MDI 210

Analyse numérique et optimisation continue

Partie 2 : Optimisation continue

Irène Charon Olivier Hudry

9 février 2016

Table des matières

Chapitre I

Programmation linéaire : l'algorithme du simplexe

I.1. Introduction

Afin d'illustrer ce qu'est la *programmation linéaire*, aussi appelée *optimisation linéaire*, commençons par un petit exemple simple. Celui-ci nous permettra en outre d'introduire certaines propriétés des problèmes relevant de ce domaine, propriétés qui seront ensuite exploitées pour fonder l'algorithme du simplexe.

Une usine fabrique deux sortes de produits, p_1 et p_2, à l'aide de deux machines m_1 et m_2. On suppose que la quantité fabriquée de ces produits n'est pas nécessairement un nombre entier, mais seulement un réel positif ou nul. Chaque unité de produit en cours de fabrication doit passer successivement sur les deux machines dans un ordre indifférent et pendant les temps suivants, exprimés en minutes :

	p_1	p_2
m_1	30	20
m_2	40	10

La machine m_1 est disponible 6000 minutes par mois et la machine m_2 est disponible 4000 minutes par mois. Le profit réalisé sur une unité du produit p_1 est de 400 €. Le profit réalisé sur une unité du produit p_2 est de 200 €.

On souhaite trouver le plan de fabrication mensuel qui maximise le profit.

Pour cela, appelons x (respectivement y) le nombre d'unités du produit p_1 (respectivement p_2) à fabriquer mensuellement ; on voit que ce problème peut s'exprimer sous la forme suivante :

$$\text{Maximiser } z = 400x + 200y$$
$$\text{avec les contraintes : } \begin{cases} 30x + 20y \leq 6000 \\ 40x + 10y \leq 4000 \\ x \geq 0, y \geq 0. \end{cases}$$

Le problème étant en deux variables, il admet une solution graphique facile à mettre en œuvre, représentée par la figure I.1.

Les points (x, y) qui satisfont les contraintes appartiennent au quadrilatère $OABC$. La famille des droites $D_\lambda = \{(x, y) \text{ avec } 400x + 200y = \lambda\}$ est une famille de droites parallèles. Parmi celles de ces droites qui ont une intersection non vide avec le quadrilatère, c'est celle qui passe par B qui correspond à la plus grande valeur de λ : elle rencontre le quadrilatère des contraintes au point de coordonnées (40, 240). La solution optimale de notre problème est donc $x = 40, y = 240$ (et $z = 64\ 000$).

Plus généralement, un problème de programmation linéaire est un problème qui peut se formuler comme suit :

- maximiser une forme linéaire de n variables $x_1, ..., x_n$,

les variables étant soumises :

- à m contraintes linéaires : pour $i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij}x_j \leq b_i$
- aux n contraintes de positivité : pour $j \in \{1, 2, ..., n\}, x_j \geq 0$.

Cette formulation s'appelle la *forme standard* d'un problème de programmation linéaire. On peut envisager d'autres formulations du problème à résoudre. Dans toute la suite, on ne considérera pas le cas d'inégalités strictes (le domaine défini par les contraintes n'étant plus alors un fermé, le problème pourrait ne pas admettre de solution optimale, même si la fonction z est majorée sur ce domaine). En revanche, des problèmes où il s'agit de minimisation, ou pour lesquels apparaissent des contraintes d'égalité ou d'inégalité large dans l'autre sens, ou encore pour lesquels les variables ont d'autres contraintes que celles d'être positives ou nulles (ou sont non contraintes) peuvent facilement se mettre sous forme standard, comme le précisent les indications suivantes :

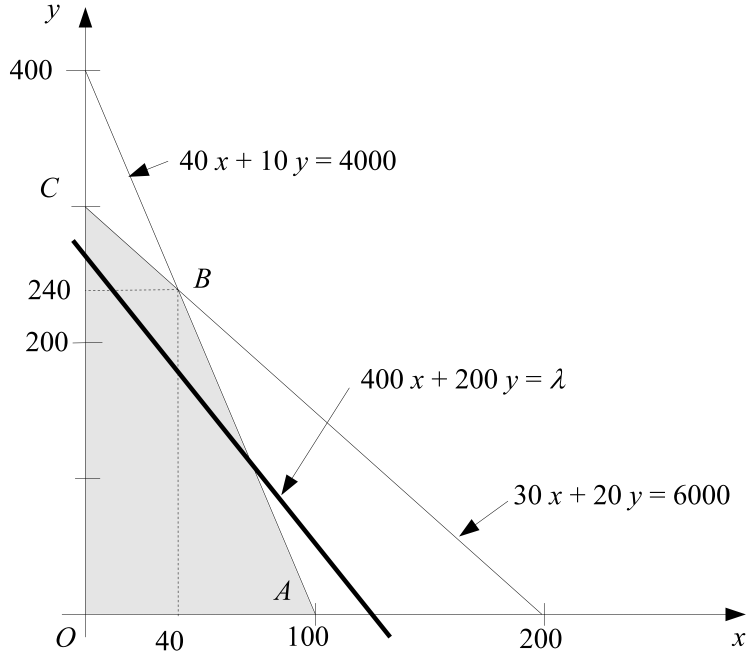

FIGURE I.1 – Illustration de l'exemple

- minimiser une fonction f (linéaire ou non) revient à maximiser $-f$, puisqu'on a la relation : minimum de $f = -$ maximum de $(-f)$;

- on transforme une inégalité du genre « $\geq$ » en une inégalité du genre « $\leq$ » en la multipliant par -1 ;

- une égalité $\sum_{j=1}^{n} a_{ij}x_j = b_i$ revient aux deux inégalités $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$ et $\sum_{j=1}^{n}(-a_{ij})x_j \leq -b_i$;

- on remplace une variable contrainte par la double inégalité $\alpha \leq x \leq \beta$ par la variable $y = x - \alpha$ et on ajoute les contraintes $y \leq \beta - \alpha$ et $y \geq 0$ (s'il existe plusieurs contraintes de ce type impliquant une même variable x, on ne considère que la plus contraignante et on élimine les autres) ;

- on remplace une variable x contrainte par l'inégalité $x \geq \alpha$ par la variable $x - \alpha$ qui devra être positive ou nulle (s'il y a plusieurs contraintes de ce type, on ne considère que celle ayant la plus grande valeur α et on élimine les autres) ;

- on remplace une variable x contrainte par l'inégalité $x \leq \beta$ par la variable $\beta - x$ qui devra être positive ou nulle (s'il y a plusieurs contraintes de ce type, on ne considère que celle ayant la plus petite valeur β et on élimine les autres) ;

- on exprime une variable x qui n'est pas contrainte par la différence de deux variables positives ou nulles : $x = x^+ - x^-$ avec $x^+ \geq 0$ et $x^- \geq 0$.

On peut alors se demander si la démarche proposée pour l'exemple précédent est susceptible d'être généralisée à la résolution de tout problème de programmation linéaire. Puisqu'il est toujours possible d'exprimer un problème de programmation linéaire (sans contrainte d'inégalité stricte) sous forme standard, considérons un problème mis sous cette forme :

$$\begin{array}{l} \text{Maximiser } z = \sum_{j=1}^{n} c_j x_j \\ \text{avec les contraintes : } \left\{ \begin{array}{l} \text{pour } i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij} x_j \leq b_i \\ \text{pour } j \in \{1, 2, ..., n\}, x_j \geq 0. \end{array} \right. \end{array}$$

L'ensemble des points de $\mathbb{R}^n$ de coordonnées $x_1, ..., x_n$ et vérifiant les $m + n$ contraintes précédentes détermine un polyèdre appelé *polyèdre des contraintes*. Ce polyèdre est *convexe*, c'est-à-dire que, pour tout point M et tout point P du polyèdre, le segment $[M, P]$ est entièrement contenu dans le polyèdre. En effet, soient $M = (x_1, ..., x_n)$ et $P = (y_1, ..., y_n)$ deux points quelconques du polyèdre déterminé par les contraintes ; alors, pour tout réel λ vérifiant $0 \leq \lambda \leq 1$, il est facile de vérifier que le point $\lambda M + (1 - \lambda)P$ (de coordonnées $\lambda x_i + (1 - \lambda)y_i$) appartient au polyèdre. Les n-uplets $(x_1, ..., x_n)$ qui satisfont les contraintes s'appellent *solutions réalisables* du problème. Ce sont les coordonnées des points intérieurs (au sens large) au polyèdre des contraintes qui, dans notre exemple, était le quadrilatère $OABC$.

On peut prouver le théorème suivant :

Théorème I.1. *Soit un problème de programmation linéaire dont le polyèdre des contraintes est non vide et dont la fonction à maximiser est majorée sur ce polyèdre. Alors le problème admet un maximum (qui est fini) atteint en au moins un sommet du polyèdre des contraintes.*

L'idée de l'algorithme du simplexe est de passer itérativement d'un sommet du polyèdre des contraintes à un sommet adjacent en suivant des arêtes du polyèdre et de façon à augmenter la valeur de la fonction à optimiser, jusqu'à trouver un sommet où le maximum est atteint. C'est grâce à la convexité du polyèdre et à la linéarité de la fonction dont on cherche le maximum que l'on peut se contenter de chercher le maximum en un sommet du polyèdre.

I.2. L'algorithme du simplexe sur un exemple

Appliquons maintenant l'algorithme du simplexe à un exemple plus sophistiqué, afin d'en illustrer le fonctionnement.

Une fabrique de tissus produit quatre types de tissus : du kelsch, du nanzouk, du shantung et du zénana. Ces tissus résultent de trois opérations principales : la filature, le tissage, la teinture. Ils sont produits en longueur variable, mesurée ici en kilomètre. La production d'un kilomètre de tissu nécessite un certain nombre d'heures de filature, de tissage et de teinture, ces nombres dépendant du tissu. Par ailleurs, la vente de ces tissus rapporte un certain bénéfice exprimé en euros. Ces données (fictives !) sont précisées dans le tableau suivant, pour un kilomètre de tissu :

	kelsch	nanzouk	shantung	zénana
filature	2	4	5	7
tissage	1	1	2	2
teinture	1	2	3	3
bénéfice	7	9	18	17

L'entreprise dispose, quotidiennement, de 42 heures de filature, 17 heures de tissage et 24 heures de teinture. On souhaite établir un plan de fabrication de façon à maximiser le bénéfice (on suppose que l'on est en régime stable de fabrication et non en phase initiale où il faut filer avant de tisser et tisser avant de teindre).

Appelons x_1, x_2, x_3, x_4 les longueurs respectives de kelsch, de nanzouk, de shantung et de zénana produites quotidiennement. Le problème admet alors la modélisation suivante :

Maximiser $z = 7x_1 + 9x_2 + 18x_3 + 17x_4$
avec les contraintes :
$$\begin{cases} 2x_1 + 4x_2 + 5x_3 + 7x_4 \leq 42 \\ x_1 + x_2 + 2x_3 + 2x_4 \leq 17 \\ x_1 + 2x_2 + 3x_3 + 3x_4 \leq 24 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{cases}$$

On reconnaît un problème de programmation linéaire sous forme standard. On va résoudre ce problème à l'aide de l'algorithme du simplexe que nous expliquerons ainsi sur cet exemple.

On introduit trois variables dites *variables d'écart* x_5, x_6, x_7, positives ou nulles, qui mesurent pour chaque ressource l'écart entre la quantité initialement disponible et la quantité consommée par le plan de fabrication donné par x_1, x_2, x_3 et x_4. On obtient ce qui s'appelle un *dictionnaire* (voir plus bas la définition générale), le premier pour la résolution de ce problème (il y en aura d'autres) :

$$\begin{array}{llrcrcrcrcr} x_5 & = 42 & - & 2x_1 & - & 4x_2 & - & 5x_3 & - & 7x_4 \\ x_6 & = 17 & - & x_1 & - & x_2 & - & 2x_3 & - & 2x_4 \\ x_7 & = 24 & - & x_1 & - & 2x_2 & - & 3x_3 & - & 3x_4 \\ \hline z & = & & 7x_1 & + & 9x_2 & + & 18x_3 & + & 17x_4 \end{array} \qquad \text{Dictionnaire I}$$

Le problème s'écrit maintenant :

Maximiser z avec $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0, x_6 \geq 0, x_7 \geq 0$.

Le polyèdre des contraintes est limité dans $\mathbb{R}^4$ par les hyperplans d'équation $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0, x_5 = 0, x_6 = 0, x_7 = 0$.

Dans ce dictionnaire, les variables x_5, x_6 et x_7 sont exprimées comme fonctions affines des variables x_1, x_2, x_3 et x_4 ; on traduit cette caractéristique en disant que les variables x_5, x_6 et x_7 sont actuellement les *variables de base* du dictionnaire et les variables x_1, x_2, x_3 et x_4 les *variables hors-base* du dictionnaire. On s'intéresse alors à ce qu'on appelle la *solution basique* associée au dictionnaire ; celle-ci est la solution obtenue en attribuant la valeur 0 à toutes les variables hors-base ; les valeurs des variables de base en découlent.

Afin de distinguer les fonctions et les variables des valeurs de ces fonctions et de ces variables, on utilisera le signe * lorsqu'il s'agit de valeurs : ainsi x^* représentera une valeur prise par la variable x. Avec cette notation, les égalités $x_1^* = 0, x_2^* = 0, x_3^* = 0, x_4^* = 0$ entraînent $x_5^* = 42, x_6^* = 17$ et $x_7^* = 24$. Les sept variables ayant des valeurs positives ou nulles dans cette solution basique, on dit que ce dictionnaire est *réalisable*. On peut remarquer que le point de coordonnées (0, 0, 0, 0) est ici un sommet du polyèdre des contraintes ; la solution basique associée au dictionnaire donne alors à z la valeur 0.

La remarque suivante est à la base de la méthode : si, choisissant une variable hors-base de coefficient strictement positif, on fait croître celle-ci à partir de 0, les autres variables hors-base restant nulles, la valeur correspondante de la fonction z croît. Dans notre exemple, choisissons la variable x_3 (on pourrait aussi choisir ici l'une quelconque des trois autres variables hors-base). Gardant x_1, x_2 et x_4 à 0, nous cherchons à augmenter x_3 au maximum, tout en conservant la propriété que le point M de $\mathbb{R}^4$ de coordonnées $(0, 0, x_3, 0)$ reste dans le polyèdre des contraintes (on se déplace sur une arête du polyèdre des contraintes issue du sommet (0, 0, 0, 0)).

Les contraintes sur l'augmentation de la variable x_3 sont :

$x_5 \geq 0$, ce qui impose $x_3 \leq 8,4$;

$x_6 \geq 0$, ce qui impose $x_3 \leq 8,5$;

$x_7 \geq 0$, ce qui impose $x_3 \leq 8$.

Le premier hyperplan que rencontre le point M est donc celui d'équation $x_7 = 0$: le point M est alors arrivé à un nouveau sommet du polyèdre des contraintes, à l'intersection des hyperplans d'équations $x_1 = 0$, $x_2 = 0$, $x_4 = 0$, $x_7 = 0$. Nous allons alors faire un changement de dictionnaire en échangeant les rôles de x_3 et x_7, pour itérer le procédé que nous venons d'employer. On utilise l'équation du dictionnaire I qui donne x_7 pour exprimer x_3 en fonction de x_1, x_2, x_4 et x_7 ; on remplace ensuite x_3 par cette expression dans les autres équations du dictionnaire :

On obtient ainsi un deuxième dictionnaire :

$$\begin{array}{rlrrrr}
x_3 & = 8 & - \frac{1}{3}x_1 & - \frac{2}{3}x_2 & - x_4 & - \frac{1}{3}x_7 \\
x_5 & = 2 & - \frac{1}{3}x_1 & - \frac{2}{3}x_2 & - 2x_4 & + \frac{5}{3}x_7 \\
x_6 & = 1 & - \frac{1}{3}x_1 & + \frac{1}{3}x_2 & & + \frac{2}{3}x_7 \\
\hline
z & = 144 & + x_1 & - 3x_2 & - x_4 & - 6x_7
\end{array} \qquad \text{Dictionnaire II}$$

On dit qu'on a fait « entrer x_3 en base » et qu'on a fait « sortir x_7 de la base », ou encore que x_3 est *variable entrante* et que x_7 est *variable sortante*. Les variables de base sont maintenant x_3, x_5 et x_6, et les variables hors-base x_1, x_2, x_4 et x_7. Dans la nouvelle solution basique, la fonction z vaut 144, valeur que l'on obtient en annulant les variables hors-base. On remarque qu'on a ainsi une nouvelle solution réalisable plus intéressante que celle associée au premier dictionnaire.

Dans la nouvelle expression de la fonction z, nous voyons que seule la variable x_1 a un coefficient strictement positif : nous décidons de faire entrer x_1 en base, et ainsi de parcourir une nouvelle arête du polyèdre des contraintes ; on a les limites suivantes sur l'augmentation possible de la valeur de x_1 à partir de la valeur nulle, les autres variables hors base restant à 0 :

$x_3 \geq 0$, ce qui impose $x_1 \leq 24$;
$x_5 \geq 0$, ce qui impose $x_1 \leq 6$;
$x_6 \geq 0$, ce qui impose $x_1 \leq 3$.

C'est la troisième limite qui est la plus contraignante ; x_6 sort de la base, ce qui conduit au dictionnaire suivant :

$$\begin{array}{llrrrrrrrr}
x_1 & = 3 & + & x_2 & & & - & 3x_6 & + & 2x_7 \\
x_3 & = 7 & - & x_2 & - & x_4 & + & x_6 & - & x_7 \\
x_5 & = 1 & - & x_2 & - & 2x_4 & + & x_6 & + & x_7 \\
\hline
z & = 147 & - & 2x_2 & - & x_4 & - & 3x_6 & - & 4x_7
\end{array} \qquad \text{Dictionnaire III}$$

La solution basique associée à ce nouveau dictionnaire donne à z la valeur 147.

De plus, nous voyons sur la dernière ligne de ce dictionnaire que, les variables x_2, x_4, x_6, x_7 étant positives ou nulles, l'optimum cherché de z est majoré par 147. La solution basique actuelle nous fournit donc une solution optimum du problème :

- il faut fabriquer chaque jour trois kilomètres de kelsch, zéro de nanzouk, sept de shantung et zéro de zénana ;
- toutes les heures de tissage et de teinture sont utilisées, alors qu'il reste une heure de filage disponible ;
- le bénéfice vaut 147.

Remarques
1. Il se trouve que la solution obtenue ici est entière alors que cela n'était pas imposée par la formulation du problème. Ceci n'a rien de général et les

problèmes de programmation linéaire en nombres entiers (c'est-à-dire des problèmes de programmation linéaire pour lesquels les variables doivent être entières) peuvent être qualitativement plus compliqués.
2. La méthode consiste, à chaque étape, à faire entrer en base une variable dont le coefficient dans la fonction z à optimiser est strictement positif. Cela ne permet cependant pas toujours d'obtenir une croissance stricte de z. Nous reviendrons sur ce phénomène dans le paragraphe consacré à la « dégénérescence ».
3. Enfin, nous avons eu la chance de trouver, sans aucune difficulté, un sommet du polyèdre des contraintes, ou autrement dit un dictionnaire réalisable, qui nous sert de point de départ. En effet, « l'origine était réalisable », autrement dit l'annulation des variables x_1, x_2, ..., x_n attribue des valeurs positives ou nulles aux variables d'écart, car les b_i étaient tous positifs ou nuls. Nous étudierons plus loin des cas moins favorables.

I.3. Définitions et terminologie

Revenons sur quelques définitions. Un problème de programmation linéaire est mis sous forme standard s'il est écrit sous la forme :

- maximiser une forme linéaire z de n variables $x_1, ..., x_n$: $z = \sum_{j=1}^{n} c_j x_j$

les variables étant soumises :

- à m contraintes linéaires : pour $i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij} x_j \leq b_i$
- aux n contraintes de positivité : pour $j \in \{1, 2, ..., n\}, x_j \geq 0$.

Tout n-uplet de valeurs $(x_1, ..., x_n)$ satisfaisant les contraintes constitue une *solution réalisable*.

La fonction z est appelée *fonction objectif*. Les variables $x_1, ..., x_n$ sont appelées *variables de décision* ou aussi *variables de choix* ou encore *variables principales* ou *initiales* ; les variables x_{n+1}, ..., x_{n+m} s'appellent les *variables d'écart*. On constate qu'une solution $x_1^*, x_2^*, ..., x_{n+m}^*$ est réalisable si et seulement si toutes ses valeurs sont positives ou nulles ; autrement dit : pour $k \in \{1, 2, ..., n+m\}, x_k^* \geq 0$. Une solution réalisable qui maximise la fonction objectif est dite *solution optimale*. Si un problème de programmation linéaire n'admet aucune solution réalisable, il est dit *infaisable* ou *non réalisable*. Si un problème admet des solutions réalisables mais n'a pas de valeur optimale finie, il est dit *non borné*. Un *dictionnaire* est un système d'équations linéaires liant $x_1, ..., x_n, x_{n+1}, ..., x_{n+m}$ et z, et satisfaisant les deux propriétés suivantes :

- les équations constituant un dictionnaire quelconque doivent exprimer z et m des $n+m$ variables $x_1, ..., x_{n+m}$ en fonction des n autres variables et ce, de manière unique ;

- tout dictionnaire est algébriquement équivalent au dictionnaire définissant les variables d'écart et la fonction objectif, c'est-à-dire au dictionnaire :

$$\begin{cases} x_{n+1} = b_1 - \sum_{j=1}^{n} a_{1j}x_j \\ ... \\ x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij}x_j \\ ... \\ x_{n+m} = b_m - \sum_{j=1}^{n} a_{mj}x_j \\ \hline z = \sum_{j=1}^{n} c_j x_j \end{cases}$$

On obtient une *base* du problème en considérant m des $n + m$ variables $x_1, x_2, ..., x_{n+m}$ (les *m variables de base* ou *en base*) de sorte que ces m variables s'expriment de manière unique à l'aide des n autres variables, appelées *variables hors-base*. Une base définit donc un dictionnaire et réciproquement. Une base étant fixée, on obtient la *solution basique* associée à cette base ou, ce qui revient au même, au dictionnaire associé à cette base, en attribuant la valeur 0 à toutes les variables hors-base. Géométriquement, une solution à la fois basique et réalisable correspond en fait à un sommet du polyèdre des contraintes.

L'algorithme du simplexe a pour objectif de déterminer une solution optimale parmi les solutions basiques et réalisables (c'est-à-dire parmi les sommets du polyèdre des contraintes).

I.4. Résumé d'une itération

Pour déterminer une telle solution basique réalisable optimale, décrivons une itération de l'algorithme du simplexe de façon générale. Pour cela, définissons deux sous-ensembles d'indices, J et I : J correspond aux indices des n variables hors-base dans le dictionnaire courant et I aux indices des m

variables de base. Plus précisément :
- $J \subset \{1, 2, ..., n+m\}$ avec $|J| = n$ (initialement, on pose $J = \{1, 2, ..., n\}$) ;
- $I = \{1, 2, ..., n+m\} \setminus J$;
- le dictionnaire courant est décrit par les égalités suivantes : pour $i \in I$, $x_i = b_i^{'} + \sum_{j \in J} a_{ij}^{'} x_j$ et $z = z^* + \sum_{j \in J} c_j^{'} x_j$; on suppose que le dictionnaire est réalisable : pour $i \in I$, $b_i^{'} \geq 0$.

L'itération courante se déroule comme suit :

- si tous les coefficients $c_j^{'}$ sont négatifs ou nuls, l'algorithme est terminé : l'annulation des variables hors-base fournit une solution optimale ;
- sinon :
  - on choisit une variable hors-base x_{j_0} pourvue d'un coefficient strictement positif dans z ; il s'agit de la variable entrante ; s'il y a plusieurs variables candidates pour entrer en base, on peut par exemple privilégier la variable ayant le plus grand coefficient dans z ou, ce qui est généralement plus efficace, privilégier la variable entraînant la plus grande augmentation de z (on verra un autre choix au paragraphe suivant, en cas de « dégénérescence ») ;
  - on détermine la variable sortante x_{i_0} comme étant la variable de base qui restreint le plus la croisssance de x_{j_0} ; pour cela, on considère, pour $i \in I$ avec $a_{ij_0}^{'} < 0$, les rapports $\frac{-b_i^{'}}{a_{ij_0}^{'}}$: i_0 est l'indice pour lequel ce rapport est le plus petit (s'il y a plusieurs variables candidates pour sortir de la base, on peut en choisir une arbitrairement ; on verra un autre choix au paragraphe suivant, là encore en cas de « dégénérescence ») ;
  - on extrait x_{j_0} de l'expression courante de x_{i_0} ;
  - on remplace x_{j_0} par sa nouvelle expression dans z et dans l'expression des autres variables de base ; on obtient ainsi le nouveau dictionnaire courant à partir duquel on applique l'itération suivante.

Remarque. Quand on passe du dictionnaire courant au dictionnaire suivant, on est sûr que celui-ci est réalisable, par le choix de la variable sortante. Autrement dit, on passe d'une solution basique réalisable à une autre solution basique réalisable. Il est donc inutile de vérifier cette propriété quand on obtient le nouveau dictionnaire.

I.5. La dégénérescence et le cyclage

Les solutions basiques réalisables avec une ou plusieurs variables de base nulles sont dites *dégénérées*, ce qu'illustre l'exemple suivant.

Exemple

Considérons le dictionnaire (non dégénéré) :

$$\begin{array}{llcrcrcr}
x_4 & = 1 & & & & & - & 2x_3 \\
x_5 & = 3 & - & 2x_1 & + & 4x_2 & - & 6x_3 \\
x_6 & = 2 & + & x_1 & - & 3x_2 & - & 4x_3 \\
\hline
z & = & & 2x_1 & - & x_2 & + & 8x_3
\end{array}$$

Choisissant de faire entrer x_3 en base, nous voyons que les relations $x_4 \geq 0$, $x_5 \geq 0, x_6 \geq 0$ imposent toutes les trois 0,5 comme limite à la croissance de x_3. Chacune des trois variables x_4, x_5, x_6 est donc candidate à quitter la base. Si nous choisissons x_4, nous obtenons comme nouveau dictionnaire :

$$\begin{array}{llcrcrcr}
x_3 & = 0,5 & & & & & - & 0,5x_4 \\
x_5 & = & - & 2x_1 & + & 4x_2 & + & 3x_4 \\
x_6 & = & & x_1 & - & 3x_2 & + & 2x_4 \\
\hline
z & = 4 & + & 2x_1 & - & x_2 & - & 4x_4
\end{array}$$

Dans la solution basique associée à ce dictionnaire, x_5 et x_6 prennent une valeur nulle. Du fait de la nullité d'au moins une des variables en base, cette solution basique est dégénérée.

Si nous faisons une itération à partir de ce dictionnaire, nous voyons que, faisant entrer x_1 en base (seule variable à avoir un coefficient positif dans z), la relation $x_5 \geq 0$ impose $x_1 \leq 0$. La plus grande valeur attribuable à x_1 vaut 0 et la valeur z^* n'augmentera donc pas au cours de cette itération.

L'inconvénient de ces inévitables itérations dégénérées est qu'elles peuvent induire un phénomène désastreux pour la convergence de l'algorithme : le cyclage. On dit qu'il y a *cyclage* lorsque, au bout d'un nombre fini d'itérations, on retrouve un dictionnaire déjà rencontré. En fait, à cause de la définition des dictionnaires, et plus précisément de l'équivalence algébrique des dictionnaires, on retrouve un dictionnaire déjà rencontré dès qu'on retrouve une même partition des $m+n$ variables en variables de base et variables hors-base (cette situation est illustrée par l'exercice I.7.5.).

On peut toujours éviter le cyclage en appliquant la règle du plus petit indice (règle de Bland) : lorsque l'on a un choix sur la variable entrante ou sur la variable sortante, on choisit toujours celle du plus petit indice. Nous allons prouver l'efficacité de cette règle.

Théorème I.2 (Théorème de Bland). *Il ne peut y avoir cyclage lorsque, à toute itération effectuée à partir d'un dictionnaire dégénéré, on choisit les variables entrante et sortante comme celles du plus petit indice parmi les candidats possibles.*

Preuve. Supposons que, appliquant la règle de Bland, l'on retrouve deux fois le même dictionnaire D_0 à l'issue d'une suite d'itérations ayant construit les dictionnaires $D_0, D_1, ..., D_k = D_0$; tous ces dictionnaires sont nécessairement dégénérés. On appelle *variable versatile* une variable qui, au cours de ces itérations, est tantôt en base, tantôt hors-base (on notera qu'il existe nécessairement des variables versatiles quand il y a cyclage) ; soit t le plus grand indice des variables versatiles. Dans la suite de dictionnaires D_0, D_1, ..., D_k, D_1, ..., D_k, il existe nécessairement un dictionnaire D' dans lequel x_t est sortante (c'est-à-dire qu'elle est de base dans D' et pas dans le dictionnaire suivant), puis un dictionnaire D'' où x_t est entrante ; nous notons x_s la variable qui entre en base lorsque, à partir de D', x_t sort (x_s n'est pas en base dans D' mais l'est dans le dictionnaire suivant) ; x_s est versatile et on a donc $s < t$.

