son second argument appartenant à T_i. Or, $G_{i_o}(\overline{x},\overline{t}) = 0$, entraîne que $\overline{t}$ appartient à D_{i_o}, par définition même de D_{i_o} et donc $\overline{t}$ appartient à l'intérieur de $D_{i_o}(\varepsilon)$, puisque $\varepsilon > 0$. Ceci cependant, est impossible car aucun point intérieur de $D_{i_o}(\varepsilon)$ ne peut être limite d'une suite de points t^k appartenant à $T_{i_o} \setminus D_{i_o}(\varepsilon)$, car sinon chaque voisinage de $\overline{t}$ appartenant à $D_{i_o}(\varepsilon)$ contiendrait des t^k appartenant à $D_{i_o}(\varepsilon)$, puisque t^k converge vers $\overline{t}$, lorsque k tend vers l'infini. On aboutit à une contradiction et dès lors, la relation (2.1.2.) doit être vraie.

Le théorème de la valeur moyenne et les relations (2.1.1.) et (2.1.2.) vont nous permettre de construire à partir de z^* un nouveau point de $X°$, vérifiant toutes les contraintes et tel que la valeur de la fonction objectif en ce point soit strictement inférieure à la valeur de la fonction objectif en $\overline{x}$, ce qui sera en contradiction avec l'hypothèse que $\overline{x}$ est minimum local du problème.

Le théorème de la moyenne appliqué à F et G_i nous fournit les relations suivantes :

$$\forall \gamma \quad \begin{aligned} &(2.1.3.) \quad F(\overline{x}+\gamma z^*) = F(\overline{x}) + \gamma\nabla_x F(\overline{x}+\theta_0(\gamma)\gamma z^*)z^* \qquad \text{et} \\ &(2.1.4.) \quad G_i(\overline{x}+\gamma z^*,t) = G_i(\overline{x},t) + \gamma\nabla_x G_i(\overline{x}+\theta_i(\gamma)\gamma z^*,t)z^* \\ &\qquad\qquad\qquad \forall\, t \in T_i,\ i=1,\dots,l \end{aligned}$$

où chaque $\theta_i(\gamma)$ est tel que $0 < \theta_i(\gamma) < 1$, $i=0,1,\dots,l$.

Nous pouvons déterminer γ de manière à aboutir à la contradiction annoncée pour ce point $\overline{x} + \gamma z^*$.

En effet, il existe $\overline{\gamma}$, strictement positif tel que, quelque soit γ avec $0 \leqslant \gamma \leqslant \overline{\gamma}$, on ait : $\overline{x}+\gamma z^* \in X°$, puisqu'il est toujours possible de trouver un tel $\overline{\gamma}$ strictement positif, car X° est un ensemble ouvert de $\mathbb{R}^n$ et $\overline{x}$ appartient à X°.

Soit $\gamma_0 = \varepsilon / \| z^* \|_{\mathbb{R}^n}$. Remarquons que, γ_0 est strictement positif.

De plus, on a : $F(\overline{x}+\gamma z^*) < F(\overline{x})$ quelque soit γ avec $0 < \gamma \leqslant \gamma_0$.

En effet, par la relation (2.1.3.) on a que :

$$F(\overline{x}+\gamma z^*) = F(\overline{x}) + \gamma\nabla_x F(\overline{x}+\theta_0(\gamma)\gamma z^*)z^*.$$

Mais la relation (2.1.1.) nous permet de trouver un δ' strictement positif tel que $\nabla_x F(\overline{x}+\theta_0(\gamma)\gamma z^*) < \delta'$, si le point $\overline{x}+\theta_0(\gamma)\gamma z^*$ est dans X°.

Soit $\delta = \delta'/\gamma z^*$. Nous avons les estimations suivantes :

$\| \overline{x}+\theta_0(\gamma)\gamma z^* - \overline{x} \|_{\mathbb{R}^n} = \| \theta_0(\gamma)\gamma z^* \|_{\mathbb{R}^n} = \theta_0(\gamma)\gamma \| z^* \|_{\mathbb{R}^n}$.

Or, par définition de γ_0 et le fait que $0 < \theta_0(\gamma) < 1$ et $\gamma \leqslant \gamma_0$, on déduit finalement : $\| \overline{x}+\theta_0(\gamma)\gamma z^* - \overline{x} \|_{\mathbb{R}^n} < \gamma_0 \| z^* \|_{\mathbb{R}^n} = \varepsilon$, ce qui revient à écrire que $\overline{x}+\theta_0(\gamma)\gamma z^*$ appartient à X°, par la définition même de X°.

Si nous parvenons à démontrer la relation (2.1.5.) ci-dessous, alors $\overline{x}+\gamma z^{*}$ sera le nouveau minimum.

(2.1.5.) $\quad G_i(\overline{x}+\gamma z^{*},t) \leqslant 0 \qquad \forall\ t \in T_i$ et quelque soit γ avec $0 \leqslant \gamma \leqslant \gamma_i$, $i=1,\dots,l$ et $\gamma \leqslant \gamma_0$.

Soit γ_i tel que :

$$\gamma_i . \max_{\substack{t \in T_i \\ x \in X^{\circ}(\varepsilon)}} |\ \nabla_x G_i(x,t) z^{*}\ | < \mu \quad \text{pour } i=1,\dots,l.$$

Remarquons d'une part que nous pouvons prendre le maximum, puisque T_i et $X^{\circ}(\varepsilon)$ sont des ensembles compacts et G_i, fonction continue sur un compact y atteint ses bornes; et que d'autre part tous les γ_i sont strictement positifs.

Montrons donc que $\overline{x}+\gamma z^{*}$ vérifie toutes les contraintes du type inégalité, en nous servant de la relation (2.1.4.) :

$$G_i(\overline{x}+\gamma z^{*},t) = G_i(\overline{x},t) + \gamma \nabla_x G_i(\overline{x}+\theta_i(\gamma)\gamma z^{*},t) z^{*} \qquad \forall\ t \in T_i .$$

En effet, d'une part on a que, quelque soit t appartenant à $D_i(\varepsilon)$,

$i=1,\dots,l$: $\qquad G_i(\overline{x}+\gamma z^{*},t) \leqslant \underbrace{G_i(\overline{x},t)}_{\leqslant 0} - \gamma\delta \leqslant -\gamma\delta$

puisque $\overline{x}+\theta_i(\gamma)\gamma z^{*}$ appartient à $X^{\circ}(\varepsilon)$ et qu'on peut donc appliquer la relation (2.1.1.). D'autre part, quelque soit t appartenant à $T_i \setminus D_i(\varepsilon)$, $i=1,\dots,l$, la relation (2.1.2.) entraîne :

$$G_i(\overline{x}+\gamma z^{*},t) \leqslant -\mu + \gamma \nabla_x G_i(\overline{x}+\theta_i(\gamma)\gamma z^{*},t) z^{*}$$

$$\leqslant -\mu + \gamma_i . \max_{\substack{t \in T_i \\ x \in X^{\circ}(\varepsilon)}} |\ \nabla_x G_i(x,t) z^{*}\ |$$

puisque $\overline{x}+\theta_i(\gamma)\gamma z^{*} \in X^{\circ}(\varepsilon)$ pour $0 \leqslant \gamma \leqslant \gamma_0$, et donc

$\leqslant -\mu + \mu = 0$, quelque soit γ avec $0 \leqslant \gamma \leqslant \gamma_i$ par la définition même des γ_i.

En regroupons les deux résultats ci-dessus, nous obtenons la relation (2.1.5.). Finalement, comme chaque contrainte du type égalité $H_j(x,s)$ est linéaire en son premier argument x, c.à.d. est de la forme:

$$H_j(x,s) = \nabla_x H_j(x,s).x + b_j(s), \qquad j=1,\ldots,m;$$

on a que, quelque soit $\gamma \geqslant 0$ et quelque soit $s \in S_j$:

$$H_j(\overline{x}+\gamma z^*,s) = H_j(\overline{x},s) + \gamma\nabla_x H_j(\overline{x},s)z^* = 0, \quad j=1,\ldots,m \;,$$

puisque $H_j(\overline{x},s) = 0$ par le système (LS) et par la linéarité de la fonction H_j en x.

En définissant γ^* comme étant le minimum de $\overline{\gamma}$, γ_0, $\gamma_1,\ldots\ldots,\gamma_l$ décrits ci-dessus et en remarquant que γ^* est strictement positif, nous obtenons de ce qui précède:

$$\forall\gamma, 0 \leqslant \gamma \leqslant \gamma^* \begin{cases} F(\overline{x}+\gamma z^*) < F(\overline{x}) & \\ G_i(\overline{x}+\gamma z^*,t) \leqslant 0 & \forall\ t \in T_i,\ i=1,\ldots,l \\ H_j(\overline{x}+\gamma z^*,s) = 0 & \forall\ s \in S_j,\ j=1,\ldots,m \\ \text{et } \overline{x}+\gamma z^* \in X^\circ & \end{cases}$$

ce qui contredit le fait que $\overline{x}$ soit minimum local du problème (PT). Dès lors, le système (LS) n'admet pas de solution $z^* \in \mathbb{R}^n$.

CQFD.

Nous pouvons à présent, énoncer le théorème des conditions nécessaires d'optimalité pour un problème d'optimisation, (PT), comportant une infinité de contraintes. La preuve de ce théorème est une application directe du lemme précédent et du théorème généralisé de Motzkin.

<u>Théorème 1</u> : Soit $\overline{x}$ un minimum local du problème (PT).
Alors, il existe des entiers s_o et s avec $0 \leqslant s_o \leqslant s \leqslant n$ tels que

(1.1.) on ait s_o indices i_k avec $1 \leqslant i_k \leqslant l$ et s_o points $t^k \in D_{i_k} = \{ t \in T_{i_k} \text{ tel que } G_{i_k}(\overline{x},t) = 0 \}$, pour $k=1,\ldots,s_o$ et tels que

(1.2.) on ait $s-s_o$ indices j_k avec $1 \leqslant j_k \leqslant m$ et $s-s_o$ points $s^k \in S_{j_k}$, pour $k=s_o+1,\ldots,s$, tels que

(1.3.) on ait $s+1$ nombres réels λ_k avec $\lambda_o > 0$ ou $s_o \geqslant 1$ et $\lambda_k > 0$, pour $k = 1,\ldots,s_o$; ayant la propriété :

(1.4.) $$\lambda_o \nabla_x F(\overline{x}) + \sum_{k=1}^{s_o} \lambda_k \nabla_x G_{i_k}(\overline{x},t^k) + \sum_{k=s_o+1}^{s} \lambda_k \nabla_x H_{j_k}(\overline{x},s^k) = 0$$

<u>Démonstration</u> : Par le lemme précédent, le système ci-dessous n'admet pas de solution dans $\mathbb{R}^n$.

$$(LS) \begin{cases} \nabla_x F(\overline{x})z < 0 & \\ \nabla_x G_i(\overline{x},t)z < 0 & \forall\ t \in D_i, \text{ pour } i=1,\ldots,l \\ \nabla_x H_j(\overline{x},s)z = 0 & \forall\ s \in S_j, \text{ pour } j=1,\ldots,m \end{cases}$$

Choisissons alors pour $\nabla_x F(\overline{x})$ le vecteur $u°$ du théorème généralisé de Motzkin (1) et remarquons que l'ensemble

(1) Ibid page 21.

$$\{ \nabla_x F(\overline{x}) \} \cup \bigcup_{i=1}^{l} \{ \nabla_x G_i(\overline{x},t) \text{ tel que } t \in D_i \}$$

est compact, puisque chaque $\nabla_{x_j} G_i(\overline{x},t)$ est continue pour $j=1,\dots,n$ et que D_i est compact. En appliquant le théorème généralisé de Motzkin, au système (LS), nous obtenons les résultats espérés (1.1.) à (1.4.), puisque si le système

$$\begin{cases} uz < 0 & \forall u \in U \\ vz \leqslant 0 & \forall v \in V \\ wz = 0 & \forall w \in W \end{cases}$$

où U et V sont des sous-ensembles compacts de $\mathbb{R}^n$ et W est un sous-ensemble arbitraire de $\mathbb{R}^m$,

n'admet pas de solution z dans $\mathbb{R}^m$, alors le système

$$\begin{cases} uz < 0 & \forall u \in U \\ vz < 0 & \forall v \in V \\ wz = 0 & \forall w \in W \end{cases}$$

n'admet pas de solution z dans $\mathbb{R}^m$.

