Optimisation

Irène Charon
Olivier Hudry

Télécom ParisTech

Octobre 2012

Les quatre premiers chapitres de ce polycopié et le corrigé des exercices s'y rapportant sont extraits du livre « Méthodes d'optimisation combinatoire », par Anne Germa, Irène Charon et Olivier Hudry, publié chez Masson, 1996.

I. L'algorithme du simplexe

I.1. Introduction

Afin d'illustrer ce qu'est la « programmation 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. Chaque unité de produit en cours de fabrication doit passer successivement sur les deux machines dans un ordre indifférent et pendant les temps suivants (en minutes) :

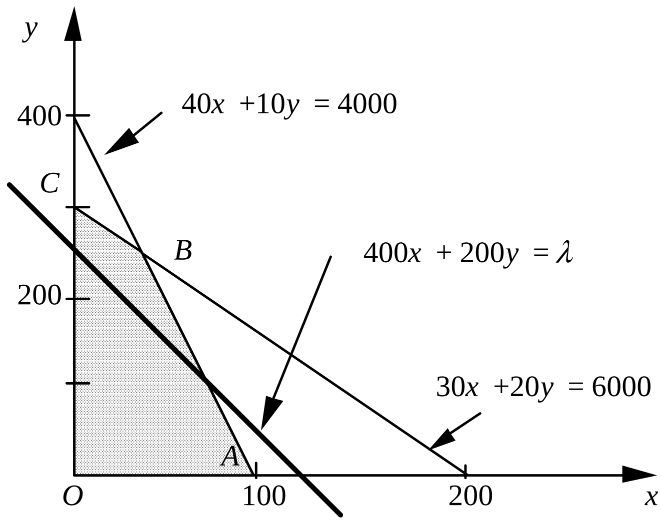
	p_1	p_2
m_1	30	20
m_2	40	10

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 :

$$\text{Maximiser } z = 400\,x + 200\,y$$

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 :



Les points (x, y) qui satisfont les contraintes appartiennent au quadrilatère *ABCO*. La famille des droites $D_\lambda = \{(x, y)$ avec $400\,x + 200\,y = \lambda\}$ 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 $z = 64\,000$).

Cet exercice est une illustration de la « programmation linéaire ». 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 : $\sum_{j=1}^{n} a_{ij}x_j \le b_i$ pour $i \in \{1, ..., m\}$,
- aux n contraintes de positivité : $x_j \ge 0$, pour $j \in \{1, ..., n\}$.

Cette formulation s'appelle la *forme standard* d'un problème de programmation linéaire. Des problèmes où il s'agit de minimisation, ou pour lesquels apparaissent des contraintes d'égalité ou d'inégalité large dans l'autre sens (on ne considérera pas d'inégalité stricte), ou encore pour lesquels les variables sont de signe non contraint ou au contraire sont contraintes à être négatives ou nulles peuvent facilement se mettre sous forme standard, comme le précisent les indications suivantes :

- 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 « $\ge$ » en une inégalité du genre « $\le$ » en la multipliant par –1 ;
- une égalité $\alpha = \beta$ revient aux deux inégalités $\alpha \le \beta$ et $-\alpha \le -\beta$;
- on remplace une variable x contrainte à être négative ou nulle par $-x$;
- on exprime une variable x qui n'a pas de signe imposé par la différence de deux variables positives ou nulles : $x = x^+ - x^-$ avec $x^+ \ge 0$ et $x^- \ge 0$.

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 sous forme standard, considérons un problème de la forme :

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

L'ensemble des points de $\mathbb{R}^n$ de coordonnées $x_1, ..., x_n$ (par rapport à une base déterminée) et vérifiant les $m + n$ contraintes précédentes détermine un polyèdre convexe appelé *polyèdre des contraintes*. Le terme *convexe* est justifié par la remarque suivante : soient $M = (x_1, ..., x_n)$ et $P = (y_1, ..., y_n)$ deux points quelconques du polyèdre déterminé par les contraintes ; alors, quel que soit le réel λ avec $0 \leq \lambda \leq 1$, 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.** *On considère une forme linéaire des n variables $x_1, ..., x_n$, soumises à des contraintes linéaires. Son maximum, qui existe si cette forme est majorée, est atteint au moins en 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 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 « monter suivant les arêtes du polyèdre de sommet en sommet » pour trouver le maximum.

I.2. L'algorithme du simplexe appliqué à un exemple

Une fabrique d'objets en terre cuite peinte produit des cendriers, des cruches, des bols et des vases. La fabrication de chacun de ces objets nécessite un certain nombre d'heures de moulage, un certain nombre d'heures de cuisson et un certain nombre d'heures de peinture. La vente de ces objets rapporte un certain bénéfice exprimé en euros. Ces données sont récapitulées dans le tableau suivant :

Objet	Cendrier	Bol	Cruche	Vase
Moulage	2	4	5	7
Cuisson	1	1	2	2
Peinture	1	2	3	3
Bénéfice	7	9	18	17

L'entreprise dispose, quotidiennement, de 42 heures de moulage, 17 heures de cuisson et 24 heures de peinture.
On souhaite établir un plan de fabrication de façon à maximiser le chiffre d'affaires (on suppose que l'on est en régime stable de fabrication et non en phase initiale où il faut mouler avant de cuire et cuire avant de peindre).
Appelant x_1, x_2, x_3, x_4 les nombres respectifs de cendriers, cruches, bols et vases fabriqués quotidiennement, le problème admet la modélisation suivante :

$$\text{Maximiser } z = 7x_1 + 9x_2 + 18x_3 + 17x_4$$

avec les contraintes : $\begin{cases} 2x_1 + 4x_2 + 5x_3 + 7x_4 \leq 42 \\ x_1 + x_2 + 2x_3 + 2x_4 \leq 17 \\ x_1 + 2x_2 + 3x_3 + 3x_4 \leq 24 \\ x_1, x_2, x_3, x_4 \text{ entières positives ou nulles} \end{cases}$

On reconnaît là, aux « contraintes d'intégrité » près, un problème de programmation linéaire sous forme standard. Nous allons « relâcher les contraintes d'intégrité » en les remplaçant par $x_i \geq 0$ pour $i \in \{1, 2, 3, 4\}$ et résoudre ce problème à l'aide de l'algorithme du simplexe que nous expliquerons donc sur cet exemple.

Introduisons 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 :

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

Le problème s'écrit maintenant :

Maximiser z avec les contraintes $x_i \geq 0$ pour $i \in \{1, ..., 7\}$.

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

Le tableau ci-dessus est appelé un *dictionnaire*. Les variables x_5, x_6 et x_7 y 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* et les variables x_1, x_2, x_3 et x_4 les *variables hors-base*. On s'intéresse alors à ce qu'on appelle la *solution basique* associée au dictionnaire ; celle-ci est la solution obtenue en donnant 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 la dernière équation du dictionnaire ci-dessus 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 :

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

On obtient ainsi :

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

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 :

$x_3 \geq 0$, 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 :

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

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 non négatives, l'optimum cherché de z est majoré par 147. La solution basique actuelle nous fournit donc la solution optimum du problème :

- il faut fabriquer chaque jour trois cendriers, aucun bol, sept cruches, aucun vase ;
- toutes les heures de cuisson et de peinture sont utilisées, alors qu'il reste une heure de moulage disponible ;
- le chiffre d'affaires est de 147.

Ayant « relâché » les contraintes d'intégrité lors de la recherche de la solution, il se trouve que la solution obtenue est entière : ceci n'a rien de général et les problèmes de programmation linéaire en nombres entiers peuvent être qualitativement plus compliqués, comme nous aurons l'occasion de le voir par la suite.
L'application de la méthode, en faisant à chaque étape entrer une variable en base lorsque son coefficient dans la fonction à optimiser est strictement positif, ne permet pas toujours d'obtenir une croissance stricte de la valeur de la fonction z. Nous reviendrons sur ce phénomène dans le paragraphe suivant, en parlant de dégénérescence.
Enfin, nous avons eu la chance de trouver, sans aucune difficulté, un sommet du polyèdre des contraintes, ou autrement dit un dictionnaire réalisable ; en effet, « l'origine était réalisable » car les b_i étaient tous positifs. Dans le paragraphe 4 de ce chapitre, nous étudierons des cas moins favorables.

I.3. La dégénérescence et le cyclage

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

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

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

Tout n-uplet de valeurs (x_1, ..., x_n) satisfaisant les contraintes constitue une *solution réalisable.*
La fonction z est dite *fonction objectif.* Les variables x_1, ..., x_n sont appelées *variables de décision* ou aussi *variables de choix* ; les variables x_{n+1}, …, x_{n+m} telles qu'elles ont été introduites dans l'exemple plus haut s'appellent les *variables d'écart*. 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.* Si un problème admet des solutions réalisables mais n'a pas de valeur optimale finie, il est dit *non*

borné.
Une solution réalisable que l'on peut exprimer à l'aide d'un dictionnaire est dite *solution de base réalisable*. 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} (les m variables en base) en fonction des n autres variables (les n variables hors-base) ;
- tout dictionnaire est algébriquement équivalent au dictionnaire définissant les variables d'écart et la fonction objectif, c'est-à-dire au dictionnaire :

$$\begin{cases} x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} x_j & \text{pour } i \in \{1, \ldots, m\} \\ z = \sum_{j=1}^{n} c_j x_j \end{cases}$$

Nous voyons que dans l'algorithme du simplexe, on ne considère que des solutions de base réalisables (ce qui revient à ne considérer que des sommets du polyèdre des contraintes).
Les solutions de base réalisables avec une ou plusieurs variables de base nulles sont dites *dégénérées*.

EXEMPLE

Considérons le dictionnaire :

$$\begin{array}{l} x_4 = 1 - 2x_3 \\ x_5 = 3 - 2x_1 + 4x_2 - 6x_3 \\ \underline{x_6 = 2 + x_1 - 3x_2 - 4x_3} \\ z \;\;=\;\; 2x_1 - x_2 + 8x_3 \end{array}$$

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 :

$$\begin{array}{l} x_3 = 0{,}5 - 0{,}5x_4 \\ x_5 = \phantom{0{,}5} - 2x_1 + 4x_2 + 3x_4 \\ \underline{x_6 = \phantom{0{,}5 - 2}x_1 - 3x_2 + 2x_4} \\ z = 4 + 2x_1 - x_2 - 4x_4 \end{array}$$

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 dite 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 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.
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 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 ; 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 donc $s < t$.
En notant B l'ensemble des indices de base de D', on peut écrire D' sous la forme :

$$x_i = b'_i - \sum_{j \notin B} a'_{ij} x_j \quad \text{pour } i \in B$$

$$z = v + \sum_{j \notin B} c'_j x_j$$

La variable x_s étant entrante, on a $c_s > 0$ et, la règle de Bland étant utilisée, on a, pour $i < s$, $c_i \leq 0$. La variable x_t étant sortante dans D', il vient $a'_{ts} > 0$.
La dernière ligne de D'' peut s'écrire :

$$z = v + \sum_{i=1}^{n+m} c''_i x_i$$

où c''_i est nul si x_i est en base, et $c''_t > 0$.
Pour toute solution (x_1^*, ..., x_{n+m}^*) du système des contraintes, on a :

$$v + \sum_{j \notin B} c'_j x_j^* = v + \sum_{i=1}^{n+m} c''_i x_i^*$$

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 :

$$c'_s x_s^* = c''_s x_s^* + \sum_{i \in B} c''_i \left(b'_i - a_{is} x_s^* \right)$$

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 $c''_r a'_{rs} < 0$.

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 : $r < t$.

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' et en choisissant x_t, avec $t > r$, nous n'avons pas appliqué la règle du plus petit indice. □

I.4. 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.

Maximiser $z = x_1 - x_2 + x_3$

avec les contraintes : $\begin{cases} 2x_1 - x_2 + 2x_3 \leq 4 \\ 2x_1 - 3x_2 + x_3 \leq -5 \\ -x_1 + x_2 - 2x_3 \leq -1 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$

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

Maximiser $w = -x_0$ (ce qui traduit : minimiser x_0)

avec les contraintes : $\begin{cases} 2x_1 - x_2 + 2x_3 - x_0 \leq 4 \\ 2x_1 - 3x_2 + x_3 - x_0 \leq -5 \\ -x_1 + x_2 - 2x_3 - x_0 \leq -1 \\ x_0 \geq 0, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$

REMARQUE

On aurait pu 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 6.

Ce problème 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.

Écrivons le dictionnaire définissant les variables d'écart du nouveau problème :

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

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 en une itération on peut se ramener 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 :

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

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 :

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

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

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

Nous voyons que le problème initial avait une solution réalisable donnée par $x_1^* = 0$; $x_2^* = 2$; $x_3^* = 1{,}6$; de plus, 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 :

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

Cette méthode est connue comme *méthode à deux phases*. Dans le chapitre III, 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 dual », ce qui permet de ne résoudre qu'un problème au lieu de deux. Cette classe de problèmes est dite *dual-réalisable*.

I.6. Exercices

Exercice 1

Résoudre l'exercice suivant par la méthode du simplexe :

Maximiser $z = 3x_1 + 2x_2 + 4x_3$

avec les contraintes : $\begin{cases} x_1 + x_2 + 2x_3 \le 4 \\ 2x_1 + 3x_3 \le 5 \\ 2x_1 + x_2 + 3x_3 \le 7 \\ x_1 \ge 0, x_2 \ge 0, x_3 \ge 0 \end{cases}$

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.

Maximiser $z = 5x_1 + 6x_2 + 9x_3 + 8x_4$

avec les contraintes : $\begin{cases} x_1 + 2x_2 + 3x_3 + x_4 \le 5 \\ x_1 + x_2 + 2x_3 + 3x_4 \le 3 \\ x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0 \end{cases}$

Exercice 3

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

Exercice 4

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

Exercice 5

Résoudre le problème représenté par le dictionnaire réalisable suivant :

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

On appliquera les règles de choix suivantes, pour les variables entrantes et sortantes, en cas de choix :

- la variable entrante est toujours la variable non en base qui a le plus grand coefficient dans la fonction objectif ;
- on fait sortir de la base, en cas de choix, la variable de plus petit indice.

Que constate-t-on ?

Exercice 6

Résoudre les deux exercices suivants :

1. Maximiser $z = 5x_1 + 3x_2$

avec les contraintes : $\begin{cases} -4x_1 + 5x_2 \le -10 \\ 5x_1 + 2x_2 \le 10 \\ 3x_1 + 8x_2 \le 12 \\ x_1 \ge 0, x_2 \ge 0 \end{cases}$

2. Maximiser $z = 5x_1 + 3x_2$

avec les contraintes : $\begin{cases} -4x_1 - 5x_2 \le -10 \\ 5x_1 + 2x_2 \le 10 \\ 3x_1 + 8x_2 \le 12 \\ x_1 \ge 0, x_2 \ge 0 \end{cases}$

II. Forme matricielle de la méthode du simplexe

II.1. Généralités

On considère un problème d'optimisation linéaire à $n + m$ variables et n contraintes qui s'écrit :

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

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

NOTATIONS

On pose :

- $A = (a_{ij})_{i=1...m,\, j=1...n+m}$
- $X = (x_1, \dots, x_{n+m})^t$
- $c = (c_1, c_2, \dots, c_{n+m})$
- $b = (b_1, \dots, b_m)^t$

Notre problème s'écrit maintenant :

$$\text{Maximiser } z = c.X \text{ avec } A.X = b \text{ et } X \geq 0.$$

Remarquons que cette définition généralise celle d'un problème de programmation linéaire mis sous forme standard, après introduction des variables d'écart.

SUITE DES NOTATIONS

$x_{j_1}, ..., x_{j_m}$ étant m des $n + m$ variables et $x_{j_{m+1}}, \dots, x_{j_{m+n}}$ les autres variables, on pose :

- $X_B = (x_{j_1}, \dots, x_{j_m})^t$
- $X_N = (x_{j_{m+1}}, \dots, x_{j_{m+n}})^t$

- B = la matrice formée par les m colonnes de A correspondant aux coefficients de $x_{j_1}, \ldots, x_{j_m}$ (en respectant l'ordre des variables)
- A_N = la matrice formée par les n colonnes de A correspondant aux coefficients de $x_{j_{m+1}}, \ldots, x_{j_{m+n}}$ (en respectant l'ordre des variables)
- c_B = matrice-ligne formée des coefficients dans z de $x_{j_1}, \ldots, x_{j_m}$ (en respectant l'ordre des variables).
- c_N = matrice-ligne formée des coefficients dans z de $x_{j_{m+1}}, \ldots, x_{j_{m+n}}$ (en respectant l'ordre des variables).

Nous avons introduit, dans le chapitre précédent, les notions de variable en base et de variable hors-base ; il est temps de donner précisément la définition d'une base.

