50376
1978
137

N° d'ordre : 734

50376
1978
137

THESE

présentée à

UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour l'obtention du titre de

DOCTEUR TROISIEME CYCLE

(Mathématiques Appliquées)

par

Michel VANBREUGEL

UN ALGORITHME GENERAL POUR L'OPTIMISATION NON LINEAIRE

BIBLIOTHEQUE UNIVERSITAIRE
Section de SCIENCES
LILLE

030 027785 3

Soutenue le 29 novembre 1978 devant la Commission d'Examen :

Monsieur P. POUZET	Président
Monsieur C. BREZINSKI	Examinateur
Monsieur J.C. FIOROT	Examinateur
Monsieur P. HUARD	Rapporteur

UNIVERSITE DES SCIENCES
ET TECHNIQUES DE LILLE

01.02.1976

DOYENS HONORAIRES de l'Ancienne Faculté des Sciences

MM. R. DEFRETIN, H. LEFEBVRE, M. PARREAU.

PROFESSEURS HONORAIRES des Anciennes Facultés de Droit et Sciences Economiques, des Sciences et des Lettres

M. ARNOULT, Mme BEAUJEU, MM. BROCHARD, CHAPPELON, CHAUDRON, CORDONNIER, CORSIN, DEHEUVELS, DEHORS, DION, FAUVEL, FLEURY, P. GERMAIN, HEIM DE BALSAC, HOCQUETTE, KAMPE DE FERIET, KOUGANOFF, LAMOTTE, LASSERRE, LELONG, Mme LELONG, MM. LHOMME, LIEBAERT, MARTINOT-LAGARDE, MAZET, MICHEL, NORMANT, PEREZ, ROIG, ROSEAU, ROUBINE, ROUELLE, SAVART, WATERLOT, WIEMAN, ZAMANSKI.

PRESIDENTS HONORAIRES DE L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

MM. R. DEFRETIN, M. PARREAU.

PRESIDENT DE L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

M. J. LOMBARD.

PROFESSEURS TITULAIRES

M.	BACCHUS Pierre	Astronomie
M.	BEAUFILS Jean-Pierre	Chimie Physique
M.	BECART Maurice	Physique Atomique et Moléculaire
M.	BILLARD Jean	Physique du Solide
M.	BIAYS Pierre	Géographie
M.	BONNEMAN Pierre	Chimie Appliquée
M.	BONNOT Ernest	Biologie Végétale
M.	BONTE Antoine	Géologie Appliquée
M.	BOUGHON Pierre	Algèbre
M.	BOURIQUET Robert	Biologie Végétale
M.	CELET Paul	Géologie Générale
M.	CONSTANT Eugène	Electronique
M.	DECUYPER Marcel	Géométrie
M.	DELATTRE Charles	Géologie Générale
M.	DELHAYE Michel	Chimie Physique
M.	DERCOURT Michel	Géologie Générale
M.	DURCHON Maurice	Biologie Expérimentale
M.	FAURE Robert	Mécanique
M.	FOURET René	Physique du Solide
M.	GABILLARD Robert	Electronique
M.	GLACET Charles	Chimie Organique
M.	GONTIER Gérard	Mécanique
M.	GRUSON Laurent	Algèbre
M.	GUILLAUME Jean	Microbiologie
M.	HEUBEL Joseph	Chimie Minérale
M.	LABLACHE-COMBIER Alain	Chimie Organique
M.	LANSRAUX Guy	Physique Atomique et Moléculaire
M.	LAVEINE Jean-Pierre	Paléontologie
M.	LEBRUN André	Electronique
M.	LEHMANN Daniel	Géométrie

.../...

Mme	LENOBLE Jacqueline	Physique Atomique et Moléculaire
M.	LINDER Robert	Biologie et Physiologie Végétales
M.	LOMBARD Jacques	Sociologie
M.	LOUCHEUX Claude	Chimie Physique
M.	LUCQUIN Michel	Chimie Physique
M.	MAILLET Pierre	Sciences Economiques
M.	MONTARIOL Frédéric	Chimie Appliquée
M.	MONTREUIL Jean	Biochimie
M.	PARREAU Michel	Analyse
M.	POUZET Pierre	Analyse Numérique
M.	PROUVOST Jean	Minéralogie
M.	SALMER Georges	Electronique
M.	SCHILTZ René	Physique Atomique et Moléculaire
Mme	SCHWARTZ Marie-Hélène	Géométrie
M.	SEGUIER Guy	Electrotechnique
M.	TILLIEU Jacques	Physique Théorique
M.	TRIDOT Gabriel	Chimie Appliquée
M.	VIDAL Pierre	Automatique
M.	VIVIER Emile	Biologie Cellulaire
M.	WERTHEIMER Raymond	Physique Atomique et Moléculaire
M.	ZEYTOUNIAN Radyadour	Mécanique

PROFESSEURS SANS CHAIRE

M.	BELLET Jean	Physique Atomique et Moléculaire
M.	BODARD Marcel	Biologie Végétale
M.	BOILLET Pierre	Physique Atomique et Moléculaire
M.	BOILLY Bénoni	Biologie Animale
M.	BRIDOUX Michel	Chimie Physique
M.	CAPURON Alfred	Biologie Animale
M.	CORTOIS Jean	Physique Nucléaire et Corpusculaire
M.	DEBOURSE Jean-Pierre	Gestion des entreprises
M.	DEPREZ Gilbert	Physique Théorique
M.	DEVRAINNE Pierre	Chimie Minérale
M.	GOUDMAND Pierre	Chimie Physique
M.	GUILBAULT Pierre	Physiologie Animale
M.	LACOSTE Louis	Biologie Végétale
Mme	LEHMANN Josiane	Analyse
M.	LENTACKER Firmin	Géographie
M.	LOUAGE Francis	Electronique
Mlle	MARQUET Simone	Probabilités
M.	MIGEON Michel	Chimie Physique
M.	MONTEL Marc	Physique du Solide
M.	PANET Marius	Electrotechnique
M.	RACZY Ladislas	Electronique
M.	ROUSSEAU Jean-Paul	Physiologie Animale
M.	SLIWA Henri	Chimie Organique

MAITRES DE CONFERENCES (et chargés d'Enseignement)

M.	ADAM Michel	Sciences Economiques
M.	ANTOINE Philippe	Analyse
M.	BART André	Biologie Animale
M.	BEGUIN Paul	Mécanique
M.	BKOUCHE Rudolphe	Algèbre
M.	BONNELLE Jean-Pierre	Chimie
M.	BONNEMAIN Jean-Louis	Biologie Végétale
M.	BOSCQ Denis	Probabilités
M.	BREZINSKI Claude	Analyse Numérique
M.	BRUYELLE Pierre	Géographie

.../...

M. CARREZ Christian	Informatique
M. CORDONNIER Vincent	Informatique
M. COQUERY Jean-Marie	Psycho-Physiologie
Mle DACHARRY Monique	Géographie
M. DEBENEST Jean	Sciences Economiques
M. DEBRABANT Pierre	Géologie Appliquée
M. DE PARIS Jean-Claude	Mathématiques
M. DHAINAUT André	Biologie Animale
M. DELAUNAY Jean-Claude	Sciences Economiques
M. DERIEUX Jean-Claude	Microbiologie
M. DOUKHAN Jean-Claude	Physique du Solide
M. DUBOIS Henri	Physique
M. DYMENT Arthur	Mécanique
M. ESCAIG Bertrand	Physique du Solide
Me EVRARD Micheline	Chimie Appliquée
M. FONTAINE Jacques-Marie	Electronique
M. FOURNET Bernard	Biochimie
M. FORELICH Daniel	Chimie Physique
M. GAMBLIN André	Géographie
M. GOBLOT Rémi	Algèbre
M. GOSSELIN Gabriel	Sociologie
M. GRANELLE Jean-Jacques	Sciences Economiques
M. GUILLAUME Henri	Sciences Economiques
M. HECTOR Joseph	Géométrie
M. JACOB Gérard	Informatique
M. JOURNEL Gérard	Physique Atomique et Moléculaire
Mle KOSMAN Yvette	Géométrie
M. KREMBEL Jean	Biochimie
M. LAURENT François	Automatique
Mle LEGRAND Denise	Algèbre
Mle LEGRAND Solange	Algèbre
M. LEROY Jean-Marie	Chimie Appliquée
M. LEROY Yves	Electronique
M. LHENAFF René	Géographie
M. LOCQUENEUX Robert	Physique Théorique
M. LOUCHET Pierre	Sciences de l'Education
M. MACKE Bruno	Physique
M. MAHIEU Jean-Marie	Physique Atomique et Moléculaire
Me N'GUYEN VAN CHI Régine	Géographie
M. MAIZIERES Christian	Automatique
M. MALAUSSENA Jean-Louis	Sciences Economiques
M. MESSELYN Jean	Physique Atomique et Moléculaire
M. MONTUELLE Bernard	Biologique Appliquée
M. NICOLE Jacques	Chimie Appliquée
M. PAQUET Jacques	Géologie Générale
M. PARSY Fernand	Mécanique
M. PECQUE Marcel	Chimie Physique
M. PERROT Pierre	Chimie Appliquée
M. PERTUZON Emile	Physiologie Animale
M. PONSOLLE Louis	Chimie Physique
M. POVY Lucien	Automatique
M. RICHARD Alain	Biologie
M. ROGALSKI Marc	Analyse
M. ROY Jean-Claude	Psycho-Physiologie
M. SIMON Michel	Sociologie
M. SOMME Jean	Géographie
Mle SPIK Geneviève	Biochimie
M. STANKIEWICZ François	Sciences Economiques
M. STEEN Jean-Pierre	Informatique

.../...

M.	THERY Pierre	Electronique
M.	TOULOTTE Jean-Marc	Automatique
M.	TREANTON Jean-René	Sociologie
M.	VANDORPE Bernard	Chimie Minérale
M.	VILLETTE Michel	Mécanique
M.	WALLART Francis	Chimie
M.	WERNIER Georges	Informatique
M.	WATERLOT Michel	Géologie Générale
Mme	ZINN-JUSTIN Nicole	Algèbre

Monsieur le Professeur POUZET a bien voulu me faire l'honneur de présider le Jury ; je le prie de croire à ma profonde gratitude.

Je sais gré à Monsieur le Professeur HUARD de m'avoir donné l'idée de ce travail et d'en avoir constamment suivi l'évolution. Je suis heureux de pouvoir lui exprimer ici toute ma reconnaissance pour les nombreux conseils et encouragements qu'il n'a cessé de me prodiguer.

Que Monsieur le Professeur BREZINSKI, dont j'ai suivi avec intérêt le Séminaire d'Analyse Numérique, veuille trouver ici l'expression de ma reconnaissance.

Je remercie Monsieur J.C FIOROT qui a bien voulu s'intéresser à ce travail et le juger.

Je remercie également les membres de l'Equipe d'Optimisation du Laboratoire de Mathématiques Appliquées et tout particulièrement J. DENEL, pour l'aide qu'ils m'ont apportée tout au long de ce travail.

J'adresse enfin mes remerciements pour la qualité de la frappe et la réalisation matérielle à Madame TAILLY et à Monsieur et Madame DEBOCK dont j'ai pu apprécier la compétence et la gentillesse.

PLAN

INTRODUCTION

NOTATIONS ET DEFINITIONS

CHAPITRE I : LES MÉTHODES DE GRADIENTS

I.1 Les principes et l'optimisation sans contraintes

I.1.1 direction améliorante p 1
I.1.2 principe des méthodes p 3
I.1.3 contrôle du pas p 4
I.1.4 contrôle de la direction p 8
I.1.5 algorithme général p 14

I.2 L'adaptation et l'optimisation avec contraintes

I.2.1 direction réalisable p 16
I.2.2 algorithmes de direction réalisable améliorante p 21
I.2.3 exemples d'algorithmes p 32

CHAPITRE II : UN ALGORITHME DE POINTS RÉALISABLES POUR L'OPTIMISATION AVEC CONTRAINTES NON LINÉAIRES

II.1 Problèmes et hypothèses p 41
II.2 Recherche de direction améliorante p 43
II.3 Recherche de direction améliorante réalisable p 45
II.4 Algorithme général p 54
II.5 Sous-optimisation finie p 59
II.6 Aplications p 64

CHAPITRE III : LES MÉTRIQUES VARIABLES ET L'OPTIMISATION AVEC CONTRAINTES

III.1 Les principes des algorithmes de métrique variable p 69
III.2 Résultats de convergence p 70
III.3 L'adaptation des algorithmes de métriques variable à l'optimisation avec contraintes p 72
III.4 L'algorithme de Han p 75
III.5 Comparaison entre l'algorithme de Han et l'algorithme général réalisable du chapitre II p 78

ANNEXE I p 80

ANNEXE II p 83

BIBLIOGRAPHIE

Le but du travail présenté est l'étude et l'extension d'un algorithme dû à P. HUARD [21], algorithme qui permet la maximisation d'une fonction concave différentiable sur un domaine convexe, par une méthode de gradient projeté discrétisé.

L'algorithme défini au chapitre II constitue une généralisation du précédent tout en en conservant les principales caractéristiques.

Méthode de gradient, l'algorithme présenté s'inscrit dans le cadre des algorithmes de direction améliorante ; ceux-ci sont étudiés en I.1 dans le contexte de l'optimisation sans contraintes et en liaison avec les conditions d'optimalité.

"Projeter une direction de gradient" est un moyen classique utilisé dans de nombreux algorithmes pour adapter, à l'optimisation avec contraintes, des algorithmes initialement construits pour l'optimisation sans contraintes. Les problèmes que pose cette adaptation sont étudiés en I.2, à cette occasion est défini un nouvel algorithme (I.2.2) utilisant les notions de direction uniformément améliorante et réalisable.
L'utilisation d'une projection à chaque itération d'un algorithme n'est pratiquement utilisable, que si le domaine sur lequel est effectuée cette projection est un polyèdre. On est donc amené, pour l'optimisation avec contraintes non linéaires, à linéariser celles-ci.
La recherche d'une direction améliorante, dans le domaine linéarisé peut alors se faire grâce à la résolution d'un sous-problème quadratique, plus général que celui de la projection (II.2) et qui ne requiert qu'un nombre fini de calculs.

Le procédé de linéarisation utilisé conduit à un inconvénient majeur, c'est que le point obtenu par projection sur le linéarisé n'est pas en général réalisable. Divers procédés avaient été utilisés pour pallier à cet inconvénient ; ils consistaient à associer, grâce à un "procédé de rentrée", un point réalisable au point non réalisable considéré.
L'ensemble de ces procédés était cependant peu satisfaisant, tant sur le plan théorique que pratique (I.2.3).
La solution apportée à ce problème par P. HUARD dans l'algorithme du gradient projeté discrétisé est à la fois simple à mettre en oeuvre et se justifie théoriquement dans le cadre des fonctions multivoques.
Ce procédé de rentrée est analysé en I.2.2, une variante très simple est mise à jour.

L'algorithme général décrit en II.4 comporte, à chaque itération, les trois étapes suivantes :

- trouver une direction améliorante par la résolution finie d'un sous-problème quadratique à contraintes linéaires (II.2)

- associer à la direction précédente une direction améliorante et réalisable, grâce au procédé de rentrée (II.3)

- utiliser, dans la direction ainsi trouvée, un ensemble discret de points réalisables qui permette la recherche de l'itéré suivant en un nombre fini de calculs (II.5)

Le sous-problème quadratique résolu à chaque itération de l'algorithme général comporte une matrice en paramètre.
Un algorithme décrit en 1976 par HAN [17] utilise le même sous-problème, la matrice paramètre étant engendrée par une formule de métrique variable. L'algorithme de HAN apparait comme une généralisation, au cas de l'optimisation avec contraintes, des algorithmes de métrique variable ; il permet de construire une suite de points qui converge localement vers un optimum avec une vitesse superlinéaire.

Dans quelle mesure l'utilisation de matrices de métrique variable, permet-elle d'obtenir un résultat analogue pour l'algorithme du chapitre II ? Ce problème est abordé au chapitre III, mais il n'a pas été possible pour l'instant de déterminer en quoi les phases de rentrée et de discrétisation perturbent la vitesse superlinéaire de convergence obtenue par HAN.

Ce problème ouvert présente un intérêt certain, car sa résolution permettrait de cumuler les résultats obtenus au chapitre II et ceux de HAN en déterminant un algorithme :

- applicable à l'optimisation avec contraintes non linéaires

- de points réalisables

- tel que chaque itération s'effectue en un nombre fini de calculs

- globalement convergent

- à vitesse superlinéaire de convergence.

*

* *

NOTATIONS ET DEFINITIONS

0 unique notation pour l'élément nul de $\mathbb{R}$ et celui de $\mathbb{R}^n$

$$\mathbb{R}^{n*} = \{x \in \mathbb{R}^n \ / \ x \neq 0\}$$

si $C \subset \mathbb{R}^n$; $\bar{C}$ la fermeture de C et $\mathring{C}$ l'intérieur de C
$C-x = \{y-x \ / \ y \in C\}$ le translaté de l'ensemble C

$x.y$ le produit scalaire de x par y dans $\mathbb{R}^n$

$||x||$ la norme euclidienne de $\mathbb{R}^n$

si f est une fonction définie sur $\mathbb{R}^n$ à valeurs réelles

∇f le gradient, $\nabla^2 f$ le Hessien

x^t, H^t le transposé d'un vecteur x, d'une matrice H

$T(x,C)$ le cône tangent à l'ensemble C au point x

$y \in T(x,C)$ si et seulement si :
$\exists \{(x_k,\lambda_k)\}_{k\in\mathbb{N}} \subset C \times \mathbb{R}^+$ telle que :
$\lim\limits_{k\to\infty} x_k = x$ et $\lim\limits_{k\to\infty} \lambda_k(x_k-x) = y$

$\Pi(T) = \{y \in \mathbb{R}^n / y.x \leq 0, \ \forall x \in T\}$ le cône polaire (négatif) d'un cône T

$X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$; $\Gamma : X \to \mathcal{P}(Y)$ une fonction multivoque telle que $\Gamma(x) \neq \phi$, $\forall x \in X$.

Γ est sup-continue en $x \in X$ si et seulement si :

$$\left.\begin{array}{l} \forall \{x_k\}_{k\in\mathbb{N}} \to x \\ \forall \{y_k\}_{k\in\mathbb{N}} \to y \\ y_k \in \Gamma(x_k), \ \forall k \in \mathbb{N} \end{array}\right\} \Longrightarrow \ y \in \Gamma(x)$$

Γ est inf-continue en $x \in X$ si et seulement si :

$$\left.\begin{array}{l} \forall \{x_k\}_{k\in\mathbb{N}} \to x \\ \forall y \in \Gamma(x) \end{array}\right\} \Longrightarrow \begin{array}{l} \exists \{y_k\}_{k\in\mathbb{N}} \to y \\ \exists \bar{k} \in \mathbb{N} \text{ telle que } \forall k \geq \bar{k} : y_k \in \Gamma(x_k) \end{array}$$

CHAPITRE I

LES MÉTHODES DE GRADIENTS

I - LES PRINCIPES ET L'OPTIMISATION SANS CONTRAINTES

I.1 .1 Direction améliorante

La notion de direction améliorante a été étudiée par différents auteurs sous des formes voisines et des dénominations diverses. Elle fût introduite à l'origine par ZOUTENDIJK [42] dans le cas différentiable. TOPKIS et VEINOT [37], DANIEL [7] puis MOKHTAR-KHARROUBI [27] utilisèrent une définition générale dans le cas d'une fonction continue. La définition donnée ici semble la moins restrictive :
soient $f : \mathbb{R}^n \to \mathbb{R}$; $x \in \mathbb{R}^n$; $s \in \mathbb{R}^n$

I.1.1.a) Définition : s est une direction améliorante pour f en x, si et seulement si :

$$\exists \bar{\lambda} \in \mathbb{R}^{+*},\ \forall \lambda \in]0,\bar{\lambda}] : f(x+\lambda s) > f(x)$$

s sera dite direction améliorante au sens large, descendante, descendante au sens large si l'on remplace dans la définition, le signe > respectivement par ≥ < ≤.
On notera $A_+^*(x)$, $A_+(x)$, $A_-^*(x)$ et $A_-(x)$ les ensembles constitués respectivement par ces directions en un point x.

Remarque 1 : pour tout x, $A_+(x)$ et $A_-(x)$ sont des cônes pointés ; $A_+^*(x)$ et $A_-^*(x)$ sont des cônes épointés (ils ne contiennent pas l'origine 0 de $\mathbb{R}^n$). Ceci permettra de caractériser une direction améliorante par un vecteur s que l'on choisira en général de norme 1.

Remarque 2 : une direction non améliorante n'est pas forcément une direction descendante. Dans le cas d'une fonction f continue aucun critère simple ne peut caractériser une direction améliorante. Par contre si f est différentiable on obtient les résultats suivants :

I.1.1.b) Proposition Soit f différentiable en x, une condition nécessaire pour que s soit une direction améliorante au sens large pour f en x est que :

$$\nabla f(x).s \geq 0$$

Dem : si f est différentiable en x le rapport $\frac{f(x+\lambda s)-f(x)}{\lambda}$ positif pour $\lambda \in]0,\bar{\lambda}]$ admet une limite quand λ tend vers 0 : c'est par définition la dérivée directionnelle de f en x selon la direction s

$$\lim_{\lambda \to 0} \left[\frac{f(x+\lambda s)-f(x)}{\lambda}\right] = f'(x+\lambda s)\Big|_{\lambda=0} = \nabla f(x)s \geq 0$$

Cette condition nécessaire n'est pas suffisante, on obtient seulement :

I.1.1.c) Proposition Soit f différentiable en x, une condition suffisante pour que s soit une direction améliorante pour f en x est que :

$$\nabla f(x).s > 0$$

Dem : on montre que :

$$\left.\begin{array}{r} f \text{ différentiable en } x \\ \\ \nabla f(x).s > 0 \end{array}\right\} \Longrightarrow \begin{array}{l} \exists \bar{\lambda} > 0,\ \forall \lambda \in]0,\bar{\lambda}] \\ f(x+\lambda s) > f(x) \end{array}$$

sinon : $\forall \bar{\lambda} > 0,\ \exists \lambda \in]0,\bar{\lambda}]$ tel que $f(x+\lambda s) \leq f(x)$
On peut alors en choisissant une suite $(\bar{\lambda}_n)_{n\in\mathbb{N}} \to 0$ trouver une suite : $(\lambda_n)_{n\in\mathbb{N}} \to 0$ avec :

$$\lim_{\substack{\lambda \to 0 \\ n}} \frac{f(x+\lambda_n.s)-f(x)}{\lambda_n} = \lim_{\lambda \to 0} \frac{f(x+\lambda s)-f(x)}{\lambda} = \nabla f(x).s \leq 0$$

d'où la contradiction.

Notations : $D_+^*(x) = \{s \in \mathbb{R}^n \ / \ \nabla f(x).s > 0\}$
$D_+(x)$, $D_-^*(x)$ et $D_-(x)$ si on remplace le signe > respectivement par ≥ < ≤.

Remarque 3 : $\mathbb{R}^n = D_+^*(x) \cup D_-(x) = D_-^*(x) \cup D_+(x)$

$D_+^*(x) \cap D_-(x) = D_-^*(x) \cap D_+(x) = \emptyset$

Les deux propositions précédentes se traduisent par : pour tout $x \in \mathbb{R}^n$, f étant différentiable en x :

$$D_+^*(x) \subset A_+^*(x) \subset A_+(x) \subset D_+(x)$$
$$D_-^*(x) \subset A_-^*(x) \subset A_-(x) \subset D_-(x)$$

mais ces inclusions sont strictes en général et l'on ne peut, même dans le cas où f est différentiable, caractériser à l'aide d'une condition nécessaire et suffisante, une direction améliorante en un point. Pratiquement les algorithmes de directions améliorantes utilisent des directions de $D_+^*(x)$; le procédé itératif s'arrêtant dès que cet ensemble est vide en un point $\overset{*}{x}$. Du fait de l'inclusion stricte il se peut alors que $A_+^*(\overset{*}{x})$ soit non vide et $\overset{*}{x}$ n'est pas un maximum local.

I.1.1.d) Proposition Soit f différentiable :

$$(x \text{ maximum local}) \Longrightarrow (A_+^*(x) = \emptyset) \Longrightarrow (D_+^*(x) = \emptyset) \Longleftrightarrow (\nabla f(x) = 0)$$

Dem :

- si $\exists s \in \mathbb{R}^n$ et $\bar{\lambda} > o$ tel que : $\forall \lambda \in]0,\bar{\lambda}]$ $f(x+\lambda s) > f(x)$ alors x n'est pas maximum local pour f.
- $D_+^*(x) \subset A_+^*(x)$ d'où la 2ème implication
- $(D_+^*(x) = \emptyset) \Longleftrightarrow (\forall s \in \mathbb{R}^n, \nabla f(x).s \leq o) \Longleftrightarrow (\nabla f(x) = 0)$ on dira dans ce cas que x est stationnaire.

Remarque 4 : les réciproques des deux implications précédentes sont fausses en général.

Dans ce qui précède dès que $\nabla f(x).s$ est nul on ne peut rien dire quant au "sens de variation" de la fonction f selon la direction s. Ce n'est que sous l'hypothèse de concavité de f que ce sens de variation peut être controlé.

I.1.1.e) Proposition Soit f différentiable en x, concave, une condition nécessaire pour que s soit une direction améliorante pour f en x est que :

$$\nabla f(x).s > 0$$

Dem : soit $s \in A_+^*(x)$

$\exists \lambda \in \mathbb{R}^{+*}$ tel que : $f(x+\lambda s) > f(x)$

par concavité de f

$\forall \theta \in [0,1]$: $f(x+\theta.\lambda s) \geq f(x) + \theta.[f(x+\lambda s)-f(x)]$

au total $\forall \theta \in]0,1]$:

$$\frac{f(x+\theta.\lambda s)-f(x)}{\theta.\lambda} \geq \frac{f(x+\lambda s)-f(x)}{\lambda}$$

et on obtient :

$$\lim_{\theta \to o} \left[\frac{f(x+\theta.\lambda x)-f(x)}{\theta.\lambda}\right] = \nabla f(x).s \geq \frac{f(x+\lambda s)-f(x)}{\lambda} > 0$$

I.1.1.f) Proposition Soit f différentiable en x, concave, une condition suffisante pour que s soit une direction descendante au sens large pour f en x est que :

$$\nabla f(x).s \leq 0$$

<u>Dem</u> : f concave : $\forall\lambda \in \mathbb{R}^+ \; f(x+\lambda s) \leq f(x) + \lambda.\nabla f(x).s$

si $\nabla f(x).s \leq 0$, $\forall\lambda \in \mathbb{R}^+ \; f(x+\lambda s) \leq f(x)$

à fortiori s est une direction descendante au sens large.

Compte tenu des propositions b, c, e, f de cette partie, on obtient dans le cas où f est concave différentiable des caractérisations nécessaires et suffisantes pour les directions améliorantes et descendantes au sens large :

$$D_+^*(x) = A_+^*(x) \subset A_+(x) \subset D_+(x)$$
$$D_-^*(x) \subset A_-^*(x) \subset A_-(x) = D_-(x)$$

On peut en déduire le résultat suivant qui donne des conditions nécessaires d'optimalité :

I.1.1.g) <u>Proposition</u> Soit f différentiable concave

$(\nabla f(x)=0) \Longleftrightarrow (D_+^*(x)=\emptyset) \Longleftrightarrow (A_+^*(x)=\emptyset) \Longrightarrow$ (x maximum)

<u>Dem</u> : seule la dernière implication est à démontrer, on montre la contraposée :

si x n'est pas maximum : $\exists y$ tel que $f(y) > f(x)$ on pose $y-x = s$ on montre que : $s \in A_+^*(x)$

f est concave : $\forall\theta \in [0,1]$

$$f(x+\theta.s) \geq f(x) + \theta.[f(x+s)-f(x)]$$

$\forall\theta \in]0,1]$: $f(x+\theta s)-f(x) \geq \theta.[f(x+s) - f(x)] > 0$

traduit que s est direction améliorante : $s \in A_+^*(x) \neq \emptyset$

<u>Remarque 4</u> : si f est différentiable concave, on a pour tout x de $\mathbb{R}^n$: $\mathbb{R}^n = A_+^*(x) \cup A_-(x)$ avec $A_+^*(x) \cap A_-(x) = \phi$ toute direction est soit améliorante soit descendante au sens large.

I.1.1.h) <u>Proposition</u> Conditions nécessaires et suffisantes d'optimalité si f est concave différentiable :

$(\nabla f(x)=0) \Longleftrightarrow (D_+^*(x) = \emptyset) \Longleftrightarrow (A_+^*(x) = \emptyset) \Longleftrightarrow (A_-(x) = \mathbb{R}^n) \Longleftrightarrow (D_-(x) = \mathbb{R}^n)$

I.1.2 Principe des méthodes

De nombreux algorithmes existants pour résoudre le problème de la maximisation d'une fonction continue de $\mathbb{R}^n$ dans $\mathbb{R}$, utilisent le principe suivant :
étant donné un point x quelconque

- trouver une direction améliorante s pour f en x
- trouver selon cette direction un point qui "améliore" la valeur de f
  soit trouver $\lambda \in \mathbb{R}^{+*}$ tel que : $f(x+\lambda s) > f(x)$

Ces algorithmes sont appelés méthodes de pente ou de descente en minimisation, méthodes de direction améliorante en maximisation ou plus généralement méthodes de gradient.

Si l'on utilise itérativement ce procédé pour construire une suite de points, tels que :

$$\forall k \in \mathbb{N} : x_{k+1} = x_k + \alpha_k.s_k$$

alors la suite des valeurs $(f(x_k))_{k\in\mathbb{N}}$ est croissante. Cependant le procédé choisi reste trop général pour assurer que la suite $(x_k)_{k\in\mathbb{N}}$ construite puisse converger vers un maximum ou même vers un point stationnaire. Ceci ne sera possible qu'avec les précautions suivantes :

- contrôler la direction améliorante s_k
- contrôler le pas α_k

Si f est différentiable les démonstrations qui permettent d'assurer la convergence des algorithmes se feront toujours en deux étapes :

- montrer que si $\lim_{k\to\infty} [f(x_{k+1})-f(x_k)] = 0$ alors $\lim_{k\to\infty} [\nabla f(x_k).s_k] = 0$
- montrer ensuite que si $\lim_{k\to\infty} [\nabla f(x_k).s_k] = 0$ alors $\lim_{k\to\infty} [\|\nabla f(x_k)\|] = 0$

si l'on considère

$$f(x_{k+1}) - f(x_k) = \alpha_k.\nabla f(x_k).s_k + \dots$$

on s'apperçoit que la 1° implication est directement liée au choix de α_k : choix du pas à chaque itération. Le problème sera de choisir α_k "judicieusement" de façon à ce que, l'amélioration $f(x_{k+1})-f(x_k)$ obtenue à chaque étape, ne tende vers 0 que lorsque la suite converge vers un point stationnaire.
La 2ème implication ne dépend que du choix de la direction améliorante, elle sera obtenue dès que l'on s'assure que la direction s_k ne peut tendre vers une position orthogonale au gradient.

I.1.3 Contrôle du pas

La notion de fonction forçante introduite par CEA [4] a permis d'étudier d'une manière générale les conditions à introduire pour la détermination du pas dans les méthodes de directions améliorantes.

I.1.3.1 <u>Définition</u> : une fonction forçante G est une fonction de $\mathbb{R}^+$ dans $\mathbb{R}^+$ telle que pour toute suite $(t_k)_{k\in\mathbb{N}}$ de $\mathbb{R}^+$ on ait :

$$[\lim_{k\to\infty} (G(t_k)) = 0] \Longrightarrow [\lim_{k\to\infty} (t_k) = 0]$$

<u>Exemple</u> : si $\alpha \in \mathbb{R}^+$ alors $G(t) = \alpha.t$ est une fonction forçante.

I.1.3.2 <u>Propriété</u> Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable bornée supérieurement, G une fonction forçante et $(x_k)_{k\in\mathbb{N}}$ une suite telle que : $\forall k \in \mathbb{N}$

$$f(x_{k+1})-f(x_k) \geq G(s_k.\nabla f(x_k))$$

alors : $\lim_{k\to\infty} [s_k.\nabla f(x_k)] = 0$

<u>Dem</u> : f étant bornée supérieurement et $f(x_{k+1}) \geq f(x_k)$ la suite $(f\ x_{k+1})-f(x_k))_{k\in\mathbb{N}}$ tend vers 0. Par définition même de la fonction forçante on trouve le résultat voulu.

Dans ce cadre on peut trouver une étude complète des différentes méthodes existantes dans DANIEL [7] et SCHWETLICK [36]. Il est ainsi montré que le résultat $\lim_{k\to\infty} (\nabla f(x_k).s_k) = 0$ est obtenu pour de nombreuses méthodes classiques, parmi lesquelles celles de CURRY [6], GOLDSTEIN [16] ainsi que divers procédés de discrétisation.
On va ici étudier deux choix particuliers du pas, on verra dans la démonstration de l'algorithme général en I.1.5 que ceux-ci permettent d'obtenir le résultat suffisant :

$$\lim_{k\to\infty} [\nabla f(x_k).s_k] \leq 0$$

I.1.3.3 La maximisation approchée

Problème Pour $\varepsilon > 0$ fixé, trouver $\bar{x} \in D(x,s) = \{x + \lambda s / \lambda \in \mathbb{R}^+\}$
tel que : $f(\bar{x}) \geq f(\hat{x}) - \varepsilon$ où $\hat{x}$ est un maximum exact de f sur $D(x,s)$.

Ce problème n'est pas plus simple que celui de la maximisation exacte puisque $f(\hat{x})$ n'est pas connu. De plus il est nécessaire pour assurer la convergence que l'erreur commise ε tende vers 0.
Cependant si l'on applique les méthodes usuelles de recherche du maximum selon une direction (méthodes dichotomiques : FIBONACCI, Section Dorée ...) on peut obtenir en un nombre fini de tests un intervalle d'incertitude : c'est à dire exhiber un intervalle $[x_1, x_2]$ qui contienne un maximum.
De plus la longueur de cet intervalle d'incertitude peut être rendue (au moins théoriquement) aussi petite que voulue.

Sous certaines hypothèses sur la fonction f à maximiser, on pourra de plus donner une majoration de l'erreur :

$$0 \leq f(\hat{x}) - f(\bar{x}) \leq \varepsilon$$

Pour ε fixé, aussi petit soit-il, on peut trouver à l'aide d'une recherche unidimensionnelle un tel $\bar{x}$ en un nombre fini de calculs. A chaque itération d'un algorithme de direction améliorante on effectuera cette recherche en choisissant des valeurs de ε qui tendent vers 0.
Un tel contrôle est donné par ARMIJO [1] sous des hypothèses fortes sur f. On va donner une majoration de l'erreur commise quand f est concave continument différentiable.

I.1.3.3a) Proposition Soit f concave continument différentiable, $[x_1, x_2]$ un intervalle d'incertitude, $\hat{x}$ un maximum de f sur la direction $(x_2 - x_1)$
$f^+ = \text{Max}\{f(x_1), f(x_2)\}$, alors : il est possible de majorer l'erreur $f(\hat{x}) - f^+$ par une quantité qui tend vers 0 avec la longueur de l'intervalle d'incertitude.