CQFD.

II.2. La qualification des contraintes

Remarquons que les conditions nécessaires d'optimalité du problème (PT) permettent au multiplicateur λ_0 de la fonction objectif d'être nul. Dans ce cas, les conditions nécessaires d'optimalité n'engendrent plus le terme $\nabla_x F(\overline{x})$ et ne fournissent plus aucun renseignement quant à la fonction objectif F du problème (PT). C'est pourquoi, nous allons imposer des conditions supplémentaires sur les contraintes pour assurer que le multiplicateur λ_0 de la fonction objectif soit strictement positif, afin que les conditions nécessaires d'optimalité nous fournissent des renseignements relatifs à la fonction objectif F.

Pour assurer en programmation non linéaire finie, c'est à dire pour des problèmes d'optimisation non linéaires avec un nombre fini de contraintes, que le multiplicateur λ_0 soit distinct de zéro, on suppose que le problème satisfait une qualification de contraintes [6] : les contraintes doivent satisfaire une condition régularisante qui assure que le multiplicateur λ_0,apparaissant dans l'énoncé des conditions nécessaires d'optimalité,de la fonction objectif , soit strictement positif.

Il est donc naturel de formuler pour des problèmes de programmation non linéaires comportant un nombre infini de contraintes, des qualifications de contraintes, qui seront des généralisations de celles de la programmation non linéaire finie.

De plus, la plupart des problèmes d'approximation vérifient une des qualification de contraintes proposées ci-dessous et permettent ainsi de formuler des conditions nécessaires d'optimalité non banales pour les problèmes considérés.

Première forme de qualification de contraintes (1).

La première qualification de contraintes est une généralisation de la condition de SLATER, aussi connue sous le nom de condition du point intérieur.

Le problème (PT) satisfait QC1, si chaque fonction G_i est pseudo-convexe en x, quelque soit t appartenant à T_i, pour i=1,...,l et que de plus, il existe un point $\overline{x}$ dans $\mathbb{R}^n$ tel que

(i) $G_i(\overline{x},t) < 0$ $\forall t \in T_i$, pour i=1,...,l et

(ii) $H_j(\overline{x},s) = 0$ $\forall s \in S_j$, pour j=1,...,m

Deuxième forme de qualification de contraintes.

La deuxième qualification de contraintes est une généralisation de la condition de ARROW - HURWICZ - UZAWA, aussi connue sous le nom condition d'inégalité stricte.

Nous dirons que le problème (PT) satisfait QC2 en un point donné, admissible, $\overline{x}$ (où $\overline{x}$ "admissible" signifie que $\overline{x} \in X$, où $X = \{ x \in X° \text{ tel que } G_i(x,t) \leq 0, \forall t \in T_i,$ pour $i=1,\ldots,l$ et $H_j(x,s) = 0, \forall s \in S_j$, pour $j=1,\ldots,m \}$), si pour tout choix des entiers s_o et s avec $0 \leq s_o \leq s \leq n$ et

(i) tout choix de s_o indices i_k avec $1 \leq i_k \leq l$ et s_o points t^k appartenant à l'ensemble $D_{i_k} = \{ t \in T_{i_k} \text{ tel que } G_{i_k}(\overline{x},t) = 0 \}$, pour $k=1,\ldots,s_o$ et

(ii) tout choix de $s-s_o$ indices j_k avec $1 \leq j_k \leq m$ et $s-s_o$ points s^k appartenant à l'ensemble S_{j_k}, pour $k=s_o+1,\ldots,s$

il existe un vecteur $y = (y_1,\ldots,y_n) \in \mathbb{R}^n$ tel que

(iii) $\sum_{q=1}^{n} y_q \nabla_{x_q} G_{i_k}(\overline{x},t^k) < 0$, pour $k=1,\ldots,s_o$ et

(iv) $\sum_{q=1}^{n} y_q \nabla_{x_q} H_{j_k}(\overline{x},s^k) = 0$, pour $k=s_o+1,\ldots,s$.

(1) Quand nous ferons dans la suite appel à la première qualification de contraintes, nous l'abrégeons comme QC1, ainsi que la seconde comme QC2.

Remarquons que pour la plupart des problèmes, il est plus facile de vérifier QC1 que QC2; ceci ést en particulier vrai pour les problème d'approximation, tels que les problèmes au sens de Chebycheff, comme nous allons le voir à propos d'un exemple au quatrième paragraphe. De plus, sous l'hypothèse que toutes les fonctions G_i soient différentiables par rapport à leur premier argument en $\overline{x}$, pour $i=1,\ldots,l$, QC1 entraîne QC2 [5].

II. 3. Conditions nécessaires d'optimalité avec $\lambda_0 > 0$

Théorème 2 : Soit $\overline{x}$ un minimum local du problème (PT). Si, soit QC1, soit QC2 est satisfaite en $\overline{x}$, alors le multiplicateur λ_0 de la fonction objectif dans l'énoncé des conditions nécessaires d'optimalité du théorème 1 est strictement positif.

Démonstration : En effet, si chaque fonction G_i est différentiable par rapport à son premier argument en $\overline{x}$, QC1 entraîne QC2.Donc il suffit de prouver le résultat, si le problème vérifie QC2.

Par le théorème 1 on a :

$$(2.1.) \quad \lambda_0 \nabla_x F(\overline{x}) + \sum_{k=1}^{s_0} \lambda_k \nabla_x G_{i_k}(\overline{x},t^k) + \sum_{k=s_0+1}^{s} \lambda_k \nabla_x H_{j_k}(\overline{x},s^k) = 0.$$

Supposons, par l'absurde, que λ_0 soit nul. Alors, de la relation (2.1.) et pour le $y \in \mathbb{R}^m$ intervenant dans l'énoncé de QC2, nous avons :

$$(2.2.) \quad \sum_{k=1}^{s_0} \lambda_k \, [\, y \, \nabla_x G_{i_k}(\overline{x},t^k) \,] + \sum_{k=s_0+1}^{s} \lambda_k \, [\, y \, \nabla_x H_{j_k}(\overline{x},s^k) \,] = 0.$$

Or, la première somme de (2.2.) est strictement négative et la deuxième est nulle; nous aboutissons donc à une contradiction évidente $(0 > 0)$, dès lors, le multiplicateur de la fonction objectif est strictement positif.

CQFD.

Puisque nous nous sommes servis explicitement de contraintes pour aboutir à des conditions nécessaires d'optimalité non banales pour des P.I.C. et puisque l'introduction de ces conditions additionnelles restreint la généralité des problèmes d'optimisation comportant une infinité de contraintes auxquelles on peut appliquer ce critère non trivial (1), il est important de remarquer que ces conditions sont réellement nécessaires pour garantir que le multiplicateur λ_o de la fonction objectif soit strictement positif.

Reprenons à cette fin, l'exemple considéré pour illustrer une variante du P.I.C. (PT) (2). Il s'agissait du problème suivant :

$$\begin{aligned} &\underset{x \in \mathbb{R}^m}{\text{minimiser}} \; -x \\ &\quad \text{s.c.} \quad : \quad x.t - t^2 \leqslant 0 \quad \forall t \in [0,1] \end{aligned}$$

qui admet évidemment la solution unique $\overline{x} = 0$. Remarquons, que la fonction contrainte $G(x,t) = x.t - t^2$ ne vérifie aucune des qualifications de contraintes QC1 et QC2 (3), au point $x = 0$. Les conditions nécessaires d'optimalité entraînent qu'il existe au plus deux constantes λ_o et λ_1 strictement positives telles que :

$$\lambda_o \nabla_x F(\overline{x}) + \lambda_1 \nabla_x G(\overline{x},\overline{t}) = 0$$

où $\overline{t} \in [0,1]$ est un point tel que $G(\overline{x},\overline{t}) = 0$. Pour le problème considéré ci-dessus, cette relation s'écrit :

$$\lambda_o(-1) + \lambda_1(0) = 0 \qquad (\alpha)$$

puisque $\nabla_x G(\overline{x},t) = t$ et $\overline{t} = 0$ est le seul point où la contrainte est active. D'autre part, la relation (α) entraîne que λ_o est nul et que λ_1 peut être un nombre positif quelconque. Donc $\lambda_o > 0$ est impossible pour ce problème. Cet exemple

(1) Ibid, théorème 2.
(2) Ibid p. 13.
(3) Pour ce problème, il suffit de considérer les qualifications de contraintes non généralisées du point intérieur et de l'inégalité stricte [6].

nous montre donc, que les qualifications de contraintes sont nécessaires, même pour le cas où la fonction objectif et les fonctions contraintes sont linéaires.

II. 4. Conditions nécessaires et suffisantes

En imposant des hypothèses légèrement plus fortes quant à la fonction objectif F et aux fonctions contraintes du problème (PT), les conditions nécessaires d'optimalité du théorème 2, sont également des conditions suffisantes d'optimalité.

Définissons à cette fin, une notion de quasi-convexité appropriée à nos contraintes d'inégalité, qui n'est en fait qu'une généralisation de celle donnée au premier chapitre (1).

Une fonction à valeurs réelles G définie sur $\mathbb{R}^n \times T$, où T est un ensemble arbitraire, est dite quasi-convexe au point $\overline{x} \in \mathbb{R}^n$, si pour tout $x \in \mathbb{R}^n$ tel que $G(x,t) \leqslant G(\overline{x},t)$ quelque soit $t \in T$, on a :

$$G[(1-\lambda)\overline{x} + \lambda x, t] \leqslant G(\overline{x},t), \quad \text{pour } 0 \leqslant \lambda \leqslant 1, \text{ quelque soit } t \in T.$$

La fonction G est dite quasi-convexe sur un ensemble $\Gamma \subset$, si elle est quasi-convexe en tout point $x \in \Gamma$.

Avant d'aboutir au théorème exprimant des conditions nécessaires et suffisantes d'optimalité pour le P.I.C. (PT), nous rappelons deux propriétés bien connues en programmation non linéaire [5], relatives aux fonctions quasi-convexes et pseudo-convexes.

---

(1) Ibid p. 15.

Soient Γ un ensemble ouvert de $\mathbb{R}^n$ et T un ensemble quelconque.
Soit G une fonction à valeurs réelles et définie sur $\mathbb{R}^n \times T$.
Soit F une fonction à valeurs réelles et définies sur $\mathbb{R}^n$.

Propriété 1 : Si x^1 et x^2 appartiennent à Γ
et si la fonction G est différentiable par rapport à son premier argument, quelque soit $t \in T$ et si elle est quasi-convexe en x^1, alors
$G(x^2,t) \leqslant G(x^1,t)$ entraîne que
$\nabla_x G(x^1,t)(x^2 - x^1) \leqslant 0$, quelque soit $t \in T$.

Propriété 2 : Soit $\overline{x} \in \Gamma$.
Supposons que F soit pseudo-convexe en $\overline{x}$
alors $\nabla_x F(\overline{x})(x - \overline{x}) \geqslant 0$, quelque soit $x \in \Gamma$
entraîne que $F(\overline{x}) = \underset{x \in \Gamma}{\text{minimum}} \, F(x)$

Connaissant ces propriétés, il est facile de démontrer le théorème de caractérisation d'un optimum pour un problème sous la forme (PT).