En notant B l'ensemble des indices des variables de base de D', on peut écrire D' sous la forme :

$$\begin{array}{l} \text{pour } i \in B, x_i = b'_i - \sum_{j \notin B} a'_{ij} x_j \\ \hline z = z^* + \sum_{j \notin B} c'_j x_j \end{array}$$

La variable x_s étant entrante, on a $c'_s > 0$ et, la règle de Bland étant utilisée, on a, pour $j < s$, $c'_j \leq 0$. La variable x_t étant sortante dans D', il vient $a'_{ts} > 0$.

La dernière ligne de D'' peut quant à elle s'écrire :

$$z = z^* + \sum_{k=1}^{n+m} c''_k x_k$$

où c''_k est nul si x_k est en base et $c''_t > 0$.

Pour toute solution $(x^*_1, ..., x^*_{n+m})$ du système des contraintes, on a :

$$z^* + \sum_{j \notin B} c'_j x^*_j = z^* + \sum_{k=1}^{n+m} c''_k x^*_k.$$

Si on définit une solution particulière du système des contraintes en donnant une valeur nulle à toutes les variables non dans B sauf à x_s et une valeur quelconque x^*_s à x_s (les valeurs des autres variables sont alors entièrement déterminées), l'égalité ci-dessus devient :

$$c'_s x^*_s = c''_s x^*_s + \sum_{i \in B} c''_i (b'_i - a'_{is} x^*_s)$$

ou encore :

$$\left(c'_s - c''_s + \sum_{i \in B} c''_i a'_{is} \right) x^*_s = \sum_{i \in B} c''_i b'_i.$$

Cette égalité étant vraie pour toute valeur x^*_s, il vient :

$$c'_s - c''_s + \sum_{i \in B} c''_i a'_{is} = 0.$$

Puisque c'est x_t qui est entrante dans D'' et non x_s alors que l'on a $s < t$, c'est que nous avons $c''_s \leq 0$. Comme nous avons remarqué l'inégalité $c'_s > 0$, il existe un indice r de B avec $c''_r a'_{rs} < 0$.

Par définition de r, la variable x_r était en base dans D' et puisque c''_r est non nul, elle n'est pas en base dans D''. Nous en déduisons que x_r est une variable versatile et donc l'inégalité $r \leq t$.

De plus, c''_t et a'_{ts} étant positifs, leur produit l'est aussi et r ne peut donc pas être égal à t, d'où $r < t$.

De plus, comme x_t entre en base dans D'' alors qu'on a $r < t$, c'est que x_r n'est pas entrante dans D'' et nous n'avons donc pas $c''_r > 0$; par conséquent c'est que nous avons $a'_{rs} > 0$.

Par ailleurs, si $D_k = D_0$, c'est que les valeurs des variables dans les solutions basiques ne varient pas au cours des itérations. Or, x_r est hors-base dans D'' : sa valeur dans la solution basique associée à D'' est 0. La valeur de x_r dans la solution basique associée à D' est donc aussi 0. En conséquence, $b'_r = 0$.

La variable x_r était donc candidate à quitter la base de D' au même titre que x_t et en choisissant x_t, avec $t > r$, nous n'avons pas appliqué la règle du plus petit indice, contradiction. ■

I.6. Recherche d'un dictionnaire réalisable

Nous allons ici encore nous appuyer sur un exemple.

Supposons que nous voulions résoudre le problème suivant, écrit sous forme standard.

$$\begin{array}{l} \text{Maximiser } z = x_1 - x_2 + x_3 \\ \text{avec les contraintes :} \\ \left\{ \begin{array}{l} 2x_1 - x_2 + 2x_3 \leq 4 \\ 2x_1 - 3x_2 + x_3 \leq -5 \\ -x_1 + x_2 - 2x_3 \leq -1 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{array} \right. \end{array}$$

Nous introduisons le *problème auxiliaire* suivant (que nous écrivons sous forme standard également).

$$\begin{array}{l} \text{Maximiser } w = -x_0 \\ \text{avec les contraintes :} \\ \left\{ \begin{array}{l} 2x_1 - x_2 + 2x_3 - x_0 \leq 4 \\ 2x_1 - 3x_2 + x_3 - x_0 \leq -5 \\ -x_1 + x_2 - 2x_3 - x_0 \leq -1 \\ x_0 \geq 0, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{array} \right. \end{array}$$

D'une façon plus générale, on obtient le problème auxiliaire en ajoutant x_0 aux b_i. On peut interpréter ceci en considérant que l'on augmente les ressources d'une quantité x_0. Il est évident que si x_0 est assez grand, les nouvelles ressources deviennent toutes positives ou nulles. D'un autre côté, si le problème initial admet une solution réalisable, on peut prendre $x_0 = 0$. La question est donc de déterminer la plus petite valeur à attribuer à x_0 pour que le problème soit réalisable. On est ainsi amené à minimiser x_0, ou encore à maximiser $-x_0$.

Remarque
On peut se contenter d'enlever x_0 aux premiers membres des inégalités correspondant à une valeur négative des seconds membres, comme il est fait dans le corrigé de l'exercice I.7.6.

Le problème auxiliaire admet des solutions réalisables puisque la solution $x_0^* = 5$, $x_1^* = x_2^* = x_3^* = 0$ en est une. Nous laissons le lecteur se convaincre que le problème initial admet une solution réalisable si et seulement si le problème auxiliaire admet 0 pour valeur optimale de la fonction objectif. De plus, si le problème auxiliaire admet 0 comme valeur optimale, toute solution optimale du problème auxiliaire donne une solution réalisable du problème initial, en « oubliant » x_0 (qui vaut 0 dans ce cas).

Revenons à l'exemple et écrivons le dictionnaire définissant les variables d'écart du problème auxiliaire :

$$\begin{array}{rcrcrcrcrcr}
x_4 & = & 4 & - & 2x_1 & + & x_2 & - & 2x_3 & + & x_0 \\
x_5 & = & -5 & - & 2x_1 & + & 3x_2 & - & x_3 & + & x_0 \\
x_6 & = & -1 & + & x_1 & - & x_2 & + & 2x_3 & + & x_0 \\
\hline
w & = & & & & & & & & - & x_0
\end{array}$$

Ce dictionnaire n'est pas réalisable puisqu'en donnant la valeur 0 aux variables hors-base x_1, x_2, x_3, x_0, les variables d'écart x_5 et x_6 prennent des valeurs négatives. Cependant on peut se ramener en une itération à un dictionnaire réalisable. Il suffit de faire entrer x_0 en base et de faire sortir de la base la variable qui est « la plus négative » (ici x_5). On obtient :

$$\begin{array}{rlcrcrcrcr}
x_0 & = 5 & + & 2x_1 & - & 3x_2 & + & x_3 & + & x_5 \\
x_4 & = 9 & & & - & 2x_2 & - & x_3 & + & x_5 \\
x_6 & = 4 & + & 3x_1 & - & 4x_2 & + & 3x_3 & + & x_5 \\
\hline
w & = -5 & - & 2x_1 & + & 3x_2 & - & x_3 & - & x_5
\end{array}$$

Dans ce dictionnaire, x_2 est variable entrante. Déterminons la variable sortante :

$x_0 \geq 0$ implique $x_2 \leq \dfrac{5}{3}$;

$x_4 \geq 0$ implique $x_2 \leq \dfrac{9}{2}$;

$x_6 \geq 0$ implique $x_2 \leq 1$.

C'est x_6 qui quitte la base. Le dictionnaire suivant est alors :

$$\begin{array}{rcrcrcrcrcr}
x_2 & = & 1 & + & 0,75x_1 & + & 0,75x_3 & + & 0,25x_5 & - & 0,25x_6 \\
x_0 & = & 2 & - & 0,25x_1 & - & 1,25x_3 & + & 0,25x_5 & + & 0,75x_6 \\
x_4 & = & 7 & - & 1,5x_1 & - & 2,5x_3 & + & 0,5x_5 & + & 0,5x_6 \\
\hline
w & = & -2 & + & 0,25x_1 & + & 1,25x_3 & - & 0,25x_5 & - & 0,75x_6
\end{array}$$

À l'étape suivante, x_3 entre en base : x_0 en sort pour donner le dernier dictionnaire.

$$\begin{array}{llllllllll}
x_3 & = 1,6 & - & 0,2x_1 & + & 0,2x_5 & + & 0,6x_6 & - & 0,8x_0 \\
x_2 & = 2,2 & + & 0,6x_1 & + & 0,4x_5 & + & 0,2x_6 & - & 0,6x_0 \\
x_4 & = 3 & - & x_1 & & & - & x_6 & + & 2x_0 \\
\hline
w & = & & & & & & & - & x_0
\end{array}$$

Nous voyons que le problème initial avait une solution réalisable donnée par $x_1^* = 0$; $x_2^* = 2,2$; $x_3^* = 1,6$. Comme indiqué plus haut, à cause de l'équivalence algébrique des dictionnaires, nous déduisons un dictionnaire pour le problème initial en « oubliant » x_0 et en choisissant comme variables en base x_3, x_2, x_4 exprimées ci-dessus en fonction de x_1, x_5, x_6. Il suffit d'exprimer z en fonction des mêmes variables. Nous avons alors comme dictionnaire pour z :

$$\begin{array}{llllllll}
x_3 & = 1,6 & - & 0,2x_1 & + & 0,2x_5 & + & 0,6x_6 \\
x_2 & = 2,2 & + & 0,6x_1 & + & 0,4x_5 & + & 0,2x_6 \\
x_4 & = 3 & - & x_1 & & & - & x_6 \\
\hline
z & = -0,6 & + & 0,2x_1 & - & 0,2x_5 & + & 0,4x_6
\end{array}$$

On peut maintenant partir de ce dictionnaire, réalisable, pour déterminer le maximum de z en appliquant une nouvelle fois l'algorithme du simplexe. Cette méthode est connue comme *méthode à deux phases*. Dans le chapitre II, nous verrons que pour certains problèmes où l'origine n'est pas réalisable (parce que certains des b_i sont négatifs), lorsque tous les coefficients c_j sont négatifs, on peut utiliser le problème appelé « dual », ce qui permet de ne résoudre qu'un problème au lieu de deux. Cette classe de problèmes est dite *dual-réalisable*.

I.7. Exercices

I.7.1. Exercice 1

Résoudre l'exercice suivant par la méthode du simplexe :

Maximiser $z = 3x_1 + 2x_2 + 4x_3$
avec les contraintes :

$$\begin{cases}
x_1 + x_2 + 2x_3 \leq 4 \\
2x_1 \qquad + 3x_3 \leq 5 \\
2x_1 + x_2 + 3x_3 \leq 7 \\
x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.
\end{cases}$$

I.7.2. Exercice 2

Résoudre l'exercice suivant par la méthode du simplexe :

1. en faisant entrer en base la variable de plus grand coefficient dans la fonction objectif ;

2. en faisant entrer en base la variable dont l'augmentation de valeur, à partir de 0, permettra d'augmenter le plus la fonction objectif.

$$\begin{array}{l} \text{Maximiser } z = 5x_1 + 6x_2 + 9x_3 + 8x_4 \\ \text{avec les contraintes :} \\ \left\{ \begin{array}{l} x_1 + 2x_2 + 3x_3 + x_4 \leq 5 \\ x_1 + x_2 + 2x_3 + 3x_4 \leq 3 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{array} \right. \end{array}$$

I.7.3. Exercice 3

Donner un exemple de problème de programmation linéaire, écrit sous forme standard, non borné.

I.7.4. Exercice 4

Même question que pour l'exercice 3, mais ici on veut un problème infaisable.

I.7.5. Exercice 5

On veut appliquer l'algorithme du simplexe au dictionnaire ci-dessous. On envisage deux stratégies lorsqu'il y a plusieurs variables candidates pour entrer dans la base ou pour en sortir.

$$\begin{array}{lllrcrcrcr} x_5 & = & - & 0,5x_1 & + & 5,5x_2 & + & 2,5x_3 & - & 9x_4 \\ x_6 & = & - & 0,5x_1 & + & 1,5x_2 & + & 0,5x_3 & - & x_4 \\ x_7 & = 1 & - & x_1 & & & & & & \\ \hline z & = & & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4 \end{array}$$

1. En cas de choix pour une variable entrante, on prend la variable pourvue du plus grand coefficient dans z et, en cas de choix pour une variable sortante, on prend la variable de plus petit indice. Qu'observe-t-on ?

2. On applique la règle de Bland : en cas de choix pour une variable entrante ou sortante, on prend la variable de plus petit indice. Qu'observe-t-on ?

I.7.6. Exercice 6

1. On considère le problème ci-dessous.

$$\begin{array}{l} \text{Maximiser } z = 5x_1 + 3x_2 \\ \text{avec les contraintes :} \\ \begin{cases} -4x_1 + 5x_2 \leq -10 \\ 5x_1 + 2x_2 \leq 10 \\ 3x_1 + 8x_2 \leq 12 \\ x_1 \geq 0, x_2 \geq 0. \end{cases} \end{array}$$

Montrer à l'aide de l'algorithme du simplexe que ce problème n'admet pas de solution réalisable.
2. On considère maintenant le problème ci-dessous (qui ne diffère du précédent que d'un signe dans la première contrainte). Le résoudre à l'aide de la méthode à deux phases issue de l'algorithme du simplexe.

$$\begin{array}{l} \text{Maximiser } z = 5x_1 + 3x_2 \\ \text{avec les contraintes :} \\ \begin{cases} -4x_1 - 5x_2 \leq -10 \\ 5x_1 + 2x_2 \leq 10 \\ 3x_1 + 8x_2 \leq 12 \\ x_1 \geq 0, x_2 \geq 0. \end{cases} \end{array}$$

Chapitre II

Dualité en programmation linéaire

II.1. Définition du problème dual

Remarque

Nous ne considérons dans ce chapitre que les **problèmes de programmation linéaire écrits sous forme standard**. Pour définir le problème dual d'un problème quelconque de programmation linéaire, on peut le mettre sous forme standard avant de déterminer le problème dual, comme indiqué dans ce chapitre. Un exemple est donné en exercice.

On considère donc le problème (P) :

Maximiser $z = \sum_{j=1}^{n} c_j x_j$

avec les contraintes : $\begin{cases} \text{pour } i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij} x_j \leq b_i \\ \text{pour } j \in \{1, 2, ..., n\}, x_j \geq 0 \end{cases}$

S'il existe m réels y_i positifs ou nuls tels que, pour tout $j \in \{1, 2, ..., n\}$, $\sum_{i=1}^{m} a_{ij} y_i \geq c_j$, alors on a, pour toute solution réalisable $(x_1, ..., x_n)$ de (P) :

$$\sum_{j=1}^{n} c_j x_j \leq \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} y_i \right) x_j = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) y_i \leq \sum_{i=1}^{m} b_i y_i.$$

D'où :

$$\sum_{j=1}^{n} c_j x_j \leq \sum_{i=1}^{m} b_i y_i$$

et cette dernière quantité donne donc un majorant de la fonction objectif.

Le *problème dual* (D) du problème (P) s'écrit :

$$\begin{array}{ll} \text{Minimiser } \sum_{i=1}^m b_i y_i & \\ \text{avec les contraintes : } & \left\{ \begin{array}{l} \text{pour } j \in \{1,2,...,n\}, \sum_{i=1}^m a_{ij} y_i \geq c_j \\ \text{pour } i \in \{1,2,...,m\}, y_i \geq 0 \end{array} \right. \end{array}$$

Le problème (P) prend alors le nom de *problème primal*. On voit que, pour toute solution réalisable $y_1^*, ..., y_m^*$ du dual (c'est-à-dire satisfaisant les contraintes de (D)), $\sum_{i=1}^m b_i y_i^*$ est un majorant de la fonction objectif du problème primal.

Remarque
On établit facilement que le problème dual de (D) est (P).

II.2. Théorème de la dualité

De la définition du problème dual, nous déduisons immédiatement la proposition suivante :

Proposition II.1. *Soient* $(x_1^*, x_2^*, ..., x_n^*)$ *une solution réalisable du problème primal et* $(y_1^*, y_2^*, ..., y_m^*)$ *une solution réalisable du problème dual. On a :*

$$\sum_{j=1}^n c_j x_j^* \leq \sum_{i=1}^m b_i y_i^*$$

De plus, si les deux quantités ci-dessus sont égales, alors $x_1^*, x_2^*, ..., x_n^*$ *constituent une solution optimale du problème primal et* $y_1^*, y_2^*, ..., y_m^*$ *une solution optimale du problème dual.*

Application

La considération du problème dual nous permet de vérifier que nous avons bien trouvé, par l'algorithme du simplexe, une solution optimale pour un problème donné. Nous allons l'expliquer sur le problème traité dans le premier chapitre. Le problème (P) est :

$$\begin{array}{l}\text{Maximiser } z = 7x_1 + 9x_2 + 18x_3 + 17x_4 \\ \text{avec les contraintes :} \\ \left\{\begin{array}{l} 2x_1 + 4x_2 + 5x_3 + 7x_4 \leq 42 \\ x_1 + x_2 + 2x_3 + 2x_4 \leq 17 \\ x_1 + 2x_2 + 3x_3 + 3x_4 \leq 24 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0 \end{array}\right.\end{array}$$

Nous avions établi que l'optimum de ce problème vaut $z^* = 147$ et est obtenu pour $x_1^* = 3, x_2^* = 0, x_3^* = 7, x_4^* = 0$. Nous voulons ici vérifier ce résultat.

Le problème dual (D) s'écrit :

$$\begin{array}{l}\text{Minimiser } 42y_1 + 17y_2 + 24y_3 \\ \text{avec les contraintes :} \\ \left\{\begin{array}{l} 2y_1 + y_2 + y_3 \geq 7 \\ 4y_1 + y_2 + 2y_3 \geq 9 \\ 5y_1 + 2y_2 + 3y_3 \geq 18 \\ 7y_1 + 2y_2 + 3y_3 \geq 17 \\ y_1 \geq 0, y_2 \geq 0, y_3 \geq 0. \end{array}\right.\end{array}$$

Rappelons que, dans le dernier dictionnaire, la fonction objectif s'écrivait :

$$z = 147 - 2x_2 - x_4 - 3x_6 - 4x_7.$$

Nous considérons les valeurs $y_1^* = 0, y_2^* = 3, y_3^* = 4$. Ces valeurs ne sont pas choisies au hasard : ce sont les opposés des coefficients respectivement de x_5, x_6, x_7 dans l'expression ci-dessus de z ; nous justifierons ce choix plus loin.

On a : $42y_1^* + 17y_2^* + 24y_3^* = 147$.

Par ailleurs, on vérifie aisément que les y_i^* satisfont les contraintes du problème dual, donc constituent une solution réalisable du dual.

La proposition ci-dessus nous permet d'affirmer que la valeur 147 est l'optimum du problème primal : ayant trouvé une solution réalisable du dual qui donne à la fonction objectif du dual la valeur que la solution trouvée pour le primal donnait à la fonction objectif du primal, nous pouvons affirmer que nous avions trouvé le maximum de la fonction objectif du primal et que nous avons également trouvé le minimum de la fonction objectif du dual. Cette vérification constitue donc un *certificat d'optimalité* de la solution trouvée pour le primal.

Cette proposition a pour corollaire ce qui suit :

Proposition II.2. *Si le problème primal admet une solution réalisable et est non borné, le problème dual n'admet pas de solution réalisable.*

Preuve. Supposons que le problème dual admette une solution réalisable et notons w^* la valeur correspondante de la fonction objectif du problème dual. La fonction objectif du problème primal est alors majorée par w^*, ce qui contredit l'hypothèse. ■

Remarques
1. En appliquant ce qui précède au problème dual, on a aussi le résultat suivant : si le problème dual admet une solution réalisable et est non borné, le problème primal n'admet pas de solution réalisable.
2. Par contraposée, on obtient l'implication suivante : si (D) (respectivement (P)) admet une solution réalisable, alors (P) (respectivement (D)) n'en admet pas – et dans ce cas (D) (respectivement (P)) n'est pas borné – ou est borné.
3. Il résulte de ce qui précède que (P) et (D) ne peuvent pas être simultanément non bornés.
4. Il existe des cas pour lesquels (P) et (D) sont simultanément non réalisables.

Le théorème suivant, parfois appelé *théorème fondamental de la dualité*, généralise les constatations de l'application faite ci-dessus.

Théorème II.3 (de la dualité)**.** *Si le problème primal a une solution optimale $x_1^*, x_2^*, ..., x_n^*$, alors le problème dual a une solution optimale $y_1^*, y_2^*, ..., y_m^*$ et $\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^*$ (autrement dit, le maximum primal est égal au minimum dual).*

Nous allons prouver ce théorème fondamental en même temps que la proposition suivante :

Proposition II.4. *Si le problème primal admet une solution optimale et si l'expression de la fonction objectif du primal dans le dernier dictionnaire* *obtenu par la méthode du simplexe s'écrit :*

$$z = z^* + \sum_{k=1}^{n+m} d_k x_k$$

(où x_{n+i} représente la i^{e} variable d'écart), alors une solution optimale du problème dual est donnée par $y_i^* = -d_{n+i}$.

Preuve du théorème de la dualité et de la proposition

Supposons le primal résolu par la méthode du simplexe, exposée dans le chapitre I. Aux n variables initiales du problème nous avions ajouté m variables d'écart $x_{n+1}, ..., x_{n+m}$. À la i^{e} contrainte du primal sont associées la variable d'écart x_{n+i} et la variable y_i du dual, ce qui établit un lien canonique entre x_{n+i} et y_i. Considérons l'expression de la fonction objectif du primal dans le dernier dictionnaire du simplexe primal :

$$z = z^* + \sum_{k=1}^{n+m} d_k x_k.$$

Les d_k sont tous négatifs ou nuls (puisqu'il s'agit du dernier dictionnaire) et les d_k associés aux variables en base sont nuls.

Par ailleurs, on a $z^* = \sum_{j=1}^{n} c_j x_j^*$ par définition de z et $x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} x_j$ par définition des variables d'écart.

Posons, pour $i \in \{1, ..., m\}$, $y_i^* = -d_{n+i}$; on a alors : $y_i^* \geq 0$. On a de plus, en distinguant dans z les variables d'écart des autres :

$$z = z^* + \sum_{j=1}^{n} d_j x_j - \sum_{i=1}^{m} \left(b_i - \sum_{j=1}^{n} a_{ij} x_j \right) y_i^*,$$

ou encore :

$$z = z^* - \sum_{i=1}^{m} b_i y_i^* + \sum_{j=1}^{n} \left(d_j + \sum_{i=1}^{m} a_{ij} y_i^* \right) x_j.$$

Mais, par définition de z, on a aussi : $z = \sum_{j=1}^{n} c_j x_j$.

À cause de l'indépendance des variables x_j, on déduit de ces égalités :

$$\begin{cases} z^* = \sum_{i=1}^{m} b_i y_i^* \\ \text{pour } j \in \{1, ..., n\}, c_j = d_j + \sum_{i=1}^{m} a_{ij} y_i^* \end{cases}$$

Les d_j ($j \in \{1, ..., n+m\}$) étant négatifs ou nuls, on obtient finalement :

$$\begin{cases} \text{pour } j \in \{1, ..., n\}, \sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \\ \text{pour } i \in \{1, ..., m\}, y_i^* \geq 0. \end{cases}$$

Les nombres $y_1^*, y_2^*, ..., y_m^*$ forment donc une solution réalisable du problème dual qui donne à la fonction objectif du problème dual la valeur z^*. La proposition du début de ce paragraphe permet de conclure. ■

II.3. Le théorème des écarts complémentaires : un certificat d'optimalité

L'application exposée dans le paragraphe précédent donne une méthode pour démontrer l'optimalité d'une solution du problème primal mais nécessite la connaissance du dernier dictionnaire de la méthode du simplexe. Nous allons voir que l'on peut aussi réussir à fournir un *certificat d'optimalité* du primal, en connaissant seulement les valeurs $x_1^*, ...x_n^*$ qui donnent son maximum à l'objectif du primal.

Théorème II.5 (des écarts complémentaires)**.** *Une solution réalisable $x_1^*, ..., x_n^*$ du primal est optimale si et seulement s'il existe des nombres $y_1^*, ..., y_m^*$ vérifiant ce qui suit :*

- *pour $i \in \{1, ..., m\}$, si $\sum_{j=1}^{n} a_{ij} x_j^* < b_i$, alors $y_i^* = 0$*
- *pour $j \in \{1, ..., n\}$, si $x_j^* > 0$, alors $\sum_{i=1}^{m} a_{ij} y_i^* = c_j$*

et constituant une solution réalisable du problème dual :

$$\begin{cases} \text{pour } j \in \{1, ..., n\}, \sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \\ \text{pour } i \in \{1, ..., m\}, y_i^* \geq 0. \end{cases}$$

De plus, ces nombres $y_1^*, ..., y_m^*$ constituent une solution optimale du problème dual.

Avant de donner la preuve de ce théorème nous allons l'appliquer à l'exemple du chapitre I. Considérons la déclaration :
« $x_1^* = 3, x_2^* = 0, x_3^* = 7, x_4^* = 0$ constituent une solution optimale du primal ».

On vérifie aisément que ces valeurs définissent bien une solution réalisable du problème primal. Cherchons donc s'il existe y_1^*, y_2^*, y_3^* vérifiant :

$$\begin{cases} y_1^* = 0 \text{ puisque la première contrainte du problème « n'est pas saturée »} \\ 2y_1^* + y_2^* + y_3^* = 7 \text{ puisque } x_1^* > 0 \\ 5y_1^* + 2y_2^* + 3y_3^* = 18 \text{ puisque } x_3^* > 0. \end{cases}$$

Utilisant la nullité de y_1^*, on obtient :

$$\begin{cases} y_2^* + y_3^* = 7 \\ 2y_2^* + 3y_3^* = 18. \end{cases}$$

La résolution de ce système donne $y_2^* = 3, y_3^* = 4$. Ces valeurs satisfont bien les inégalités du problème dual. En effet :

$$\begin{aligned} & 4y_1^* + y_2^* + 2y_3^* = 11 \geq 9 \\ \text{et} \quad & 7y_1^* + 2y_2^* + 3y_3^* = 18 \geq 17. \end{aligned}$$

Les deux autres inégalités du même type résultent du système définissant y_1^*, y_2^* et y_3^*. Enfin y_1^*, y_2^*, y_3^* sont positifs ou nuls.

La solution proposée pour le primal est donc bien optimale. On vérifie d'autre part que y_1^*, y_2^*, y_3^* donnent le maximum primal à la fonction objectif duale : ces y_1^*, y_2^*, y_3^* constituent donc bien une solution optimale du problème dual.

Preuve du théorème des écarts complémentaires. La preuve se déduit immédiatement du résultat que nous énonçons puis prouvons ci-dessous.

Si on connaît une solution réalisable (x_j^*) du primal et une solution réalisable (y_i^*) du dual, ces solutions sont optimales si et seulement si :

- pour $j \in \{1, ..., n\}$, $x_j^* = 0$ ou $\sum_{i=1}^m a_{ij} y_i^* = c_j$
- et, pour $i \in \{1, ..., m\}$,$y_i^* = 0$ ou $\sum_{j=1}^n a_{ij} x_j^* = b_i$.

En effet, d'après le théorème de la dualité, on sait que si on connaît une solution réalisable (x_j^*) du primal et une solution réalisable (y_i^*) du dual, ces solutions sont optimales si et seulement si on a :

$$\sum_{j=1}^n c_j x_j^* = \sum_{i=1}^m b_i y_y^*.$$

Or, on a les inégalités suivantes (conséquences de la réalisabilité des solutions) :

$$\sum_{j=1}^{n} c_j x_j^* \leq \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} y_i^* \right) x_j^* = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j^* \right) y_i \leq \sum_{i=1}^{m} b_i y_i^*.$$

Si, avant sommation, une des inégalités était stricte, il en serait de même après sommation. On voit donc qu'il y a égalité entre les bornes de cette suite si et seulement si on a :

- pour $j \in \{1, ..., n\}$, $x_j^* = 0$ ou $\sum_{i=1}^{m} a_{ij} y_i^* = c_j$
- et, pour $i \in \{1, ..., m\}$, $y_i^* = 0$ ou $\sum_{j=1}^{n} a_{ij} x_j^* = b_i$.

■

On peut remarquer que, si l'on peut déterminer les y_i de façon unique, dès lors que l'une des inégalités requises (y compris les contraintes de signe) n'est pas vérifiée, on peut en déduire que la solution n'est pas optimale.