Dem : $\phi(\lambda) = f[x_1 + \lambda(x_2 - x_1)]$ est une fonction concave, continument différentiable sur $[0,1]$
$\phi'(0) = (x_2 - x_1).\nabla f(x_1)$ et $\phi'(1) = (x_2 - x_1).\nabla f(x_2)$ si $\nabla f(x_1)(x_2 - x_1)$ ou $\nabla f(x_2)(x_2 - x_1)$ sont nuls alors l'un des deux points est un maximum et l'erreur est nulle, sinon :
$\hat{x} \in]x_1, x_2[$, les directions $(\hat{x} - x_1)$ et $(\hat{x} - x_2)$ sont améliorantes pour f en x_1 et x_2 respectivement ; par I.1.1.a)

$$\phi'(0) > 0 \text{ et } \phi'(1) < 0$$

on note : $\phi'(0) = g_1 > 0$ et $-\phi'(1) = g_2 > 0$; $\phi(0) = f(x_1) = f_1$ et $\phi(1) = f(x_2) = f_2$
par concavité de ϕ : $\forall \lambda \in [0,1]$
$\phi(\lambda) \leq g_1.\lambda + f_1$ et $\phi(\lambda) \leq g_2(1-\lambda) + f_2$
on en déduit :

$$\underset{\lambda \in [0,1]}{\text{Max}}\ \phi(\lambda) \leq \underset{\lambda \in [0,1]}{\text{Max}}\ \{\text{Inf}(g_1\lambda + f_1 \ ; \ g_2(1-\lambda) + f_2)\}$$

soit :

$$f(\hat{x}) \leq \frac{g_1 f_2 + g_2 f_1 + g_1 g_2}{g_1 + g_2} = m$$

$$m = f_2 + \frac{g_1 g_2 - g_2.(f_2 - f_1)}{g_1 + g_2} = f_1 + \frac{g_1 g_2 - g_1(f_1 - f_2)}{g_1 + g_2}$$

si on note par + celui des deux indices tel que
$f(x_1) = \text{Max}\{f(x_1), f(x_2)\} = f^+$
on note - l'aute indice, on obtient :

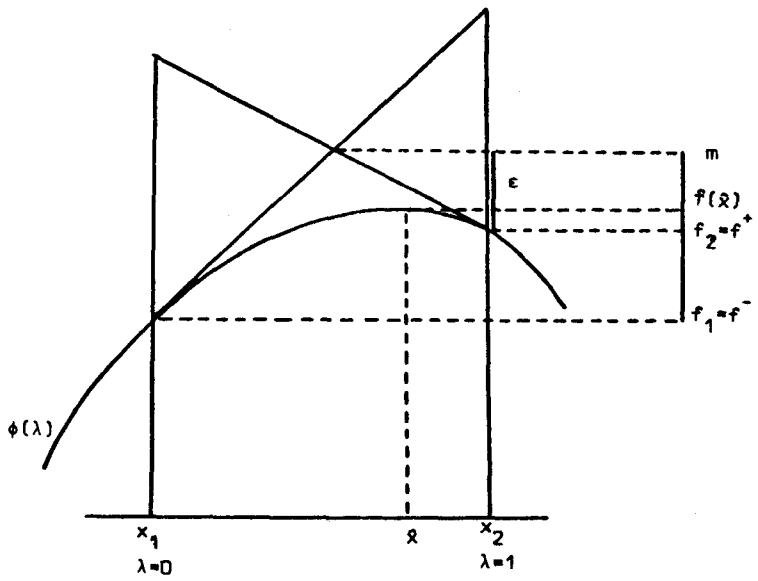
$$m = f^+ + \frac{g^- g^+ - g^+ . (f^+ - f^-)}{g^+ + g^-}$$

$$f(\hat{x}) - f^+ \leq m - f^+ = \frac{g^- g^+ - g^+ . (f^+ - f^-)}{g^+ + g^-} = \varepsilon$$

on dispose ainsi d'une majoration ε de l'erreur commise quand on remplace la valeur $f(\hat{x})$ du maximum exact par le maximum f^+ des valeurs obtenues en x_1 et x_2

$$\varepsilon = \frac{g^- + f^- - f^+}{1 + \frac{g^-}{g^+}}$$

quand la longueur de l'intervalle d'incertitude tend vers 0 ; x_1 et x_2 tendent tous deux vers $\hat{x}$
par continuité de f : $(f^- - f^+)$ tend vers 0
par continuité de ∇f : g^- tend vers 0
$\frac{g^-}{g^+}$ étant positif le dénominateur reste ≥ 1 donc ε tend vers 0 quand la longueur de $[x_1, x_2]$ tend vers 0. C.Q.F.D.



I.1.3.4 <u>La discrétisation</u> HUARD [21]

Les algorithmes de maximisation unidimensionnelle ne peuvent pratiquement chercher le maximum que parmi un ensemble discret de points.
On va ici choisir dans une direction un ensemble discret défini une fois pour toutes, le pas consistant à chaque itération à prendre un maximum de f parmi l'ensemble discret défini.
Soit $(\alpha_i)_{i \in \mathbb{N}}$ une suite de $\mathbb{R}^+$ strictement décroissante qui converge vers 0, on lui associe :

$$D_d(x,s) = \{x + \alpha_i . s \,/\, \alpha_i \in (\alpha_i)_{i \in \mathbb{N}}\} \cup \{x\}$$

La convergence des algorithmes utilisant un tel pas est montrée en I.1.5.

<u>Remarque 1</u> : $x \in D_d(x,s) = \overline{D_d}(x,s) \subset D(x,s)$

<u>Remarque 2</u> : $T[x,D_d(x,s)] = T[x,D(x,s)] = \{y \in \mathbb{R}^n \ / \ y = \lambda.s \ ; \ \lambda \in \mathbb{R}^+\}$

<u>Dem</u> : $[D_d(x,s) \subset D(x,s)] \Longrightarrow T[x,D_d(x,s)] \subset T[x,D(x,s)]$

soit $y = \lambda s \in T[x,D(x,s)] \ ; \ \lambda \geq 0$

$$\begin{cases} y = \lim\limits_{i\to\infty} \dfrac{\lambda}{\alpha_i}(x + \alpha_i s - x) \text{ avec } \dfrac{\lambda}{\alpha_i} \geq 0 \ ; \ \forall i \in \mathbb{N} \\ \lim\limits_{i\to\infty} (x+\alpha_i s) = x \end{cases}$$

$y \in T[x,D_d(x,s)]$ par définition.

I.1.3.4a) <u>Proposition</u> Soit f continue différentiable en x, si $\nabla f(x).s > 0$, un maximum de f sur $D_d(x,s)$ peut être obtenu en un nombre fini de calculs.

<u>Dem</u> : on montre $\exists j \in \mathbb{N}$ tel que un maximum de f sur $D_d(x,s)$ soit dans : $\{x + \alpha_i s / i \leq j\}$ sinon :

$\forall j, \ \exists i > j$ tel que

$$f(x+\alpha_i s) \geq \mathrm{Max}\{f(x+\alpha_k.s)/k \leq j\} \geq f(x+\alpha_j s)$$

on peut extraire de $\mathbb{N}$ une suite croissante $(i(p))_{p \in \mathbb{N}}$ telle que la suite $(f(x+\alpha_{i(p)}.s))_{p \in \mathbb{N}}$ soit monotone non décroissante.

Par continuité de f :

$$\lim_{p\to\infty} (f(x+\alpha_{i(p)}.s)) = \lim_{i\to\infty} (f(x+\alpha_i s)) = f(x)$$

la suite étant non décroissante :

$$\forall p \in \mathbb{N} : f(x+\alpha_{i(p)}.s) \leq f(x)$$

$$\forall p \in \mathbb{N} : \frac{f(x+\alpha_{i(p)}.s)-f(x)}{\alpha_{i(p)}} \leq 0$$

$$\lim_{p\to\infty} \left[\frac{f(x+\alpha_{i(p)}.s)-f(x)}{\alpha_{i(p)}}\right] = \nabla f(x).s \leq 0$$

contradiction avec l'hypothèse

donc :

$\exists j$ tel que $\{x+\alpha_i.s/i \leq j\}$ contienne un maximum de f sur $D_d(x,s)$; celui-ci sera trouvé en un nombre fini de calculs.

<u>Remarque 3</u> : l'indice j de la proposition précédente est inconnu, aussi quel que soit le nombre de points testés on ne peut être assuré d'avoir atteint j et donc un maximum de f sur $D_d(x,s)$.

I.1.3.4b) <u>Proposition</u> Soit f continue différentiable en x et quasi-concave, si $\nabla f(x).s > 0$, $\exists j$ tel que :

$$f(x+\alpha_{j+1}.s) < f(x+\alpha_j.s)$$

si j est plus petit indice qui vérifie cette relation $x+\alpha_j.s$ est un maximum de f sur $D_d(x,s)$

<u>Dem</u> : la 1° partie se montre comme dans la proposition précédente : sinon la suite $(f(x+\alpha_i s))_{i\in\mathbb{N}}$ monotone non décroissante converge vers x et on aboutit à la même contradiction.
Soit j le plus petit indice tel que : $f(x+\alpha_{j+1}.s) < f(x+\alpha_j.s)$

$$f(x+\alpha_j.s) \geq f(x+\alpha_{j-1}.s) \ldots \geq f(x+\alpha_0.s)$$
$$f(x+\alpha_j.s) \geq f(x+\alpha_i.s) \ ; \ \forall i \leq j-1$$

On montre de plus que :

$$f(x+\alpha_j s) > f(x+\alpha_i.s) \ ; \ \forall i \geq j+1$$

c'est vrai pour j+1 par hypothèse, si ce n'était pas vérifié pour tout $i \geq j+2$ on aurait :

$$\exists i \geq j+2 \text{ tel que : } f(x+\alpha_i s) \geq f(x+\alpha_j.s) > f(x+\alpha_{j+1}.s)$$

avec : $(x+\alpha_{j+1}.s) \in [x+\alpha_i s \ ; \ x+\alpha_j.s]$
contradiction avec la quasi-concavité de f qui assure :

$$f(x+\alpha_{j+1}.s) \geq \text{Min}\ \{f(x+\alpha_i s),\ f(x+\alpha_j.s)\}$$
$$f(x+\alpha_{j+1} s) \geq f(x+\alpha_j.s)$$

<u>Remarque 4</u> : l'indice j de la proposition précédente est inconnu, mais cette fois on dispose d'un "test d'arrêt" pour le déterminer. Pratiquement il suffira donc d'examiner les points de $D_d(x,s)$ pour les valeurs croissantes de i et d'arrêter la recherche dès que :

$$f(x+\alpha_{j+1}.s) < f(x+\alpha_j.s)$$

ceci se produisant en un nombre fini de calculs.

I.1.4 Contrôle de la direction

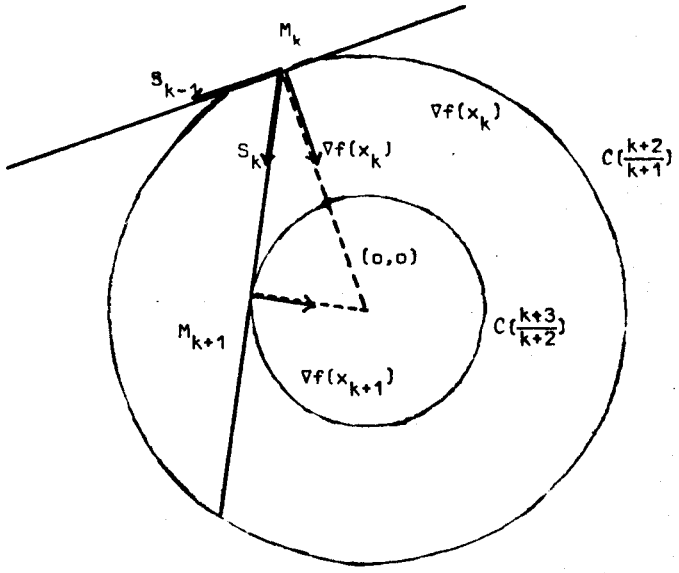
On peut penser que le fait de choisir à chaque itération une direction améliorante et une maximisation exacte selon cette direction assure la convergence vers un point stationnaire.
Ce n'est pas le cas.

I.1.4.a) <u>Exemple</u> TOPKIS et VEINOTT [37]
Maximiser $-\frac{1}{2}(x^2+y^2)$; $(x,y) \in \mathbb{R}^2$
(0,0) est le seul point stationnaire. Les équipotentielles sont des cercles concentriques.
On note C(n) le cercle de centre 0 de rayon r. $(M_k)_{k\in\mathbb{N}}$ la suite telle que :

$$M_k \in C(\tfrac{k+2}{k+1})$$

$[M_{k-1}, M_k]$ est tangente à ce cercle en M_k.

On vérifie que : $\bar{M} = \lim_{k\to\infty} (M_k) \in C(1)$ n'est pas stationnaire. Pourtant chaque direction $s_k = M_{k+1}-M_k$ est améliorante et M_{k+1} maximise exactement f sur $D(M_k, s_k)$



On peut vérifier dans cet exemple que : $\lim_{k\to\infty} [\nabla f(x_k).s_k] = 0$ en concordance avec les résultats de I.1.3 et pourtant $\lim_{k\to\infty}[||\nabla f(M_k)||] = ||\nabla f(\bar{M})|| = 1$, ceci est dû au fait que la direction s_k tend vers une position normale au gradient.

I.1.4.b) <u>Proposition</u> MOKHTAR-KHARROUBI. Soit f continument différentiable, bornée sur $\mathbb{R}^n$ et $\Gamma : \mathbb{R}^n \to P(\mathbb{R}^n)$ telle que, $\forall x$:

H1 : $\Gamma(x) \subset D_+^*(x)$

H2 : $[\Gamma(x) = \emptyset] \Rightarrow [D_+^*(x) = \emptyset]$

H3 : Γ sup-continue en tout point non stationnaire pour f.

<u>si</u> :

$$\lim_{k\to\infty} [\nabla f(x_k).s_k] = 0 \text{ et } \forall k \in \mathbb{N} : ||s_k|| = 1 \; ; \; s_k \in \Gamma(x_k)$$

et que la suite $(f(x_k))_{k\in\mathbb{N}}$ est non décroissante : la suite $(x_k)_{k\in\mathbb{N}}$ possède des points d'accumulation. Si $\overset{*}{x}$ est un tel point :

$$\nabla f(\overset{*}{x}) = 0$$

<u>Dem</u> : $\forall k \in \mathbb{N}$; $x_k \in \{y \in \mathbb{R}^n \,/\, f(y) \geq f(x_o)\}$ compact
soit $\overset{*}{x}$ un point d'accumulation :

$$\exists N' \subset \mathbb{N} \text{ tel que : } (x_n)_{n\in N'} \to \overset{*}{x}$$

$(s_n)_{n\in N'}$ située dans un compact y admet au moins un point d'accumulation $\overset{*}{s}$

$$\exists N'' \subset N' \text{ tel que : } (s_n)_{n\in N''} \to \overset{*}{s}$$

$$\left.\begin{array}{l}(x_n,s_n)_{n\in N^*} \to (\overset{*}{x},\overset{*}{s}) \\ \\ \nabla f \text{ continue}\end{array}\right\} \Rightarrow \lim_{n\in N^*} (\nabla f(x_n).s_n) = \nabla f(\overset{*}{x}).\overset{*}{s}$$

$$\lim_{n\in N^*} [\nabla f(x_n)s_n] = \lim_{n\in \mathbb{N}} [\nabla f(x_n).s_n] = 0 = \nabla f(\overset{*}{x}).\overset{*}{s}$$

si $\overset{*}{x}$ n'est pas stationnaire

$$\left.\begin{array}{l}\Gamma \text{ sup-continue} \\ \\ s_n \in \Gamma(x_n),\ \forall n \in \mathbb{N}\end{array}\right\} \Rightarrow (\overset{*}{s} \in \Gamma(\overset{*}{x}) \subset D_+^*(\overset{*}{x})) \Longrightarrow (\nabla f(\overset{*}{x}).\overset{*}{s} > 0)$$

on aboutit donc à une contradiction.
$\overset{*}{x}$ est stationnaire.

I.1.4.c) <u>Application de I.1.4.b)</u>

Pour tout $\rho \in]0,1]$ on définit :

$$\begin{cases}\Gamma_\rho(x) = \{s \in \mathbb{R}^n / ||s|| = 1 \ ;\ \nabla f(x).s \geq \rho.||\nabla f(x)||\} \text{ si } \nabla f(x) \neq 0 \\ \\ \Gamma_\rho(x) = \emptyset \text{ sinon}\end{cases}$$

Cette fonction multivoque vérifie H1, H2, H3.

<u>Dem</u> : H1 : évident

H2 : si x est stationnaire $\Gamma_\rho(x) = \emptyset$

or $[\nabla f(x) = 0] \Longrightarrow [D_+^*(x) = \emptyset]$ par I.1.1.d)

si x n'est pas stationnaire on montre que $\Gamma_\rho(x) \neq \emptyset$

si $\nabla f(x) \neq 0$; soit $\rho \in]0,1]$

$s = \dfrac{\nabla f(x)}{||\nabla f(x)||} \in \Gamma_\rho(x)$ car :

$\nabla f(x).s = ||\nabla f(x)|| \geq \rho.||\nabla f(x)||$ et $\Gamma_\rho(x) \neq \emptyset$

H3 : f étant continument différentiable si $\nabla f(x) \neq 0$ alors ∇f est non nulle dans un voisinage de x.

$$\Gamma_\rho(x) = \{s \in \mathbb{R} / ||s|| - 1 = 0\} \cap \{s \in \mathbb{R}^n / \nabla f(x).s - \rho.||\nabla f(x)|| \geq 0\}$$

$$\Gamma_\rho(x) = \{s \in \mathbb{R}^n / g_1(x,s) = 0\} \cap \{s \in \mathbb{R}^n / g_2(x,s) \geq 0\}$$

sous les hypothèses faites g_1 et g_2 sont continues ; Γ_ρ intersection de deux fonctions sup-continues est sup-continue.

On appelle parfois directions uniformement améliorantes, des directions telles que : $\forall k \in \mathbb{N}$ $s_k \in \Gamma_\rho(x_k)$

I.1.4.d) Application de I.1.4.b) MOKHTAR-KHARROUBI [27]

Soit H l'application de Y compact non vide de $\mathbb{R}^n$ dans l'ensemble des matrices carrées d'ordre n, telle que :

- H est continue sur Y
- $\forall y \in Y$, H(y) est définie positive

alors :

$$\Gamma(x) = \begin{cases} \{H(y).\nabla f(x) \ / \ y \in Y\} \text{ si } \nabla f(x) \neq 0 \\ \emptyset \quad \text{sinon} \end{cases}$$

vérifie H1, H2, H3.

Dem : H1 : H(y) est définie positive

$$\nabla f(x).H(y).\nabla f(x) > 0 \text{ dès que } \nabla f(x) \neq 0$$

H2 : évident car Y est non vide

H3 : soient $(x_n)_{n \in \mathbb{N}} \to \bar{x}$ non stationnaire

$(s_n)_{n \in \mathbb{N}} \to \bar{s}$; $s_n \in \Gamma(x_n)$, $\forall n \in \mathbb{N}$

$\forall n \in \mathbb{N}$, $\exists y_n \in Y$ tel que : $s_n = H(y_n).\nabla f(x_n)$

$(y_n)_{n \in \mathbb{N}} \subset Y$ compact, admet un point d'accumulation $\bar{y}$

$\exists N' \subset \mathbb{N}$ tel que : $(x_n, s_n, y_n)_{n \in N'} \to (\bar{x}, \bar{s}, \bar{y})$

par continuité de H et de ∇f :

$\bar{s} = H(\bar{y}).\nabla f(\bar{x})$

$\bar{s} \in \Gamma(\bar{x})$ ce qui montre la sup-continuité

Remarque : la condition Y compact est essentielle dans ce qui précède, sinon on pourrait avoir

$$\left.\begin{array}{l} (y_n) \to \bar{y} \in \bar{Y}-Y \\ \\ \forall y \in Y : H(y) \\ \text{définie positive} \end{array}\right\} \Rightarrow H(\bar{y}) \text{ semi-définie-positive}$$

alors : $\bar{s}.\nabla f(\bar{x}) = \nabla f(\bar{x}).H(\bar{y}).\nabla f(\bar{x}) = 0$

n'entraine plus : $\nabla f(\bar{x}) = 0$

si par contre on peut assurer que $H(\bar{y})$ est définie positive : $\nabla f(\bar{x}).H(\bar{y}).\nabla f(\bar{x}) = \rho > 0$ dès que $\bar{x}$ est non stationnaire et par continuité à partir d'un rang : $\nabla f(x_n).H(y_n).\nabla f(x_n) \geq \frac{\rho}{2} > 0$ on retrouvera en conformité avec I.1.4.c) des directions uniformément améliorantes à partir d'un certain rang.

I.1.4.e) Application de I.1.4.b) : cas discret.

Si on choisit associée à $(x_k)_{k \in \mathbb{N}}$ une suite de matrices $(H_k)_{k \in \mathbb{N}}$ définies positives;
H un point d'accumulation de cette suite sera une matrice semi-définie-positive et non définie positive.
Si l'on veut que les directions $H_k . f(x_k)$ soient uniformément améliorantes, il faut, $||\nabla f(x_k)||$ étant supposé non nul :

$$\forall k \in \mathbb{N} : \frac{\nabla f(x_k).H_k.\nabla f(x_k)}{||\nabla f(x_k)||^2} \geq \rho > 0$$

ce sera le cas si :

$$\forall k \in \mathbb{N} : \frac{y.H_k.y}{||y||^2} \geq \rho \ ; \ \forall y \in \mathbb{R}^{n*}$$

on dit que les matrices H_k sont uniformément définies positives. Dans tous les algorithmes qui utilisent de telles suites de matrices cette propriété sera fondamentale. On verra au chapitre III que c'est en particulier le cas pour les algorithmes de métriques variables.
Une autre possibilité pour obtenir H définie positive est de choisir dans $(H_k)_{k \in \mathbb{N}}$ une sous suite $(H_k)_{k \in N'}$ telle que :

$$\forall k \in N' \ ; \ H_k = H \text{ matrice constante définie positive}$$

alors :

$$\bullet \ \overset{*}{H} = \lim_{k \in N'} (H_k) = H \text{ est définie positive.}$$

I.1.4.f) Direction de plus grande pente.

Le choix de la direction améliorante défini en I.1.4.b) ou I.1.4.c) reste "large" et l'on peut penser à le particulariser. Il est logique de choisir en chaque point non pas une direction améliorante, mais la (ou une) "meilleure direction améliorante".
Le problème est de définir un critère de "meilleure direction améliorante". Selon le choix de ce critère les réponses restent multiples.
Soit x tel que $\nabla f(x) \neq 0$; $\theta \in \mathbb{R}^+$ fixé et $|| \ ||$ une norme quelconque la résolution du problème :

$$\boxed{\underset{||y||^2 \leq \theta}{\text{Max}} [f(x+y)-f(x)]}$$

n'est pas plus aisé que celui de la maximisation de f dans $\mathbb{R}^n$. On va approcher celui-ci, quand f est différentiable par une approximation au 1° ordre de l'accroissement de la fonction

$$\boxed{\underset{||y||^2 \leq \theta}{\text{Max}} \nabla f(x).y}$$

Pour résoudre ce problème convexe à fonction économique linéaire, il reste à choisir une norme. Dans ZOUTENDIJK [43] différents choix de normes sont analysés. On va ici faire l'étude associée à la norme : $||y||^2 = y.Ay$ ou A est une matrice définie positive le problème devient :

$$\boxed{\begin{array}{l} \text{Max}[\nabla f(x).y] \\ \theta - y.Ay \geq 0 \end{array}}$$

on sait résoudre un tel problème dont l'unique solution est donnée par les conditions nécessaires et suffisantes de Kühn et Tucker [20].

$$\exists \overset{*}{u} \in \mathbb{R}^+ \text{ tel que : } \begin{cases} -\overset{*}{u}(A+A^t)\overset{*}{y} + \nabla f(x) = 0 \\ \overset{*}{u}(\theta - \overset{*}{y}A\overset{*}{y}) = 0 \end{cases}$$

si : $\nabla f(x) \neq 0$ alors $\overset{*}{u} > 0$

la solution est : $\overset{*}{y} = \frac{1}{u^*}(A+A^t)^{-1}.\nabla f(x)$

la meilleure direction améliorante étant choisie :

$$\boxed{s = (A+A^t)^{-1}.\nabla f(x)}$$

Remarque 1 : la direction obtenue est améliorante

$$\nabla f(x).s = \nabla f(x).(A+A^t)^{-1}.\nabla f(x) > 0$$

car $(A+A^t)^{-1}$ est définie positive dès que A l'est.

Remarque 2 : si A = I la matrice unité de rang n la direction améliorante est la direction de gradient, CAUCHY [3]. Une telle méthode a l'avantage d'être simple à mettre en oeuvre ; mais a l'inconvénient de donner une vitesse de convergence linéaire qui peut être très lente. Ceci étant essentiellement dû au fait qu'à chaque itération tout renseignement ayant pu être obtenu au cours des itérations précédentes est "oublié" ; les calculs en un point ne dépendent jamais des itérés précédents.

Remarque 3 : si f est 2 fois différentiable de Hessien H(x) on trouve $s = [H(x)]^{-1}.\nabla f(x)$ direction de la méthode de NEWTON-RALPHSON. Celle ci assure une vitesse de convergence superlinéaire avec différents choix de pas POLAK [30]. Par contre elle requiert à chaque itération l'inversion de la matrice H(x) ce qui est en général prohibitif.

Remarque 4 : si H_k est une matrice définie positive symétrique on peut considérer $H_k.\nabla f(x_k) = s_k$ comme la meilleure direction améliorante pour la norme :

$$||y||^2 = y.H_k^{-1}y$$

Les algorithmes de métriques variables (chapitre III) sont donc des algorithmes de plus grande pente. Le problème étant de choisir les matrices H_k de façon a avoir avec une simplicité analogue à la méthode de Cauchy, les performances de la méthode de Newton-Ralphson.

I.1.5 Algorithme général

Cet algorithme est analogue à celui décrit par HUARD [21].
Soit $f : \mathbb{R}^n \to R$ une fonction continument différentiable. On suppose qu'il existe un point $x_o \in \mathbb{R}$ tel que :

$$E_o = \{x \in \mathbb{R}^n \ / \ f(x) \geq f(x_o)\}$$

soit compact.

Soient : B un compact de $\mathbb{R}^{n*}$

$\Gamma : \mathbb{R}^n \to P(B)$

$\Delta : \mathbb{R}^n \times B \to P(\mathbb{R}^n)$

$M : \mathbb{R}^n \times B \times \mathbb{R}^+ \to P(\mathbb{R}^n)$

ces fonctions multivoques vérifiant pour tout x :

$$\begin{cases} \Gamma(x) \subset D_+^*(x) \\ [\Gamma(x) = \emptyset] \Rightarrow [D_+^*(x) = \emptyset] \\ \Gamma \text{ sup-continue en tout point non stationnaire pour } f \end{cases}$$

$$\begin{cases} x \in \Delta(x,s) = \overline{\Delta(x,s)} \subset D(x,s) \\ T[x,\Delta(x,s)] = T[x,D(x,s)] \\ \Delta \text{ continue} \end{cases}$$

$$\{ \ M(x,s,\varepsilon) = \{y \in \Delta(x,s) / y \text{ maximise } f \text{ à } \varepsilon \text{ près sur } \Delta(x,s)\}$$

<u>Algorithme</u> : itération 0 : x_o est connu, ε_o connu

itération k :

si $\Gamma(x_k) = \phi$ STOP

Sinon : $s_k \in \Gamma(x_k)$

si $x_k \in M(x_k,s_k,\varepsilon_k)$ alors $x_{k+1} = x_k$

sinon : $x_{k+1} \in M(x_k,s_k,\varepsilon_k)$

$k = k+1$

I.1.5.a) <u>Proposition</u> Si l'algorithme précédent est fini, le dernier point obtenu est stationnaire. S'il est infini et si la suite $(\varepsilon_k)_{k \in \mathbb{N}}$ tend vers 0, la suite $(x_k)_{k \in \mathbb{N}}$ engendrée par l'algorithme possède au moins un point d'accumulation. Un tel point est stationnaire

<u>Dem</u> : si l'algorithme est fini : $\exists k \in \mathbb{N}$ tel que : $\Gamma(x_k) = \emptyset$
alors : $[\Gamma(x_k) = \emptyset] \Rightarrow [D_+^*(x_k) = \emptyset] \Rightarrow [\nabla f(x_k) = 0]$ d'après I.1.1.d. Sinon, soit $(x_k)_{k \in \mathbb{N}}$ la suite engendrée. Par construction $(f(x_k))_{k \in \mathbb{N}}$ est une suite non décroissante. La suite $(x_k)_{k \in \mathbb{N}}$ est donc dans le compact E_o et y possède au moins un point d'accumulation $\overset{*}{x}$.

$\exists N \subset \mathbb{N}$ tel que : $(x_n)_{n \in N} \to \overset{*}{x}$

$$\left.\begin{array}{l} \forall k \in \mathbb{N},\ s_k \in B \text{ compact} \\ (\varepsilon_k)_{k \in \mathbb{N}} \to 0 \end{array}\right\} \quad \exists N' \subset N \text{ tel que : } (x_k, s_k, \varepsilon_k)_{k \in N'} \to (\overset{*}{x}, \overset{*}{s}, o)$$

$\forall k \in \mathbb{N},\ s_k \in \Gamma(x_k)$

Si $\overset{*}{x}$ n'est pas stationnaire :

$$\left.\begin{array}{l} s_k \in \Gamma(x_k),\ \forall k \in \mathbb{N} \\ \Gamma \text{ sup-continue en } \overset{*}{x} \end{array}\right\} \Rightarrow \overset{*}{s} \in \Gamma(\overset{*}{x})$$

1° - On montre que : $\overset{*}{x} \in M(\overset{*}{x}, \overset{*}{s}, o)$ ZANGWILL [15], soit $\overset{**}{x}$ un point d'accumulation de $(x_{k+1})_{k \in N'}$ $\exists N'' \subset N' \subset N \subset \mathbb{N}$ tel que :

$$(x_n)_{n \in N''} \to \overset{*}{x}$$

$$(x_{n+1})_{n \in N''} \to \overset{**}{x}$$

la suite $(f(x_n))_{n \in \mathbb{N}}$ est monotone non décroissante, les suites extraites $(f(x_n))_{n \in N''}$ et $(f(x_{n+1}))_{n \in N''}$ admettent la même limite. Par continuité de f :

$$f(\overset{*}{x}) = f(\overset{**}{x}) \qquad (\alpha)$$

de plus :

$$\left.\begin{array}{l} \Delta \text{ continue} \\ \text{définition de } M(x_k, s_k, \varepsilon_k) \end{array}\right\} \Rightarrow M \text{ est sup-continue} \qquad [20]$$

$$\left.\begin{array}{l} \forall k \in \mathbb{N} : x_{k+1} \in M(x_k, s_k, \varepsilon_k) \\ M \text{ est sup-continue} \\ (\varepsilon_k)_{k \in N''} \to 0 \ ;\ (x_{k+1})_{k \in N''} \to \overset{**}{x} \ ;\ (x_k)_{k \in N''} \to \overset{*}{x} \end{array}\right\} \Rightarrow \overset{**}{x} \in M(\overset{*}{x}, \overset{*}{s}, o) \qquad (\beta)$$

(β) et $(\alpha) \Rightarrow \overset{*}{x} \in M(\overset{*}{x}, \overset{*}{s}, o)$

2° - On montre que : $\overset{*}{s}.\nabla f(\overset{*}{x}) \leq 0$ résultat en concordance avec I.1.3.

$$\begin{aligned} \overset{*}{x} \in M(\overset{*}{x}, \overset{*}{s}, 0) &\Rightarrow \overset{*}{x} \text{ maximise f exactement sur } \Delta(\overset{*}{x}, \overset{*}{s}) \\ &\Rightarrow \nabla f(\overset{*}{x}) \in \Pi[T[\overset{*}{x}, \Delta(\overset{*}{x}, \overset{*}{s})]] \\ &\Rightarrow \nabla f(\overset{*}{x}) \in \Pi[T[\overset{*}{x}, \Delta(\overset{*}{x}, \overset{*}{s})]] \\ &\Rightarrow \nabla f(\overset{*}{x}).\overset{*}{s} \leq 0 \end{aligned}$$

où Π désigne le cône polaire du cône T.

3° - On montre alors que : $\nabla f(\overset{*}{x}) = 0$ en concordance avec I.1.4.

$$\left.\begin{array}{l} \Gamma(\overset{*}{x}) \subset D_+^*(\overset{*}{x}) \\ \\ \overset{*}{s} \in \Gamma(\overset{*}{x}) \quad . \end{array}\right\} \Rightarrow \nabla f(\overset{*}{x}).\overset{*}{s} > 0$$

on aboutit à une contradiction avec le résultat du 2°. C'est donc que $\overset{*}{x}$ est stationnaire et :

$$\nabla f(\overset{*}{x}) = 0$$

Remarque 1 : si $\Delta(x,s) = D(x,s)$ les hypothèses sont vérifiées. Si f est de plus concave, on obtient par I.1.3.a) un contrôle de l'erreur commise et chaque itération peut se faire en un nombre fini de calculs tant que ε n'est pas nul.

Remarque 2 : si $\Delta(x,s) = D_d(x,s)$ défini en I.1.3.4 les hypothèses sont vérifiées. Si f est de plus quasi-concave on peut effectuer chaque itération en un nombre fini de calculs, même si ε est nul. On pourra pratiquement dans ce cas, maximiser f exactement à chaque itération sur $D_d(x,s)$, c'est à dire choisir $\varepsilon_k = 0$, $\forall_k \in \mathbb{N}$.

I.2 - L'ADAPTATION À L'OPTIMISATION AVEC CONTRAINTES.

Pour résoudre le problème de la maximisation avec contraintes :

$$\boxed{\begin{array}{l} \text{Max } f(x) \\ x \in C \end{array}} \qquad f : \mathbb{R}^n \to \mathbb{R} \quad ; \quad C \subset \mathbb{R}^n$$

on peut chercher à adapter les principes des algorithmes de direction améliorante, utilisés pour l'optimisation dans $\mathbb{R}^n$. Un moyen simple pour tenir compte du domaine C sera d'associer à tout point un successeur selon les principes décrits en I.1, mais de façon à ce que ce successeur soit de plus réalisable.

I.2.1 - Direction réalisable ZOUTENDIJK [43].

Soient $C \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$; $s \in \mathbb{R}^n$

I.2.1.1 Définition : s est une direction réalisable pour C en x si et seulement si :

$$\exists \lambda \in \mathbb{R}^{+*} \text{ tel que } [x, x+\lambda s] \subset C$$

Soit $R_C(x)$ l'ensemble des directions réalisables pour C en x

I.2.1.2 Propriétés

(1) : $[x \in C] \Rightarrow R_C(x)$ est un cône pointé

(2) : $[x \in \overset{\circ}{C}] \Rightarrow [R_C(x) = \mathbb{R}^n]$

(3) : $[x \notin C] \Rightarrow [R_C(x) = \emptyset]$

(4) : $\overline{R_C(x)} \subset T(x,C)$

(5) : $$C \text{ convexe} \Rightarrow \left\{\begin{array}{l} R_C(x) \text{ est convexe} \\ \\ \overline{R_C(x)} = T(x,c) \end{array}\right.$$

<u>Dem</u> : (1) (2) et (3) sont évidents. La structure de cône permettra, comme pour les directions améliorantes, de caractériser une direction réalisable par un vecteur de norme 1.

(4) : $T(x,C)$ étant fermé il suffit de montrer :

$$R_C(x) \subset T(x,C)$$

$$s \in R_C(x) \Longleftrightarrow \exists\lambda > 0,\ [x,x+\lambda s] \subset C$$

$$\forall n \in \mathbb{N} \text{ on pose : } x_n = x + \frac{\lambda.s}{n}$$

$$\left.\begin{array}{l} \left.\begin{array}{l} (x_n)_{n\in\mathbb{N}} \to x \\ (n(x_n-x))_{n\in\mathbb{N}} \to \lambda.s \end{array}\right\} \Rightarrow \begin{array}{l} \lambda.s \in T(x,C) \\ T(x,C) \text{ cône} \\ \lambda > 0 \end{array} \end{array}\right\} \Rightarrow s \in T(x,C)$$

(5) :

$$s_1 \in R_C(x) \Longleftrightarrow \exists\lambda_1 > 0 : [x,x+\lambda_1 s_1] \in C$$

$$s_2 \in R_C(x) \Longleftrightarrow \exists\lambda_2 > 0 : [x,x+\lambda_2 s_2] \in C$$

soit $\lambda = \mathrm{Min}(\lambda_1,\lambda_2)$

$x+\lambda s_1 \in C$ et $x+\lambda.s_2 \subset C$

C convexe $\Rightarrow$ $\forall\theta \in [0,1]$

$\theta(x+\lambda s_1)+(1-\theta)(x+\lambda s_2) \in C$

$x+\lambda.[\theta s_1+(1-\theta).s_2] \in C$

$\exists\lambda > 0$ tel que : $[x,x+\lambda(\theta s_1+(1-\theta)s_2] \subset C$

$\forall\theta \in [0,1]$: $\theta s_1+(1-\theta)s_2 \in R_C(x)$

$R_C(x)$ est convexe C.Q.F.D

d'après (4) il suffit de montrer que si C est convexe :

$$T(x,C) \subset \overline{R_C(x)}$$

soit $s \in T(x,C)$

$$\left.\begin{array}{l} \exists(x_n)_{n\in\mathbb{N}} \to x \ ;\ \forall n \in \mathbb{N}, x_n \in C \\ \exists(\lambda_n)_{n\in\mathbb{N}} \ ;\ \forall n \in \mathbb{N}\ \lambda_n > 0 \end{array}\right\}$$

telles que : $(\lambda_n(x_n-x))_{n\in\mathbb{N}} \to s$

$$\text{C convexe} \Rightarrow \left.\begin{array}{l} [x,x_n] \subset C \\ \lambda_n > 0 \end{array}\right\} \Rightarrow s_n = \lambda_n(x_n-x) \in R_C(x)$$

donc : $s \in \overline{R_C(x)}$ C.Q.F.D

Si dans le cas le plus général on ne sait expliciter ni $R_C(x)$, ni $\overline{R_C(x)}$ cela devient possible dans le cas particulier où C est un polyèdre.