Théorème 3 : Outre les hypothèses propres au problème (PT), (1) Supposons que, la fonction F soit pseudo-convexe sur X° et que chaque fonction G_i soit quasi-convexe, ainsi que le problème satisfasse soit QC1, soit QC2 au point $\overline{x}$.

Alors $\overline{x}$ résout le problème (PT) si et seulement si il existe des entiers s_o et s avec $1 \leqslant s_o \leqslant s \leqslant n$ tels que

(3.1.) on ait s_o indices i_k avec $1 \leqslant i_k \leqslant l$ et s_o points t^k appartenant à $T_{i_k} = \{ t \in D_{i_k}$ tel que $G_{i_k}(\overline{x},t) = 0 \}$, pour $k = 1,\dots,s_o$ et tels que

(3.2.) on ait $s-s_o$ indices j_k avec $1 \leqslant j_k \leqslant m$ et $s-s_o$ points dans S_{j_k}, pour $k=s_o+1,\dots,s$ tels que

(3.3.) on ait s nombres réels $\lambda_k > 0$ pour $k=1,\dots,s_o$, ayant la propriété

(3.4.) $$\nabla_x F(\overline{x}) + \sum_{k=1}^{s_o} \lambda_k \nabla_x G_{i_k}(\overline{x},t^k) + \sum_{k=s_o+1}^{s} \lambda_k \nabla_x H_{j_k}(\overline{x},s^k) = 0 .$$

Démonstration : En effet, les conditions sont évidemment nécessaires par le théorème 2.

Pour démontrer la suffisance, supposons que x soit un point admissible du problème (PT), c'est à dire que x appartienne à l'ensemble $\{ x \in X°$ tel que $G_i(x,t) \leqslant 0, \forall t \in T_i, i=1,\dots,l$ et $H_j(x,s) = 0, \forall s \in S_j, j=1,\dots,m \}$

---

(1) Ibid p. 6.

Comme chaque fonction G_i est quasi-convexe et différentiable, la proriété 1,(1), permet d'affirmer que:
$\nabla_x G_i(x^*,t^*)(x - x^*) \leqslant 0$, $\forall x^* \in X°$ et t^* vérifiant $G_i(x^*,t^*) = 0$, pour $i=1,\dots,l$.

Donc, puisque $\lambda_k > 0$, pour $k=1,\dots,s_o$, on a:

$$\sum_{k=1}^{s_o} \lambda_k \nabla_x G_{i_k}(\overline{x},t^k)(x - \overline{x}) \leqslant 0. \tag{3.5.}$$

D'autre part, de la linéarité de H_j, pour $j=1,\dots,m$, on déduit que:

$\nabla_x H_j(x^*,s^*)(x - x^*) = 0$, $\forall x^* \in X°$ et s^* vérifiant $H_j(x^*,s^*) = 0$;
donc que:

$$\sum_{k=s_o+1}^{s} \lambda_k \nabla_x H_{j_k}(\overline{x},t^k)(x - \overline{x}) = 0. \tag{3.6.}$$

En introduisant les relations (3.5.) et (3.6.) dans (3.4.), on obtient finalement : $\nabla_x F(\overline{x})(x - \overline{x}) \geqslant 0$.

Or, ceci par la proriété 2, entraîne : $F(x) \geqslant F(\overline{x})$,(1), à cause de la pseudo-convexité de F.
Puisque x est un point quelcoque du domaine admissible du problème (PT), on en conclut que $\overline{x}$ résout le problème (PT).

CQFD.

---

(1) Ibid p.43 .

II.5. Application

Dans un dernier paragraphe nous allons reprendre le problème d'approximation au sens de Chebycheff comportant un nombre fini de contraintes d'interpolation, qui nous a servi d'exemple introductif des P.I.C. (1), en lui appliquant les critères développés dans les paragraphes précédents.

Rappelons que ce problème d'approximation au sens de Chebycheff conduisait au problème d'optimisation comportant une infinité de contraintes de la forme (PT) suivante :

$$
\text{(CC)}\qquad
\begin{array}{lll}
\underset{x_1,\dots,x_n,m}{\text{minimiser}}\ m & & \\
\text{s.c. : (a)} & -f(t) + \sum_{i=1}^{n} x_i\phi_i(t) - m \leqslant 0 & \\
 & & \text{quelque soit } t \in T \\
\text{(b)} & +f(t) - \sum_{i=1}^{n} x_i\phi_i(t) - m \leqslant 0 & \\
\text{(c)} & f(t^j) - \sum_{i=1}^{n} x_i\phi_i(t^j) = 0 & \text{pour } j=1,\dots,q
\end{array}
$$

Remarquons, que la fonction objectif définie par $m = \sup_{t \in T} |f(t)-\sum_{i=1}^{n} x_i\phi_i(t)$

et toutes les contraintes du problème (CC) sont linéaires en les variables m et x et sont donc à fortiori convexes en ces arguments.

---

(1) Ibid p. 2.

Vérifions d'abord, si le problème au sens de Chebycheff vérifie l'une des deux qualifications de contraintes QC1 ou QC2 (1).

Supposons, qu'il existe un vecteur, soit $\overline{x} = (\overline{x}_1,\ldots,\overline{x}_n)$ de $\mathbb{R}^m$ qui vérifie la contrainte (c) du problème d'optimisation (CC). Alors, puisque la fonction objectif f et toutes les fonctions ϕ_i, pour i=1,...,q sont continues sur le compact T, elles y atteignent leurs bornes et par conséquent

$\sum_{i=1}^{n} \overline{x}_i\phi_i(t)$ est borné. D'où en choisissant un $\overline{m} \in \mathbb{R}$ "assez grand" les contraintes (a) et (b) du problème (CC) seront satisfaites comme inégalités strictes. Dès lors, pour ce choix de $\overline{x} \in \mathbb{R}^m$ et $\overline{m} \in \mathbb{R}$, on aura :

$$-f(t) + \sum_{i=1}^{n} \overline{x}_i\phi_i(t) - \overline{m} < 0$$

$$f(t) - \sum_{i=1}^{n} \overline{x}_i\phi_i(t) - \overline{m} < 0 \qquad \text{quelque soit } t \in T \text{ et}$$

$$f(t^j) - \sum_{i=1}^{n} \overline{x}_i\phi_i(t^j) = 0, \qquad \text{pour } j=1,\ldots,q$$

Le problème d'optimisation vérifie donc QC1, c'est à dire la condition généralisée de SLATER (1). De plus, comme la fonction objectif est linéaire, donc convexe, elle est pseudo-convexe par les implications rappelées au paragraphe I.4.5.. De même, comme chaque contrainte est linéaire, donc convexe elle est aussi quasi-convexe.

On est donc dans les conditions d'applicabilité du théorème de caractérisation d'une solution d'un P.I.C..

---

(1) Ibid p. 37.

En appliquant dès lors le théorème 3 de caractérisation au problème (CC), nous obtenons les conditions ci-dessous pour le problème de Chebycheff comportant un nombre fini de contraintes d'interpolation.

Théorème 4 : Soit T, un sous-ensemble compact d'un espace métrique complet. Soient f et $\phi_1,\ldots,\phi_n$ des fonctions continues sur T à valeurs réelles.

Soit $D = \{ t^1,\ldots,t^q \}$ l'ensemble des q points d'interpolation choisis dans T.

Supposons, que la fonction objectif f n'appartienne pas à l'espace engendré par les fonctions $\phi_1,\ldots,\phi_n$.

Alors un vecteur $x^* = (x_1^*,\ldots,x_n^*)$ de $\mathbb{R}^n$ résout le problème d'approximation au sens de Chebycheff si et seulement si, on peut exprimer l'origine de $\mathbb{R}^n$ comme une combinaison linéaire d'au plus n+1 points appartenant aux ensembles :

$$\left\{ e(t)\begin{bmatrix} \phi_1(t) \\ \vdots \\ \phi_n(t) \end{bmatrix} \text{ tel que } |e(t)| = \|e\|_T \right\} \text{ où } e = \sum_{i=1}^{n} x_i^*\phi_i - f \text{ est}$$

la fonction erreur du problème d'approximation, $\|e\|_T = \sup\limits_{t \in T} |e(t)|$

$$\text{et } \left\{ \begin{bmatrix} \phi_1(t^j) \\ \vdots \\ \phi_n(t^j) \end{bmatrix} \text{ tel que } j=1,\ldots,q \right\}.$$

De plus, au moins un point de cette combinaison linéaire doit appartenir au premier ensemble, et tout point de cette combinaison appartenant au premier ensemble est affecté d'un coéfficient positif.

Démonstration: Le problème (CC) vérifie une des qualifications de contraintes exigées par le théorème 3, comme on vient de voir.

Dès lors, le théorème 3 et par la remarque sur l'équivalence des qualifications de contraintes (1), on a que (x^*,m^*) constitue une solution du problème (CC) si et seulement si :

$$(5.1.)\quad \begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix} + \sum_{k=1}^{s_o} \lambda_k \begin{bmatrix}-1\\(-1)^{\varepsilon_k}\phi_1(t^k)\\ \vdots\\(-1)^{\varepsilon_k}\phi_n(t^k)\end{bmatrix} + \sum_{k=s_o+1}^{s} \mu_k \begin{bmatrix}0\\ \phi_1(t^{j_k})\\ \vdots\\ \phi_n(t^{j_k})\end{bmatrix} = \begin{pmatrix}0\\0\\ \vdots\\0\end{pmatrix}$$

où $\varepsilon_k = \begin{cases}0, \text{ si la contrainte (a) de (CC) est active en } t^k\\ 1, \text{ si la contrainte (b) de (CC) est active en } t^k\end{cases}$

et $\lambda_k > 0$, pour $k=1,\dots,s_o$.

D'autre part, on a que :

$$e(t^k) = \sum_{i=1}^{n} x_i^*\phi_i(t^k) - f(t^k) = \begin{cases} m^*, \text{ si la contrainte (a) de (CC) est active en } t^k, \text{càd } \varepsilon_k=0\\ -m^*, \text{ si la contrainte (b) de (CC) est active en } t^k, \text{càd } \varepsilon_k=0\end{cases}$$

et $m^* > 0$, car la fonction objectif f n'appartient pas à l'espace engendré par les fonctions $\phi_1,\dots,\phi_n$.

D'où, en définissant $\mu_k = \lambda_k/m^*$, pour $k=1,\dots,s_o$ la relation (5.1.) peut se récrire :

(1) Ibid p. 38.

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \sum_{k=1}^{s_o} \mu_k \, e(t^k) \begin{bmatrix} -1 \\ \phi_1(t^k) \\ \vdots \\ \phi_n(t^k) \end{bmatrix} + \sum_{k=s_o+1}^{s} \mu_k \begin{bmatrix} 0 \\ \phi_1(t^{j_k}) \\ \vdots \\ \phi_n(t^{j_k}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

où $\mu_k > 0$, pour $k=1,\dots,s_o$ et $s \leqslant n+1$.

CHAPITRE III

PARTIE ALGORITHMIQUE

Dans un dernier chapitre, nous considèrons les P.I.C. d'un point de vue algorithmique. Nous nous proposons, partant d'un probleme d'optimisation comportant une infinité de contraintes, d'établir une méthode de résolution, permettant d'aboutir à sa solution. Nous formulerons deux algorithmes, que nous analyserons ensuite quant à leur convergence.