II.4. La signification économique du dual

Nous allons montrer ici que la connaissance de la solution du problème dual peut permettre de prendre en compte des données économiques.

Nous allons considérer que :

- b_i représente la quantité totale de la ressource i ;
- a_{ij} représente la quantité de la ressource i consommée par la fabrication d'une unité de produit j ;
- x_j représente la quantité fabriquée de produit j ;
- c_j représente la valeur unitaire du produit j.

La relation à l'optimum : $z^* = \sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_y^*$ induit que y_i doit représenter la « valeur unitaire de la ressource i ». Ces variables duales y_i sont souvent appelées *prix implicite*. La valeur de y_i donne le montant maximum que l'on serait prêt à payer pour obtenir une unité supplémentaire de la ressource i.

Les relations $\sum_{i=1}^{m} a_{ij}y_i \geq c_j$, $(j \in \{1, ..., n\})$ peuvent se comprendre à l'aide du schéma suivant : supposons qu'une personne étrangère à l'entreprise souhaite acquérir les ressources de l'entreprise ; elle doit proposer pour les ressources un prix tel que ce soit plus intéressant pour l'entreprise de lui vendre ses ressources que de fabriquer elle-même les produits (or, c_j est le profit escompté sur le produit j) et bien sûr elle désire faire cet achat des ressources à un prix minimum. Rappelons que le coefficient a_{ij} représente la quantité de la ressource i requise pour fabriquer une unité de produit j de sorte que $\sum_{i=1}^{m} a_{ij}y_i$ représente la somme à dépenser pour acquérir les ressources nécessaires à la fabrication d'une unité du produit j.

Nous allons donner un second éclairage à l'aide de notre exemple.

Problème

Le fabricant de tissus du chapitre 1 a la possibilité de faire faire à ses ouvriers spécialisés dans la teinture quelques heures supplémentaires à un prix horaire de t euros. A-t-il ou non intérêt à utiliser cette possibilité ?

Pour résoudre ce problème, nous allons énoncer un théorème, que nous démontrerons après avoir résolu notre problème.

Théorème II.6. *On considère le problème (P) :*

$$\begin{array}{l} \textit{Maximiser } z = \sum_{j=1}^{n} c_j x_j \\ \textit{avec les contraintes : } \left\{ \begin{array}{l} \textit{pour } i \in \{1, ..., m\}, \sum_{j=1}^{n} a_{ij}x_j \leq b_i \\ \textit{pour } j \in \{1, 2, ..., n\}, x_j \geq 0. \end{array} \right. \end{array}$$

On suppose que la base optimale de (P) est non dégénérée. Pour des variations δb_i des b_i, on considère le problème (P_δ) défini par :

$$\begin{array}{l} \textit{Maximiser } z = \sum_{j=1}^{n} c_j x_j \\ \textit{avec les contraintes : } \left\{ \begin{array}{l} \textit{pour } i \in \{1, ..., m\}, \sum_{j=1}^{n} a_{ij}x_j \leq b_i + \delta b_i \\ \textit{pour } j \in \{1, 2, ..., n\}, x_j \geq 0. \end{array} \right. \end{array}$$

On suppose que les variations δb_i sont suffisamment faibles pour que la base optimale de (P) soit encore réalisable pour (P_δ). La variation de la valeur optimum de la fonction objectif du programme linéaire vaut alors $\sum_{i=1}^{m} \delta b_i y_i^*$ où $(y_1^*, ..., y_m^*)$ est solution optimale du problème dual de (P).

Notons que la non dégénérescence de la base optimale de (P) implique l'unicité des y_i^*.

Pour notre problème, appelons u le nombre d'heures supplémentaires pour la teinture (avec u petit). La variation du second membre est $(0, 0, u)$. La solution optimale du problème dual est $(0, 3, 4)$. La variation de la fonction objectif est donc égale à $4u$. Il s'agit donc de la variation du chiffre d'affaires que le patron peut espérer de u heures supplémentaires, mais ce n'est pas là un bénéfice net puisqu'elles lui coûteront $t.u$ euros.

On voit qu'il a intérêt à recourir à cette solution dès que $t \leq 4$ euros (il y a peu de chances qu'il puisse convaincre ses ouvriers de faire des heures supplémentaires à ce prix !). On retrouve là l'interprétation de y_i^* : valeur unitaire de la ressource.

Preuve du théorème. On considère la suite des dictionnaires obtenus lorsqu'on résout (P) par la méthode du simplexe. Quand on change b pour $b + \delta b$, seules les constantes des seconds membres sont changées. Si le dernier dictionnaire reste réalisable (c'est-à-dire si les constantes des seconds membres des égalités exprimant les variables de base restent positives ou nulles), alors ce dernier dictionnaire reste optimal. On suppose qu'on est dans ce cas. Les coefficients des variables hors-bases dans la ligne exprimant la fonction z étant inchangés, la solution du problème dual est inchangée. La valeur optimale commune des nouveaux problèmes primal et dual vaut : $\sum_{i=1}^m (b_i + \delta b_i) y_i^* = \sum_{i=1}^m b_i y_i^* + \sum_{i=1}^m \delta b_i y_i^*$ où $(y_1^*, ..., y_m^*)$ est solution optimale du problème dual de P. La variation de la valeur optimum de la fonction objectif vaut $\sum_{i=1}^m \delta b_i y_i^*$. ■

On admet que si la base optimale de (P) est non dégénérée, pour des raisons de continuité, il existe des variations non nulles des δb_i assez petites pour conserver le fait que la base optimale de (P) reste réalisable.

II.5. Problème dual-réalisable

L'utilisation du problème dual permet, sans utiliser l'algorithme à deux phases décrit dans le premier chapitre, de résoudre un problème de programmation linéaire où la solution nulle n'est pas réalisable, pourvu que les coefficients c_j de la fonction objectif du problème écrit sous forme standard soient tous négatifs ou nuls. Un tel problème est dit dual-réalisable.

Exemple

Considérons le problème de programmation linéaire :

$$\text{Minimiser } x_1 + x_2$$
$$\text{avec les contraintes : } \begin{cases} 3x_1 + x_2 \geq 4 \\ -7x_1 + x_2 \geq -7 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

dont l'écriture, sous forme standard est :

$$\text{Maximiser } -x_1 - x_2$$
$$\text{avec les contraintes : } \begin{cases} -3x_1 - x_2 \leq -4 \\ 7x_1 - x_2 \leq 7 \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

Le problème dual s'écrit :

$$\text{Minimiser } -4y_1 + 7y_2$$
$$\text{avec les contraintes : } \begin{cases} -3y_1 + 7y_2 \geq -1 \\ -y_1 - y_2 \geq -1 \\ y_1 \geq 0, y_2 \geq 0. \end{cases}$$

ou encore :

$$\text{Maximiser } 4y_1 - 7y_2$$
$$\text{avec les contraintes : } \begin{cases} 3y_1 - 7y_2 \leq 1 \\ y_1 + y_2 \leq 1 \\ y_1 \geq 0, y_2 \geq 0 \end{cases}$$

Les variables d'écarts constituent maintenant une base réalisable : la méthode du simplexe ne nécessite qu'une seule phase.

II.6. Exercices

II.6.1. Exercice 1

On considère le problème :

$$\text{Maximiser } z = 4x_1 + 3x_2$$
$$\text{avec les contraintes : } \begin{cases} 5x_1 + 3x_2 \leq 30 \\ 2x_1 + 3x_2 \leq 24 \\ x_1 + 3x_2 \leq 18 \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

1. Résoudre graphiquement ce problème.
2. Utiliser le théorème des écarts complémentaires pour prouver que la solution graphique est exacte.
3. La fonction z donne un profit en euros. On envisage de se procurer une unité de plus de la première ressource à un prix unitaire de t euros. Jusqu'à quelle valeur de t cela semble-t-il intéressant ?
4. On suppose qu'on se procure a unités supplémentaires de la première ressource. Jusqu'à quelle valeur de a la base optimale du problème initial, c'est-à-dire la base (x_1, x_2, x_4), reste-t-elle réalisable (auquel cas cette base reste optimale) ?

II.6.2. Exercice 2

On propose $x_1^* = 0, x_2^* = \dfrac{4}{3}, x_3^* = \dfrac{2}{3}, x_4^* = \dfrac{5}{3}, x_5^* = 0$ comme solution optimale du problème suivant :

$$\text{Maximiser } z = 7x_1 + 6x_2 + 5x_3 - 2x_4 + 3x_5$$
$$\text{avec les contraintes : } \begin{cases} x_1 + 3x_2 + 5x_3 - 2x_4 + 2x_5 \leq 4 \\ 4x_1 + 2x_2 - 2x_3 + x_4 + x_5 \leq 3 \\ 2x_1 + 4x_2 + 4x_3 - 2x_4 + 5x_5 \leq 5 \\ 3x_1 + x_2 + 2x_3 - x_4 - 2x_5 \leq 1 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0. \end{cases}$$

Est-ce correct ?

II.6.3. Exercice 3

Donner un exemple de problème (P) tel que ni le problème (P) et le problème dual de (P) n'admettent de solution réalisable.

II.6.4. Exercice 4

1. On considère le problème (P) ci-dessous :

$$\begin{array}{ll} \text{Minimiser} & z = \sum_{j=1}^{n} c_j x_j \\ \text{avec :} & \begin{cases} \text{pour } i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij} x_j \geq b_i \\ \text{pour } i \in \{m+1, ..., m+p\}, \sum_{j=1}^{n} a_{ij} x_j = b_i. \end{cases} \end{array}$$

Exprimer le problème dual du problème (P).

2. On considère le problème (Q) ci-dessous :

$$\begin{array}{ll} \text{Maximiser} & w = \sum_{i=1}^{m+p} b_i y_i \\ \text{avec :} & \text{pour } j \in \{1, 2, ..., n\}, \sum_{i=1}^{m+p} a_{ij} y_i = c_j \\ \text{et} & \text{pour } i \in \{1, 2, ..., m\}, y_i \geq 0. \end{array}$$

Exprimer le problème dual du problème (Q).

3. Établir le théorème suivant (théorème de Farkas et Minkowski) :
Les notations étant celles de ce chapitre, les deux propositions ci-dessous sont équivalentes.

a. Soit $x \in \mathbb{R}^n$;

si on a : $\left\{ \begin{array}{l} \text{pour } i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij} x_j \geq 0 \\ \text{pour } i \in \{m+1, ..., m+p\}, \sum_{j=1}^{n} a_{ij} x_j = 0 \end{array} \right.$

alors : $\sum_{j=1}^{n} c_j x_j \geq 0$;

b. il existe $y \in \mathbb{R}^{m+p}$ tel que :
pour $j \in \{1, 2, ..., n\}, \sum_{i=1}^{m+p} a_{ij} y_i = c_j$
et pour $i \in \{1, 2, ..., m\}, y_i \geq 0$.

Chapitre III

Optimisation non linéaire sans contrainte

III.1. Généralités

On considère ici des fonctions f définies dans $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}$. On cherche à déterminer les points où f atteint des extrema locaux ou globaux. Pour cela, nous avons besoin de quelques définitions.

III.1.1. Gradient

Soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ admettant en un point x des dérivées partielles du premier ordre. On posera $x = (x_1, x_2, ..., x_n)^{\mathrm{t}}$ (les éléments de $\mathbb{R}^n$ sont assimilés à des vecteurs-colonnes).

On note $\nabla f(x)$ et on appelle *gradient* de f au point x le vecteur-colonne :

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), ..., \frac{\partial f}{\partial x_n}(x) \right)^{\mathrm{t}}.$$

Si $F(x) = (f_1(x), ..., f_p(x))$ est un vecteur-ligne où $f_1, ..., f_p$ sont des fonctions réelles de n variables réelles dérivables au point x, alors $\nabla F(x)$ est la matrice dont la j^{e} colonne est $\nabla f_j(x)$.

Les formules suivantes nous seront utiles ultérieurement : si A est une matrice carrée constante d'ordre n, si $u(x)$ et $v(x)$ sont deux vecteurs-colonnes dépendant de x, alors :

$\nabla \left(u^{\mathrm{t}}.A \right) = \nabla \left(u^{\mathrm{t}} \right).A$

et $\nabla(u^{t}.v) = \nabla(u^{t}).v + \nabla(v^{t}).u$

Si f admet en x^0 des dérivées partielles continues, on peut lui appliquer la formule de Taylor à l'ordre 1 :

$f(x) = f(x^0) + (x - x^0)^{t}.\nabla f(x^0) + ||x - x^0||.\epsilon(x)$

où $\epsilon(x)$ est une fonction qui tend vers 0 quand x tend vers x^0.

Remarques

1. Supposons f de classe C^1. Si on considère la surface S de $\mathbb{R}^{n+1}$ d'équation $x_{n+1} = f(x_1, ..., x_n)$, alors $x_{n+1} = f(x^0) + (x - x^0)^{t}.\nabla f(x^0)$ est l'équation de l'hyperplan tangent à S au point $(x^0, f(x^0))$.

2. Nous nous intéresserons par la suite aux variations de f dans une direction d de $\mathbb{R}^n$ donnée, à partir d'un point x^0 de $\mathbb{R}^n$ donné. Pour $s \in \mathbb{R}$, posons $g(s) = f(x^0 + s.d)$. On obtient alors : $g'(s) = d^{t}.\nabla f(x^0 + s.d)$ et $g'(0) = d^{t}.\nabla f(x^0)$.

III.1.2. Matrice hessienne

Si maintenant f admet des dérivées partielles d'ordre 2 en x, on pose :

$$\nabla^2 f(x) = \nabla\left(\nabla f(x)^{t}\right),$$

c'est-à-dire :

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & ... & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & ... & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ ... & ... & ... \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & ... & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix};$$

$\nabla^2 f$ s'appelle la *matrice hessienne* de f.

Si f est une fonction de classe C^2 (autrement dit, f admet des dérivées partielles à l'ordre 2 continues), la matrice hessienne de f est une matrice symétrique.

Si f est de classe C^2 en x^0, on peut écrire la formule de Taylor d'ordre 2 :

$$f(x) = f(x^0) + (x - x^0)^{t}.\nabla f(x^0) + \frac{1}{2}(x - x^0)^{t}\nabla^2 f(x^0).(x - x^0) + ||x - x^0||^2.\epsilon(x),$$

où $\epsilon(x)$ est une fonction qui tend vers 0 quand x tend vers x^0.

III.2. Condition nécessaire et condition suffisante d'optimalité locale

On suppose ici que f est de classe C^2.

Théorème III.1 (condition nécessaire d'optimalité). *Si f admet un minimum local en x^*, alors :*

1. $\nabla f(x^*) = 0$

2. $\nabla^2 f(x^*)$ *est une matrice positive* ($\forall h \in \mathbb{R}^n, h^t.\nabla^2 f(x^*).h \geq 0$).

Preuve. D'après le développement de Taylor à l'ordre 1 en x^*, on a :

$$f(x) = f(x^*) + (x - x^*)^t.\nabla f(x^*) + ||x - x^*||.\epsilon(x),$$

où $\epsilon(x)$ est une fonction qui tend vers 0 quand x tend vers x^*. En particulier, en choisissant $x = x^* - \theta.\nabla f(x^*)$, avec $\theta \in \mathbb{R}$, on obtient :

$$f(x) - f(x^*) = -\theta.||\nabla f(x^*)||^2 + \theta.\epsilon_1(\theta) = \theta\left(-||\nabla f(x^*)||^2 + \epsilon_1(\theta)\right),$$

où $\epsilon_1(\theta)$ est une fonction qui tend vers 0 quand θ tend vers 0. Pour θ positif, $f(x) - f(x^*)$ est du signe de $-||\nabla f(x^*)||^2 + \epsilon_1(\theta)$. Si $\nabla f(x^*) \neq 0$, il existe dans tout voisinage de x^* des points x vérifiant $f(x) < f(x^*)$ (pour θ positif petit, $f(x) - f(x^*)$ est du signe de $-||\nabla f(x^*)||^2$, si on suppose ce terme non nul), contradiction avec l'optimalité locale de x^*. D'où le résultat 1.

Supposons maintenant qu'il existe $h \in \mathbb{R}^n$ tel qu'on ait la relation : $h^t.\nabla^2 f(x^*).h < 0$. On a alors, d'après le développement de Taylor d'ordre 2 :

$$f(x^* + \theta h) - f(x^*) = \theta^2.\left(\frac{1}{2}h^t.\nabla^2 f(x^*).h + \epsilon_2(\theta)\right),$$

où $\epsilon_2(\theta)$ est une fonction qui tend vers 0 quand θ tend vers 0. Pour θ assez petit, la différence $f(x^* + \theta h) - f(x^*)$ serait négative, ce qui contredit l'hypothèse sur x^*. ■

Théorème III.2 (condition suffisante d'optimalité). *Si une fonction f vérifie en x^* :*

1. $\nabla f(x^*) = 0$

2. $\nabla^2 f(x^*)$ *est une matrice définie positive* ($\forall h \in \mathbb{R}^n \backslash \{0\}, h^t \nabla^2 f(x^*).h > 0$),

alors f *admet un minimum local en* x^*.

Preuve. La matrice $\nabla^2 f(x^*)$ étant définie positive, il existe $a > 0$ tel que :

$$\forall h \in \mathbb{R}^n, h^t.\nabla^2 f(x^*).h \geq a||h||^2.$$

En effet, plaçons-nous sur la sphère S de centre 0 et de rayon 1 et définissons a par $a = \inf\{h^{\mathrm{t}}.\nabla^2 f(x^*).h$ pour $h \in S\}$. La sphère étant un compact, la valeur a est atteinte : $\exists h_0 \in S$ tel que $a = h_0^{\mathrm{t}}.\nabla^2 f(x^*).h_0 > 0$. On en déduit aisément le résultat précédent. Soit $x \in \mathbb{R}^n$. Appliquons la formule de Taylor à l'ordre 2 en posant $h = x - x^*$:

$$f(x) - f(x^*) = f(x^* + h) - f(x^*) \geq ||h||^2. \left(\frac{a}{2} + \epsilon(h)\right),$$

où $\epsilon(h)$ est une fonction qui tend vers 0 quand h tend vers 0, ce qui montre le théorème car, pour h de norme assez petite, $\frac{a}{2} + \epsilon(h)$ est du signe de a, c'est-à-dire positif. On a par conséquent $f(x) \geq f(x^*)$ quand x tend vers x^* : x^* est un minimum local de f. ■

III.3. Fonctions quadratiques

Soient A une matrice symétrique d'ordre n, b un vecteur-colonne d'ordre n et c un nombre réel. L'application q de $\mathbb{R}^n$ dans $\mathbb{R}$ définie par :

$$q(x) = c + b^{\mathrm{t}}.x + \frac{1}{2}x^{\mathrm{t}}.A.x$$

s'appelle *fonction quadratique*.

Remarque. La partie polynomiale du développement de Taylor d'ordre 2 d'une fonction f est la fonction quadratique q telle que la surface d'équation $x_{n+1} = q(x)$ soit « la plus proche » de la surface d'équation $x_{n+1} = f(x)$ au voisinage du point considéré.

On a, en utilisant les formules données dans le paragraphe 1 :

$$\nabla q(x) = \nabla(x^{\mathrm{t}}).b + \frac{1}{2}[\nabla(x^{\mathrm{t}}).A.x + \nabla((A.x)^{\mathrm{t}}).x].$$

Or, $\nabla(x^t)$ est la matrice identité. On a de plus les égalités suivantes :

$$\nabla(A.x)^t = \nabla(x^t.A^t) = \nabla(x^t).A^t = A^t = A.$$

D'où l'expression du gradient : $\nabla q(x) = b + A.x$.

De plus : $\nabla^2 q(x) = \nabla((\nabla q(x))^t) = \nabla(b^t + x^t.A^t) = A^t = A$. On a donc finalement : $\nabla^2 q(x) = A$.

Les dérivées d'ordre au moins 3 de q sont nulles. Une fonction quadratique coïncide avec son développement de Taylor à l'ordre 2.

III.4. Fonctions convexes

On dit qu'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ est convexe si, pour tout x et tout y de $\mathbb{R}^n$ et pour tout λ de $]0,1[$, on a : $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$. Si cette inégalité est stricte, on dit que f est strictement convexe.

Théorème III.3. *Si f est une fonction convexe et admet des dérivées partielles, alors f admet un minimum global en x^* si et seulement si on a la relation $\nabla f(x^*) = 0$.*

Preuve. D'après le théorème III.1, si x^* est un minimum local, alors on a $\nabla f(x^*) = 0$. Montrons la réciproque : si $\nabla f(x^*)$ vaut 0, alors f admet un minimum global en x^*. Soit $x \in \mathbb{R}^n$. Posons $g(s) = f(x^* + s(x - x^*))$. On a alors : $g(0) = f(x^*)$, $g(1) = f(x)$ et $g'(0) = (x - x^*)^t.\nabla f(x^*) = 0$. Par ailleurs, on vérifie facilement que g est une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$. La dérivée d'une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$ étant croissante, on a, pour $s \geq 0, g'(s) \geq 0$; g est donc croissante pour $s \geq 0$, d'où $g(1) \geq g(0)$: f admet bien un minimum global en x^*. ■

Théorème III.4. *Si f est convexe et admet un minimum local en x^*, alors f admet un minimum global en x^*.*

Preuve. Si f admet un minimum local en x^*, alors $\nabla f(x^*) = 0$. Si de plus f est convexe, le théorème précédent permet de conclure qu'elle admet un minimum global en x^*. ■

On admettra le théorème suivant.

Théorème III.5. *Si f est deux fois continûment dérivable, les propositions suivantes sont équivalentes :*
a. f est convexe.
b. Pour tout x et tout x^0 de $\mathbb{R}^n$, $f(x) \geq f(x^0) + (\nabla f(x^0))^t.(x - x^0)$ (autrement dit, la surface de $\mathbb{R}^{n+1}$ d'équation $x_{n+1} = f(x)$ est au-dessus de ses hyperplans tangents).
c. Pour tout x, $\nabla^2 f(x)$ est positive.

On déduit qu'une fonction quadratique $q(x) = \frac{1}{2}x^{\mathrm{t}}.A.x + b^{\mathrm{t}}.x + c$ est convexe si et seulement si A est positive. Par ailleurs, si A est définie positive, alors q est strictement convexe et admet un minimum global unique.

III.5. Généralités sur les méthodes d'optimisation sans contrainte

Même si on s'intéresse le plus souvent à des extrema globaux, on cherchera en général des extrema locaux, quitte à examiner ensuite (si possible) s'il s'agit d'extrema globaux.

Toutes les méthodes que nous indiquerons concernent des minima. Les modifications pour la recherche d'un maximum sont toujours immédiates. Par ailleurs, quand nous considérerons des fractions dans ce qui suit, nous supposerons que les dénominateurs sont non nuls (les adaptations étant immédiates sinon).

Pour déterminer un point où une fonction f atteint un minimum local, les méthodes consistent très souvent à construire une suite $x^0, x^1, ..., x^k, ...$ qui doit converger vers un point x^* vérifiant une condition nécessaire d'optimalité. Cette condition (par exemple, $\nabla f(x^*) = 0$) n'est en général pas suffisante et le comportement de f au voisinage de x^* doit donc faire l'objet d'une étude supplémentaire (pouvant porter entre autres sur la matrice hessienne de f en x^*).

Enfin, les méthodes que nous utiliserons seront quasiment toutes des méthodes de descente ; on appelle *méthode de descente* toute méthode où, à chaque étape, on pose $x^{k+1} = x^k + s_k d^k$, où $s_k \in \mathbb{R}^+$ et d^k est une direction de $\mathbb{R}^n$ qui vérifie $(d^k)^{\mathrm{t}}.\nabla f(x^k) < 0$. Cette dernière condition signifie que $f(x^k + sd^k)$ a une dérivée négative pour $s = 0$: partant de x^k dans la direction d^k, f décroît (« on descend »). La différence entre les diverses méthodes de descente porte sur le choix de s_k et de d^k, choix qui doit au minimum assurer $f(x^{k+1}) \leq f(x^k)$.

Lorsque la convergence d'un algorithme a été établie, une qualité importante de cet algorithme est sa *vitesse de convergence*.

- Si on a $\dfrac{||x^{k+1} - x^*||}{||x^k - x^*||} \leq \alpha < 1$ pour k assez grand, on dit que la convergence est *linéaire* de taux α.
- Si $\dfrac{||x^{k+1} - x^*||}{||x^k - x^*||}$ tend vers 0 quand k tend vers l'infini, on dit que la *convergence* est *superlinéaire*.
- Si $\dfrac{||x^{k+1} - x^*||}{||x^k - x^*||^\gamma}$ est borné, avec $\gamma > 1$, on dit que la *convergence* est *superlinéaire* d'ordre γ.
  Dans le cas $\gamma = 2$, on dit que la convergence est *quadratique*.

III.6. Méthodes de gradient

III.6.1. Principe

Il s'agit d'une famille de méthodes itératives qui s'appliquent à des fonctions dérivables et qui utilisent l'idée ci-dessous. Soient d un vecteur de $\mathbb{R}^n$ et x^k un point de $\mathbb{R}^n$ avec $\nabla f(x^k) \neq 0$. Posons, pour $s \in \mathbb{R}$: $g(s) = f(x^k + sd)$.

On dit que d est une *direction de descente* si $g'(0) < 0$. Nous avons vu la relation $g'(0) = d^t.\nabla f(x^k)$. D'où, en notant θ l'angle entre $\nabla f(x^k)$ et d :

$$g'(0) = ||\nabla f(x^k)||.||d||.\cos\theta.$$

En supposant d unitaire, $g'(0)$ est minimum si $\cos\theta = -1$, c'est-à-dire si d est donnée par l'opposé du gradient : $d = -\dfrac{\nabla f(x^k)}{||\nabla f(x^k)||}$. Cette dernière direction donne ce qu'on appelle la *direction de plus grande pente descendante*. C'est ce choix qui est fait dans les méthodes de gradient.

III.6.2. Méthode de la plus forte pente à pas optimal

La méthode de la plus forte pente à pas optimal est la méthode de gradient la plus utilisée. On choisit ici $-\nabla f(x^k)$ pour d^k : il s'agit de la descente de plus forte pente. On pose ensuite $g(s) = f(x^k - s\nabla f(x^k))$ et on calcule s_k de façon à minimiser g pour $s \geq 0$. On est alors ramené à un problème d'optimisation unidimensionelle.

L'algorithme de la plus forte pente peut s'écrire de la façon suivante :

- Choisir un point de départ x^0 ;
- $k \leftarrow 0$
- répéter
  - $d^k \leftarrow -\nabla f(x^k)$
  - déterminer s_k vérifiant $f(x^k + s_k d^k) = \mathrm{Min}_{s \geq 0} f(x^k + sd^k)$
  - $x^{k+1} \leftarrow x^k + s_k d^k$
  - $k \leftarrow k+1$

  tant qu'un test d'arrêt donné n'est pas vérifié.

Le test d'arrêt peut être par exemple :

- on a épuisé un nombre d'itérations fixé à l'avance ;
- le gradient est très petit : $\sum_{i=1}^{n} (\frac{\partial f}{\partial x_i}(x^k))^2 \leq \epsilon$, où ϵ est un paramètre donné ;
- la suite x^k est « presque » stationnaire : $|f(x^{k+1}) - f(x^k)| \leq \epsilon$, où ϵ est un paramètre donné.

On peut aussi exiger que l'un de ces tests soit vérifié sur plusieurs itérations ou que plusieurs tests soient satisfaits simultanément.

On peut montrer que si $f(x)$ est une fonction de classe C^1 qui tend vers l'infini quand $||x||$ tend vers l'infini, cet algorithme converge vers un point stationnaire (point où le gradient s'annule).

L'inconvénient de cette méthode est que la vitesse de convergence peut être très faible (linéaire avec un taux proche de 1). Cette lenteur peut s'expliquer de la façon suivante : l'égalité $\frac{d}{ds}[f(x^k - s.\nabla f(x^k))](s_k) = 0$ s'écrit : $[\nabla f(x^k)]^{\mathrm{t}}.\nabla f(x^{k+1}) = 0$; les directions de déplacement successives sont orthogonales. Sur la figure III.1, on a représenté quelques courbes de niveau et les déplacements.

Il y a convergence en zig-zag.