I.2.1.3 <u>Proposition</u> Si C est un polyèdre et $x \in C$:

$$R_C(x) = \overline{R_C(x)} = T(x,C)$$

<u>Dem</u> :

$$C = \{x \in \mathbb{R}^n \ / \ A_i.x \geq b_i \ ; \ i = 1 \text{ à } m\}$$

$$E(x) = \{i \in \{1,\dots,m\} \ / \ A_i.x = b_i\}$$

$$\overline{E(x)} = \{i \in \{1,\dots,m\} \ / \ A_i.x > b_i\}$$

ce sont respectivement les indices des contraintes actives et des contraintes inactives en x, alors [20] :

$$T(x,C) = \{s \in \mathbb{R}^n \ / \ A_i.s \geq 0, \ \forall i \in E(x)\}$$

pour montrer cette proposition il suffit de montrer que :

$$T(x,C) \subset R_C(x)$$

soit $s \in T(x,C)$ par continuité des fonctions linéaires :

$$\exists \theta > 0 \ ; \ \forall i \in \overline{E(x)} \ ; \ \forall \lambda \in [0,\theta] \ : \ A_i(x+\lambda s) > b_i$$

alors : $\forall i \in E(x) \ : \ A_i(x+\lambda s) = A_i x + \lambda . A_i s \geq b_i$

$\exists \theta$ tel que : $[x, x+\theta s] \subset C$

donc : $s \in R_C(x)$ C.Q.F.D

La définition introduite d'une direction réalisable a un aspect essentiellement linéaire. Ceci est à l'origine des remarques suivantes :

<u>Remarque 1</u> :

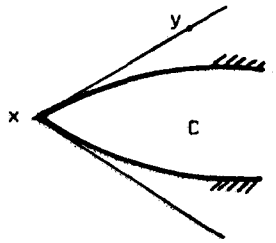
$$\left.\begin{array}{l} g : \mathbb{R}^n \to \mathbb{R} \text{ non linéaire} \\ G = \{y \in \mathbb{R}^n \ / \ g(y) = 0\} \\ x \in G \end{array}\right\} \Rightarrow R_G(x) = \{0\}$$

Les algorithmes qui utilisent les directions réalisables ne peuvent traiter des problèmes d'optimisation où figurent des contraintes non linéaires en égalité.

<u>Remarque 2</u> : il est plus facile d'expliciter T(x,C) que $R_C(x)$, aussi est-ce dans ce premier ensemble que pratiquement, sera cherchée une direction s.
Mais dès que le domaine C est défini à l'aide d'inégalités non linéaires, même s'il est convexe, on peut alors déterminer une direction s qui est dans T(x,C) sans être dans $R_C(x)$. La partie réalisable de la demi-droite D(x,s) est réduite au seul point x.

Exemple dans $\mathbb{R}^2$:

$$s = y-x \in T(x,C)$$
$$s \notin R_C(x)$$



La définition de $R_C(x)$ utilisée s'attache à un aspect uniquement local du domaine C. Dans le cas où C est un polyèdre, $R_C(x)$ s'explicite en utilisant seulement les contraintes actives en x.

Remarque 3 : même si C est convexe compact ; R_C n'est pas en général une fonction multivoque sup-continue de x sur C.

Exemple : $C = \{(x_1,x_2) \in \mathbb{R}^2 \ / \ 0 \le x_1 \le 1 \ ; \ 0 \le x_2 \le 1\}$

si $\forall n \in \mathbb{N}$: $x_n = (1 - \frac{1}{n},0)$ et $s_n = (1,0)$ alors $s_n \in R_C(x_n)$

$$\overset{*}{x} = \lim_{n\to\infty} (x_n) = (1,0) \qquad s = \lim_{n\to\infty} (s_n) = (1,0)$$

$$s \notin R_C(\overset{*}{x})$$

ce qui montre que R_C n'est pas sup-continue sur C. La propriété I.2.1.2 (2) montre que R_C est continue sur $\overset{o}{C}$ puisque constante, égale à $\mathbb{R}^n$ sur cet ensemble ; les seuls points de C où R_C ne sera pas sup-continue seront les points frontières.

On a vu en I.1.1 la liaison entre direction améliorante et conditons d'optimalité dans $\mathbb{R}^n$. On obtient ici des résultats analogues à I.1.1.d) et I.1.1.g) qui marquent la liaison entre direction améliorante réalisable et conditions d'optimalité dans un domaine C.

I.2.1.4 Conditions nécessaires d'optimalité.

Soit f différentiable en $x \in C$

$$\left\{\begin{array}{l}\text{x maximum local}\\ \text{de f sur C}\end{array}\right\} \underset{(1)}{\Longrightarrow} (A_+^*(x) \cap T(x,C) = \emptyset) \underset{(2)}{\Longrightarrow} (D_+^*(x) \cap T(x,C) = \emptyset) \underset{(3)}{\Longrightarrow} (D_+^*(x) \cap \overline{R_C(x)} = \emptyset)$$

$$\Updownarrow (4) \qquad\qquad \Updownarrow (5)$$

$$\nabla f(x) \in \Pi[T(x,C)] \qquad \nabla f(x) \in \Pi[\overline{R_C(x)}] = \Pi[R_C(x)]$$

Dem : (1) on démontre la contraposée

$$s \in A_+^*(x) \cap T(x,C)$$

$$\exists \lambda_1 > 0 \text{ tel que } [x,x+\lambda_1 s] \subset C$$
$$\exists \lambda_2 > 0 \text{ tel que } f(x+\theta.s) > f(x), \ \forall \theta \in [0,\lambda_2]$$
$$\lambda = \mathrm{Min}(\lambda_1,\lambda_2)$$
$$[x,x+\lambda s] \subset C \text{ et } \forall y \in [x,x+\lambda s], \ f(y) > f(x)$$

à fortiori x n'est pas un maximum local de f sur C

(2) évident car $D_+^*(x) \subset A_+^*(x)$ (I.1.1 remarque 3)

(3) évident car $\overline{R_C(x)} \subset T(x,C)$ (I.2.1.2 (4))

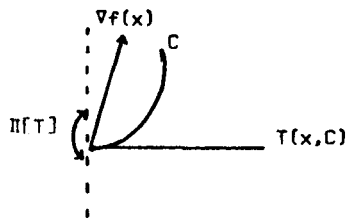
(4) et (5) : d'une manière générale soit K un cône de sommet x, $\Pi(K)$ le cône polaire :

$$(D_+^*(x) \cap K = \emptyset) \iff (\forall s \in K \; ; \; \nabla f(x).s \leq 0) \iff (\nabla f(x) \in \Pi(K))$$

de plus $\Pi(K) = \Pi(\bar{K})$ est une propriété générale des cônes polaires.

Remarque 4 : la réciproque de l'implication (3) est fausse en général

Exemple dans $\mathbb{R}^2$: $\nabla f(x) \notin \Pi[T(x,C)]$ à fortiori x n'est pas un maximum local de f sur C. Or $R_C(x)=\overline{R_C(x)}=\{0\}$
$\nabla f(x) \in \Pi[\overline{R_C(x)}] = \mathbb{R}^n$



Donc la condition nécessaire d'optimalité $\nabla f(x) \in \Pi[\overline{R_C(x)}]$ est une condition plus faible que :
$\nabla f(x) \in \Pi[T(x,C)]$ condition usuelle de Kühn et Tucker.
Dans le cas où C est convexe et f concave les conditions précédentes deviennent équivalentes et sont de plus suffisantes.

I.2.1.5 Conditions suffisantes d'optimalité

Soit f concave différentiable en $x \in C$ convexe, alors les réciproques des implications (1) (2) et (3) de la proposition I.2.1.4 sont vérifiées.

Dem :
C convexe $\Rightarrow$ $(\overline{R_C(x)} = T(x,C))$ par I.2.1.2 (5)
la réciproque de (3) est vraie

f concave $\Rightarrow$ $(D_+^*(x) = A_+^*(x))$ par I.1.1.e)
la réciproque de (2) est vraie

On montre :

$$\begin{pmatrix} x \text{ non maximum} \\ \text{de f sur C} \end{pmatrix} \Rightarrow (A_+^*(x) \cap T(x,C) \neq \emptyset)$$

$\forall v(x)$ voisinage de x dans $\mathbb{R}^n$:

$\exists y \in v(x) \cap C$ tel que : $f(y) > f(x)$

$$\left.\begin{array}{l} f \text{ concave} \\ \\ f(y) > f(x) \end{array}\right\} \Rightarrow (\forall z \in]x,y] : f(z) > f(x)) \Rightarrow (y-x \in A_+^*(x)) \qquad (\alpha)$$

$$\left.\begin{array}{l}\left.\begin{array}{l} y \in C,\ x \in C \\ \\ C \text{ convexe}\end{array}\right\} \Rightarrow ([x,y] \subset C) \\ \\ C \text{ convexe} \Rightarrow (C \subset T(x,C))\end{array}\right\} \Rightarrow y-x \in T(x,C) \qquad (\beta)$$

$$(\alpha) \text{ et } (\beta) \Rightarrow (A_+^*(x) \cap T(x,C) \neq \emptyset) \text{ C.Q.F.D.}$$

Remarque 5 :

$D_+^*(x)$ étant un ouvert de $\mathbb{R}^n$:

$$[D_+^*(x) \cap \overline{R_C(x)} = \emptyset] \Longleftrightarrow [D_+^*(x) \cap R_C(x) = \emptyset]$$

Soient $f : \mathbb{R}^n \to \mathbb{R}$ différentiable et $C \subset \mathbb{R}^n$, d'après I.2.1.4 chacune de ces deux égalités équivalentes constitue une condition nécessaire d'optimalité plus faible que la condition usuelle de Kühn et Tucker ; pour la distinguer de celle-ci on l'appelera par la suite "condition d'optimalité faible".
Dès que f est concave et C convexe cette condition d'optimalité devient suffisante et équivalente à celle de Kühn et Tucker.

I.2.2 Algorithmes de direction réalisable améliorante.

L'adaptation des méthodes de gradient au cas de la maximisation avec contraintes consiste à tenir compte du domaine C dans les deux phases de chaque itération : recherche de direction, recherche du pas.
Dès que f est différentiable, le principe des algorithmes est d'utiliser la condition d'optimalité faible : $D_+^*(x) \cap R_C(x) = \emptyset$, de la manière suivante :
soit $x \in C$ on cherche un successeur à x :

- si $D_+^*(x) \cap R_C(x) = \emptyset$ STOP x vérifie une condition nécessaire d'optimalité ; c'est un point stationnaire.

- sinon : soit $s \in D_+^*(x) \cap R_C(x)$
  choisir $y \in D(x,s) \cap C$ tel que $f(y) > f(x)$
  ce choix étant toujours possible car $s \in D_+^*(x) \cap R_C(x)$.

Un tel algorithme ne peut fournir au plus qu'un point stationnaire et non en général un maximum de f sur C.
De plus les algorithmes utilisant ce principe de direction améliorante et réalisable engendrent des suites de points réalisables. C'est un des avantages fondamentaux de cette classe de méthodes par rapport à d'autres dites de "points extérieurs".
En effet, si théoriquement on peut construire une suite de points $(x_k)_{k \in \mathbb{N}}$ qui converge vers un point stationnaire ou un maximum x, en pratique on ne connait pas x mais seulement une approximation x_k plus ou moins bonne.
Dans le cas présent x_k sera réalisable.

Pour assurer la convergence des algorithmes de gradient dans le cas sans contraintes, on a vu en I.1 les limitations à apporter au choix de la direction et au choix de pas à chaque itération. Celles-ci sont à fortiori nécessaires dans le cas de la maximisation avec contraintes. On peut obtenir l'adaptation suivante de l'algorithme en I.1.5, dans le cas d'un domaine C :

I.2.2.1 Adaptation de l'algorithme décrit en I.1.5

Soient : f continument différentiable ; C un fermé de $\mathbb{R}^n$; B compact de $\mathbb{R}^{n*}$

$x_o \in C$; $E_o = \{x \in C / f(x) \geq f(x_o)\}$ est compact

$\Gamma_C : \mathbb{R}^n \to \mathcal{P}(B)$

$\Delta_C : \mathbb{R}^n \times B \to \mathcal{P}(\mathbb{R}^n)$; $\varepsilon \in \mathbb{R}^+$

telles que :

$$\begin{cases} 1C & : \quad \Gamma_C(x) \subset D_+^*(x) \cap R_C(x) \\ 2C & : \quad [\Gamma_C(x) = \emptyset] \Rightarrow [D_+^*(x) \cap R_C(x) = \emptyset] \\ 3C & : \quad \Gamma_C \text{ sup-continue en tout point tel que : } D_+^*(x) \cap R_C(x) \neq \emptyset \end{cases}$$

$$\begin{cases} 1'_C & : \quad x \in \Delta_C(x,s) = \overline{\Delta_C(x,s)} \subset D(x,s) \cap C \\ 2'_C & : \quad T(x,\Delta_C(x,s)) = T(x,D(x,s)) \\ 3'_C & : \quad \Delta_C \text{ continue} \end{cases}$$

posons : $M(x,s,\varepsilon)$ ensemble des maximum de f sur $\Delta_C(x,s)$ à ε près.

Algorithme et résultat : les mêmes qu'en I.1.5 où Γ_C remplace Γ. Un point stationnaire étant ici tel que :

$$D_+^*(x) \cap R_C(x) = \emptyset$$

Dem : elle reste analogue à celle du cas sans contraintes

si l'algorithme est fini : $\exists k \in \mathbb{N}$ tel que

$$D_+^*(x_k) \cap R_C(x_k) = \emptyset$$

sinon, soit $(\overset{*}{x},\overset{*}{s},o)$ un point d'accumulation de $(x_n,s_n,\varepsilon_n)_{n \in \mathbb{N}}$ si $\overset{*}{x}$ n'est pas stationnaire

Γ_C sup-continue $\Rightarrow \overset{*}{s} \in \Gamma_c(\overset{*}{x}) \subset D_+^*(\overset{*}{x}) \Rightarrow \nabla f(\overset{*}{x}).\overset{*}{s} > 0$

$$\left.\begin{array}{l} \left.\begin{array}{l} (f(x_n))_{n \in \mathbb{N}} \text{ non décroissante} \\ \\ M \text{ sup-continue} \end{array}\right\} \Rightarrow \left.\begin{array}{l} \overset{*}{x} \in M(\overset{*}{x},\overset{*}{s},0) \\ \\ \text{relation sur} \\ \text{les cônes} \end{array}\right\} \Rightarrow \nabla f(\overset{*}{x}).\overset{*}{s} \leq 0 \end{array}\right.$$

On aboutit à la même contradiction qu'en I.1.5.a) donc $\overset{*}{x}$ est stationnaire

L'intérêt de l'algorithme décrit n'existe que s'il est possible de déterminer des fonctions multivoques vérifiant les propriétés requises.

I.2.2.2 Recherche de Γ_C

On cherche à remplacer une fonction Γ vérifiant :

(1) : $\Gamma(x) \subset D_+^*(x)$

(2) : $[\Gamma(x) = \emptyset] \Rightarrow [D_+^*(x) = \emptyset]$

(3) : Γ sup-continue en tout point tel que $\nabla f(x) \neq 0$

par une fonction Γ_C qui vérifie cette fois :

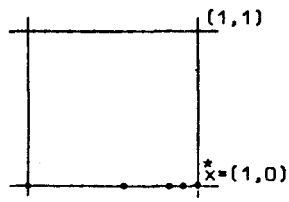
$$(1C) \quad : \quad \Gamma_C(x) \subset D_+^*(x) \cap R_C(x)$$
$$(2C) \quad : \quad [\Gamma_C(x) = \emptyset] \Rightarrow [D_+^*(x) \cap R_C(x) = \emptyset]$$
$$(3C) \quad : \quad \Gamma_C \text{ sup-continue en tout point tel que : } D_+^*(x) \cap R_C(x) \neq \emptyset$$

I.2.2.2.a) <u>Première tentative.</u>

Connaissant Γ le choix le plus simple pour Γ_C est :

$$\forall x \in \mathbb{R}^n : \Gamma_C(x) = \Gamma(x) \cap R_C(x)$$

si (1C) est vérifié ; (2C) ne l'est pas et surtout (3C) n'est pas vérifié, car R_C n'est pas en général sup-continue au point x (I.2.1 remarque 3).
Ceci implique que pour assurer la convergence des algorithmes de direction améliorante réalisable, il n'est pas suffisant de choisir des directions réalisables quelconques même si celles-ci sont uniformément améliorantes.
Si l'on reprend l'exemple de la remarque 3 de I.2.1 en considérant : $f(x_1,x_2) = x_1+x_2$



$\nabla f(x) = (1,1)$
$s_n.\nabla f(x_n) = 1 > 0$

Les directions $s_n = (1,0)$ sont uniformément améliorantes et réalisables ; cependant la suite de points $(x_n)_{n \in \mathbb{N}}$ considérée converge vers $\overset{*}{x} = (1,0)$ qui n'est pas stationnaire.
Soit :

$$\Gamma(x) = \{\alpha.s/\alpha \in \mathbb{R}^+\}$$
$$\Gamma_C(x) = \Gamma(x) \cap R_C(x)$$
$$\forall n \in \mathbb{N} : s_n = s \in \Gamma_C(x_n) \text{ mais } s \notin \Gamma_C(\overset{*}{x})$$
$$\Gamma_C(\overset{*}{x}) = \emptyset \text{ mais } \overset{*}{x} \text{ n'est pas stationnaire.}$$

Les propriétés (2C) et (3C) ne sont pas vérifiées en $\overset{*}{x}$. Il se produit un phénomène de "blocage" à la rencontre de la contrainte $x_1 \leq 1$.
WOLFE [40] a fourni un exemple dans $\mathbb{R}^3$, où :

- les directions sont uniformément améliorantes : ce sont celles du gradient
- les directions sont réalisables
- chaque successeur est obtenu par maximisation sur la partie réalisable de la direction choisie
- la suite des itérés converge vers un point non stationnaire.

Dans cet exemple le blocage sur une contrainte est remplacé par un blocage alternatif sur deux contraintes avec des déplacements dont la longueur tend vers 0. C'est le phénomène de "Zig-Zag" étudié initialement par ZOUTENDIJK cause de non convergence de nombreux algorithmes de direction améliorante réalisable.

I.2.2.2b) Deuxième tentative

Pour tout δ de $\mathbb{R}^{+*}$ on définit : $R_C(x,\delta) = \{s \in B_1 \,/\, [x,x+\delta s] \subset C\}$ où $B_1 = \{s \in \mathbb{R}^n \,/\, ||s|| = 1\}$

Proposition : $\forall\delta > 0$; $R_C(x,\delta)$ est une fonction multivoque sup-continue sèlon x dans C fermé

Dem :

$$\left.\begin{array}{l} (x_n)_{n\in\mathbb{N}} \to x \\ (s_n)_{n\in\mathbb{N}} \to s \\ \forall n,\ s_n \in R_C(x_n,\delta) \end{array}\right\} \Rightarrow [x,x+\delta s] = \lim_{n\to\infty} ([x_n,x_n+\delta.s_n])$$

or $\forall\alpha \in [0,1]$

$$\left.\begin{array}{l} x+\alpha\delta s = \lim_{n\to\infty} (x_n+\alpha.\delta s_n) \\ C \text{ fermé} \end{array}\right\} \Rightarrow [x,x+\delta s] \subset C$$

donc : $s \in R_C(x,\delta)$ C.Q.F.D

Comme $||s|| = 1$, les directions de $R_C(x,\delta)$ sont réalisables sur un segment de longueur supérieure à δ ; à ce titre on peut les appeler "uniformément réalisables".
Etant donné Γ vérifiant (1)(2) et (3) et $\delta > 0$, si :

$$\Gamma_C(x) = \Gamma(x) \cap R_C(x,\delta)$$

Γ_C ainsi définie vérifie (1C) et (3C) comme intersection de deux fonctions sup-continues ; mais elle ne vérifie pas (2C) car :

$$\left.\begin{array}{l} [\Gamma(x) = \emptyset] \Rightarrow [D_+^*(x) = \emptyset] \\ \Gamma_C(x) = \Gamma(x) \cap R_C(x,\delta) \\ \Gamma_C(x) = \emptyset \end{array}\right\} \underset{\text{en général}}{\not\Rightarrow} D_+^*(x) \cap R_C(x) = \emptyset$$

Si l'on avait $\Gamma = D_+^*$ (2C) serait vérifiée par $\Gamma_C = D_+^* \cap R_C$, mais Γ_C n'est alors plus sup-continue car Γ n'est plus sup-continue, et c'est (3) qui n'est pas vérifiée.

I.2.2.2c) Troisième tentative.

Soit $\delta \in \mathbb{R}^{+*}$

$$D_+(x,\delta) = \begin{cases} \{s \in \mathbb{R}^n \,/\, \nabla f(x).s \geq \delta.||\nabla f(x)||.||s|| & \text{si } \nabla f(x) \neq 0 \\ \emptyset \text{ si } \nabla f(x) = 0 \end{cases}$$

$$R_C(x,\delta) = \{s \in B_1 \,/\, [x,x+\delta s] \subset C\}$$

$$\Gamma_\delta(x) = D_+^*(x,\delta) \cap R_C(x,\delta)$$

<u>Proposition</u> Γ_δ ainsi défini vérifie les propriétés (1C) et (3C) et :

$$[\forall\delta > 0,\ \Gamma_\delta(x) = \emptyset \Longrightarrow [D_+^*(x) \cap R_C(x) = \emptyset]$$

<u>Dem</u> : Γ_δ intersection de deux fonctions sup-continues est sup-continue en tout point non stationnaire.
Si $\nabla f(x) = 0$ alors $D_+^*(x) = \emptyset$ et l'implication est trivialement vérifiée. Si $D_+^*(x) \cap R_C(x) \neq \emptyset$ et $\nabla f(x) \neq 0$

$$\exists s\ ;\ ||s|| = 1 \quad \text{tel que :}$$

$$\left.\begin{array}{l} \dfrac{\nabla f(x).s}{||s||.||\nabla f(x)||} = \delta_1 > 0 \\[2ex] \exists\delta_2 \text{ tel que } [x, x+\delta_2 s] \subset C \end{array}\right\}$$

soit $\delta = \text{Min}\ \{\delta_1, \delta_2\} > 0$

$$s \in D_+^*(x,\delta) \cap R_C(x,\delta) = \Gamma_\delta(x)$$

donc $\exists\delta > 0$ tel que $\Gamma_\delta(x) \neq \emptyset$ C.Q.F.D

Le lemme suivant est une propriété analogue à (2C) et servira en I.2.2.4 à assurer la convergence d'un algorithme général. DENEL utilise le même lemme dans un cadre plus général [10].

Soit N une sous-suite quelconque de $\mathbb{N}$, $(\delta_n)_{n\in N}$ une suite de scalaires positifs et $(x_n)_{n\in N}$ une suite de points de C.

<u>Lemme</u> : si : $(\delta_n)_{n\in N} \to 0$; $(x_n)_{n\in N} \to \overset{*}{x}$; $\Gamma_{\delta_n}(x_n) = \emptyset\ \forall n \in N$
alors : $D_+^*(\overset{*}{x}) \cap R_C(\overset{*}{x}) = \emptyset$

<u>Dem</u> : d'après le résultat de la proposition précédente il suffit de montrer que sous les hypothèses faites : $\forall\delta > 0,\ \Gamma_\delta(\overset{*}{x}) = \emptyset$
Si ce n'est pas le cas :

$$\exists\delta > 0 \text{ tel que : } \Gamma_\delta(\overset{*}{x}) \neq \emptyset$$

$$\exists\delta > 0,\ \exists s \in \mathbb{R}^n,\ ||s|| = 1 \text{ tels que :}$$

$$\frac{\nabla f(\overset{*}{x}).s}{||s||.||\nabla f(\overset{*}{x})||} \geq \delta > 0 \text{ et } [\overset{*}{x}, \overset{*}{x}+\delta s] \subset C$$

a) $$\frac{\nabla f(\overset{*}{x}).s}{||s||.||\nabla f(\overset{*}{x})||} > \frac{\delta}{2} > 0$$

soit pour tout $n \in N$: $\tilde{s}_n = \dfrac{\overset{*}{x}+\delta.s-x_n}{\delta}$

par construction $(\tilde{s}_n)_{n \in N} \to s$ quand $(x_n)_{n \in \mathbb{N}} \to \overset{*}{x}$, $||\tilde{s}_n||$ est donc non nul à partir d'un certain rang.
Par continuité de ∇f et du produit scalaire :

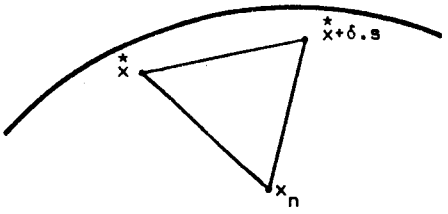
$\exists n_1 \in N$ tel que $\forall n \in N$, $n \geq n_1$: $s_n = \dfrac{\tilde{s}_n}{||\tilde{s}_n||}$ donc $||s_n|| = 1$

$$\nabla f(x_n).s_n \geq \frac{\delta}{2}.||\nabla f(x_n)||.||s_n|| > o$$

$\exists n_1 \in N$ tel que, $\forall n \in N$, $n \geq n_1$:

$$\boxed{s_n \in D^*_+(x_n, \delta/2)}$$

b) $$\left.\begin{array}{l} C \text{ convexe} \\ x_n \in C \\ \overset{*}{x} + \delta s \in C \end{array}\right\} \Rightarrow [x_n, \overset{*}{x}+\delta s] \subset C, \ \forall n \in N$$



le choix fait de s_n assure que :

$\forall n \geq n_1$

$$x_n+\delta.\tilde{s}_n = x_n+\delta.||\tilde{s}_n||.\frac{\tilde{s}_n}{||\tilde{s}_n||} = x_n+\delta.||\tilde{s}_n||.s_n = \overset{*}{x}+\delta.s$$

Soit $\delta_n = \delta.||\tilde{s}_n||$

$$\lim_{n\to\infty} (\delta_n) = \delta.\lim_{n\to\infty} ||\tilde{s}_n|| = \delta.||s|| = \delta$$

donc $\exists n_2 \in \mathbb{N}$ tel que $\forall n \geq n_2$: $\delta_n > \dfrac{\delta}{2}$

Soit $n_3 = \max(n_1, n_2)$; $\forall n \geq n_3$:

$$[x_n, x_n + \frac{\delta}{2}.s_n] \subset [x_n, x_n + \delta_n.s_n] \subset C$$

Soit :

$$\boxed{s_n \in R_C(x_n, \frac{\delta}{2})}$$

a) et b) entrainent alors : $\forall n \in N,\ n \geq n_3$:

$$s_n \in D_+^*(x_n, \frac{\delta}{2}) \cap R_C(x_n, \frac{\delta}{2})$$

$$s_n \in \Gamma_{\delta/2}(x_n)$$

en contradiction avec l'hypothèse $\Gamma_{\delta_n}(x_n) = \emptyset,\ \forall n \in N$
c'est donc que : $\forall \delta > 0,\ \Gamma_\delta(\overset{*}{x}) = \emptyset$ ce qui entraine :

$$D_+^*(\overset{*}{x}) \cap R_C(\overset{*}{x}) = \emptyset$$

I.2.2.3 Recherche de Δ_C

Le problème est de pouvoir remplacer une fonction multivoque Δ vérifiant :

(1)' : $x \in \Delta(x,s) = \overline{\Delta(x,s)} \subset D(x,s)$
(2)' : $T[x,\Delta(x,s)] = T[x,D(x,s)]$
(3)' : Δ continue

par une fonction Δ_C vérifiant :

(1C)' : $x \in \Delta_C(x,s) = \overline{\Delta_C(x,s)} \subset D(x,s) \cap C$
(2C)' : $T[x,\Delta_C(x,s)] = T[x,D(x,s)]$
(3C)' : Δ_C continue

I.2.2.3a) Première tentative

La solution la plus simple consiste à choisir :

$$\Delta_C(x,s) = \Delta(x,s) \cap C$$

alors (1C)' est trivialement vérifié, mais (2C)' et surtout (3C)' ne sont pas vérifiés en général.

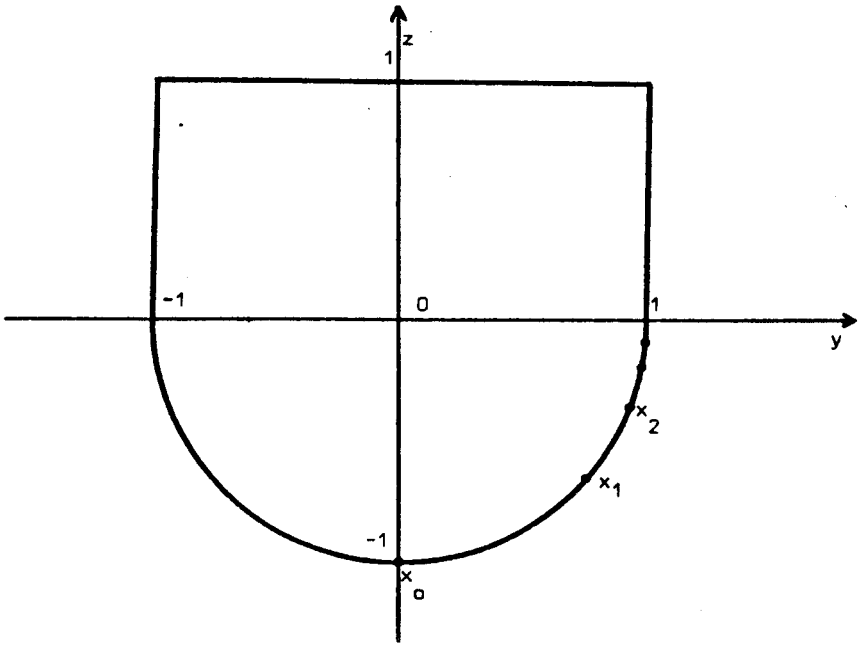
L'intersection de deux fonctions multivoques inf-continues n'est pas en général inf-continue [20] et l'on ne peut assurer l'inf-continuité de : $\Delta \cap C$.
C'est en particulier vrai pour l'intersection d'une demi-droite avec un domaine C, même si celui-ci est convexe compact. Ceci est mis en évidence dans l'exemple suivant dû à TOPKIS et VEINOTT [37]

Exemple : Maximiser $f(y,z) = z$ sur

$$C = \{(y,z) \in \mathbb{R}^2 / |y| \leq 1,\ |z| \leq 1,\ z+(1-y^2)^{1/2} \geq 0\}$$

soit :

$$x_{n-1} = (\frac{1}{n}.(n^2-1)^{1/2},\ -\frac{1}{n})$$



à chaque itération :

- la direction de recherche est uniformément améliorante
- la direction est réalisable
- le successeur d'un point résulte d'une maximisation sur la partie réalisable de la demi-droite choisie

Cependant la suite des itérés converge vers (1,0) qui n'est pas stationnaire.
Si $\Delta_C(x,s) = D(x,s) \cap C$ on peut vérifier que cette fonction n'est pas inf-continue en $x = (1,0)$

I.2.2.3b) Deuxième tentative

Soit $\delta \in \mathbb{R}^{+*}$ fixé alors :

$$[x,x+\delta s] \text{ et } \{x + \frac{\delta}{2^i} s / i \in \mathbb{N}\} \cup \{x\}$$

sont deux fonctions multivoques qui vérifient (3C)' et (2C)'. De plus si on note Δ_δ indifféremment l'une ou l'autre de ces fonctions, elle vérifie la propriété suivante analogue à (1C)' : $\forall s \in \Gamma_\delta(x)$ définie en I.2.2.2c)

$$x \in \Delta_\delta(x,s) = \overline{\Delta_\delta(x,s)} \subset D(x,s) \cap C$$

I.2.2.4 Un algorithme général de direction améliorante réalisable.

On va utiliser les résultats de I.2.2.2c) et I.2.2.3b) obtenus pour Γ_C et Δ_C respectivement, afin de construire un algorithme utilisant la notion de direction uniformément améliorante et uniformément réalisable.

Hypothèses f continument différentiable
$x_o \in C$ convexe fermé de $\mathbb{R}^n$
$E_o = \{y \in C / f(y) \geq f(x_o)\}$ compact

Notations

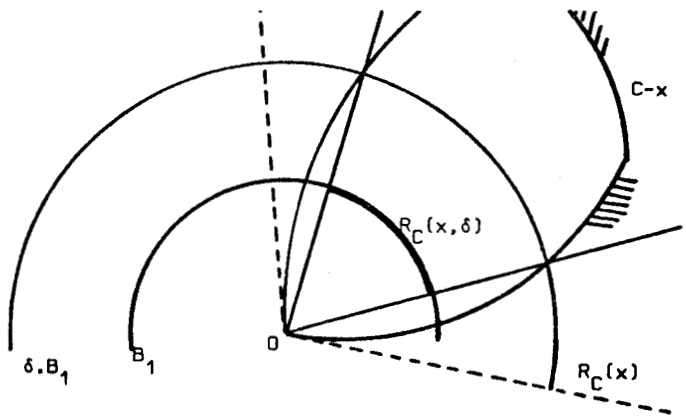
$$D_+^*(x,\delta) = \begin{cases} \{s \in \mathbb{R}^n \ / \ \nabla f(x).s \geq \delta. \ ||\nabla f(x)||.||s||\} \text{ si } \nabla f(x) \neq 0 \\ \phi \text{ si } \nabla f(x) = 0 \end{cases}$$

$R_C(x,\delta) = \{s \in B_1 \ / \ [x,x+\delta s] \subset C\}$

$\Gamma_\delta(x) = D_+^*(x,\delta) \cap R_C(x,\delta)$

$\Delta_\delta(x,s) = [x,x+\delta.s]$

$M_\delta(x,s,\varepsilon)$ = ensemble des maximum de f sur $\Delta_\delta(x,s)$ à ε près.



Algorithme itération 0 : $x_o \in C$, $\delta_o > 0$ et $\varepsilon_o \geq 0$

itération k : connus x_k, δ_k, ε_k

(α) si $\Gamma_{\delta k}(x_k) = \emptyset$ faire $x_{k+1} = x_k$

$\delta_{k+1} = \delta_k/2$

$\varepsilon_{k+1} = \varepsilon_k$ retour en (α)

sinon : soit $s_k \in \Gamma_{\delta k}(x_k)$

(β) si $x_k \in M_{\delta k}(x_k,s_k,\varepsilon_k)$ faire $x_{k+1} = x_k$

$\varepsilon_{k+1} = \varepsilon_k/2$

$\delta_{k+1} = \delta_k$ retour en (β)

sinon : $x_{k+1} \in M_{\delta k}(x_k,s_k,\varepsilon_k)$

faire $\varepsilon_{k+1} = \varepsilon_k/2$

$\delta_{k+1} = \delta_k$ retour en (α).

Résultat la suite $(x_n)_{n \in \mathbb{N}}$ engendrée par l'algorithme contient une sous-suite $(x_n)_{n \in \mathbb{N}}$ parfaitement déterminée dont les points d'accumulation sont stationnaires.
Si de plus f est concave tout point d'accumulation de la suite $(x_n)_{n \in \mathbb{N}}$ est un maximum de f sur C.