Considèrons à cette fin le P.I.C. suivant :

$$\min_{x \in \mathbb{R}^n} \text{imiser } F(x)$$

$$\text{s.c. :} \quad G_i(x,t) \leq 0, \ \forall t \in T; \text{ pour } i=1,\dots,l$$

$$K_j(x) \leq 0; \text{ pour } j=1,\dots,m$$

$$x \in X°$$

où $X°$ est un ensemble compact de $\mathbb{R}^n$,

T est un espace métrique compact,

$F : \begin{array}{l}\mathbb{R}^n \longrightarrow \mathbb{R}\\ x \rightsquigarrow F(x)\end{array}$ est une fonction continue en x,

$G_i : \begin{array}{l}\mathbb{R}^n \times T \longrightarrow \mathbb{R}\\ (x,t) \rightsquigarrow G_i(x,t)\end{array}$ est une fonction continue en x et t, pour i=1,...,l

$K_j : \begin{array}{l}\mathbb{R}^n \longrightarrow \mathbb{R}\\ x \rightsquigarrow K_j(x)\end{array}$ est une fonction continue en x, pour j=1,...,m.

L'idée de base des deux algorithmes est la suivante: partant du problème d'optimisation comportant une infinité de contraintes ci-dessus, nous nous ramenons à des problèmes d'optimisation n'en comportant plus qu'un nombre fini, ceux-ci étant résolus par des méthodes connues de la programmation non linéaire finie [6].

Pour simplifier davantage la déscription des algorithmes, posons $G = \text{maximum}(G_1, \ldots, G_l)$; alors le problème ci-dessus peut se récrire :

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimiser}}\ F(x) \\ &\quad \text{s.c. :} \quad G(x,t) \leqslant 0,\ \forall t \in T \\ &\qquad\qquad\quad K_j(x) \leqslant 0, \text{ pour } j=1,\ldots,m \\ &\qquad\qquad\quad x \in X^\circ \end{aligned}$$

Nous avons adopté cette forme d'un P.I.C., en suivant K.R. GEHNER [1], pour faire apparaître la distinction entre des fonctions-contraintes K_j, pour $j=1,\ldots,m$, qui seront satisfaites pour toute solution d'un sous problème fini et des fonctions-contraintes G_i, pour $i=1,\ldots,l$.

III.1. DESCRIPTION DES DEUX ALGORITHMES

Comme annoncé dans le chapitre introductif, les deux algorithmes engenderont une suite, dont chaque élément constitue une solution d'un sous-problème ne comportant plus qu'un nombre fini de contraintes du problème (PA), tel que la suite des sous-problèmes tende en un certain sens vers le problème original et dès lors tel que la suite des solutions correspondant à ces sous-problèmes finis tende vers une solution du problème original (PA).

A. Algorithme I

Nous nous proposons d'exposer d'abord le premier algorithme en terminant par une esquisse de son organigramme.

Le premier pas de l'algorithme I, consiste en une initialisation : nous choisissons un nombre entier, fini, soit q_1, de points appartenant à l'ensemble T du problème original (PA). En rénumérotant ces points de un à q_1, nous obtenons un ensemble : $T^{(1)} = \{ t^q \in T, q=1,\dots,q_1 \}$, qui est inclus dans l'ensemble T. On calcule dès lors une solution du problème :

$$(P_1) \quad \begin{array}{ll} \underset{x \in \mathbb{R}^m}{\text{minimiser}}\ F(x) & \\ \text{s.c. :} & G(x,t) \leqslant 0,\ \forall t \in T^{(1)} \\ & K_j(x) \leqslant 0, \text{ pour } j=1,\dots,m \\ & x \in X^\circ \end{array}$$

Ce faisant, nous obtenons un "sous-problème" de minimisation (P_1), du problème original (PA), ne comportant plus qu'un nombre fini de contraintes. Une solution x^1 de (P_1) peut dès lors être obtenue par des méthodes de la programmation non linéaire finie [6].

Supposons qu'à l'étape k, nous ayons un ensemble fini $T^{(k)}$, formé de q_k points de T. En outre, supposons que x^k soit une solution du sous-problème fini correspondant, c'est à dire que x^k soit tel que :

$$(P_k) \quad \begin{aligned} F(x^k) = \underset{x \in \mathbb{R}^n}{\text{minimum}}\ & F(x) \\ \text{s.c. :}\ & G(x,t) \leqslant 0,\ \forall t \in T^{(k)} = \{\ t^q \in T,\ q=1,\dots,q_k\ \} \\ & K_j(x) \leqslant 0,\ \text{pour } j=1,\dots,m \\ & x \in X^\circ \end{aligned}$$

Le deuxième pas de l'algorithme I teste si x^k, solution de (P_k), est une solution du problème original (PA). Pour cela, il faut que x^k, solution de (P_k), vérifie toutes les contraintes du P.I.C. (PA). Par la structure même du problème original, nous savons que les contraintes $K_j(x) \leqslant 0$, pour $j=1,\dots,m$, se retrouvent dans chaque sous-problème de minimisation. Dès lors il suffit de vérifier que x^k satisfait les contraintes :

$$G(x,t) \leqslant 0 \qquad \forall t \in T$$

Si x^k satisfait toutes ces contraintes, alors x^k, solution du sous-problème fini (P_k), résout le problème original (PA); sinon nous passons au troisième pas.

Au troisième pas, nous passons du sous-problème fini (P_k) à un nouveau sous-problème fini, (P_{k+1}), en lui ajoutant une des contraintes du problème (PA), sélectionnée parmi celles non satisfaites par x^k. Nous cherchons donc t^k appartenant à T, tel que :

$$G(x^k,t^k) > 0 \text{ et}$$
$$G(x^k,t^k) + \varepsilon_k \geq G(x^k,t) \qquad \text{avec } G(x^k,t) > 0,\ \forall t \in T$$
$$\text{où } \varepsilon_k \geq 0 \text{ et } \lim_{k\to\infty} \varepsilon_k = 0$$

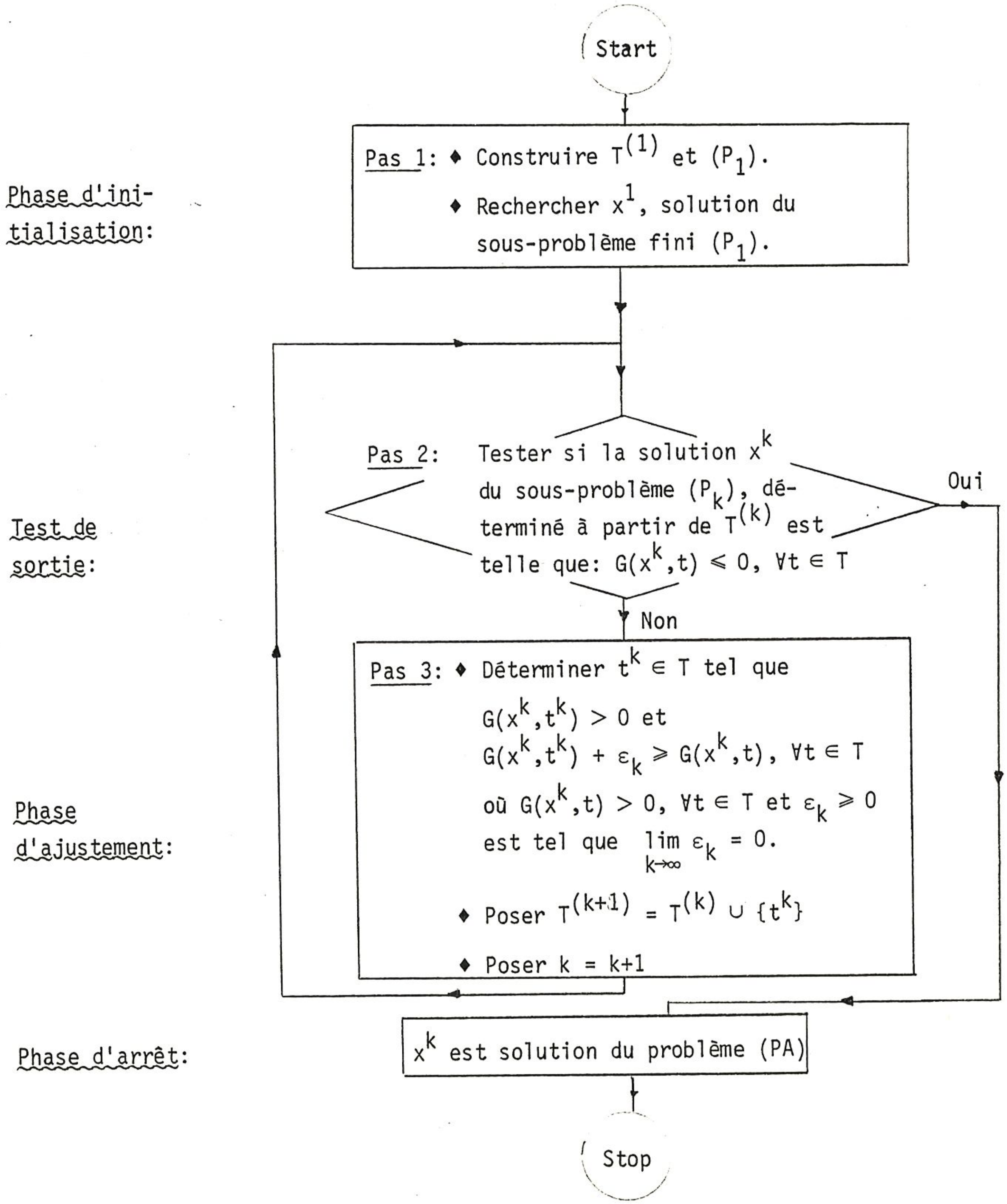
En d'autres mots, on cherche la "pire à ε_k près" des contraintes, soit $G(x,t^k) \leq 0$, parmi l'ensemble des contraintes du problème original (PA), non satisfaites par la solution x^k du sous-problème fini (P_k); la "pire", car elle est, à ε_k près, la contrainte du problème original qui prend la plus grande valeur en x^k. Nous ajoutons à l'ensemble $T^{(k)}$, le point t^k décrit ci-dessus, pour obtenir un nouvel ensemble, noté $T^{(k+1)}$:

$$T^{(k+1)} = T^{(k)} \cup \{t^k\}.$$

Nous retournons ensuite au deuxième pas, en cherchant une solution du nouveau sous-problème (P_{k+1}) déterminé à partir de l'ensemble $T^{(k+1)}$.

En appliquant l'algorithme I au problème original (PA), nous engendrons dès lors une suite $(x^k)_{k \in \mathbb{N}_0}$ de solutions des sous-problèmes (P_k), qui comme nous allons le montrer dans un deuxième paragraphe, tend vers une solution du P.I.C. (PA).

Résumons ce premier algorithme en donnant une esquisse de son organigramme.



B. Algorithme II

Le deuxième algorithme possède la même structure générale que l'algorithme I. Cependant à chacune des itérations, nous omettrons un certain nombre de contraintes qui ne sont pas cruciales dans la détermination du minimum du sous-problème considéré.

L'algorithme II comporte en effet, les trois pas de l'algorithme I et un quatrième pas décrit ci-dessous.

La phase d'ajustement de l'étape k débouche en la construction de l'ensemble $T^{(k+1)}$ et de la recherche de x^{k+1}, solution du sous-problème correspendant (P_{k+1}). Dans un quatrième pas propre à l'algorithme II, nous omettrons certaines contraintes du type $G(x,t^q) \leqslant 0$, pour $q=1,\ldots,q_{k+1}$, ce qui correspond à la suppression de certains éléments t^q de $T^{(k+1)}$, tout en nous assurant que x^{k+1}, solution de (P_{k+1}), reste solution du sous-problème (PD_{k+1}) résultant de cette suppression. Le quatrième pas consiste donc en la recherche d'un sous-ensemble, noté $D^{(k+1)}$, de $T^{(k+1)}$ tel que :

$$
(PD_{k+1}) \qquad
\begin{aligned}
F(x^{k+1}) = \underset{x \in \mathbb{R}^n}{\operatorname{minimum}} F(x) & \\
\text{s.c.} \quad : \quad & G(x,t) \leqslant 0, \quad \forall t \in D^{(k+1)} \subset T^{(k+1)} \\
& K_j(x) \leqslant 0, \text{ pour } j=1,\ldots,m \\
& x \in X^\circ
\end{aligned}
$$

où x^{k+1} est solution du sous-problème fini (P_{k+1}).