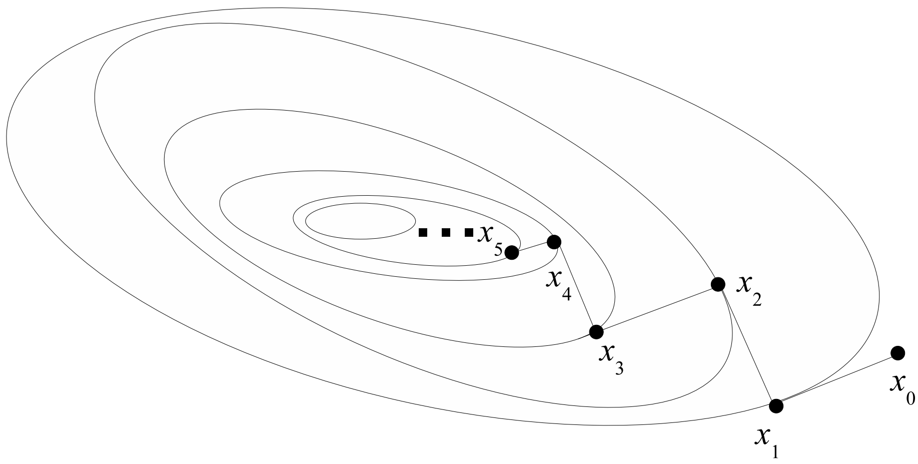

FIGURE III.1 – Convergence en zig-zag

III.7. Méthode de la plus forte pente accélérée

La méthode de la plus forte pente accélérée est une méthode de descente qui s'appuie sur la méthode de la plus forte pente.

Soit p un entier fixé. À partir d'un point x^k, on effectue p itérations de la méthode de la plus forte pente ; on obtient un point y^k et on pose $d^k = y^k - x^k$. Le point x^{k+1} est le point où la fonction $f(x^k + sd^k)$ admet un minimum pour $s > 0$.

Cette méthode peut gagner beaucoup de temps par rapport à la méthode précédente. La figure III.2 illustre cette méthode dans le cas $p = 2$. Pour $p = 1$, on retrouve la méthode précédente.

III.8. Méthode des gradients conjugués

III.8.1. Cas d'une fonction quadratique

Soit $q(x) = \frac{1}{2}x^{\mathrm{t}}.A.x + b^{\mathrm{t}}.x + c$ une fonction quadratique, où A est une matrice symétrique définie positive.

La méthode consiste, à partir d'un point x^0, à minimiser q suivant n directions $d^0, d^1, ..., d^{n-1}$ mutuellement conjuguées par rapport à A, c'est-à-dire vérifiant : pour $1 \leq i < j \leq n, (d^i)^{\mathrm{t}}.A.d^j = 0$.

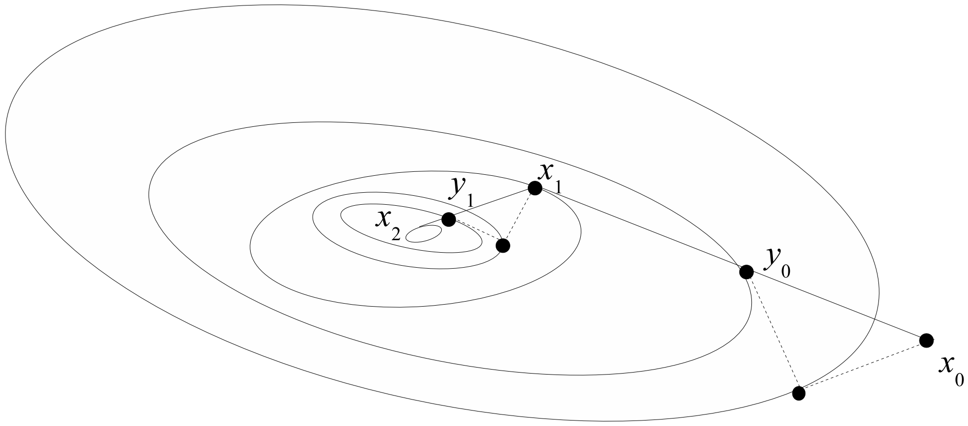

FIGURE III.2 – Convergence accélérée

Soient n telles directions : $d^0, d^1, ..., d^{n-1}$.

Ayant déterminé x^k, le point x^{k+1} est le point : $x^{k+1} = x^k + s_k d^k$ où s_k est choisi de façon à minimiser $q(x^k + s_k d^k)$.

On a donc : $(d^k)^t.\nabla q(x^k + s_k d^k) = 0$ ou encore : $(d^k)^t.[A(x^k + s_k d^k) + b] = 0$ d'où l'on déduit : $s_k = -\dfrac{(d^k)^t.(Ax^k + b)}{(d^k)^t.A.d^k}$.

Lemme III.6. *Si $d^0, d^1, ..., d^{k-1}$ sont mutuellement conjuguées par rapport à A, alors on a, pour tout $i < k$: $(d^i)^t.\nabla q(x^k) = 0$.*

Preuve. On a en effet :

$$\begin{aligned}
(d^i)^t.\nabla q(x^k) &= (d^i)^t.(Ax^k + b) \\
&= (d^i)^t[A(x^i + \sum_{j=i}^{k-1} s_j d^j) + b] \\
&= (d^i)^t.(Ax^i + b) + s_i.(d^i)^t.A.d^i \\
&= 0 \text{ d'après la valeur de } s_i \text{ calculée ci-dessus.}
\end{aligned}$$

■

Théorème III.7. *Si les directions $d^0, d^1, ..., d^{n-1}$ sont mutuellement conjuguées, le point x^n est l'optimum de $q(x)$ sur $\mathbb{R}^n$.*

Preuve. Les directions $d^0, d^1, ..., d^{n-1}$ étant mutuellement conjuguées, elles forment une base de $\mathbb{R}^n$. D'après le lemme III.6, on a, pour tout i vérifiant

$1 \leq i \leq n-1$, $(d^i)^{\mathrm{t}}.\nabla q(x^n) = 0$ d'où $\nabla q(x^n) = 0$, d'où le résultat d'après le théorème III.2. ■

La méthode de Fletcher et Reeves engendre au fur et à mesure les directions d^i ; on l'explicite ci-dessous en posant : $g^k = \nabla q(x^k) = Ax^k + b$.

- Choisir un point de départ x^0.
- $d^0 \leftarrow -g^0$
- $s^0 \leftarrow -\dfrac{(d^0)^{\mathrm{t}}.g^0}{(d^0)^{\mathrm{t}}.A.d^0}$
- $x^1 \leftarrow x^0 + s_0.d^0$
- Pour k variant de 0 à $n-2$ faire
  - ▶ $b_k \leftarrow \dfrac{(d^k)^{\mathrm{t}}.A.g^{k+1}}{(d^k)^{\mathrm{t}}.A.d^k}$
  - ▶ $d^{k+1} \leftarrow -g^{k+1} + b_k.d^k$
  - ▶ $s^{k+1} \leftarrow := -\dfrac{(d^{k+1})^{\mathrm{t}}.g^{k+1}}{(d^{k+1})^{\mathrm{t}}.A.d^{k+1}}$.
  - ▶ $x^{k+2} \leftarrow x^{k+1} + s_{k+1}.d^{k+1}$

Pour justifier la méthode, il suffit de vérifier que $d^0, d^1, ...d^{n-1}$ sont mutuellement conjuguées. Montrons par récurrence sur k que, pour $k \geq 0$, $d^0, d^1, ...d^k$ sont mutuellement conjuguées. Il n'y a rien à vérifier pour $k = 0$. Supposons que cela soit vrai pour un certain k, $0 \leq k \leq n-2$. On a alors pour $k+1$:

$$\begin{aligned}(d^k)^{\mathrm{t}}.A.d^{k+1} &= (d^k)^{\mathrm{t}}.A.(-g^{k+1} + b_k.d^k)\\ &= -(d^k)^{\mathrm{t}}.A.g^{k+1} + b_k.(d^k)^{\mathrm{t}}.A.d^k = 0 \text{ d'après le choix de } b_k.\end{aligned}$$

Pour $i < k$, $(d^{k+1})^{\mathrm{t}}.A.d^i = -(g^{k+1})^{\mathrm{t}}.A.d^i + b_k.(d^k)^{\mathrm{t}}.A.d^i = -(g^{k+1})^{\mathrm{t}}.A.d^i$.

Or : $A.d^i = A\left(\dfrac{x^{i+1} - x^i}{s_i}\right) = \dfrac{Ax^{i+1} - Ax^i}{s_i} = \dfrac{g^{i+1} - g^i}{s_i}$.

D'autre part :

- $g^{i+1} = -d^{i+1} + b_i.d^i$
- si $i \geq 1$, $g^i = -d^i + b_{i-1}.d^{i-1}$ et $g^0 = -d^0$.

D'après le lemme, g^{k+1} est orthogonal à d^{i+1}, d^i et d^{i-1} ; $A.d^i$ étant combinaison linéaire de ces trois vecteurs, $(g^{k+1})^{\mathrm{t}}.A.d^i = 0$, ce qui montre l'égalité $(d^{k+1})^{\mathrm{t}}.A.d^i = 0$.

Pour terminer, nous démontrons une formule qui nous sera utile dans le paragraphe suivant.

On a : $g^{k+1} - g^k = A(x^{k+1} - x^k) = s_k A.d^k$.

D'où : $(g^{k+1})^{\mathrm{t}}.A.d^k = \dfrac{(g^{k+1})^{\mathrm{t}}.(g^{k+1} - g^k)}{s_k}$.

Comme $g^k = -d^k + b_{k-1}.d^{k-1}$, le lemme montre l'égalité $(g^{k+1})^{\mathrm{t}}.g^k = 0$.

D'où : $b_k = \dfrac{1}{s_k} \dfrac{(g^{k+1})^{\mathrm{t}}.g^{k+1}}{(d^k)^{\mathrm{t}}.A.d^k} = -\dfrac{(g^{k+1})^{\mathrm{t}}.g^{k+1}}{(d^k)^{\mathrm{t}} g^k}$.

Or : $(d^k)^{\mathrm{t}}.g^k = (-g^k + b_{k-1}.d^{k-1})^{\mathrm{t}}(g^k) = -(g^k)^{\mathrm{t}}.g^k$ d'après le lemme.

On en déduit le résultat : $b_k = \dfrac{||g^{k+1}||^2}{||g^k||^2}$.

III.8.2. Cas d'une fonction quelconque

L'algorithme de Fletcher et Reeves pour une fonction quelconque est le suivant :

- partir d'un point x^0 ;
- faire $d^0 \leftarrow -\nabla f(x^0)$ et $k \leftarrow 0$;
- répéter
  - ▶ choisir s_k minimisant $f(x^k + sd^k)$, par rapport à s
  - ▶ $x^{k+1} \leftarrow x^k + s_k.d^k$
  - ▶ $b_k \leftarrow \dfrac{||\nabla f(x^{k+1})||^2}{||\nabla f(x^k)||^2}$
  - ▶ $d^{k+1} \leftarrow -\nabla f(x^{k+1}) + b_k.d^k$
  - ▶ $k \leftarrow k + 1$

  jusqu'à ce qu'un test d'arrêt soit vérifié.

Cette méthode a deux avantages : elle nécessite le stockage de très peu d'informations et sa vitesse de convergence est très supérieure à celle des algorithmes de gradient classiques.

III.9. Méthode de Newton

On suppose ici que f est de classe C^2.

Au voisinage d'un point x^k, on approche f par la fonction quadratique donnée par la formule de Taylor d'ordre 2 :

$$q(x) = f(x^k) + (x - x^k)^t.\nabla f(x^k) + \frac{1}{2}(x - x^k)^t\nabla^2 f(x^k).(x - x^k).$$

On peut alors choisir pour x^{k+1} le point, s'il existe, qui minimise q ; pour que ce point minimisant q existe, il est suffisant que $\nabla^2 f(x^k)$ soit définie positive ; il est alors déterminé par l'équation $\nabla q(x) = 0$ qui s'écrit :

$$\nabla f(x^k) + \nabla^2 f(x^k).(x - x^k) = 0\,;$$

d'où :

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1}.\nabla f(x^k).$$

Proposition III.8. *On suppose la fonction f de classe C^3. Si x^0 est choisi suffisamment proche d'un minimum local x^* où la matrice hessienne de f est définie positive, alors la suite (x^k) a une convergence quadratique vers x^*.*

Preuve.

On va établir une condition suffisante portant sur x^0 pour assurer une convergence quadratique de la suite (x^k) construite à partir de x^0 par la méthode de Newton. Pour cela, on considère les éléments suivants :

- On sait que $\nabla^2 f(x^*)$ est définie positive. Par continuité de la fonction $\nabla^2 f$, il existe une boule B_0 de centre x^* et de rayon r_0 sur laquelle $\nabla^2 f(x)$ est définie positive et donc inversible. On note alors M un majorant de $||(\nabla^2 f)^{-1}||$ sur B_0 (un tel majorant existe puisque $(\nabla^2 f)^{-1}$ est continue et que B_0 est un compact).
- Pour tout $k \geq 0$, on définit une fonction ϕ_k par $\phi_k(x^k) = 0$ et, pour $x \neq x^k$:

$$\nabla f(x) - \nabla f(x^k) = \nabla^2 f(x^k)(x - x^k) + \phi_k(x)||x - x^k||^2. \qquad (1)$$

On note $\nabla^3 f$ le tableau tridimensionnel $(n \times n \times n)$ formé des dérivées troisièmes de la fonction f. Soit N un majorant de $||\nabla^3 f||$ dans la boule B_1 centrée en x^* et de rayon 1 (un tel majorant existe puisque, f étant de classe C^3, $||\nabla^3 f||$ est continue et que B_1 est un compact). Pour x dans B_1, on peut montrer (à l'aide de l'inégalité de Taylor-Lagrange appliquée à ∇f à l'ordre 1) que l'on a, pour tout $k \geq 0$, $||\phi_k(x)|| \leq N/2$.
- On pose $M' = \frac{MN}{2}$.
- On considère un réel r vérifiant simultanément : $r < 1$, $r \leq r_0$, $r \leq \frac{1}{M'}$ et on appelle B la boule centrée en x^* et de rayon r (on notera que B est donc incluse dans B_0 et dans B_1).

On va montrer que la suite (x^k) converge quadratiquement vers x^* si on part d'un point x^0 appartenant à B. On fait donc désormais cette hypothèse : $x^0 \in B$.

Montrons d'abord, par récurrence, que la suite (x^k) est constituée de points de B. Par choix de x^0, cette propriété est vraie pour $k = 0$. Supposons-la vraie jusqu'à l'ordre k : pour $k \geq 0$, $x^k \in B$, ou encore : $||x^k - x^*|| \leq r$. On a :

$$x^{k+1} - x^* = (x^{k+1} - x^k) + (x^k - x^*). \tag{2}$$

Par construction : $x^{k+1} - x^k = -[\nabla^2 f(x^k)]^{-1}.\nabla f(x^k)$. (3)

Appliquons la formlule (1) avec $x = x^*$ et en utilisant $\nabla f(x^*) = 0$ (conséquence de la minimalité de x^*) :

$$\nabla^2 f(x^k)(x^k - x^*) = \nabla f(x^k) + \phi_k(x)||x^* - x^k||^2. \tag{4}$$

La matrice $\nabla^2 f(x^k)$ étant inversible (puisque x^k appartient à B et donc à B_0), on multiplie les deux membres de (4) à gauche par $[\nabla^2 f(x^k)]^{-1}$:

$$x^k - x^* = [\nabla^2 f(x^k)]^{-1}\nabla f(x^k) + [\nabla^2 f(x^k)]^{-1}\phi_k(x)||x^* - x^k||^2. \tag{5}$$

En utilisant les égalités (2), (3) et (5), on obtient :

$$\begin{aligned} x^{k+1} - x^* &= -[\nabla^2 f(x^k)]^{-1}.\nabla f(x^k) \\ &\qquad + [\nabla^2 f(x^k)]^{-1}\nabla f(x^k) + [\nabla^2 f(x^k)]^{-1}\phi_k(x)||x^* - x^k||^2 \\ &= [\nabla^2 f(x^k)]^{-1}\phi_k(x)||x^* - x^k||^2. \end{aligned}$$

D'où : $||x^{k+1} - x^*|| \leq ||[\nabla^2 f(x^k)]^{-1}||.||\phi_k(x)||.||x^* - x^k||^2$, ce qui entraîne $||x^{k+1} - x^*|| \leq M.N.||x^* - x^k||^2/2 = M'||x^* - x^k||^2$.
En utilisant l'hypothèse de récurrence, on obtient : $||x^{k+1} - x^*|| \leq M'.r^2$, ce qui donne, puisque r est inférieur ou égal à $\frac{1}{M'}$, $||x^{k+1} - x^*|| \leq r$. Par conséquent, l'hypothèse de récurrence est vérifiée à l'ordre $k + 1$.

Montrons maintenant que la suite (x^k) converge vers x^*. On a l'inégalité $||x^{k+1} - x^*|| \leq M'||x^k - x^*||^2$, d'où : $||x^k - x^*|| \leq M'||x^0 - x^*||^{2k} \leq M'r^{2k}$.

Comme r est plus petit que 1, la suite (r^{2k}) converge vers 0 et il en est donc de même pour la suite $(||x^k - x^*||)$: (x^k) converge bien vers x^*.

Enfin, $||x^{k+1} - x^*|| \leq M'||x^k - x^*||^2$ peut se récrire $\frac{||x^{k+1}-x^*||}{||x^k-x^*||^2} \leq M'$, ce qui est la définition d'une convergence quadratique. ■

La contrainte de choisir x^0 proche de x^* est forte ; on peut éventuellement appliquer d'abord une autre méthode pour s'approcher de x^*, puis appliquer la méthode de Newton.

III.10. Optimisation unidimensionnelle

On considère ici une application f de $\mathbb{R}$ dans $\mathbb{R}$. Les méthodes d'optimisation de telles fonctions ont d'autant plus d'importance qu'elles servent d'outils pour l'optimisation de fonctions de plusieurs variables.

III.10.1. Méthode de Newton

La méthode de Newton que nous avons étudiée dans le cas des fonctions de plusieurs variables réelles peut s'appliquer ici. À partir d'un point x_k, on approche f par :

$q(x) = f(x_k) + f'(x_k)(x - x_k) + \dfrac{1}{2}f''(x_k)(x - x_k)^2$.

On remarque la relation suivante : $q'(x) = f'(x_k) + f''(x_k)(x - x_k)$.

Si $f''(x_k) > 0$ (cas où f est strictement convexe autour de x_k), on pose :

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

qui est le point où q atteint son minimum ($q'(x_{k+1}) = 0$).

Si $f''(x_k) \leq 0$, la méthode échoue.

III.10.2. Dichotomie pour une fonction dérivable

On suppose que f est dérivable.

Définition III.1. *On dit qu'une fonction est unimodale s'il existe un réel* x^* *pour lequel la fonction est strictement décroissante sur* $]-\infty, x^*]$ *et strictement croissante sur* $[x^*, +\infty[$.

Le point x^* est alors minimum global de f. Il vérifie la relation $f'(x^*) = 0$.

On suppose ici que f est unimodale. La première étape consiste en la recherche de x_{min} et x_{max} tels qu'on ait les deux relations $f'(x_{min}) < 0$ et $f'(x_{max}) > 0$.

Après cette première étape, on pose : $x = \dfrac{1}{2}(x_{min} + x_{max})$; si $f'(x) > 0$, on remplace x_{max} par x, sinon on remplace x_{min} par x ; on répète l'opération jusqu'à un critère d'arrêt à préciser.

La longueur de l'intervalle étant à chaque itération divisée par 2, on montre que la convergence est linéaire de taux 0,5.

Pour déterminer x_{min} et x_{max}, une bonne méthode est la suivante (on suppose dans ce qui suit que $f'(0)$ n'est pas nul ; dans le cas contraire, 0 est la solution du problème !) :

• définir un pas de déplacement $h > 0$.

• si $f'(0) < 0$, faire :

$x_{min} \leftarrow 0$
tant que $f'(h) < 0$, faire
$x_{min} \leftarrow h$
$h \leftarrow 2h$
$x_{max} \leftarrow h$

sinon si $f'(0) > 0$, faire :

$h \leftarrow -h$
$x_{max} \leftarrow 0$
tant que $f'(h) > 0$, faire
$x_{max} \leftarrow h$
$h \leftarrow 2h$
$x_{min} \leftarrow h$

III.10.3. Interpolation quadratique

On suppose ici que f est une fonction qui peut ne pas être dérivable ou bien dont on ne connaît pas la dérivée.

La méthode part du principe suivant : on choisit d'abord, à l'aide d'un algorithme préliminaire, x_1, x_2 et x_3 vérifiant : $x_1 < x_2 < x_3$ ainsi que les inégalités $f(x_2) \leq f(x_1)$ et $f(x_2) \leq f(x_3)$. On approche f par une fonction quadratique q ayant les mêmes valeurs que f en x_1, x_2 et x_3 :

$$q(x) = f(x_1).\frac{(x-x_2).(x-x_3)}{(x_1-x_2).(x_1-x_3)} + f(x_2).\frac{(x-x_1).(x-x_3)}{(x_2-x_1).(x_2-x_3)} + f(x_3).\frac{(x-x_1).(x-x_2)}{(x_3-x_1).(x_3-x_2)}.$$

Le minimum de q est atteint sur $[x_1, x_3]$ en un point dont l'abscisse s'exprime facilement en fonction de $x_1, x_2, x_3, f(x_1), f(x_2)$ et $f(x_3)$; on note x_4 ce point. La mise à jour des points x_1, x_2 et x_3 se fait selon les règles suivantes :

- si $f(x_4) \leq f(x_2)$

si $x_4 \leq x_2$, le nouveau triplet est (x_1, x_4, x_2)
sinon le nouveau triplet est (x_2, x_4, x_3)

- sinon

  si $x_4 \leq x_2$, le nouveau triplet est (x_4, x_2, x_3)

  sinon le nouveau triplet est (x_1, x_2, x_4)

On peut montrer que, si f est assez régulière, la convergence est superlinéaire d'ordre 1,3.

III.10.4. Dichotomie sans dérivation

On suppose ici que f est unimodale. Au départ, à l'aide d'un algorithme préliminaire, on choisit a et b tels que le minimum de f soit atteint entre a et b. On partage alors, à l'aide de points d, c et e, l'intervalle $[a, b]$ en quatre sous-intervalles égaux : $c = \dfrac{a+b}{2}, d = \dfrac{a+c}{2}, e = \dfrac{c+b}{2}$.

En comparant les valeurs prises par f en a, b, c, d et e, on peut éliminer deux des sous-intervalles définis par ces points et affirmer que le minimum de f est atteint dans l'union de deux sous-intervalles contigus $[a_1, c_1]$ et $[c_1, b_1]$. La figure III.3 illustre un tel cas. On recommence alors avec l'intervalle $[a_1, b_1]$. À chaque étape, la longueur de l'intervalle est divisée par 2. La vitesse de convergence est linéaire.

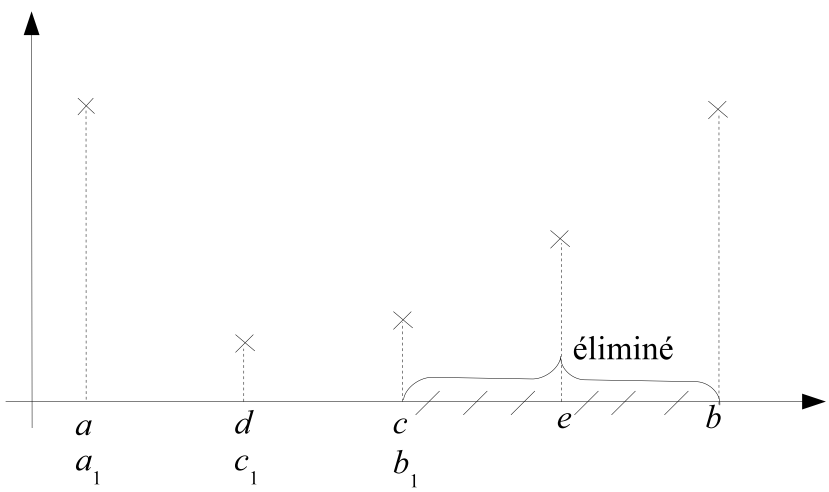

FIGURE III.3 – Dichotomie sans dérivation

III.11. Exercices

III.11.1. Exercice 1

On s'intéresse au minimum de la fonction f définie sur $\mathbb{R}^2$ par :

$$f(x, y) = e^{x+y} + x^2 + 2y^2.$$

Appliquer trois itérations de la méthode du gradient à pas optimal.

Chapitre IV

Optimisation non linéaire avec contraintes

IV.1. Généralités

On considère dans tout le chapitre des fonctions f, g_i $(1 \leq i \leq m)$ et h_j $(1 \leq j \leq p)$ de $\mathbb{R}^n$ dans $\mathbb{R}$ **de classe $\mathbf{C}^1$**.

Soit (P) le problème :

$$\begin{array}{ll} & \text{Minimiser } f(x) \\ \text{avec} & \left\{ \begin{array}{ll} \text{pour } i \in \{1, ..., m\}, & g_i(x) \leq 0 \\ \text{pour } j \in \{1, ..., p\}, & h_j(x) = 0 \end{array} \right. \\ \text{et} & x \in \mathbb{R}^n. \end{array} \qquad (P)$$

On pose $I = \{1, ..., m\}$ et $J = \{1, ..., p\}$.

Les conditions $g_i(x) \leq 0$ et $h_j(x) = 0$ s'appellent les *contraintes*. Tout x appartenant à $\mathbb{R}^n$ qui vérifie le système des contraintes s'appelle *solution réalisable*. L'ensemble des solutions réalisables est le *domaine réalisable*. On note X le domaine réalisable. Si, pour $i \in \{1, 2, ..., m\}$ et pour $x \in X$, on a $g_i(x) = 0$, on dit que la contrainte g_i est *saturée* en x.

Par sa définition, le domaine réalisable est un fermé de $\mathbb{R}^n$. On supposera dans tout ce chapitre que **le domaine réalisable est non vide**.

Théorème IV.1. *Si le domaine réalisable est borné, alors le problème (P) admet au moins une solution.*

Preuve. Avec les hypothèses, le domaine réalisable est un fermé, borné non vide de $\mathbb{R}^n$. Le résultat résulte immédiatement d'un théorème de topologie :

« toute fonction continue sur un fermé borné non vide de $\mathbb{R}^n$ est bornée et atteint ses bornes ».

Définition IV.1. *Soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. On dit que f est coercive si elle tend vers $+\infty$ quand $||x||$ tend vers $+\infty$.*

Théorème IV.2. *Si la fonction f est coercive, le problème (P) admet au moins une solution.*

Preuve. Soit x une solution réalisable. Puisque f est coercive, il existe une boule B de $\mathbb{R}^n$ centrée sur l'origine telle que, pour tout x' non dans B, $f(x') > f(x)$. L'ensemble $B \cap X$ étant un ensemble fermé et borné de $\mathbb{R}^n$, la fonction f atteint son minimum sur $B \cap X$. Ce minimum est aussi un minimum pour le problème (P). ■

Définition IV.2. *On dit qu'une direction d est admissible en $x^0 \in X$ s'il existe une fonction ϕ de $\mathbb{R}$ dans $\mathbb{R}^n$ telle que :*
1. *$\phi(0) = x^0$*
2. *pour tout $t > 0$ assez petit, $\phi(t) \in X$*
3. *la dérivée à droite de ϕ en 0 est d.*

Soit $x^0 \in X$. On note $A(x^0)$ l'ensemble des directions admissibles en x^0 ; on pose $I_0(x^0) = \{i \in I \text{ vérifiant } g_i(x^0) = 0\}$.

Proposition IV.3. *Si d est une direction admissible en x^0, alors :*
(i) pour $i \in I_0(x^0)$, $d^t.\nabla g_i(x^0) \leq 0$
(ii) pour $j \in J, d^t.\nabla h_j(x^0) = 0$.

Preuve. Soit ϕ une fonction correspondant à la définition IV.2.
(i) Si $g_i(x^0) = 0$, on a : $g_i(\phi(t)) = td^t.\nabla g_i(x^0) + t\epsilon(t)$ où $\epsilon(t) \to 0$ quand $t \to 0$. Pour $t > 0$ assez petit, $g_i(\phi(t)) \leq 0$ et donc : $d^t.\nabla g_i(x^0) + \epsilon(t) \leq 0$, ce qui donne le résultat en passant à la limite quand $t \to 0$.
(ii) On a : $h_j(\phi(t)) = h_j(x^0) + td^t.\nabla h_j(x^0) + t\epsilon(t)$ où $\epsilon(t) \to 0$ quand $t \to 0$.

Pour t assez petit, $h_j(\phi(t)) = 0$ et $h_j(x^0) = 0$; on a donc pour $t > 0$ assez petit :

$$d^t.\nabla h_j(x^0) + \epsilon(t) = 0,$$

ce qui donne le résultat en passant à la limite quand $t \to 0$. ■

On note $B(x^0)$ l'ensemble des directions d vérifiant :
- pour $i \in I_0(x^0)$, $d^t.\nabla g_i(x^0) \leq 0$

- pour $j \in J, d^t.\nabla h_j(x^0) = 0$.

La proposition IV.3 se réécrit : $A(x^0) \subseteq B(x^0)$.

Définition IV.3. *On dit que les contraintes sont qualifiées en $x^0 \in X$ si toute direction dans $B(x^0)$ est limite d'une suite de directions de $A(x^0)$.*

Les propositions suivantes donnent des conditions suffisantes pour que des contraintes soient qualifiées.