Définition. *Une base est constituée de m variables qui s'expriment de façon unique, et affine, en fonction des n autres variables, cette expression étant algébriquement équivalente aux m contraintes d'égalité initiales.*

Propriété. $x_{i_1}, \ldots, x_{j_m}$ *est une base si et seulement si la matrice B correspondante est inversible.*

Preuve. Le système $AX = b$ s'écrit : $BX_B + A_NX_N = b$
ou encore :

$$BX_B = b - A_NX_N.$$

Si B est inversible, $AX = b$ est équivalent à $X_B = B^{-1}b - B^{-1}A_NX_N$: X_B est bien une base. Réciproquement, si $x_{j_1}, \ldots, x_{j_m}$ est une base, alors les valeurs des variables hors-base étant fixées, les valeurs de $x_{j_1}, \ldots, x_{j_m}$ sont déterminées de manière unique par $AX = b$; pour un X_N donné, $BX_B = b - A_NX_N$ admet une solution unique. Ce qui implique que B est inversible. □

Comme dans le chapitre précédent, si V est une variable quelconque, V^* représente une valeur prise par cette variable. Une base étant choisie, la solution basique associée est obtenue en donnant à toutes les variables hors-base la valeur 0, c'est-à-dire par $X_N^* = 0$. Les valeurs prises par les variables de base sont alors données par :

$$X_B^* = B^{-1}b$$

La base est dite réalisable si toutes les coordonnées de X_B^* sont positives ou nulles. Trouver une base réalisable revient à déterminer une matrice carrée B de dimension (n, n), extraite de A, inversible et telle que $B^{-1}b$ ait tous ses coefficients positifs ou nuls. Une telle base nous permettra d'initialiser l'algorithme du simplexe.
On peut remarquer que, lorsque le problème initial est donné sous forme standard, la matrice carrée formée des coefficients des variables d'écart est l'identité ; c'est donc une matrice inversible : on peut prendre comme base l'ensemble des variables d'écart. Cette base est réalisable si le vecteur b n'a aucune composante négative.

II.2. Version matricielle d'une itération de la méthode du simplexe

On suppose que l'on dispose d'une base réalisable $X_B = (x_{j_1}, \ldots, x_{j_m})^t$, vecteur-colonne de variables de base, associée à une matrice inversible B, et l'on veut effectuer une étape de la méthode du simplexe, en exploitant au maximum les propriétés algébriques du problème. Nous cherchons à échanger une variable, dite sortante, de X_B par une variable non dans X_B, dite entrante, de façon à obtenir une nouvelle base réalisable tout en augmentant la valeur de l'objectif.
Au départ de l'étape, on dispose de $X_B = (x_{j_1}, \ldots, x_{j_m})^t$ (dont on peut déduire B, A_N, c_B, c_N) et de X_B^*.

II.2.1. Recherche d'une variable entrante

Pour rechercher une variable entrante, il faut calculer, dans le dictionnaire actuel (non explicité), les coefficients des variables hors-base dans la ligne exprimant z et voir si au moins un de ces coefficients est strictement positif. Si tous les coefficients sont négatifs ou nuls, la solution basique actuelle est optimale.
On a :

$$X_B = B^{-1}b - B^{-1}A_N X_N$$

$$\boxed{X_B = X_B^* - B^{-1}A_N X_N}$$

Par ailleurs : $z = c.X = c_B X_B + c_N X_N$

D'où : $z = c_B(X_B^* - B^{-1} A_N.X_N) + c_N.X_N$

$$\boxed{z = c_B X_B^* + (c_N - c_B B^{-1} A_N) X_N}$$

Les deux formules ci-dessus encadrées forment le dictionnaire actuel.
Posons : $y = c_B B^{-1}$; y est un vecteur-ligne à m composantes que l'on détermine en pratique en résolvant le système :

$$yB = c_B.$$

Soit x_j une variable hors-base ; on cherche d'abord le coefficient de x_j dans $c_B B^{-1} A_N X_N$, c'est-à-dire dans $yA_N X_N$:

$$yA_N X_N = (y_1, ..., y_m) \begin{pmatrix} . & . & . & a_{1j} & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & a_{mj} & . & . \end{pmatrix} \begin{pmatrix} . \\ . \\ . \\ x_j \\ . \\ . \end{pmatrix}$$

Posons : $a_j = \begin{pmatrix} a_{1j} \\ . \\ . \\ a_{mj} \end{pmatrix}$.

Le coefficient de x_j dans $c_B B^{-1} A_N X_N$ est $y.a_j$. Le coefficient de x_j dans z vaut donc $c_j - ya_j$.
En conséquence, dans le cas où on cherche un maximum, trouver une variable entrante, c'est trouver une colonne a_j de A_N telle que : $c_j - ya_j > 0$. La variable x_j correspondante sera entrante.

II.2.2. Recherche d'une variable sortante

Nous supposons maintenant que l'on a choisi une variable entrante x_j correspondant à une colonne a_j de A. Il faut maintenant déterminer une variable sortante.
On considère que toutes les variables hors-base restent nulles sauf x_j. X_B vaut alors :

$$X_B = X_B^* - B^{-1} A_N (0, 0, \ldots, 0, x_j, 0, \ldots, 0)^t$$

On pose : $d = B^{-1} a_j$; d est un vecteur-colonne à m composantes que l'on détermine en pratique en résolvant $Bd = a_j$; on a :

$$X_B = X_B^* - x_j . d$$

L'accroissement de x_j (qui fait croître la valeur de la fonction objectif) est limité par la condition : $X_B \geq 0$. On cherche alors la plus grande valeur x_j^* de x_j telle que :

$$X_B = X_B^* - x_j^* . d \geq 0.$$

La première coordonnée de X_B qui s'annule définit la variable sortante.

II.2.3. Actualisation

Pour pouvoir, à l'étape suivante, recommencer dans les mêmes conditions, il faut actualiser X_B et X_B^*. On obtient le nouveau vecteur-colonne $X_{B'}$ associé à la nouvelle base B' en remplaçant dans X_B la variable sortante par la variable entrante. Les valeurs des variables qui n'ont pas quitté la base sont calculées par la formule $X_{B'} = X_B^* - x_j^* . d$ et, par définition, la valeur de la variable entrante est x_j^*.

II.2.4. Résumé

Nous allons maintenant résumer l'algorithme du simplexe sous sa forme révisée, tel que nous l'avons développé dans ce chapitre.
Considérons le problème :

Maximiser $c.X$ sous les contraintes $A.X = b$ et $X \geq 0$.

et supposons connue une base réalisable. On appelle c_B et B les parties de c et de A correspondant à cette base réalisable X_B. Il nous faut initialiser X_B^* qui vaut $B^{-1}b$.
Une étape se déroule alors comme suit :

1. Résoudre le système $yB = c_B$.

2. Choisir une colonne entrante s'il en existe une, c'est-à-dire une colonne a_j de A non dans B, telle que $y.a_j$ soit plus petit que c_j (s'il s'agissait de minimiser, on chercherait une colonne a_j

telle que $y.a_j$ soit plus grand que c_j). S'il n'existe pas de colonne entrante, la base est optimale et l'algorithme s'arrête.

3. Résoudre le système $Bd = a_j$.

4. Trouver le plus grand t tel que $X_B^* - td \geq 0$. S'il n'existe pas de telle valeur de t, le problème est non borné. Sinon, lorsque t atteint cette valeur maximum, une au moins des composantes de $X_B^* - td$ vaut 0 et la variable correspondante peut quitter la base.

5. Remplacer dans X_B la variable sortante par la variable entrante. Mettre à t dans X_B^* la valeur de la variable entrante, et pour les autres variables remplacer leur valeur dans X_B^* par leur valeur dans $X_B^* - td$. Dans B, remplacer la colonne sortante, c'est-à-dire la colonne associée à la variable sortante, par la colonne entrante.

REMARQUE

Lorsque le problème est donné sous forme standard, si la solution nulle est réalisable, on peut initialiser l'algorithme avec $B = I$ et $X_B^* = b$, ceci correspondant au choix de l'ensemble des variables d'écart pour constituer la première base.

II.3. Un exemple d'application de la méthode

Nous reprenons l'exemple du premier chapitre.
Le problème s'écrit :

$$\text{Maximiser } z = 7x_1 + 9x_2 + 18x_3 + 17x_4$$

$$\text{avec les contraintes : } \begin{cases} 2x_1 + 4x_2 + 5x_3 + 7x_4 + x_5 = 42 \\ x_1 + x_2 + 2x_3 + 2x_4 + x_6 = 17 \\ x_1 + 2x_2 + 3x_3 + 3x_4 + x_7 = 24 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0, x_6 \geq 0, x_7 \geq 0 \end{cases}$$

On a ici :

$$A = \begin{pmatrix} 2 & 4 & 5 & 7 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 2 & 3 & 3 & 0 & 0 & 1 \end{pmatrix}, \; b = \begin{pmatrix} 42 \\ 17 \\ 24 \end{pmatrix}, \; c = (7 \;\; 9 \;\; 18 \;\; 17 \;\; 0 \;\; 0 \;\; 0).$$

On choisit comme base de départ les variables d'écart et donc :

$$X_B = \begin{pmatrix} x_5 \\ x_6 \\ x_7 \end{pmatrix} \qquad X_B^* = \begin{pmatrix} 42 \\ 17 \\ 24 \end{pmatrix}$$

D'où :

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \; A_N = \begin{pmatrix} 2 & 4 & 5 & 7 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 3 & 3 \end{pmatrix}, \; c_B = (0 \;\; 0 \;\; 0), \quad c_N = (7 \;\; 9 \;\; 18 \;\; 17).$$

ÉTAPE 1

1. On résout $y.B = c_B$: $y = (0\ \ 0\ \ 0)$.

2. $y.a_1 = 0 < c_1$, $y.a_2 = 0 < c_2$, $y.a_3 = 0 < c_3$, $y.a_4 = 0 < c_4$: les quatre variables sont candidates à entrer. On choisit la variable x_3 pour suivre un cheminement parallèle à celui adopté pour la résolution de ce même problème dans le premier chapitre.

3. On résout $B.d = a_3$: $d = a_3$.

4. $X_B^* - t.d \geq 0$ s'écrit : $\begin{cases} 42 - 5t \geq 0 \\ 17 - 2t \geq 0 \\ 24 - 3t \geq 0 \end{cases}$

La plus grande valeur possible de t est 8, et donner cette valeur à la variable x_3 annule la troisième variable de base, c'est-à-dire x_7. Cette dernière est donc la variable sortante.

5. On actualise ; on a maintenant :

$$X_B = \begin{pmatrix} x_5 \\ x_6 \\ x_3 \end{pmatrix} \quad \text{et} \quad X_B^* = \begin{pmatrix} 42 - 5\times 8 \\ 17 - 2\times 8 \\ 8 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 8 \end{pmatrix}$$

ÉTAPE 2

1. On résout $y.B = c_B$:

$$(y_1\ \ y_2\ \ y_3)\begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix} = (0\ \ 0\ \ 18)$$

d'où le système : $\begin{cases} y_1 & = 0 \\ y_2 & = 0 \\ 5y_1 + 2y_2 + 3y_3 & = 18 \end{cases}$

dont la solution est : $y = (0\ \ 0\ \ 6)$

2. $y.a_1 = 6 < c_1$, $y.a_2 = 12 > c_2$, $y.a_4 = 18 > c_4$, $y.a_7 = 6 > c_7$: seule la variable x_1 est candidate à entrer en base, c'est donc la variable entrante.

3. On résout $B.d = a_1$:

$$\begin{cases} d_1 + 5d_3 = 2 \\ d_2 + 2d_3 = 1 \\ 3d_3 = 1 \end{cases}$$

On obtient :

$$d = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}^t$$

4. $X_B^* - t.d \geq 0$ s'écrit :

$$\begin{cases} 2-\dfrac{1}{3}t \geq 0 \\ 1-\dfrac{1}{3}t \geq 0 \\ 8-\dfrac{1}{3}t \geq 0 \end{cases}$$

La plus grande valeur possible de t est 3 et donner cette valeur à la variable x_1 annule la deuxième variable de base, c'est-à-dire x_6. Cette dernière est donc la variable sortante.

5. On actualise : on a maintenant

$$X_B = \begin{pmatrix} x_5 \\ x_1 \\ x_3 \end{pmatrix} \text{ et } \quad X_B^* = \begin{pmatrix} 2-\dfrac{1}{3}\times 3 \\ 3 \\ 8-\dfrac{1}{3}\times 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix}.$$

ÉTAPE 3

1. On résout y.B = c_B :

$$(y_1 \;\; y_2 \;\; y_3) \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix} = (0 \;\; 7 \;\; 18).$$

d'où le système :

$$\begin{cases} y_1 & = 0 \\ 2y_1 + \; y_2 \; + \; y_3 & = 7 \\ 5y_1 + 2y_2 + 3y_3 & = 18 \end{cases}$$

dont la solution est :

$$y = (0 \;\; 3 \;\; 4).$$

2. $y.a_2 = 11 > c_2$, $y.a_4 = 18 > c_4$, $y.a_6 = 3 > c_6$, $y.a_7 = 4 > c_7$; aucune variable n'est entrante. La base actuelle est la base optimale.
La méthode du simplexe est terminée ; la conclusion sur la solution du problème se trouve dans le chapitre I.

I.4. Application au problème de découpe

Une entreprise achète du tissu par rouleaux de grandes longueurs et de 1 m de largeur. Elle a besoin des bandes de tissu suivantes :

97 m en 45 cm de large
610 m en 36 cm de large
395 m en 31 cm de large
211 m en 14 cm de large.

Ces bandes peuvent être morcelées dans leur longueur mais non dans leur largeur (ainsi, pour constituer la première bande, on peut mettre bout à bout 50 m en 45 cm de large et 47 m en la même largeur, mais pas une pièce de 97 m de long et de 30 cm de large avec une pièce de

même longueur et de 15 cm de large).
Il s'agit de minimiser la longueur de tissu acheté : il est clair que pour atteindre cet objectif, on s'impose de découper le tissu de façon à avoir exactement la longueur nécessaire dans chacune des largeurs.
On appelle « modèle de découpe » tout vecteur $a = (a_i)_{1 \leq i \leq 4}$ où les a_i sont des entiers positifs ou nuls mais pas tous nuls simultanément, tels que :

$$45\, a_1 + 36\, a_2 + 31\, a_3 + 14\, a_4 \leq 100.$$

Les vecteurs suivants donnent quatre modèles de découpe possibles :

$$\begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 7 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

On pourrait montrer, mais cela est sans intérêt, qu'il y a 37 modèles de découpe différents. Supposons que tous les modèles de découpe soient répertoriés et numérotés ; on note $(a_{ij})_{1 \leq i \leq 4}$ le j^{e} modèle. Les variables du problème sont les longueurs x_j de rouleau que l'on découpe suivant le modèle j. Le problème s'écrit :

$$\text{Minimiser } \sum_{j=1}^{37} x_j$$

avec les contraintes : $$\begin{cases} \sum_{j=1}^{37} a_{ij} x_j = b_i \text{ pour } i = 1, 2, 3, 4 \\ x_j \geq 0 \text{ pour } j = 1, \ldots, 37 \end{cases}$$

où $b_1 = 97$, $b_2 = 610$, $b_3 = 395$, $b_4 = 211$.
Il est difficile de traiter, en tout cas à la main, des dictionnaires aussi larges. C'est pour cela que nous allons employer la variante matricielle de la méthode du simplexe en ne calculant pas, à chaque étape, tout le dictionnaire.
Ici, la matrice A a pour colonnes les 37 modèles de découpe et on a :

$$X = \begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_{37} \end{pmatrix}, \; b = \begin{pmatrix} 97 \\ 610 \\ 395 \\ 211 \end{pmatrix} \text{ et } c = (1\ 1\ \ldots\ 1).$$

Pour déterminer une base, il suffit de trouver 4 modèles de découpe formant une matrice inversible et la base est formée des variables x correspondantes. Il faut bien entendu s'assurer, après le calcul de X_B^*, que cette base est réalisable.
Prenons :

$$B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

car cette matrice a l'avantage d'être diagonale, inversible et la base associée réalisable.

En effet, appelons X_B le vecteur-colonne formé des quatre variables de base et X_N le vecteur-colonne formé par les 33 autres variables.
Cherchons la solution basique correspondante : $X_B^* = B^{-1}.b$. Ici on a :

$$X_B^* = \left(\frac{97}{2} \quad 305 \quad \frac{395}{3} \quad \frac{211}{7}\right)^t ;$$

cette solution basique est réalisable.
Nous avons un bon point de départ pour la méthode du simplexe. La solution basique actuelle donne à z la valeur

$$\frac{21643}{42} \approx 515{,}31.$$

À chaque étape, on mémorise ici B plutôt que X_B car on n'attribue pas explicitement d'indices aux différentes variables. Les éléments que l'on actualise étape après étape sont donc B et X_B^*.