C. Quelques remarques

Remarquons que dans le cas particulier, où le sous-problème (P_{k+1}) satisfait une qualification de contraintes de la programmation non linéaire en dimension finie [5] , et si la fonction objectif F (respectivement la fonction G), est différentiable et convexe en x, (respectivement par rapport au premier argument quelque soit t dans $T^{(k+1)}$), alors les théorèmes de KUHN et TUCKER [7] et de CARATHEODORY (1) nous permettent de mettre en évidence un résultat relatif au nombre de contraintes à omettre afin de construire l'ensemble $D^{(k+1)}$.

En effet, si la fonction objectif F et la fonction-contrainte G sont différentiables et convexes, alors par le théorème de Kuhn et Tucker de la programmation non linéaire finie, on sait que x^{k+1} est solution du sous problème (P_{k+1}) si et seulement si :

$$\lambda_o \nabla_x F(x^{k+1}) + \sum_{t \in T^{(k+1)}} \lambda_t \nabla_x G(x^{k+1},t) + \sum_{j=1}^{m} \lambda_j \nabla_x K_j(x^{k+1}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

où $\lambda_o + \sum_{t \in T^{(k+1)}} \lambda_t + \sum_{j=1}^{m} \lambda_j = 1$; $\lambda_o > 0$, $\lambda_t \geqslant 0$, $\forall t \in T^{(k+1)}$ et

$\lambda_j \geqslant 0$ pour $j=1,\dots,m$.

---

(1) Ibid p. 20.

En appliquant le théorème de Carathéodory à cette combinaison convexe du point zéro, nous savons qu'il existe au plus n+1 multiplicateurs non nuls tels que :

$$\mu_0 \nabla_x F(x^{k+1}) + \sum_{t \in T^{(k+1)}} \mu_t \nabla_x G(x^{k+1},t) + \sum_{j=1}^{m} \mu_j K_j(x^{k+1}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

où $\mu_0 + \sum_{t \in T^{(k+1)}} \mu_t + \sum_{j=1}^{m} \mu_j = 1$; $\mu_j \geqslant 0$, pour $j=0,\ldots,m$; $\mu_t \geqslant 0$

$\forall t \in T^{(k+1)}$ et où au plus n+1 de ces multiplicateurs sont non nuls.

En outre, remarquons que le multiplicateur μ_0 de la fonction objectif doit être strictement positif, puisque sinon la deuxième expression entraîne une contradiction de l'hypothèse affirmant que le problème (P_{k+1}) satisfait une qualification de contraintes. Dès lors, il existe au plus n multiplicateurs μ_t, pour $t \in T^{(k+1)}$, non nuls. Nous pouvons donc affirmer que dans ce cas $D^{(k+1)}$ est au plus constitué par n éléments t appartenant à $T^{(k+1)}$.

Remarquons d'autre part, que lors de l'application de l'algorithme I au problème (PA), nous ajoutons à chaque itération une contrainte du problème original au sous-problème (P_k), lorsque la solution x^k de celui-ci ne vérifie pas toutes les contraintes du problème (PA). Cela revient à ajouter un élément t^k de T à l'ensemble $T^{(k)}$. Dès lors, à chaque itération, l'ensemble des contraintes à considérer pour les problèmes finis de minimisation croît, aussi longtemps qu'on n'a pas trouvé une solution du problème (PA), ce qui entraîne l'inclusion suivante :

$$\text{(i)} \qquad T^{(k)} \subsetneqq T^{(k+1)} \qquad \forall k \in \mathbb{N}_0$$

Cette inclusion n'est plus valide pour les ensembles $D^{(k)}$ résultant du quatrième pas de l'algorithme II et on ne peut dès lors expliciter aucune relation entre les ensembles successifs qui sont construits. On sait seulement que :

$$\text{(ii)} \qquad D^k \subset T^{(k)} \qquad \forall k \in \mathbb{N}_0$$

La démonstration des théorèmes de convergence des deux algorithmes nécessitera la connaissance des relations existant entre les domaines admissibles des différents sous-problèmes considérés. A cette fin, introduisons les notations suivantes :

$X = \{ x \in X° \text{ tel que } G(x;t) \leqslant 0,\ \forall t \in T,\ K_j(x) \leqslant 0 \text{ pour } j=1,\dots,m \}$
est l'ensemble des points admissibles du problème original (PA)

$X_{k+1} = \{ x \in X° \text{ tel que } G(x,t) \leqslant 0,\ \forall t \in T^{(k+1)},\ K_j(x) \leqslant 0, \text{ pour } j=1,\dots,$
est l'ensemble des points admissibles du problème (P_{k+1}) et

$Y_{k+1} = \{ x \in X° \text{ tel que } G(x,t) \leqslant 0,\ \forall t \in D^{(k+1)},\ K_j(x) \leqslant 0, \text{ pour } j=1,\dots,$
est l'ensemble des points admissibles du problème (PD_{k+1}).

Dès lors, en se servant des inclusions (i) et (ii) on obtient :

$\forall k \in \mathbb{N}_0$

$$\text{(iii)} \qquad X_{k+1} \subsetneqq X_k \,,$$

$$\text{(iv)} \qquad X_{k+1} \subset Y_{k+1} \,,$$

$$\text{(v)} \qquad X_{k+2} \subset Y_{k+1} \text{ et}$$

$$\text{(vi)} \qquad X \subset X_k \text{ et } X \subset Y_k .$$

III. 2. CONVERGENCE DES ALGORITHMES

Liste des hypothèses

Avant de considérer les théorèmes relatifs à la convergence, nous présentons une liste des hypothèses qui seront utilisées dans la suite.

H_1 : Chaque fonction G_i : $\begin{array}{ccc} \mathbb{R}^n \times T & \longrightarrow & \mathbb{R} \\ (x,t) & & G_i(x,t) \end{array}$ est uniformément Lipschitz-continue en son premier argument; c'est à dire qu'il existe M_i strictement positif tel que :

$$\begin{array}{l} \forall t \in T \\ \forall x^1, x^2 \end{array} \quad : \quad G_i(x^1,t) - G_i(x^2,t) \leq M_i \; \| \, x^1 - x^2 \, \| \qquad \text{pour } i=1,\dots,l$$

H_2 : La fonction objectif F est quasi-convexe (1)

H_3 : Il existe un entier positif L tel que, à chaque itération k, pour k=1,2,..., on ait au plus L points dans l'ensemble $D^{(k+1)}$ de l'algorithme II. (2)

H_4 : Tout sous-problème (P_{k+1}) ou (PD_{k+1}) admet une et une seule solution.

---

(1) Ibid p. 15

(2) Ibid p. 57

L'hypothèse H_1 nous permettra de démontrer la convergence de l'algorithme I, alors que l'algorithme II convergera sous les hypothèses H_1 à H_4.

Dans la suite du chapitre nous convenons de noter $(x^k)_{k \in \mathbb{N}}$, la suite des solutions communes aux sous-problèmes successifs (P_k) ou (PD_k), commune , puisque nous avons exigé que x^k, solution de (P_k), l'est également pour (PD_k).

Remarquons que l'hypothèse H_1 entraîne que la fonction G définie par $G = \text{maximum}\ (G_1,\dots,G_l)$ reste uniformément Lipschitz continue en son premier argument. Montrons en effet, qu'il existe M strictement positif tel que :

$$\begin{array}{l}\forall t \in T \\ \forall x^1, x^2 \in \mathbb{R}^m\end{array} : \quad G(x^1,t) - G(x^2,t) \leqslant M \parallel x^1 - x^2 \parallel \qquad (\alpha)$$

Il suffit pour cela de poser $M = \text{maximum}(M_1,\dots,M_l)$ où M_i est la constante de Lipschitz de la fonction G_i, pour $i=1,\dots,l$.

Dès lors, par la définition de G, il existe i_1 et i_2 avec $1 \leqslant i_1, i_2 \leqslant l$ tels que :

$$\left.\begin{array}{l} G(x^1,t) = G_{i_1}(x^1,t) \quad \text{et} \\ G(x^2,t) = G_{i_2}(x^2,t) \end{array}\right. \qquad (\beta)$$

D'où :

$$\begin{aligned}
G(x^1,t) - G(x^2,t) &= G_{i_1}(x^1,t) - G_{i_2}(x^2,t) \\
&= G_{i_1}(x^1,t) - G_{i_1}(x^2,t) + G_{i_1}(x^2,t) - G_{i_2}(x^2,t) \\
&\leqslant M_{i_1} \| x^1-x^2 \| + G_{i_1}(x^2,t) - G_{i_2}(x^2,t) \text{ par } H_1 \\
&\leqslant M_{i_1} \| x^1-x^2 \| \text{, puisque } G_{i_1}(x^2,t) \leqslant G_{i_2}(x^2,t) \text{ par } (\beta) \\
&\leqslant M \| x^1-x^2 \| \text{, par la définition de } M.
\end{aligned}$$

Avant de passer au théorème relatif à la convergence de l'algorithme II nous exposerons deux lemmes, dont le premier, moyennant une remarque, constituera une démonstration de la convergence de l'algorithme I. Ces deux lemmes, nous permettront d'aborder la convergence du second algorithme.

<u>Lemme 1</u> : Supposons que l'hypothèse H_1 soit satisfaite. Si le point x^* est la limite d'une sous-suite $(x^{k_i})_{i \in \mathbb{N}_0}$ quelconque de $(x^k)_{k \in \mathbb{N}_0}$ telle que x^* soit point d'adhérence de la sous-suite $(x^{k_i+1})_{i \in \mathbb{N}_0}$ de $(x^k)_{k \in \mathbb{N}_0}$, alors il existe une sous-suite convergente $(x^{k_q})_{q \in \mathbb{N}_0}$ telle que

a) $\lim\limits_{q \to \infty} x^{k_q} = x^*$,

b) x^* résout le problème original (PA) et

c) $\lim\limits_{k \to \infty} F(x^k) = F(x^*)$

avec $F(x^k) \leqslant F(x^{k+1}) \leqslant F(x^*)$.

Démonstration :

Puisque x^* est un point d'adhérence de la suite $(x^{k_i+1})_{i \in \mathbb{N}_0}$,
il existe une sous-suite $(x^{k_q+1})_{q \in \mathbb{N}_0}$ de $(x^{k_i+1})_{i \in \mathbb{N}_0}$ telle que $\lim_{q \to \infty} x^{k_q+1} = x^*$.

Posons $\quad F^* = \operatorname*{infimum}_{x \in X} F(x)$

où, rappelons-le, X est l'ensemble des points admissibles du problème original (PA), (1).

♦ Montrons d'abord : $\quad F(x^*) = \lim_{q \to \infty} F(x^{k_q}) \leqslant F^*. \qquad$ (3.1.1.)

En effet, puisque $\lim_{q \to \infty} x^{k_q} = x^*$ et comme F est une fonction continue on a :

$$\lim_{q \to \infty} F(x^{k_q}) = F(x^*).$$

D'autre part : $\quad F(x^{k_q}) \leqslant F(x^*) \qquad \forall q \in \mathbb{N}_0$

puisque $X \subset X_{k_q}$, (1), entraîne que

$$\operatorname*{infimum}_{x \in X_{k_q}} F(x) \leqslant F^*, \qquad \forall q \in \mathbb{N}_0$$

et donc : $F(x^{k_q}) \leqslant F^*$, car x^{k_q} est solution du sous-problème correspondant.