Dem :

a) si l'éventualité $\Gamma_{\delta_k}(x_k) = \emptyset$ ne se produit qu'un nombre fini de fois

$$\exists K \in \mathbb{N} \text{ tel que } \forall k \in \mathbb{N},\ k \geq K : \delta_k = \delta > 0$$

$$\forall k \geq K : \quad \begin{array}{l} s_k \in \Gamma_\delta(x_k) \\ x_{k+1} \in M_\delta(x_k, s_k, \varepsilon_k) \\ \varepsilon_{k+1} = \varepsilon_k/2 \end{array}$$

soit $(\overset{*}{x}, \overset{*}{s}, 0)$ un point d'accumulation quelconque de $(x_n, s_n, \varepsilon_n)_{n \in \mathbb{N}}$. Il existe de tels points car la suite $(f(x_n))_{n \in \mathbb{N}}$ est non décroissante et E_0 compact.
Supposons $\overset{*}{x}$ non stationnaire :

$$\left.\begin{array}{l} s_k \in \Gamma_\delta(x_k) \\ \Gamma_\delta \text{ sup-continue} \\ \text{par I.2.2.2b)} \end{array}\right\} \Longrightarrow (\overset{*}{s} \in \Gamma_\delta(\overset{*}{x})) \Longrightarrow (\nabla f(\overset{*}{x}).\overset{*}{s} \geq \delta.||\nabla f(\overset{*}{x})||.||\overset{*}{s}||)$$

$$\left.\begin{array}{l} x_{k+1} \in M(x_k, s_k, \varepsilon_k) \\ M_\delta \text{ est sup-continue} \\ \text{car } \Delta_\delta \text{ est continue} \\ (f(x_k))_{k \in N} \text{ non décroissante} \\ \forall k \in \mathbb{N},\ x_k \in \overset{0}{E} \text{ compact} \end{array}\right\} \Longrightarrow (\overset{*}{x} \in M_\delta(\overset{*}{x}, \overset{*}{s}, 0)) \Longrightarrow (\nabla f(\overset{*}{x}).\overset{*}{s} \leq 0)$$

en remarquant que $||\nabla f(\overset{*}{x})||$ et $||\overset{*}{s}||$ sont non nuls, les deux inégalités obtenues sont en contradiction. C'est donc que $\overset{*}{x}$ est stationnaire et dans ce cas le résultat est démontré, la suite parfaitement déterminée étant ici $\mathbb{N}$.

b) si l'éventualité $\Gamma_{\delta_k}(x_k) = \emptyset$ se produit un nombre infini de fois ; soit $N' \subset \mathbb{N}$ la suite infinie telle que :

$$\forall k \in N' \quad : \quad \Gamma_{\delta_k}(x_k) = \emptyset$$

$$\forall k \in N' \quad : \quad \delta_{k+1} = \delta_k/2$$

par construction la suite $(\delta_k)_{k \in \mathbb{N}}$ tend vers 0.
On montre alors que tout point d'accumulation de la suite $(x_k)_{k \in N'}$ vérifie :

$$\forall \delta > 0,\ \Gamma_\delta(\overset{*}{x}) = \emptyset$$

Soit $\overset{*}{x}$ un tel point d'accumulation : $\exists N \subset N' \subset \mathbb{N}$ sous-suite infinie telle que : $(x_k)_{k \in N} \to \overset{*}{x}$ alors :

$$(\delta_k)_{k \in N} \to 0 \ ;\ (x_k)_{k \in N} \to \overset{*}{x} : \Gamma_{\delta_k}(x_k) = \emptyset,\ \forall k \in N$$

toutes les hypothèses du lemme démontré en I.2.2.2c sont vérifiées, l'application de celui-ci donne le résultat :

$$D_+^*(x) \cap R_C(\overset{*}{x}) = \emptyset$$

et le point $\overset{*}{x}$ est stationnaire.

Le résultat énoncé est également vrai dans ce cas la suite parfaitement déterminée étant celle constituée des indices tels que : $\Gamma_{\delta_k}(x_k) = \emptyset$ c'est à dire N'.

c) si de plus f est concave

Soit $\overset{**}{x}$ un point d'accumulation de $(x_n)_{n \in \mathbb{N}}$

$\overset{*}{x}$ un point d'accumulation défini en (a) ou (b)

la suite : $(f(x_n))_{n \in \mathbb{N}}$ est monotone non décroissante toutes les suites extraites ont même limite

$$f(\overset{**}{x}) = f(\overset{*}{x})$$

$\overset{*}{x}$ est tel que : $D_+^*(\overset{*}{x}) \cap R_C(\overset{*}{x}) = \emptyset$ quand f est concave cette condition est une condition suffisante d'optimalité (I.2.1 remarque 5)

$$\left.\begin{array}{l} \overset{*}{x} \text{ maximum de } f \text{ sur } C \\ \\ f(\overset{*}{x}) = f(\overset{**}{x}) \end{array}\right\} \Rightarrow \overset{**}{x} \text{ maximum de } f \text{ sur } C \qquad \text{C.Q.F.D}$$

Remarque 1. Les mêmes résultats sont obtenus si dans l'algorithme, on fait :

$$\Delta_\delta(x,s) = \{x + \frac{\delta}{2^i} s / i \in \mathbb{N}\} \cup \{x\}$$

les propriétés utilisées étant vérifiées en I.2.2.3b)

Remarque 2 La recherche du maximum selon la direction s se limite dans l'algorithme à un ensemble $\Delta_\delta(x,s)$ avec :

$$\Delta_\delta(x,s) = [x, x+\delta s] \text{ dans le cas continu}$$
$$\Delta_\delta(x,s) = \{x + \frac{\delta}{2^i} s / i \in \mathbb{N}\} \cup \{x\} \text{ dans le "cas discret"}$$

Si on a limité la recherche du maximum à de tels ensembles c'est pour assurer la sup-continuité de $M_\delta(x,s,\varepsilon)$ associé.

En fait on peut considérer

$$\Delta'_\delta(x,s) = D(x,s) \cap C$$

ou

$$\Delta'_\delta(x,s) = \{x\} \cup [\{x + \frac{\delta}{2^i}.s \ / \ i \in \mathbb{Z}\} \cap C]$$

si $\bar{y}$ est un successeur de x maximisant f à ε près sur l'un de ces ensembles alors que y est un successeur trouvé dans $M_\delta(x,s,\varepsilon)$:

à chaque itération $f(\bar{y}) \geq f(y)$, la suite $(f(y))$ converge vers la valeur maximum de f sur C par le résultat de l'algorithme, il en est de même pour $(f(\bar{y}))$.

Remarque 3 On peut obtenir un algorithme très voisin de cet algorithme en utilisant le formalisme des familles ρ-décroissantes introduites par DENEL [10]. Cet algorithme et les résultats qui s'y rapportent sont décrits en annexe II.

I.2.3 Exemples d'algorithmes

Les algorithmes décrits en I.2.2.1 et I.2.2.4 sont, par construction, des adaptations simples de l'algorithme donné en I.1.5 pour la maximisation sans contraintes.
Ceux-ci présentent essentiellement un intérêt théorique. En effet, le premier requiert la connaissance de deux fonctions multivoques Γ_C et Δ_C vérifiant les propriétés (1C) (2C) (3C) et (1'C) (2'C) (3'C), or on ne pourra pas toujours mettre en évidence des fonctions possèdant ces propriétés, dans les algorithmes particuliers.
Le second algorithme utilise la notion de direction uniformément réalisable, mais l'ensemble $\Gamma_\delta(x) = D_+^*(x,\delta) \cap R_C(x,\delta)$ ne peut en général, s'expliciter "facilement".
Cependant les problèmes rencontrés en I.2.2, tels que le blocage ou le phénomène de zig-zag, se retrouveront dans tous les algorithmes de direction améliorante et réalisable.
Les principes des différents algorithmes seront ici décrits par analogie avec I.2.2, sans que soient données les démonstrations de convergence, celles-ci se trouvant dans de nombreux ouvrages généraux tels que : ZANGWILL [41] ou LUENBERGER [25]. Une approche des problèmes de convergence est faite par HUARD [20] à l'aide d'un algorithme général ; alors que KLESSIG étudie d'une manière globale les procédés anti-zig-zag [23].

Les algorithmes de direction améliorante réalisable se divisent en deux familles qui diffèrent par la manière dont est recherchée une direction améliorante réalisable.

- les méthodes de plus forte pente sont des généralisations au cas de la maximisation avec contraintes, des algorithmes de I.1.4.f
  Dans ces méthodes il sera nécessaire, à chaque itération, de résoudre un problème linéaire afin de trouver "une meilleure direction" compte tenu des contraintes.
- les méthodes de projection consistent à projeter sur le domaine une direction améliorante afin de la rendre réalisable.

Dans la plupart de ces algorithmes, lorsqu'une direction améliorante et réalisable sera déterminée, la deuxième phase de chaque itération consistera à maximiser la fonction économique sur la partie réalisable dans la direction choisie.

I.2.3.1 Méthode de FRANK et WOLFE [14]

Elle résoud des problèmes de la forme :

$$P : \quad \boxed{\begin{array}{l} \text{Max } f(x) \\ x \in C \end{array}} \qquad \begin{array}{l} f \text{ continument différentiable} \\ C \text{ polyèdre compact} \end{array}$$

à tout point réalisable x, on associe le sous-problème :

$$P(x) : \quad \boxed{\begin{array}{l} \text{Max } \nabla f(x).z \\ z \in C \end{array}}$$

soit $M(x)$ l'ensemble des solutions de ce programme linéaire et

$$\Gamma_C(x) = \begin{cases} \{(\overset{*}{z}-x) \ / \ \overset{*}{z} \in M(x)\} \text{ si } x \notin M(x) \\ \emptyset \text{ si } x \in M(x) \end{cases}$$

Γ_C ainsi définie sur C, vérifie (1C) (2C) et (3C).

I.2.3.2 Méthode de TOPKIS et VEINOTT [37]

Elle résoud des problèmes de la forme :

$$P : \quad \boxed{\begin{array}{l} \text{Max } f(x) \\ g_i(x) \geq 0, \ \forall i \in I \end{array}}$$

I un ensemble fini d'indices
f, g_i continument différentiable $\forall i \in I$

à tout x réalisable on associe le sous-problème :

$$P(x) : \quad \boxed{\begin{array}{l} \text{Max } w \\ \nabla g_i(x).s-\omega \geq g_i(x), \ \forall i \in I \\ \nabla f(x).s-w \geq 0 \end{array}}$$

soit $\overset{*}{\omega}$ la valeur optimale de $P(x)$ et $M(x)$ l'ensemble des maximum

$$\text{on note } \Gamma_C(x) = \begin{cases} \{M(x) \text{ si } \overset{*}{\omega} \neq 0 \\ \emptyset \text{ si } \overset{*}{\omega} = 0 \end{cases}$$

soit $I(x) = \{i \in I \ / \ g_i(x) = 0\}$ indices des contraintes actives en x, on dira par abus de langage que les contraintes actives sont indépendantes si et seulement si les vecteurs $(\nabla g_i(x))_{i \in I(x)}$ sont indépendants.

On note H l'hypothèse suivante appelée, hypothèse de qualification des contraintes ou de non-dégénérescence

H : $\forall x$ réalisable, les contraintes actives en x sont indépendantes

Sous l'hypothèse H on peut alors montrer que :

Γ_C vérifie (1C) (2C) et (3C).

I.2.3.3 Méthode de ZOUTENDIJK [42]

Sous les mêmes hypothèses, avec les mêmes notations que dans la méthode précédente, une tentative naturelle est d'essayer de réduire la taille du sous-problème $P(x)$ en ne considérant que les contraintes actives en x. Une telle tentative conduit à un algorithme soumis au phénomène de zig-zag. Zoutendijk dans l'un de ses algorithmes utilise les éléments suivants :

- $I(x,\varepsilon) = \{i \in I / g_i(x) \leq \varepsilon\}$ indices des contraintes actives à ε près

- $P(x)$:

$$\boxed{\begin{array}{l} \text{Max } w \\ \nabla g_i(x).s - w \geq 0, \ \forall i \in I(x,\varepsilon) \\ \nabla f(x).s - \omega \geq 0 \\ |s_i| \leq 1, \ \forall i = 1,\dots,n \end{array}}$$

- Γ_ε fonction définie comme Γ_C dans la méthode précédente, mais qui dépend ici de ε.

- l'assurance que ε tend vers zéro au cours des itérations est obtenue en divisant ε par 2, à chaque fois que la valeur optimale ω de $P(x)$ devient inférieure à ε.

alors on montre que :

$$\forall \varepsilon > 0, \ \Gamma_\varepsilon \text{ vérifie (1C)}$$

et une propriété analogue à celle démontrée en I.2.2.2c) :

$$[\exists \varepsilon > 0, \ \Gamma_\varepsilon(x) = \emptyset] \Rightarrow x \text{ est stationnaire.}$$

<u>Dem</u> : le principe de la démonstration est le suivant :
pour ε "suffisamment petit" $I(x,\varepsilon) = I(x,0) = I(x)$. 0 est alors la valeur optimale de $P(x)$ où l'ensemble des indices est celui des contraintes actives en x. On montre alors, sous l'hypothèse H que les conditions nécessaires de Kühn et Tucker sont vérifiées en x.

I.2.3.4 <u>Algorithme de Projection.</u>

Le problème à résoudre est :

$$P : \quad \boxed{\begin{array}{l} \text{Max } f(x) \\ x \in C \end{array}}$$

C est un compact convexe de $\mathbb{R}^n$
f continument différentiable

soit $\alpha > 0$ une constante fixée, on définit :

$y(x) = x + \alpha.\nabla f(x)$

$z(x) = \text{Proj}_C\{y(x)\}$ l'unique projection euclidienne du point $y(x)$ sur le convexe C

$$\Gamma_C(x) = \begin{cases} \{z(x)-x\} & \text{si } z(x) \neq x \\ \emptyset & \text{sinon} \end{cases}$$

$\Delta_C(x) = [x, z(x)]$

<u>alors</u> : Γ_C vérifie (1C) (2C) (3C)
Δ_C vérifie (1'C) (2'C) (3'C)

ces résultats sont essentiellement dûs aux propriétés suivantes qui sont démontrées au chapitre II :

- $[z(x)=x] \Longleftrightarrow$ x est stationnaire
- z(x) est une fonction univoque continue de x

Remarque : la détermination de z(x) peut s'effectuer dans le cas général où C est convexe, grâce à un algorithme établi par LAURENT [24]. Cet algorithme construit une suite de points qui converge vers z(x), aussi pratiquement on obtiendra ; en un nombre fini d'itérations, une approximation non réalisable de z(x). Ceci est peu satisfaisant car la direction déterminée risque alors, de ne plus être réalisable. Dès que C est un polyèdre, z(x) peut être obtenu en un nombre fini de calculs et l'algorithme I.2.2.1 s'applique exactement.

I.2.3.5 Méthode de ROSEN [34]

Toutes les méthodes précédemment décrites ont un inconvénient commun, c'est d'avoir à résoudre à chaque itération un problème d'optimisation auxiliaire qui est au moins linéaire.
La méthode de Rosen, appliquée à un programme non linéaire à contraintes linéaires, a pour but d'engendrer à chaque itération une direction améliorante et réalisable sans avoir à résoudre un sous-problème d'optimisation : on utilisera à chaque itération une matrice de projection qui se calcule de proche en proche.
Soit

$$P : \quad \boxed{\begin{array}{l} \text{Max } f(x) \\ A_i.x \geq b_i \\ i \in E \end{array}}$$

E ensemble d'indices fini
A la matrice dont les lignes sont les A_i quand $i \in I$,
C le domaine $\subset \mathbb{R}^n$

pour tout x réalisable on définit :

$$I(x) = \{i \in E / A_i.x = b_i\}$$
$$\bar{I}(x) = \{i \in E / A_i.x > b_i\}$$

A_I la matrice dont les lignes sont les A_i quand $i \in I(x)$
b_I le vecteur de composantes b_i, quand $i \in I(x)$
Pour θ constante positive fixée on considère :

$$P(x) : \quad \boxed{\begin{array}{l} \text{Max } \nabla f(x).s \\ A_I(x+s) = b_I \\ ||s||^2 = s.s \leq \theta \end{array}}$$

que l'on peut écrire en remarquant que seules les contraintes actives en x interviennent :

$$P(x) : \quad \boxed{\begin{array}{l} \text{Max } \nabla f(x).s \\ A_I.s = 0 \\ ||s||^2 = s.s \leq \theta \end{array}}$$

ce problème analogue à ceux rencontrés en I.1.4.f), est une recherche de meilleure direction améliorante, mais relativement à la variété affine définie par l'intersection des contraintes actives :

$$\{y \in \mathbb{R}^n \ / \ A_I.y = b_I\}$$

Sous l'hypothèse H d'indépendance des contraintes actives en x, on peut exprimer la solution $\overset{*}{s}$ de P(x) à l'aide des conditions de Kühn et Tucker :

$$\exists u \leq 0$$
$$\exists v = (v_i)_{i \in I(x)}$$
$$\nabla f(x) + 2u.\overset{*}{s} + \sum_{i \in I(x)} v_i.A_i = 0$$
$$u(\theta - \overset{*}{s}.\overset{*}{s}) = 0$$

<u>si</u> u = 0 alors

$$\nabla f(x) = - \sum_{i \in I(x)} v_i.A_i = -A_I^t.v$$
$$\nabla f(x).\overset{*}{s} = -v.A_i\overset{*}{s} = 0 \text{ car } \overset{*}{s} \text{ est réalisable pour } P(x)$$
$$\nabla f(x).\overset{*}{s} = 0$$

et 0 est solution de P(x)

<u>sinon</u> :

$$\overset{*}{s} = - \frac{1}{2u}(\nabla f(x) + A_I^t.v)$$

$$A_I.\overset{*}{s} = - \frac{1}{2u} (A_I.\nabla f(x) + A_I.A_I^t.v) = 0$$

sous l'hypothèse H la matrice $A_I.A_I^t$ de rang maximum est inversible ; on obtient :

$$v = (A_I.A_I^t)^{-1}.A_I.\nabla f(x)$$
$$\overset{*}{s} = \nabla f(x) + A_I^t.v = [U - A_I^t(A_I.A_I^t)^{-1}.A_I].\nabla f(x)$$

où U désigne la matrice unité de rang n.
$P_I = U - A_I^t(A_I.A_I^t)^{-1}.A_I$ est la matrice de projection associée à la variété linéaire $\{y \in \mathbb{R}^n / A_I.y = 0\}$

$$\nabla f(x) = \overset{*}{s} - A_I^t.v = \overset{*}{s} + \overset{*}{r}$$

est la décomposition du gradient sur les sous-espaces vectoriels supplémentaires orthogonaux :

$$\{y \in \mathbb{R}^n \ / \ A_I.y = 0\} \text{ et } \{\sum_{i \in I(x)} \lambda_i \ A_i / \lambda_i \in \mathbb{R}\}$$

<u>Propriété</u> Si $\overset{*}{s}$ est non nul, alors $\overset{*}{s} \in D_+^*(x) \cap R_C(x)$

<u>Dem</u> : $\overset{*}{s}$ réalisable pour P(x) $\Rightarrow A_I.\overset{*}{s} = 0$

$\Rightarrow \forall i \in I(x) \ ; \ A_i.\overset{*}{s} \geq 0$

$\Rightarrow \overset{*}{s} \in T(x,C)$

$\Leftrightarrow \overset{*}{s} \in R_C(x)$ d'après I.2.1.3 car C est un polyèdre

$\nabla f(x).\overset{*}{s} = (\overset{*}{s}+\overset{*}{r}).\overset{*}{s} = \overset{*}{s}.\overset{*}{s} > 0$ donc $\overset{*}{s} \in D_+^*(x)$

<u>Propriété</u> Si $\overset{*}{s}$ est nul, alors :

soit x est un point stationnaire pour P
soit on exhibe $\overset{**}{s} \in D_+^*(x) \cap R_C(x)$

<u>Dem</u> : $(\overset{*}{s} = 0) \Rightarrow \nabla f(x) = -A_I^t.v = \sum_{i \in I(x)} -v_i.A_i$

- si $\forall i \in I(x)$; $-v_i \leq 0$

$$\left.\begin{array}{l} \nabla f(x) = \sum_{i \in I(x)} -v_i.A_i \\ \forall i \in I(x), -v_i \leq 0 \end{array}\right\} \Longleftrightarrow \nabla f(x) \in \Pi[T(x,C)]$$

conditions nécessaires d'optimalité (I.2.1.4)

- sinon $\exists j \in I(x)$ tel que $-v_j > 0$

on note : $I(x) - j = \{i \in I(x) \ / \ i \neq j\}$

A_{I-j} = la matrice de lignes A_i, $i \in I(x)-j$

$P_{I-j} = U - A_{I-j}^t (A_{I-j}.A_{I-j}^t)^{-1}.A_{I-j}$

la matrice de projection associée

$\overset{**}{s} = P_{I-j}.\nabla f(x)$

$\nabla f(x) = \overset{**}{s} + \overset{**}{r}$

$\overset{**}{s} \neq 0$ sinon : $\overset{*}{s}+\overset{*}{r} = \overset{****}{s+r} = \overset{*}{r} = \overset{**}{r}$

la dernière égalité donnant une combinaison linéaire nulle à coefficients non tous nuls $(-v_j > 0)$ des lignes A_i, $i \in I(x)$: contraire à *H*.
On en déduit :

$$\nabla f(x).\overset{**}{s} = (\overset{****}{s+r}).\overset{**}{s} = \overset{****}{s.s} > 0 \text{ et } \overset{**}{s} \in D_+^*(x)$$

$$\nabla f(x).\overset{**}{s} = (\overset{*}{s}+\overset{*}{r}).\overset{**}{s} = \overset{*}{r}.\overset{**}{s} = -A_I^t v\overset{**}{s} = -(v_j A_j + \sum_{i \in I(x)-j} v_i A_i).\overset{**}{s}$$

$$= -v_j A_j \overset{**}{s}$$

or $-v_j > 0$ avec ce qui précède : $\left.\begin{array}{l} A_j.\overset{**}{s} > 0 \\ \text{avec} \quad A_{I-j}\overset{**}{s} = 0 \end{array}\right\}$

au total : $(A_I.\overset{**}{s} \geq 0) \Rightarrow (\overset{**}{s} \in R_C(x))$

La méthode de Rosen fournit donc en tout point x non stationnaire une direction réalisable améliorante notée s (c'est $\overset{*}{s}$ ou $\overset{**}{s}$). La simplicité de la méthode de Rosen et son avantage par rapport à toutes les autres méthodes de direction améliorante réalisable, réside dans le fait qu'il ne faudra pas effectivement résoudre le problème intermédiaire P(x). La matrice de projection P_I associée à I(x), peut se calculer d'une manière itérative quand un seul indice s'ajoute ou se retranche à I(x) ; on utilise pour cela les formules d'inversion d'une matrice partitionnée. Quand une direction améliorante et réalisable est déterminée la deuxième phase de chaque itération consiste à maximiser la fonction économique sur la partie réalisable dans cette direction.

Si on pose :

$$\Gamma_C(x) = \begin{cases} \{s\} \text{ si } \overset{*}{s} \neq 0 \text{ et } v \not\leq 0 \\ \emptyset \text{ sinon} \end{cases}$$

les démonstrations faites montrent que Γ_C vérifie (1C) et (2C).
Par contre (3C) n'est pas vérifiée car s ne varie pas continument avec x : en effet, dès qu'une nouvelle contrainte devient active, s peut varier brusquement.
Il en résulte que, sous cette forme, la convergence de la méthode de Rosen n'est pas assurée.
Celle-ci est obtenue d'une part par POLAK [29] grâce à un procédé anti-zig-zag où l'on ne considère que les contraintes actives à ε près, avec ε qui tend vers zéro ; soit par HUARD [19] où ε reste constant et qui consiste à conserver la même direction de gradient dès que le déplacement obtenu à une itération devient inférieure à ε.
Par sa simplicité la méthode du gradient projeté de Rosen reste très employée et jusqu'à ce jour aucun exemple de zig-zag n'a été mis en évidence.

I.2.4 Contraintes non linéaires et procédés de rentrée.

Les méthodes de direction améliorante et réalisable qui utilisent le principe de la projection sont particulièrement bien adaptées au cas de programmes à contraintes linéaires.
En effet dès que C est un polyèdre, étant donné x réalisable :

- on peut obtenir une direction améliorante dans T(C,x) à l'aide d'une projection sur C ou sur T(C,x) lui-même
- d'autre part $T(C,x) = R_C(x)$ (I.2.1.3)

on peut donc déterminer en tout point x non stationnaire une direction améliorante et réalisable.

Si C est un convexe quelconque déterminé par des contraintes non linéaires on peut adapter le premier des points précédents de la manière suivante :

Soit $C = \{y \in \mathbb{R}^n / g_i(y) \geq 0,\ i \in I\}$ g_i concave continument différentiable $\forall i \in I$
on suppose $\exists y_0 \in \mathbb{R}^n$ tel que $g_i(y_0) > 0,\ \forall i \in I$
sous cette hypothèse, si on désigne par C'(x) le linéarisé du domaine C relativement à x :

$$C'(x) = \{y \in \mathbb{R}^n / g_i'(x,y) = g_i(x) + g_i(x).(y-x) \geq 0,\ \forall i \in I\}$$

alors :

$$T(C'(x),x) = T(\mathcal{E},x)$$

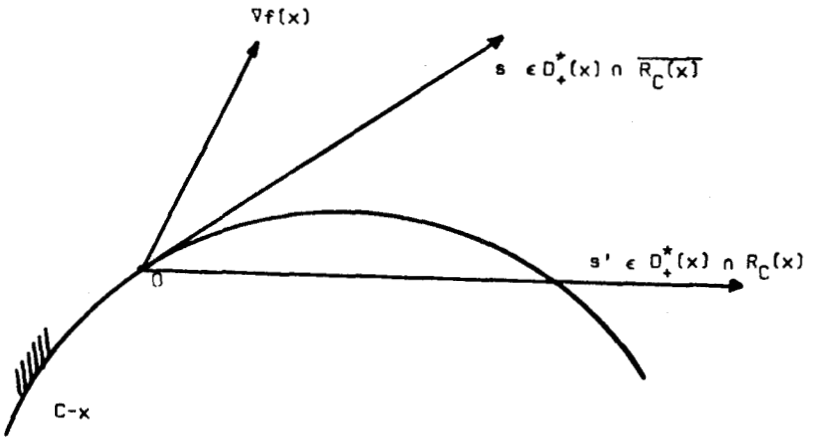
on peut alors à l'aide d'une projection sur un polyèdre trouver une direction améliorante dans T(C,x).

Par contre le deuxième point utilisé dans le cas d'un polyèdre n'est plus vérifié pour un convexe quelconque :

$$T(C,x) \neq R_C(x) \qquad (I.2.1.2)$$
$$T(C,x) = \overline{R_C(x)}$$

Finalement l'adaptation, par linéarisation du domaine, d'une méthode de projection utilisée dans le cas d'un polyèdre, permet seulement de déterminer en un point une direction améliorante mais pas forcément réalisable.

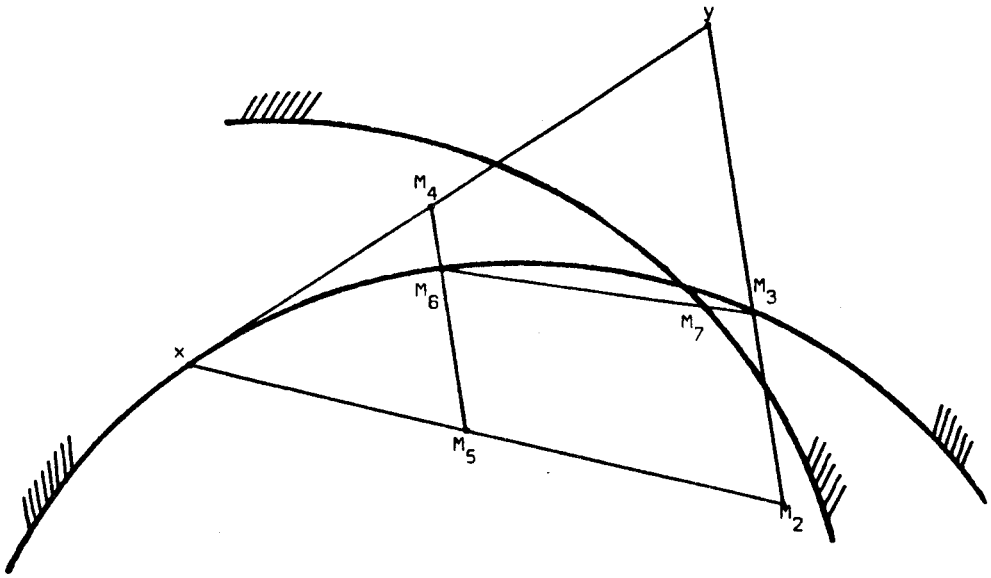
<u>Problème</u> : soit $s \in D_+^*(x) \cap \overline{R_C(x)}$
trouver $s' \in D_+^*(x) \cap R_C(x)$



Le principe "du procédé de rentrée" a été initialement utilisé par ROSEN [35] pour adapter son algorithme au cas de contraintes non linéaires. Il consiste en un point y de la demi-droite D(x,s), à utiliser une direction "pointant" vers le domaine.

cela conduit à des "déplacements" en "dents-de-scie" ou "hemistiches". Ce principe fut repris par MURTAGH et SARGENT [28] ainsi que DAVIS [9]. Sans entrer dans les détails on peut remarquer que le procédé décrit nécessite une série d'interpolations, de recherches de maximum selon une direction et de recherches de solutions d'équations non linéaires. On peut en remarquer la complexité dans la figure suivante :

BUS LILLE



La direction (x,M_7) étant finalement utilisée si elle est améliorante.
Au total le "procédé de rentrée" peut être considéré comme un artifice, dont la complexité alourdit considérablement à la fois l'étude théorique de la convergence, mais aussi les "codes" existants.
Une solution beaucoup plus simple est apportée à ce même problème, dans l'algorithme décrit au chapitre II.

BUS LILLE

CHAPITRE II

UN ALGORITHME DE POINTS RÉALISABLES POUR L'OPTIMISATION AVEC CONTRAINTES NON LINÉAIRES

II.1. Problème et hypothèses

Considérons le problème suivant :

$$P : \quad \boxed{\begin{array}{l} \text{Max } f(x) \\ x \in A \cap B = C \end{array}}$$

avec les hypothèses :

- $f : \mathbb{R}^n \to \mathbb{R}$ continument différentiable
- $A \subset \mathbb{R}^n$, convexe fermé
- $B \subset \mathbb{R}^n$, convexe compact
- $\exists \overset{o}{x} \in \overset{o}{C}$

Soit C' une fonction multivoque de C dans $\mathcal{P}(B)$ telle que :

- C' est continue
- $\forall x \in C \qquad C'(x) \supset C$
  $C'(x)$ convexe fermé
  $T(x,C'(x)) = T(x,C)$

Les hypothèses précédentes sont vérifiées dans le cas particulier où :

$A = \{y \in \mathbb{R}^n \ / \ a_i(y) \geq 0, \ \forall i \in E\}$

E sous-ensemble fini de $\mathbb{N}$

$a_i : \mathbb{R}^n \to \mathbb{R}$ concave continument différentiable pour tout i de E

$B = \{y \in \mathbb{R}^n \ / \ b_i(y) \geq 0, \ \forall i \in E'\}$; B borné

$b_i : \mathbb{R}^n \to \mathbb{R}$ concave continument différentiable pour tout i de E'

$\exists \overset{o}{x} \in \mathbb{R}^n$ tel que : $a_i(\overset{o}{x}) > 0, \ \forall i \in E$ et $b_i(\overset{o}{x}) > 0, \ \forall i \in E'$

$\forall x \in C = A \cap B$; $C'(x) = A'(x) \cap B$ avec :

$A'(x) = \{y \in \mathbb{R}^n \ / \ a_i^*(x,y) = a_i(x) + \nabla a_i(x).(y-x) \geq 0, \ \forall i \in E\}$

$A'(x)$ est le linéarisé de A par rapport au point x

La vérification des hypothèses dans ce cas particulier se fait à partir des résultats suivants :

II.1.a) <u>Lemme</u> [21]

$X = \{y \in \mathbb{R}^n / g_i(y) \geq 0, \ \forall i \in I\}$

$g_i : \mathbb{R}^n \to \mathbb{R}$ concave différentiable pour tout i de I

soit x réalisable et $I(x) = \{i \in I / g_i(x) = 0\}$ les indices des contraintes actives en x

<u>si</u> : $\exists \overset{o}{x}$ tel que : $g_i(\overset{o}{x}) > 0, \ \forall i \in I(x)$

<u>alors</u> : $\quad T(x,X) = \{y \in \mathbb{R}^n / \nabla g_i(x).y \geq 0, \ \forall i \in I(x)\}$

II.1.b) Lemme [21]

$A = \{y \in \mathbb{R}^n / a_i(y) \geq 0, \forall i \in E\}$ convexe

$A'(x) = \{y \in \mathbb{R}^n / a'_i(x,y) = a_i(x) + \nabla a_i(x).(y-x) \geq 0, \forall i \in E\}$

B un fermé de $\mathbb{R}^n$ convexe

si : $\{y \in B / a'_i(x,y) > 0, \forall i \in E\} \neq \phi$ pour tout $x \in A \cap B$

Alors : $\Gamma_1 : A \cap B \to \mathbb{R}^n$ telle que $\Gamma_1(x) = A'(x) \cap B$ est une fonction continue sur $A \cap B$.

On peut remarquer l'importance pour l'utilisation des lemmes précédents, de l'hypothèse :

$$\exists \overset{o}{x} \in \mathbb{R}^n \text{ tel que : } a_i(\overset{o}{x}) > 0, \forall i \in E \text{ et } b_i(\overset{o}{x}) > 0, \forall i \in E'$$

Cette hypothèse qui assure la non vacuité de l'intérieur du domaine C, ne peut être vérifiée dès que C est défini à l'aide de contraintes en égalités.
Cependant au cas où le domaine est défini par des inégalités et des égalités linéaires, on peut à l'aide d'une élimination de variables se ramener à un problème équivalent sous la forme choisie : c'est à dire uniquement avec des contraintes en inégalité. Cette forme sera seule utilisée par la suite.

Remarque 1 : si A est un convexe compact de $\mathbb{R}^n$, le linéarisé $A'(x)$ en un point x est un convexe fermé, mais n'est pas en général borné dans $\mathbb{R}^n$. Les hypothèses choisies et les définitions de C et de C' assurent que pour tout x, $C'(x)$ reste inclus dans le borné B ; cette propriété sera essentielle pour les résultats qui seront démontrés.

Remarque 2 : de manière pratique on peut avoir à résoudre en programmation mathématique des problèmes sous une forme générale :

$$\boxed{\begin{array}{l} \text{Max } f(x) \\ x \in A \end{array}}$$ où A est convexe fermé

Si toutes les solutions optimales sont à distance finie pour un tel programme, alors il existe un compact convexe B tel que le programme :

$$\boxed{\begin{array}{l} \text{Max } f(x) \\ x \in A \cap B \end{array}}$$ soit équivalent au premier

Dans la pratique si l'on connait une évaluation d'une solution optimale, on choisira pour B un pavé de $\mathbb{R}^n$ "suffisamment grand".

Un algorithme va être construit pour résoudre tout problème qui se met sous la forme (P) avec les hypothèses faites. Cet algorithme utilise les principes décrits au chapitre I et est fondé sur la notion de direction réalisable améliorante. Il comporte à chaque itération les phases suivantes :

- utilisation de $C'(x) = A'(x) \cap B$
- calcul d'une direction améliorante de $T(x,C)$ à l'aide de la résolution d'un sous-programme quadratique (II.2)
- recherche d'une ou plusieurs directions améliorantes réalisables. Construction à partir de celles-ci d'un sous-ensemble de C. Maximisation sur ce sous-ensemble pour trouver un successeur à x. (II.3).

L'algorithme sera décrit en (II.4) ainsi que les résultats de convergence associés.

II.2. Recherche de direction améliorante

On note $\mathcal{M}$ l'ensemble des matrices carrées d'ordre n.
Etant donnés $M \in \mathcal{M}$ et $x \in C$ un point réalisable pour le problème (P) ; on associe à ce couple le programme mathématique :

$$P(x,M) : \quad \boxed{\begin{array}{l} \text{Max } \phi_x(z) = \nabla f(x).(z-x) - \frac{1}{2}(x-z).M(x-z) \\ \\ z \in C'(x) = A'(x) \cap B \end{array}}$$

Si B est un polyèdre, le programme P(x,M) est un problème quadratique en contraintes linéaires. De nombreuses méthodes existent pour résoudre un tel problème en un nombre <u>fini</u> de calculs [32].

<u>Notations</u> SP(x,M) l'ensemble (jamais vide) des solutions optimales de P(x,M)
ϕ_x^* la valeur maximale de la fonction économique ϕ_x sur C'(x).