ÉTAPE 1

1. On calcule y défini par $yB = c_B$. Ici, en posant $y = (y_1, y_2, y_3, y_4)$, on trouve le système d'équations :

$$\begin{cases} 2y_1 = 1 \\ 2y_2 = 1 \\ 3y_3 = 1 \\ 7y_4 = 1 \end{cases}$$

On obtient donc :

$$y = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{7}\right).$$

2. Cherchons alors une colonne entrante, c'est-à-dire $(a_1, a_2, a_3, a_4)^t$ vérifiant :

$$\begin{cases} \frac{1}{2}a_1 + \frac{1}{2}a_2 + \frac{1}{3}a_3 + \frac{1}{7}a_4 > 1 \\ \text{avec } 45a_1 + 36a_2 + 31a_3 + 14a_4 \leq 100 \\ \text{et } a_1, a_2, a_3, a_4 \text{ entiers} \end{cases}$$

On ne risque pas de trouver un modèle qui soit déjà « dans la base » (variable correspondante dans la base) car, pour les modèles a « de base », on a, par définition de y : $y.a = 1$.
Un examen exhaustif des possibilités indique que les candidats à entrer en base sont :

$$\begin{pmatrix}1\\1\\0\\1\end{pmatrix} \qquad \begin{pmatrix}0\\2\\0\\2\end{pmatrix} \qquad \begin{pmatrix}0\\2\\0\\1\end{pmatrix} \qquad \begin{pmatrix}0\\1\\0\\4\end{pmatrix} \qquad \begin{pmatrix}0\\1\\2\\0\end{pmatrix} \qquad \begin{pmatrix}0\\1\\1\\2\end{pmatrix}$$

$$y.a = \frac{8}{7} \qquad y.a = \frac{9}{7} \qquad y.a = \frac{8}{7} \qquad y.a = \frac{15}{14} \qquad y.a = \frac{7}{6} \qquad y.a = \frac{47}{42}$$

N'importe laquelle des variables correspondantes est candidate à entrer dans la base. Malgré tout, de manière à choisir la variable qui, dans le dictionnaire sous-jacent actuel, a le plus petit coefficient (ce qui correspond à un des critères possibles de choix puisqu'on cherche à minimiser), on peut choisir a de façon à ce que $1 - y.a$ soit le plus petit possible, c'est-à-dire de façon à ce que

$$\frac{1}{2}a_1 + \frac{1}{2}a_2 + \frac{1}{3}a_3 + \frac{1}{7}a_4$$

soit le plus grand possible. Le problème général correspondant s'appelle le problème du « sac à dos ». Ici, ayant déjà établi la liste exhaustive des candidats, il est facile de constater que le « meilleur » des modèles est $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix}$.

3. On détermine d par $B.d = a$. En notant d_1, d_2, d_3, d_4 les composantes de d, il vient :

$$\begin{cases} 2d_1 = 0 \\ 2d_2 = 2 \\ 3d_3 = 0 \\ 7d_4 = 2 \end{cases} \qquad \text{d'où} \qquad d = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \frac{2}{7} \end{pmatrix}$$

4. On cherche la variable sortante. Lorsque toutes les variables hors-base sauf x sont nulles, X_B s'exprime de la façon suivante :

$$X_B = \begin{pmatrix} \frac{97}{2} \\ 305 - x \\ \frac{395}{3} \\ \frac{211}{7} - \frac{2}{7}x \end{pmatrix}$$

d'où $x \leq \frac{211}{2}$, et quand $x = \frac{211}{2}$, c'est la quatrième variable de l'ancienne base qui est nulle, et qui donc sort de la base.

5. On actualise B et X_B^* :

$$B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \text{et} \quad X_B^* = \begin{pmatrix} \frac{97}{2} \\ 305 - \frac{211}{2} \\ \frac{395}{3} \\ \frac{211}{7} \end{pmatrix} = \begin{pmatrix} \frac{97}{2} \\ \frac{399}{2} \\ \frac{395}{3} \\ \frac{211}{7} \end{pmatrix}$$

La valeur de z correspondante vaut $\frac{2911}{6} \approx 485{,}17$.

ÉTAPE 2

1. On cherche y tel que $y.B = (1\ \ 1\ \ 1\ \ 1)$:

$$\begin{cases} 2y_1 = 1 \\ 2y_2 = 1 \\ 3y_3 = 1 \\ 2y_2 + 2y_4 = 1 \end{cases} \quad \text{d'où} \quad y = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{3}, 0\right)$$

2. On cherche un modèle a tel que $y.a > 1$ c'est-à-dire :

$$\frac{1}{2} a_1 + \frac{1}{2} a_2 + \frac{1}{3} a_3 > 1$$

avec bien sûr : $45\ a_1 + 36\ a_2 + 31 a_3 + 14\ a_4 \leq 100$.

Si $a_1 = 2$: $a_2 = a_3 = a_4 = 0$ et $y.a = 1$.
Si $a_1 = 1$ et $a_2 = 1$: $a_3 = 0$ et $y.a = 1$.
Si $a_1 = 1$ et $a_2 = 0$: $a_3 \leq 1$ et $y.a \leq 1$.
Si $a_1 = 0$ et $a_2 = 2$: $a_3 = 0$ et $y.a = 1$.
Si $a_1 = 0$ et $a_2 = 1$, on a un modèle pour lequel $y.a > 1$: $(0\ \ 1\ \ 2\ \ 0)^t$
Si $a_1 = 0$ et $a_2 = 0$: $a_3 \leq 3$ et $y.a \leq 1$.

On fait entrer dans la base la variable correspondant au modèle $a = (0\ \ 1\ \ 2\ \ 0)^t$.

3. On cherche d tel que $B.d = a$:

$$\begin{cases} 2d_1 = 0 \\ 2d_2 + 2d_4 = 1 \\ 3d_3 = 2 \\ 2d_4 = 0 \end{cases} \quad \text{d'où} \quad d = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{2}{3} \\ 0 \end{pmatrix}$$

4. On cherche la variable sortante. Quand toutes les variables hors-base sont nulles sauf la variable x correspondant à a, on obtient :

$$X_B = \begin{pmatrix} \frac{97}{2} \\ \frac{399}{2} - \frac{1}{2} x \\ \frac{395}{3} - \frac{2}{3} x \\ \frac{211}{2} \end{pmatrix}$$

d'où $x \leq \frac{395}{2}$ et quand $x = \frac{395}{2}$, c'est la troisième variable de base qui s'annule et donc sort de la base.

5. On réactualise : $$B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{et} \quad X_B^* = \begin{pmatrix} 48{,}5 \\ 100{,}75 \\ 197{,}5 \\ 105{,}5 \end{pmatrix}$$

La valeur de z correspondant à cette nouvelle solution basique est 452,25.

ÉTAPE 3

1. On cherche y tel que $y.B = (1, 1, 1, 1)$:

$$\begin{cases} 2y_1 = 1 \\ 2y_2 = 1 \\ y_2 + 2y_3 = 1 \\ 2y_2 + 2y_4 = 1 \end{cases} \quad \text{et} \quad y = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{4}, 0\right)$$

2. On cherche une colonne entrante : $y.a = \frac{1}{2}a_1 + \frac{1}{2}a_2 + \frac{1}{4}a_3$.

L'étude faite à l'étape précédente montre que, pour tout modèle, on a $y.a \leq 1$. Dans le dictionnaire actuel, tous les coefficients des variables hors-base dans la ligne exprimant z sont positifs ou nuls ; la solution basique actuelle est optimale.

CONCLUSION

Il sera acheté 452,25 m de tissu. Le tissu sera utilisé de la façon suivante :

- dans 48,5 m seront découpées deux bandes de 45 cm de large ;
- dans 100,75 m seront découpées deux bandes de 36 cm de large ;
- dans 197,5 m seront découpées une bande de 36 cm de large et deux bandes de 31 cm de large ;
- dans 105,5 m seront découpées deux bandes de 36 cm de large et deux bandes de 14 cm de large.

II.5. Exercice

Résoudre le problème de découpe pour des rouleaux de tissu de 100 cm de large lorsque l'on a besoin de :

- 600 m de largeur 52 cm ;
- 600 m de largeur 29 cm ;
- 600 m de largeur 27 cm ;
- 1200 m de largeur 21 cm.

III. Dualité

III.1. Définition du problème dual

REMARQUE

Nous ne considérons dans ce chapitre que les problèmes de programmation linéaire écrits sous forme standard. On pourrait néanmoins définir le problème dual de tout problème de programmation linéaire. Cette généralisation, facile à effectuer grâce aux indications données au chapitre I à propos de la forme standard, est laissée au lecteur.

On considère donc le problème **(*P*)** :

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

sous les contraintes $\begin{cases} \sum_{j=1}^{n} a_{ij} x_j \le b_i \text{ pour } i = 1,\dots,m \\ x_j \ge 0 \text{ pour } j = 1,\dots,n \end{cases}$

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

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

D'où :

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

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

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

$$\text{Minimiser } \sum_{i=1}^{m} b_i y_i$$

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

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.

III.2. Théorème de la dualité

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

Proposition. *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 :

$$\text{Maximiser } z = 7x_1 + 9x_2 + 18x_3 + 17x_4$$

avec les contraintes $\begin{cases} 2x_1 + 4x_2 + 5x_3 + 7x_4 \leq 42 \\ x_1 + x_2 + 2x_3 + 2x_4 \leq 17 \\ x_1 + 2x_2 + 3x_3 + 3x_4 \leq 24 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0 \end{cases}$

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 :

$$\text{Minimiser } 42y_1 + 17y_2 + 24y_3$$

avec les contraintes
$$\begin{cases} 2y_1 + \ y_2 + \ \ y_3 \geq 7 \\ 4y_1 + \ y_2 + 2y_3 \geq 9 \\ 5y_1 + 2y_2 + 3y_3 \geq 18 \\ 7y_1 + 2y_2 + 3y_3 \geq 17 \\ y_1 \geq 0, y_2 \geq 0, y_3 \geq 0 \end{cases}$$

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 durant la fin de ce paragraphe.

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.

Théorème de la dualité. *Si le problème primal a une solution optimale $x_1^*, x_2^*, \ldots, x_n^*$, alors le problème dual a une solution optimale $y_1^*, y_2^*, \ldots, y_m^*$ et* $\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^*$.

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

Proposition. *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 donc ajouté m variables d'écart $x_{n+1} \ldots x_{n+m}$. 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 $z^* = \sum_{j=1}^{n} c_j x_j^*$ et $x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} x_j$ par définition des variables d'écart.

Posons : $y_i^* = -\, d_{n+i} \geq 0$ pour $i = 1, ..., m$. On a :

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

et aussi : $z = \sum_{j=1}^{n} c_j x_j$

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

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

Les d_j ($j = 1, ..., n + m$) étant négatifs ou nuls, on obtient :

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

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. □

REMARQUE

Supposons que l'on applique la méthode matricielle du simplexe à un problème de programmation linéaire mis sous forme standard et reprenons les notations du chapitre précédent. Nous avons constaté que le coefficient de la variable x_j ($j = 1, ..., n + m$) dans l'expression de la fonction objectif en fonction des variables hors-base de la dernière étape est donné par $c_j - \sum_{i=1}^{m} a_{ij} y_i$, où y est la solution du système $y.B = c_B$. Si x_j est une variable d'écart, alors $c_j = 0$, $a_{ij} = 0$ si $i \neq j$ *et* $a_{ii} = 1$. Le coefficient de x_{n+i} ($i = 1, ..., m$) est donc égal à $-y_i$; la proposition ci-dessus permet d'affirmer que le vecteur y déterminé dans cette dernière étape n'est donc rien d'autre que la solution optimum du problème dual.

III.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 pas 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 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* tels que :*

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

et tels que $\begin{cases} \sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \ \ pour\ j = 1, \dots, n \\ y_i^* \geq 0 \ pour\ i = 1, \dots, m \end{cases}$

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 requises, en effet :

$$4y_1* + y_2* + 2y_3* = 11 > 9$$

et
$$7y_1* + 2y_2* + 3y_3* = 18 > 17$$

Les deux autres inégalités du même type résultent du système vérifié par y_1*, y_2* et y_3*. Enfin y_1*, y_2*, y_3* sont non négatifs.
La solution proposée pour le primal est bien optimale.

Preuve du théorème des écarts complémentaires.— Elle 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 :

$$\left(x_j^* = 0 \text{ ou } \sum_{i=1}^{m} a_{ij} y_i^* = c_j\right) \text{ et } \left(y_i^* = 0 \text{ ou } \sum_{i=1}^{m} a_{ij} x_j^* = b_i\right)$$

En effet, d'après le théorème de 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 :

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

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

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

On voit donc qu'il y a égalité entre les bornes de cette suite si et seulement si :

$$\left(x_j^* = 0 \text{ ou } \sum_{i=1}^{m} a_{ij} y_i^* = c_j\right) \text{ et } \left(y_i^* = 0 \text{ ou } \sum_{i=1}^{m} a_{ij} x_j^* = b_i\right)$$

En effet, si, avant sommation, une des inégalités était stricte, il en serait de même après sommation. □

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

III.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 le nombre d'unités de la ressource i consommées par la fabrication d'une unité de produit j ;
- x_j représente le nombre d'unités de produit j fabriquées ;
- c_j représente la valeur unitaire du produit j.

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

Les relations $\sum_{i=1}^{m} a_{ij} y_i \geq c_j$, ($j$ = 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.

PROBLEME

Le fabricant d'objets de terre cuite a la possibilité de faire faire à ses ouvriers spécialisés dans la peinture quelques heures supplémentaires à un prix horaire t. 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. *On considère le problème* ***(P)***

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

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

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

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

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

On suppose que les variations δb_i *sont suffisamment faibles pour que la base optimale de* ***(P)*** *soit encore réalisable pour* ***(Pδ)***. *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 peinture (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 tu euros.
On voit qu'il a intérêt à recourir à cette solution dès que $t \leq 4$ €/heure (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.— Nous reprenons pour cette preuve les notations du chapitre sur la forme matricielle de la méthode du simplexe. Nous avons remarqué plus haut que la solution du problème dual est donnée par le vecteur y vérifiant $y.B = c_B$ dans la dernière étape de la méthode matricielle. La valeur du produit scalaire $y.b$ n'est autre que $c_B.B^{-1}.b$, c'est-à-dire la valeur à l'optimum de la fonction objectif du primal.
Supposons que b soit transformé en $b + \delta b$, cette variation étant suffisamment petite pour que la base déterminée à la dernière étape reste une base réalisable. L'expression de X_B^* où la matrice B est inchangée nous donne la limite sur les variations admissibles de b. Aucun coefficient n'étant changé dans la matrice A ni dans le vecteur c, la base optimale pour **(*P*)** l'est aussi pour **(*Pδ*)** et la valeur optimale de **(*Pδ*)** est $c_B.B^{-1}.(b + \delta b) = c_B.B^{-1}.b + c_B.B^{-1}.\delta b$: le théorème est ainsi démontré. □

III.5. Problème dual-réalisable

L'utilisation du 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. Un tel problème est dit *dual-réalisable*.

EXEMPLE

Considérons le problème de programmation linéaire

$$\text{Minimiser } x_1 + x_2$$

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 :

$$\text{Maximiser } -x_1 - x_2$$

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 :

$$\text{Minimiser } -4y_1 + 7y_2$$

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

Maximiser $4y_1 - 7y_2$

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.

III.6. Exercices

Exercice 1

On propose la solution $x_1^* = 0$; $x_2^* = \frac{4}{3}$; $x_3^* = \frac{2}{3}$; $x_4^* = \frac{5}{3}$; $x_5^* = 0$ pour le problème

Maximiser $7x_1 + 6x_2 + 5x_3 - 2x_4 + 3x_5$

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 ?

Exercice 2

Un fabricant de fils téléphoniques produit trois types de fils (F_1, F_2 et F_3) de différentes sections et obtenus à l'aide de cuivre enrichi de cadmium pour F_1 et F_2 (en proportions différentes) ou de cadmium et d'étain pour F_3. Le tableau suivant donne la masse de cuivre (exprimée en kilogrammes) et celles de cadmium et d'étain (exprimées en décagrammes) nécessaires pour fabriquer 100 mètres de chacun de ces fils téléphoniques.