---

(1) Ibid p. 60

♦ Montrons à présent que : $F^* \leqslant F(x^*)$

Il suffit pour cela, de prouver que x^* appartient à X, car F^* étant définie comme $\inf_{x \in X}$ imum F(x), on aura que :

$$F^* \leqslant F(x^*).$$

Montrons donc que x^* satisfait les contraintes du problème original (PA).

a) $G(x^*,t) \leqslant 0,\ \forall t \in T$

En effet : $\forall t \in T$: $G(x^*,t) = G(x^*,t) - G(x^{k_q},t) + G(x^{k_q},t)$

$\leqslant M \| x^* - x^{k_q} \| + G(x^{k_q},t)$ par la relation (α)(1)

$\leqslant M \| x^* - x^{k_q} \| + G(x^{k_q},t^{k_q}) + \varepsilon_{k_q}$

avec $\lim \varepsilon_{k_q} = 0$, par la définition de la "pire à ε_{k_q}" des contraintes au point x^{k_q} (2)

Passons à la limite inférieur sur q dans les deux membres de cette dernière inégalité :

$$\liminf_{q\to\infty} G(x^*,t) \leqslant M \| x^* - \liminf_{q\to\infty} x^{k_q} \|$$

$$+ \liminf_{q\to\infty} G(x^{k_q},t^{k_q}) + \liminf_{q\to\infty} \varepsilon_{k_q}$$

ou $$G(x^*,t) \leqslant M \| x^* - x^* \| + \liminf_{q\to\infty} G(x^{k_q},t^{k_q}) + 0$$

ou $$G(x^*,t) \leqslant \liminf_{q\to\infty} G(x^{k_q},t^{k_q})$$

---

(1) Ibid p. 62

(2) Ibid p. 55

Dès lors, il ne reste plus qu'à montrer que :

$$\liminf_{q\to\infty} G(x^{k_q},t^{k_q}) \leqslant 0$$

Sinon, il existe ε strictement positif tel que $0 < \varepsilon \leqslant G(x^{k_q},t^{k_q})$,

pour q entier "assez grand". Dans ce cas :

$$\begin{aligned}\varepsilon &\leqslant G(x^{k_q},t^{k_q}) \leqslant G(x^{k_q},t^{k_q}) - G(x^{k_q+1},t^{k_q}) + G(x^{k_q+1},t^{k_q})\\ &\leqslant M \| x^{k_q} - x^{k_q+1} \| + G(x^{k_q+1},t^{k_q})\end{aligned}$$

à cause de la Lipschitz-continuïté de G, (1)

$$\leqslant M \| x^{k_q} - x^{k_q+1} \|,$$

puisque $G(x^{k_q+1},t^{k_q}) \leqslant 0$, car x^{k_q+1} étant solution du problème (P_{k_q+1}) (2), la contrainte $G(x^{k_q+1},t)$ est satisfaite au point t^{k_q}.

Or, puisque :

$$\lim_{q\to\infty} x^{k_q} = x^* \quad \text{et}$$

$$\lim_{q\to\infty} x^{k_q+1} = x^*,$$

il existe q_0 entier "assez grand", tel que :

$$\| x^{k_q} - x^{k_q+1} \| < \varepsilon/M \qquad \forall q \geqslant q_0.$$

Et en regroupant ces résultats on trouve $\varepsilon < \varepsilon$, ce qui est absurde.

Dès lors :

$$\liminf_{q\to\infty} G(x^{k_q},t^{k_q}) \leqslant 0$$

---

(1) Ibid p. 62

(2) ou solution de (PD_{k_q+1})

b) Le point x^* vérifie les contraintes $K_j(x) \leqslant 0$ pour $j=1,\ldots,m$

En effet, $K_j(x^{k_q}) \leqslant 0$, où x^{k_q} est solution du sous-problème quelconque (P_{k_q}) (1). D'où puisque $\lim_{q\to\infty} x^{k_q} = x^*$ et que la fonction K_j est

est continue , pour $j=1,\ldots,m$, on déduit que $K_j(x^*) \leqslant 0$, pour $j=1,\ldots,m$.

Dès lors, x^* appartient à X et donc :

$$F^* = \text{infimum } F(x) \leqslant F(x^*). \qquad (3.1.2.)$$

En regroupant les relations (3.1.1.) et (3.1.2.) on prouve que x^* résout le problème original (PA).

Finalement remarquons que, comme x^k est solution du problème (P_k),(2), et puisque X_{k+1} est inclu dans Y_k (3), on a :

$$F(x^k) \leqslant \inf_{x \in Y_k} F(x) \leqslant \inf_{x \in X_{k+1}} F(x) \leqslant F(x^{k+1}) \qquad \forall k \in \mathbb{N}_0$$

En outre, X partie de Y_k, (3), entraîne que :

$$F(x^k) \leqslant F(x^*), \qquad \forall k \in \mathbb{N}_0$$

Dès lors, le caractère monotone de la suite $(F(x^k))_{k \in \mathbb{N}_0}$ et le fait que $\lim_{q\to\infty} F(x^{k_q}) = F(x^*)$ entraînent que :

$$\lim_{q\to\infty} F(x^k) = F(x^*).$$

(1) ou solution de (PD_{k_q}), par la structure même du problème original (PA).

(2) ou solution de (PD_k)

(3) Ibid p. 60

Le fait que la suite $(x^k)_{k \in \mathbb{N}_0}$ admette au moins une valeur d'adhérence vient de l'appartenance de celle-ci à X°. En effet, puisque x^k est solution du sous-problème (P_k) (1), x^k vérifie la contrainte d'appartenance à X°, ensemble compact de $\mathbb{R}^m$ pour k=1,2,... Et par le théorème de Bolzano-Weierstrass, on déduit dès lors que la suite $(x^k)_{k \in \mathbb{N}_0} \subset X° \subset \mathbb{R}^m$ admet au moins une valeur d'adhérence dans X°.

CQFD.

<u>Théorème 1</u> : (Convergence de l'algorithme I)

Supposons que l'hypothèse H_1 soit satisfaite (2).
Alors la suite $(x^k)_{k \in \mathbb{N}_0}$ construite par l'algorithme I, admet au moins un point d'adhérence et tout point d'adhérence résout le problème original (PA).
De plus, observons que :

$$F(x^k) \leqslant F(x^{k+1}) \leqslant F(x^*) \qquad \forall k \in \mathbb{N}_0 \qquad \text{et}$$

$$\lim_{k\to\infty} F(x^k) = F(x^*)$$

<u>Démonstration</u> : En réexaminant la preuve du lemme 1, nous remarquons que l'hypothèse, que la sous-suite $(x^{k_i+1})_{i \in \mathbb{N}_0}$ de $(x^k)_{k \in \mathbb{N}_0}$

---

(1) ou solution de (PD_k)

(2) Ibid p. 61

admette x^* comme valeur d'adhérence sert seulement à démontrer que :

$$\liminf_{q\to\infty} G(x^{k_q}, t^{k_q}) \leqslant 0$$

Cependant pour l'algorithme I, cette assertion est toujours satisfaite, puisque :

$$G(x^k, t) \leqslant 0, \quad \forall t \in T^{(p)}, \quad \forall p \geqslant (k+1)$$

Par la construction même des ensembles $T^{(p)}$, p=1,2,... (1)

Dès lors, s'il existe ε strictement positif tel que :

$$\varepsilon \leqslant G(x^{k_q}, t^{k_q}) \qquad \forall q \in \mathbb{N}_0 \quad ,$$

on a :

$$0 \leqslant \varepsilon \leqslant G(x^{k_q}, t^{k_q}) = G(x^{k_q}, t^{k_q}) - G(x^{k_{q+1}}, t^{k_q}) + G(x^{k_{q+1}}, t^{k_q})$$
$$\leqslant M \| x^{k_q} - x^{k_{q+1}} \| \qquad \forall q \in \mathbb{N}_0 \, ,$$

par la Lipschitz-continuité uniforme de G, et par la construction des sous-problèmes.

D'autre part, puisque : $\lim\limits_{q\to\infty} x^{k_q} = x^*$

la suite $(x^{k_q})_{q \in \mathbb{N}_0}$ est de Cauchy. Il existe donc un rang $Q(\varepsilon)$ tel que, quelque soit q supérieur à $Q(\varepsilon)$, on ait :

$$\| x^{k_q} - x^{k_{q+1}} \| \leqslant M/\varepsilon \text{ , quelque soit } \varepsilon \text{ strictement positif.}$$

---

(1) Ibid p. 55

Dès lors, en groupant ces résultats on obtient $\varepsilon < \varepsilon$, ce qui est absurde. Il faut donc que :

$$\liminf_{q\to\infty} G(x^{k_q}, t^{k_q}) \leq 0.$$

Moyennant cette remarque, la démonstration du lemme précédent constitue également une preuve du théorème de convergence de l'algorithme I.

CQFD.

Lemme 2 : Supposons que les hypothèses H_1 et H_2 soient satisfaites (1). Dans ce cas, si x^* est un point d'adhérence d'une sous-suite convergente $(x^{k_q})_{q \in}$ quelconque de $(x^k)_{k \in}$ et si x^* ne résout pas le problème original (PA); alors tout point d'adhérence $\bar{x}$ de la sous-suite $(x^{k_q+1})_{q \in}$ de $(x^k)_{k \in}$ - et il en existe au moins un - est tel que :

$$\bar{x} \neq x^* \quad \text{et}$$

$$F(\bar{x}) = F(\alpha\bar{x} + (1-\alpha)x^*) = F(x^*) \ , \quad \forall\alpha,\ 0 \leq \alpha \leq 1.$$

Démonstration : Puisque x^* n'est pas solution du problème (PA), x^* ne peut être point d'adhérence de (x^{k_q+1}) par le lemme 1.

---

(1) Ibid p. 61

Cependant, comme par la structure même du problème, tous les éléments de la suite $(x^{k_q+1})_{q \in \mathbb{N}}$, appartienent au compact X°, cette suite y admet au moins une valeur d'adhérence, soit $\overline{x}$, et on a donc que $\overline{x}$ est différent de x^*.

En utilisant les inclusions des ensembles admissibles (1) et comme la solution x^{k_q+1} du sous-problème fini (PD_{k_q+1}) appartient à l'ensemble X_{k_q+1}, lui-même partie de Y_{k_q}, et comme l'ensemble Y_{k_q} est convexe, on a :

$$\alpha x^{k_q} + (1-\alpha)x^{k_q+1} \in Y_{k_q} \qquad \forall\alpha,\ 0 \leqslant \alpha \leqslant 1$$

On en déduit que :

$$F(x^{k_q}) \leqslant F(\alpha x^{k_q} + (1-\alpha)x^{k_q+1}) \qquad \forall\alpha,\ 0 \leqslant \alpha \leqslant 1$$

puisque x^{k_q} est solution du sous-problème fini (PD_{k_q}).

La continuité de F et le passage à la limite dans cette dernière inégalité entraînent :

$$F(x^*) \leqslant F(\alpha x^* + (1-\alpha)\overline{x}) \qquad \forall\alpha,\ 0 \leqslant \alpha \leqslant 1$$

D'autre part, puisque :

$$F(x^{k_q}) \leqslant F(x^k) \leqslant F(x^{k_{q+1}}), \qquad k_q \leqslant k \leqslant k_{q+1} \qquad (2)$$

et

$$\lim_{q\to\infty} F(x^{k_q}) = F(x^*)$$

on a :

$$\lim_{k\to\infty} F(x^k) = F(x^*).$$

(1) Ibid p. 60

(2) A cause du caractère monotone de la suite $(F(x^k))_{k \in \mathbb{N}}$.

Puisque $(x^{k+1})_{k \in \mathbb{N}_0}$ est une sous-suite de $(x^k)_{k \in \mathbb{N}_0}$ admettant le point $\overline{x}$ comme valeur d'adhérence, il existe une sous-suite $(x^{k_q+1})_{q \in \mathbb{N}_0}$ convergent vers $\overline{x}$.