Les notations introduites au chapitre I sont conservées, en particulier :

$$D_+^*(x) = \{s \in \mathbb{R}^n \ / \ \nabla f(x).s > 0\}$$

$$R_C(x) = \text{ensemble des directions réalisables pour C en x}$$

On considère la fonction multivoque $\Gamma : C \times \mathcal{M} \to P(\mathbb{R}^n)$ définie par :

$$\Gamma(x,M) = \begin{cases} \{(z-x)/z \in SP(x,M)\} \text{ si } x \notin SP(x,M) \\ \\ \phi \text{ si } x \in SP(x,M) \end{cases}$$

Les propositions II.2.1 et II.2.2 suivantes montrent que la fonction multivoque Γ vérifie les propriétés (2C) et (3C) définies en I.2.2.1 pour les algorithmes de direction améliorante réalisable :

II.2.1 <u>Proposition</u> pour tout x de C :

$[\phi_x^* = 0] \Longleftrightarrow [x \in SP(x,M)] \Longleftrightarrow [\Gamma(x,M) = \phi] \Longleftrightarrow$ x stationnaire *pour* (P)

<u>Dem</u> : $\forall x$, $\phi_x(x) = 0$ donc $0 = \phi_x^*$ entraine que x est solution optimale de P(x,M). Réciproquement : $[\phi_x^* = \phi_x(x)] \Longrightarrow [\phi_x^* = 0]$. Par définition même de $\Gamma(x,M)$, cet ensemble est vide si et seulement si x est solution de P(x,M).
Si x est solution de P(x,M) il vérifie la condition nécessaire suivante, condition d'optimalité pour P(x,M) :

$$\nabla\phi_x(x) \in \Pi[T(x,C'(x))]$$

or

$$\nabla\phi_x(z) = \nabla f(x) + \frac{1}{2}(M+M^t).(x-z)$$

$$\nabla\phi_x(x) = \nabla f(x)$$

avec l'hypothèse faite : T(x,C'(x)) = T(x,C) on obtient :

$$\nabla f(x) \in \Pi[T(x,C)]$$

condition nécessaire d'optimalité pour le problème initial (P), x est donc stationnaire.

On peut ainsi reconnaitre à l'aide d'un test très simple si un point réalisable x est stationnaire pour le programme (P) ; il suffit en fait de calculer une solution optimale quelconque de P(x,M) et la valeur associée de la fonction économique.

II.2.2 <u>Proposition</u> pour tout x de C non stationnaire de (P) : Γ est sup-continue en (x,M)

<u>Dem</u> : C'(x) le domaine de P(x,M) est une fonction multivoque continue en x indépendante de M, donc continue en (x,M). La fonction économique de P(x,M) notée ϕ_x est en fait une fonction réelle définie sur : $(C \times \mathcal{M}) \times C'(x)$ qui dépend du couple ((x,M),z) où (x,M) est un paramètre et z la variable ; ϕ_x est une fonction continue sur $(x,M) \times C'(x)$.
On peut alors appliquer un théorème classique sur la paramétrisation des programmes mathématiques [20] qui permet d'affirmer que SP est une fonction multivoque sup-continue au point (x,M).
Montrons la sup-continuité de Γ en (x,M) si x non stationnaire, soient :

$$\left.\begin{array}{c} (x_n,M_n)_{n \in \mathbb{N}} \to (x,M) \\ (y_n)_{n \in \mathbb{N}} \to y \\ \forall n \in \mathbb{N},\ y_n \in \Gamma(x_n,M_n) \end{array}\right\}$$

$\forall n \in \mathbb{N},\ \exists z_n \in SP(x_n,M_n)$ tel que : $y_n = z_n - x_n$

$\forall n \in \mathbb{N},\ z_n \in SP(x_n,M_n) \subset C'(x_n) \subset B$

la suite $(z_n)_{n \in \mathbb{N}}$ située dans le compact B y admet au moins un point d'accumulation z, il existe une sous-suite N' de $\mathbb{N}$ pour laquelle :

$$(z_n,x_n,y_n)_{n \in N'} \to (z,x,y)$$

$$\left.\begin{array}{l} \forall n \in N',\ z_n \in SP(x_n,M_n) \\ (x_n,M_n)_{n \in N'} \to (x,M) \\ (z_n)_{n \in N'} \to z \\ SP \text{ sup-continue en } (x,M) \end{array}\right\} \Longrightarrow \left.\begin{array}{l} z \in SP(x,M) \\ y = z-x \end{array}\right\} \Longrightarrow y \in \Gamma(x,M)$$

C.Q.F.D

Les deux prépositions précédentes montrent que Γ vérifie les propriétés (2C) et (3C). Par contre la propriété (1C) de I.2.2.1 n'est pas vérifiée par Γ. On montre seulement un résultat plus faible : si x n'est pas stationnaire pour (P), si M est une matrice semi-définie-positive, alors la résolution de P(x,M) permet de trouver une direction dans $\overline{R_C(x)}$, mais non dans $R_C(x)$.
C étant ici convexe $\overline{R_C(x)} = T(x,C)$ (propriété I.2.1.2). Pour toute solution z de P(x,M) la direction (z-x) est améliorante, située dans le cône tangent à C en x, elle n'est pas forcément réalisable pour C en x.

II.2.3 <u>Proposition</u> Soit M une matrice de $\mathcal{M}$ semi-définie-positive et x un point quelconque de C :

$$\Gamma(x,M) \subset D_+^*(x) \cap \overline{R_C(x)}$$

<u>Dem</u> : la proposition est trivialement vérifiée si $\Gamma(x,M)$ est vide, sinon soit $(z-x) \in \Gamma(x,M)$

$$\left.\begin{array}{l} z \in SP(x,M) \\ x \notin SP(x,M) \end{array}\right\} \Longrightarrow \phi_x(z) > \phi_x(x) = 0$$

$$\phi_x(z) = \nabla f(x).(z-x) - \frac{1}{2}(x-z).M(x-z) > 0$$

$$\nabla f(x).(z-x) > \frac{1}{2}(x-z).M(x-z) \geq 0 \qquad \text{car M est semi-définie-positive}$$

donc : $\nabla f(x).(z-x) > 0$ et $(z-x) \in D_+^*(x)$

d'autre part :

$$\left.\begin{array}{l} C'(x) \text{ convexe} \\ \\ z \in C'(x) \end{array}\right\} \Longrightarrow [z-x \in T(x,C'(x))] \text{ voir } [20]$$

avec les hypothèses faites : $T(x,C'(x)) = T(x,C) = \overline{R_C(x)}$

donc : $(z-x) \in \overline{R_C(x)}$

L'utilisation du linéarisé $A'(x)$ pour définir le domaine de $P(x,M)$ conduit à un inconvénient déjà rencontré en I.2, principalement dans les algorithmes de projection ; c'est qu'on obtient par résolution de $P(x,M)$ une direction améliorante mais qui peut ne pas être réalisable si elle se trouve dans $T(x,C)$ sans être dans $R_C(x)$.

Il est nécessaire de retrouver à partir de cette direction au moins une direction améliorante réalisable.

II.3 Recherche de direction améliorante réalisable

Dans toute cette partie x est un point fixé de C. Sous les hypothèses faites au début de chapitre $R_C(x)$ est un convexe.

Les procédés de rentrée définis en I.2 apportaient une solution au problème :

étant donné $s \in D_+^*(x) \cap \overline{R_C(x)}$

trouver $s' \in D_+^*(x) \cap R_C(x)$

HUARD [21] apporte à ce problème une réponse à la fois plus générale et plus simple que les procédés usuels de rentrée ; elle utilise essentiellement la convexité de $R_C(x)$ et repose sur les deux propriétés des ensembles convexes [20] :

<u>Lemme 1</u> : X convexe, $a \in X$, $b \in \overset{\circ}{X} \Longrightarrow]a,b] \subset \overset{\circ}{X}$

<u>Lemme 2</u> : X convexe, $\overset{\circ}{X} \neq \phi \Longrightarrow \bar{X} = \overline{\overset{\circ}{X}}$

L'utilisation de ces deux lemmes va permettre, dès que sera connue une direction de $\overset{\circ}{R}_C(x)$ l'intérieur du cône des directions réalisables, de donner une réponse au problème de rentrée.

On rappelle la définition de $R_C(x)$ donnée au chapitre I :

$$R_C(x) = \{s \in \mathbb{R}^n \ / \ \exists \lambda > 0, \ [x,x+\lambda s] \subset C\}$$

On va donner plusieurs formulations équivalentes de l'intérieur de cet ensemble :

II.3.1 <u>Proposition</u> Sous l'hypothèse faite en II.1 : C est un convexe d'intérieur non vide

$$\mathring{R}_C(x) = \{s \in \mathbb{R}^n \ / \ \exists\lambda > 0,\ ]x, x + \lambda s] \subset \mathring{C}\}$$

$$= \{s \in \mathbb{R}^n \ / \ \exists\lambda > 0,\ x + \lambda s \in \mathring{C}\}$$

$$= \{\lambda(y-x) \ / \ \lambda > 0,\ y \in \mathring{C}\}$$

$$= \bigcup_{\lambda>0} \lambda(\mathring{C}-x)$$

<u>Dem</u> : On montre d'abord l'équivalence des quatre formulations : l'équivalence entre la 2° et la 3° formulation est immédiate, l'équivalence entre la 1° et la 2° forme se déduit du lemme 1 de II.3 :

$$\left.\begin{array}{l} s \in \mathbb{R}^n,\ \exists\lambda > 0,\ x + \lambda s \in \mathring{C} \\ \\ x \in C \end{array}\right\} \Longrightarrow]x, x+\lambda s] \subset \mathring{C}$$

la dernière formulation est obtenue à partir de la 3° en utilisant la notation ensembliste :

$$C-x = \{z-x \ / \ z \in C\}$$

alors : $$\{\lambda(y-x) \ / \ \lambda > 0,\ y \in \mathring{C}\} = \bigcup_{\lambda>0} \lambda(\mathring{C}-x)$$

Soit E l'ensemble défini par l'une des quatre formulations équivalentes.
On montre que E est un ouvert :

$$E = \bigcup_{\lambda>0} \lambda(\mathring{C}-x)$$

pour tout λ positif, l'ensemble $\lambda(\mathring{C}-x)$ est obtenu par translation puis homothétie appliquées à l'ouvert $\mathring{C}$, c'est un ouvert.
E réunion de ces ouverts est un ouvert.
Dès que C est convexe on peut écrire :

$$\begin{aligned} R_C(x) &= \{s \in \mathbb{R}^n \ / \ \exists\lambda > 0,\ [x, x+\lambda s] \subset C\} \\ &= \{s \in \mathbb{R}^n \ / \ \exists\lambda > 0,\ x+\lambda s \in C\} \\ &= \{s \in \mathbb{R}^n \ / \ \exists\lambda > 0,\ \exists y \in C \text{ tels que } s = \tfrac{1}{\lambda}(y-x)\} \\ &= \bigcup_{\lambda>0} \tfrac{1}{\lambda}(C-x) \\ &= \bigcup_{\lambda>0} \lambda(C-x) \end{aligned}$$

on peut alors en déduire :

$$\mathring{R}_C(x) = \overbrace{\bigcup_{\lambda>0} \lambda(C-x)}^{\circ} \supset \bigcup_{\lambda>0} \overbrace{\lambda(C-x)}^{\circ} = \bigcup_{\lambda>0} \lambda(\mathring{C}-x) = E$$

cependant l'inclusion réciproque est fausse en général

On la montre néanmoins dans ce cas particulier :

$$\left.\begin{array}{l} C \text{ convexe} \\ \mathring{C} \neq \phi \\ \text{lemme 2} \end{array}\right\} \Longrightarrow C \subset \bar{C} = \bar{\mathring{C}}$$

$$\begin{aligned} R_C(x) = \bigcup_{\lambda>0} \lambda(C-x) &\subset \bigcup_{\lambda>0} \lambda(\bar{\mathring{C}}-x) \\ &= \bigcup_{\lambda>0} \overline{\lambda(\mathring{C}-x)} \\ &\subset \overline{\bigcup_{\lambda>0} \lambda(\mathring{C}-x)} = \bar{E} \end{aligned}$$

au total on a montré que E est un ouvert inclus dans $R_C(x)$ et que $\bar{E}$ contient $R_C(x)$, c'est suffisant pour obtenir :

$$E = \bigcup_{\lambda>0} \lambda(\mathring{C}-x) = \overbrace{\bigcup_{\lambda>0} \lambda.(C-x)}^{\circ} = \mathring{R}_C(x)$$

C.Q.F.D

<u>Remarque 1</u> : les hypothèses faites en II.1 assurent la connaissance d'un point intérieur à C : $\mathring{x}$. Pour tout point réalisable x : $(\mathring{x}-x) \in \mathring{R}_C(x)$.

II.3.2 <u>Proposition</u> Soient $s_o \in \mathring{R}_C(x)$ et $s \in \overline{R_C(x)} \cap D^*_+(x)$ il existe $s_1 \in\ ]s, s_o]$ tel que :

$$]s,s_1] \subset D^*_+(x) \cap \mathring{R}_C(x)$$

<u>Dem</u> :

$$\left.\begin{array}{l} R_C(x) \text{ convexe} \\ \mathring{s} \in \mathring{R}_C(x) \\ \text{lemme 2} \end{array}\right\} \Longrightarrow \left.\begin{array}{l} s \in R_C(x) \\ \overline{R_C(x)} = \overline{\mathring{R}_C(x)} \end{array}\right\} \Longrightarrow \left.\begin{array}{l} s \in \overline{\mathring{R}_C(x)} \\ s_o \in \mathring{R}_C(x) \\ \text{lemme 1} \\ \overline{\mathring{R}_C(x)} = \mathring{R}_C(x) \end{array}\right\} \Longrightarrow]s,s_o] \subset \mathring{R}_C(x) \quad \text{(a)}$$

$$\left.\begin{array}{l} s \in D^*_+(x) \\ D^*_+(x) \text{ ouvert} \\ D^*_+(x) \text{ convexe} \end{array}\right\} \Longrightarrow \begin{array}{l} \exists s_1 \in\ ]s,s_o] \text{ tel que :} \\ [s,s_1[\ \subset D^*_+(x) \end{array} \quad \text{(b)}$$

a) et b) $\Longrightarrow \exists s_1 \in\ ]s,s_o]$ tel que : $]s,s_1] \subset D^*_+(x) \cap \mathring{R}_C(x)$

C.Q.F.D

<u>Remarque 2</u> : en général : $[X \text{ convexe}, a \in \bar{X}, b \in X] \not\Longrightarrow []a,b] \subset X]$ il n'est donc pas possible de simplifier la procédure précédente en choisissant :

$$s \in \overline{R_C(x)},\ s_o \in R_C(x)$$

car alors $]s,s_o]$ peut ne pas être dans l'ensemble $R_C(x)$ des directions réalisables en x. Il est effectivement nécessaire de connaître un élément s_o à l'intérieur de $R_C(x)$ pour utiliser, cette direction afin de construire des directions réalisables.

Remarque 3 :

$$\left.\begin{array}{l} (C \text{ convexe}) \Longrightarrow (T(x,C) = \overline{R_C(x)}) \\ \left.\begin{array}{l} \mathring{R}_C \neq \phi \\ R_C \text{ convexe} \end{array}\right\} \Longrightarrow \overline{\mathring{(R_C(x)}} = \mathring{R}_C(x) \end{array}\right\} \Longrightarrow (\mathring{T}(x,C) = \mathring{R}_C(x))$$

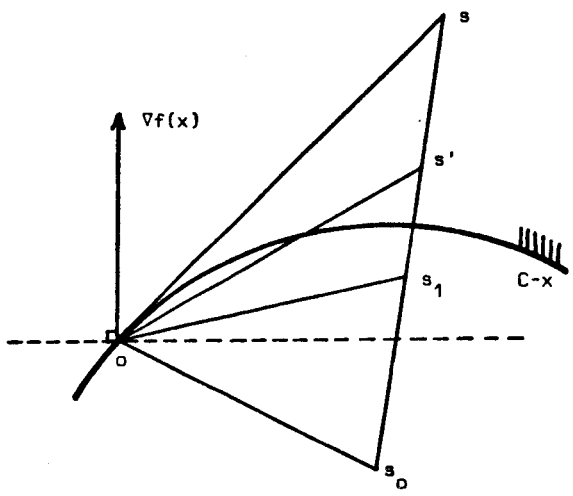
l'intérieur du cône tangent et l'intérieur du cône des directions réalisables coïncident.

Compte tenu du résultat de la proposition II.3.2 deux approches vont être utilisées :

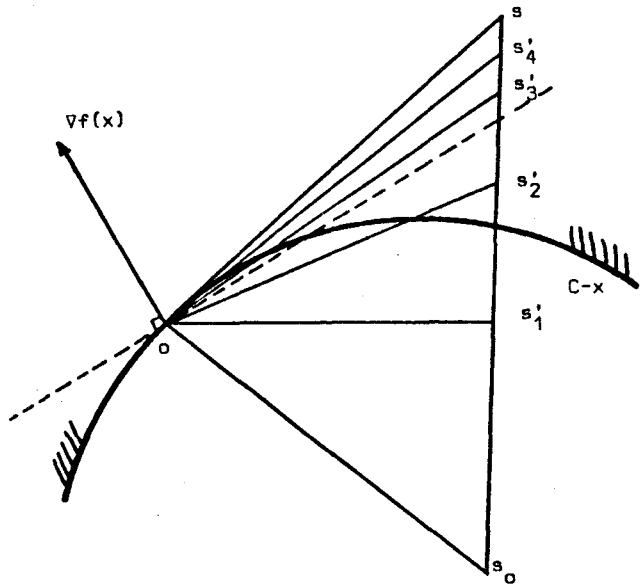
- la première correspond à la solution usuelle du problème de rentrée. Soit $s_o \in \mathring{R}_C(x)$ et $s \in D_+^*(x) \cap \overline{R_C(x)}$ on cherche une direction améliorante et réalisable ; en fait on va même déterminer une direction améliorante de $\mathring{R}_C(x)$.
  Il suffit de pouvoir déterminer une valeur de s_1 qui satisfait au résultat de II.3.2, alors par application :

  $$\forall s' \in \,]s,s_1], \; s' \in D_+^*(x) \cap \mathring{R}_C(x)$$

  on peut alors utiliser <u>une et une seule</u> direction s'.
  Cette approche est la plus simple et la plus habituelle.



- La deuxième approche consiste non pas à chercher une direction améliorante réalisable, mais à utiliser <u>un ensemble de directions</u> qui en contiendra au moins une.
  Par exemple on peut considérer toutes les directions $[s,s_o]$, d'après ce qui précède cet ensemble contient une infinité de directions améliorantes dans $R_C(x)$. Si $(s_i')_{i \in \mathbb{N}}$ est une suite de $]s,s_o]$ qui converge vers s, par application de II.3.2, il existe $s_1 \in \,]s,s_o]$ et un rang à partir duquel les directions s_i' sont dans $]s,s_1[$; ces directions sont alors améliorantes et réalisables, car dans $\mathring{R}_C(x)$.



On peut se demander à priori, quel est l'intérêt de considérer tout un ensemble de directions au lieu d'une seulement. La complexité plus grande qui en résulte risque d'être compensée par la remarque suivante :

- choisir une seule direction réalisable améliorante risque de donner une direction qui n'est que "faiblement" améliorante. Cet inconvénient n'existe plus quand on considère tout un faisceau de directions qui convergent vers la direction améliorante s.

BUS LILLE

Pour chacune de ces deux approches on va associer une fonction multivoque utilisant la direction (ou le faisceau de directions) choisie. Cette fonction multivoque inf-continue sera utilisée à chaque itération de l'algorithme général décrit en II.4.

II.3.3 Utilisation d'une direction

Le résultat de la proposition II.3.2 assure l'existence de $s_1 \in]s,s_o]$ tel que :

$$]s,s_1] \subset D_+^*(x) \cap \mathring{R}_C(x).$$

On montre que l'on peut déterminer très facilement une telle valeur ; on choisira pour s_1 l'intersection de $[s,s_o]$ avec l'hyperplan $\{y \in \mathbb{R}^n / \nabla f(x).y = 0\}$ si celle-ci existe, sinon il suffira de choisir $s_1 = s_o$.

Par hypothèse : $\nabla f(x).s > 0$ <u>si</u> $\nabla f(x).s_o \leq 0$ alors $\exists s_1 \in]s,s_o[$ tel que : $\nabla f(x) s_1 = 0$

Soit : $s_1 = s + \beta_1 (s_o - s)$

$$\nabla f(x).s_1 = 0 = \nabla f(x).s + \beta_1 . \nabla f(x).(s_o - s)$$

$$\beta_1 = \frac{\nabla f(x).s}{\nabla f(x).(s - s_o)}$$

en remarquant que : $\left.\begin{array}{l} \nabla f(x).s > 0 \\ -\nabla f(x).s_o \geq 0 \end{array}\right\}$ ⟹ le dénominateur est strictement positif et $\beta_1 \in]0,1]$

si $\nabla f(x).s_o > 0$ la direction s_o est améliorante et $s_1 = s_o$ convient. On réunit ces deux cas en posant :

$$\beta = \begin{cases} \dfrac{\nabla f(x).s}{\nabla f(x).(s-s_o)} & \text{si } \nabla f(x).s_o \leq 0 \\ 1 & \text{si } \nabla f(x).s_o > 0 \end{cases}$$

Soit

$$\boxed{\beta = \frac{\nabla f(x).s}{\nabla f(x).s - \mathrm{Min}\{\nabla f(x).s_o, 0\}}}$$

β est un scalaire de $]0,1]$ qui dépend de x et de s et uniquement de ceux-ci quand s_o est fixé.
On obtient :

$$s_1 = s + \beta(s_o - s)$$
$$\nabla f(x).s_1 = \mathrm{Max}\{\nabla f(x).s_o, 0\}$$
$$]s, s_1[\subset D_+^*(x) \cap \mathring{R}_C(x)$$

c'est à dire :

$$\forall \theta \in]0,1[, \boxed{s' = s + \theta.\beta(s_o - s)} \text{ est une direction améliorante de } \mathring{R}_C(x)$$

Soient $\theta \in]0,1[$ et $(\alpha_i)_{i \in \mathbb{N}} \to 0$ un scalaire et une suite de scalaires qui tend vers zéro, fixés indépendants de x et de s. A tout couple (x,s) on associe la direction :

$$s' = s + \theta.\beta\,(s_o - s) = s + \theta(s_1 - s)$$

et les ensembles suivants :

$$D(x.s') = \{y = x + \alpha.s' \;/\; \alpha \in \mathbb{R}^+\}$$
$$D_d(x)s') = \{y = x + \alpha_i.s' \;/\; \alpha_i \in (\alpha_i)_{i \in \mathbb{N}}\}$$

cet ensemble est un "discretise" du précédent. La notation est ici différente de celle de I.1.3.4 car x n'appartient pas à cet ensemble.

Soient $\Delta_0^1(x,s) = D(x,s') \cap \mathring{C}$

et $\Delta_0^2(x,s) = D_d(x,s') \cap \mathring{C}$

Pour montrer l'inf-continuité des fonctions multivoques Δ_0^1 et Δ_0^2, on utilise le lemme suivant qui pallie au fait que l'intersection de deux fonctions multivoques inf-continues n'est pas en général inf-continue.

II.3.3.a) Lemme si U est une fonction multivoque inf-continue et V un ouvert, alors $U \cap V$ est une fonction inf-continue

Dem :

$$\left.\begin{array}{l} (x_n)_{n \in \mathbb{N}} \to \bar{x} \\ \bar{y} \in U(\bar{x}) \\ U \text{ inf-continue} \end{array}\right\} \Longrightarrow \left.\begin{array}{l} \exists (y_n)_{n \in \mathbb{N}} \to \bar{y} \\ \forall n \in \mathbb{N},\ y_n \in U(x_n) \\ V \text{ ouvert} \end{array}\right\} \Longrightarrow \begin{array}{l} \text{un rang à partir duquel} \\ y_n \in V \end{array}$$

C.Q.F.D

II.3.3.b) Proposition Δ_o^1 et Δ_o^2 sont inf-continues

Dem : β est une fonction continue de (x,s) dans $]0,1]$. θ étant fixé, il en est de même pour $s' = s + \theta.\beta(s_o-s)$ il en résulte que D et Dd sont toutes deux des fonctions multivoques inf-continues de (x,s), et par application du lemme précédent avec l'ouvert $\mathring{C}$ on obtient $\Delta_o^1 = D \cap \mathring{C}$ et $\Delta_o^2 = D_d \cap \mathring{C}$ sont inf-continues.

Remarque 4 : le lemme précédent justifie l'emploi de l'intérieur de C pour définir les ensembles $\Delta_o^1(x,s)$ et $\Delta_o^2(x,s)$. En effet si l'on utilisait C au lieu de $\mathring{C}$ on pourrait montrer que $D \cap C$ est inf-continue par contre $D_d \cap C$ n'est pas inf-continue, car D_d est un ensemble discret.

$\Delta_o^1(x,s)$ et $\Delta_o^2(x,s)$ sont pour tout couple (x,s) des sous-ensembles de la demi-droite issue de x et de direction s', on montre qu'on obtient de plus l'identité des cônes tangents en x :

II.3.3.c) Proposition Pour tout couple (x,s) :

$$T(x, \Delta_o^2(x,s)) = T(x, \Delta_o^1(x,s)) = T(x, D(x,s'))$$

démonstration évidente.

II.3.3.d) Corollaire si x est un maximum de f sur $\Delta_o^1(x,s)$ ou sur $\Delta_o^2(x,s)$, alors :

$$\nabla f(x).s' \leq 0$$

Dem : si x max de f sur $\Delta_o^1(x,s)$ pour i = 1 ou 2

$$\nabla f(x) \in \Pi(T(x, \Delta_o^i(x,s)))$$

$$\nabla f(x) \in \Pi(T(x,D(x,s'))) \text{ par identité des cônes tangents}$$

$$\nabla f(x).s' \leq 0$$

Remarque 5 : pour tout x de C par construction : $x \in \overline{\Delta_o^2(x,s)} \subset \overline{\Delta_o^1(x,s)}$.

Quand on considère l'expression : $s' = s + \theta.\beta(s_o-s) = s + \theta(s_1-s)$; s' apparait comme un barycentre d'une direction améliorante s et d'une direction réalisable s_1. Le coefficient barycentrique θ pondère dans s', la "réalisabilité" de s' et son caractère de direction améliorante.

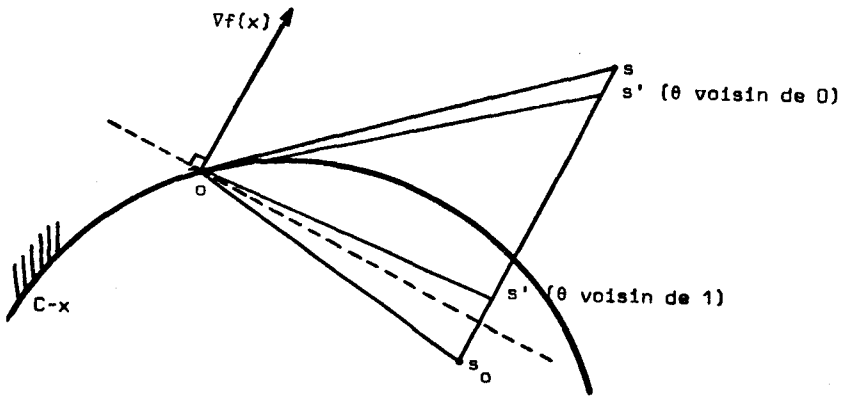
$$\begin{aligned}\nabla f(x).s' &= \nabla f(x).[s+\theta(s_1-s)] \\ &= (1-\theta).\nabla f(x).s + \theta.\nabla f(x).s_1 \\ &= (1-\theta).\ f(x).s + \theta.\text{Max}\ \{\nabla f(x).s_o, 0\}\end{aligned}$$

dans l'hypothèse où s_o n'est pas une direction améliorante pour f en x, on a vu que le deuxième terme de cette somme est nul, donc :

$$\nabla f(x).s' = (1-\theta)\ \nabla f(x).s \text{ dans ce cas particulier.}$$

Ce scalaire est la dérivée directionnelle de f en x selon la direction s', l'accroissement de la fonction f localement dans la direction s' est d'autant plus grand que ce scalaire $\nabla f(x).s'$ est grand. En remarquant que : $\nabla f(x).s > 0$ une valeur de θ voisine de zéro donne une direction s' qui risque d'être "fortement" améliorante alors qu'un choix de θ voisin de 1 ne donnera qu'une direction "faiblement" améliorante.

Par contre s n'étant pas en général une direction réalisable une valeur de θ voisine de zéro risque de donner une direction s' "faiblement" réalisable : c'est à dire que la partie réalisable de la demi-droite issue de x de direction s' peut avoir une longueur infiniment petite. Une valeur de θ proche de 1 fournit une direction qui pénètre "bien" à l'intérieur du domaine.



L'algorithme décrit en II.4 utilise à chaque itération la direction s' précédemment définie, la convergence de l'algorithme sera montrée indépendamment de toute valeur fixée de θ dans $]0,1[$.
Par contre le choix de ce paramètre devrait, en fonction de ce qui précède, avoir une forte influence sur la vitesse de convergence de l'algorithme. Il semble cependant difficile d'établir un critère pour permettre un choix "judicieux" de ce paramètre, en l'absence d'un tel critère on peut en pratique choisir pour θ la valeur moyenne $\frac{1}{2}$.

II.3.4 <u>Utilisation d'un faisceau de directions.</u>

Le résultat de la proposition II.3.2 permet d'affirmer que étant donnés $s_o \in \mathring{R}_C(x)$ et $s \in D_+^*(x) \cap \overline{R_C(x)}$ alors $]s,s_o]$ contient une infinité de directions améliorantes et dans $R_C(x)$.
On va ici construire deux sous-ensembles de $]s,s_o]$ qui possèdent cette même propriété.
On peut d'abord considérer toutes les directions de $]s,s_o]$ cet ensemble à fortiori convient. On peut aussi choisir un ensemble discret de directions constitué par une suite de directions dans $]s,s_o]$ qui converge vers s. A chacun de ces deux faisceaux de directions on associera un sous-ensemble de C.
Soit $(\theta_j)_{j \in \mathbb{N}}$ une suite d'éléments de $]0,1]$ qui converge vers 0 ; on supposera pour simplifier les notations que $\theta_o = 1$.

<u>Notations</u> :

$$\forall\theta \in]0,1] \qquad s_\theta = s + \theta(s_o - s)$$

$$\forall\theta_j \in (\theta_j)_{j \in \mathbb{N}} \qquad s_j = s + \theta_j(s_o - s)$$

$$D(x,s_\theta) = \{y = x + \lambda.s_\theta \ / \ \lambda \in \mathbb{R}^+\}$$

et pour $(\alpha_i)_{i \in \mathbb{N}}$ une suite de scalaires de $]0,1]$ qui converge vers 0

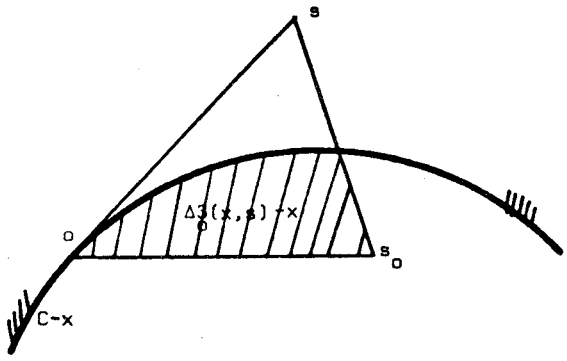
$$D_d(x,s) = \{y = x + \alpha_i.s \ / \ \alpha_i \in (\alpha_i)_{i \in \mathbb{N}}\}$$

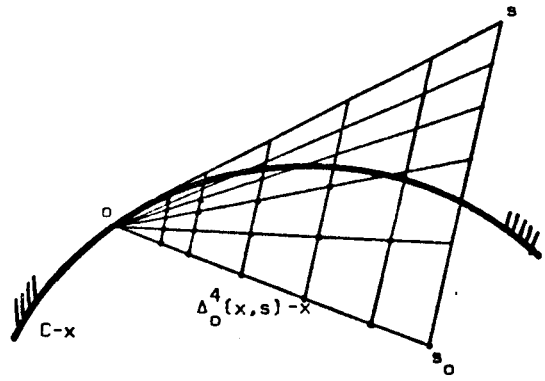
On définit les sous-ensembles de $\mathring{C}$ suivants :

$$\Delta_o^3(x,s) = \bigcup_{\theta \in]0,1]} \{D(x,s_\theta) \cap \mathring{C}\}$$

$$\Delta_o^4(x,s) = \bigcup_{j \in \mathbb{N}} \{D_d(x,s_j) \cap \mathring{C}\}$$

on remarquera la similitude avec les définitions de Δ_o^1 et Δ_o^2 dans II.3.3. L'ensemble $\Delta_o^4(x,s)$ correspond à une double discrétisation : discrétisation selon les directions apportée par la suite $(\theta_j)_{j \in \mathbb{N}}$, discrétisation selon chaque direction apportée par la suite $(\alpha_i)_{i \in \mathbb{N}}$.





<u>Remarque 6</u> : par construction $\forall x \in C,\ x \in \overline{\Delta_o^4(x,s)} \subset \overline{\Delta_o^3(x,s)}$

$$\forall (x,s) : \Delta_o^4(x,s) \subset \Delta_o^3(x,s)$$

II.3.4.a) <u>Proposition</u> $\forall x \in C,\ \forall s \in D_+^*(x) \cap \overline{R_C(x)}$
$s \in T\ (x,\ \Delta_o^4(x,s))$

<u>Dem</u>

$\forall j \in \mathbb{N},\ s_j \in\ ]s,s_o] \subset \mathring{R}_C(x)$ d'après le (a) de II.3.2

$$\left.\begin{array}{l} \text{formulation de } \mathring{R}_C(x) \text{ en II.3.1} \\ (\alpha_i)_{i \in \mathbb{N}} \to o \end{array}\right\} \Longrightarrow$$

$\forall j \in \mathbb{N}, \exists$ rang $\bar{i}$ tel que $\forall i \geq \bar{i}$ on ait : $x + \alpha_i . s_j \in \mathring{C}$

On pose pour tout $j \in \mathbb{N}$: $i(j) = \max\{j,\bar{i}\}$ par construction :

la suite $(i(j))_{j \in \mathbb{N}}$ tend vers l'infini avec j

la suite $(\alpha_{i(j)})_{j \in \mathbb{N}}$ tend vers 0 quand j tend vers l'infini

$\forall j \in \mathbb{N} : (x + \alpha_{i(j)} . s_j) \in \Delta_o^4(x,s)$

si l'on admet que les vecteurs s_j sont bornés et l'on peut toujours se ramener à ce cas ils ne représentent que des directions

$$(x + \alpha_{i(j)} . s_j)_{j \in \mathbb{N}} \to x$$

comme par définition : $(s_j)_{j \in \mathbb{N}} \to s$

il en résulte que : $s \in T(x, \Delta_o^4(x,s))$

<u>Remarque 7</u> : comme : $\Delta_o^4(x,s) \subset \Delta_o^3(x,s)$

$$T(x, \Delta_o^4(x,s)) \subset T(x, \Delta_o^3(x,s))$$

en conséquence de la proposition précédente : $s \in T(x, \Delta_o^3(x,s))$

II.3.4.b) <u>Proposition</u> Δ_o^3 et Δ_o^4 sont inf-continues.

<u>Dem</u> : $\forall \theta \in]0,1]$ s_θ est une fonction continue de s

$\forall \theta_j \in (\theta_j)_{j \in \mathbb{N}}$ s_j est une fonction continue de s

donc $D(x, s_\theta)$ et $D_d(x,s_j)$ sont des fonctions multivoques inf-continues de (x,s).

Par application du lemme II.3.3.a) :

pour tout $\theta \in]0,1]$; $D(x,s_\theta) \cap \mathring{C}$ est inf-continue

pour tout $\theta_j \in (\theta_j)_{j \in \mathbb{N}}$; $D_d(x,s_j) \cap \mathring{C}$ est inf-continue

$$\Delta_o^3(x,s) = \bigcup_{\theta \in]0,1]} \{D(x,s_\theta) \cap \mathring{C}\}$$

$$\Delta_o^4(x,s) = \bigcup_{j \in \mathbb{N}} \{D_d(x,s_j) \cap \mathring{C}\}$$

sont des unions infinies de fonctions inf-continues, elles sont inf-continues.

II.4 Algorithme général

Sous les hypothèses posées en début de chapitre on va construire un algorithme général en vue de la résolution du problème (P).

Chaque itération de cet algorithme consistera à :

- résoudre le problème P(x,M) afin de trouver une direction $z-x=s$ de $D_+^*(x) \cap \overline{R_C(x)}$
- associer à cette direction s et à $s_o=\mathring{x}-x$ de $\mathring{R}_C(x)$ l'un des sous-ensembles $\Delta_o^i(x,s)$ $i=1,2,3,4$ définis en II.3
- maximiser f sur $\Delta(x,s)$ un ensemble fermé contenu dans C et contenant un ensemble $\Delta_o^i(x,s)$ pour $i=1,2,3,4$. Cette sous-optimisation fournissant le successeur de x.