	F_1	F_2	F_3
cuivre	9	5	6
cadmium	2	1	2
étain	0	0	1

L'entreprise dispose de 600 kilogrammes de cuivre, 150 décagrammes de cadmium et 60 décagrammes d'étain. Elle emploie en outre des ouvriers pour couler puis tréfiler les alliages. Il faut une journée de travail pour fabriquer 100 mètres de fil F_1, F_2 ou F_3. La force de travail disponible pour la production des fils s'élève à 90 jours. Enfin, les profits relatifs à la fabrication de 100 mètres de fil s'élèvent à 4200 €, 3900 € et 5200 € respectivement pour F_1, F_2 et F_3.

a. Écrire le problème de maximisation du profit de l'entreprise puis résoudre celui-ci à l'aide de l'algorithme du simplexe (si plusieurs variables sont candidates pour entrer dans la base ou pour en sortir, on choisira celle de plus grand indice pour entrer en base et celle de plus petit indice pour en sortir).

b. Écrire le problème dual du problème précédent. En donner une solution.

c. Le profit sur le fil F_1 a été mal évalué. La solution optimale est-elle modifiée si celui-ci passe de 4200 € à 4800 € (pour 100 mètres) ? s'il passe de 4200 € à 5400 € (toujours pour 100 mètres) ? Dans chaque cas, on donnera, le cas échéant, la nouvelle solution optimale (ici, en cas de choix, on fera entrer ou sortir la variable ayant le plus grand indice).

d. De même, à cause des fluctuations sur le marché de l'étain, le profit associé à F_3 varie. En supposant le profit associé à F_1 de nouveau égal à 4200 € à l'hectomètre, montrer que la solution trouvée à la première question reste optimale si le profit associé à F_3 est compris entre 4200 € et 7800 € inclus.

e. Pour cette question, on suppose que le profit associé au fil F_1 vaut 5400 €, comme il était envisagé dans la question c. On peut de plus obtenir de l'étain supplémentaire au prix de 100 € le décagramme ; on peut aussi, moyennant une prime de 2000 € par jour supplémentaire, augmenter le nombre de jours que les ouvriers consacrent à la production des fils. Conseilleriez-vous d'avoir recours à l'une ou l'autre de ces possibilités, ou les deux ? On donnera, le cas échéant, une indication sur la quantité d'étain à acheter ou sur le nombre (éventuellement fractionnaire) de jours supplémentaires.

f. On revient aux données initiales. En fait, l'entreprise n'achète pas d'étain et ne fait pas faire de jour supplémentaire, mais envisage de produire un nouveau fil, un peu plus gros, nécessitant 8 kilogrammes de cuivre et 2 décagrammes de cadmium pour 100 mètres (et pas d'étain). Comme pour les autres fils, il faut un jour de travail pour en produire 100 mètres. Le profit vaut 5500 € à l'hectomètre. Est-il intéressant de fabriquer ce fil ? Si oui, quel nouveau plan de fabrication faut-il envisager ?

g. Finalement, l'entreprise ne produit pas le fil de la question précédente, mais décide de protéger les fils F_2 et F_3 en les gainant avec du polythène. Il faut 100 mètres de polythène pour gainer 100 mètres des fils F_2 ou F_3. Sachant qu'on dispose de 80 hectomètres de gaine, quelle est la nouvelle solution optimale ?

IV. Relaxation lagrangienne

Nous allons dans ce chapitre examiner une partie de ce qu'on appelle les méthodes de relaxation lagrangienne ; ces techniques sont très utilisées pour la résolution de problèmes de programmation linéaire en nombres entiers de grande dimension.

IV.1. Position du problème

On considère un ensemble fini S et un ensemble de fonctions réelles $f(x)$, $g_1(x)$, ..., $g_m(x)$ définies sur S. On s'intéresse alors à un problème (P) du type :

$$\text{Minimiser } f(x)$$

avec les contraintes $\begin{cases} g_i(x) \leq 0 \text{ pour } i = 1, 2, ..., m \\ x \in S \end{cases}$

c'est-à-dire à la recherche, parmi les éléments x de S qui donnent à tous les $g_i(x)$ ($1 \leq i \leq m$) des valeurs négatives ou nulles, d'un élément x^* qui rende $f(x)$ le plus petit possible ; si ce minimum existe, il sera alors noté f^*.

Si le cardinal de S est petit, il suffit de passer en revue tous ses éléments : le problème est alors élémentaire. Nous nous intéressons ici au cas où le cardinal de S est trop grand pour qu'il soit réaliste de considérer successivement tous ses éléments.

EXEMPLE DU PLUS COURT CHEMIN A UNE CONTRAINTE

Soit $G = (X, U, c, t)$ un graphe orienté doublement valué où c et t sont des fonctions de U dans $\mathbb{R}^+$; pour tout arc $u \in U$, $c(u)$ représente le coût à payer pour utiliser l'arc u et $t(u)$ le temps nécessaire pour le parcourir (temps de transmission). Étant donnés deux sommets a et b, on cherche le chemin de moindre coût de a à b de telle sorte que le temps total de transmission ne dépasse pas un temps limite T. À un chemin C de G d'origine a et d'extrémité b, on associe la fonction x définie de U dans $\{0, 1\}$ par :

$$\begin{cases} x(u) = 1 \text{ si } u \text{ appartient à } C \\ x(u) = 0 \text{ sinon} \end{cases}$$

On note S l'ensemble des fonctions ainsi obtenues (par abus de langage, un élément x de S est appelé chemin de *a* à *b*). Notre problème peut s'écrire :

$$\text{Minimiser } \sum_{u\in U} c(u)x(u)$$

avec les contraintes $\begin{cases} \sum_{u\in U} t(u)x(u) \leq T \\ x \in S \end{cases}$

La méthode de relaxation lagrangienne que nous allons utiliser donne une minoration de ce minimum.

Définitions. *1. On appelle fonction de Lagrange associée au problème (P) la fonction L définie sur S* ∞ *$(R^+)^m$ par :*

$$L(x, \Lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

avec $\Lambda = (\lambda_1, ..., \lambda_m)$ et $\lambda_i \geq 0$ pour tout i dans {1, ..., m}.

2. Les λ_i sont appelés multiplicateurs de Lagrange.

3. On appelle fonction duale la fonction w définie sur $(R^+)^m$ par :

$$w(\Lambda) = \operatorname{Min}_{x\in S} L(x, \Lambda)$$

Dans notre exemple, on obtient :

$$w(\lambda) = \operatorname{Min}_{x\in S} \left\{ \sum_{u\in U} c(u)x(u) + \lambda \left[\sum_{u\in U} t(u)x(u) - T \right] \right\}$$

avec λ réel positif ou nul. Le multiplicateur de Lagrange λ peut s'interpréter comme un poids que l'on affecte à la contrainte afin de pénaliser les chemins qui ne la respectent pas en augmentant leur coût. En faisant varier λ, on accorde une importance plus ou moins grande à la contrainte : la recherche de $w(0)$ est celle du plus court chemin au sens des coûts, et lorsque λ tend vers l'infini, le problème devient celui d'un plus court chemin au sens des temps.

Théorème. *Si $\Lambda \in (\mathbb{R}^+)^m$ et si x est un élément de S qui vérifie les contraintes de (P), alors on a $w(\Lambda) \leq f(x)$.*

Preuve.— Par définition de la fonction duale, on a, pour tout $\Lambda \in (\mathbb{R}^+)^m$ et tout $x \in S$:

$$w(\Lambda) \leq L(x, \Lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

Or, x étant supposé réalisable, $g_i(x) \leq 0$ pour tout i ; comme de plus $\lambda_i \geq 0$ aussi pour tout i, on a bien $w(\Lambda) \leq f(x)$. □

Définitions. *Le problème dual (D) du problème (P) consiste à maximiser $w(\Lambda)$ pour $\Lambda \in (\mathbb{R}^+)^m$. (P) s'appelle alors le problème primal.*

Du théorème précédent, on déduit que, lorsqu'on résout (*D*), le maximum obtenu est un minorant du minimum du problème (*P*). Le théorème suivant indique quelle est l'allure de *w*.

Théorème. *La fonction w est concave, affine par morceaux.*

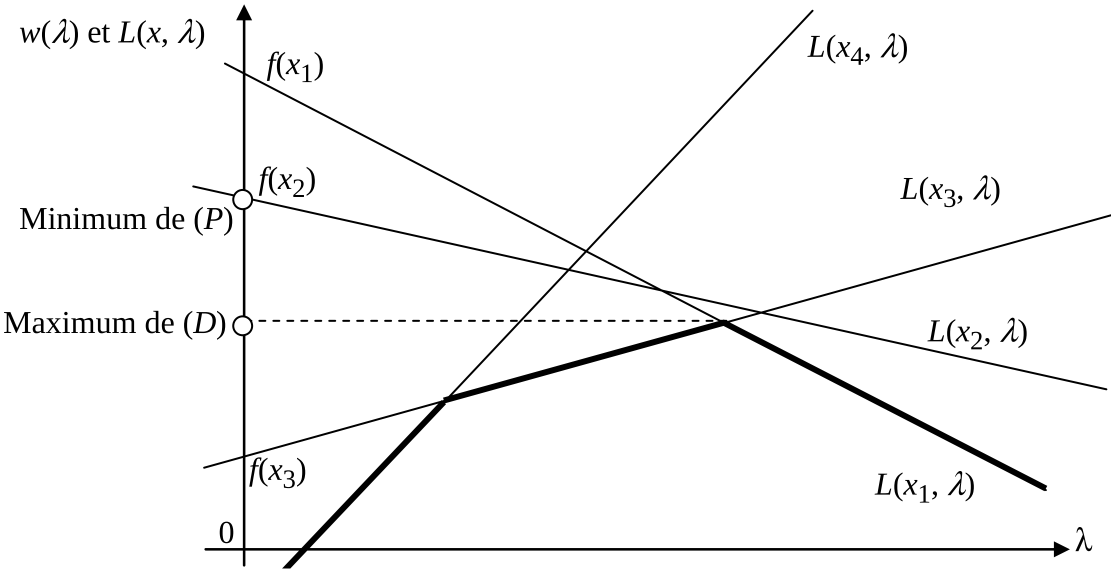
Preuve.— Chaque fonction $\Lambda \to L(x, \Lambda)$ est une fonction affine, donc concave. Or, le minimum d'une famille finie de fonctions affines (respectivement concaves) est affine par morceaux (respectivement concave). □

REMARQUE

La concavité de w permet éventuellement de lui appliquer les techniques d'optimisation des fonctions concaves.

EXEMPLE

Si $m = 1$ et si $|S| = 4$, les fonctions $\lambda \to L(x, \lambda)$ et $\lambda \to w(\lambda)$ peuvent se présenter comme ci-dessous où $w(\lambda)$ est représentée en traits gras.



Pour cet exemple, les solutions réalisables sont x_1 et x_2.

IV.2. Résolution du problème dual

IV.2.1. Principe de la méthode

Il s'agit de trouver : $\displaystyle \operatorname*{Max}_{\Lambda \in (R^+)^m} w(\Lambda) = \operatorname*{Max}_{\Lambda \in (R^+)^m} \left[\operatorname*{Min}_{x \in S} L(x, \Lambda) \right]$

Notons w^* ce maximum. On verra plus loin que le problème ci-dessus est un problème d'optimisation linéaire ayant autant de contraintes que d'éléments dans S ; la taille de S en interdit la résolution directe.

On commence par choisir un « petit » nombre de points de S : x_1, ..., x_{k-1}. Une étape de la méthode se déroule comme décrit ci-dessous.

1. On résout par programmation linéaire le problème (D_{k-1}) défini par :

$$\underset{\Lambda \in (\mathrm{R}^+)^m}{\text{Maximiser}}\, w_k(\Lambda)$$

avec $$w_k(\Lambda) = \operatorname{Min}\left[L(x_1, \Lambda), \ldots, L(x_{k-1}, \Lambda)\right].$$

Le maximum est atteint pour $\Lambda = \Lambda_k$ et prend une valeur w_k^*.

2. On calcule alors :

$$w(\Lambda_k) = \underset{x \in S}{\operatorname{Min}}\, L(x, \Lambda_k)$$

et on note x_k un élément de S qui atteint ce minimum. On remarque ici qu'il est essentiel que l'on puisse résoudre (en un temps raisonnable) ce dernier problème. Il arrive souvent que celui-ci se résolve de la même façon que celui de l'optimisation de la fonction f lorsqu'on n'a pas de contrainte.

3. On conclut cette étape. On a :

a. $w_k^* \geq w^*$ puisque ce sont les maxima de deux fonctions telles que la première fonction soit pour toute valeur de Λ au moins égale à la seconde. En effet, la famille des fonctions $\Lambda \to L(x_1, \Lambda)$, …, $\Lambda \to L(x_{k-1}, \Lambda)$ est incluse dans la famille des fonctions $\Lambda \to L(x, \Lambda)$ quand x décrit S et donc, le minimum d'une fonction sur un ensemble Ω étant supérieur ou égal au minimum de la même fonction sur tout ensemble contenant Ω, on a $\operatorname{Min}\left[L(x_1, \Lambda), \ldots, L(x_{k-1}, \Lambda)\right] \geq \underset{x \in S}{\operatorname{Min}}\, L(x, \Lambda)$ pour tout Λ.

b. $w_k^* \geq w(\Lambda_k)$ puisque w_k^* est le maximum de w.

Si l'égalité $w_k^* = w(\Lambda_k)$ est vérifiée, alors $w_k^* = w^* = w(\Lambda_k)$ et le problème dual est résolu. Sinon, on continue en considérant le problème (D_k), c'est-à-dire en résolvant :

$$\underset{\Lambda \in (\mathrm{R}^+)^m}{\text{Maximiser}}\left\{\operatorname{Min}\left[L(x_1, \Lambda), \ldots, L(x_k, \Lambda)\right]\right\}$$

Autrement dit, on passe à l'étape suivante en tenant compte d'un point x_k supplémentaire.

On peut remarquer que la suite w_k^* est décroissante, mais qu'en revanche la suite $w(\Lambda_k)$ n'est pas nécessairement croissante ; on a alors plus précisément à chaque étape :

$$\underset{i \in \{1,2,..,k\}}{\operatorname{Max}}\, w(\Lambda_i) \leq w^* \leq w_k^*$$

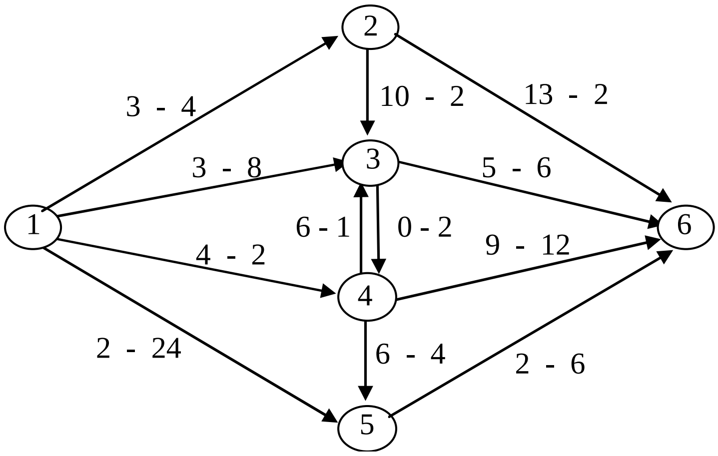
Par ailleurs, en ce qui concerne le problème primal (P), si on note f^* le minimum de (P) et f_k^* celui de la famille $\{f(x_i) \,/\, 1 \leq i \leq k$ et x_i vérifie les contraintes de $(P)\}$, il vient :

$$\underset{i \in \{1, \ldots, k\}}{\operatorname{Max}}\, w(\Lambda_i) \leq w^* \leq f^* \leq f_k^*.$$

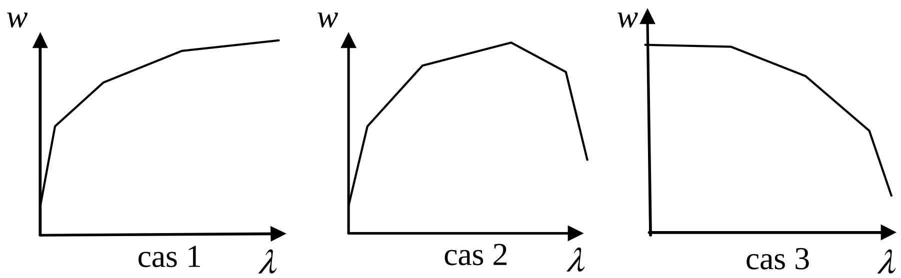
On arrête la résolution du problème dual soit quand on a obtenu l'optimum, c'est-à-dire si on constate que $\underset{i \in \{1, \ldots, k\}}{\operatorname{Max}}\, w(\Lambda_i) = w_k^*$, soit sur un critère d'arrêt à définir ; ce critère peut être une borne sur le nombre total d'étapes, cela peut aussi être que l'encadrement de w^* donne la valeur de w^* à moins de ε près, où ε a été fixé.