Dès lors, par l'unicité de la limite de la suite $(x^k)_{k \in \mathbb{N}_0}$ et par la continuité de F, on déduit que :

$$F(\overline{x}) = F(x^*).$$

Or F est une fonction quasi-convexe, par l'hypothèse H_2, et dès lors :

$$F(x^*) = F(\overline{x}) = \text{maximum} \ (F(\overline{x}), F(x^*)) \geqslant F(\alpha x^* + (1-\alpha)\overline{x}) \qquad \forall\alpha,\ 0 \leqslant \alpha \leqslant 1$$

On en clonclu, en conbinant ces résultats que :

$$F(x^*) = F(\alpha x^* + (1-\alpha)\overline{x}) = F(\overline{x}) \qquad \forall\alpha,\ 0 \leqslant \alpha \leqslant 1.$$

CQFD.

Sur base de ces deux lemmes, nous pouvons à présent démontrer le théorème de convergence pour l'algorithme II.

Théorème 2 : (Convergence de l'algorithme II)

Admettons que les hypothèses H_1 à H_4 soient satisfaites (1).
Alors tout point d'adhérence de la suite $(x^k)_{k \in \mathbb{N}_0}$ construite par l'algorithme II - et il en existe au moins un - résout le problème original (PA).

En outre, nous avons que :

$$F(x^k) \leqslant F(x^{k+1}) \leqslant F(x^*) \text{ et}$$

$$\lim_{k \to \infty} F(x^k) = F(x^*)$$

où x^* est un point d'adhérence de la suite $(x^k)_{k \in \mathbb{N}_0}$

Démonstration :

Comme la suite $(x^k)_{k \in \mathbb{N}_0}$ engendrée par l'algorithme II est contenue dans le compact $X°$ de $\mathbb{R}^m$, il existe au moins une sous-suite convergente $(x^{k_q})_{q \in \mathbb{N}_0}$ de $(x^k)_{k \in \mathbb{N}_0}$ telle que :

$$\lim_{q \to \infty} x^{k_q} = x^*.$$

Supposons que x^*, point d'adhérence de $(x^k)_{k \in \mathbb{N}_0}$, ne soit pas solution du problème original (PA). Alors tout point d'adhérence $\overline{x}$ de la sous-suite $(x^{k_q+1})_{q \in \mathbb{N}_0}$ de $(x^k)_{k \in \mathbb{N}_0}$ doit vérifier :

$$x^* \neq \overline{x} \quad \text{et}$$

$$F(\overline{x}) = F(\alpha\overline{x} + (1-\alpha)x^*) = F(x^*) \qquad \forall\alpha,\ 0 \leqslant \alpha \leqslant 1$$

sur base du lemme 2. Montrons que l'hypothèse que x^* n'est pas solution du problème (PA), conduit à une contradiction.

(1) Ibid p. 61

A cette fin, analysons la structure asymptotique des problèmes de minimisation finis, en particulier, considérons la relation existant entre les ensembles admissibles X_{k+1} et Y_k et leur forme limite.

Par l'hypothèse H_3, l'ensemble $D^{(k)}$ comprend au plus L éléments, quelque soit k dans $\mathbb{N}_0$. Rangeons les éléments de l'ensemble $D^{(k)}$ dans l'ordre de leur apparition dans l'ensemble $D^{(k)}$:

$$D^{(k)} = \{ t_k^{(1)},\ldots,t_k^{(s_k)} \} \text{ avec } 1 \leqslant s_k \leqslant L.$$

La compacité de T et la possibilité d'extraire successivement des sous-suites convergentes, nous permet d'affirmer qu'il existe une sous-suite $(t_{k_i}^{(s_{k_i})})_{i \in \mathbb{N}_0}$ et au plus L points $d^1,\ldots,d^{L_0}$ dans l'ensemble T, avec L_0 inférieur ou égal à L tels que :

$$\lim_{i\to\infty} t_{k_i}^{(j)} = d^j \quad \text{pour } j=1,\ldots,L_0 \quad \text{et tels que}$$

$$\lim_{i\to\infty} x^{k_i} = x^*, \text{ par ce que nous venons de montrer ci-dessus.}$$

Dès lors, nous avons que :

$$G(x^*,d^j) = G(x^*,d^j) - G(x^*,t_{k_i}^{(j)}) - G(x^{k_i},t_{k_i}^{(j)}) + G(x^{k_i},t_{k_i}^{(j)})$$

$$\leqslant [\, G(x^*,d^j) - G(x^*,t_{k_i}^{(j)}] + M \,\|\, x^* - x^{k_i} \,\|, \text{ par la}$$

Lipschitz-continuité uniforme de G, et puisque $G(x^{k_i},t_{k_i}^{(j)}) \leqslant 0$ pour $j=1,\ldots,s_k$, car x^{k_i} est solution du sous-problème fini de minimisation (PD_{k_i}).

Par la continuité de la fonction G et par le fait que :

$$\lim_{i\to\infty} t_{k_i}^{(j)} = d^j \quad \text{et que}$$

$$\lim_{i\to\infty} x^{k_i} = x^* \quad ,$$

on déduit que :

$$G(x^*,d^j) \leqslant 0 \quad \text{pour } j=1,\dots,L_o.$$

Dès lors,

$$x^* \in Q = \{ x \in X°,\ G(x,d^k) \leqslant 0 \ , \text{ pour } k=1,\dots,L_o \text{ et } K_j(x) \leqslant 0 \text{ pour } j=1,\dots,m \}$$

et Q est un ensemble limite de Y_{k_i} (1).

De plus, remarquons que tout point d'adhérence $\overline{x}$ de la suite $(x^{k_i+1})_{i \in \mathbb{N}_o}$ de $(x^{k_i})_{i \in \mathbb{N}_o}$ doit appartenir à l'ensemble Q, puisque :

$$x^{k_i+1} \in \{ x \in X° \text{ tel que } G(x,t^{k_i}) \leqslant 0 \} \cap Y_{k_i}.$$

Finalement x^* est solution du problème

$$\underset{x \in Q}{\text{minimiser}}\ F(x).$$

En effet, comme x^{k_i} est la solution du sous-problème (PD_{k_i}) on a que :

$$F(x^{k_i}) \leqslant F(x) \quad \forall x \in Y_{k_i},$$

et comme tout point $x° \in Q$ est un point d'adhérence d'une sous-suite dans Y_{k_i}, pour i dans $\mathbb{N}_o$, la continuité de la fonction F entraîne que :

$$F(x^*) \leqslant F(x°) \quad \forall x° \in Q.$$

---

(1) Ibid p. 61.

x^* résout donc le problème de minimisation fini de la fonction F sur l'ensemble Q.

Dès lors, puisque x^* est tel que $F(x^*) = \underset{x \in Q}{\text{minimum}}\ F(x)$ et que $\overline{x}$ appartient à l'ensemble Q, le lemme 2 entraîne que :

$$\overline{x} \neq x^* \text{ et que}$$

$$F(\overline{x}) = F(x^*),$$

sous l'hypothèse que x^* ne résout pas le problème original (PA). Donc $\overline{x}$ résout également le problème de minimisation c'est à dire est tel que :

$$F(\overline{x}) = \underset{x \in Q}{\text{minimum}}\ F(x).$$

Or, ce problème de minimisation a la forme des problèmes de minimisation du quatrième pas du deuxième algorithme (1) et admet deux solutions distinctes $\overline{x}$ et x^*. Ceci est impossible, étant donnée la quatrième hypothèse (2). Dès lors, notre hypothèse de départ, que x^* ne soit pas solution du problème original (PA), est fausse et x^* résout donc le P.I.C. (PA).

CQFD.

Remarquons, qu'après avoir vu la démonstration de ce théorème de convergence pour l'algorithme II, il est clair que l'hypothèse H_4 (2) d'unicité de la solution des sous-problèmes engendrés par l'algorithme II est une condition suffisante pour prouver ce théorème ; mais qu'en fait, il suffirait de faire l'hypothèse que le problème limite $\underset{x \in Q}{\text{minimiser}}\ F(x)$ admet une solution.

---

(1) Ibid p. 57

(2) Ibid p. 61.

De plus, notons qu'aucune hypothèse de convexité n'a été nécessaire pour les fonctions-contraintes des deux algortithmes et que nous avons seulement exigé la quasi-convexité de la fonction objectif.

III.3. VITESSE DE CONVERGENCE DES ALGORITHMES

Nous démontrons d'abord le lemme ci-dessous avant de passer aux résultats relatifs à la vitesse de convergence des deux algorithmes.

Lemme 3 : Considérons une suite $(p_k)_{k \in \mathbb{N}}$ de nombres réels vérifiant :

$p_{k+1} \leqslant p_k - \mu p_k^2$ où $\mu \in \mathbb{R}$

et chaque p_k est strictement positif.

Alors on a :

$$p_k \leqslant \frac{1}{k\mu} .$$

Démonstration : En effet, on a:

$$p_{k+1} \leqslant p_k - \mu p_k^2$$

ou $$p_{k+1} \leqslant p_k(1 - \mu p_k)$$

ou $$\frac{1}{p_{k+1}} \geqslant \frac{1}{p_k} \quad \frac{1}{(1 - \mu p_k)} \text{, puisque } p_k > 0, \forall k \in \mathbb{N} .$$

D'autre part :

$$0 < p_{k+1} \leqslant p_k(1 - \mu p_k)$$

ou $0 < 1 - \mu p_k$, puisque $p_k > 0, \forall k \in \mathbb{N}$

ou $$p_k < \frac{1}{\mu} .$$

Nous pouvons remplacer $\dfrac{1}{(1 - \mu p_k)}$ par la progression géométrique $\sum_{i=0}^{\infty} (\mu p_k)^i$, de raison μp_k telle que $0 < \mu p_k < 1$.

Dès lors,

$$\frac{1}{p_{k+1}} \geqslant \frac{1}{p_k} \left(\sum_{i=0}^{\infty} (\mu p_k)^i \right) \geqslant \frac{1}{p_k} (1 + \mu p_k) = \frac{1}{p_k} + \mu .$$

Par récurrence, on déduit :

$$\frac{1}{p_{k+1}} \geqslant \frac{1}{p_o} + (k+1)\,\mu$$

Dès lors :

$$p_k \leqslant \frac{1}{\frac{1}{p_o} + k\mu} \leqslant \frac{1}{k\mu} \text{ , puisque } p_o > 0.$$

CQFD.

Théorème 3 : (Vitesse de convergence)

Supposons que : - nous n'ayons pas de fonctions-contraintes additionnelles K_j à vérifier à chaque itération.

- la fonction objectif F soit une fonction
  - différentiablen fortement convexe, c-à-d qu'il existe une constante γ strictement positive telle que :
    $\forall x^1, x^2 \in \mathbb{R}^n : F(x^1)-F(x^2) \geqslant \nabla_x F(x^2)(x^1-x^2)+\gamma \| x^1-x^2 \|^2$
  - Lipschitz-continue, c-à-d qu'il existe une constante L strictement positive :
    $\forall x^1, x^2 \in \mathbb{R}^n \; F(x^1) - F(x^2) \leqslant L \| x^1 - x^2 \|$
- chaque fonction-contrainte G_i, pour i=1,...,l soit une fonction :
  - uniformément Lipschitz-continue en son premier argument (1)
  - convexe en son premier argument et qu'il existe $\overline{x} \in X^{\circ}$ tel que $G_i(\overline{x},t) < 0$, quelque soit $t \in T$ et
- nous utilisons dans le cadre de l'algorithme II, la "pire contrainte à $\varepsilon_k = 0$ près" c-à-d
  $G(x^k,t^k) \geqslant G(x^k,t) \qquad \forall t \in T.$

Alors il existe des constantes C_1 et C_2 telles que :

$$\forall k \in \mathbb{N}_0 \quad \begin{array}{l} F(x^*) - F(x^k) \leqslant C_1/k \quad \text{et} \\ \| x^* - x^k \| \leqslant C_2/\sqrt{k} \end{array}$$

où x^* est la solution unique du problème (PA).