La démonstration de la convergence utilisera le résultat suivant :

soient :

X, Z deux sous-ensembles de $\mathbb{R}^n$

$f : X \to \mathbb{R}$ continue

$U,V : X \times Z \to \mathcal{P}(X)$ deux fonctions multivoques telles que : U est inf-continue

$\forall x \in X, \exists z \in Z$ tel que : $x \in \overline{U(x,z)} \subset V(x,z) = \overline{V(x,z)}$

on notera : SU(x,z) et SV(x,z) l'ensemble des maximum de f sur U(x,z), respectivement V(x,z).

$(x_n,z_n)_{n \in \mathbb{N}}$ une suite d'éléments de $X \times Z$ telle que :

$$\forall n \in \mathbb{N},\ x_{n+1} \in SV(x_n,z_n)$$

On démontre un résultat qui est une généralisation du théorème de point fixe de ZANGWILL [41] utilisé en I.1.5 (1° de la démonstration). Celle-ci est due à MEYER [26] qui considère une suite :

$$\forall n \in \mathbb{N} \; ; \; x_{n+1} \in SU(x_n, z_n)$$

hypothèse peu réaliste puisque U peut être un ouvert.
HUARD [21] corrige ce fait en considérant $V = \bar{U}$, la démonstration est ici faite dans le cas général où V est un fermé de X qui contient U.

II.4.1 <u>Lemme</u> Si X et Z sont compacts, tout point d'accumulation (x,z) de la suite $(x_n, z_n)_{n \in \mathbb{N}}$ est tel que :

$$\boxed{\overset{*}{x} \in SU(\overset{*}{x}, \overset{*}{z})}$$

<u>Dem</u> : la suite $(x_n, z_n)_{n \in \mathbb{N}}$ dans le compact $X \times Z$ admet au moins un point d'accumulation $(\overset{*}{x}, \overset{*}{z})$

$$\exists N \subset \mathbb{N} \text{ tel que : } (x_n, z_n)_{n \in N} \to (\overset{*}{x}, \overset{*}{z})$$

$\forall n \in \mathbb{N}$:

$$\left.\begin{array}{l} x_{n+1} \in SV(x_n, z_n) \\ \\ x_n \in V(x_n, z_n) \end{array}\right\} \Longrightarrow f(x_{n+1}) \geq f(x_n)$$

la suite $(f(x_n))_{n \in \mathbb{N}}$ monotone non décroissante, bornée car X est compact, converge.

$$\left.\begin{array}{l} U \text{ inf-continue} \\ \forall y \in U(\overset{*}{x}, \overset{*}{z}) \\ (x_n, z_n)_{n \in N} \to (\overset{*}{x}, \overset{*}{z}) \end{array}\right\} \Longrightarrow \begin{array}{l} \exists (y_n)_{n \in N} \to y \\ \forall n \in N, \; y_n \in U(x_n, z_n) \end{array}$$

$$\left.\begin{array}{l} \forall n \in N, \; y_n \in U(x_n, z_n) \subset V(x_n, z_n) \\ \\ x_{n+1} \in SV(x_n, z_n) \end{array}\right\} \Longrightarrow \forall n \in N, \; f(y_n) \leq f(x_{n+1})$$

la suite $(f(x_{n+1}))_{n \in N}$ est une suite extraite de la suite convergente $(f(x_n))_{n \in \mathbb{N}}$, elle admet la même limite qui est f(x) par continuité de f. L'inégalité précédente donne pour tout y par continuité de f :

$$f(y) \leq f(\overset{*}{x})$$

on a ainsi montré :

$$\forall y \in U(\overset{*}{x}, \overset{*}{z}) \; ; \; f(y) \leq f(\overset{*}{x})$$

soit :

$$\boxed{\overset{*}{x} \in SU(\overset{*}{x}, \overset{*}{z})}$$

C.Q.F.D

Remarque 1 : la liberté de choix de $V(x,z)$ dans le lemme précédent est très grande puisque les seules conditions requises sont :

$\forall(x,z) \in X \times Z$ • $V(x,z)$ fermé, inclus dans X

• $V(x,z) \supset U(x,z)$

aucune hypothèse de continuité n'est nécessaire sur la fonction multivoque V, en particulier la définition de V peut varier avec (x,z).

II.4.2 Application

En II.3 ont été définies les fonctions multivoques Δ_o^i pour $i = 1,2,3,4$. Si $z \in SP(x,M)$ défini en II.2 où M est semi-définie-positive

$$s = z-x \in D_+^*(x_o) \cap \overline{R_C(x)}$$
$$s_o = x_o-x \in \mathring{R}_C(x)$$

on peut alors appliquer les constructions de II.3 et considérer les fonctions multivoques Δ_o^i pour $i = 1,2,3,4$ comme des fonctions du couple (x,z) chacune de ces fonctions vérifie :

$$x \in \overline{\Delta_o^i(x,z)} \text{ et } \Delta_o^i \text{ est inf-continue en } (x,z)$$
$$\Delta_o^i : C \times B \to P(C)$$

on applique alors le lemme II.4.1 avec :

$$X = C \; ; \; Z = B \; ; \; U = \Delta_o^i \text{ pour } i = 1,2,3 \text{ ou } 4 \; ; \; V = \Delta$$

avec Δ une fonction multivoque, telle que :

$$\exists i \in \{1,2,3,4\} \text{ tel que } \forall x \in C : \Delta_o^i(x,z) \subset \Delta(x,z) = \overline{\Delta(x,z)} \subset C$$

On verra en application de l'algorithme général que pour chaque fonction multivoque Δ_o^i, $i = 1,2,3,4$ on peut faire un choix particulier de Δ, afin de faciliter la maximisation s'effectuant sur $\Delta(x,z)$.

II.4.3 Algorithme HUARD [22]

Pour simplifier l'écriture on notera Δ_o l'une quelconque des quatre fonctions Δ_o^i, $i = 1$ à 4 définies en II.3 ; on distinguera les valeurs de i quand cela sera nécessaire dans une démonstration.

ϕ_n au lieu de ϕ_{x_n} la fonction économique du programme $P(x_n,M_n)$

ϕ^* au lieu de ϕ_{x^*} de la même manière . Soit

$$\Delta : C \times B \to P(C) \text{ telle que : } \forall(x,z) \in C \times B, \; \overline{\Delta(x,z)} = \Delta(x,z) \supset \Delta_o(x,z)$$

On note $S\Delta(x,z)$ l'ensemble des maximum de f sur $\Delta(x,z)$.

Itération 0 : connus $x_o = \mathring{x}$; M_o semi-définie-positive

Itération n : connus x_n, M_n semi-définie-positive

soit : $z_n \in SP(x_n,M_n)$

si : $\phi_n(z_n) = 0$ alors STOP

sinon : $x_{n+1} \in S\Delta(x_n,z_n)$

$n = n+1$

II.4.3.a) <u>Proposition</u> Si l'algorithme précédent est fini le dernier point obtenu est stationnaire pour (P). Sinon

si l'hypothèse (H') suivante est vérifiée :

(H') : $\exists N1 \in \mathbb{N}$ telle que $(M_n)_{n \in N_1}$ converge vers M^*

<u>alors</u> :

tout point d'accumulation de $(x_n)_{n \in N_1}$ (et il en existe au moins un) est stationnaire.

<u>Dem</u> : si l'algorithme est fini, $\exists n \in \mathbb{N}$ tel que :

$$\left.\begin{array}{l} \phi_n(z_n) = 0 \\ z_n \in SP(x_n, M_n) \end{array}\right\} \Longrightarrow \text{la valeur optimale } \phi_n^* \text{ de } \phi_n \text{ est nulle}$$

par application de II.2.1 : x_n est stationnaire pour (P)

si l'algorithme est infini ; soit N1 la sous suite définie par (H') :

$$\left.\begin{array}{l} x_n \in C \\ z_n \in C'(x_n) \subset B \\ C \text{ compact, } B \text{ compact} \end{array}\right\} \Rightarrow \begin{array}{l}\text{la suite } (x_n, z_n)_{n \in N_1} \text{ admet au moins} \\ \text{un point d'accumulation } (\overset{*}{x}, \overset{*}{z})\end{array}$$

$\exists N2 \subset N1 \subset \mathbb{N}$ telle que :

$$((x_n, z_n), M_n)_{n \in N2} \to ((\overset{*}{x}, \overset{*}{z}), M^*)$$

on suppose que $\overset{*}{x}$ n'est pas stationnaire pour (P) on montre une contradiction :

$$\left.\begin{array}{l} \forall n \in \mathbb{N},\ x_{n+1} \in S\Delta(x_n, s_n) \\ \forall n \in \mathbb{N},\ x_n \in \Delta_o(x_n, z_n) \subset \Delta(x_n, z_n) = \overline{\Delta(x_n, z_n)} \subset C \\ \text{lemme II.4.1} \end{array}\right\} \Longrightarrow \boxed{\overset{*}{x} \in S\Delta_o(\overset{*}{x}, \overset{*}{z})} \quad \text{(a)}$$

$$\left.\begin{array}{l} \forall n \in \mathbb{N},\ z_n \in SP(x_n, z_n) \\ ((x_n, z_n), M_n)_{n \in N2} \to ((\overset{*}{x}, \overset{*}{z}), M^*) \\ SP \text{ sup-continue en } (\overset{*}{x}, M^*) \\ \text{car } \overset{*}{x} \text{ non stationnaire pour (P)} \\ \text{(II.2.2)} \end{array}\right\} \Longrightarrow \boxed{\overset{*}{z} \in SP(\overset{*}{x}, \overset{*}{M})} \quad \text{(b)}$$

$$\left.\begin{array}{l} \left.\begin{array}{l} \overset{*}{x} \in S\Delta_o(\overset{*}{x}, \overset{*}{z}) \\ \text{si } \Delta_o \equiv \Delta_o^1 \text{ ou } \Delta_o^2 \\ \text{corollaire II.3.3.d)} \end{array}\right\} \Longrightarrow \nabla f(\overset{*}{x}).\overset{*}{s}' \leq 0 \\ \nabla f(\overset{*}{x}).\overset{*}{s}' = (1-\theta)\ \nabla f(\overset{*}{x}).\overset{*}{s} + \theta.\mathrm{Max}\ \{\nabla f(\overset{*}{x}).\overset{*}{s}_o, 0\} \quad \text{avec } \overset{*}{s} = \overset{*}{z} - \overset{*}{x} \\ \text{donc : } \dfrac{\nabla f(\overset{*}{x}).\overset{*}{s}'}{(1-\theta)} \geq \nabla f(\overset{*}{x}).\overset{*}{s} = \nabla f(\overset{*}{x}).(\overset{*}{z} - \overset{*}{x}) \qquad \overset{*}{s}_o = x_o - \overset{*}{x} \end{array}\right\} \boxed{\nabla f(\overset{*}{x}).(\overset{*}{z} - \overset{*}{x}) \leq 0} \quad \text{(c)}$$

$$\left.\begin{array}{l}\text{Si } \Delta_o \equiv \Delta_o^3 \text{ ou } \Delta_o^4 \\ \\ \text{II.3.4a)}\end{array}\right\} \begin{array}{l} \Longrightarrow (\overset{*}{z}-\overset{*}{x}) \in T(\overset{*}{x},\Delta_o(\overset{*}{x},\overset{*}{z})) \\ \\ \overset{*}{x} \in S\Delta_o(\overset{*}{x},\overset{*}{z}) \end{array} \Bigg\} \Longrightarrow \boxed{\nabla f(\overset{*}{x}).(\overset{*}{z}-\overset{*}{x}) \leq 0}$$

(c)

M^* limite de matrices semi-définies-positives est semi-définie-positive

$$\phi^*(z) = \nabla f(\overset{*}{x}).(\overset{*}{z}-\overset{*}{x}) - \frac{1}{2}(x-z).M^*(\overset{*}{x}-\overset{*}{z})$$

$$\left.\begin{array}{l}\overset{*}{x} \text{ non stationnaire} \\ \\ \text{II.2.1}\end{array}\right\} \Longrightarrow [\overset{*}{s} \notin SP(\overset{*}{x},M^*)] \Longleftrightarrow [\phi^*(\overset{*}{z}) > \phi^*(\overset{*}{x}) = 0]$$

$$\phi^*(\overset{*}{z}) = \nabla f(\overset{*}{x})(\overset{*}{z}-\overset{*}{x}) - \frac{1}{2}(\overset{*}{x}-\overset{*}{z})\, M^*(\overset{*}{x}-\overset{*}{z}) > 0$$

$$\nabla f(\overset{*}{x}).(\overset{*}{z}-\overset{*}{x}) > \frac{1}{2}(\overset{*}{x}-\overset{*}{z})\, M^*(\overset{*}{x}-\overset{*}{z}) \geq 0$$

car M^* est semi-définie-positive

On obtient :

$$\boxed{\nabla f(\overset{*}{x}).(\overset{*}{z}-\overset{*}{x}) > 0}$$

(d)

on obtient une contradiction entre (c) et (d).

$\overset{*}{x}$ est stationnaire pour (P). C.Q.F.D.

<u>Remarque 2</u> : sous l'hypothèse supplémentaire de <u>concavité</u> de f on obtient le résultat suivant :

si (H') : $\exists N1 \subset \mathbb{N}$ telle que : $(M_n)_{n \in N1}$ converge vers M^*

alors

tout point d'accumulation de $(x_n)_{n \in \mathbb{N}}$ est solution optimale de (P).

<u>Dem</u> : $\overset{*}{x}$ un point d'accumulation de $(x_n)_{n \in N1}$; $\bar{x}$ un point d'accumulation de $(x_n)_{n \in \mathbb{N}}$

$$\left.\begin{array}{l} \left.\begin{array}{l} (f(x_n))_{n \in \mathbb{N}} \text{ monotone} \Longrightarrow [f(\overset{*}{x}) = f(\bar{x})] \\ \\ \text{II.4.3.a)} \Longrightarrow \overset{*}{x} \text{ stationnaire} \\ \\ \qquad\qquad\qquad f \text{ concave, } C \text{ convexe} \end{array}\right\} \overset{*}{x} \text{ solution optimale} \end{array}\right\} \Longrightarrow \bar{x} \text{ solution optimale}$$

II.5 Sous-optimisation finie

L'algorithme général décrit en II.4 nécessite à chaque itération la résolution du programme quadratique $P(x_n,M_n)$. Sous les hypothèses faites M_n est une matrice semi-définie-positive, la fonction économique ϕ_n est concave et le programme $P(x_n,M_n)$ est à chaque itération un programme convexe dont le domaine est un polyèdre borné. On pourra par différentes méthodes spécifiques à l'optimisation quadratique [38] trouver une solution optimale à ce sous-problème en un nombre fini de calculs.
L'algorithme général engendre toute une famille d'algorithmes selon le choix des matrices M_n et selon la définition de Δ.
Δ doit vérifier :

$$\Delta : C \times B \to \mathcal{P}(C)$$

$$\exists i \in \{1,2,3,4\} \text{ tel que : } \forall (x,z) \in C \times B,\ \overline{\Delta(x,z)} = \Delta(x,z) \supset \Delta_o^i(x,z)$$

où les fonctions Δ_o^i sont celles définies en II.3.
Sous l'hypothèse de <u>concavité de f</u>, différents choix simples d'ensembles discrets Δ permettent une sous-optimisation en un nombre <u>fini</u> de calculs : à chaque itération la recherche de x_{n+1} élément de $S\Delta(x_n,z_n)$ est finie.

II.5.1 On définit un sous-ensemble discret et fermé de C à partir de Δ_o^2 ; qui utilise donc une seule direction améliorante et réalisable de recherche.
Rappel des notations :

$$s = z-x \ ;\ s_o = x_o-x$$

$$\beta = \frac{\nabla f(x).s}{\nabla f(x).s-\mathrm{Min}\{\nabla f(x).s_o,0\}}$$

$$s' = s + \theta.\beta\ (s_o-s)$$

on pose : $s' = t \ominus x$

$$t = z + \theta.\beta\ (x_o-z)$$

pour toute suite <u>fixée</u> (indépendante de l'itération) strictement décroissante et qui converge vers zéro, on a :

$$D_d(x,s') = \{x + \alpha_i.s' \ /\ i \in \mathbb{N}\} = \bigcup_{i\in\mathbb{N}} \{x+\alpha_i(t-x)\}$$

$$\Delta_o^2(x,z) = \bigcup_{i\in\mathbb{N}} \{x+\alpha_i.(t-x)\} \cap \mathring{C}$$

Soit Δ défini par :

$$\Delta(x,z) = \overline{\bigcup_{i\in\mathbb{N}} \{x + \alpha_i(t-x)\}} \cap C$$

$$\Delta(x,z) = [\{x\} \cup \bigcup_{i\in\mathbb{N}} \{x+\alpha_i(t-x)\}] \cap C$$

alors Δ vérifie toutes les hypothèses requises dans l'algorithme général

- $\Delta(x,z)$ est un fermé
- $\Delta(x,z)$ est inclus dans C
- $\Delta(x,z)$ contient $\Delta_o^2(x,z)$

d'autre part, par construction même de l'algorithme la direction (t-x) est améliorante pour f :

$$\nabla f(x)(t-x) = (1-\theta).\nabla f(x)(z-x) + \theta.\ \mathrm{Max}\ \{\nabla f(x)(x_o-x),\ 0\}$$

$$\nabla f(x)(t-x) \geq (1-\theta).\nabla f(x).(z-x) > 0$$

On peut alors appliquer le résultat de I.1.3.4.b) : la maximisation de f concave dans la direction améliorante (t-x), sur l'ensemble discret $\{x\} \cup \bigcup_{i\in\mathbb{N}} \{x+\alpha_i(t-x)\}$ se fait en un nombre fini de calculs.
A fortiori la maximisation sur la partie réalisable de cet ensemble, c'est à dire sur $\Delta(x,z)$ ne requiert qu'une recherche finie.

Remarque 1 : soit i_o le plus petit indice pour lequel $x+\alpha_i(t-x)$ soit réalisable ; pour i supérieur à i_o la recherche du maximum se fait comme dans le cas sans contraintes : on examine successivement les points $x + \alpha_i(t-x)$ dans l'ordre de i croissant, la recherche s'arrêtant dès que :

$$f[x + \alpha_{i+1}(t-x)] \leq f[x + \alpha_i(t-x)]$$

Remarque 2 : compte tenu de la remarque 1 de II.4 il peut être intéressant de redéfinir localement $\Delta(x,z)$. Supposons par exemple qu'on obtienne à une itération le maximum de f sur $\Delta(x,z)$ au point : $x+\alpha_o(t-x)$; cela peut signifier que la fonction f concave reste croissante dans la direction s=t-x à partir de ce point. On a intérêt dans ce cas, pour obtenir un accroissement maximal de la fonction à cette itération à modifier localement la définition de $\Delta(x,z)$: on peut alors continuer à explorer la partie réalisable de la demi-droite considérée à l'aide d'un nombre fini de points.
Cet aspect est important pour la rapidité de la convergence, car il permet une adaptation en fonction de l'itération (locale) de la suite de paramètres $(\alpha_i)_{i\in\mathbb{N}}$.

II.5.2 Quand on utilise, à une itération de l'algorithme général, un faisceau de directions de recherche, sur chacune de celle-ci on peut d'après II.5.1 trouver un maximum en un nombre fini de calculs ; le nombre de directions à explorer étant infini, la recherche globale d'un maximum dans l'ensemble de ces directions risque de nécessiter une recherche infinie. HUARD [22] a cependant montré que même dans ce cas, un choix judicieux de Δ et un ordre d'exploration adéquat permettent de trouver un maximum sur Δ en un nombre fini de calculs. Ce nombre fini pourra pratiquement être déterminé grâce à un test d'arrêt.

$(\theta_j)_{j\in\mathbb{N}}$ et $(\alpha_i)_{i\in\mathbb{N}}$ suites strictement décroissantes qui tendent vers 0

$$\theta_o = 1$$

$$s_j = s + \theta_j(s_o-s) = z-x+\theta_j(x_o-z)$$

$$D_d(x,s_j) = \bigcup_{i\in\mathbb{N}} \{x + \alpha_i.s_j\}$$

$$D_d(x,s_j) = \{x + \alpha_i(z-x) + \alpha_i.\theta_j(x_o-z)\ /\ i \in \mathbb{N}\}$$

$$\Delta_o^4(x,z) = \bigcup_{j\in\mathbb{N}} \{D_d(x,s_j) \cap \mathring{C}\}$$

$$\Delta_o^4(x,z) = \{x + \alpha_i(z-x) + \alpha_i.\theta_j(x_o-z)\ /\ (i,j) \in \mathbb{N}^2\} \cap \mathring{C}$$

On définit Δ par :

$$\Delta(x,z) = \overline{\{x + \alpha_i(z-x) + \alpha_i\theta_j(x_o-z)\ /\ (i,j) \in \mathbb{N}^2\}} \cap C$$

$$\Delta(x,z) = [\{x\} \cup \{x+\alpha_i(z-x)\ /\ i \in \mathbb{N}\} \cup \{x + \alpha_i(z-x) + \alpha_i.\theta_j(x_o-z)/(i,j) \in \mathbb{N}^2\}] \cap C$$

cet ensemble est par construction un fermé, inclus dans C et qui contient $\Delta_o^4(x,z)$.

pour tout $i \in \mathbb{N}$ on pose :

$$M_i = x + \alpha_i(z-x)$$

$$N_i = x + \alpha_i(x_o-x)$$

$$\Delta(x,z) = \bigcup_{i\in \mathbb{N}} ([\{M_i\} \cup \{M_i + \theta_j(N_i-M_i)/j \in \mathbb{N}\}] \cap C) \cup \{x\}$$

On montre dans un premier temps que la recherche d'un maximum sur la partie de Δ restreinte à $[M_iN_i]$ est finie.

Soit $\phi_d(\alpha_i) = \text{Max } \{f(t)/t \in (\{M_i\} \cup \{M_i+\theta_j(N_i-M_i)/j \in \mathbb{N}\}) \cap C\}$ pour tout i, $\phi_d(\alpha_i)$ peut être obtenu en un nombre fini de calculs.

<u>1° cas</u> : $M_i \notin C$ alors : $[\{M_i\} \cup \{M_i + \theta_j(N_i-M_i) / j \in \mathbb{N}\} \cap C$ ne contient qu'un nombre fini de points et la recherche d'un maximum est finie.

<u>2° cas</u> : $M_i \in C$ alors :
$[\{M_i\} \cup \{M_i+\theta_j(N_i-M_i)/j \in \mathbb{N}\}] \cap C = \{M_i\} \cup \{M_i + \theta_j(N_i-M_i)/j \in \mathbb{N}\}$

<u>si</u> : $\nabla f(M_i).(N_i-M_i) \leq 0$ alors M_i est un maximum de la fonction concave f sur la demi-droite issue de M_i passant par N_i, et à fortiori sur l'ensemble discret considéré

<u>sinon</u> : $\nabla f(M_i).(N_i-M_i) > 0$ on applique alors I.1.3.4.b) à la fonction f concave : un maximum est atteint en examinant un nombre fini connu de points.

Avec les notations définies :

$$\text{Max } \{f(t) / t \in \Delta(x,z)\} = \text{Max } \{\text{Max}_{i\in\mathbb{N}} \{\phi_d(\alpha_i)\}, f(x)\}$$

on a vu que pour tout i, $\phi_d(\alpha_i)$ se calcule en un temps fini. On montre maintenant qu'il suffit d'en calculer un nombre fini <u>connu</u>.

<u>Lemme</u> : soient ϕ_d, ϕ, ψ : $\mathbb{R}^+ \to \mathbb{R}$ telles que : $(\alpha_i)_{i \in \mathbb{N}}$ étant une suite strictement décroissante de scalaires qui converge vers 0

(1) • $\exists p \in \mathbb{N}$ tel que : $\phi_d(\alpha_p) > \phi_d(o)$

(2) • ϕ est concave

(3) • $\lim_{i\in\mathbb{N}} \psi(\alpha_i) = \phi_d(o)$

(4) • $\forall i \in \mathbb{N} : \psi(\alpha_i) \geq \phi(\alpha_i) \geq \phi_d(\alpha_i)$

<u>alors</u> : le maximum de ϕ_d sur l'ensemble $\{o\} \cup \bigcup_{i\in\mathbb{N}} \{\alpha_i\}$ peut être atteint en un nombre fini de calculs. Ce nombre fini peut être déterminé pratiquement grâce à un test d'arrêt.

<u>Dem</u> : • $\exists j$ tel que : $\underset{i\leq j-1}{\text{Max}} \{\phi_d(\alpha_i)\} \geq \psi(\alpha_j)$

sinon :

$\forall j \in \mathbb{N}$; $\underset{i\leq j-1}{\text{Max}} \{\phi_d(\alpha_i)\} < \psi(\alpha_j)$

à la limite quand $j \to \infty$, $\alpha_j \to 0$ et $\psi(\alpha_j) \to \phi_d(o)$

on obtient : $\underset{i\in\mathbb{N}}{\text{Max}} \{\phi_d(\alpha_i)\} \leq \phi_d(o)$

et donc $\phi_d(\alpha_p) \leq \phi_d(o)$

contraire à l'une des hypothèses faites (1).

• Soit j le plus petit indice tel que :

$$\underset{i\leq j-1}{\text{Max}} \{\phi_d(\alpha_i)\} \geq \psi(\alpha_j)$$

soit k un indice tel que : $\phi_d(\alpha_k) = \underset{i\leq j-1}{\text{Max}} \{\phi_d(\alpha_i)\}$

$$\left.\begin{array}{l} (k < j) \Longrightarrow (\alpha_j < \alpha_k) \\ \phi(\alpha_j) \leq \psi(\alpha_j) \leq \phi_d(\alpha_k) \leq \phi(\alpha_k) \\ \phi \text{ concave} \end{array}\right\} \Longrightarrow \phi \text{ est monotone non décroissante sur }]0,\alpha_j]$$

donc : $\underset{i\in\mathbb{N}}{\text{Max}} \{\phi_d(\alpha_i)\} = \underset{i\leq j-1}{\text{Max}} \{\phi_d(\alpha_i)\}$

où l'indice j fini est connu grâce au test d'arrêt :

$$\underset{i\leq j-1}{\text{Max}} \{\phi_d(\alpha_i)\} \geq \psi(\alpha_j)$$

<u>Remarque 1</u> : ce lemme peut être considéré comme une généralisation de la proposition I.1.3.4.b), car si f est une fonction concave dont 0 n'est pas un maximum, en posant : $\psi \equiv \phi \equiv \phi_d \equiv f$ toutes les hypothèses sont vérifiées. La détermination de j se faisant par le test d'arrêt : j le plus petit indice tel que :

$$f(\alpha_{i+1}) \leq f(\alpha_i)$$

On va appliquer ce lemme dans le cas particulier de la maximisation sur Δ, en définissant des fonctions ϕ et ψ qui vérifient les hypothèses.

<u>Vérification de (1)</u> : on montre que si $\nabla f(x).(z-x) > 0 \Rightarrow x$ n'est pas un maximum de f sur $\Delta(x,z)$.

Sinon :

$$\left.\begin{array}{l} x \text{ maximise } f \text{ sur } \Delta(x,z) \\ \Delta(x,z) \supset \overline{\Delta_o^4(x,z)} \end{array}\right\} \Longrightarrow x \text{ maximise } f \text{ sur } \overline{\Delta_o^4(x,z)}$$

$$\Longrightarrow \quad \left.\begin{array}{l} \nabla f(x) \in \Pi[T(x, \Delta_o^4(x,z)] \\ (z-x) \in T(x, \Delta_o^4(x,z) \end{array}\right\} \Longrightarrow \nabla f(x).(z-x) \leq 0$$

on aboutit à une contradiction, donc :

$$x \notin S\ \Delta(x,z) \text{ et } \exists p \in \mathbb{N} \text{ tel que : } \phi_d(\alpha_p) > \phi_d(o) = f(x)$$

Vérification de (2) : on exhibe une fonction concave qui de plus vérifie (4) soit :

$$\phi : \mathbb{R}^+ \to \mathbb{R} \text{ telle que : } \phi(\alpha_i) = \text{Max } \{f(t)/t \in [M_i N_i] \cap C\}$$

on pose : $\forall \alpha \in \mathbb{R}^+$

$$\psi(\alpha) = \text{Max } \{f(t)/t \in [MN] \cap C\}$$

où

$$M = x + \alpha(z-x)$$
$$N = x + \alpha(x_o - x)$$

par construction $\forall i \in \mathbb{N}$, $\phi(\alpha_i) \geq \phi_d(\alpha_i)$.
De plus ϕ maximum d'une fonction continue en α sur le domaine $[M\ N] \cap C$ qui est continu en α, est une fonction concave. Le test d'arrêt qui intervient dans le lemme ne fait pas intervenir ϕ, qui n'a pas à être calculée.

Vérification de (3) : on va exhiber des fonctions ψ qui vérifient (3) en même temps que l'inégalité de (4).
Toute fonction majorante de ϕ, qui tend vers $\phi_d(o) = f(x)$ quand α tend vers 0 convient.
La concavité de f permet de trouver de telles fonctions.

- $\psi(\alpha) = f(M) + \text{Max } \{0, \nabla f(M).(N-M)\}$

  quand $\alpha \to 0$, les points N et M tendent vers x et par continuité : $\psi(0) = f(x)$

  de plus :
$$\begin{aligned}\psi(\alpha_i) &= f(M_i) + \text{Max } \{0, \nabla f(M_i)(N_i - M_i)\} \\ &= \underset{\theta \in [0,1]}{\text{Max}} \ [f(M_i) + \theta.\nabla f(M_i)(N_i - M_i)] \\ &\geq \underset{\theta \in [0,1]}{\text{Max}} \ [f(M_i + \theta(N_i - M_i))] \\ &\geq \phi(\alpha_i)\end{aligned}$$

- la détermination de $\phi_d(\alpha_i)$ c'est à dire la maximisation de f sur :

$$[\{M_i\} \cup \{M_i + \theta_j(N_i - M_i)/j \in \mathbb{N}\}] \cap C$$

  peut fournir un intervalle d'incertitude pour le maximum de f sur $[M_i N_i] \cap C$
  On peut alors par I.1.3.3.a) ou d'autres méthodes, déterminer une majoration de $\phi(\alpha_i)$, de la forme :

$$\psi(\alpha_i) = \phi(\alpha_i) + \varepsilon_i$$

  "l'erreur" commise ε_i tend vers 0 avec la longueur de l'intervalle d'incertitude et donc avec α_i ; c'est à dire $\lim\limits_{i\ \mathbb{N}} \psi(\alpha_i) = \phi(0) = f(x)$

- $\psi(\alpha) = f(x) + \alpha.\ \text{Max } \{\nabla f(x)(x_o - x), \nabla f(x)(z-x)\}$ vérifie évidemment

$$\lim_{i \in \mathbb{N}} \psi(\alpha_i) = \lim_{\alpha \to o} \psi(\alpha) = f(x) = \phi_d(o)$$

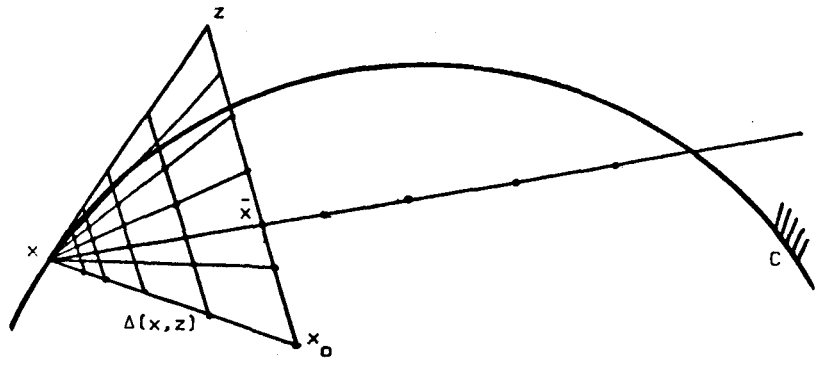
  si :
$$t = x + \alpha(1-\theta)(z-x) + \alpha.\theta\ (x_o - x)$$
$$f(t) \leq f(x) + \nabla f(x).[\alpha(1-\theta)(z-x) + \alpha.\theta(x_o - x)]$$
$$\underset{\substack{\theta \in [0,1] \\ \alpha = \alpha_i}}{\text{Max}} \{f(t)\} \leq f(x) + \alpha_i.\ \underset{\theta}{\text{Max}} \ \{(1-\theta)\ \nabla f(x)(z-x) + \theta.\nabla f(x)(x_o - x)\}$$
$$\phi(\alpha_i) \leq f(x) + \alpha_i.\ \text{Max } \{\nabla f(x)(x_o - x), \nabla f(x)(z-x)\}.$$

Si de nombreuses majoration de $\phi(\alpha)$ = Max $\{f(t) \ / \ \theta \in [0,1],\ \alpha$ fixé $\}$ conviennent, il est cependant intéressant de pouvoir utiliser une majoration la plus "fine" possible : en effet le test d'arrêt

$$\max_{i \leq j-1} \{\phi_d(\alpha_i)\} \geq \psi(\alpha_j)$$

risque d'être satisfait plus rapidement (pour une plus petite valeur de j) dans l'ordre d'exploration des j croissant.

Remarque 2 : la même remarque qu'en II.5.1 peut être faite quant à une adaptation de la définition de $\Delta(x,z)$ à une itération déterminée. Si $\bar{x}$ est un maximum de f sur $\Delta(x,z)$ initialement fixé, qui est obtenu "en butée" (pour la valeur α_o du paramètre), on peut continuer à explorer la direction $\bar{x}$-x par un ensemble fini de points réalisables.



En conclusion chaque itération de l'algorithme principal consiste à résoudre les deux problèmes :

- trouver $z \in SP(x,M)$
- trouver $\bar{x} \in S\Delta(x,z)$

les choix indiqués de Δ permettent le traitement d'une itération en un nombre fini parfaitement déterminé de calculs.

II.6 Applications

II.6.1 On montre dans cette partie la liaison entre l'algorithme général II.4 et les algorithmes de projection ou les algorithmes utilisant le critère de meilleure direction améliorante.
Soit M une matrice définie positive et $||u||_M^2 = u\ M\ u$ la norme associée. Il a été démontré en I.1.4.f) que la meilleure direction améliorante pour f dans le cas sans contraintes et relativement à la norme associée à M, est :

$$\left[\frac{M+M^t}{2}\right]^{-1} . \nabla f(x)$$

On considère la projection de $(x + (\frac{M+M^t}{2})^{-1} . \nabla f(x))$ sur le linéarisé du domaine en x : C'(x).
Il s'agit de résoudre le problème suivant dans C'(x) :

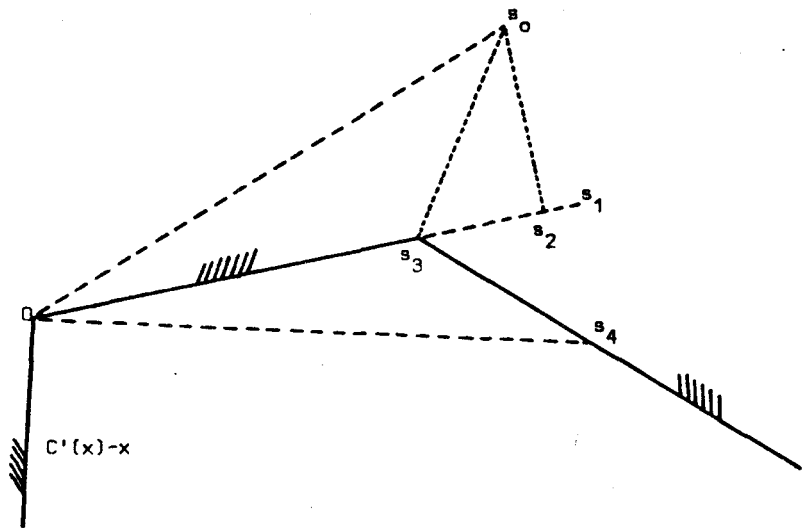
$$\text{Min } \{(x-z+2(M+M^t)^{-1}.\nabla f(x)).M(x-z+2(M+M^t)^{-1}.\nabla f(x))\}$$

$$\equiv \text{ Min } \{2\ \nabla f(x).(x-z) + (x-z).M(x-z) + \text{const}\}$$

$$\equiv \text{ Min } \{\nabla f(x)(x-z) + \frac{1}{2}\ (x-z).M(x-z)\}$$

$$\equiv \text{ Max } \{\nabla f(x).(z-x) - \frac{1}{2}\ (z-x)\ M(z-x)\}$$

on retrouve ainsi le problème P(x,M). Quand la matrice M est défini positive la direction (z-x) utilisée dans l'algorithme général peut être considérée comme une projection (selon la norme associée à M) de la meilleure direction améliorante trouvée dans le cas sans contrainte.
On remarquera que la direction (z-x) n'est pas en général, la meilleure direction améliorante du cône tangent à C en x.
On peut résumer les liaisons entre les différentes projections et les différentes notions de "meilleure direction améliorante" dans le tableau suivant :



C'(x) le linéarisé du domaine C en x ; $C'_o(x) = C'(x)-x$ le translaté

T(x,C) le cône tangent à C en x

$||u||^2_M = u.M.u$ on note Proj la projection associée

$||u||$ la norme euclidienne

x fixé ; θ varie dans $\mathbb{R}^+$

1	Définition	s_i	$\frac{s_i}{\|\|s_i\|\|}$
0	s_0 solution optimale de $\boxed{\begin{array}{l}\text{Max } \nabla f(x).s \\ \|\|s\|\|_M^2 \leq \theta\end{array}}$	varie avec θ	indépendant de θ
1	s_1 solution optimale de $\boxed{\begin{array}{l}\text{Max } \nabla f(x).s \\ \|\|s\|\|_M^2 \\ s \in T(x,C)\end{array}}$	varie avec θ	indépendant de θ
2	s_2 projection de s_0 sur $T(x,C)$	varie avec θ	indépendant de θ
3	s_3 projection de s_0 sur $C_0'(x)$	varie avec θ	varie avec θ
4	s_4 solution optimale de $\boxed{\begin{array}{l}\text{Max } \nabla f(x).s \\ \|\|s\|\|_M^2 \leq \theta \\ s \in C_0'(x)\end{array}}$	varie avec θ	varie avec θ

$\left.\right\} \frac{s_1}{\|\|s_1\|\|} = \frac{s_2}{\|\|s_2\|\|}$

$\left.\right\} \frac{s_3}{\|\|s_3\|\|} \neq \frac{s_4}{\|\|s_4\|\|}$ en général

BUS LILLE

Par contre, soit $B(\varepsilon) = \{u \in \mathbb{R}^n \ / \ ||u||_M^2 = u.M.u \leq \varepsilon\}$
alors $\exists \varepsilon > o$ suffisamment petit tel que :

$$B(\varepsilon) \cap C'_o(x) = B(\varepsilon) \cap T(x,C)$$

donc $\exists \theta_o > o$ suffisamment petit tel que pour tout $\theta \leq \theta_o$ on ait :

$$\frac{s_1}{||s_1||} = \frac{s_2}{||s_2||} = \frac{s_3}{||s_3||} = \frac{s_4}{||s_4||}$$

et dans ce cas il est possible de parlerde meilleure direction améliorante compte-tenu des contraintes.