IV.2.2. Étude du problème de chemin de coût minimum à une contrainte

Reprenons l'exemple cité plus haut et considérons le graphe suivant, où sur chaque arc sont indiqués coût (à gauche) et temps de transmission (à droite). On cherche le plus court chemin du sommet $a = 1$ au sommet $b = 6$ de temps de transmission au plus $T = 10$.



La fonction $w(\lambda)$ présente une des trois formes suivantes, puisqu'elle est concave et affine par morceaux :



Dans le premier cas, w n'est pas bornée ; le problème n'admet aucune solution réalisable, d'après le premier théorème. Dans le deuxième cas, w est bornée et le plus court chemin (au sens des coûts) n'est pas réalisable en ce qui concerne le temps de transmission. Dans le troisième cas, le maximum de $w(\lambda)$ est : $w(0) = \underset{x \in S}{\text{Min}}\, L(x,0) = \underset{x \in S}{\text{Min}} \sum_{u \in U} c(u)x(u)$.

Dans ce dernier cas, le maximum de w est égal au coût d'un plus court chemin au sens des coûts.

Nous allons appliquer la méthode générale expliquée dans le paragraphe précédent pour résoudre le problème dual. Avant de débuter, voyons s'il est possible de résoudre la partie 2 de la méthode. Il s'agit de déterminer, avec λ_k fixé, $w(\lambda_k) = \underset{x \in S}{\text{Min}}\, L(x, \lambda_k)$. On a :

$$w(\lambda_k) = \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} c(u)x(u) + \lambda_k \left[\sum_{u \in U} t(u)x(u) - T \right] \right\}$$

$$= \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} [c(u) + \lambda_k . t(u)] x(u) \right\} - \lambda_k T$$

Résoudre ce problème revient à chercher le plus court chemin de a à b avec comme valuation sur chaque arc la quantité $c(u) + \lambda_k.t(u)$, problème que l'on sait résoudre avec un algorithme tel que l'algorithme de Dijkstra, puisque les valuations sont toutes positives ou nulles. On a :

$$L(x, \lambda) = \sum_{u \in U} c(u)x(u) + \lambda \left[\sum_{u \in U} t(u)x(u) - T \right]$$

et $$w(\lambda) = \operatorname*{Min}_{x \in S} L(x, \lambda).$$

Notons, pour $x \in S$, $\Delta(x)$ la droite représentant la fonction $\lambda \rightarrow L(x, \lambda)$. La courbe $\lambda \rightarrow w(\lambda)$ est l'enveloppe inférieure de ces droites.

Nous choisissons d'abord deux chemins x_1 et x_2 comme indiqué ci-dessous.

Soit x_1 le plus court chemin au sens des coûts de (1) à (6). D'une manière générale, l'intersection de $\Delta(x)$ avec l'axe *Oy* donne le coût de *x* ; la droite $\Delta(x)$ qui coupe *Oy* « le plus bas » est $\Delta(x_1)$; pour λ assez petit, $w(\lambda)$ est confondue avec $L(x_1, \lambda)$.

Lorsque λ tend vers l'infini, $w(\lambda)$ est donné par la droite de plus petite pente ; or, d'une manière générale, la pente de $\Delta(x)$ est le temps de transmission de *x* moins *T* (elle indique donc si *x* est réalisable ou non) ; il est intéressant de choisir comme chemin x_2 le plus court chemin au sens des temps : c'est avec $L(x_2, \lambda)$ que se confond $w(\lambda)$ pour λ assez grand.

Connaissant $\Delta(x_1)$ et $\Delta(x_2)$, on sait dans laquelle des trois situations précédentes on se trouve.

DETERMINATION DE $\Delta(x_1)$

Appliquons un algorithme de plus court chemin : on trouve que le chemin x_1 est composé des arcs (1, 5) et (5, 6). Son coût vaut $\sum_{u \in U} c(u)x_1(u) = 4$ et son temps de transmission $\sum_{u \in U} t(u)x_1(u) = 30$; $\Delta(x_1)$ est la droite représentant la fonction $\lambda \rightarrow 4 + 20\,\lambda$; la pente est positive : on est dans le cas 1 ou 2.

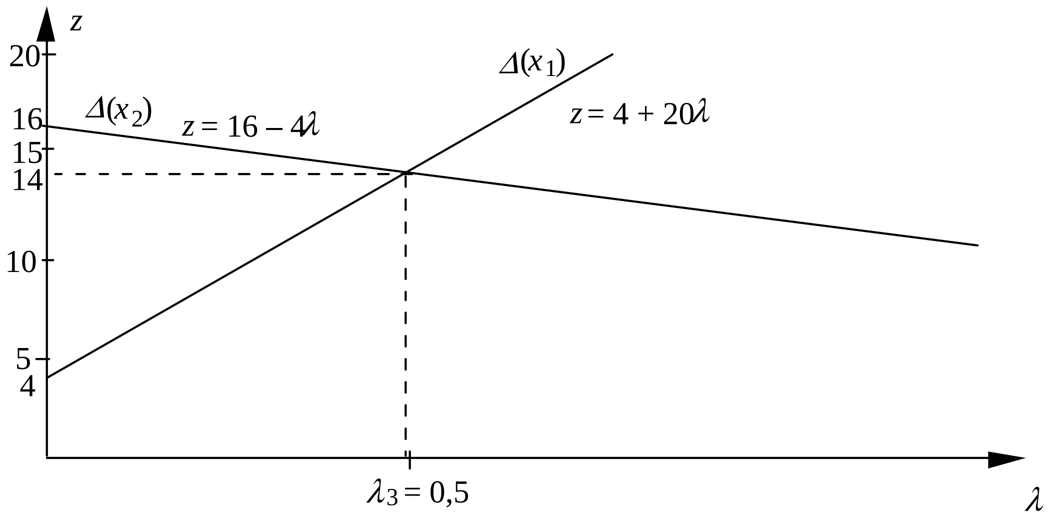
DETERMINATION DE $\Delta(x_2)$

Au sens des temps, le plus court chemin est composé des arcs (1, 2) et (2, 6), et est de coût $\sum_{u \in U} c(u)x_2(u) = 16$ et de temps de transmission $\sum_{u \in U} t(u)x_2(u) = 6$. Ce chemin est réalisable (si ce n'était pas le cas, aucun chemin ne le serait). On peut donc ici remarquer l'encadrement $4 \leq f^* \leq 16$. La droite $\Delta(x_2)$ correspond à la fonction $\lambda \rightarrow 16 - 4\,\lambda$; la pente est négative : on est dans le cas 2.

PREMIERE ETAPE

1. Résolution de : Maximiser $\text{Min}\{L(x_1, \lambda), L(x_2, \lambda)\}$ pour $\lambda \geq 0$.

Pour résoudre ce problème, on peut appliquer la méthode évoquée plus haut et décrite plus loin consistant à le formuler comme un problème de programmation linéaire. Comme il n'y a qu'une contrainte, on peut aussi effectuer une résolution graphique ; c'est ce que nous ferons.



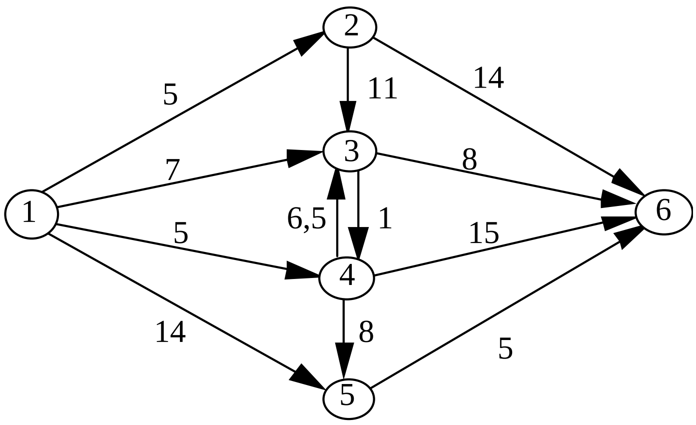
$\Delta(x_1)$ et $\Delta(x_2)$ se coupent pour λ tel que : $4 + 20\lambda = 16 - 4\lambda$ d'où $\lambda = 0{,}5$. Avec les notations définies dans la section précédente, cela s'écrit : $\lambda_3 = 0{,}5$ et le maximum $w_3^* = 14$.

2. Résolution de : Minimiser $L(x, 0{,}5)$ pour $x \in S$.
On cherche :

$$w(0{,}5) = \underset{x \in S}{\text{Min}}\ L(x, 0{,}5) = \underset{x \in S}{\text{Min}} \left\{ \sum_{u \in U} [c(u) + 0{,}5t(u)]x(u) \right\} - 0{,}5T \ .$$

On cherche pour cela le plus court chemin dans le graphe ci-contre, valué par $c(u) + 0{,}5.t(u)$. On trouve comme plus court chemin x_3 le chemin composé des arcs (1, 3) et (3, 6) ; on en déduit :

$$w(0{,}5) = 15 - 5 = 10.$$



3. Conclusion de l'étape.

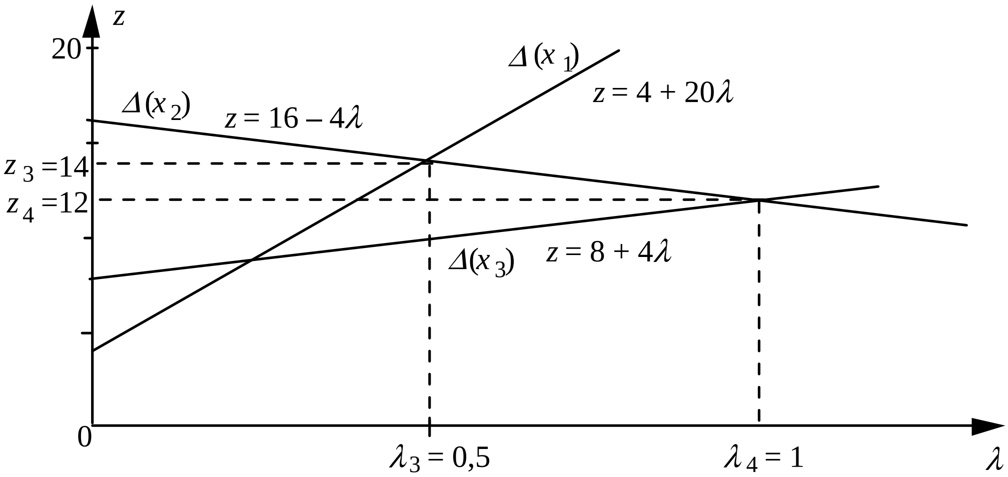
On a : $w(0{,}5) = 10 < 14 = w_3^*$. D'où : $10 \le w^* \le 14$.

On peut remarquer que le temps de transmission de x_3 est 14 : x_3 n'est pas réalisable. On a donc actuellement obtenu, en ce qui concerne le problème primal : $10 \le f^* \le 16$.

Nous allons passer à l'étape suivante en ajoutant la droite $\Delta(x_3)$ représentant la fonction : $\lambda \to 8 + 4\,\lambda$.

DEUXIEME ETAPE

1. Résolution de : Maximiser $\text{Min}\{L(x_1, \lambda), L(x_2, \lambda), L(x_3, \lambda)\}$ pour $\lambda \ge 0$.

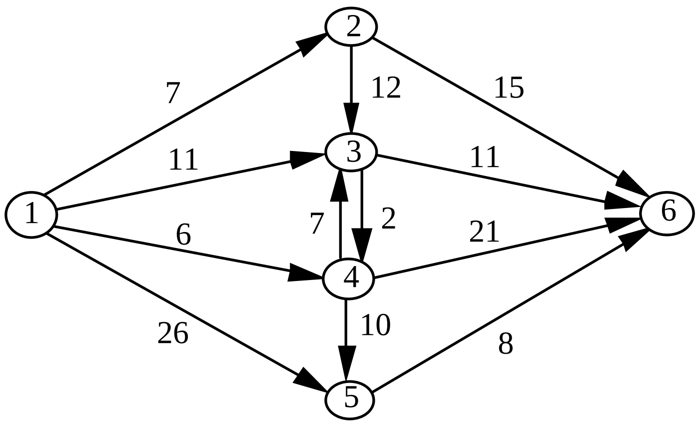


Le maximum est obtenu à l'intersection de $\Delta(x_2)$ et $\Delta(x_3)$ pour une valeur de λ, notée λ_4, qui vérifie $8 + 4\lambda = 16 - 4\lambda$: $\lambda_4 = 1$ et le maximum vaut $w_4^* = 12$.

2. Résolution de : Minimiser $L(x, 1)$ pour $x \in S$.
On cherche :

$$w(1) = \underset{x \in S}{\mathrm{Min}}\ L(x,1) = \underset{x \in S}{\mathrm{Min}}\left\{\sum_{u \in U}[c(u) + t(u)]x(u)\right\} - T \ .$$

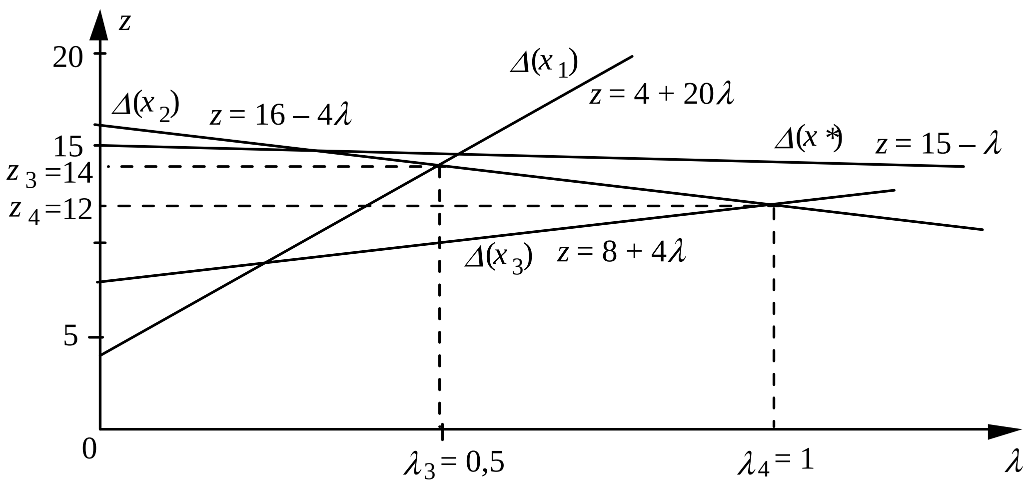
On cherche pour cela le plus court chemin dans le graphe valué par $c(u) + t(u)$, représenté ci-dessous. Il s'agit du chemin composé des arcs (1, 3) et (3, 6) ou du chemin composé des arcs (1, 2) et (2, 6) ; d'où $w(\lambda_4) = 22 - 10 = 12$.



3. Conclusion de l'étape.

Puisque $w(\lambda_4) = w_4^*$, la résolution du problème dual est terminée. La fonction duale est maximum pour $\lambda = 1$; la valeur du maximum est $w^* = 12$. En ce qui concerne le problème primal, on a : $12 \leq f^* \leq 16$.

En fait, la solution optimale est le chemin x^* constitué des arcs (1, 4), (4, 3) et (3, 6), de coût 15 et de temps de transmission 9 ; $\Delta(x^*)$ est au-dessus du graphe de $w(\lambda)$, ce qui explique que la résolution du problème dual n'ait pu l'exhiber.



La relaxation lagrangienne ne permet donc d'obtenir qu'un encadrement du plus petit coût avec contrainte. On a cependant pu constater, sur une série de 40 tests, pour des graphes ayant entre 60 et 200 sommets, que la relaxation lagrangienne trouve la solution dans 32 cas, et pour 5 des 8 cas restants, la solution approchée est à moins de 10 % de la solution exacte.
On peut aussi remarquer qu'il peut être très utile d'avoir un bon minorant du minimum recherché ; cette information peut par exemple servir à effectuer l'évaluation dans un problème traité par séparation et évaluation (voir chapitre XI).

IV.3. Maximum du minimum d'une famille de fonctions linéaires

Dans le cas où il existe plus d'une contrainte, une résolution graphique du problème dual comme celle qui a été appliquée plus haut n'est plus envisageable. On peut cependant résoudre les problèmes appelés plus haut (D_k) à l'aide de l'algorithme du simplexe, en exprimant ces problèmes sous forme de problèmes de programmation linéaire. D'une façon plus générale, le lemme suivant montre comment transformer la recherche du maximum du minimum d'une famille de fonctions linéaires en un tel problème de programmation linéaire.