Proposition IV.4. *Si :*
- *les fonctions g_i sont convexes,*
- *les fonctions h_j sont affines,*
- *et il existe $\tilde{x} \in X$ avec,*
  pour tout $i \in I, g_i(\tilde{x}) < 0$
  pour tout $j \in J, h_j(\tilde{x}) = 0$

alors les contraintes sont qualifiées en tout point de X.

Proposition IV.5. *On suppose que, pour $j \in J$, les fonctions h_j sont affines. Si, en $x^0 \in X$, l'ensemble des gradients*
- *$\nabla g_i(x^0)$ pour $i \in I_0(x^0)$*
- *$\nabla h_j(x^0)$ pour $j \in J$*

sont linéairement indépendants, alors les contraintes sont qualifiées en x^0.

Avant de prouver ces propositions, on établit deux lemmes :

Lemme IV.6. *On suppose que, pour $j \in J$, les fonctions h_j sont affines. Soient $x^0 \in X$ et d une direction vérifiant :*
- *pour $i \in I_0(x^0), d^t.\nabla g_i(x^0) < 0$*
- *pour $j \in J$, $d^t.\nabla h_j(x^0) = 0$.*

Alors d est une direction admissible en x^0.

Preuve. Pour $t \geq 0$, on pose : $\phi(t) = x^0 + td$. On a $\phi(0) = x^0$ et $\phi'(0) = d$, autrement dit les points 1 et 3 de la définition d'une direction admissible sont satisfaits.

Pour $j \in J$, les fonctions h_j étant supposées affines, on peut écrire : $h_j(\phi(t)) = h_j(x^0) + td^t.\nabla h_j(x^0)$. Par hypothèse sur x^0, on a $h_j(x^0) = 0$ et, par hypothèse sur d, on a $d^t.\nabla h_j(x^0) = 0$. D'où $h_j(\phi(t)) = 0$.

De plus, pour $i \in I_0(x^0)$, on peut écrire :
$g_i(\phi(t)) = g_i(x^0) + t(d^t.\nabla g_i(x^0) + \epsilon(t))$, où $\epsilon(t) \to 0$ quand $t \to 0$.

De $g_i(x^0) = 0$ et $d^t.\nabla g_i(x^0) < 0$, on déduit $g_i(\phi(t)) \leq 0$ pour t positif assez petit.

Ainsi, la direction d est admissible. ■

Lemme IV.7. *On suppose que, pour $j \in J$, les fonctions h_j sont affines. Soit $x^0 \in X$. S'il existe une direction $\tilde{d}$ telle que :*
- pour $i \in I_0(x^0)$, $\tilde{d}^t.\nabla g_i(x^0) < 0$
- pour $j \in J, \tilde{d}^t.\nabla h_j(x^0) = 0$,
alors les contraintes sont qualifiées en x^0.

Preuve.

Soient $d \in B(x^0)$ et $\tilde{d}$ vérifiant les hypothèses du lemme.

Pour $\lambda \in [0, 1[$, soit $d_\lambda = \lambda d + (1 - \lambda)\tilde{d}$.

Pour $i \in I_0(x^0)$: $d_\lambda^t.\nabla g_i(x^0) = \lambda d^t.\nabla g_i(x^0) + (1 - \lambda)\tilde{d}^t.\nabla g_i(x^0) < 0$.

Pour $j \in J$: $d_\lambda^t.\nabla h_j(x^0) = \lambda d^t.\nabla h_j(x^0) + (1 - \lambda)\tilde{d}^t.\nabla h_j(x^0) = 0$.

Le lemme IV.6 indique que, pour tout $\lambda \in [0, 1[$, d_λ est une direction admissible. En considérant une suite λ_n de nombres tendant vers 1 par valeurs inférieures, on obtient une suite d_{λ_n} de directions admissibles qui tend vers d : cela montre que les contraintes sont qualifiées en x^0. ■

Preuve de la proposition IV.4.

Soit $\tilde{x}$ vérifant les hypothèses de la proposition et x^0 un point quelconque de X.

En utilisant la convexité des g_i, on a pour $i \in I_0(x_0)$:

$0 > g_i(\tilde{x}) \geq g_i(x^0) + (\tilde{x} - x^0)^t.\nabla g_i(x^0)$.

D'où, comme $g_i(x^0) = 0, (\tilde{x} - x^0)^t.\nabla g_i(x^0) < 0$.

On pose $\tilde{d} = \tilde{x} - x^0$; on a donc $\tilde{d}^t.\nabla g_i(x^0) < 0$.

Pour $j \in \{1, ..., p\}$, les h_j étant affines : $0 = h_j(\tilde{x}) = h_j(x^0) + \tilde{d}^t.\nabla h_j(x^0)$ d'où on déduit : $\tilde{d}^t.\nabla h_j(x^0) = 0$.

On utilise le lemme IV.7 : les contraintes sont qualifiées en x^0. Comme x^0 est quelconque, on conclut que les contraintes sont qualifiées en tout point de X. ■

Preuve de la proposition IV.5.

On considère les deux problèmes (Q) et (R) d'optimisation linéaire ci-dessous :

Maximiser $z = \sum_{i \in I_0(x^0)} \lambda_i$

avec $\begin{cases} \sum_{j \in J} \mu_j \nabla h_j(x^0) - \sum_{i \in I_0(x^0)} \lambda_i \nabla g_i(x^0) = 0 & (Q) \\ \text{pour } i \in I_0(x^0), \lambda_i \geq 0 \end{cases}$

Minimiser $w = 0$

avec $\begin{cases} \text{pour } i \in I_0(x^0), d^{\mathrm{t}}.\nabla g_i(x^0) \leq -1 \\ \text{pour } j \in J, d^{\mathrm{t}}.\nabla h_j(x^0) = 0. & (R) \end{cases}$.

On peut facilement vérifier que les problèmes (Q) et (R) sont duaux l'un de l'autre.

Le problème (Q) est réalisable puisque la solution nulle l'est ; montrons qu'il est majoré par 0. Supposons qu'il puisse prendre une valeur strictement positive. Alors, dans cette solution, au moins un λ_i $(i \in I_0(x^0))$ est non nul et les vecteurs $\nabla g_i(x^0)$ $(i \in I_0(x^0))$ et $\nabla h_j(x^0)$ $(j \in J)$ sont linéairement dépendants, ce qui est contraire à l'hypothèse. Le maximum du problème (Q) vaut donc 0.

On utilise le théorème de la dualité pour l'optimisation linéaire : le problème (R) est réalisable. On note $\tilde{d}$ une solution réalisable de (R). Il suffit alors d'utiliser le lemme IV.7 pour conclure. ■

Remarque.

On peut montrer que la proposition IV.5 est encore exacte si on retire l'hypothèse que les fonctions h_j sont affines.

Un exemple de points où les contraintes ne sont pas qualifiées

On considère dans $\mathbb{R}^2$ le domaine représenté sur la figure IV.1. défini par :

$$\begin{cases} x^2 + (y-1)^2 \geq 1 \\ x \leq 1 \\ y \geq 0. \end{cases}$$

On pose $g_1(x, y) = 1 - x^2 - (y-1)^2$, $g_2(x, y) = x - 1$, $g_3(x, y) = -y$; les contraintes s'écrivent : $g_1(x, y) \leq 0$, $g_2(x, y) \leq 0$, $g_3(x, y) \leq 0$.

Au point (0, 0), les contraintes g_1 et g_3 sont saturées.

$\nabla g_1(x, y) = \begin{pmatrix} -2x \\ 2(1-y) \end{pmatrix}$, $\nabla g_1(0, 0) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ et $\nabla g_3(0, 0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$.

La direction $d = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ vérifie $d^{\mathrm{t}}.\nabla g_1(0, 0) = 0$ et $d^{\mathrm{t}}.\nabla g_3(0, 0) = 0$, la direction d appartient à $B(0, 0)$. Or, la seule direction admissible en (0, 0) est

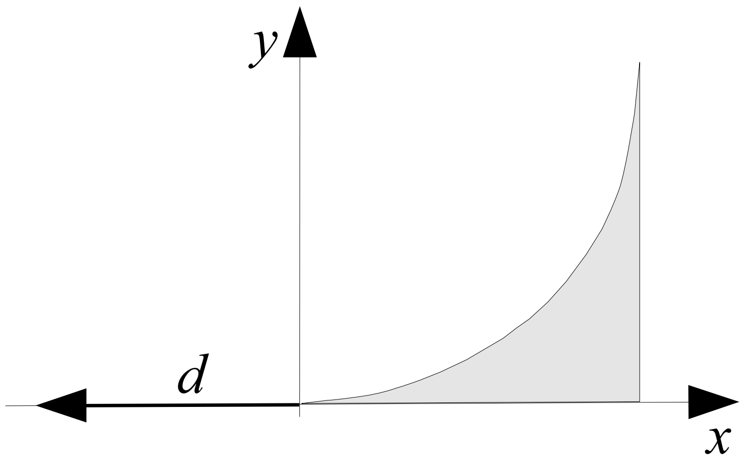

FIGURE IV.1 – Contraintes non qualifiée en (0, 0)

la direction $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. La direction d n'est pas limite d'une suite de directions admissibles.

On établit enfin le théorème suivant :

Théorème IV.8. *On suppose que le problème admet un minimum local en un point x^* où les contraintes sont qualifiées. Alors, si $d \in B(x^*)$:*

$$d^t.\nabla f(x^*) \geq 0$$

(par conséquent, aucune direction admissible en x^* n'est de descente).

Preuve. Soit d_k une suite de directions admissibles tendant vers d et soit ϕ_k la fonction associée à d_k. Soit $t > 0$. Il vient :

$f[\phi_k(t)] = f(x^*) + td_k^{\mathrm{t}}.\nabla f(x^*) + t\epsilon(t)$ où $\epsilon(t) \to 0$ quand $t \to 0$.

Si t est assez petit pour que $\phi_k(t)$ appartienne à X : $f[\phi_k(t)] \geq f(x^*)$. On a alors : $t[d_k^{\mathrm{t}}.\nabla f(x^*) + \epsilon(t)] \geq 0$ et donc $d_k^{\mathrm{t}}.\nabla f(x^*) + \epsilon(t) \geq 0$.

Par passage à la limite quand t tend vers 0, on obtient $d_k^{\mathrm{t}}.\nabla f(x^*) \geq 0$. ■

IV.2. Condition de Lagrange

On s'intéresse ici au problème :

$$\begin{array}{l} \text{Minimiser } f(x) \\ \text{avec } x \in \mathbb{R}^n \\ \text{et, pour } 1 \leq j \leq p, \quad h_j(x) = 0 \end{array}$$

où les fonctions f et h_j $(1 \leq j \leq p)$ sont de classe C^1. La condition de Lagrange, que donne le théorème suivant, fournit une condition nécessaire pour qu'un élément de $\mathbb{R}^n$ soit un minimum local de (P).

Théorème IV.9. *Soit x^* un minimum local du problème. On suppose que les contraintes sont qualifiées en x^*. Alors il existe des réels $\mu_1, \mu_2, ..., \mu_p$ tels que : $\nabla f(x^*) = \sum_{j=1}^{p} \mu_j \nabla h_j(x^*)$.*

Preuve. Notons E le sous-espace de $\mathbb{R}^n$ engendré par les vecteurs $\nabla h_j(x^*)$ et $E^\perp$ le sous-espace orthogonal à E. On a :

$\nabla f(x^*) = y + z$ avec $y \in E$ et $z \in E^\perp$.

Pour $1 \leq i \leq p$, $(-z)^{\mathrm{t}}.\nabla h_j(x^*) = 0$ puisque $-z$ appartient à $E^\perp$. Par conséquent, $-z$ est une direction admissible ; d'après le théorème du paragraphe précédent, il vient : $(-z)^{\mathrm{t}}.\nabla f(x^*) \geq 0$.

Or : $(-z)^{\mathrm{t}}.\nabla f(x^*) = (-z)^{\mathrm{t}}.y + (-z)^{\mathrm{t}}.z = (-z)^{\mathrm{t}}.z = -||z||^2$.

La relation $-||z||^2 \geq 0$ entraîne : $||z||^2 = 0$ et donc $z = 0$. D'où le théorème. ■

Le théorème IV.10, conséquence directe du théorème IV.12 démontré plus loin, donne des hypothèses pour lesquelles la condition de Lagrange est suffisante.

Théorème IV.10. *La condition de Lagrange est suffisante lorsque f est convexe dans un ouvert contenant X et que les h_j $(1 \leq i \leq p)$ sont affines.*

IV.3. Condition de Karush, Kuhn et Tucker

On reprend le problème (P) initial :

$$\begin{array}{lll} & & \text{Minimiser } f(x) \\ (P) & \text{avec} & \begin{cases} \text{pour } 1 \leq i \leq m, & g_i(x) \leq 0 \\ \text{pour } 1 \leq j \leq p, & h_j(x) = 0 \end{cases} \\ & \text{et} & x \in \mathbb{R}^n \end{array}$$

où les fonctions f, g_i $(1 \leq i \leq m)$ et h_j $(1 \leq j \leq p)$ sont supposées de classe C^1. La condition suivante, appelée condition de Karush, Kuhn et Tucker, donne une condition suffisante d'optimalité et généralise la condition de Lagrange :

Théorème IV.11. *On suppose que les contraintes sont qualifiées en x^* et que x^* est un minimum local du problème ; alors il existe :*

- *m nombres réels positifs ou nuls $\lambda_1, \lambda_2, ..., \lambda_m$*
- *p nombres réels $\mu_1, \mu_2, ..., \mu_p$*

tels que :

$\nabla f(x^*) = \sum_{j \in J} \mu_j \nabla h_j(x^*) - \sum_{i \in I_0(x^*)} \lambda_i \nabla g_i(x^*)$.

Preuve. On utilise d'abord le théorème IV.8.

L'hypothèse $d \in B(x^*)$ peut s'écrire :

- pour tout $i \in I_0(x^*)$, $\sum_{k=1}^n \dfrac{\partial g_i}{\partial x_k}(x^*) d_k \leq 0$

- pour tout $j \in J$, $\sum_{k=1}^n \dfrac{\partial h_j}{\partial x_k}(x^*) d_k = 0$.

La conclusion $d^t.\nabla f(x^*) \geq 0$ peut s'écrire : $\sum_{k=1}^n \dfrac{\partial f}{\partial x_k}(x^*) d_k \geq 0$.

On pose :

- pour tout $i \in I_0(x^*)$ et $k \in \{1, 2, ..., n\}$, $a_{ik} = -\dfrac{\partial g_i}{\partial x_k}(x^*)$

- pour tout $j \in J$ et $k \in \{1, 2, ..., n\}$, $b_{jk} = \dfrac{\partial h_j}{\partial x_k}(x^*)$

- pour tout $k \in \{1, 2, ..., n\}$, $c_k = \dfrac{\partial f}{\partial x_k}(x^*)$.

Avec ces notations, le théorème IV.8 s'écrit :

pour tout $d \in \mathbb{R}^n$, si on a :

- pour tout $i \in I_0(x^*)$ $\sum_{k=1}^n a_{ik} d_k \geq 0$
- pour tout $j \in J$, $\sum_{k=1}^n b_{jk} d_k = 0$,

alors $\sum_{k=1}^n c_k d_k \geq 0$.

On utilise le théorème de Farkas et Minkowski qui fait l'objet de l'exercice II.6.4. :

il existe $\{\lambda_i \text{ pour } i \in I_0(x^*)\}$ et $\mu_1, ..., \mu_p$ tels que :

- pour $k \in \{1, ..., n\}$, $\sum_{i \in I_0(x^*)} a_{ik} \lambda_i + \sum_{j \in J} b_{jk} \mu_j = c_k$

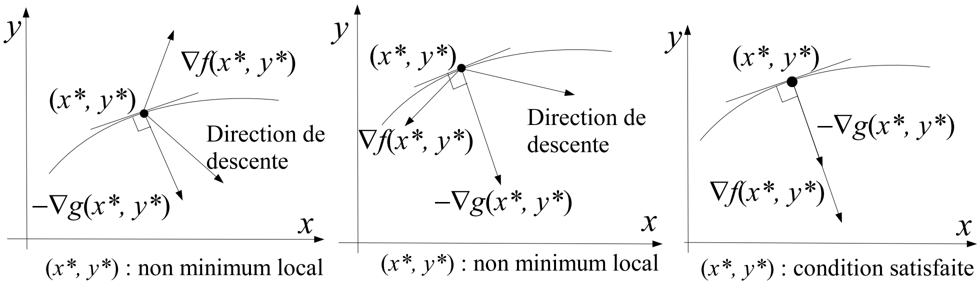

FIGURE IV.2 – Dans le plan, une seule contrainte saturée

- pour $i \in I_0(x^*)$, $\lambda_i \geq 0$.

La première ligne s'écrit :

pour $k \in \{1, ..., n\}$, $\sum_{j \in J} \mu_j \dfrac{\partial h_j}{\partial x_k}(x^*) - \sum_{i \in I_0(x^*)} \lambda_i \dfrac{\partial g_i}{\partial x_k}(x^*) = \dfrac{\partial f}{\partial x_k}(x^*)$

ou enfin : $\sum_{j \in J} \mu_j \nabla h_j(x^*) - \sum_{i \in I_0(x^*)} \lambda_i \nabla g_i(x^*) = \nabla f(x^*)$.

Avec la positivité des λ_i, on obtient l'énoncé du théorème de Karush, Kuhn et Tucker. ■

Nous illustrons ci-dessous deux cas où il n'y a pas de contrainte d'égalité ; la condition de Karush, Kuhn et Tucker exprime alors qu'il est nécessaire que $\nabla f(x^*)$ se décompose sur $\{-\nabla g_i(x^*) | i \in I_0(x^*)\}$ avec des coefficients positifs ou nuls.

Illustration

$\star$ *Cas* $n = 2, p = 0$ *et une seule contrainte d'inégalité saturée*

On suppose qu'on est dans $\mathbb{R}^2$; on note x et y les coordonnées d'un point. On suppose que seule la contrainte $g(x, y) \leq 0$ est saturée en (x^*, y^*) : $g(x^*, y^*) = 0$. Le vecteur $-\nabla g(x^*, y^*)$ est perpendiculaire à la courbe d'équation $g(x, y) = 0$ et dirigée vers l'intérieur du domaine. Si le vecteur $\nabla f(x^*, y^*)$ fait un angle non nul avec $-\nabla g(x^*, y^*)$, on peut trouver une direction de descente pour f dirigée vers l'intérieur du domaine et le point (x^*, y^*) n'est donc pas un minimum local.

La figure IV.2 illustre ce cas ; on rappelle qu'une direction est une direction de descente si elle fait un angle obtus avec $\nabla f(x^*, y^*)$ et qu'une direction admissible fait un angle aigu avec $-\nabla f(x^*, y^*)$.

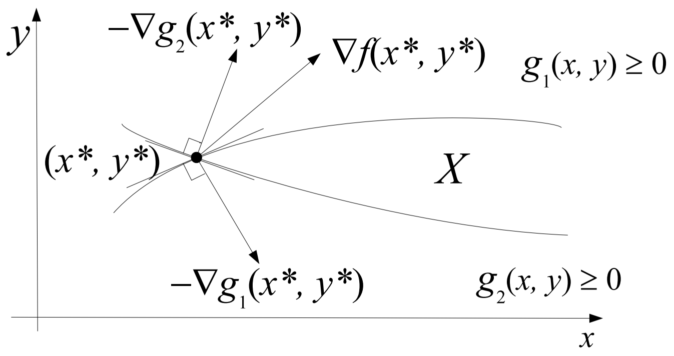

FIGURE IV.3 – Dans le plan, deux contraintes saturées

⋆ *Cas $n = 2, p = 0$ et deux contraintes d'inégalité saturées*

Sur la figure IV.3, on voit que pour qu'aucune direction de descente ne soit dirigée vers le domaine, il est nécessaire que le vecteur $\nabla f(x^*, y^*)$ se situe dans le secteur formé par $-\nabla g_1(x^*, y^*)$ et $-\nabla g_2(x^*, y^*)$. C'est le cas sur la figure.

Le théorème IV.12 donne des hypothèses sous lesquelles la condition de Karush, Kuhn et Tucker est suffisante pour un minimum local.

Théorème IV.12. *On suppose que les contraintes sont qualifiées en un point x^*. La condition de Karush, Kuhn et Tucker en x^* est suffisante pour avoir un minimum local lorsque simultanément f est convexe dans un voisinage de x^*, les g_i ($i \in I_0(x^*)$) sont convexes dans un voisinage de x^* et les h_j ($1 \leq j \leq p$) sont affines dans un voisinage de x^*.*

Preuve.

On suppose qu'il existe des nombres réels positifs ou nuls λ_i ($i \in I_0(x^*)$) et des nombres réels μ_j ($1 \leq j \leq p$) tels que :

$$\nabla f(x^*) = \sum_{j \in J} \mu_j \nabla h_j(x^*) - \sum_{i \in I_0(x^*)} \lambda_i \nabla g_i(x^*).$$

On considère une boule B de centre x^* dans laquelle f et les g_i sont convexes et les h_j affines. Soit $x \in B \cap X$, on va montrer que $f(x) \geq f(x^*)$, ce qui prouvera le théorème.

La convexité de f dans B induit : $f(x) \geq f(x^*) + (x - x^*)^{\mathrm{t}}.\nabla f(x^*)$.
En utilisant la condition de Karush, Kuhn et Tucker :

$f(x) \geq f(x^*) + \sum_{j \in J} \mu_j (x - x^*)^t . \nabla h_j(x^*) - \sum_{i \in I_0(x^*)} \lambda_i (x - x^*)^t . \nabla g_i(x^*)$.
Pour $j \in J, (x - x^*)^t . \nabla h_j(x^*) = h_j(x) - h_j(x^*) = 0$.
Soit $i \in I_0(x^*)$; la fonction g_i étant convexe dans B :
$g_i(x) \geq g_i(x^*) + (x - x^*)^t . \nabla g_i(x^*)$.
On a donc :
$(x - x^*)^t . \nabla g_i(x^*) \leq g_i(x) - g_i(x^*)$.
Or, on a $\lambda_i \geq 0$; de plus, par hypothèse, $g_i(x^*) = 0$ et $g_i(x) \leq 0$, d'où :
$\lambda_i (x - x^*)^t . \nabla g_i(x^*) \leq 0$.
On obtient finalement $f(x) \geq f(x^*)$: f admet un minimum local en x^*. ■

IV.4. Méthode de descente

Dans cette section IV.4., nous nous intéressons au problème suivant :

$$\begin{array}{ll} & \text{Minimiser } f(x) \\ \text{avec,} & \text{pour } 1 \leq j \leq m, \ g_i(x) \leq 0. \end{array}$$

Pour tenter de résoudre ce problème, on choisit un point de départ $x^0 \in X$ et on construit de façon itérative une suite x^k de X vérifiant $f(x^{k+1}) \leq f(x^k)$, jusqu'à ce qu'on estime avoir obtenu une approximation satisfaisante.

À partir de x^k, on recherche une direction de descente d qui ne fasse pas sortir « immédiatement » de X. On cherche alors, en se déplaçant dans la direction d, un point x^{k+1} de X meilleur que x^k (par exemple, en minimisant $f(x^k + sd)$ pour $s > 0$, avec la contrainte que $x^k + s.d$ appartienne à X, si on sait résoudre ce nouveau problème). On recommence à partir de x^{k+1} tant qu'un certain critère d'arrêt n'est pas vérifié. Pour choisir d, on peut résoudre le problème :

$$\begin{array}{ll} & \text{Minimiser } d^t . f(x) \\ \text{avec} & d^t . \nabla g_i(x^k) \leq 0 \text{ pour tout } i \text{ tel que } g_i(x^k) = 0 \\ \text{et} & d^t . d = 1. \end{array}$$

On obtient alors pour d la direction dans $B(x^k)$ de plus grande descente.

On peut choisir de remplacer la condition $d^t . d = 1$ par la condition : $-1 \leq d_i \leq 1$ $(1 \leq i \leq n)$ pour avoir un problème linéaire ; dans ce cas, la direction retenue ne sera pas exactement la direction de plus grande descente.

La méthode, telle qu'elle vient d'être exposée, doit être accompagnée de compléments. Considérons l'exemple représenté sur la figure IV.4. Tout déplacement dans la direction d fait sortir de X. Il faut alors une procédure de projection pour que x^{k+1} soit dans X, procédure qui peut être schématiquement représentée par la figure IV.5.

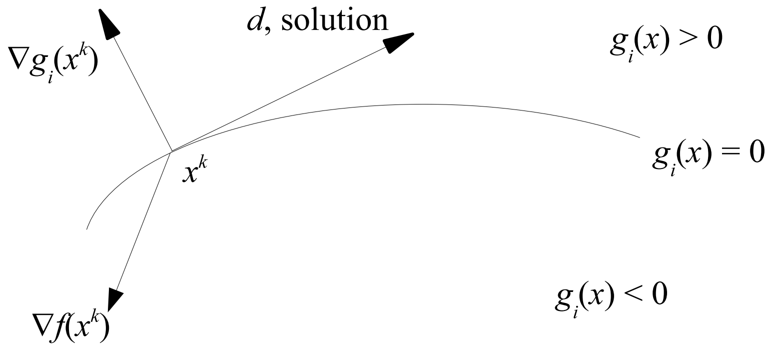

FIGURE IV.4 – Sortie du domaine réalisable

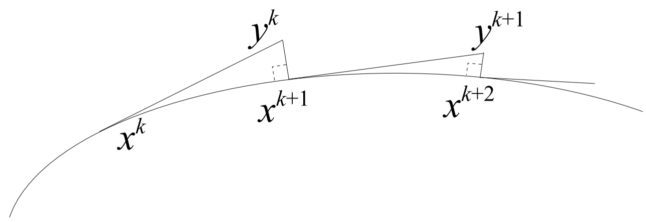

FIGURE IV.5 – Projection sur le domaine réalisable

Remarquons néanmoins que ce dépassement n'est pas utile, par exemple, dans le cas où les contraintes sont affines.

IV.5. Cas des fonctions convexes

IV.5.1. Généralités

On suppose dans toute cette partie que :

- les fonctions g_i $(1 \leq i \leq m)$ sont convexes,
- les fonctions h_j $(1 \leq j \leq p)$ sont affines,
- la fonction f est convexe.

Définition IV.4. *Soit C une partie de $\mathbb{R}^n$. On dit que C est convexe si, pour tout couple (x, x') de points de C, le segment d'extrémités x et x' est contenu dans C.*

Remarque

On vérifie aisément que le domaine réalisable du problème (P), c'est-à-dire :

$$\{x \in \mathbb{R}^n | g_i(x) \leq 0 \text{ pour } 1 \leq i \leq m \text{ et } h_j(x) = 0 \text{ pour } 1 \leq j \leq p\}$$

est convexe.

Théorème IV.13. *Si f est strictement convexe, le problème (P) admet au plus une solution optimale.*

Preuve

Supposons qu'il existe dans C deux solutions optimales x et y. Posons $z = \dfrac{x+y}{2}$. La convexité du domaine implique que $z \in C$ et la stricte convexité de f implique $f(z) < \dfrac{f(x)+f(y)}{2}$. On a donc soit $f(z) < f(x)$, soit $f(z) < f(y)$, contradiction avec l'optimalité supposée de x et y. ■

Des théorèmes IV.13 et IV.1, on déduit :

Théorème IV.14. *Si le domaine réalisable est borné et si f est strictement convexe, le problème (P) admet une unique solution.*

Des théorèmes IV.13 et IV.2, on déduit :

Théorème IV.15. *Si f est strictement convexe et coercive, le problème (P) admet une unique solution.*

Théorème IV.16. *On suppose qu'on a un minimum local en un point x^* où les contraintes sont qualifiées. Alors le problème (P) admet un minimum global en x^*.*

Preuve. Soit x une solution réalisable du problème (P).

Soit ϕ la fonction définie sur l'intervalle $[0, 1]$ par : $\phi(t) = f[x^* + t(x - x^*)]$.

On a : $\phi(0) = f(x^*)$ et $\phi(1) = f(x)$. De plus, la convexité de f entraîne celle de ϕ.

Par ailleurs : $\phi'(0) = (x - x^*)^t.\nabla f(x^*)$.

La direction $d = x - x^*$ appartient à $B(x^*)$ car, si ce n'était pas le cas, on sortirait du domaine réalisable en suivant la direction d à partir de x^*. Le théorème IV.8 montre que $\phi'(0) \geq 0$. La fonction ϕ étant convexe sur [0, 1], $\phi'(0) \geq 0$ implique $\phi(1) \geq \phi(0)$, c'est-à-dire : $f(x) \geq f(x^*)$. ■

On peut maintenant donner des conditions pour que la condition de Karush, Kuhn et Tucker soit suffisante pour un minimum global en s'appuyant sur les théorèmes IV.12 et IV.16.