II.6.2 La suite de matrices $(M_n)_{n \in \mathbb{N}}$ utilisée dans l'algorithme doit vérifier deux conditions :

- $\forall n \in \mathbb{N}$, M_n semi-définie-positive
- $\exists N1 \in \mathbb{N}$ telle que : $(M_n)_{n \in N1} \to M^*$

différents choix simples de cette suite de matrices assurent ces deux hypothèses et établissent la liaison avec des algorithmes connus.

- $\forall n \in \mathbb{N}$: $M_n = 0$ matrice nulle
  alors

$$P(x,M) : \boxed{\begin{array}{l} \text{Max } \nabla f(x).(z-x) \\ z \in C'(x) \end{array}}$$

l'algorithme consiste alors en une adaptation de l'algorithme de Frank et Wolfe (I.2.3.1) dans le cas de contraintes non linéaires.

- $\forall n \in \mathbb{N}$; $M_n = m_n.I$; I matrice identité, m_n situé dans un borné de $\mathbb{R}^+$

$$P(x,M) : \boxed{\begin{array}{l} \text{Max } \nabla f(x).(z-x) - \frac{1}{2}m_n.(z-x)(z-x) \\ z \in C'(x) \end{array}}$$

l'algorithme est alors un algorithme de gradient projeté selon la norme euclidienne.
Des résultats numériques sont donnés en annexe 1, avec :
$M_n = m.I$; $\forall n \in \mathbb{N}$ pour différents choix du paramètre $m > 0$, avec $\Delta(x,z)$ défini en II.5.2 à partir d'un faisceau de directions.

- Le choix pour la famille $(M_n)_{n \in \mathbb{N}}$ de matrices de métriques variables fait l'objet du chapitre III.

D'une manière générale on peut remarquer que l'hypothèse

$$(H') : \exists N1 \in \mathbb{N} \text{ telle que : } (M_n)_{n \in N1} \to M^*$$

peut être assurée par le procédé simple connu sous le terme de "restart" et qui consiste à choisir pour M_n une matrice constante toutes les p itérations.

Soit $N1 = p.\mathbb{N} = \{p.n \ / \ n \in \mathbb{N}\}$ où $p \in \mathbb{N}^*$
alors $\forall j \in N1$, $M_j = M$ matrice semi-définie-positive
et l'hypothèse (H') est ainsi "forcée".

CHAPITRE III

LES MÉTRIQUES VARIABLES
ET
L'OPTIMISATION AVEC CONTRAINTES

Le but de ce chapitre est d'établir la liaison entre l'algorithme général décrit au chapitre II et les algorithmes de métrique variable.
Après avoir rappelé les principes (III.1) et les résultats (III.2) connus pour l'optimisation sans contraintes, le problème de l'adaptation des méthodes de métrique variable au cas de l'optimisation avec contraintes non linéaires est exposé (III.3). Malgré l'interférence entre l'algorithme décrit au chapitre II et un algorithme donné par Han en 1976 (II.4), les approches et les résultats afférents restent nettement distincts pour ces deux algoritmes (III.5).

III.1 Les principes des algorithmes de métrique variable

Pour résoudre le problème sans contraintes :

$$\boxed{\begin{array}{l} \text{Max } f(x) \\ x \in \mathbb{R}^n \end{array}}$$

où f est 2 fois continument différentiable de gradient ∇f et Hessien $\nabla^2 f$, toute une classe d'algorithmes utilise des directions de recherche qui sont engendrées de la manière suivante :

$$s_k = -H_k.\nabla f(x_k)$$

on passe alors itérativement d'un point à son successeur par une formule de la forme :

$$\boxed{x_{k+1} = x_k - \alpha_k.H_k.\nabla f(x_k)}$$

où α_k est scalaire positif qui caractérise le pas à chaque itération.
Si les matrices $(-H_k)$ sont définies positives de tels algorithmes peuvent être interprétés (I.1.4.f remarque 4) comme des algorithmes de plus grande pente par rapport à la norme

$$||y||^2 = y(-H_k)^{-1}.y$$

les directions s_k sont, dans ce cas, améliorantes.

Les algorithmes de quasi-Newton construisent une suite de matrices $(H_k)_{k \in \mathbb{N}}$ en liaison avec l'inverse du Hessien : $[\nabla^2 f(x)]^{-1}$
Ces algorithmes ont pour but d'obtenir les avantages de la méthode de Newton, en utilisant seulement des renseignements au 1° ordre sur la fonction économique.
La mise à jour de la matrice H_k se fait à partir des valeurs du gradient aux points déjà rencontrés, de manière à ce que la condition suivante soit vérifiée :

$$\boxed{H_{k+1}.\Delta g_k = \Delta x_k}$$

où :

$$\Delta g_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$
$$\Delta x_k = x_{k+1} - x_k$$

cette condition souvent désignée comme "condition de quasi-Newton" trouve sa justification dans le fait que si f est quadratique, alors :

$$[\Delta^2 f(x)]^{-1}.\Delta g_k = \Delta x_k$$

Le premier algorithme de ce type fût défini par DAVIDON [8] en 1959, transformé par la suite par FLETCHER et POWELL [13].

Dans le DAVIDON-FLETCHER-POWELL (D.F.P) algorithme la matrice H_k est révisée à chaque itération par la formule :

$$H_{k+1} = H_k - \frac{H_k.\Delta g_k.\Delta g_k H_k}{\Delta g_k.H_k\Delta g_k} + \frac{\Delta x_k.\Delta x_k^t}{\Delta x_k.\Delta g_k}$$ qui vérifie l'équation de quasi-Newton

le coefficient de pas α_k étant choisi par maximisation exacte dans la direction $s_k = H_k.\nabla f(x_k)$.

Tout au long des années 1960, différentes méthodes furent décrites chacune proposant une formule particulière de "mise à jour" de la matrice H_k.

En 1970, HUANG [18] publia une approche plus générale dans le but de construire des algorithmes tels que :

- seules des recherches unidimensionnelles soient utilisées
- appliqué à une fonction quadratique le maximum est atteint en un nombre d'itérations inférieur ou égal au nombre de variables
- à chaque itération ne sont utilisées que des informations portant sur le point considéré et l'itéré précédent
- seuls la fonction et le gradient sont utilisés

Huang obtînt ainsi une formule de mise à jour générale où, selon la valeur de coefficients, on retrouve la plupart des formules existantes.

III.2 Résultats de convergence

Jusqu'à 1971, hormis la propriété de terminaison quadratique aucun résultat de convergence des algorithmes de métrique variable n'avait été obtenu, la justification de ceux-ci ne se faisant que par l'appréciation des résultats numériques.

En 1971 POWELL [30] montre le résultat suivant :

<u>si</u> : f est 2 fois continument différentiable de Hessien uniformément défini négatif :
$\exists m \in \mathbb{R}^{+*}$ tel que $\forall x$: $0 \le m\,||y||^2 < -y.\nabla^2 f(x).y$ qui vérifie une condition de Lipschitz en $\overset{*}{x}$ unique maximum de f sur $\mathbb{R}^n$:
$\exists L > 0$ tel que : $||\nabla^2 f(x)-\nabla^2 f(\overset{*}{x})|| \le L.||x-\overset{*}{x}||$ si à chaque itération α_k est obtenu par maximisation exacte sur la direction $H_k.\nabla f(x_k)$

<u>alors</u> :
$\forall x_0 \in \mathbb{R}^n$, si $(-H_0)$ est symétrique définie positive la méthode de D.F.P converge Q superlinéairement vers x, soit :

$$\lim_{k\to\infty} \frac{||x_{k+1}-\overset{*}{x}||}{||x_k-\overset{*}{x}||} = 0$$

En 1972, POWELL montre sous des hypothèses plus faibles

- f 2 fois continument différentiable, concave
- $\{x \in \mathbb{R}^n \,/\, f(x) \ge f(x_0)\}$ borné

la convergence vers un maximum de f de la méthode de D.F.P toujours avec maximisation unidimensionnelle exacte.

Le résultat démontré par DIXON [12] en 1972 : toutes les méthodes de quasi-Newton construites à l'aide d'une formule issue de la classe de Huang, et utilisant à chaque itération une maximisation exacte engendrent les mêmes points ; cette propriété permet, alors d'étendre, à tous les algorithmes de la classe de Huang, le résultat précédent de convergence démontré par POWELL.

Les propriétés essentielles sur les méthodes de quasi-Newton ont été obtenues quand à chaque itération le pas α_k est posé égal à 1, ou résulte d'une optimisation exacte. Néanmoins on peut trouver dans POLAK [30] en 1971 et LENARD [44] en 1975 des études de convergence de ces méthodes, sans utiliser de maximisation exacte. Sur la vitesse de convergence les travaux les plus importants à ce jour, restent ceux de DENIS et MORE [11] (1974) qui donnent une caractérisation de la vitesse Q-superlinéaire de convergence.

<u>Théorème</u>

Soit $\overset{*}{x} \in D$ ouvert convexe de $\mathbb{R}^n$

$F : \mathbb{R}^n \to \mathbb{R}^n$ différentiable sur D

F' continue en $\overset{*}{x}$, $[F'(\overset{*}{x})]^{-1}$ existe

$\{H_k\}_{k \in \mathbb{N}}$ une suite de matrices régulières

$\{x_k\}_{k \in \mathbb{N}}$ une suite de points de D convergent vers $\overset{*}{x}$, telle que :

$$\boxed{x_{k+1} = x_k - H_k^{-1}.F(x_k)}$$

<u>alors</u> :

$\{x_k\}_{k \in \mathbb{N}}$ converge Q-superlinéairement vers $\overset{*}{x}$ et $F(\overset{*}{x}) = 0$ si et seulement si :

$$\boxed{\lim_{k\to\infty} \frac{\|(H_k - F'(\overset{*}{x}))(x_{k+1} - x_k)\|}{\|x_{k+1} - x_k\|} = 0}$$

De plus DENIS et MORE montrent que sous les mêmes conditions, si :

- $\lim_{k\to\infty} \dfrac{\|(H_k - F'(\overset{*}{x}))(x_{k+1} - x_k)\|}{\|x_{k+1} - x_k\|} = 0$

- $x_{k+1} = x_k - \alpha_k.H_k^{-1}.F(x_k)$

alors :

la suite engendrée converge vers $\overset{*}{x}$ ($F(\overset{*}{x}) = 0$) Q-superlinéairement si et seulement si :

$$\lim_{k\to\infty} (\alpha_k) = 1$$

Ces résultats de DENIS et MORE sont fondamentaux tant pour l'optimisation sans contraintes que par la possibilité de les appliquer à l'optimisation avec contraintes.

III.3 L'adaptation des algorithmes de métrique variable à l'optimisation avec contraintes

Soit le problème

$$P : \boxed{\begin{array}{l} \text{Max } f(x) \\ g_i(x) \geq 0 \\ i = 1 \text{ à } m \end{array}} \qquad f,\ g_i : \mathbb{R}^n \to \mathbb{R}$$

trouver un point stationnaire pour ce problème consiste à trouver un couple $(\overset{*}{x},\overset{*}{u})$ de $\mathbb{R}^{n+m}$, constitué d'une solution et du multiplicateur associé, qui vérifie les conditions nécessaires de Kühn et Tucker :

$$\begin{cases} \overset{*}{u} \geq 0 \\ \overset{*}{u}.\nabla g(\overset{*}{x}) + \nabla f(\overset{*}{x}) = 0 \\ \overset{*}{u}^i.g_i(\overset{*}{x}) = 0,\ \forall i = 1,2,\ldots m \end{cases}$$

<u>avec</u>

$\forall i = 1$ à m

$\overset{*}{u}^i = i^{\text{ème}}$ composante de $\overset{*}{u}$

$g_i(\overset{*}{x}) = i^{\text{ème}}$ composante de $g(\overset{*}{x})$

$\nabla^t g_i(\overset{*}{x}) = i^{\text{ème}}$ ligne de $\nabla g(x)$

Soit L le Lagrangien associé au problème (P) :

$$L(x,u) = f(x) + u.g(x) = f(x) + \sum_{i=1}^{m} u^i.g_i(x)$$

$$\nabla_x L(x,u) = \nabla f(x) + u.\nabla g(x) = \nabla f(x) + \sum_{i=1}^{m} u^i.\nabla g_i(x)$$

en posant

$$F : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m} \quad F(x,u) = \begin{bmatrix} \nabla x L(x,u) \\ u^i.g_i(x) \end{bmatrix}$$

trouver un point stationnaire pour (P) consiste à trouver une solution $t^* = (\overset{*}{x},\overset{*}{u})$ où $\overset{*}{u} \geq 0$ au système :

$$\boxed{F(t) = F(x,u) = 0}$$

Si l'on cherche une solution t^* de ce système par la méthode de Newton, on construit une suite $(t_k)_{k \in \mathbb{N}}$ de $\mathbb{R}^{n+m}$ telle que :

$$① : \boxed{F(t_k) + \nabla_t F(t_k).(t_{k+1} - t_k) = 0}$$

où $\nabla_t F$ désigne le Jacobien de F :

$$\nabla_t F(t) = \nabla_t F(x,u) = \begin{array}{c|c|c} n & m & \\ \hline \nabla_{xx} L(x,u) & \nabla g^t(x) & n \\ \hline u_i.\nabla g_i(x) & g_i(x) & m \end{array}$$

avec $\nabla_{xx} L(x,u)$ le Hessien du Lagrangien par rapport à la variable x.

Si on développe le système ① on obtient :

$$\begin{cases} \nabla_x L(t_k) + \nabla_{xx} L(t_k)(x_{k+1}-x_k) + \nabla_g(x_k)(u_{k+1} - u_k) = 0 \\ u_k^i.g_i(x_k) + u_k^i \nabla g_i(x_k)(x_{k+1}-x_k)+g_i(x_k)(u_{k+1}^i-u_k^i) = 0 \end{cases}$$

que l'on met sous la forme

$$①' : \begin{cases} \nabla_x L(x_k,u_{k+1})+\nabla_{xx} L(x_k,u_k)(x_{k+1}-x_k) = 0 \\ u_{k+1}^i[g_i(x_k)+\nabla g_i(x_k)(x_{k+1}-x_k)] - (u_{k+1}^i-u_k^i).\nabla g_i(x_k)(x_{k+1}-x_k) = 0 \quad i = 1 \text{ à } m \end{cases}$$

En fait on ne va pas résoudre le système ① à chaque itération, mais par une méthode de quasi-Newton :

$$② : \boxed{F(t_k) + H_k(t_{k+1}-t_k) = 0}$$

où H_k sera une approximation de la matrice Jacobienne $\nabla_t F(t_k)$. Le problème est de déterminer une suite de matrices $(H_k)_{k \in \mathbb{N}}$ telles que :

- à chaque itération on puisse effectivement déterminer une solution $t_{k+1} = (x_{k+1}, u_{k+1})$ au système ②
- la suite $(t_k)_{k \in \mathbb{N}}$ converge vers t^* tel que $F(t^*) = 0$

Différents algorithmes, dont ceux de WILSON [32], ROBINSON [33] ainsi que GARCIA-PALOMARES et MANGASARIAN [15] utilisent cette approche pour résoudre le problème (P).
Soit :

$$P(x_k,-H_k) : \boxed{\begin{array}{l} \text{Max } \nabla f(x_k)(x-x_k) + \frac{1}{2}(x-x_k)H_k(x-x_k) \\ g_i(x_k)+\nabla g_i(x_k)(x-x_k) \geq 0 \\ i = 1 \text{ à } m \end{array}}$$

ce problème quadratique où les contraintes sont les linéarisées en x_k des contraintes du problème (P) est celui étudié au chapitre II (II.2).
On supposera par commodité d'écriture que H_k est symétrique mais cette condition n'est nullement nécessaire dans ce qui suit.

Conditions de Kühn et Tucker pour $P(x_k,-H_k)$: soient x_{k+1} un point stationnaire et $u_{k+1} \in \mathbb{R}^m$ le multiplicateur associé :

$$\begin{cases} u_{k+1} \geq 0 \\ u_{k+1}.\nabla g(x_k)+\nabla f(x_k) + H_k(x_{k+1}-x_k) = 0 \\ u^i_{k+1}[g_i(x_k) + \nabla g_i(x_k)(x_{k+1}-x_k)] = 0 \quad i = 1 \text{ à } m \end{cases}$$

la résolution du problème quadratique à contraintes linéaires $P(x_k,-H_k)$ fournit à partir de x_k un couple (x_{k+1},u_{k+1}) de $\mathbb{R}^{n+m}$ solution du système :

$$② : \begin{cases} \nabla_x L(x_k,u_{k+1}) + H_k(x_{k+1}-x_k) = 0 \\ u^i_{k+1} [g_i(x_k)+\nabla g_i(x_k).(x_{k+1}-x_k)] = 0 \quad i = 1 \text{ à } m \end{cases}$$

système que l'on peut mettre sous la forme matricielle :

$$② : \boxed{F(t_k) + \mathcal{H}_k(t_{k+1}-t_k) = 0}$$

avec :

$$\mathcal{H}_k = \begin{array}{c|c|c} \overset{n}{} & \overset{m}{} & \\ \hline H_k & \nabla^t_g(x_k) & n \\ \hline u^i.\nabla g_i(x_k) & g_i(x_k)+\nabla g_i(x_k)(x_{k+1}-x_k) & m \end{array}$$

si l'on compare $\mathcal{H}_k$ et $\nabla_t F(t_k)$, on remarque que si $\lim_{k\to\infty} (x_{k+1}-x_k) = 0$ il suffit que H_k soit une approximation de $\nabla_{xx} L(x_k,\mu_k)$ pour que $\mathcal{H}_k$ constitue une approximation de $\nabla_t F(t_k)$.

$$\mathcal{H}_k-\nabla F_t(t_k) = \begin{array}{c|c|c} \overset{n}{} & \overset{m}{} & \\ \hline H_k-\nabla_{xx}L(t_k) & 0 & n \\ \hline 0 & R_k & m \end{array}$$

où R_k est la matrice diagonale :

$$R_k = \boxed{\nabla g_i(x_k)(x_{k+1}-x_k)}$$

La structure particulière de $H_k - \nabla F_t(t_k)$ sera utilisée pour simplifier les démonstrations de convergence des différents algorithmes construits à partir de ce principe.

C'est le cas par exemple de la condition nécessaire et suffisante de DENIS et MORE (III.2) qui peut être transformée en une condition suffisante ne portant que sur H_k.
On suppose que toutes les conditions d'application du théorème de DENIS et MORE sont vérifiées, notamment que la suite $(t_k)_{k \in \mathbb{N}}$ converge vers t^*.

$$\frac{(H_k - \nabla_t F(t_k))(t_{k+1}-t_k)}{t_{k+1}-t_k} = \frac{(H_k - \nabla_{xx} L(t_k))(x_{k+1}-x_k),\ R_k(u_{k+1}-u_k)}{t_{k+1}-t_k}$$

$$\frac{||(H_k - \nabla_t F(t_k))(t_{k+1}-t_k)||}{||t_{k+1}-t_k||} \leq \frac{||(H_k - \nabla_{xx} L(t_k))(x_{k+1}-x_k)||}{||t_{k+1}-t_k||} + \frac{||R_k(u_{k+1}-u_k)||}{||t_{k+1}-t_k||}$$

$$\leq \frac{||(H_k - \nabla_{xx} L(t_k))(x_{k+1}-x_k)||}{||x_{k+1}-x_k||} + \frac{||R_k(u_{k+1}-u_k)||}{||u_{k+1}-u_k||}$$

$$\lim_{k\to\infty} \frac{||(H_k - \nabla_t F(t_k))(t_{k+1}-t_k)||}{||t_{k+1}-t_k||} \leq \lim_{k\to\infty} \frac{||(H_k - \nabla_{xx} L(t_k))(x_{k+1}-x_k)||}{||x_{k+1}-x_k||} + \lim_{k\to\infty} ||R_k||$$

en remarquant que lorsque $k \to \infty$; $t_k \to t^*$ donc $x_k \to \overset{*}{x}$ chaque terme de la matrice diagonale R_k tend vers 0. donc :

$$\lim_{k\to\infty} ||R_k|| = 0$$

$$\boxed{\lim_{k\to\infty} \frac{||(H_k - \nabla_{xx} L(x_k, u_k))(x_{k+1}-x_k)||}{||x_{k+1}-x_k||} = 0}$$

est une condition suffisante pour que, avec les hypothèses formulées, la convergence de $(t_k)_{k \in \mathbb{N}}$ vers t^* ait une vitesse superlinéaire. Cette condition n'est cependant pas nécessaire et POWELL [32] en a récemment utilisée une autre.

III.4 L'algorithme de Han [17]

Pour trouver $t^* = (\overset{*}{x}, \overset{*}{u})$ un point stationnaire de (P) et le multiplicateur associé, Han décrit l'algorithme suivant :

- soit (x_0, u_0) une approximation de $(\overset{*}{x}, \overset{*}{u})$
- soit H_0 une approximation de $\nabla_{xx} L(\overset{*}{x}, \overset{*}{u})$

Itération k

Trouver un point stationnaire x_{k+1} et le multiplicateur associé u_{k+1} au sous-problème quadratique à contraintes linéaires $P(x_k,-H_k)$. S'il existe plusieurs solutions on choisit pour x_{k+1} la solution la plus proche de x_k (au sens de la norme euclidienne)

si : $x_{k+1} = x_k$ STOP

sinon : x_{k+1} est le successeur de x_k
calcul de H_{k+1} par une formule de métrique variable

Les hypothèses suivantes sont faites par Han, elles assurent que $[\nabla_t F(t^*)]^{-1}$ existe :

- $I(\overset{*}{x}) = \{i \in \{1,\dots m\} / g_i(\overset{*}{x}) = 0\}$; $i \in I(\overset{*}{x}) \Longrightarrow \overset{*}{u}_i > 0$
  condition de stricte complémentarité

- $\{\nabla g_i(\overset{*}{x})\}_{i \in I(\overset{*}{x})}$ linéairement indépendant

- $\forall y \in \mathbb{R}^{n*}$, tel que : $\nabla g_i(\overset{*}{x}).y = 0$ pour $i \in I(\overset{*}{x})$ alors :

$$y.\nabla_{xx}L(\overset{*}{x},\overset{*}{u}).y < 0$$

On note $h \in LC^2(\overset{*}{x})$ une fonction dont la dérivée seconde vérifie la condition le Lipschitz suivante :

$$\exists C > 0,\ \exists v(\overset{*}{x}) \text{ voisinage de } \overset{*}{x} \text{ tel que : } \forall y \in v(\overset{*}{x})$$
$$||h''(y)-h''(\overset{*}{x})|| \leq C||y-\overset{*}{x}||$$

Han suppose que $f \in LC^2(\overset{*}{x})$ et $\forall i = 1$ à m, $g_i \in LC^2(\overset{*}{x})$

Sous ces hypothèses la démonstration de Han utilise essentiellement les propriétés suivantes :

Ⓐ $$\lim_{k\to\infty} \frac{||(\nabla_{xx}L(t^*)-H_k)(x_{k+1}-x_k)||}{||t_{k+1}-t_k||} = 0$$

Ⓑ $$\sum_k ||t_k-t^*|| < \infty$$

Ⓒ $$||H_{k+1}-\nabla_{xx}L(t^*)||' \leq (1+\alpha_1||t_k-t^*||).||H_k-\nabla_{xx}L(t^*)||' + \alpha_2.||t_k-t^*||$$

où $||\ \ ||'$ désigne une norme quelconque

Ⓓ $$||H_{k+1} - \nabla_{xx}L(t^*)||' \leq (1-\sigma_k+\alpha_1||t_k-t^*||).||H_k-\nabla_{xx}L(t^*)||'+\alpha_2||t_k-t^*||$$

Ⓔ $||H_k-\nabla_{xx}L(t^*)|| \to 0$ ou $\sigma_k \to 0$

σ_k est un scalaire qui dépend de H_k, $x_{k+1}-x_k$, $\nabla_{xx}L(t^*)$ et M(voir Ⓖ)

Ⓕ pour simplifier l'écriture on n'écrira pas l'indice k qui devrait figurer à toutes les variables se trouvant à droite de l'égalité

$$\text{successeur de } H_k = H - \frac{(\Delta g+H.\Delta x)c^t+c(\Delta g+H\Delta x)^t}{c.\Delta x} + \frac{\Delta^t x(\Delta g+H\Delta x)cc^t}{(c.\Delta x)^2}$$

où $\Delta g = \Delta g_k = \nabla_x L(x_{k+1},u_{k+1}) - \nabla_x L(x_k,u_{k+1})$

$\Delta x = \Delta x_k = x_{k+1} - x_k$

$c = c_k$ un vecteur quelconque tel que $\underline{c}_k . \Delta x_k \neq 0$

Ⓖ $\exists M$ inversible, telle que :

$$\frac{||Mc_k - M^{-1}.\Delta x_k||}{||M^{-1}.\nabla x_k||} \leq \mu.\max\{||t_k - t^*||, ||t_{k+1} - t^*||\}$$

Convergence Q-linéaire locale

$\exists \varepsilon_1 > 0$, $\exists \varepsilon_2 > 0$ tels que si : $||t_0 - t^*|| \leq \varepsilon_1$; $||H_0 - \nabla_{xx} L(t^*)|| \leq \varepsilon_2$ alors la suite $(t_k)_{k \in \mathbb{N}}$ construite par l'algorithme vérifie :

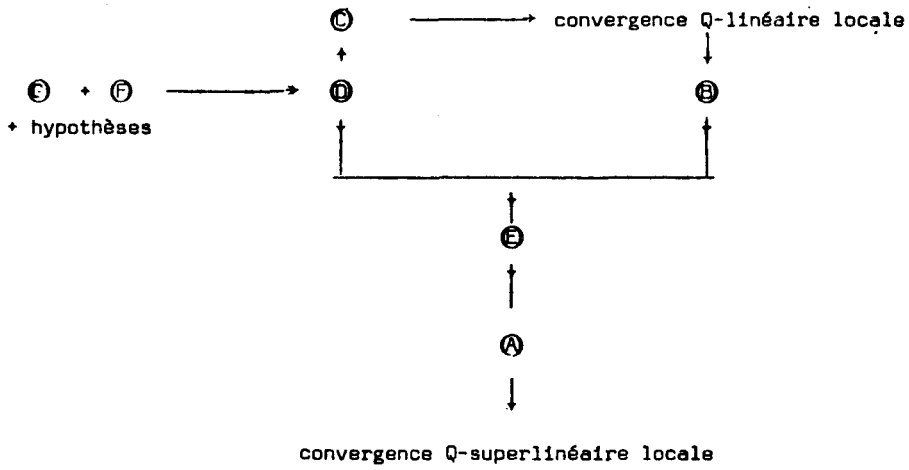
$$\lim_{k \to \infty} \frac{||t_{k+1} - t^*||}{||t_k - t^*||} \in [0,1[$$

Convergence Q-superlinéaire locale

idem sauf :

$$\lim_{k \to \infty} \frac{||t_{k+1} - t^*||}{||t_k - t^*||} = 0$$

A partir de ces propriétés les démonstrations de Han peuvent être schématisées de la manière suivante :



La plupart des démonstrations sont analogues à celles données dans le cas sans contraintes par BROYDEN , DENIS et MORE [2] (1973), DENIS et MORE [11] (1974), mais aussi très liées à l'étude faite par GARCIA-PALOMARES et MANGASARIAN [15] (1976).
Cependant Han n'utilise pas le résultat de DENIS et MORE pour montrer comme en III.3 :

A ==> convergence Q-superlinéaire locale

il redémontre cette propriété en introduisant des conditions suffisantes intermédiaires particulières aux algorithmes à contraintes linéarisées.

III.5 Comparaison entre l'algorithme de Han et l'algorithme général réalisable du chapitre II

La liaison entre les deux algorithmes est très étroite, puisqu'ils sont tous deux basés sur la résolution répétitive du même sous problème quadratique $P(x_k,-H_k)$.
Cependant Han prend comme successeur de x_k une solution de ce sous-problème, alors que l'algorithme du chapitre II introduit ensuite une perturbation par le procédé de rentrée décrit à cette occasion : ce qui permet de trouver x_{k+1} à partir de z_k.
Le parallèle entre les 2 méthodes avec les différences entre hypothèses, algorithmes, calculs intermédiaires et résultats peut être résumé de la manière suivante :

ALGORITHME DE HAN	ALGORITHME REALISABLE DU CHAPITRE II
successeur de (x_k) = la solution de $P(x_k,-H_k)$ la plus proche de x_k.	successeur de (x_k) : soit z_k une solution quelconque de $P(x_k,-H_k)$; x_{k+1} est un maximum de f sur $\Delta(x_k,z_k)$
méthode par points non forcément réalisables	méthode par points réalisables
l'assurance de l'existence d'une solution de $P(x_k,-H_k)$ à distance finie est assurée localement	l'existence d'une solution de $P(x_k,-H_k)$ à distance finie est assurée par : $C'(x) \subset B$ compact
nécessité théorique de calculer toutes les solutions de $P(x_k,-H_k)$ puis la plus proche de x_k	calcul d'une solution quelconque z_k mais nécessité d'effectuer en plus une sous-optimisation sur $\Delta(x_k,z_k)$
H_k calculée grâce à Ⓕ $\|\|H_0-\nabla_{xx}L(t^*)\|\| \leq \epsilon$	$(-H_k)$ semi-définie positive $\exists N1 \in \mathbb{N}$ telle que : $(H_k)_{k \in N_1} \to \overset{*}{H}$
la fonction économique de $P(x_k,-H_k)$ est quadratique	la fonction économique de $P(x_k,-H_k)$ est quadratique concave
directions non forcément améliorantes	algorithme de directions améliorantes
convergence locale Q-superlinéaire	convergence globale vitesse de convergence non connue

Le problème essentiel qui reste posé est le suivant :

en quoi le procédé de rentrée peut-il perturber la vitesse de convergence superlinéaire obtenue par Han ?

Pour l'instant aucune réponse n'est apportée à cette question. Il semble qu'une approche directe de l'étude de la vitesse de convergence pour l'algorithme du chapitre II soit difficile : on remarquera à ce propos que tous les résultats de vitesse superlinéaire de convergence obtenus pour les algorithmes de métrique variable appliquée à l'optimisation avec contraintes sont locaux. C'est le cas pour tous les travaux qui s'échelonnent de ceux de Broyden, Denis et Moré en 1972 à ceux plus récents de POWELL [32] en 1977.

L'approche qui semble la plus prometteuse pour démontrer l'éventuelle convergence globale superlinéaire de l'algorithme du chapitre II est d'essayer de déduire ce résultat de celui de Han. Cependant, bien que le procédé de construction présente une analogie, on ne retrouve pas, générée par l'algorithme la suite de points obtenus par Han et la connexion entre les deux méthodes reste à établir.

Les approches restent à ce jour très différentes et les résultats peuvent être considérés comme complémentaires :

- Han obtient une convergence locale mais superlinéaire
- l'algorithme du chapitre II construit une suite réalisable qui converge globalement mais avec une vitesse de convergence inconnue actuellement

*

* *

L'algorithme exposé au chapitre II a été programmé dans le cas particulier, simple suivant :

x_k itéré considéré

$y_k = x_k + \alpha.\nabla f(x_k)$

$\alpha \in \mathbb{R}^{+*}$ paramètre fixé

$z_k = \mathrm{Prog}_{C'(x_k)}\{y_k\}$

la projection euclidienne du point y_k sur le linéarisé du domaine en x_k : $C'(x_k)$

$$\Delta(x_k,z_k) = \overline{\{x_k + \frac{z_k-x_k}{2^i} + \frac{x_o-z_k}{2^{i+j}} \;/\; (i,j) \in \mathbb{N}^2\}} \cap C$$

x_o un point fixé, intérieur au domaine

chaque itération de l'algorithme considéré comporte ainsi deux phases principales :

- la projection euclidienne d'un point sur un polyèdre
- une sous-optimisation sur la partie réalisable d'une grille construite à partir de x_k, z_k et x_o.

On trouvera ci-après :

- le programme réalisé en FORTRAN
- les résultats obtenus sur les trois premiers problèmes tests de COLVILLE [5] par traitement sur l'ordinateur 10070 CII du Laboratoire de Calcul de l'Université de Lille I

Le critère de comparaison entre algorithmes, défini par COLVILLE et utilisé depuis, consiste à mesurer pour certains problèmes tests, le temps de calcul nécessaire à l'obtention d'une précision fixée.
Pour tenir compte des performances différentes des ordinateurs, c'est en fait un "temps standard" qui est mesuré : rapport du temps réel de calcul au temps mis par l'ordinateur utilisé pour traiter un programme standard fixé par COLVILLE (ce temps est de 117 s pour le 10070 CII).

Les résultats obtenus sont regroupés dans le tableau suivant :

Problèmes de COLVILLE	I	II	III
degré de la fonction économique	3	3	2
nombre de variables	5	15	5
nombre de contraintes linéaires	20	30	10
nombre de contraintes non linéaires	0	5	8
précision du rapport COLVILLE	- 32.348	- 32.349	-30665.5
meilleur temps standard du rapport COLVILLE	0.0058	0.0840	0.0051
temps standard obtenu ici	0.0068	0.4478	0.0025
valeur du paramètre pour l'obtention du temps standard	1.6	60	1
nombre d'itérations	5	64	3
meilleur résultat obtenu	- 32.34867897	- 32.348679	- 30665.5387
valeur du paramètre pour obtenir le meilleur résultat	1.7	30	1

BUS LILLE

On peut analyser ces résultats sous deux aspects :

- la précision obtenue : elle est aussi bonne sur les 3 problèmes que les meilleurs résultats cités dans le rapport COLVILLE

- le temps standard : les résultats sont variables selon les 3 problèmes : très bon pour le problème III, bon pour le problème I, médiocre pour le problème II.

Le bilan global est néanmoins satisfaisant, compte tenu du fait que la programmation est certainement améliorable et que le domaine de sous-optimisation a été choisi, dans ce cas, de la manière la plus élémentaire.