Lemme. *Soient $f_1, f_2, ..., f_r$, r fonctions affines de $\mathbb{R}^m$ dans $\mathbb{R}$. Soit le problème (D_r) :*

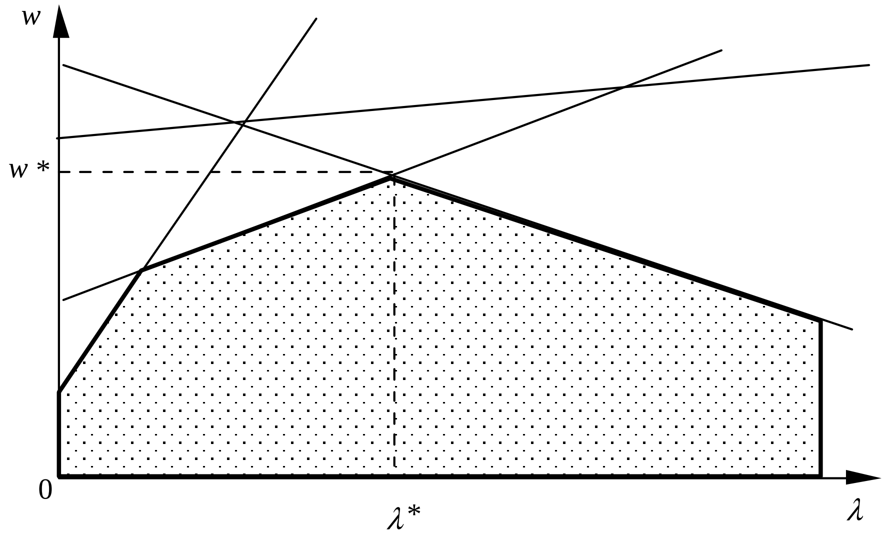
$$\underset{\Lambda \in (R^+)^m}{\text{Maximiser}} \ \text{Min} \ \{f_1(\Lambda), ..., f_r(\Lambda)\}$$

Λ^* réalise ce maximum et w^* est la valeur de ce maximum si et seulement si (Λ^*, w^*) est solution optimum du problème de programmation linéaire (P_r) défini par :

Maximiser w

avec les contraintes $\begin{cases} w \le f_1(\Lambda) \\ w \le f_2(\Lambda) \\ \dots \\ w \le f_r(\Lambda) \\ w \in \mathbb{R}, \Lambda \in (\mathbb{R}^+)^m \end{cases}$

Preuve.— Afin de ne pas alourdir, la preuve sera informelle ; nous considérons le cas $m = 1$ et représentons les droites $\lambda \varnothing f_i(\lambda)$.



Une solution de (D_r) est un couple $(\lambda^*,\ z^*)$ qui donne le point de plus grande ordonnée de la courbe : $\lambda \to \mathrm{Min}\{f_1(\lambda), \ldots , f_r(\lambda)\}$.
Une solution de (P_r) est un couple $(\lambda^*,\ z^*)$ qui donne le point de plus grande ordonnée parmi les points qui sont en dessous de toutes les droites : $\lambda \to f_i(\lambda)$ (zone grisée).
On obtient évidemment les mêmes solutions.
Ce lemme permet de transformer les problèmes (D_k) rencontrés dans la résolution en problèmes de programmation linéaire ; (D_k) devient :

$$\text{Maximiser } w$$

avec les contraintes
$$\begin{cases} w \leq L(x_1, \Lambda) \\ w \leq L(x_2, \Lambda) \\ \ldots \\ w \leq L(x_r, \Lambda) \\ w \in \mathbb{R},\ \Lambda \in (\mathbb{R}^+)^m \end{cases}$$

À chaque étape, on ajoute une contrainte. Remarquons que, dans la base optimale correspondant à la solution de (D_{k-1}), le problème (D_k) est dual-réalisable ; on utilisera cette base de départ pour la résolution de (D_k) : si on a résolu à l'itération précédente le dual de (D_{k-1}), seule la colonne (duale) associée à la nouvelle contrainte (primale) peut éventuellement être entrante dans la résolution du dual de (D_k).

IV.4. Exercice

Un réseau téléinformatique est constitué de deux parties : un réseau d'interconnexion reliant les calculateurs entre eux et des réseaux arborescents reliant les terminaux au calculateur auquel ils sont affectés (réseaux d'accès).

Nous allons nous intéresser à la seconde partie du problème : étant donnés l'emplacement du site central (calculateur ou concentrateur), l'emplacement des sites terminaux et l'emplacement des « nœuds interurbains de liaisons spécialisées » (« points d'éclatement » des liaisons multipoints), comment constituer un réseau de coût minimum respectant certaines contraintes du type : limitation du nombre de sorties sur un point d'éclatement (contrainte de degré) ? En supposant les capacités des liaisons suffisamment élevées, le réseau optimal est alors un arbre de poids minimum respectant les contraintes de degré en chacun de ses sommets. Le but de cet exercice est de déterminer un tel arbre par la relaxation lagrangienne.

On dispose d'un graphe $G = (V, U)$ à n sommets représentant le réseau. Chaque arête u est munie d'un coût $c(u)$ positif ou nul. À un arbre A de G, on associe la fonction x de U dans $\{0, 1\}$ définie par :

$$\begin{cases} x(u) = 1 \text{ si } u \in A, \\ x(u) = 0 \text{ sinon.} \end{cases}$$

On note S l'ensemble des fonctions ainsi obtenues (par abus de langage, un élément x de S est appelé arbre).

On appellera $d_v(x)$ le degré du sommet v dans l'arbre x, et pour chaque sommet v du graphe on définit une borne D_v que ne devra pas dépasser $d_v(x)$ dans la solution finale. On notera enfin $i(u)$ et $j(u)$ les extrémités de l'arête u. Avec ces notations, le problème que l'on veut résoudre s'écrit :

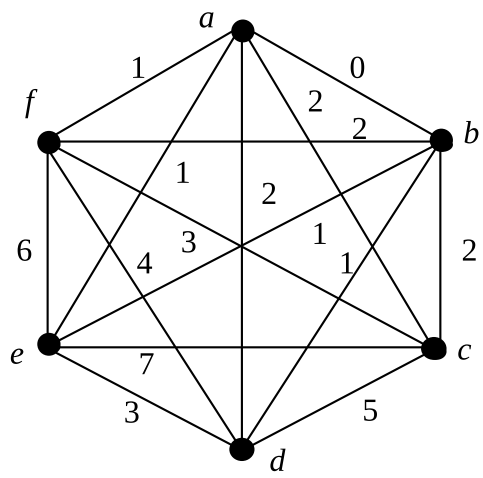
$$\underset{x \in S}{\text{Minimiser}} \quad \sum_{u \in U} c(u)x(u)$$

avec les contraintes : $\forall\, v \in V,\ d_v(x) \leq D_v$.

1. Écrire le problème dual obtenu en relâchant les contraintes (on notera λ_v les multiplicateurs de Lagrange).

2. En remarquant la relation $\sum_{v \in V} \lambda_v d_v(x) = \sum_{u \in U} x(u)\left[\lambda_i(u) + \lambda_j(u)\right]$, mettre le problème dual sous une forme ne faisant plus intervenir les sommets v, mais seulement les arêtes u. En déduire que l'on peut déterminer les valeurs prises par la fonction duale grâce à un algorithme calculant un arbre (couvrant) de poids minimum (« APM ») ; on précisera la valeur des valuations des arêtes en fonction de $c(u)$ et des $\lambda_{i(u)}$.

3. Application : trouver un APM dans le graphe ci-dessous, avec des contraintes sur les degrés de a et b données par $D_a = 2$ et $D_b = 2$.



La solution trouvée par la relaxation est-elle la solution optimale du problème primal ?

V. Optimisation non linéaire sans contrainte

V.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.

V.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 $x = (x_1, x_2, \ldots, x_n)^t$ (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 :

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

Si $F(x) = (f_1(x), \ldots, f_p(x))$ est un vecteur-ligne où $f_1, \ldots, f_p$ 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 $$\nabla\left(u^{\mathrm{t}}.v\right) = \nabla\left(u^{\mathrm{t}}\right).v + \nabla\left(v^{\mathrm{t}}\right).u\ .$$

Si f admet en x^0 des dérivées partielles continues, on peut lui appliquer la formule de Taylor à l'ordre 1 : $f(x) = f(x^0) + (x - x^0)^{\mathrm{t}}.\nabla f(x^0) + \| x - x^0\|.\varepsilon(x)$, où $\varepsilon(x)$ 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 $x_{n+1} = f(x_1, \ldots, x_n)$, 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 ; x^0 et d étant fixés, posons, pour $s \in \mathbb{R}$: $g(s) = f(x^0 + s.d)$

On obtient alors : $g'(s) = d^t . \nabla f(x^0 + s.d)$

et $g'(0) = d^t . \nabla f(x^0)$.

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

$\nabla^2 f$ s'appelle la matrice hessienne de f.
Si f est une fonction de classe C^2 (admet des dérivées partielles d'ordre 2 continues), le hessien 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 :

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

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

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

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

Théorème (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 :

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

où $\varepsilon(x)$ 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 :

$$f(x) - f(x^*) = -\theta.\|\nabla f(x^*)\|^2 + \theta.\varepsilon_1(\theta) = \theta\left(-\|\nabla f(x^*)\|^2 + \varepsilon_1(\theta)\right),$$

où $\varepsilon_1(\theta)$ est une fonction qui tend vers 0 quand θ tend vers 0. Pour θ positif, $f(x) - f(x^*)$ est du signe de $-\|\nabla f(x^*)\|^2 + \varepsilon_1(\theta)$. Si $\nabla f(x^*) \neq 0$, il existe dans tout voisinage de x^* des points x vérifiant $f(x) < f(x^*)$ (pour θ 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 : $h^t.\nabla^2 f(x^*).h < 0$. On a alors, d'après le développement de Taylor d'ordre 2 :

$$f(x^*+\theta h) - f(x^*) = \theta^2.\left(\frac{1}{2}h^t.\nabla^2 f(x^*).h + \varepsilon_2(\theta)\right),$$

où $\varepsilon_2(\theta)$ 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 (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 ($\forall h \in \mathbb{R}^n$ avec $h \neq 0$, $h^t.\nabla^2 f(x^*).h > 0$),
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 $a = \inf\{h^t.\nabla^2 f(x^*).h \text{ pour } h \in S\}$. La sphère étant un compact, a est atteint : $\exists\, h_0 \in S$ tel que $a = h_0^t.\nabla^2 f(x^*).h_0 > 0$. On en déduit aisément la proposition précédente. D'où, en appliquant la formule de Taylor avec $h = x - x^*$:

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

où $\varepsilon(h)$ est une fonction qui tend vers 0 quand h tend vers 0, ce qui montre le théorème car, pour h de norme assez petite, $\frac{a}{2} + \varepsilon(h)$ est du signe de a, c'est-à-dire positif. ♦

V.3. Fonctions quadratiques.

Soit 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

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

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 :
$\nabla q(x) = \nabla(x^t).b + \frac{1}{2}[\nabla(x^t).A.x + \nabla((A.x)^t).x]$.
Or, $\nabla(x^t)$ est la matrice identité et :
$\nabla[(A.x)^t] = \nabla(x^t.A^t) = \nabla(x^t).A^t = A^t$.
D'où : $\nabla q(x) = b + \frac{1}{2}(A + A^t).x$.
et donc : $\boldsymbol{\nabla q(x) = b + A.x}$.
De plus : $\nabla^2 q(x) = \nabla(\nabla q^t(x)) = \nabla(b^t + x^t.A^t) = A^t = \mathrm{A}$.
On a donc finalement : $\boldsymbol{\nabla^2 q(x) = A}$
Les dérivées d'ordre au moins 3 de q sont nulles. Une fonction quadratique coïncide avec son développement de Taylor à l'ordre 2.

V.3. 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. 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 $\nabla f(x^*) = 0$.

Preuve. Il faut montrer que si $\nabla f(x^*) = 0$, alors f admet un minimum global en x^*. Soit $x \in \mathbb{R}^n$. Posons $g(s) = f(x^* + s(x - x^*))$. On a : $g(0) = f(x^*)$ et $g(1) = f(x)$. De plus, on a la relation $g'(0) = (x - x^*)^t.\nabla f(x^*) = 0$. 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 croissante pour $s \geq 0$ et donc $g(1) \geq g(0)$: f admet donc un minimum global en x^*. ♦

Théorème. 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. 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 $\mathbb{R}^n$, $f(x) \geq f(x^0) + \left(\nabla f\left(x^0\right)\right)^{\mathrm{t}}.(x - x^0)$ (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.

De ce théorème, 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 admet un minimum global unique.

V.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 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 le hessien 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 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, on « descend ».
Lorsque la convergence d'un algorithme a été établie, une qualité importante de cet algorithme est sa vitesse de convergence.

- Si $\dfrac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|} \leq \alpha < 1$ pour k assez grand, on dit que la convergence est *linéaire* de taux α.
- Si $\dfrac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|}$ tend vers 0 quand k tend vers l'infini, on dit que la convergence est *superlinéaire*.
- Si $\dfrac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|^{\gamma}}$ 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*.

V.6. Méthodes de gradient

V.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 $\mathbb{R}^n$ tel que $\nabla f(x^k) \neq 0$. Posons, pour $s \in$ 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 :

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

En supposant d unitaire, $g'(0)$ est minimum si $\cos\theta = -1$, c'est-à-dire si $d = -\dfrac{\nabla f\left(x^k\right)}{\left\|\nabla f\left(x^k\right)\right\|}$.

Cette dernière direction donne ce qu'on appelle la *direction de plus grande pente*.
La différence entre les diverses méthodes de gradient porte sur le choix de s_k et de d^k, choix qui doit au minimum assurer $f(x^{k+1}) \leq f(x^k)$.

V.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 $-\nabla f\left(x^k\right)$ pour d^k : il s'agit de la descente de plus forte pente. On pose ensuite $g(s) = f\left(x^k - s.\nabla f\left(x^k\right)\right)$ 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 tel que $f(x^k + s_k d^k) = \underset{s \geq 0}{\text{Min}}\ f(x^k + sd^k)$
    - $x^{k+1} \leftarrow x^k + s_k d^k$
    - $k \leftarrow k + 1$

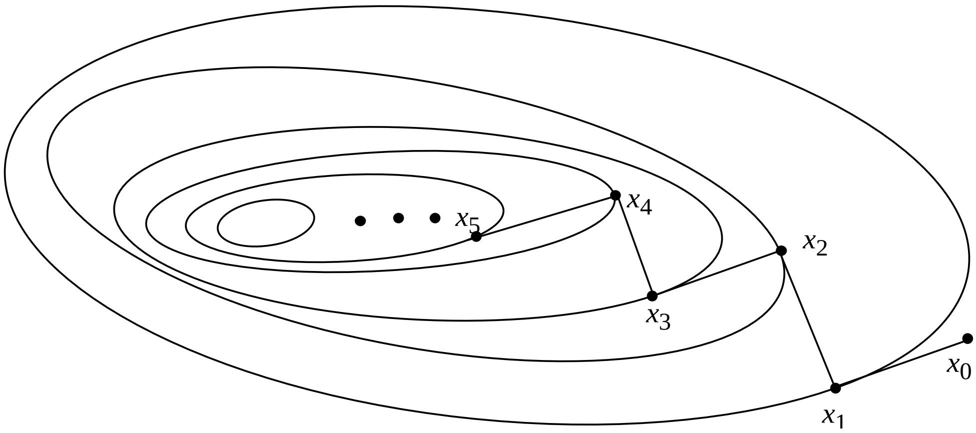
  tant que le test d'arrêt n'est pas vérifié.

Le test d'arrêt peut être par exemple :

- le gradient est très petit : $\sum_{i=1}^{n}\left(\dfrac{\partial f}{\partial x_i}\left(x^k\right)\right)^2 \leq \varepsilon$, où ε est un paramètre donné ;
- la suite x^k est « presque » stationnaire : $|f(x^{k+1}) - f(x^k)| \leq \varepsilon$ (ε donné).

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

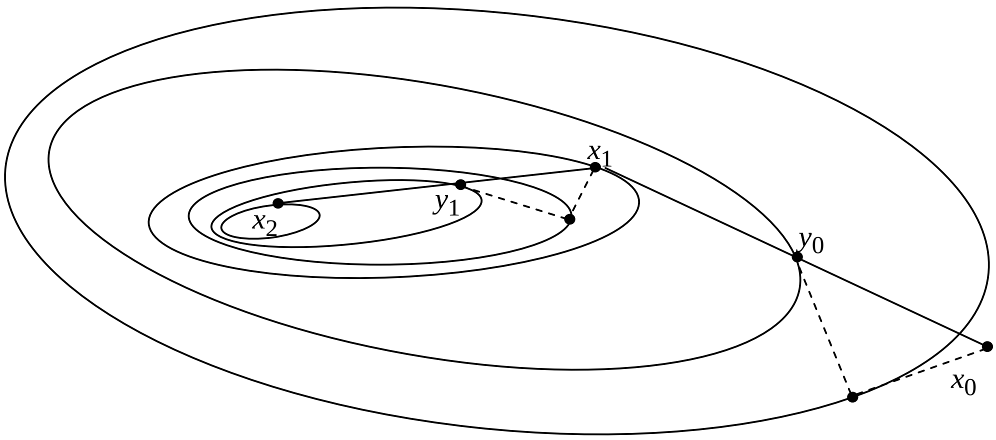
On peut montrer que si $f(x)$ est une fonction de classe C^1 qui tend vers l'infini quand $\|x\|$ tend vers l'infini, cet algorithme converge vers un point stationnaire (point où le gradient s'annule).
L'inconvénient de cette méthode est que la vitesse de convergence peut être très faible (linéaire avec un coefficient 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 : $[\nabla f(x^k)]^t.\nabla f(x^{k+1}) = 0$; les directions de déplacement successives sont orthogonales.
Sur le dessin ci-dessous, on a représenté quelques courbes de niveau et les déplacements.