Théorème IV.17. *On suppose que la condition de Karush, Kuhn et Tucker est vérifiée en un point x^* où les contraintes sont qualifiées. Avec les hypothèses de cette partie, le point x^* est un minimum global de (P). De plus, si f est strictement convexe, x^* est l'unique point où (P) atteint le minimum global.*

Preuve. D'après le théorème IV.12, le problème (P) atteint un minimum local en x^*. D'après le théorème IV.16, le problème (P) admet le minimum global en x^*. Si, de plus, f est strictement convexe, le théorème IV.13 permet de conclure que (P) atteint son minimum global uniquement en x^*.

IV.5.2. Linéarisation : introduction

On s'intéresse au problème de la recherche du minimum d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}$ de classe C^1 sur un domaine fermé convexe X. On va « linéariser » f, c'est-à-dire approcher f par son développement de Taylor à l'ordre 1 en une suite de points construite en utilisant la linéarisation. Ceci conduit à un algorithme simple décrit ci-dessous, dont on va cependant constater les limites.

On considère l'algorithme suivant :

- $x^0 \leftarrow$ un point quelconque
- $k \leftarrow 0$
- répéter

  $x^{k+1} \leftarrow$ un point qui minimise $f(x^k) + (x - x^k)^t.\nabla f(x^k)$ sur X.

  $k \leftarrow k + 1$

jusqu'à ce qu'un test d'arrêt à préciser soit vérifié.

Remarques.

1) Le point x^{k+1} est aussi un point qui minimise $x^t.\nabla f(x^k)$ sur X puisque cette fonction ne diffère de la fonction $f(x^k) + (x - x^k)^t.\nabla f(x^k)$ que par une constante.

2) Si le domaine est un polyèdre, la détermination de x^{k+1} est un problème d'optimisation linéaire.

Appliquons cet algorithme au problème dans $\mathbb{R}^2$ suivant :

$$(P_0) \qquad \begin{array}{l} \text{Minimiser } f(x,y) = (x-3)^2 + (y-5)^2 \\ \text{avec les contraintes : } \left\{ \begin{array}{l} y - 2x \leq 0 \\ 2x + y - 20 \leq 0 \\ -2x + 3y - 4 \leq 0 \\ y \geq 0. \end{array} \right. \end{array}$$

La figure IV.6 illustre ce problème. La fonction objectif est constante sur des cercles centrés sur le point C de coordonnées (3, 5). On peut donc anticiper sur le fait qu'elle est minimum pour le point du domaine le plus proche du centre de C, c'est-à-dire le point $(\frac{49}{13}, \frac{50}{13})$. On note X le domaine réalisable, grisé sur la figure.

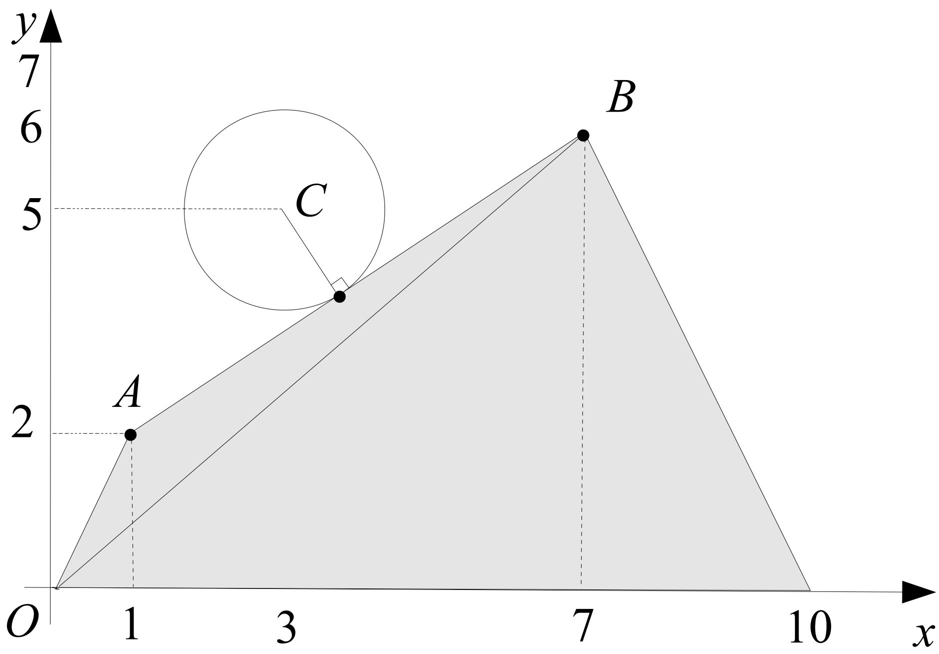

FIGURE IV.6 – Problème avec la linéarisation

On a : $\nabla f(x,y) = \begin{pmatrix} 2x-6 \\ 2y-10 \end{pmatrix}$.

On démarre l'algorithme à partir de l'origine O, avec $\nabla f(0,0) = \begin{pmatrix} -6 \\ -10 \end{pmatrix}$. On cherche le minimum sur X de $-6x-10y$. Les considérations développées dans le chapitre I montrent que le minimum est atteint en un des quatre sommets de X ; il est facile de montrer qu'il s'agit du point $B=(7,6)$.

On recommence à partir du sommet B, avec $\nabla f(7,6) = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$. On cherche le minimum sur le X de $8x+2y$. Il est atteint en un des quatre sommets de X, on vérifie qu'il s'agit du point $A=(1,2)$.

Si on recommence à partir de A, , avec $\nabla f(1,2) = \begin{pmatrix} -4 \\ -6 \end{pmatrix}$. On cherche le minimum sur le X de $-4x-6y$. On retrouve le sommet B et ainsi de suite. La méthode ne converge donc pas.

IV.5.3. Linéarisation : méthode de Frank et Wolfe

La méthode de Frank et Wolfe s'applique à la minimisation d'une fonction de classe C^1 convexe sur un domaine convexe et compact.

L'algorithme est le suivant :

- $x^0 \leftarrow$ un point quelconque
- $k \leftarrow 0$
- répéter
  - $\tilde{x^k} \leftarrow$ un point qui minimise $x^{\mathrm{t}}.\nabla f(x^k)$ sur X
  - $x^{k+1} \leftarrow$ un point qui minimise f sur le segment $[x^k, \tilde{x^k}]$
  - $k \leftarrow k+1$

jusqu'à ce qu'un test d'arrêt à préciser soit vérifié.

Appliquons cette méthode au problème précédent (P_0) à partir de l'origine O : $(x_0, y_0) = (0,0)$. Le déroulement de l'algorithme est illustré par la figure IV.7.

On a : $\nabla f(x,y) = \left\{ \begin{array}{l} 2x-6 \\ 2y-10) \end{array} \right.$.

Étape 1 : à partir du point (0, 0).

On cherche le minimum de $l_1(x,y) = (x,y)^{\mathrm{t}}.\nabla f(0,0) = -6x-10y$ sur X. Le point $(\tilde{x_0}, \tilde{y_0})$ qui atteint ce minimum est le point $B=(7,6)$.

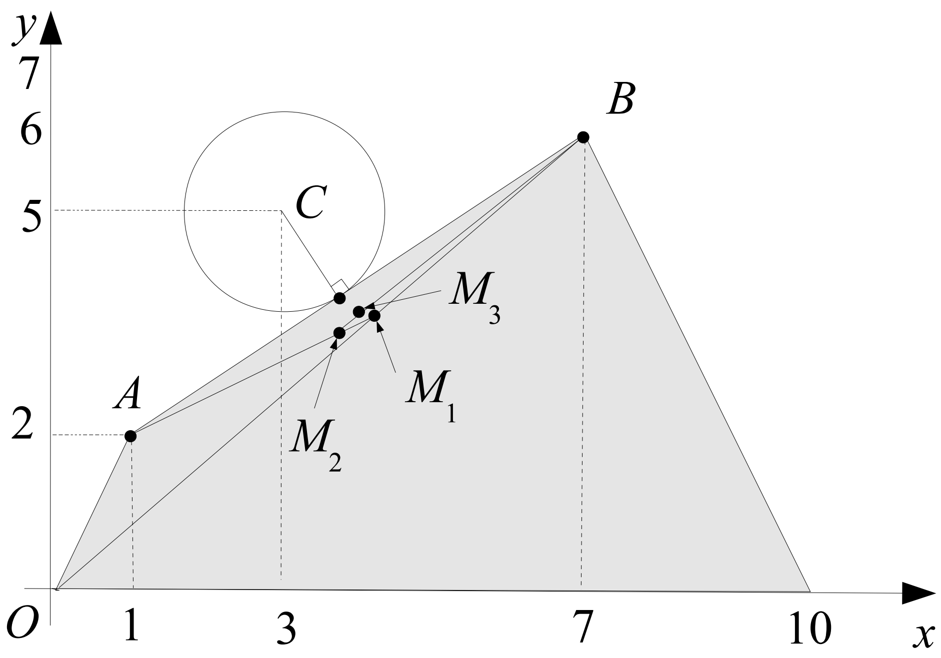

FIGURE IV.7 – Résolution par la méthode de Frank et Wolfe

On cherche le minimum de f sur le segment $[O, B]$. On paramètre ce segment par $x = 7t, y = 6t$ $(0 \leq t \leq 1)$ et on pose : $\phi_1(t) = f(7t, 6t)$, ce qui donne :

$$\begin{aligned}
\phi_1(t) &= (7t-3)^2 + (6t-5)^2 \\
&= 49t^2 - 42t + 9 + 36t^2 - 60t + 25 \\
&= 85t^2 - 102t + 34. \\
D'o : \phi_1'(t) &= 170t - 102.
\end{aligned}$$

Le minimum de ϕ_1 est obtenu pour $t_1 = \dfrac{3}{5} = 0,6$.

La fonction f atteint donc son minimum sur le segment $[O, B]$ au point : $M_1 = (x_1, y_1) = (7 \times 0,6\ ; 6 \times 0,6) = (4,2\ ; 3,6)$.

Étape 2 : à partir du point (4,2 ; 3,6).

On cherche le minimum de $l_2(x, y) = (x, y)^{\mathrm{t}}.\nabla f(4,2\,; 3,6)$ sur X.

La fonction $l_2(x, y) = 2,4x - 2,8y$ a son minimum sur X au point $(\tilde{x_1}, \tilde{y_1}) = A = (1, 2)$.

On cherche le minimum de f sur le segment $[M_1, A]$.

On paramètre ce segment par : $\left\{\begin{array}{l} x = t + 4,2(1-t) \\ y = 2t + 3,6(1-t) \end{array}\right. \quad (0 \leq t \leq 1)$

ou encore : $\left\{\begin{array}{l} x = -3,2t + 4,2 \\ y = -1,6t + 3,6 \end{array}\right. \quad (0 \leq t \leq 1).$

On pose : $\phi_2(t) = f(-3,2t + 4,2\,; -1,6t + 3,6)$, ce qui donne :

$$\begin{aligned} \phi_2(t) &= (-3,2t + 1,2)^2 + (-1,6t - 1,4)^2 \\ &= 10,24t^2 - 7,68t + 1,44 + 2,56t^2 + 4,48t + 1,96 \\ &= 12,8t^2 - 3,2t + 3,4. \\ D'o : \phi_2'(t) &= 25,6t - 3,2. \end{aligned}$$

Le minimum de ϕ_2 est obtenu pour $t = \dfrac{3,2}{25,6} = 0,125$.

La fonction f atteint donc son minimum sur le segment $[M_1, A]$ au point $M_2 = (x_2, y_2) = (-0,125 \times 3,2 + 4,2; -0,125 \times 1,6 + 3,6) = (3,8\,; 3,4)$.

Étape 3 : à partir du point (3,8 ; 3,4).

On cherche le minimum sur X de $l_3(x, y) = (x, y)^{\mathrm{t}}.\nabla f(3,8\,; 3,4)$ sur X.

La fonction $l_3(x, y) = 1,6x - 3,2y$ a son minimum au sur X point $(\tilde{x_2}, \tilde{y_2}) = B = (7, 6)$.

On cherche le minimum de f sur le segment M_2B.

On paramètre ce segment par : $\left\{\begin{array}{l} x = 3,8t + 7(1-t) \\ y = 3,4t + 6(1-t) \end{array}\right. \quad (0 \leq t \leq 1)$

ou encore : $\left\{\begin{array}{l} x = -3,2t + 7 \\ y = -2,6t + 6 \end{array}\right. \quad (0 \leq t \leq 1).$

On pose : $\phi_3(t) = f(-3,2t + 7\,; -2,6t + 6)$, ce qui donne :

$$\begin{aligned} \phi_3(t) &= (-3,2t + 4)^2 + (-2,6t + 1)^2 \\ &= 10,24t^2 - 25,6t + 16 + 6,76t^2 - 5,2t + 1 \\ &= 17t^2 - 30,8t + 17. \\ D'o : \phi_3'(t) &= 34t - 30,8. \end{aligned}$$

Le minimum de ϕ_3 est obtenu pour $t = \dfrac{30,8}{34} \approx 0,9059$.

La fonction f atteint donc son minimum sur le segment $[M_2, B]$ au point $M_3 = (x_3, y_3) = (-\dfrac{30,8}{34} \times 3,2 + 7, -\dfrac{30,8}{34} \times 2,6 + 6) \approx (4,101; 3,64)$.

On peut continuer ainsi. Le théorème IV.18 ci-dessous montrera que la suite des points (x_k, y_k) converge vers l'optimum du problème.

On aurait pu néanmoins choisir un autre point de départ. Par exemple, On recommençons l'algorithme en débutant au point A.

Étape 1. On cherche le minimum de $l_1(x, y) = (x, y)^{\mathrm{t}}.\nabla f(1; 2)$ sur X.

$l_1(x, y) = -4x - 6y$ et est minimum au point $B = (7, 6)$.

On cherche le minimum de f sur le segment $[A, B]$.

On paramètre ce segment par : $\left\{\begin{array}{l} x = t + 7(1-t) \\ y = 2t + 6(1-t) \end{array}\right. \quad (0 \leq t \leq 1)$

ou encore : $\left\{\begin{array}{l} x = -6t + 7 \\ y = -4t + 6 \end{array}\right. \quad (0 \leq t \leq 1).$

On pose : $\phi_1(t) = f(-6t + 7, -4t + 6)$.

$$\phi = 36t^2 - 48t + 16 + 16t^2 - 8t + 1$$

$$= 52t^2 - 56t + 17.$$

$$D'o : \phi_1'(t) = 104t - 56.$$

Le minimum de ϕ_1 est obtenu pour $t = \dfrac{56}{104} = \dfrac{7}{13} \approx 0,538$.

La fonction f atteint donc son minimum sur le segment $[A, B]$ au point $M_2' = (x_2', y_2') = (-\dfrac{7}{13} \times 6 + 7, -\dfrac{7}{13} \times 4 + 6) = (\dfrac{49}{13}, \dfrac{50}{13}) \approx (4,101; 3,64)$.

Étape 2 : à partir du point (1, 2).

On cherche le minimum de $l_2(x, y) = (x, y)^{\mathrm{t}}.\nabla f(\dfrac{49}{13}; \dfrac{50}{13})$ sur X.

$l_2(x, y) = \dfrac{20}{13}x - \dfrac{30}{13}y = \dfrac{10}{13}(2x - 3y)$ et est minimum au point $B = (7, 6)$.

La fonction l_2 est constante sur le segment A, B (elle vaut $-\dfrac{40}{3}$) et son minimum sur X est obtenu sur tout ce segment. On peut choisir un point de ce segment comme point où l_2 est minimum. On choisit le point M_2'.

On cherche le minimum de f sur le segment $[M_2', M_2']$, réduit à un point : la suite des points (M_k') est stationnaire (en M_2') et l'algorithme est terminé. On peut vérifier immédiatement que la condition de Karush, Kuhn et Tucker est vérifiée en M_2'. La condition étant ici suffisante, on a obtenu l'optimum.

Théorème IV.18. *Soit f une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, de classe C^1, strictement convexe. Soit X un polyèdre convexe compact de $\mathbb{R}^n$. La méthode de*

Frank et Wolfe appliqué au problème (P) *de la minimisation de* f *sur* X *converge vers le minimim global.*

On établit d'abord le lemme suivant.

Lemme IV.19. *Avec les hypothèses du théorème, soit* $(x^k)_{k\in\mathbb{N}}$ *une suite de points de* X *telle que la suite* $f(x^k)_{k\in\mathbb{N}}$ *soit décroissante. Supposons que la suite* $(x^k)_{k\in\mathbb{N}}$ *admette une sous-suite* $(x^k)_{k\in K\subset\mathbb{N}}$ *convergeant vers un minimum global de* f*. Alors la suite* $(x^k)_{k\in\mathbb{N}}$ *converge aussi vers le minimum global.*

Preuve du lemme.
Notons x^* la limite de la sous-suite $(x^k)_{k\in K}$; supposons que la suite $(x^k)_{k\in\mathbb{N}}$ ne converge pas vers x^*. Alors, il existe $\epsilon > 0$ tel que, pour tout $N \in \mathbb{N}$, il existe $k \geq N$ pour lequel $||x^k - x^*|| \geq \epsilon$; on en déduit qu'il existe une sous-suite infinie $(x^k)_{k\in U\subset\mathbb{N}}$ de la suite $(x^k)_{k\in\mathbb{N}}$ vérifiant, pour tout $k \in U, ||x^k - x^*|| \geq \epsilon$. Comme X est un compact, on peut extraire de la suite $(x^k)_{k\in U}$ une sous-suite $(x^k)_{k\in V\subset U}$ convergente ; soit y^* la limite de cette sous-suite.

La fonction f étant continue, les suites $(f(x^k))_{k\in K}$ et $(f(x^k))_{k\in V}$ sont convergentes respectivement vers $f(x^*)$ et $f(y^*)$. Le point x^* donnant le minimum global de f, on a : $f(y^*) \geq f(x*)$. Supposons $f(y^*) > f(x*)$. Il existe $k_0 \in K$ tel que $f(x^{k_0}) < f(y^*)$. Soit $k \in V$ vérifiant : $k \geq k_0$. La fonction f étant décroissante, $f(x^k) \leq f(x^{k_0})$, qui implique : $f(x^k) < f(y^*)$, ce qui contredit le fait que la suite $(f(x^k))_{k\in V}$ converge en décroissant vers $f(y^*)$.

On a donc : $f(y^*) = f(x^*)$, ce qui est impossible puisque la fonction f admet un unique minimum global sur X (cf. le théorème IV.14). ■

Preuve du théorème
Si la suite construite par la méthode devient stationnaire après un nombre fini d'étapes, on voit aisément que ce point stationnaire est le minimum global.

On suppose donc que ce n'est pas le cas et on note $(x^k)_{k\in\mathbb{N}}$ la suite construite par la méthode de Frank et Wolfe. Cette suite étant dans X qui est compact, elle admet une sous-suite convergente $(x^k)_{k\in K\subset\mathbb{N}}$; notons x^* la limite de cette suite. La suite $(\tilde{x}^k)_{k\in K}$ obtenue par la linéarisation prend chacune de ses valeurs dans un des sommets du polyèdre ; le nombre de sommets du polyèdre étant fini, on peut extraire de la suite $(x^k)_{k\in K}$ une sous-suite $(x^k)_{k\in H\subset K}$ telle que la suite $(\tilde{x}^k)_{k\in H}$ soit constante ; on note $\tilde{x}$ cette valeur constante.

La suite $(x^k)_{k \in H}$ converge vers x^* ; montrons que x^* donne l'optimum du problème (P).

Soient $t \in [0,1]$ et $k \in H$. La linéarisation en x^k prend son minimum en $\tilde{x}$; la méthode cherche alors le minimum de f sur le segment $[x^k, \tilde{x}^k] = [x^k, \tilde{x}]$ et obtient le point x^{k+1} ; on a donc : $f(x^k + t(\tilde{x} - x^k)) \geq f(x^{k+1})$.

Soit $h(k) \in H$ minimum parmi les indices $h \in H$ vérifiant $h \geq k+1$; la suite $f(x^k)_{k \in \mathbb{N}}$ étant décroissante par construction des x^k : $f(x^k + t(\tilde{x} - x^k)) \geq f(x^{h(k)+1})$.

En passant à la limite pour k dans H qui tend vers l'infini, on obtient :

$$f(x^* + t(\tilde{x} - x^*)) \geq f(x^*).$$

Écrivons la formule de Taylor de la fonction $t \rightarrow f(x^* + t(\tilde{x} - x^*))$ au voisinage de 0, à l'ordre 1 :

$$f(x^* + t(\tilde{x} - x^*)) = f(x^*) + t(\tilde{x} - x^*)^{\mathrm{t}}.\nabla f(x^*) + t\epsilon(t)$$

où $\epsilon(t)$ tend vers 0 quand t tend vers 0.

En utilisant l'inégalité obtenue plus haut, il vient :

$$t(\tilde{x} - x^*)^{\mathrm{t}}.\nabla f(x^*) + t\epsilon(t) \geq 0.$$

ou encore, pour $t > 0$: $(\tilde{x} - x^*)^{\mathrm{t}}.\nabla f(x^*) + \epsilon(t) \geq 0$.

En faisant tendre t vers 0, on a : $(\tilde{x} - x^*)^{\mathrm{t}}.\nabla f(x^*) \geq 0$, ce qui s'écrit :

$$\tilde{x}^{\mathrm{t}}.\nabla f(x^*) \geq (x^*)^{\mathrm{t}}.\nabla f(x^*). \qquad (1)$$

Soit $k \in H$ et $x \in X$. Par construction de $\tilde{x}^k = \tilde{x}$: $x^{\mathrm{t}}.\nabla f(x^k) \geq \tilde{x}^{\mathrm{t}}.\nabla f(x^k)$.

En passant à la limite quand $k \in H$ tend vers l'infini, on a :

$$x^{\mathrm{t}}.\nabla f(x^*) \geq \tilde{x}^{\mathrm{t}}.\nabla f(x^*). \qquad (2)$$

En utilisant les inégalités (1) et (2), on obtient :

$$x^{\mathrm{t}}.\nabla f(x^*) \geq (x^*)^{\mathrm{t}}.\nabla f(x^*),$$

ou encore : $(x - x^*)^{\mathrm{t}}.\nabla f(x^*) \geq 0$.

La fonction f étant convexe : $(x - x^*)^{\mathrm{t}}.\nabla f(x^*) \leq f(x) - f(x^*)$.

On a donc maintenant : $f(x) - f(x^*) \geq 0$, ce qui montre que x^* est solution optimale du problème (P).

Le lemme montre que $(x^k)_{k \in \mathbb{N}}$ converge aussi vers x^*. ■

IV.5.4. Linéarisation : méthode des plans sécants de Kelley

La méthode de Kelley s'applique à la recherche du minimum d'une fonction convexe sur un domaine convexe. La résolution du problème se fait par l'intermédiaire de problèmes d'optimisation linéaire.

Remarque. Considérons le problème :

$$\begin{array}{l} \text{Minimiser } f(x) \\ \text{avec, pour } 1 \leq i \leq m, g_i(x) \leq 0 \end{array}$$

où les fonctions f et $g_i, 1 \leq i \leq m$, sont convexes.

Ce problème peut se résoudre en ajoutant une variable réelle y sous la forme suivante :

$$\begin{array}{l} \text{Minimiser } y \\ \text{avec les contraintes : } \left\{ \begin{array}{l} \text{pour } 1 \leq i \leq m, g_i(x) \leq 0 \\ f(x) - y \leq 0. \end{array} \right. \end{array}$$

On ramène ainsi le problème à l'optimisation d'une fonction linéaire sur un domaine convexe.

En s'appuyant sur la remarque ci-dessus, la méthode des plans sécants de Kelley sera décrite comme la minimisation d'une fonction linéaire sur un domaine défini par des inégalités de type $g_i(x) \leq 0$ où les fonctions g_i sont convexes :

$$(P) \qquad \begin{array}{l} \text{Minimiser } f(x) \\ \text{avec, pour } 1 \leq i \leq m, g_i(x) \leq 0 \end{array}$$

où la fonction f est affine et les fonctions $g_i, 1 \leq i \leq m$, sont convexes.

On note X le domaine réalisable.

À chaque étape k, l'algorithme inclut X dans un polyèdre Q_k et cherche le minimum de f dans Q_k ; on résout ainsi un problème d'optimisation linéaire. En fait en sorte d'avoir toujours : $Q_{k+1} \subset Q_k$. L'étape k est décrite plus bas.

Initialisation de l'algorithme.

On choisit un point x^0 quelconque dans $\mathbb{R}^n$.

On considère le polyèdre Q_0 défini par :

$$\text{pour } 1 \leq i \leq m, g_i(x^0) + (x - x^0)^{\mathrm{t}}.\nabla g_i(x^0) \leq 0.$$

Soit $x \in X$ et i compris entre 1 et m. On a $g_i(x) \leq 0$ et, puisque la fonction g_i est convexe :

$$g_i(x) \geq g_i(x^0) + (x - x^0)^{\mathrm{t}}.\nabla g_i(x^0).$$

Ainsi : $g_i(x^0) + (x - x^0)^{\mathrm{t}}.\nabla g_i(x^0) \leq 0$. On a donc : $X \subset Q_0$.

On résout le problème de la minimisation de la fonction f sur le domaine Q_0. Il s'agit d'un problème d'optimisation linéaire que l'on sait résoudre avec l'algorithme du simplexe.

Si le problème d'optimisation linéaire n'est pas borné, l'initialisation rencontre une difficulté. On rend alors le domaine borné en choisissant une « grande » valeur M et en ajoutant les $2n$ inégalités $-M \leq x_i \leq M$ pour $1 \leq i \leq n$. Si, au cours de la méthode engendrant la suite x^k, un point de la suite sature une de ces contraintes ajoutées, on recommence à partir de l'initialisation avec une plus grande valeur de M, par exemple en doublant la précédente valeur de M. On peut remarquer que si on sait au départ que X est borné, on peut avoir intérêt à ajouter au départ de telles contraintes en assurant que X est strictement inclus dans le pavé défini par les contraintes ajoutées. On peut aussi, lorsque l'initialisation a conduit à un problème non borné, essayer un autre point de départ.

Supposons donc maintenant que le problème d'optimisation linéaire admet une solution, qu'on note x^1. On termine l'initialisation en posant $k = 1$.

Description de l'étape k ($k \geq 1$).

L'étape précédente a défini un nouveau point x^k qui donne l'optimum du problème sur le polyèdre Q_{k-1}.

Si x^k appartient à X, x^k donne la solution du problème puisque la relation $X \subset Q_{k-1}$ entraîne que le minimum de f sur X est au moins égal au minimum de f sur Q_{k-1}. L'algorithme se termine. Supposons donc que x^k n'appartient pas à X.

Il existe au moins une des m contraintes du problème qui n'est pas vérifiée. Parmi les contraintes non vérifiées, on en choisit une, g_{i_k}, qui maximise $g_i(x^k)$: $g_{i_k}(x^k) > 0$ et, pour $1 \leq i \leq m$, $g_i(x^k) \leq g_{i_k}(x^k)$.

On considère l'inégalité $g_{i_k}(x^k) + (x - x^k)^{\mathrm{t}}.\nabla g_{i_k}(x^k) \leq 0$ qui définit un demi-espace E_k. Comme pour l'initialisation, le domaine X appartient à E_k. En revanche, le point x^k n'appartient pas à E_k : le plan qui a pour équation $g_{i_k}(x^k) + (x - x^k)^{\mathrm{t}}.\nabla g_{i_k}(x^k) = 0$ sépare le point x^k du domaine X.

On considère le polyèdre $Q_k = Q_{k-1} \cap E_k$.

On résout le problème de la minimisation de la fonction f sur le domaine Q_k. Le problème est minoré puisque le minimum sur Q_k est minoré par le minimum sur Q_{k-1}. Soit x^{k+1} un point qui atteint le minimum. On termine l'étape en incrémentant k de 1.

Illustration.

On considère dans $\mathbb{R}^2$ le problème :

$$\begin{array}{l} \text{Minimiser } f(x,y) = y - x \\ \text{avec } \left\{ \begin{array}{l} x^2 + (y+3)^2 \leq 10 \\ x^2 + (y-2)^2 \leq 5. \end{array} \right. \end{array}$$

On pose : $g_1(x,y) = x^2 + (y+3)^2 - 10$ et $g_2(x,y) = x^2 + (y-2)^2 - 5$. Il vient :

$$\nabla g_1(x,y) = \begin{pmatrix} 2x \\ 2y+6 \end{pmatrix}, \ \nabla g_2(x,y) = \begin{pmatrix} 2x \\ 2y-4 \end{pmatrix}.$$

Le domaine est illustré sur la figure IV.8.