Le seul inconvénient majeur qui limite la portée de cet algorithme est introduit par le contrôle du paramètre α. On constate dans l'étude qui suit la forte influence de ce paramètre sur la vitesse de convergence : variation avec α du temps exact (en 1/10 de seconde) et du nombre d'itérations requis pour obtenir la précision -32.348 dans le problème de COLVILLE I :

α	1.3	1.4	1.5	1.6	1.65	1.70	1.75	1.80	1.90	2.00	3.00
temps exact	12	15	13	8	12	14	9	19	16	13	22
nb. d'itérations	6	7	6	5	7	6	6	8	6	7	9

Or aucun critère ne semble utilisable pour la détermination de ce paramètre, qui sur les 3 problèmes étudiés varie de 1 à 60.
Par contre, l'utilisation dans l'algorithme d'une formule de métrique variable, au lieu de cette méthode de gradient pur projeté, fournirait :

$$y_k = x_k - \alpha . H_k . \nabla f(x_k)$$

et le problème du contrôle du paramètre, serait sans doute moins délicat : on pourrait espérer, en liaison avec le résultat de DENIS et MORE (III.2), obtenir la convergence la plus rapide avec des valeurs de α voisines de 1.

```
      SUBROUTINE PROJ (N, M,C,D,BDL,BDU,X)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION Y(15),Z(15),QC(15),CH(15),E(15),QLS(15),LT(35)
      DIMENSION BDL(15),BDU(15),X(15),D(5)
      DIMENSION C(15,5),H(15,15),Q(15,15)
      NN=N+N
1     DO 2 I=1,M
2     LT(I)=1
      K=0
      DO 3 I=1,N
      DO 3 J=1,N
      Q(I,J)=0.
3     H(I,J)=0.
      DO 4 I=1,N
4     H(I,I)=1.
C
C     CALCUL DE Z PROJECTION DU GRADIENT
  100 DO 101 I=1,N
      Z(I)=0.
      DO 101 J=1,N
  101 Z(I)=Z(I)-H(I,J)*X(J)
C
C     TEST SI Z EST NUL
      DO 102 I=1,N
  102 IF(ABS(Z(I)).GT.1E-6) GO TO 103
      GO TO 20
C
C     CALCUL DE Y
  103 DO 104 I=1,N
  104 Y(I)=Z(I)+X(I)
C     TEST DE REALISABILITE
      DO 55 I=1,M
      IF(LT(I).EQ.0) GO TO 55
      IF(I.GT.N) GO TO 51
      GAMA=Y(I)-BDL(I)
      GOTO 54
   51 IF(I.GT.NN) GO TO 52
      GAMA=BDU(I-N)-Y(I-N)
      GOTO 54
   52 J=I-NN
      GAMA=0.
      DO 53 L=1,N
   53 GAMA=GAMA+C(L,J)*Y(L)
      GAMA=GAMA-D(J)
   54 IF(GAMA.LT.0.) GOTO 60
   55 CONTINUE
C     POINT REALISABLE
      DO 105 I=1,N
  105 X(I)=Y(I)
      GOTO 20
```

BUS LILLE

```
C       TEST D'OPTIMALITE
   20 U=-1E50
      DO 22 I=1,K
      RES=0.
      DO 21 J=1,N
   21 RES=RES-Q(I,J)*X(J)
      IF(RES.LE.U) GOTO 22
      U=RES
      II=I
      INDS=E(II)
   22 CONTINUE
      IF(U.LE.0.) RETURN
C     DIMINUTION DE LA BASE
      IF(K.EQ.1) GOTO 1
      ALFA=0.
      DO 31 I=1,N
   31 ALFA=ALFA+Q(II,I)*Q(II,I)
      ALFA=1./ALFA
      DO 32 I=1,N
      DO 32 J=1,I
      H(I,J)=H(I,J)+ALFA*Q(II,I)*Q(II,J)
   32 H(J,I)=H(I,J)
      IF(II.EQ.K) GOTO 34
      DO 33 I=1,N
      QLS(I)=Q(II,I)
   33 Q(II,I)=Q(K,I)
   34 K=K-1
      DO 36 I=1,K
      DO 36 J=1,N
      VAL=0.
      DO 35 L=1,N
35    VAL=VAL+Q(I,L)*QLS(L)*QLS(J)
   36 Q(I,J)=Q(I,J)-ALFA*VAL
      LT(INDS)=1
      DO 37 I=II  ,K
37    E(I)=E(I+1)
      GOTO 100
      END
```

```
C       CALCUL DE ALFA
   60   ALFA=1E50
        DO 65 I=1,M
        IF(LT(I).EQ.0) GOTO 65
        IF(I.GT.N) GOTO 61
        IF(Z(I).GE.0.) GOTO 65
        CC=(BDL(I)-X(I))/Z(I)
        GOTO 64
   61   IF(I.GT.NN) GOTO 62
        IF(Z(I-N).LE.0.) GOTO 65
        CC=(BDU(I-N)-X(I-N))/Z(I-N)
        GOTO 64
   62   J=I-NN
        V1=0.
        DO 63 L=1,N
63      V1=V1+C(L,J)*Z(L)
        IF(V1.GE.0.) GOTO 65
        V2=0.
        DO 66 L=1,N
   66   V2=V2+C(L,J)*X(L)
        CC=(D(J)-V2)/V1
   64   IF(CC.GT.ALFA) GOTO 65
        ALFA=CC
        IAL=I
   65   CONTINUE
C
C       DEPLACEMENT
        DO 106 I=1,N
  106   X(I)=X(I)+ALFA*Z(I)
C
C       AUGMENTATION CAS SPECIAL D'UNE SEULE CONTRAINTE
        IF(K.GT.0) GO TO 80
        IF (IAL.GT.N) GOTO 10
        Q(1,IAL)=1.
        H(IAL,IAL)=0.
        E(1)=IAL
        GOTO 200
10      IB=IAL-N
        IF (IB.GT.N) GOTO 72
        Q(1,IB)=-1.
        H(IB,IB)=0.
        E(1)=IAL
        GOTO 200
   72   IB=IAL-NN
        V3=0.
        DO 73 L=1,N
   73   V3=V3+C(L,IB)*C(L,IB)
        DO 74 I=1,N
   74   Q(1,I)=C(I,IB)/V3
        DO 75 I=1,N
        DO 75 J=1,I
        H(I,J)=H(I,J)-C(I,IB)*C(J,IB)/V3
   75   H(J,I)=H(I,J)
        E(1)=IAL
        GOTO 200
```

BUS LILLE

```
C       AUGMENTATION CAS GENERAL
C       ON UTILISE QC(/K/);CH(/N/);CHC=SCALAIRE
   80   IF(IAL.GT.N) GOTO 83
        DO 81 J=1,K
   81   QC(J)=Q(J,IAL)
        DO 82 J=1,N
   82   CH(J)=H(IAL,J)
        CHC=CH(IAL)
        GOTO 90
   83   IB=IAL-N
        IF(IB.GT.N) GOTO 86
        DO 84 J=1,K
   84   QC(J)=-Q(J,IB)
        DO 85 J=1,N
   85   CH(J)=-H(IB,J)
        CHC=-CH(IB)
        GOTO 90
   86   IB=IB-N
        DO 87 J=1,K
        QC(J)=0.
        DO 87 L=1,N
   87   QC(J)=QC(J)+Q(J,L)*C(L,IB)
        DO 88 J=1,N
        CH(J)=0.
        DO 88 L=1,N
   88   CH(J)=CH(J)+C(L,IB)*H(L,J)
        CHC=0.
        DO 89 I=1,N
   89   CHC=CHC+CH(I)*C(I,IB)
   90   CONTINUE
C       CALCUL DE Q
        DO 151 J=1,N
        BETA=CH(J)/CHC
        Q(K+1,J)=BETA
        DO 151 I=1,K
  151   Q(I,J)=Q(I,J)-QC(I)*BETA
C       CALCUL DE H
        IF(K.LT.N-1) GOTO 162
        DO 161 I=1,N
        DO 161 J=1,N
  161   H(I,J)=0.
        E(N)=IAL
        GOTO 200
  162   DO 163 I=1,N
        DO 163 J=1,I
        H(I,J)=H(I,J)-CH(I)*CH(J)/CHC
  163   H(J,I)=H(I,J)
C
C       INCREMENTATION
  200   LT(IAL)=0
        K=K+1
        E(K)=IAL
        GOTO 100
```

```
      SUBROUTINE GRILLE(NV,NC,FECO,DDFECO,KPPART,X0,X,Z,MAX,COOMAX)
C     MAXIMISATION SUR LA GRILLE DE SOMMETS X,X0,Z
C     LE POINT OU LE MAXIMUM EST ATTEINT: COOMAX
C     LE MAXIMUM DE FECO EST: MAX
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION X(NV),X0(NV),Z(NV),COOMAX(NV)
      DIMENSION POINT(15)
      DOUBLE PRECISION M(15),N(15),MAX
      ITEST=3
      K=0
C     RECHERCHE DU POINT DE BASE
      JB=0
      DO 1 I=1,NV
1     N(I)=X0(I)
2     JB=JB+1
      DO 3 I=1,NV
3     N(I)=(N(I)+Z(I))/2.
      IF (DDFECO(X,N,NV).LE.0.) GO TO 2
      IF (KPPART(N,NV,NC,1).LT.0 ) GO TO 4
      IB=0
      GO TO 7
4     IB=0
5     IB=IB+1
      DO 6 I=1,NV
6     N(I)=(N(I)+X(I))/2.
      IF (KPPART(N,NV,NC,1).GT.0 ) GO TO 7
      GO TO  5
7     CONTINUE
C
C
      JMAX=0.
      IMAX=0.
      MAX=FECO(X)
      DO 21 I=1,NV
21    COOMAX(I)=X(I)
      S=FECO(N)
      IF(S.LE.MAX) GOTO 23
      MAX=S
      DO 22 I=1,NV
22    COOMAX(I)=N(I)
23    CONTINUE
      KV=2**IB
      DO 8 I=1,NV
8     M(I)=X(I)+(Z(I)-X(I))/FLOAT(KV)
      IC=0
9     IF (KPPART(M,NV,NC,1).GT.0 ) GO TO 200
C
C
C     M NON REALISABLE
100   S=FECO(N)
      IF (S.LE.MAX) GO TO 11
      MAX = S
      DO 10 I=1,NV
10    COOMAX(I)=N(I)
      IMAX=IC
      JMAX=0
11    JL=1
      DO 12 I=1,NV
12    POINT(I)=N(I)
13    DO 14 I=1,NV
14    POINT(I)=(POINT(I)+M(I))/2.
      IF(JL.LT.K) GOTO 15
      IF (KPPART(POINT,NV,NC,1).LT.0 ) GOTO 18
      K=JL
```

BUS LILLE

```
15      IF (FECO(POINT).LE.S) GOTO 18
        S=FECO(POINT)
        IF (S.LE.MAX) GOTO 17
        MAX=S
        DO 16 I=1,NV
16      COOMAX(I)=POINT(I)
        IMAX=IC
        JMAX=JL
17      JL=JL+1
        GOTO 13
18      CONTINUE
        IC=IC+1
        IF (IC.GE.ITEST) GOTO 50
        DO 19 I=1,NV
        M(I)=(M(I)+X(I))/2.
19      N(I)=(N(I)+X(I))/2.
        GOTO 9
C
C
C       M REALISABLE
200     IF (DDFECO(M,N,NV).LE.0.) GOTO 39
        S=FECO(N)
        IF (S.LE.MAX) GOTO 31
        MAX=S
        DO 30 I=1,NV
30      COOMAX(I)=N(I)
        IMAX=IC
        JMAX=0
31      JL=1
        DO 32 I=1,NV
32      POINT(I)=N(I)
33      DO 34 I=1,NV
34      POINT(I)=(POINT(I)+M(I))/2.
        IF (FECO(POINT).LE.S) GOTO 41
        S=FECO(POINT)
        IF (S.LE.MAX) GOTO 37
        MAX=S
        DO 36 I=1,NV
36      COOMAX(I)=POINT(I)
        IMAX=IC
        JMAX=JL
37      JL=JL+1
        GOTO 33
39      JL=100
        S=FECO(M)
        IF (S.LE.MAX) GOTO 41
        MAX=S
        DO 40 I=1,NV
40      COOMAX(I)=M(I)
        IMAX=IC
        JMAX=99
41      CONTINUE
        IC=IC+1
        IF (IC.GE.ITEST) GOTO 50
        DO 42 I=1,NV
        M(I)=(M(I)+X(I))/2.
42      N(I)=(N(I)+X(I))/2.
        GO TO 200
C
C
50      CONTINUE
        RETURN
        END
```

```
      FUNCTION KPPART(X,NV,NC,MODE)
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION X(NV)
      DIMENSION A(5)
      DIMENSION BDL(15),BDU(15)
      MM=NC-2*NV
      KPPART=1
      IF (MODE.EQ.1) GOTO 5
C
C     BORNES
      DO 4 I=1,NV
      BDL(I)=0.
      BDU(I)=80.
      IF ((X(I)-BDL(I)).LT.0.) GOTO 2
4     IF ((X(I)-BDU(I)).GT.0.) GOTO 2
C     CONTRAINTES LINEAIRES
5     CONTINUE
C     CONTRAINTES NON LINEAIRES
      DO 1 K=1,MM
1     IF (A(K).LT.0.) GO TO 2
      RETURN
2     KPPART=-1
3     RETURN
      END

      FUNCTION FECO(X)
      RETURN
      END

      FUNCTION DDFECO(X,Y,NV)
      DIMENSION X(NV),Y(NV)
      RETURN
      END
```

BUS LILLE

```
C       PROGRAMME PRINCIPAL
C       NV EST LE NOMBRE DE VARIABLES
C       NC EST LE NOMBRE DE CONTRAINTES
C       LA FONTION ECONOMIQUE EST NOTEE :FECO
        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        DIMENSION X0(15),X(15),Y(15),Z(15),P(15),GRAD(15)
        DIMENSION A(15,5),B(5),B0(5)
        DIMENSION BDL(15),BDU(15),BDL0(15),BDU0(15),P0(15)
        EXTERNAL FECO,DDFECO,KPPART
        READ 502,ALFA
        READ 501,NV,NC
501     FORMAT(I2,2X,I2)
        MM=NC-2*NV
        DO 20  I=1,NV
20      X0(I)=0.0001
        X0(7)=60.
502     FORMAT(8 (F7.4,2X))
        DO 30 I=1,NV
        BDL(I)=0.
30      BDU(I)=80.
        DO 1 I=1,NV
1       X(I)=X0(I)
C       INTRODUCTION DES CONTRAINTES LINEARISEES
        JC=0
        CALL UCTIME (0,INTER)
2        CONTINUE
C       INTRODUCTION DU GRADIENT DE FECO
        GN=0.
        DO 12 I=1,NV
12      GN=GN+GRAD(I)*GRAD(I)
        GN=SQRT(GN)
        DO 3 I=1,NV
3       IF(ABS(GRAD(I)).GT.1.E-7) GO TO 11
        GO TO 10
11      DO 4 I=1,NV
4       Y(I)=X(I)+ALFA*GRAD(I) /GN
        IF(KPPART(Y,NV,NC,0).GT.0 ) GO TO 6
        DO 22 I=1,MM
        TOT=0.
        DO 18 J=1,NV
18      TOT=TOT+A(J,I)*Y(J)
22      B0(I)=B(I)-TOT
        DO 23 I=1,NV
        BDL0(I)=BDL(I)-Y(I)
        BDU0(I)=BDU(I)-Y(I)
23      P0(I)=X(I)-Y(I)
        CALL PROJ(NV,NC,A,B0,BDL0,BDU0,P0)
        DO 24 I=1,NV
24      Z(I)=P0(I)+Y(I)
        DO 5 I=1,NV
5       IF(ABS(Z(I)-X(I)).GT.1.E-4) GO TO 9
        GO TO 10
9       CONTINUE
        CALL GRILLE(NV,NC,FECO,DDFECO,KPPART,X0,X,Z,VAL,P)
        GOTO 7
6       CALL GRILLE(NV,NC,FECO,DDFECO,KPPART,X0,X,Y,VAL,P)
7       DO 8 I=1,NV
8       X(I)=P(I)
        JC=JC+1
        IF(ABS(VAL+32.349).LT.0.0005) GOTO 10
        OUTPUT VAL
        GOTO 2
10      CONTINUE
        CALL UCTIME (1,INTER)
        OUTPUT INTER
        PRINT 509 ,JC
509     FORMAT(40X,16H ITERE NUMERO :  ,2I)
        OUTPUT VAL
         STOP
        END
```

Le formalisme des familles ρ-décroissantes a été introduit par J. DENEL afin d'élargir certaines propriétés des fonctions multivoques et leur application à l'optimisation.
On trouve dans cette annexe une application simple de ce formalisme à la construction d'un algorithme de direction réalisable et améliorante très voisin de celui présenté en I.2.2.4.
Toutes les démonstrations ici présentées sont dues à J. DENEL, elles consistent essentiellement à la vérification d'hypothèses H'_1, H'_2, H'_3 nécessaires pour l'obtention d'un résultat décrit en [10] ; toutes les définitions et propriétés utilisées pourront être trouvées dans cette publication.

C un convexe compact de $\mathbb{R}^n$,

$f : C \to \mathbb{R}$ une fonction continument différentiable telle que ∇f est non nul sur C

$B = \{s \in \mathbb{R}^n \ / \ ||s|| = 1\}$

pour tout $\rho \in \mathbb{R}^{+*}$ on définit l'application multivoque :

$$F_\rho : C \to \mathcal{P}(B)$$
$$F_\rho(x) = \{s \in B / \nabla f(x).s \geq \rho.||\nabla f(x)|| \text{ et } [x, x+\rho.s] \subset C\}$$

on remarque l'analogie de définition entre F_ρ et Γ_ρ définie en I.2.2.2.C ; la seule différence étant que les vecteurs s ici considérés sont normalisés.
Soit :

$$F_o : C \to \mathcal{P}(B)$$
$$F_o(x) = \{s \in B / \nabla f(x).s > 0 \text{ et } \exists \lambda > 0, [x, x+\lambda s] \subset C\}$$

la liaison avec les ensembles et les notations employées en I.2 :

$$D^*_+(x) \cap R_C(x) = \{s \in \mathbb{R}^n \ / \ \nabla f(x).s > 0 \text{ et } \exists \lambda > 0 \text{ tel que : } [x, x+\lambda s] \subset C\}$$

est évidente en remarquant que :

$$F_o = D^*_+ \cap R_C \cap B$$

pour tout $\theta \in \mathbb{R}^{+*}$ on définit l'application multivoque :

$$G_\theta : C \times B \to \mathcal{P}(C)$$
$$G_\theta(x,s) = \{y \in D(x,s) \ / \ f(y) \geq f(x) + \theta\}$$

où $\quad D(x,s) = \{x + \lambda.s / \lambda \in \mathbb{R}^+\} \cap C$

cette fonction multivoque correspond à la recherche de points de D(x,s) qui assurent une amélioration de la valeur de la fonction économique d'au moins $\theta > 0$.
La liaison entre cette notion et l'optimisation exacte ou approchée est obtenue en remarquant que si $\bar{x}$ est un maximum de f sur D(x,s) alors :

pour $\theta = f(\bar{x}) - f(x)$

$G_\theta(x,s) = \{y \in D(x,s) \ / \ f(y) \geq f(\bar{x})\}$ est l'ensemble des maximum de f sur D(x,s)

pour $\theta = f(\bar{x}) - f(x) - \varepsilon$ où $\varepsilon > 0$

$G_\theta(x,s) = \{y \in D(x,s) \ / \ f(y) \geq f(\bar{x}) - \varepsilon\}$ est l'ensemble des maximum à ε près de f sur $D(x,s)$

soit :

$$G_o : C \times B \to \mathcal{P}(C)$$

$$G_o(x,s) = \{y \in D(x,s) \ / \ f(y) > f(x)\}$$

La terminologie de famille ρ-décroissante provient du fait que les familles $\{F_\rho\}_{\rho \geq o}$ et $\{G_\theta\}_{\theta \geq o}$ diminuent par inclusion quand les paramètres ρ ou θ augmentent.
On vérifie aisément que :

$$\forall x \in C, [0 < \rho_1 \leq \rho_2] \Longrightarrow [F_{\rho_2}(x) \subset F_{\rho_2}(x) \subset F_o(x)]$$

$$\forall (x,s) \in C \times B \ ; [0 < \theta_1 \leq \theta_2] \Longrightarrow [G_{\theta_2}(x,s) \subset G_{\theta_1}(x,s) \subset G_o(x,s)]$$

On utilisera, les ensembles :

$$M_1 = \{x \in C \ / \ F_o(x) = \phi\}$$

$$M_2 = \{x \in C \ / \ \exists s \in F_o(x) \text{ tel que } G_o(x,s) = \phi\}$$

On vérifie que les famille décrites possèdent les propriétés suivantes ;

<u>Vérification de H_1'</u>

a) $M = M_1 \cup M_2 \neq \phi$

$$\left.\begin{array}{r} C \text{ compact} \Longrightarrow \exists \overset{*}{x} \text{ max de } f \\ \text{sur } C \\ \text{I-2-1-4} \end{array}\right\} \Longrightarrow \left.\begin{array}{c} D_+^*(x) \cap \overline{R_C}(x) = \phi \\ \text{I.2.1 remarque 5} \end{array}\right\} \Longrightarrow \left.\begin{array}{c} D_+^*(x) \cap R_C(x) = \phi \\ F_o = D_+^* \cap R_C \cap B \end{array}\right\} \Longrightarrow \left.\begin{array}{c} F_o(\overset{*}{x}) = \phi \\ \text{définition de } M_1 \end{array}\right\} \Longrightarrow \overset{*}{x} \in M_1 \neq \phi$$

b) $[M_1 \neq \phi] \Longrightarrow [M_2 = \phi]$

On a vu en (a) que $M_1 \neq \phi$, on vérifie ici que $M_2 = \phi$.
$[s \in F_o(x)] \Longrightarrow s \in D_+^*(x) \cap R_C(x)] \Longrightarrow [\exists y \in D(x,s) \cap C$ tel que : $f(y) > f(x)]$
donc :

$$[s \in F_o(x) \Longrightarrow G_o(x,s) \neq \phi] \Longrightarrow [M_2 = \phi]$$

<u>Vérification de H_2'</u>

a) Uniforme régularité de la famille $\{F_\rho\}_{\rho \geq o}$

$\forall \bar{x} \in C$; $\forall \{x_n\}_{n \in \mathbb{N}} \to \bar{x}$ on a :

$$[F_o(\bar{x}) \neq \phi] \Longrightarrow [\exists \rho > o, \exists N \in \mathbb{N} \text{ tel que } \forall n \geq N, F_\rho(x_n) \neq \phi]$$

<u>Dem</u> : $[F_o(\bar{x}) \neq \phi] \Longrightarrow [\exists \bar{s} \in B$ tel que : $\nabla f(\bar{x})\ \bar{s} > 0$ et $\exists \bar{\lambda} > 0$ tel que : $[\bar{x}, \bar{x}+\bar{\lambda}.\bar{s}] \subset C]$

Soit :

$$\rho = \text{Min}\left\{\frac{\bar{\lambda}}{2}, \frac{1}{2}\ \frac{\nabla f(\bar{x}).\bar{s}}{||\nabla f(\bar{x})||}\right\} > 0$$

Soit

$$\tilde{s}_n = \frac{\bar{x}+\bar{\lambda}.\bar{s}-x_n}{\bar{\lambda}}$$

il existe un rang n_o à partir duquel $||\tilde{s}_n|| \neq 0$

$\forall n \geq n_o$ on pose :

$$s_n = \frac{\tilde{s}_n}{||\tilde{s}_n||} \text{ et } \lambda_n = ||\tilde{s}_n||.\bar{\lambda} > 0$$

par construction :

$$\{\lambda_n\}_{n\in\mathbb{N}} \to \bar{\lambda} \geq 2\rho$$

$\exists$ un rang n_1, tel que : $\forall n \geq n_1$, $\lambda_n \geq \rho$ ⓐ

par construction :

$$\{s_n\}_{n\in\mathbb{N}} \to \bar{s}$$

$$x_n + \lambda_n.s_n = x_n + \bar{\lambda}||\tilde{s}_n||.\frac{\bar{x}+\bar{\lambda}.\bar{s}-x_n}{||\tilde{s}_n||.\bar{\lambda}} = \bar{x} + \bar{\lambda}.\bar{s}$$

or $\bar{x} + \bar{\lambda}.\bar{s} \in C$ par hypothèse et comme C est convexe :

$$\forall n \geq n_o,\ [x_n,\ x_n+\lambda_n.s_n] \subset C \qquad ⓑ$$

par continuité de ∇f :

$$\left\{\frac{\nabla f(x_n).s_n}{||\nabla f(x_n)||}\right\}_{n\in\mathbb{N}} \to \frac{\nabla f(\bar{x}).\bar{s}}{||\nabla f(\bar{x})||}$$

$\exists$ un rang n_2 tel que : $\forall n \geq n_2$

$$\frac{\nabla f(x_n).s_n}{||\nabla f(x_n)||} \geq \rho \qquad ⓨ$$

ⓐ ⓑ et ⓐ $\Longrightarrow \exists N = \max\{n_o, n_1, n_2\}$ tel que :

$$F_\rho(x_n) \neq \phi,\ \forall n \geq N$$

C.Q.F.D

b) Uniforme régularité de la famille $\{G_\theta\}_{\theta\geq 0}$

Soient :

$$\{x_n, s_n\}_{n\in\mathbb{N}} \to (\bar{x},\bar{s}) \text{ et } G_0(\bar{x},\bar{s}) \neq \phi$$

$$\exists y \in D(\bar{x},\bar{s}) \text{ tel que : } f(y) > f(\bar{x})$$

Soit :

$$\theta = \frac{f(y)-f(\bar{x})}{2} > 0 \text{ et } \forall n \in \mathbb{N} : y_n = x_n + ||y-\bar{x}||.s_n$$

$$\{y_n\}_{n\in\mathbb{N}} \to y \text{ par construction}$$

et

$f(y) > f(\bar{x}) + \theta$ par continuité de f à partir d'un certain rang :

$$f(y_n) \geq f(x_n) + \theta$$

soit

$$y_n \in G_\theta(x_n, s_n) \neq \phi$$

C.Q.F.D

c) Pseudo-sup-continuité de la famille $\{F_\rho\}_{\rho\geq 0}$

$$\left.\begin{array}{l} \forall \bar{x} \in C \ ; \ \forall \{x_n\}_{n\in\mathbb{N}} \to \bar{x} \ ; \ x_n \in C \\ \forall \{s_n\}_{n\in\mathbb{N}} \to \bar{s} \\ \exists \rho > 0, \ \exists N \in \mathbb{N} \text{ tel que : } \forall n \geq N : s_n \in F_\rho(x_n) \end{array}\right\} \Longrightarrow \bar{s} \in F_0(\bar{x})$$

<u>Dem</u> : immédiate en remarquant

$$[\forall n \in \mathbb{N}, \ n \geq N, \ \nabla f(x_n).s_n \geq \rho.||\nabla f(x_n)||] \Longrightarrow \nabla f(\bar{x}).\bar{s} > 0$$

$$[\forall n \in \mathbb{N}, \ n \geq N, \ [x_n, \ x_n + \rho.s_n] \subset C] \Longrightarrow [\bar{x}, \bar{x}+\rho.\bar{s}] \subset C$$

<u>Vérification de H_3'</u>

Cette hypothèse nécessite l'existence de deux fonctions :

$$h : C \to \mathbb{R} \text{ et } \alpha : \mathbb{R}^+ \times C \to \mathbb{R}$$

qui vérifient certaines propriétés. Ces propriétés sont ici vérifiées quand $h \equiv f$ et $\alpha(\theta,x) = f(x) + \theta$

a) f est borné sur C compact par hypothèse

b) $[s \in F_0(x) \text{ et } \forall\theta, \ y \in G_\theta(x,s)] \Longrightarrow [f(y) \geq f(x) + \theta]$

c) $\forall \{x_n\}_{n\in\mathbb{N}}$ suite de C telle que :

$$\{f(x_n)\}_{n\in\mathbb{N}} \to f^* \text{ alors } \forall\theta > 0 : \lim_n[\alpha(\theta,x_n)] = f^* + \theta > f^*$$

ALGORITHME

Soit : $\beta \in]0,1[$

Itération n connus : x_n, $\rho_n > 0$, $\theta_n > 0$

1° étage : si $F_{\rho_n}(x_n) = \phi$ alors : $x_{n+1} = x_n$

$0 < \rho_{n+1} \leq \beta.\rho_n$

$0 < \theta_{n+1} \leq \theta_n$

sinon, soit $s_n \in F_{\rho_n}(x_n)$

2° étage : si $G_{\theta_n}(x_n,s_n) = \phi$ alors : $x_{n+1} = x_n$

$\rho_{n+1} = \rho_n$

$0 < \theta_{n+1} \leq \beta.\theta_n$

sinon, $x_{n+1} \in G_{\theta_n}(x_n,s_n)$

RESULTATS.

Sous les hypothèses H_1', H_2' et H_3' vérifiées, si $(x_n)_{n \in \mathbb{N}}$ est la suite engendrée par l'algorithme alors il existe une sous-suite parfaitement déterminée dont les points d'accumulation sont dans M.
Or :

$$[\overset{*}{x} \in M] \Longrightarrow [F_o(\overset{*}{x}) = \phi] \Longrightarrow [D_+^*(\overset{*}{x}) \cap R_C(\overset{*}{x}) = \phi]$$

un tel point d'accumulation vérifie donc la condition de stationnarité faible décrite en I.2.1.
Si de plus f est concave tout point d'accumulation de $(x_n)_{n \in \mathbb{N}}$ est dans M, c'est d'après I.2.1.4 un maximum de f sur C.
Ces résultats correspondent exactement à ceux obtenus en I.2.2.4.

[1] L. ARMIJO
"minimization of functions having continuous partial derivatives"
Pacific J. of Mathematics VOL 16, p 1.3, (1966)

[2] C.G. BROYDEN, J.E. DENIS and J.J. MORE
"On the local and superlinear convergence of quasi-Newton methods"
J. Inst. Maths. Applic. VOL 12, p 223.245 (1973)

[3] A. CAUCHY
"Méthode générale pour la résolution des sytèmes d'équations simultanées"
Compte-rendu de l'Académie des Sciences VOL25, p 536.538 (1847)

[4] J. CEA
"Fondements des méthodes numériques d'optimisation"
Cours d'analyse numérique J.L LIONS : chapitre 10

[5] A.R. COLVILLE
"A comparative study on nonlinear programming codes"
IBM New-York Scientific Center Report n° 320-2949 (1968)

[6] H.B. CURRY
"The method of steepest descent for nonlinear minimization problems"
Quart. Appl. Math. VOL 2, p 258-263 (1944)

[7] J.W. DANIEL
"Convergent step-sizes for gradient like feasible directions algorithms for constrained optimization"
in "Nonlinear Programming" J.B. ROSEN Ed. p 245-274 Academic Press (1970)

[8] W.C. DAVIDON
"Variable metric method for minimization"
Argonne National Laboratory Report ANL5990 (1959)

[9] D. DAVIS
"Review of constraint optimization"
in "Optimization" R. FLETCHER Ed. Academic Press (1969)

[10] J. DENEL
"Extensions of the continuity of point-to-set maps, applications to fixed points algorithms"
Math. Prog. Studies VOL 10, à paraître (1979)

[11] J.E. DENIS and J.J. MORE
"A characterization of superlinear convergence and its application to quasi-Newton methods"
Mathematics of Computation VOL 28, n° 126, p 549.560 (1974)

[12] L.C.W DIXON
"All the quasi-Newton family generate identical points"
J.O.T.A VOL 10, p 34.40 (1972)

[13] R. FLETCHER and M. POWELL
"A rapidly convergent descent method for minimization"
Computer J. VOL 6, n° 2, p 163.168 (1963)

[14] M. FRANK and P. WOLFE
"An algorithm for quadratic programming"
Naval Research Log. Quaterly VOL 3, p 95.120 (1956)

[15] U.M. GARCIA-PALOMARES and MANGASARIAN
"Superlinearly convergent quasi-Newton algorithms for nonlinearly constrained optimization problems"
Math. Prog. VOL 11, p 1.13 (1976)

[16] A.A. GOLDSTEIN
"On steepest descent"
S.I.A.M J. of Control VOL 3, p 147.151 (1965)

[17] S.P. HAN
"Superlinearly convergent variable metric algorithms for general nonlinear programming problems"
Math. Prog. VOL 11, p 263.282 (1976)

[18] H.Y. HUANG
"Unified approach to quadratically convergent algorithms for function minimization"
J.O.T.A VOL 5, n° 6, p 405-423 (1970)

P. HUARD
[19] "Méthode des centres et méthode des centres par majoration"
Bulletin de la Direction des Etudes et Recherches E.D.F
Série C. Mathématiques Informatique n°2, p 33.52 (1970)

[20] "Cours de D.E.A de traitement de l'information"
Optimisation dans $\mathbb{R}^n$: éléments théoriques, algorithmes généraux
Université des Sciences et Techniques - Lille (1972)

[21] "Implementation of gradient methods by tangential discretization"
Publication n° 73 - Université des Sciences et Techniques - Lille (1976)

[22] "Implementation of gradient methods by tangential discretization"
J.O.T.A à paraître (1978)

[23] **R. KLESSIG**
"A general theory of convergence for algorithms that use antizigzaging provisions"
S.I.A.M J. of Control VOL 12, n° 4, p 598.608 (1974)

[24] **P.J. LAURENT**
"Un algorithme dual pour le calcul de la distance entre deux convexes"
Séminaire d'Analyse Numérique n° 220
Université Scientifique et Médicale de Grenoble (1975)

[25] **D.G. LUENBERGER**
"Introduction to linear and nonlinear programming"
Addison Wesley (1973)

[26] **R. MEYER**
"The validity of a family of optimization methods"
S.I.A.M J. of Control VOL 8, n°1, p 41.54 (1970)

[27] **M. MOKHTAR-KHARROUBI**
"Etude de quelques méthodes de direction réalisable"
Thèse de 3° cycle - Université P. et M. Curie - Paris (1976)

[28] **B.A. MURTAGH and R.W.H. SARGENT**
"A constaint minimization mehtod with quadratic convergence"
in "Optimization" R. FLETCHER Ed. p 215.246, Academic Press (1969)

E. POLAK
[29] "On the convergence of optimization algorithms"
R.I.R.O. VOL 16, p 17.34 (1969)

[30] "Computational methods in optimization : a unified approach"
Academic Press (1971)

M.J.D. POWELL
[31] "Variable metric methods for constrained optimization"
Troisième Colloque International sur les méthodes de calcul scientifique et technique I.R.I.A Rocquencourt (1977)

[32] "The convergence of variable metric methods for nonlinearly constrained optimization calculations"
Argonne National Laboratory Report n° 315 (1977)

[33] **S.M. ROBINSON**
"A quadratically convergent algorithm for general nonlinear programming problems"
Math. Prog. VOL 3, p 145-156 (1972)

J.B. ROSEN
[34] "The gradient projection method for nonlinear programming"
Part I : "Linear constraints"
S.I.A.M. J. of Appl. Math. VOL 8, n° 1, p 181-217 (1960)

[35] Part II : "Nonlinear constraints"
S.I.A.M. J. of Appl. Math. VOL 9, n° 4, p 514-532 (1961)

[36] H. SCHWETLICK
"Minimizing nonlinear functions by descent methods"
Information Processing North-Holland Publishing Compagny (1974)

[37] D.M. TOPKIS and A.F. VEINOTT
"On the convergence of some feasible direction algorithms for nonlinear programming"
S.I.A.M J. of Control VOL 5, n° 2, (1967)

[38] C. VAN DE PANNE
"Methods for linear and quadratic programming"
North-Holland, Amsterdam (1975)

[39] R.B. WILSON
"A simplicial method for concave programming"
Ph. D. dissertation - Harvard University (1963)

[40] P. WOLFE
"On the convergence of gradient methods under constraints"
IBM J. of Research and Development VOL 16, n° 4, p 407-411 (1972)

[41] W.I. ZANGWILL
"Nonlinear programming : a unified approach"
Prentice Hall, Englewood Cliffs, New-Jersey (1969)

G. ZOUTENDIJK
[42] "Methods of feasible direction"
Elsevier Publishing C°, Amsterdam (1960)

[43] "Some algorithms based on the principle of feasible directions"
in "Nonlinear Programming" J.B. ROSEN Ed. p 93-121, Academic Press (1970)

[44] M. LENARD
"Practical convergence conditions for the DAVIDON-FLETCHER-POWELL method"
Math. Prog. VOL 9, p 69.86 (1975)

BIBLIOTHEQUE UNIVERSITAIRE
Section de SCIENCES
LILLE