Il y a convergence en zigzag.

V.6.3. 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. Le dessin ci-dessous illustre cette méthode dans le cas $p = 2$.



V.7. Méthode des gradients conjugués

V.7.1. Cas d'une fonction quadratique

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

La méthode consiste, à partir d'un point x^0, à minimiser q suivant n directions $d^0, d^1, \dots, d^{n-1}$ mutuellement conjuguées par rapport à A, c'est-à-dire vérifiant :

pour $1 \leq i < j \leq n$, $(d^i)^t.A.d^j = 0$.

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

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 : Si $d^0, d^1, \dots, d^{k-1}$ sont mutuellement conjuguées, alors on a pour tout $i < k$ la relation : $(d^i)^t.\nabla q(x^k) = 0$,.

Preuve : on a en effet $(d^i)^t.\nabla q(x^k)$

$$\begin{aligned} &= (d^i)^t.(Ax^k + b) \\ &= (d^i)^t[A\,(x^i + \sum_{j=i}^{k-1} s_j d^j) + b] \\ &= (d^i)^t.(Ax^i + b) + s_i.(d^i)^t.A.d^i \end{aligned}$$

$= 0$ d'après la valeur de s_i calculée ci-dessus.

Théorème : Le point x^n est l'optimum de $q(x)$ sur $\mathbb{R}^n$.

Preuve : Les directions $d^0, d^1, \dots, d^{n-1}$ étant mutuellement conjuguées, elles forment une base de $\mathbb{R}^n$. D'après le lemme, $\nabla q(x^n) = 0$, ce qui démontre le théorème.

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 ; poser $d^0 := -g^0$.
- Pour k variant de 0 à n faire :
  - o $s_k \leftarrow -\dfrac{(d^k)^t.g^k}{(d^k)^t.A.d^k}$
  - o $x^{k+1} \leftarrow x^k + s_k.d^k$
  - o $b_k \leftarrow \dfrac{(g^{k+1})^t.A.d^k}{(d^k)^t.A.d^k}$
  - o $d^{k+1} \leftarrow -g^{k+1} + b_k.d^k$

(on pourra remarquer l'égalité suivante : $(d^k)^t.g^k = -\| g^k \|^2$)

Pour justifier la méthode, il suffit de vérifier que $d^0, d^1 \dots d^{n-1}$ sont mutuellement conjuguées. Nous montrons qu'elles le sont à l'aide d'une récurrence. Pour cela, soit k compris entre 0 et $n - 2$; on suppose que les directions $d^0, d^1 \dots d^k$ sont mutuellement conjuguées. On a alors pour $k + 1$:

$(d^k)^t.A.d^{k+1} \quad = (d^k)^t.A.(-g_{k+1} + b_k.d^k)$

$\quad = -(d^k)^t.A.g^{k+1} + b_k.(d^k)^t.A.d^k = 0$ d'après le choix de b_k.

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

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

D'autre part : $g^{i+1} = -d^{i+1} + b_i.d^i$ et $g^i = -d^i + b_{i-1}.d^{i-1}$.

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, $(g^{k+1})^t.A.d^i = 0$, ce qui montre l'égalité $(d^{k+1})^t.A.d^i = 0$.
Pour terminer, nous démontrons une formule qui nous sera utile dans le paragraphe suivant. On a : $g^{k+1} - g^k = A(x^{k+1} - x^k) = s_k\, A.d^k$.

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

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

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

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

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

V.7.2. 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 $f(x^k + s\, d^k)$, par rapport à s
    - $x^{k+1} \leftarrow x^k + s_k.d^k$
    - $b_k \leftarrow \dfrac{\left\|\nabla f(x^{k+1})\right\|^2}{\left\|\nabla f(x^k)\right\|^2}$
    - $d^{k+1} \leftarrow -\nabla f(x^{k+1}) + b_k.d^k$
    - $k \leftarrow k + 1$

  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.

V.8. Méthode de Newton

On suppose ici que f est deux fois continûment dérivable et que l'on sait calculer ses dérivées secondes.
Au voisinage d'un point x^k, on approche f par la fonction quadratique donnée par la formule de Taylor d'ordre 2 :

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

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 : Si x^0 est choisi suffisamment proche d'un minimum local x^* où le hessien de f est défini positif, alors la suite (x^k) a une convergence quadratique vers x^*.

Preuve : On a : $(\nabla f)(x^k + h) = (\nabla f)(x^k) + (\nabla^2 f)(x^k).h + O(\|h\|^2)$ où $\frac{O(\|h^2\|)}{\|h^2\|}$ est une fonction bornée de h.
Pour $h = x^* - x^k$, on obtient :
$0 = (\nabla f)(x^*) = (\nabla f)(x^k) - (\nabla^2 f)(x^k).(x^k - x^*) + O(\|x^k - x^*\|^2)$.
Au voisinage de x^*, le hessien de f est inversible. On multiplie l'égalité ci-dessus par $[(\nabla^2 f)(x^k)]^{-1}$. On admet $[(\nabla^2 f)(x^k)]^{-1}.O(\|x^k - x^*\|^2) = O(\|x^k - x^*\|^2)$ (résultat qui découle de la continuité de $[(\nabla^2 f)(x^k)]^{-1}$). Comme on a $x^{k+1} - x^k = -[(\nabla^2 f)(x^k)]^{-1} . \nabla f(x^k)$, on obtient :

$- (x^{k+1} - x^k) - (x^k - x^*) + O(\|x^k - x^*\|^2) = 0$,

c'est-à-dire : $x^{k+1} - x^* = O(\|x^k - x^*\|^2)$.
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. ♦

V.9. 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.

V.9.1. 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 :

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

On remarque : $q'(x) = f'(x_k) + f''(x_k)(x - x_k)$.
Si $f''(x_k) > 0$ (cas où f est 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.

V.9.2. Méthode par dichotomie pour une fonction dérivable

Définition : 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.
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 : $x = \frac{1}{2}(x_{min} + x_{max})$; 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 :

- 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
    - $x_{max} \leftarrow h$
    - $h \leftarrow 2h$
  - $x_{min} \leftarrow h$

V.9.3. 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 tels que : $x_1 < x_2 < x_3$ avec $f(x_2) \leq f(x_1)$ 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 :

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

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.

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)

Si $f(x_4) > f(x_2)$

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)

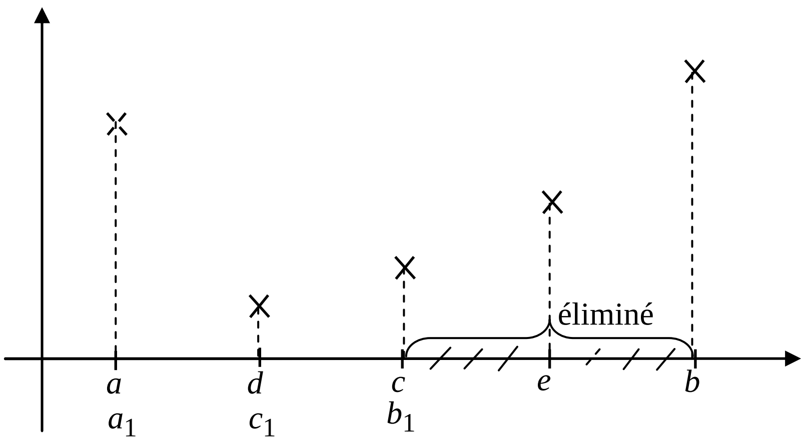
On peut montrer que, si f est assez régulière, la convergence est superlinéaire d'ordre 1,3.

V.9.4. Méthode par 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 ci-dessous 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.



IV.10. Exercice

Déterminer le minimum de la fonction f définie sur $\mathbb{R}^2$ par :

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

VI. Optimisation non linéaire avec contraintes

VI.1. Généralités

Dans cette partie, beaucoup de résultats ne seront pas démontrés.
Soit (P) le problème :

$$(\mathrm{P}) \qquad \text{Minimiser } f(x) \text{ avec } \begin{cases} g_i(x) = 0 \text{ pour } 1 \le i \le m \\ h_j(x) \ge 0 \text{ pour } 1 \le j \le p \end{cases} \text{ et } x \in \mathbb{R}^n.$$

Les fonctions f, g_i ($1 \le i \le m$) et h_j ($1 \le j \le p$) sont supposées de classe C^1. Les conditions $g_i(x) = 0$ et $h_j(x) \ge 0$ s'appellent les *contraintes*. Tout x appartenant à $\mathbb{R}^n$ qui vérifie le système des contraintes s'appelle *solution réalisable*. On note X l'ensemble des solutions réalisables. Si, pour $j \in \{1, 2,..., p\}$ et pour $x \in X$, on a $h_j(x) = 0$, on dit que la contrainte h_j est *saturée* en x.
Certains énoncés de ce chapitre nécessiteraient, pour être tout à fait exacts, des conditions supplémentaires difficiles à énoncer (et qui correspondent à des cas « pathologiques ») ; afin de simplifier, nous n'avons pas précisé celles-ci.

Définition : On dit qu'une direction d est admissible en $x^0 \in X$ si :

- pour $i \in \{1, 2, ..., m\}$, $d^t . \nabla g_i(x^0) = 0$
- pour $j \in \{1, 2, ..., p\}$, si $h_j(x^0) = 0$ alors $d^t . \nabla h_j(x^0) \ge 0$.

Suivre une direction admissible à partir d'un point de X permet de rester dans X ou de le quitter « tangentiellement ».

Théorème. On suppose que le problème admet un minimum local en x^*. Alors, si d est une direction admissible en x^* :

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

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

VI.2. Condition de Lagrange

On s'intéresse ici au problème :

$$(P) \quad \begin{array}{ll} & \text{Minimiser } f(x) \\ \text{avec} & g_i(x) = 0 \quad (1 \le i \le m) \\ \text{et} & x \in \mathbb{R}^n \end{array}$$

où les fonctions f et g_i ($1 \le i \le m$) 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. Soit x^* un minimum local du problème. Alors il existe des réels $\lambda_1, \lambda_2, \ldots, \lambda_m$ tels que : $\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*)$.

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

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

Pour $1 \le i \le m$, $z^t.\nabla g_i(x^*) = 0$ puisque z appartient à $E^\perp$. Par conséquent, z est une direction admissible ; d'après le théorème du paragraphe précédent, il vient : $z^t.\nabla f(x^*) \ge 0$. De plus, pour $1 \le i \le m$, $(-z)^t.\nabla g_i(x^*) = 0$; on a donc aussi $(-z)^t.\nabla f(x^*) \ge 0$. D'où :

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

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

Par conséquent $\|z\|^2 = 0$ et donc $z = 0$. D'où le théorème. ♦

La condition de Lagrange n'est généralement pas suffisante. Elle l'est cependant dans le cas suivant, ce que nous ne prouvons pas.

Théorème. La condition de Lagrange est suffisante lorsque f est convexe dans un ouvert contenant X et que les g_i ($1 \le i \le m$) sont linéaires.

VI.3. Condition de Kuhn et Tucker (ou de Karush, Kuhn et Tucker)

On reprend le problème (P) initial :

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

où les fonctions f, g_i ($1 \le i \le m$) et h_j ($1 \le j \le p$) sont supposées de classe C^1. La condition suivante, appelée condition de Kuhn et Tucker, généralise la condition de Lagrange :

Théorème. Si x^* est un minimum local du problème, alors il existe :

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

tels que :

1) $\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*)$

2) si $h_j(x^*) \neq 0$, alors $\mu_j = 0$ (dans la seconde somme, seuls les gradients des contraintes saturées en x^* peuvent figurer).

Comme la condition de Lagrange, la condition de Kuhn et Tucker n'est généralement pas suffisante. Elle l'est cependant dans le cas suivant. Nous ne prouvons pas le théorème précédent, ni le suivant.

Théorème. La condition de Kuhn et Tucker est suffisante lorsque simultanément f est convexe dans un ouvert contenant X, les g_i ($1 \le i \le m$) sont linéaires et les h_j ($1 \le j \le p$) sont concaves dans un ouvert contenant X.

VI.4. Méthode des directions admissibles

Dans cette dernière partie, nous nous intéressons au problème suivant :

$$
\begin{array}{ll}
 & \text{Minimiser } f(x) \\
\text{avec} & h_j(x) \ge 0 \text{ pour } 1 \le j \le p \\
\text{et} & x \in \mathbb{R}^n.
\end{array}
$$

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 $f(x^{k+1}) \le f(x^k)$, 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 $f(x^k+s.d)$ 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.

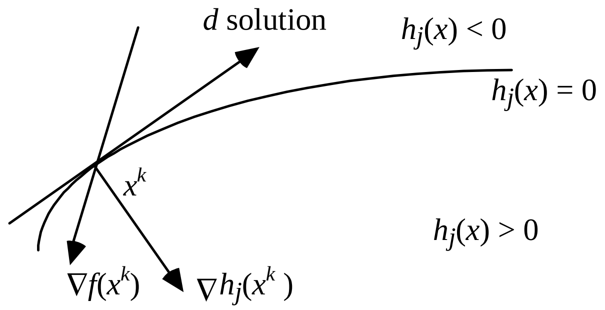
Pour choisir d, on peut résoudre le problème :

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

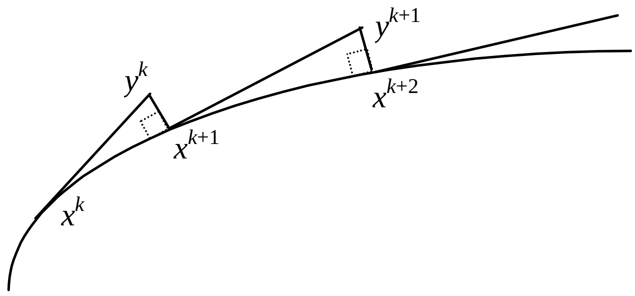
On obtient alors pour d la direction admissible de plus grande pente.

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

La méthode, telle qu'elle vient d'être exposée, doit être accompagnée de compléments. Considérons l'exemple suivant :



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 le dessin ci-dessous.



Remarquons néanmoins que ce dépassement n'apparaît pas, par exemple, dans le cas où les contraintes sont linéaires.

VI.5. 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 quelle(s) valeur(s) de a peut-on affirmer que le minimum est atteint au point (1, 1) ?

2. Résoudre :

$$\text{Minimiser } 2x^2 + y^4$$

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

Exercice 2

On considère le problème (P_α) suivant :

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

où α est un paramètre réel positif ou nul.

1. Indiquer pour quelles valeurs de α le point (1, 0) est un minimum de local (*P*) (pour ces mêmes valeurs, il s'agit en fait d'un minimum global, mais on ne demande pas de le prouver).

On considère maintenant le problème (P_2) suivant :

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

2. On se propose de relâcher la contrainte $x^2 + (y + 1)^2 \le 2$ au sens de la relaxation lagrangienne. Écrire la fonction de Lagrange correspondante, puis la fonction duale *w* et enfin le problème dual.