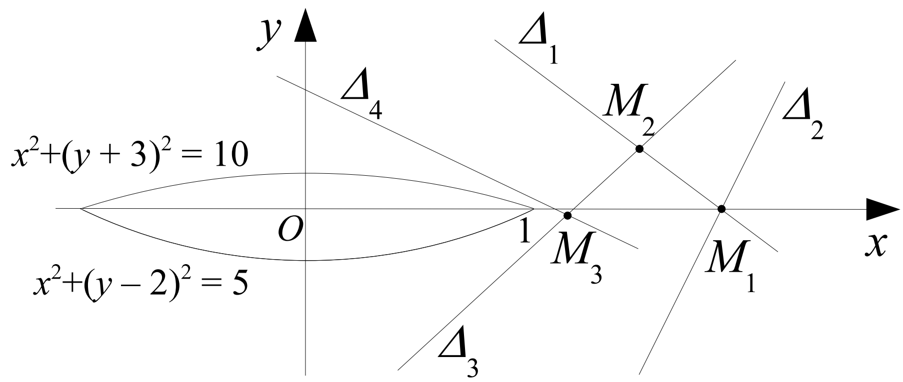

FIGURE IV.8 – Méthode des plans sécants de Kelley

Pour initialiser la méthode on essaie le point $O = (0, 0)$. On obtient :

$g_1(0,0) + ((x,y) - (0,0))^{\mathrm{t}}.\nabla g_1(0,0) = 6y - 1$
et $g_2(0,0) + ((x,y) - (0,0))^{\mathrm{t}}.\nabla g_2(0,0) = -4y - 1$.

Le domaine Q_1 est alors défini par : $y \leq 1/6$ et $y \geq 1/4$ et la fonction f n'est pas minorée dans ce domaine.

On essaie un autre point de départ, le point $(x_0, y_0) = (3,1)$.

Initialisation

$g_1(3,1) + ((x,y) - (3,1))^{\mathrm{t}}.\nabla g_1(3,1) = 6x + 8y - 11$
$g_2(3,1) + ((x,y) - (3,1))^{\mathrm{t}}.\nabla g_2(3,1) = 6x - 2y - 11$.

On note Δ_1 la droite d'équation $6x + 8y - 11$ et Δ_2 la droite d'équation $6x - 2y - 11$.

On résout alors le problème :

$$\begin{array}{l} \text{Minimiser } f(x,y) = y - x \\ \text{avec } \left\{ \begin{array}{l} 6x + 8y \leq 11 \\ 6x - 2y \leq 11. \end{array} \right. \end{array}$$

La solution est $(x_1, y_1) = (\dfrac{11}{6}, 0)$. Le point (x_1, y_1) est noté M_1 sur la figure IV.8.

Étape 1

On a : $g_1(x_1, y_1) = g_2(x_1, y_1) = \dfrac{85}{36}$.

Les deux contraintes sont violées au point (x_1, y_1) et y prennent la même valeur. On peut choisir l'une ou l'autre comme étant la contrainte la plus violée. On choisit g_2.

$$\begin{aligned} g_2(\frac{11}{6}, 0) + ((x - \frac{11}{6}, y))^{\mathrm{t}}.\nabla g_2(\frac{11}{6}, 0) &= \frac{85}{36} + \frac{11}{3}(x - \frac{11}{6}) - 4y \\ &= \frac{11}{3}x - 4y - \frac{157}{36}. \end{aligned}$$

On note Δ_3 la droite d'équation $\dfrac{11}{3}x - 4y = \dfrac{157}{36}$.

On résout alors le problème :

$$\begin{array}{l} \text{Minimiser } f(x,y) = y - x \\ \text{avec } \left\{ \begin{array}{l} 6x + 8y \leq 11 \\ 6x - 2y \leq 11 \\ \dfrac{11}{3}x - 4y \leq \dfrac{157}{36}. \end{array} \right. \end{array}$$

La solution est à l'intersection des droites Δ_1 et Δ_3, c'est-à-dire le point $(x_2, y_2) = (\frac{71}{48}, \frac{17}{64}) \approx (1,4791; 0,2656)$. Le point (x_2, y_2) est noté M_2 sur la figure IV.8.

Étape 2

On a : $g_1(x_2, y_2) \approx 2,8522$ et $g_2(x_2, y_2) \approx 0,1962$.

On continue avec les valeurs approchées pour simplifier les calculs.

Les deux contraintes sont violées au point (x_2, y_2) et la contrainte la plus violée est la contrainte g_1, qu'on retient donc.

$g_1(x_2, y_2) + (x - x_2, y - y_2)^{\mathrm{t}}.\nabla g_1(x_2, y_2) = 2,9582x + 6,5312y - 3,2580.$

On note Δ_4 la droite d'équation $2,9582x + 6,5312y = 3,2582$.

On résout alors le problème :

$$\text{Minimiser } f(x, y) = y - x$$
$$\text{avec } \begin{cases} 6x + 8y \leq 11 \\ 6x - 2y \leq 11 \\ 3,6667x - 4y \leq 4,3611 \\ 2,9582x + 6,5312y \leq 3,2580 \end{cases}$$

La solution est à l'intersection des droites Δ_3 et Δ_4, c'est-à-dire le point $(x_3, y_3) \approx (1,1603; -0,0267)$. Le point (x_3, y_3) est noté M_3 sur la figure IV.8.

On arrête ici l'algorithme. L'optimum du problème est au point (1, 0), ce qu'on peut vérifier avec la condition de Karush, Kuhn et Tucker, qui est ici suffisante. On constate donc que la suite de points s'approche du point qui donne l'optimum du problème.

Les points engendrés par la méthode des plans sécants de Kelley fournissent des minorants du problème alors que la méthode de Frank et Wolfe en donnent des majorants.

Théorème IV.20. *Si le problème (P) admet un minimum à distance finie, alors tout point d'accumulation de la suite (x^k) engendrée par la méthode des plans sécants de Kelley est une solution du problème (P).*

Preuve.

On se place dans le cas où la méthode ne trouve pas la solution après un nombre fini d'étapes.

Soit $\tilde{x}$ un point d'accumulation de la suite $(x^k)_{k\in\mathbb{N}}$. On peut extraire de cette suite une sous-suite $(x^k)_{k\in K\subset\mathbb{N}}$ convergeant vers $\tilde{x}$.

Soit x^* un point de X atteignant le minimum.

Pour tout $k \in K$, on a $f(x^k) \leq f(x^*)$ puisque $X \subset Q_k$. En passant à la limite quand $k \in K$ tend vers l'infini, on obtient : $f(\tilde{x}) \leq f(x^*)$. Nous allons montrer que $\tilde{x} \in X$.

On note i_k l'indice de la contrainte retenue à l'étape k, c'est-à-dire une contrainte la plus violée dans cette étape. La suite $(i_k)_{k \in K}$ prend toutes ses valeurs dans l'ensemble fini $\{1, 2, ..., m\}$. On peut donc en extraire une sous-suite $(i_h)_{h \in H \subset K}$ constante i_0.

Soit $h \in H$. On note $r(h)$ le plus petit indice r vérifiant : $r \in H, r > h$. Le point $x^{r(h)}$ est dans le demi-espace E_h :

$g_{i_0}(x^h) + (x^{r(h)} - x^h)^{\mathrm{t}}.\nabla g_{i_0}(x^h) \leq 0$

$0 < g_{i_0}(x^h) \leq -(x^{r(h)} - x^h)^{\mathrm{t}}.\nabla g_{i_0}(x^h)$

$0 < g_{i_0}(x^h) \leq ||(x^{r(h)} - x^h||.||\nabla g_{i_0}(x^h)||$.

On fait tendre $h \in H$ vers l'infini ; $r(h)$ tend aussi vers l'infini.

De plus, $x^h \to \tilde{x}$ et $x^{r(h)} \to \tilde{x}$. D'où : $||x^{r(h)} - x^h|| \to 0$.

Par ailleurs $\nabla g_{i_0}(x^h) \to \nabla g_{i_0}(\tilde{x})$, ce qui implique que $||\nabla g_{i_0}(x^h)||$ est borné. Par conséquent :

$||x^{r(h)} - x^h||.||\nabla g_{i_0}(x^h)|| \to 0$, doù $g_{i_0}(x^h) \to 0$.

Comme par ailleurs $g_{i_0}(x^h) \to g_{i_0}(\tilde{x})$, on a : $g_{i_0}(\tilde{x}) = 0$. Le point $\tilde{x}$ vérifie la contrainte i_0.

Soit maintenant $i \in \{1, 2, ..., m\}$. La contrainte g_{i_0} étant la plus violée, pour $h \in H, g_i(x^h) \leq g_{i_0}(x^h)$. En passant à la limite quand $h \in H$ vers l'infini, on obtient : $g_i(\tilde{x}) \leq 0$. Le point $\tilde{x}$ vérifie toutes les contraintes, il appartient à X. Comme $f(\tilde{x}) \leq f(x^*)$, $\tilde{x}$ donne le minimum de f sur X. ■

IV.6. Exercices

Exercice 1

On s'intéresse au problème d'optimisation sur $\mathbb{R}^2$ suivant :

$$\text{Minimiser } 2x^2 + y^4$$
$$\text{avec les contraintes} \begin{cases} x \geq 1 \\ x + ay \geq a + 1 \end{cases}$$

où a est un paramètre réel.

1. Pour quelles valeurs de a peut-on affirmer que le minimum est atteint au point (1, 1) ? Pour ces valeurs, le minimum est-il aussi atteint en un autre point ?

2. Résoudre le problème suivant à l'aide de la méthode de descente en partant du point (1, 1) :

$$\text{Minimiser } 2x^2 + y^4$$
$$\text{avec les contraintes} \begin{cases} x \geq 1 \\ 2x + y \geq 3. \end{cases}$$

Exercice 2

On considère le problème (P_α) suivant :

$$\text{Minimiser } f(x, y) = (x-2)^2 + \alpha(y-1)^2$$
$$\text{avec les contraintes} \begin{cases} x^2 + (y+1)^2 \leq 2 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

où α est un paramètre réel positif ou nul.

Indiquer pour quelles valeurs de α le minimum de (P_α) est atteint au point (1, 0).

Chapitre V

Corrigés des exercices

V.1. Corrigé des exercices du chapitre 1

Corrigé de l'exercice I.7.1.

Introduisons les variables d'écart du problème. On obtient comme premier dictionnaire :

$$\begin{array}{rcrcrcrcr} x_4 & = & 4 & - & x_1 & - & x_2 & - & 2x_3 \\ x_5 & = & 5 & - & 2x_1 & & & - & 3x_3 \\ x_6 & = & 7 & - & 2x_1 & - & x_2 & - & 3x_3 \\ \hline z & = & & & 3x_1 & + & 2x_2 & + & 4x_3 \end{array}$$

Chacune des trois variables hors-base étant candidate à entrer en base, cherchons celle dont la croissance à partir de 0 permet d'augmenter le plus la valeur de la fonction objectif, actuellement égale à 0. Si x_1 entre en base, comme son augmentation est bornée par $5/2$, la fonction objectif augmente de $15/2$. Si x_2 entre en base, la fonction objectif augmente de 8, enfin si c'est x_3, l'objectif augmente de $20/3$. On choisit donc de faire entrer x_2. La variable en base x_4, qui, parce qu'elle doit rester positive, est la variable en base qui a le plus contraint l'accroissement de x_2, quitte la base. On obtient le nouveau dictionnaire :

$$\begin{array}{rcrcrcrcr} x_2 & = & 4 & - & x_1 & - & 2x_3 & - & x_4 \\ x_5 & = & 5 & - & 2x_1 & - & 3x_3 & & \\ x_6 & = & 3 & - & x_1 & - & x_3 & + & x_4 \\ \hline z & = & 8 & + & x_1 & & & - & 2x_4 \end{array}$$

Cette fois, nous n'avons plus le choix de la variable entrante, puisque seule x_1 a un coefficient positif dans z, et x_5 quitte la base. Le nouveau dictionnaire est le suivant :

$$\begin{array}{rcrcrcrcr}
x_1 & = & \frac{5}{2} & - & \frac{3}{2}x_3 & & & - & \frac{1}{2}x_5 \\
x_2 & = & \frac{3}{2} & - & \frac{1}{2}x_3 & - & x_4 & + & \frac{1}{2}x_5 \\
x_6 & = & \frac{1}{2} & + & \frac{1}{2}x_3 & + & x_4 & + & \frac{1}{2}x_5 \\
\hline
z & = & \frac{21}{2} & - & \frac{3}{2}x_3 & - & 2x_4 & - & \frac{1}{2}x_5
\end{array}$$

Ce dictionnaire est le dernier puisque maintenant il n'existe plus de variable hors-base dont le coefficient dans z soit positif. Le maximum cherché pour z est donc de $21/2$, et il est obtenu pour les valeurs suivantes des variables :

$$x_1 = \frac{5}{2}\,;\; x_2 = \frac{3}{2}\,;\; x_3 = 0.$$

Corrigé de l'exercice I.7.2.

Introduisons les variables d'écart du problème. On obtient comme premier dictionnaire :

$$\begin{array}{rcrcrcrcrcr}
x_5 & = & 5 & - & x_1 & - & 2x_2 & - & 3x_3 & - & x_4 \\
x_6 & = & 3 & - & x_1 & - & x_2 & - & 2x_3 & - & 3x_4 \\
\hline
z & = & & & 5x_1 & + & 6x_2 & + & 9x_3 & + & 8x_4
\end{array}$$

1. D'après le critère retenu ici pour faire entrer une variable en base, c'est tout d'abord la variable x_3 qui entre en base. La variable sortante est x_6. Le nouveau dictionnaire est le suivant :

$$\begin{array}{rcrcrcrcrcr}
x_3 & = & 1,5 & - & 0,5x_1 & - & 0,5x_2 & - & 1,5x_4 & - & 0,5x_6 \\
x_5 & = & 0,5 & + & 0,5x_1 & - & 0,5x_2 & + & 3,5x_4 & + & 1,5x_6 \\
\hline
z & = & 13,5 & + & 0,5x_1 & + & 1,5x_2 & - & 5,5x_4 & - & 4,5x_6
\end{array}$$

Si on choisit encore la variable entrante de plus grand coefficient, il s'agit de x_2. La variable sortante est alors x_5. On obtient le dictionnaire ci-dessous :

$$\begin{array}{rcrcrcrcrcr} x_2 & = & 1 & + & x_1 & + & 7x_4 & - & 2x_5 & + & 3x_6 \\ x_3 & = & 1 & - & x_1 & - & 5x_4 & + & x_5 & - & 2x_6 \\ \hline z & = & 15 & + & 2x_1 & + & 5x_4 & - & 3x_5 & & \end{array}$$

La variable x_4 entre maintenant en base et la variable x_3 en sort. D'où :

$$\begin{array}{rcrcrcrcrcr} x_4 & = & 0,2 & - & 0,2x_1 & - & 0,2x_3 & + & 0,2x_5 & - & 0,4x_6 \\ x_2 & = & 2,4 & - & 0,4x_1 & - & 1,4x_3 & - & 0,6x_5 & + & 0,2x_6 \\ \hline z & = & 16 & + & x_1 & - & x_3 & - & 2x_5 & - & 2x_6 \end{array}$$

Enfin, la variable x_1 entre en base et x_4 en sort. Le dernier dictionnaire est :

$$\begin{array}{rcrcrcrcrcr} x_1 & = & 1 & - & x_3 & - & 5x_4 & + & x_5 & - & 2x_6 \\ x_2 & = & 2 & - & x_3 & + & 2x_4 & - & x_5 & + & x_6 \\ \hline z & = & 17 & - & 2x_3 & - & 5x_4 & - & x_5 & - & 4x_6 \end{array}$$

2. Envisageons maintenant, à l'aide du tableau ci-dessous, les quatre possibilités pour le choix de la variable entrante :

variable entrante	x_1	x_2	x_3	x_4
accroissement maximum de la variable	3	5/2	3/2	1
accroissement correspondant de z	15	15	27/2	8

Le critère actuel conduit à choisir x_1 ou x_2. Faisons par exemple entrer x_1 ; c'est alors la variable x_6 qui sort ; le nouveau dictionnaire est :

$$\begin{array}{rcrcrcrcrcr} x_1 & = & 3 & - & x_2 & - & 2x_3 & - & 3x_4 & - & x_6 \\ x_5 & = & 2 & - & x_2 & - & x_3 & + & 2x_4 & + & x_6 \\ \hline z & = & 15 & + & x_2 & - & x_3 & - & 7x_4 & - & 5x_6 \end{array}$$

Seule la variable x_2 est candidate à entrer en base, la variable x_5 sort ; on obtient le dictionnaire suivant :

$$\begin{array}{rcrcrcrcrcr} x_2 & = & 2 & - & x_3 & + & 2x_4 & - & x_5 & + & x_6 \\ x_1 & = & 1 & - & x_3 & - & 5x_4 & + & x_5 & - & 2x_6 \\ \hline z & = & 17 & - & 2x_3 & - & 5x_4 & - & x_5 & - & 4x_6 \end{array}$$

Tous les coefficients de z sont négatifs ou nuls : la base constituée de $\{x_1, x_2\}$ est donc optimale, avec $x_1^* = 1$ et $x_2^* = 2$.

Nous remarquons que, avec la première stratégie sur le choix de la variable entrante, le nombre d'étapes est de quatre alors qu'avec la seconde stratégie, ce nombre est de deux. Sur ce cas particulier, la seconde stratégie est plus avantageuse.

Corrigé de l'exercice I.7.3.

Considérons le problème :

$$\text{Maximiser } z = x_1 + x_2$$
$$\text{avec les contraintes : } \begin{cases} x_1 - x_2 \leq 1 \\ 2x_1 - 3x_2 \leq 2 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Toute solution de la forme $x_1 = 0, x_2 = t \geq 0$ est réalisable et pour ces valeurs on a $z = t$. Comme t n'est pas borné, le problème n'est pas borné.

Corrigé de l'exercice I.7.4.

Considérons le problème :

$$\text{Maximiser } z = x_1 + x_2$$
$$\text{avec les contraintes : } \begin{cases} -2x_1 + 3x_2 \leq -4 \\ x_1 - x_2 \leq 1 \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

Supposons la seconde contrainte satisfaite, on a alors $-x_1 + x_2 \geq -1$, d'où : $-3x_1 + 3x_2 \geq -3$.

Puisque x_1 est positif, on obtient : $-2x_1 + 3x_2 \geq -3x_1 + 3x_2 \geq -3$, contradiction avec la première contrainte. Le problème est donc infaisable.

Corrigé de l'exercice I.7.5.

1. On part du dictionnaire donné.

$$\begin{array}{llrrrrrrr}
x_5 & = & - & 0,5x_1 & + & 5,5x_2 & + & 2,5x_3 & - & 9x_4 \\
x_6 & = & - & 0,5x_1 & + & 1,5x_2 & + & 0,5x_3 & - & x_4 \\
x_7 & = 1 & - & x_1 & & & & & & \\
\hline
z & = & & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4
\end{array}$$

On fait entrer x_1 et sortir x_5. Après la première itération :

$$\begin{array}{lllrcrcrcr}
x_1 & = & & 11x_2 & + & 5x_3 & - & 18x_4 & - & 2x_5 \\
x_6 & = & - & 4x_2 & - & 2x_3 & + & 8x_4 & + & x_5 \\
x_7 & =1 & - & 11x_2 & - & 5x_3 & + & 18x_4 & + & 2x_5 \\
\hline
z & = & & 53x_2 & + & 41x_3 & - & 204x_4 & - & 20x_5
\end{array}$$

On fait entrer x_2 et sortir x_6. Après la deuxième itération :

$$\begin{array}{lllrcrcrcr}
x_2 & = & - & 0,5x_3 & + & 2x_4 & + & 0,25x_5 & - & 0,25x_6 \\
x_1 & = & - & 0,5x_3 & + & 4x_4 & + & 0,75x_5 & - & 2,75x_6 \\
x_7 & =1 & + & 0,5x_3 & - & 4x_4 & - & 0,75x_5 & - & 2,75x_6 \\
\hline
z & = & & 14,5x_3 & - & 98x_4 & - & 6,75x_5 & - & 13,25x_6
\end{array}$$

On fait entrer x_3 et sortir x_1. Après la troisième itération :

$$\begin{array}{lllrcrcrcr}
x_3 & = & - & 2x_1 & + & 8x_4 & + & 1,5x_5 & - & 5,5x_6 \\
x_2 & = & & x_1 & - & 2x_4 & - & 0,5x_5 & + & 2,5x_6 \\
x_7 & =1 & - & x_1 & & & & & & \\
\hline
z & = & - & 29x_1 & + & 18x_4 & + & 15x_5 & - & 93x_6
\end{array}$$

On fait entrer x_4 et sortir x_2. Après la quatrième itération :

$$\begin{array}{lllrcrcrcr}
x_4 & = & & 0,5x_1 & - & 0,5x_2 & - & 0,25x_5 & + & 1,25x_6 \\
x_3 & = & & 2x_1 & - & 4x_2 & - & 0,5x_5 & + & 4,5x_6 \\
x_7 & =1 & - & x_1 & & & & & & \\
\hline
z & = & - & 20x_1 & - & 9x_2 & + & 10,5x_5 & - & 70,5x_6
\end{array}$$

On fait entrer x_5 et sortir x_3. Après la cinquième itération :

$$\begin{array}{lllrcrcrcr}
x_5 & = & & 4x_1 & - & 8x_2 & - & 2x_3 & + & 9x_6 \\
x_4 & = & - & 0,5x_1 & + & 1,5x_2 & + & 0,5x_3 & - & x_6 \\
x_7 & =1 & - & x_1 & & & & & & \\
\hline
z & = & & 22x_1 & - & 93x_2 & - & 21x_3 & + & 24x_6
\end{array}$$

On fait entrer x_6 et sortir x_4. Après la sixième itération :

$$\begin{array}{lllrcrcrcr}
x_5 & = & - & 0,5x_1 & + & 5,5x_2 & + & 2,5x_3 & - & 9x_4 \\
x_6 & = & - & 0,5x_1 & + & 1,5x_2 & + & 0,5x_3 & - & x_4 \\
x_7 & =1 & - & x_1 & & & & & & \\
\hline
z & = & & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4
\end{array}$$

On retrouve le dictionnaire de départ : on observe qu'il y a cyclage.

2. La règle des indices donne les mêmes cinq premières itérations, mais pas la sixième. On reprend les calculs précédents après la cinquième itération :

$$\begin{array}{rcl}
x_5 &=& 4x_1 - 8x_2 - 2x_3 + 9x_6 \\
x_4 &=& -0,5x_1 + 1,5x_2 + 0,5x_3 - x_6 \\
x_7 &=& 1 - x_1 \\
\hline
z &=& 22x_1 - 93x_2 - 21x_3 + 24x_6
\end{array}$$

On fait entrer x_1 (et non plus x_6) et sortir x_4. Après la sixième itération :

$$\begin{array}{rcl}
x_1 &=& 3x_2 + x_3 - 2x_4 - 2x_6 \\
x_5 &=& 4x_2 + 2x_3 - 8x_4 + x_6 \\
x_7 &=& 1 - 3x_2 - x_3 + 2x_4 + 2x_6 \\
\hline
z &=& -27x_2 + x_3 - 44x_4 - 20x_6
\end{array}$$

On fait entrer x_3 et sortir x_7. Après la septième itération :

$$\begin{array}{rcl}
x_3 &=& 1 - 3x_2 + 2x_4 + 2x_6 - x_7 \\
x_1 &=& 1 - x_7 \\
x_5 &=& 2 - 2x_2 - 4x_4 + 5x_6 - 2x_7 \\
\hline
z &=& 1 - 30x_2 - 42x_4 - 18x_6 - 2x_7
\end{array}$$

Tous les coefficients dans z sont négatifs ou nuls, la méthode s'arrête. On constate que l'application de la règle de Bland a permis d'éviter le cyclage.

Corrigé de l'exercice I.7.6.

Les deux problèmes d'optimisation de cet exercice sont mis sous forme standard. On s'aperçoit que, dans les deux cas, la solution obtenue en mettant à zéro les deux variables x_1 et x_2 n'est pas réalisable. Il faut donc utiliser la méthode du simplexe à deux phases. La première phase débute par l'écriture du problème auxiliaire. Pour cela, on peut soustraire une variable x_0 dans les trois premiers membres des inégalités, comme indiqué dans le cours ; on peut aussi se contenter de soustraire cette variable x_0 dans les premiers membres des inégalités ayant un second membre négatif. C'est cette méthode que nous choisissons ici.

1. Le problème auxiliaire s'écrit :

$$\text{Maximiser } w = -x_0$$
$$\text{avec les contraintes : } \begin{cases} -4x_1 + 5x_2 - x_0 \leq -10 \\ 5x_1 + 2x_2 \leq 10 \\ 3x_1 + 8x_2 \leq 12 \\ x_1 \geq 0, x_2 \geq 0, x_0 \geq 0. \end{cases}$$

On en déduit le dictionnaire initial :

$$\begin{array}{rcrcrcrcr} x_3 & = & -10 & + & x_0 & + & 4x_1 & - & 5x_2 \\ x_4 & = & 10 & & & - & 5x_1 & - & 2x_2 \\ x_5 & = & 12 & & & - & 3x_1 & - & 8x_2 \\ \hline w & = & & - & x_0 & & & & \end{array}$$

Ce dictionnaire n'est pas réalisable, mais on passe immédiatement à un dictionnaire réalisable en faisant entrer la variable x_0 et en faisant sortir la variable x_3. On obtient le dictionnaire ci-dessous :

$$\begin{array}{rcrcrcrcr} x_0 & = & 10 & - & 4x_1 & + & 5x_2 & + & x_3 \\ x_4 & = & 10 & - & 5x_1 & - & 2x_2 & & \\ x_5 & = & 12 & - & 3x_1 & - & 8x_2 & & \\ \hline w & = & -10 & + & 4x_1 & - & 5x_2 & - & x_3 \end{array}$$

On fait maintenant entrer la variable x_1 et sortir la variable x_4 ; on obtient :

$$\begin{array}{rcrcrcrcr} x_1 & = & 2 & - & \frac{2}{5}x_2 & & & - & \frac{1}{5}x_4 \\ x_0 & = & 2 & + & \frac{33}{5}x_2 & + & x_3 & + & \frac{4}{5}x_4 \\ x_5 & = & 6 & - & \frac{34}{5}x_2 & & & + & \frac{3}{5}x_4 \\ \hline w & = & -2 & - & \frac{33}{5}x_2 & - & x_3 & - & \frac{4}{5}x_3 \end{array}$$

Il n'y a plus de variable entrante ; l'optimum de w est -2 et n'est donc pas nul : le problème étudié n'admet pas de solution réalisable.