---

(1) Ibid p. 61 et 62.

<u>Démonstration</u> : Soit la fonction $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ avec $h(x) = \sup_{t \in T} G(x,t)$.

Remarquons que h est convexe sur $\mathbb{R}^n$, puisque $G : \mathbb{R}^n \times T \longrightarrow \mathbb{R}$ est convexe en son premier argument, quelque soit t appartenant à T. En outre, h est continue, car G est continue et dès lors atteint son suprémum, puisque l'ensemble T est compact. En considérant le point $\overline{x}$ appartenant à X°, mentionné dans les hypothèses, nous obtenons que $h(\overline{x})$ est strictement négatif et d'autre part :

$h(x^k) > 0, \quad \forall k \in \mathbb{N}_0$,

où x^k est solution du problème (P_k) , (1).

Dès lors, comme h est continue, le théorème de la valeur intermédiaire assure qu'il existe λ_k avec $0 < \lambda_k < 1$ tel que :

$h(\overline{x} + \lambda_k(x^k - \overline{x})) = 0.$

D'autre part, comme h est convexe, nous avons les majorations :

$$\begin{aligned} 0 = h(\overline{x} + \lambda_k(x^k - \overline{x}) &= h(\lambda_k x^k + (1 - \lambda_k)\overline{x}) \\ &\leqslant \lambda_k\, h(x^k) + (1 - \lambda_k)\, h(\overline{x}) . \end{aligned}$$

Donc :

$$h(x^k) \geqslant \frac{1 - \lambda_k}{\lambda_k} (-h(\overline{x})) \geqslant (1 - \lambda_k)(-h(\overline{x})), \text{ puisque } 0 < \lambda_k < 1.$$

Si nous définissons :

$$\rho(x,Q) = \inf_{y \in Q} \|x - y\| , \qquad \forall x \in \mathbb{R}^n$$

quelque soit Q, ensemble fermé de $\mathbb{R}^n$, alors :

---

(1) ou solution du problème (PD_k).

$$\|x^k - (\overline{x} + \lambda_k(x^k - \overline{x}))\| \geqslant \rho(x^k,X)$$

où, rappelons-le, X est l'ensemble admissible du problème (PA).
Dès lors, puisque la suite $(x^k)_{k \in \mathbb{N}_0}$ est bornée, car contenue dans le compact X°, on a :

$$(1 - \lambda_k) \geqslant \frac{\rho(x^k,X)}{\|x^k-\overline{x}\|} \geqslant c\ \rho(x^k,X) \qquad \text{avec } c > 0 .$$

En regroupant ces résultats, on a :

$$h(x^k) \geqslant (1 - \lambda_k)(-h(\overline{x})) \geqslant c\ \rho(x^k,X)(-h(\overline{x})) = K\ \rho(x^k,X),$$

avec K strictement positif, puisque $(-h(\overline{x}))$ est strictement positif, par hypothèse.
De plus :

$$h(x^k) \leqslant M \| x^{k+1} - x^k \| \text{ où } M = \max(M_1,\dots,M_l)$$

et les M_i , constantes de Lipschitz des fonctions G_i, pour i=1,...,l.

En effet :

$h(x^k) = G(x^k,t^k)$, pour un certain t^k de T, puisque le suprémum de h est atteint, et dès lors :

$G(x^k,t^k) - G(x^{k+1},t^k) \leqslant M \| x^k - x^{k+1} \|$ avec $G(x^{k+1},t^k) \leqslant 0$,

à cause de la définition du troisième pas des algorithmes.

Nous déduisons finalement :

$$K\ \rho(x^k,X) \leqslant h(x^k) \leqslant M \| x^{k+1} - x^k \|$$

ou $\rho(x^k,X) \leqslant (M/K) \| x^{k+1} - x^k \|$ (3.1.)

D'autre part, comme X est un ensemble fermé, il existe un point $\hat{x}^k$ appartenant à X tel que :

$$\|x^k - \hat{x}^k\| = \rho(x^k,X).$$

Dès lors :

$$F(x^*) - F(x^k) = F(x^*) - F(\hat{x}^k) + F(\hat{x}^k) - F(x^k)$$
$$\leqslant F(\hat{x}^k) - F(x^k),$$

puisque par hypothèse x^* est solution du problème original (PA).

Or, nous avons également :

$$F(\hat{x}^k) - F(x^k) \leqslant L \; \| \hat{x}^k - x^k \| \; ,$$

puisque F est une fonction Lipschitz-continue et dès lors :

$$F(x^*) - F(x^k) \leqslant L \; \rho(x^k, X) \qquad (3.2.)$$

En groupant les relations (3.1.) et (3.2.), on trouve que :

$$F(x^*) - F(x^k) \leqslant (ML/K) \; \| x^{k+1} - x^k \| .$$

Si nous définissons :

$$\phi_k = F(x^*) - F(x^k),$$

cette dernière relation s'écrit :

$$\phi_k \leqslant (ML/K) \; \| x^{k+1} - x^k \| \qquad (3.3.)$$

Mais, comme la fonction F est différentiable, fortement convexe,

on a :

$$\phi_k - \phi_{k+1} = F(x^{k+1}) - F(x^k) \geqslant \gamma \; \| x^{k+1} - x^k \|^2 \qquad (3.4.),$$

puisque :

$$\nabla_x F(x^k)(x - x^k) \leqslant 0, \qquad \forall x \in X_k \qquad (1)$$

Dès lors, par les relations (3.3.) et (3.4.) :

$$\phi_{k+1} \leqslant \phi_k - \frac{\gamma K^2}{L^2 M^2} \phi_k^2 \; .$$

---

(1) ou aussi $\forall x \in Y_k$

Si nous appliquons le lemme 3 à cette dernière relation, nous obtenons :

$$F(x^*) - F(x^k) = \phi_k \leqslant \frac{M^2L^2}{\gamma\ K^2}\ \frac{1}{k} \qquad (3.5.).$$

Finalement, par le caractère de convexité forte de la fonction objectif F :

$$F(x^*) - F(x^k) \geqslant \gamma \parallel x^* - x^k \parallel^2$$

et par la relation (3.5.) :

$$\parallel x^* - x^k \parallel \leqslant \frac{L\ M}{\gamma\ K}\ \frac{1}{k}\ .$$

CQFD.

Remarquons, que le caractère d'unicité de la solution du problème original (PA) est dû au caractère de convexité forte de la fonction objectif F.

En effet, supposons que x°, distinct de x^*, soit aussi solution du problème (PA). Comme F est une fonction différentiable, fortement convexe, on a :

$$(3.6.) \qquad F(x°) - F(x^*) \geqslant \nabla_x F(x^*)(x°-x^*) + \gamma \parallel x° - x^* \parallel^2.$$

Or, par hypothèses, x^* et x° résolvent le P.I.C. (PA), d'où :

$$(3.7.) \qquad F(x°) = \underset{x \in X}{\text{minimum}}\ F(x) = F(x^*).$$

D'autre part, puisque x^* est solution du problème (PA), et que :

$$\forall x \in X_k \qquad \nabla_x F(x^k)(x - x^k) \leqslant 0,$$

on déduit que :

$$\nabla_x F(x^*)(x - x^*) \leqslant 0 \qquad \forall x \in X_k,$$

puisque $x^* \in X \subset X_k \qquad \forall k \in \mathbb{N}_0$,(1).

Mais, $x°$ étant également solution de (PA) on a :

$$x° \in X \subset X_k \qquad \forall k \in \mathbb{N}_0 \qquad (1)$$

D'où :

(3.8.) $\qquad \nabla_x F(x^*)(x° - x^*) \leqslant 0.$

En groupant ces réulstats et en les appliquant à la relation (3.6.) on déduit :

$$0 = F(x°) - F(x^*) \geqslant \nabla_x F(x^*)(x° - x^*) + \gamma \| x° - x^* \|^2$$

ou $\qquad 0 \geqslant \gamma \| x° - x^* \|^2 \geqslant 0 \quad \text{avec } \gamma > 0$

ou $\qquad \| x° - x^* \|^2 = 0$

ou $\qquad x° = x^*.$

CQFD.

---

(1) ou $X \subset Y_k \quad \forall k \in \mathbb{N}_0$

CONCLUSIONS ET EXTENSIONS

Esquissons ce qu'ont été les grandes lignes de ce mémoire. Le point de départ a été le problème de Fritz John auquel nous avons ajouté une infinité de contraintes du type égalité. Pour le problème obtenu, nous avons formulé des conditions nécessaires et suffisantes d'optimalité, pour appliquer ensuite ces critères à des problèmes de l'approximation au sens de Chebycheff. Dans une ultime étape, nous avons exposé deux algorithmes permettant la résolution du problème d'optimisation comportant une infinité de contraintes. Bien que ces algorithmes, étant donné leur formulation dans un cadre fort général, s'appliquent au vaste domaine des problèmes de la programmation non linéaire, il ne faut pas perdre de vue que leur application effective bute sur un certain nombre de détails techniques. Citons par exemple la recherche de la "pire contrainte à ε_k près".

Enfin, il faudrait passer à la programmation de ces algorithmes dans un cadre évidemment beaucoup plus restrictif que celui adopté ici et comparer leurs résultats avec ceux d'autres algorithmes quant à la convergence et à la vitesse de convergence.

La recherche de méthodes d'accélération de la convergence et la vérification qu'une solution d'un des sous-problèmes est également solution du problème original de ces algorithmes, constituent une extension possible de ce mémoire.

Une dernière étape à franchir serait l'application des deux algorithmes aux problèmes de l'approximation au sens de Chebycheff et l'identification de ceux-ci avec les algorithmes de Remes [9].

REFERENCES BIBLIOGRAPHIQUES

[1] K.R. GEHNER (1971)

Optimization problems with an infinite number of constraints and application to constrained approximation problems.

Ph. D. thesis, University of Wisconsin, Madison, WI.

[2] F. JOHN (1948)

Extremum problems with inequalities as subsidiary conditions.

dans "Studies et Essays, Courant Anniversary Volume"
K.O. Friedricks, O.E. Neugebauer et J.J. Stocker, Interscience
New York, pp. 187-204.

[3] KUHN and TUCKER (1950)

Non-linear Programming.

in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (Ed. J. Neyman)
University California Press, Berkeley, California, pp. 481-492.

[4] K.R. GEHNER (1974)

Necessary and sufficient optimality conditions for the Fritz John problem with linear equality constraints.

SIAM J. Control 12, pp. 140-149.

[5] O.L. MANGASARIAN (1969)
Non-linear Programming.
Mc Graw-Hill, New York

[6] B. MARTOS (1975)
Non-linear Programming - Theory and Methods.
Akademiai Kiado, Budapest et North Holland Publishing Company, Amsterdam.

[7] M.R. HESTENES (1975)
Optimization Theory.
John Wiley & Sons, Inc. , New York.

[8] T. BONNESEN & W. FENCHEL (1974)
Theorie der konvexen Körper.
Springer Verlag, Berlin - Heidelberg - New York.

[9] E. W. CHENEY (1966)
Introduction to Approximation Theory.
Mc Graw-Hill, New York.

[10] G.G. LORENTZ (1966)
Approximation of Functions.
Holt, Rinehart and Winston, Inc., New York.

[11] E.W. CHENEY and A.A. GOLDSTEIN (1959)

Newton's Method for convex Programming and Chebycheff Approximation.

Numer. Math. 1, pp 253-268.

TABLE DES MATIERES