2. De la même façon que pour la question précédente, le problème auxiliaire s'écrit :

$$\text{Maximiser } w = -x_0$$
$$\text{avec les contraintes : } \begin{cases} -4x_1 - 5x_2 - x_0 \leq -10 \\ 5x_1 + 2x_2 \leq 10 \\ 3x_1 + 8x_2 \leq 12 \\ x_1 \geq 0, x_2 \geq 0, x_0 \geq 0. \end{cases}$$

On obtient le dictionnaire initial suivant :

$$\begin{array}{rcrcrcrcr} x_3 & = & -10 & + & x_0 & + & 4x_1 & + & 5x_2 \\ x_4 & = & 10 & & & - & 5x_1 & - & 2x_2 \\ x_5 & = & 12 & & & - & 3x_1 & - & 8x_2 \\ \hline w & = & & - & x_0 & & & & \end{array}$$

Ce dictionnaire n'est pas réalisable mais, ici encore, on passe immédiatement à un dictionnaire réalisable en faisant entrer la variable x_0 et en faisant sortir la variable x_3. On obtient le dictionnaire ci-dessous :

$$\begin{array}{rcrcrcrcr} x_0 & = & 10 & - & 4x_1 & - & 5x_2 & + & x_3 \\ x_4 & = & 10 & - & 5x_1 & - & 2x_2 & & \\ x_5 & = & 12 & - & 3x_1 & - & 8x_2 & & \\ \hline w & = & -10 & + & 4x_1 & + & 5x_2 & - & x_3 \end{array}$$

On fait maintenant entrer la variable x_1 et sortir la variable x_4 ; on obtient :

$$\begin{array}{rcrcrcrcr} x_1 & = & 2 & - & \frac{2}{5}x_2 & & & - & \frac{1}{5}x_4 \\ x_0 & = & 2 & - & \frac{17}{5}x_2 & + & x_3 & + & \frac{4}{5}x_4 \\ x_5 & = & 6 & - & \frac{34}{5}x_2 & & & + & \frac{3}{5}x_4 \\ \hline w & = & -2 & + & \frac{17}{5}x_2 & - & x_3 & - & \frac{4}{5}x_4 \end{array}$$

La variable x_2 est ici entrante alors que la variable x_0 sort. Le dictionnaire obtenu est :

$$\begin{array}{rcrcrcrcr}
x_2 & = & \frac{10}{17} & + & \frac{5}{17}x_3 & + & \frac{4}{17}x_4 & - & \frac{5}{17}x_0 \\
x_1 & = & \frac{30}{17} & - & \frac{2}{17}x_3 & - & \frac{5}{17}x_4 & + & \frac{2}{17}x_0 \\
x_5 & = & 2 & - & 2x_3 & - & x_4 & + & 2x_0 \\
\hline
w & = & & & & & & - & x_0
\end{array}$$

L'optimum du problème auxiliaire vaut zéro : le problème initial est réalisable. On peut maintenant commencer la seconde phase de la méthode. Pour obtenir un dictionnaire réalisable du problème initial, on reprend le dernier dictionnaire ci-dessus, dans lequel on supprime la variable x_0 et on remplace la fonction w par la fonction z exprimée en fonction des variables hors-base, c'est-à-dire en fonction de x_3 et x_4. On obtient le dictionnaire ci-dessous :

$$\begin{array}{rcrcrcr}
x_2 & = & \frac{10}{17} & + & \frac{5}{17}x_3 & + & \frac{4}{17}x_4 \\
x_1 & = & \frac{30}{17} & - & \frac{2}{17}x_3 & - & \frac{5}{17}x_4 \\
x_5 & = & 2 & - & 2x_3 & - & x_4 \\
\hline
z & = & \frac{180}{17} & + & \frac{5}{17}x_3 & - & \frac{13}{17}x_4
\end{array}$$

La variable x_3 entre en base alors que la variable x_5 en sort. Le dictionnaire devient :

$$\begin{array}{rcrcrcr}
x_3 & = & 1 & - & \frac{1}{2}x_4 & - & \frac{1}{2}x_5 \\
x_2 & = & \frac{15}{17} & + & \frac{3}{34}x_4 & - & \frac{5}{34}x_5 \\
x_1 & = & \frac{28}{17} & - & \frac{4}{17}x_4 & + & \frac{1}{17}x_5 \\
\hline
z & = & \frac{185}{17} & - & \frac{31}{34}x_4 & - & \frac{5}{34}x_5
\end{array}$$

Ce dernier dictionnaire est optimal ; la solution optimale est donc donnée par :

$x_1^* = \frac{28}{17}, x_2^* = \frac{15}{17}$ pour les variables de décision ;

$x_3^* = 1, x_4^* = x_5^* = 0$ pour les variables d'écart ;

$z^* = \frac{185}{17}$ pour la fonction objectif.

V.2. Corrigé des exercices du chapitre 2

Corrigé de l'exercice II.6.1.

1. On représente graphiquement le problème sur la figure V.1.

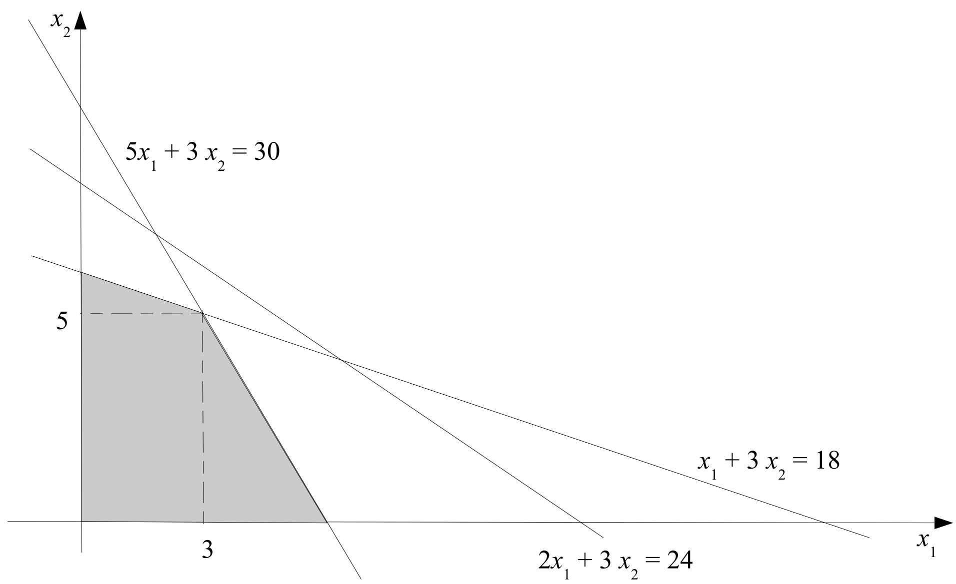

FIGURE V.1 – Solution graphique

La solution graphique semble être : $x_1^* = 3, x_2^* = 5$.

2. On vérifie la solution graphique en utilisant le théorème des écarts complémentaires.

   La solution $x_1^* = 3, x_2^* = 5$ est bien une solution réalisable.

   On cherche y_1^*, y_2^* et y_3^* vérifiant les conditions du théorème des écarts complémentaires.

   Avec $x_1^* = 3$ et $x_2^* = 5$, on a : $2x_1^* + 3x_2^* = 21 < 24$, ce qui entraîne : $y_2^* = 0$.

   Comme x_1^* et x_2^* sont non nuls, on doit avoir :

$$\begin{cases} 5y_1^* + y_3^* = 4 \\ 3y_1^* + 3y_3^* = 3 \end{cases}$$

   qui a pour unique solution : $y_1^* = 3/4, y_3^* = 1/4$.

   Il reste à vérifier que les valeurs $y_1^* = 3/4, y_2^* = 0, y_3^* = 1/4$ constituent une solution réalisable du problème dual, ce qui est immédiat. La solution déterminée graphiquement est bien optimale.

3. La valeur marginale de la première ressource est égale à 3/4. Se procurer une unité de plus de la première ressource est intéressant si le prix unitaire de celle-ci est inférieure à 0,75 euros.

4. La solution basique associée à la base (x_1, x_2, x_4) vérifie (puisque x_3 et x_5 sont nulles) :

$$\begin{cases} 5x_1 + 3x_2 = 30 + a \\ 2x_1 + 3x_2 + x_4 = 24 \\ x_1 + 3x_2 = 18. \end{cases}$$

   En s'appuyant sur la solution numérique du problème initial, posons $x_1 = 3 + \delta x_1, x_2 = 5 + \delta x_2, x_4 = 3 + \delta x_4$. On a alors :

$$\begin{cases} 5\delta x_1 + 3\delta x_2 = a \\ 2\delta x_1 + 3\delta x_2 + \delta x_4 = 0 \\ \delta x_1 + 3\delta x_2 = 0. \end{cases}$$

   Ce système a pour solution : $\delta x_1 = \dfrac{a}{4}, \delta x_2 = -\dfrac{a}{12}, \delta x_4 = -\dfrac{a}{4}$.

La solution est réalisable si :

$$\begin{cases} 3 + \dfrac{a}{4} \geq 0 \\ 5 - \dfrac{a}{12} \geq 0 \\ 3 - \dfrac{a}{4} \geq 0 \end{cases}$$

ce qui équivaut à : $a \leq 12$.

Corrigé de l'exercice II.6.2.

La vérification se fait comme suit. On examine d'abord si la solution proposée est réalisable.

- La solution proposée est positive ou nulle.
- Vérifions qu'elle satisfait les autres contraintes et repérons simultanément les contraintes saturées et celles qui ne le sont pas.
  $x_1^* + 3x_2^* + 5x_3^* - 2x_4^* + 2x_5^* = 4$: contrainte saturée.
  $4x_1^* + 2x_2^* - 2x_3^* + x_4^* + x_5^* = 3$: contrainte saturée.
  $2x_1^* + 4x_2^* + 4x_3^* - 2x_4^* + 5x_5^* = 14/3 < 5$: contrainte vérifiée mais non saturée.
  $3x_1^* + x_2^* + 2x_3^* - x_4^* - 2x_5^* = 1$: contrainte saturée.
  Nous écrivons les égalités que doivent vérifier les nombres y_i^* $(i = 1, 2, 3, 4)$.
- Puisque la troisième contrainte n'est pas saturée, $y_3^* = 0$.
- Puisque $x_2^* > 0$, $3y_1^* + 2y_2^* + 4y_3^* + y_4^* = 6$.
- Puisque $x_3^* > 0$, $5y_1^* - 2y_2^* + 4y_3^* + 2y_4^* = 5$.
- Puisque $x_4^* > 0$, $-2y_1^* + y_2^* - 2y_3^* - y_4^* = -2$.

Calculons les y_i^* $(i = 1, 2, 3, 4)$:

$$\begin{cases} 3y_1^* + 2y_2^* + y_4^* = 6 \\ 5y_1^* - 2y_2^* + 2y_4^* = 5 \\ -2y_1^* + y_2^* - y_4^* = -2 \end{cases}$$

La solution de ce système est : $y_1^* = y_2^* = y_4^* = 1$.

Regardons si les y_i^* $(i = 1, 2, 3, 4)$ constituent une solution réalisable du problème dual.

- Ils sont tous positifs ou nuls.
- Il reste à vérifier la première et la cinquième contrainte du problème dual puisque les autres contraintes sont saturées par définition des y^* :

$$\begin{aligned} &y_1^* + 4y_2^* + 2y_3^* + 3y_4^* = 8 \geq 7 \\ &2y_1^* + y_2^* + 5y_3^* - 2y_4^* = 1 < 3 \end{aligned}$$

La dernière contrainte du dual n'est pas vérifiée : la solution actuelle n'est pas optimale.

On peut remarquer que, si on veut maintenant rechercher la solution optimale, il serait judicieux de partir de la base x_2, x_3, x_4, x_8, où x_8 représente la troisième variable d'écart ; cette base correspond à la solution proposée.

Corrigé de l'exercice II.6.3.

On considère le problème (P) suivant :

$$\text{Maximiser } z = 2x_1 - x_2$$
$$\begin{cases} x_1 - x_2 \leq 1 \\ -x_1 + x_2 \leq -2 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Le dual (Q) de (P) est :

$$\text{Minimiser } w = y_1 - 2y_2$$
$$\begin{cases} y_1 - y_2 \geq 2 \\ -y_1 + y_2 \geq -1 \\ y_1 \geq 0, y_2 \geq 0 \end{cases}$$

On vérifie aisément que le problèmes (P) et (Q) n'admettent pas de solution réalisable.

Corrigé de l'exercice II.6.4.

1. On commence par se rapprocher de la forme standard :

$$\text{Maximiser } \sum_{j=1}^{n} -c_j x_j$$
$$\text{avec : } \begin{cases} \text{pour } i \in \{1,2,...,m\}, \sum_{j=1}^{n} -a_{ij}x_j \leq -b_i \\ \text{pour } i \in \{m+1,...,m+p\}, \sum_{j=1}^{n} -a_{ij}x_j \leq -b_i \\ \text{pour } i \in \{m+1,...,m+p\}, \sum_{j=1}^{n} a_{ij}x_j \leq b_i \end{cases}$$

Mettons maintenant le problème sous forme standard. Pour $j \in \{1,2,...,n\}$, on pose : $x_j = x_j^1 - x_j^2$. On obtient :

$$\text{Maximiser } \sum_{j=1}^{n} -c_j x_j^1 + \sum_{j=1}^{n} c_j x_j^2$$
$$\text{avec : } \begin{cases} \text{pour } i \in \{1,2,...,m\}, \sum_{j=1}^{n} -a_{ij}x_j^1 + \sum_{j=1}^{n} a_{ij}x_j^2 \leq -b_i \\ \text{pour } i \in \{m+1,...,m+p\}, \sum_{j=1}^{n} -a_{ij}x_j^1 + \sum_{j=1}^{n} a_{ij}x_j^2 \leq b_i \\ \text{pour } i \in \{m+1,...,m+p\}, \sum_{j=1}^{n} a_{ij}x_j^1 + \sum_{j=1}^{n} -a_{ij}x_j^2 \leq b_i \end{cases}$$
$$\text{et pour } j \in \{1,2,...,n\}, x_j^1 \geq 0, x_j^2 \geq 0$$

Le problème dual est :

$$\text{Minimiser } \sum_{i=1}^{m} -b_i y_i - \sum_{i=m+1}^{m+p} b_i y_i^1 + \sum_{i=m+1}^{m+p} b_i y_i^2$$
$$\text{avec : } \begin{cases} \text{pour } j \in \{1,2,...,n\}, \\ \qquad \sum_{i=1}^{m} -a_{ij}y_i + \sum_{i=m+1}^{m+p} -a_{ij}y_i^1 + \sum_{i=m+1}^{m+p} a_{ij}y_i^2 \geq -c_j \\ \text{pour } j \in \{1,2,...,n\}, \\ \qquad \sum_{j=1}^{n} a_{ij}y_i + \sum_{j=m+1}^{m+p} a_{ij}y_i^1 + \sum_{j=m+1}^{m+p} -a_{ij}y_i^2 \geq c_j \end{cases}$$
$$\text{et pour } i \in \{1,2,...,m\}, y_i \geq 0$$
$$\text{pour } i \in \{m+1,...,m+p\} y_i^1 \geq 0, y_i^2 \geq 0$$

qui peut se réécrire :

$$\text{Maximiser } \sum_{i=1}^{m} b_i y_i + \sum_{i=m+1}^{m+p} b_i (y_i^1 - y_i^2)$$
$$\text{avec : } \begin{cases} \text{pour } j \in \{1,2,...,n\}, \\ \qquad \sum_{i=1}^{m} a_{ij}y_i + \sum_{i=m+1}^{m+p} a_{ij}(y_i^1 - y_i^2) \leq c_j \\ \text{pour } j \in \{1,2,...,n\}, \\ \qquad \sum_{i=1}^{m} a_{ij}y_i + \sum_{i=m+1}^{m+p} a_{ij}(y_i^1 - y_i^2) \geq c_j \end{cases}$$
$$\text{et pour } i \in \{1,2,...,m\}, y_i \geq 0$$
$$\text{pour } i \in \{m+1,...,m+p\} y_i^1 \geq 0, y_i^2 \geq 0$$

En posant, pour $i \in \{m+1,...,m+p\}$, $y = y_i^1 - y_i^2$, la variable y est non borné et on peut encore écrire ce problème dual :

$$\text{Maximiser } \sum_{i=1}^{m+p} b_i y_i$$
$$\text{avec : pour } j \in \{1,2,...,n\}, \sum_{i=1}^{m+p} a_{ij}y_i = c_j$$
$$\text{et pour } i \in \{1,2,...,m\}, y_i \geq 0$$

2. Le dual obtenu à la question précédente est le problème (Q). Comme le dual du dual est le primal, le problème dual de (Q) est (P).
3. On utilise les questions précédentes en choisissant, pour $i \in \{1, ..., m+p\}$, $b_i = 0$. Les problèmes (P) et (Q) deviennent :

$$(P_0) \quad \begin{array}{l} \text{Minimiser } z = \sum_{j=1}^{n} c_j x_j \\ \text{avec : } \left\{ \begin{array}{l} \text{pour } i \in \{1, 2, ..., m\}, \sum_{j=1}^{n} a_{ij} x_j \geq 0 \\ \text{pour } i \in \{m+1, ..., m+p\}, \sum_{j=1}^{n} a_{ij} x_j = 0 \end{array} \right. \end{array}$$

et

$$(Q_0) \quad \begin{array}{l} \text{Maximiser } w = 0 \\ \text{avec : pour } i \in \{1, 2, ..., m+p\}, \sum_{i=1}^{p} a_{ij} y_i = c_j \\ \text{et pour } i \in \{1, ..., m\}, y_i \geq 0 \end{array}$$

Si la proposition a. est vérifiée, le minimum du problème (P_0) est 0. D'après le théorème de la dualité, le problème (Q_0) est réalisable, ce qui signifie que la proposition b. est vérifiée.

Si Si la proposition b. est vérifiée, le problème (Q_0) est réalisable, de maximum 0. D'après le théorème de la dualité, le problème (P_0) est réalisable de minimum 0, ce qui signifie que la proposition a. est vérifiée.

V.3. Corrigé des exercices du chapitre 3

Corrigé de l'exercice III.11.1.

La fonction f est de classe C^∞. Commençons par déterminer le gradient et la matrice hessienne de f :

$$\nabla f(x,y) = \begin{pmatrix} e^{x+y} + 2x \\ e^{x+y} + 4y \end{pmatrix}, \nabla^2 f(x,y) = \begin{pmatrix} e^{x+y} + 2 & e^{x+y} \\ e^{x+y} & e^{x+y} + 4 \end{pmatrix}.$$

Le déterminant de la hessienne (produit des valeurs propres) ainsi que sa trace (somme des valeurs propres) étant strictement positifs, les valeurs propres de $\nabla^2 f(x,y)$ sont strictement positives et $\nabla^2 f(x,y)$ est définie positive : f est donc convexe.

On en déduit que tout minimum local est global, et une condition nécessaire et suffisante pour que x^* soit un minimum est $\nabla f(x^*) = 0$. Comme par ailleurs f tend vers l'infini à l'infini, f admet un minimum global, que l'on cherche par la méthode du gradient à pas optimal.

On part de $P_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$; $\nabla f(P_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $d_0 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

On cherche λ tel que $g(\lambda) = f(P_0 + \lambda d_0)$ soit minimum :

$g(\lambda) = f(-\lambda, -\lambda) = e^{-2\lambda} + 3\lambda^2$.

On minimise par dichotomie ou par la méthode de Newton et on trouve $\lambda = 0,216$. D'où :

$P_1 = \begin{pmatrix} -0,216 \\ -0,216 \end{pmatrix}$; $\nabla f(P_1) = \begin{pmatrix} 0,216 \\ -0,216 \end{pmatrix}$; $d_1 = \begin{pmatrix} -0,216 \\ 0,216 \end{pmatrix}$.

On constate que d_1 est orthogonal à d_0.

$$\begin{aligned} g(\lambda) &= f(P_1 + \lambda d_1) \\ &= f(-0,216(1+\lambda), -0,216(1-\lambda)) \\ &= e^{-2\times 0,216} + (0,216)^2[(1+\lambda)^2 + 2(1-\lambda)^2] \\ &= e^{-2\times 0,216} + (0,216)^2 h(\lambda) \end{aligned}$$

avec $h(\lambda) = [(1+\lambda)^2 + 2(1-\lambda)^2]$.
Cherchons à optimiser h : son minimum est atteint pour $\lambda = 1/3$. D'où :
$P_2 = \begin{pmatrix} -0,216(1+1/3) \\ -0,216(1-1/3) \end{pmatrix} = \begin{pmatrix} -0,288 \\ -0,144 \end{pmatrix}$; $\nabla f(P_2) = \begin{pmatrix} 0,0732 \\ 0,0732 \end{pmatrix}$;
$d_2 = \begin{pmatrix} -0,0732 \\ -0,0732 \end{pmatrix}$ (on constate de nouveau que d_1 est orthogonal à d_0...).

On considère maintenant :

$$\begin{aligned} g(\lambda) &= f(-0,288 - \lambda, -0,144 - \lambda) \\ &= e^{-0,432-2\lambda} + (0,288+\lambda)^2 + 2(0,144+\lambda)^2. \end{aligned}$$

On minimise par la méthode de Newton, et on trouve $\lambda = 0,07119$, d'où :

$P_3 = \begin{pmatrix} -0,305 \\ -0,161 \end{pmatrix}$.

On peut continuer ainsi pour avoir plus de précision.

V.4. Corrigé des exercices du chapitre 4

Corrigé de l'exercice IV.6.

1. Soit $f(x,y) = 2x^2 + y^4, h_1(x,y) = 1 - x$ et $h_2(x,y) = a + 1 - x - ay$. Le problème s'écrit :

$$\begin{array}{l} \text{Minimiser } f(x,y) \\ \text{avec les contraintes :} \\ \left\{ \begin{array}{l} h_1(x,y) \leq 0 \\ h_2(x,y) \leq 0 \end{array} \right. \end{array}$$

$\nabla^2 f(x,y) = \begin{pmatrix} 4 & 0 \\ 0 & 12y^2 \end{pmatrix}$ qui est définie positive : la fonction f est convexe sur $\mathbb{R}$.

Par ailleurs, une fonction affine est convexe (et d'ailleurs aussi concave) : h_1 et h_2 sont convexes.

Il faut et il suffit donc que la condition de Karush, Kuhn et Tucker soit vérifiée au point (1, 1) pour que ce point réalise le minimum de notre problème. En ce point les deux contraintes sont saturées ; dire que la condition de Karush, Kuhn et Tucker est vérifiée, c'est montrer qu'il existe deux réels positifs ou nuls μ_1 et μ_2 tels que :

$\nabla f(1,1) = -\mu_1 \nabla h_1(1,1) - \mu_2 \nabla h_2(1,1)$.

$\nabla f(x,y) = \begin{pmatrix} 4x \\ 4y^3 \end{pmatrix}$, $\nabla f(1,1) = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$

$\nabla h_1(1,1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, $\nabla h_2(1,1) = \begin{pmatrix} -1 \\ -a \end{pmatrix}$

Cherchons des coefficients μ_1 et μ_2 tels que :

$\begin{pmatrix} 4 \\ 4 \end{pmatrix} = \mu_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 \\ a \end{pmatrix}$

ce qui s'écrit :

$$\left\{ \begin{array}{l} \mu_1 + \mu_2 = 4 \\ a\mu_2 = 4 \end{array} \right.$$

Pour qu'il y ait une solution, il faut et il suffit que $a \neq 0$ et alors :

$$\mu_1 = 4(1 - \frac{1}{a}), \mu_2 = \frac{4}{a}.$$

La condition de Kuhn et Tucker est vérifiée si et seulement si μ_1 et μ_2 sont positifs ou nuls c'est-à-dire si et seulement si $\boxed{a \geq 1}$.

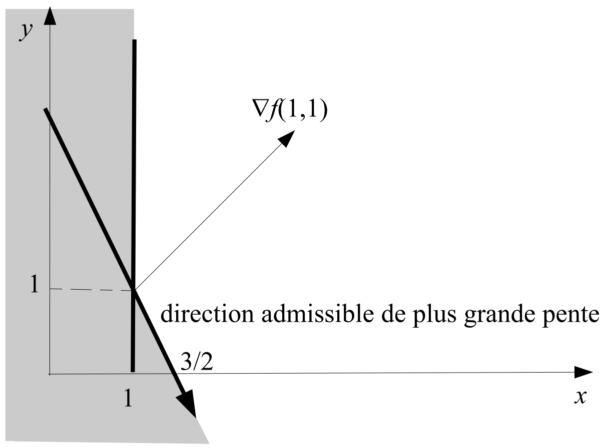

FIGURE V.2 – Solution graphique

2. C'est le problème de la question 1. avec $a = 1/2$: le minimum n'est pas atteint au point (1,1).

Nous appliquons la méthode des directions admissibles, en partant du point (1,1) : nous cherchons la direction admissible de plus grande pente pour f en ce point (voir la figure V.2).

On voit graphiquement que cette direction est $d = (1, -2)$. Nous cherchons alors le minimum de $g(t) = f[(1,1) + t(1,-2)]$ pour t positif ; nous remarquons en effet qu'ainsi nous ne sortons pas du domaine.

$g(t) = 2(1+t)^2 + (1-2t)^4$. Donc, $g'(t) = 4(1+t) - 8(1-2t)^3$.

Nous remarquons que $g''(t) = 4 + 48(1-2t)^2 > 0$ donc $g'(t)$ est croissante.

$g'(0) = -4 < 0; g'(1/2) = 6 > 0$. Nous allons déterminer t par dichotomie.
$g'(0,25) > 0$; $g'(0,125) > 0$; $g'(0,06) < 0$; $g'(0,09) < 0$; $g'(0,1) > 0$;
$g'(0,095) > 0$; $g'(0,0925) > 0$; $g'(0,092) > 0$; $g'(0,091) < 0$;
$g'(0,0915) > 0$; $g'(0,0913) < 0$; $g'(0,0914) < 0$; $g'(0,09145) > 0$;
$g'(0,09142) > 0$; $g'(0,09141) > 0$:

$$0,09140 < t_{min} < 0,09141.$$

Le minimum de f dans la direction d est donc atteint au point : (1,0914 ; 0,8172).

Seule la contrainte h_2 est saturée en ce point. Regardons si la condition de Kuhn et Tucker est maintenant vérifiée.

$\nabla f(1,0914, 0,8172) = \begin{pmatrix} 4,3656 \\ 2,18296 \end{pmatrix}$

$\nabla h_2((1,0914;0,8172)) = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = \dfrac{1}{4,3656}\begin{pmatrix} 4,3656 \\ 2,18296 \end{pmatrix}$.

$\nabla f(1,0914,0,8172)$ et $\nabla h_2((1,0914;0,8172))$ sont colinéaires et de même sens : la condition de Kuhn et Tucker est vérifiée.

Corrigé de l'exercice IV.6.

Remarquons d'abord que la fonction $f_\alpha(x,y)$ est convexe. En effet, sa matrice hessienne est : $\begin{pmatrix} 2 & 0 \\ 0 & 2\alpha \end{pmatrix}$, dont les valeurs propres sont 2 et 2α, qui sont des nombres positifs.

Posons : $h_1(x,y) = x^2 + (y+1)^2 - 2, h_2(x,y) = -x, h_3(x,y) = -y$.

Le problème s'écrit :

$$\begin{array}{l} \text{Minimiser } f_\alpha(x,y) \\ \text{avec } \begin{cases} h_1(x,y) \leq 0 \\ h_2(x,y) \leq 0 \\ h_3(x,y) \leq 0 \end{cases} \end{array}$$

La matrice hessienne de la fonction h_1 vaut : $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, et a donc ses valeurs propres, 2 et 2, strictement positives. La fonction h_1 est convexe, les fonctions h_2 et h_3 sont convexes car affines. La condition de Karush, Kuhn et Tucker pour être un minimum local est donc ici nécessaire et suffisante.

On a : $h_1(1,0) = 0, h_2(1,0) = -1, h_3(1,0) = 0$. Seuls les gradients des fonctions h_1 et h_3 au point (1, 0) interviennent dans la condition de Karush, Kuhn et Tucker en ce point.

$\nabla f(x,y) = \begin{pmatrix} 2(x-2) \\ 2\alpha(y-1) \end{pmatrix}$, $\nabla f(1,0) = \begin{pmatrix} -2 \\ -2\alpha \end{pmatrix}$.

$\nabla h_1(x,y) = \begin{pmatrix} 2x \\ 2(y+1) \end{pmatrix}$, $\nabla h_1(1,0) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

$\nabla h_3(x,y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, $\nabla h_3(1,0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$.

Décomposons $\nabla f(1,0)$ dans la base constituées de $-\nabla h_1(1,0)$ et $-\nabla h_3(1,0)$ en écrivant $\nabla f(1,0) = \mu_1(-\nabla h_1(1,0)) + \mu_2(-\nabla h_3(1,0))$. Le point (1, 0) est un minimum local si et seulement si les coefficients μ_1 et μ_2 sont positifs ou nuls.

La relation $\nabla f(1,0) = \mu_1(-\nabla h_1(1,0)) + \mu_2(-\nabla h_3(1,0))$ s'écrit :

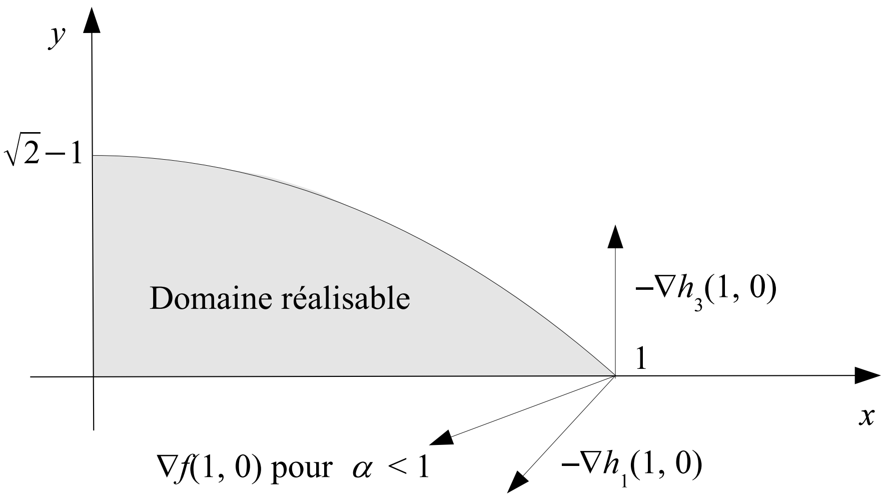

FIGURE V.3 – Solution graphique

$$\begin{cases} -2 = -2\mu_1 \\ -2\alpha = -2\mu_1 + \mu_3 \end{cases} \text{ ou encore : } \begin{cases} \mu_1 = 1 \\ \mu_3 = 2(1-\alpha) \end{cases}.$$

Le point (1, 0) est un minimum local (et même un minimum global) si et seulement si : $\boxed{\alpha \leq 1}$ (en gardant l'hypothèse $\alpha \geq 0$) (voir figure V.3).

Index