3. Calculer *w*(1) et *w*(3). Indiquer ce qu'on peut en déduire pour le problème (P_2).

Corrigés des exercices

1. Chapitre I

1.1. Corrigé de l'exercice 1

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

$$\begin{array}{l}
x_4 = 4 - (x_1 + x_2 + 2x_3) \\
x_5 = 5 - (2x_1 \quad\quad + 3x_3) \\
x_6 = 7 - (2x_1 + x_2 + 3x_3) \\
\hline
z \;\;=\;\; 3x_1 + 2x_2 + 4x_3
\end{array}$$

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 $\frac{5}{2}$, la fonction objectif augmente de $\frac{15}{2}$. Si x_2 entre en base, la fonction objectif augmente de 8, enfin si c'est x_3, l'objectif augmente de $\frac{20}{3}$. 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 :

$$\begin{array}{l}
x_2 = 4 - x_1 - 2x_3 - x_4 \\
x_5 = 5 - 2x_1 - 3x_3 \\
x_6 = 7 - 2x_1 - (4 - x_1 - 2x_3 - x_4) - 3x_3 = 3 - x_1 - x_3 + x_4 \\
\hline
z \;\;=\;\; 3x_1 + 2(4 - x_1 - 2x_3 - x_4) + 4x_3 = 8 + x_1 - 2x_4
\end{array}$$

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 :

$$x_1 = \frac{5}{2} - \frac{3}{2}x_3 - \frac{1}{2}x_5$$
$$x_2 = 4 - \left(\frac{5}{2} - \frac{3}{2}x_3 - \frac{1}{2}x_5\right) - 2x_3 - x_4$$
$$x_6 = 3 - \left(\frac{5}{2} - \frac{3}{2}x_3 - \frac{1}{2}x_5\right) - x_3 + x_4$$
$$z = 8 + \frac{5}{2} - \frac{3}{2}x_3 - \frac{1}{2}x_5 - 2x_4 = \frac{21}{2} - \frac{3}{2}x_3 - \frac{1}{2}x_5 - 2x_4$$

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 $\frac{21}{2}$, et il est obtenu pour les valeurs suivantes des variables :

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

1.2. Corrigé de l'exercice 2

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

$$x_5 = 5 - x_1 - 2x_2 - 3x_3 - x_4$$
$$x_6 = 3 - x_1 - x_2 - 2x_3 - 3x_4$$
$$z = 5x_1 + 6x_2 + 9x_3 + 8x_4$$

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 :

$$x_3 = 1{,}5 - 0{,}5x_1 - 0{,}5x_2 - 1{,}5x_4 - 0{,}5x_6$$
$$x_5 = 0{,}5 + 0{,}5x_1 - 0{,}5x_2 + 3{,}5x_4 + 1{,}5x_6$$
$$z = 13{,}5 + 0{,}5x_1 + 1{,}5x_2 - 5{,}5x_4 - 4{,}5x_6$$

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 :

$$x_2 = 1 + x_1 + 7x_4 - 2x_5 + 3x_6$$
$$x_3 = 1 - x_1 - 5x_4 + x_5 - 2x_6$$
$$z = 15 + 2x_1 + 5x_4 - 3x_5$$

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

$$x_4 = 0{,}2 - 0{,}2x_1 - 0{,}2x_3 + 0{,}2x_5 - 0{,}4x_6$$
$$x_2 = 2{,}4 - 0{,}4x_1 - 1{,}4x_3 - 0{,}6x_5 + 0{,}2x_6$$
$$z = 16 + x_1 - x_3 - 2x_5 - 2x_6$$

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

$$x_1 = 1 - x_3 - 5x_4 + x_5 - 2x_6$$
$$x_2 = 2 - x_3 + 2x_4 - x_5 + x_6$$
$$z = 17 - 2x_3 - 5x_4 - x_5 - 4x_6$$

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	$\frac{5}{2}$	$\frac{3}{2}$	1
accroissement correspondant de z	15	15	$\frac{27}{2}$	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 :

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

Seule la variable x_2 est candidate à entrer, la variable x_5 sort ; la base est alors constituée de $\{x_1, x_2\}$ et est donc la base optimale, déterminée ci-dessus.

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.

1.3. Corrigé de l'exercice 3

Considérons le problème :

$$\text{Maximiser } z = x_1 + x_2$$

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 $z = t$. Comme t n'est pas borné, le problème n'est pas borné.

1.4. Corrigé de l'exercice 4

Considérons le problème :

$$\text{Maximiser } z = x_1 + x_2$$

avec les contraintes : $\begin{cases} -2x_1 + 3x_2 \leq -4 \\ \quad x_1 - x_2 \leq 1 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$

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

$$-3x_1 + 3x_2 \geq -3 \text{ ;}$$

comme $-3 > -4$, la première contrainte $-2x_1 + 3x_2 \leq -4$ n'est pas satisfaite. Le problème est donc infaisable.

1.5. Corrigé de l'exercice 5

Une seule variable est candidate à entrer : x_1. Les variables x_5 et x_6 peuvent sortir mais, d'après le critère de choix retenu, c'est la variable x_5 qui sort. On obtient le dictionnaire suivant :

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

La variable entrante ayant le plus grand coefficient est maintenant x_2. La variable sortante est x_6. On obtient le dictionnaire suivant :

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

La variable entrante est maintenant x_3 et la variable sortante, candidate de plus petit indice, est x_1. On obtient le dictionnaire ci-dessous :

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

La variable entrante est maintenant x_4 et la variable sortante est x_2. On obtient le dictionnaire ci-dessous :

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

La variable entrante est maintenant x_5 et la variable sortante est x_3. D'où :

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

La variable entrante est maintenant x_6 et la variable sortante est x_4. D'où:

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

La variable entrante est maintenant x_1 et la variable sortante est x_5. On retrouve alors le dictionnaire initial : il y a cyclage, l'algorithme ne se termine pas.

1.6. Corrigé de l'exercice 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 de décision 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 :

$$\text{Maximiser } w = -x_0$$

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 :

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

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 :

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

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

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

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 :

Maximiser $w = -x_0$

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

On obtient le dictionnaire initial suivant :

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

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 :

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

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

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

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

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

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 :

$$\begin{array}{l} x_2 = \frac{10}{17} + \frac{5}{17}x_3 + \frac{4}{17}x_4 - \frac{5}{17}x_0 \\ x_1 = \frac{30}{17} - \frac{2}{17}x_3 - \frac{5}{17}x_4 + \frac{2}{17}x_0 \\ x_5 = 2 - 2x_3 - x_4 + 2x_0 \\ \hline z = \frac{180}{17} + \frac{5}{17}x_3 - \frac{13}{17}x_4 \end{array}$$

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

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

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.

2. Chapitre II

2. Corrigé de l'exercice

Pour constituer la base, prenons les modèles de découpe suivants :

$$\begin{pmatrix}1\\0\\1\\1\end{pmatrix}, \begin{pmatrix}0\\3\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\3\\0\end{pmatrix} \text{ et } \begin{pmatrix}0\\0\\0\\4\end{pmatrix}$$

$$\begin{pmatrix}1&0&0&0\\0&3&0&0\\1&0&3&0\\1&0&0&4\end{pmatrix}$$

B est donc la matrice $\begin{pmatrix}1&0&0&0\\0&3&0&0\\1&0&3&0\\1&0&0&4\end{pmatrix}$ qui est inversible. X_B^* vaut alors $\begin{pmatrix}600\\200\\0\\150\end{pmatrix}$: la solution proposée est bien une solution basique réalisable.

ÉTAPE 1

1. $y.B = (1\ \ 1\ \ 1\ \ 1)$ conduit à $y = \begin{pmatrix} \frac{5}{12} & \frac{1}{3} & \frac{1}{3} & \frac{1}{4} \end{pmatrix}$.

2. On cherche un modèle de découpe a tel que $y.a > 1$: $a = (0\ \ 2\ \ 0\ \ 2)^t$ convient.

3. On résout $B.d = a$; on obtient $d = \begin{pmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{2} \end{pmatrix}^t$.

4. X_B^* devient $X_B^* - t.d = \begin{pmatrix} 600 \\ 200 - \frac{2}{3}t \\ 0 \\ 150 - \frac{1}{2}t \end{pmatrix}$ où t représente le nombre de mètres à découper selon le modèle entrant en base. On a $t \le 300$, et quand t vaut 300, le nombre de mètres à découper selon les modèles 2 et 4 de la base actuelle devient nul. On choisit comme modèle sortant le modèle 4.

5. Par conséquent, on a maintenant : $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix}$ et $X_B^* = \begin{pmatrix} 600 \\ 0 \\ 0 \\ 300 \end{pmatrix}$.

ÉTAPE 2

1. $y.B = (1\ \ 1\ \ 1\ \ 1)$ conduit à $y = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$.

2. On cherche un modèle de découpe a tel que $y.a > 1$: aucun modèle de découpe ne convient. La solution actuelle est donc optimale ; elle consiste à acheter 900 m de tissu que l'on découpe de la façon suivante :
   - dans 600 m sont découpées une bande de 52 cm de large, une autre de 27 cm de large et une troisième de 21 cm de large ;
   - dans les 300 m restants sont découpées deux bandes de 29 cm de large et deux de 21cm de large.

3. Chapitre III

3.1. Corrigé de l'exercice 1

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.

$2x_1^* + 4x_2^* + 4x_3^* - 2x_4^* + 5x_5^* = \frac{14}{3} < 5$: 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^* :

$$y_1^* + 4y_2^* + 2y_3^* + 3y_4^* = 8 \geq 7$$
$$2y_1^* + y_2^* + 5y_3^* - 2y_4^* = 1 < 3$$

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. Appliquant la forme matricielle de la méthode du simplexe, le calcul de y donne le vecteur y^* déterminé ci-dessus, et la dernière inégalité obtenue indique que la variable x_5 est variable entrante. Il reste à poursuivre la méthode.

3.2. Corrigé de l'exercice 2

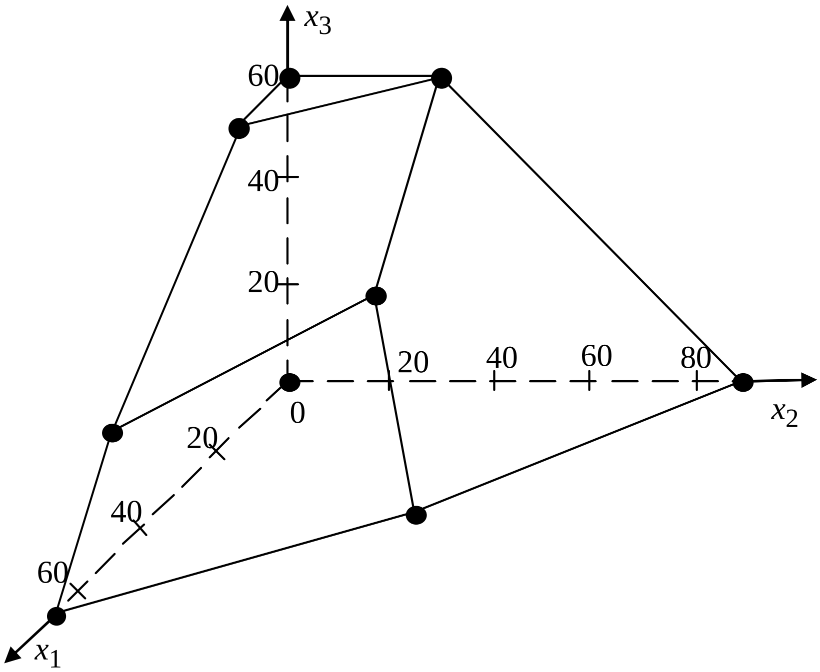
a. Nous appliquons la forme matricielle de la méthode du simplexe.
En adoptant le kilogramme comme unité pour le cuivre, le décagramme pour le cadmium et l'étain, la journée pour le travail des ouvriers, la centaine d'euros pour le profit et l'hectomètre pour la longueur des fils, le problème proposé s'écrit :

Maximiser $c.X$ avec $AX = b$ et $X \geq 0$

où : $A = \begin{pmatrix} 9 & 5 & 6 & 1 & 0 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$ $b = \begin{pmatrix} 600 \\ 150 \\ 60 \\ 90 \end{pmatrix}$ $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}$

et $c = (42, 39, 52, 0, 0, 0, 0)$

Le polyèdre défini par $x_j \geq 0$ pour $1 \leq j \leq 7$ est représenté ci-dessous.



Soit B la matrice correspondant aux colonnes 4, 5, 6, 7 de la matrice A ; elle est inversible (en fait c'est la matrice identité) et par suite $\{x_4, x_5, x_6, x_7\}$ est une base du problème. La solution basique correspondante (il s'agit de b) ayant toutes ses composantes positives ou nulles, cette base est réalisable : on peut l'utiliser comme point de départ de la méthode du simplexe.

ÉTAPE 1

On commence donc avec la base $\{x_4, x_5, x_6, x_7\}$.

$$X_B = \begin{pmatrix} x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} X_B^* = \begin{pmatrix} 600 \\ 150 \\ 60 \\ 90 \end{pmatrix} X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

On actualisera en fait à chaque étape uniquement X_B, B et X_B^*.

1. Calculons $y = (y_1, y_2, y_3, y_4)$ tel que $yB = c_B = (0, 0, 0, 0)$. On trouve $y = (0, 0, 0, 0)$.

2. Cherchons le plus grand indice i d'une variable hors-base telle qu'on ait $c_i - ya_i > 0$: $i = 3$ convient puisqu'il s'agit du plus grand indice des variables hors-base et qu'on a $ya_3 = 0$ et $c_3 = 52$.

3. On a alors $a = \begin{pmatrix} 6 \\ 2 \\ 1 \\ 1 \end{pmatrix}$; cherchons $d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}$ tel que $Bd = a$. Il vient $d = \begin{pmatrix} 6 \\ 2 \\ 1 \\ 1 \end{pmatrix}$.

4. On trouve donc $X_B^* - td = \begin{pmatrix} 600-6t \\ 150-2t \\ 60-t \\ 90-t \end{pmatrix}$; d'où $\begin{cases} t \le 100 \\ t \le 75 \\ t \le 60 \\ t \le 90 \end{cases}$. On prend donc x_6 comme variable quittant la base et on attribue à t la valeur 60.

5. On a maintenant : $X_B = \begin{pmatrix} x_4 \\ x_5 \\ x_3 \\ x_7 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 0 & 6 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ $X_B^* = \begin{pmatrix} 240 \\ 30 \\ 60 \\ 30 \end{pmatrix}$

ÉTAPE 2

1. Le vecteur y vérifie : $y_1 = 0$, $y_2 = 0$, $6y_1 + 2y_2 + y_3 + y_4 = 52$, $y_4 = 0$. On obtient donc : $y = (0, 0, 52, 0)$.

2. Cherchons le plus grand indice i d'une variable hors-base telle qu'on ait $c_i - ya_i > 0$: $i = 6$ ne convient pas mais $i = 2$ convient, puisqu'on a $ya_2 = 0$ et $c_2 = 39$.

3. On cherche alors d tel que $Bd = a_2$, ou encore :

$$\begin{cases} d_1 + 6d_3 = 5 \\ d_2 + 2d_3 = 1 \\ d_3 = 0 \\ d_3 + d_4 = 1 \end{cases} \quad \text{soit} \quad d = \begin{pmatrix} 5 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

4. On a :

$$X_B^* - td = \begin{pmatrix} 240-5t \\ 30-t \\ 60 \\ 30-t \end{pmatrix}; \quad \text{d'où} \quad \begin{cases} t \le 48 \\ t \le 30 \\ t \text{ quelconque} \\ t \le 30 \end{cases}$$

On voit donc que l'on doit prendre $t = 30$ et que c'est la deuxième variable (x_5) qui sort (x_7 aurait pu sortir aussi, mais la règle fixée dans l'énoncé nous fait choisir x_5).

5. L'actualisation conduit à :

$$X_B = \begin{pmatrix} x_4 \\ x_2 \\ x_3 \\ x_7 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 5 & 6 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \quad X_B^* = \begin{pmatrix} 90 \\ 30 \\ 60 \\ 0 \end{pmatrix}$$

ÉTAPE 3

1. Le vecteur y vérifie : $\begin{cases} y_1 = 0 \\ 5y_1 + y_2 + y_4 = 39 \\ 6y_1 + 2y_2 + y_3 + y_4 = 52 \\ y_4 = 0 \end{cases}$

On trouve $y = (0, 39, -26, 0)$.

2. Cherchons le plus grand indice i d'une variable hors-base telle qu'on ait $c_i - ya_i > 0$: $i = 6$ convient, puisqu'on a $ya_6 = -26$ et $c_6 = 0$.

3. On cherche alors d tel que $Bd = a_6$, ou encore :

$$\begin{cases} d_1 + 5d_2 + 6d_3 = 0 \\ d_2 + 2d_3 = 0 \\ d_3 = 1 \\ d_2 + d_3 + d_4 = 0 \end{cases} \quad \text{soit} \quad d = \begin{pmatrix} 4 \\ -2 \\ 1 \\ 1 \end{pmatrix}.$$

4. On a :

$$X_B^* - td = \begin{pmatrix} 90-4t \\ 30+2t \\ 60-t \\ 0 \end{pmatrix}; \qquad \text{d'où} \quad \begin{cases} t \leq 22{,}5 \\ t \text{ quelconque} \\ t \leq 60 \\ t \leq 0 \end{cases}$$

On voit donc que l'on doit prendre $t = 0$ et que c'est la quatrième variable qui sort, soit x_7.

5. L'actualisation conduit à :

$$X_B = \begin{pmatrix} x_4 \\ x_2 \\ x_3 \\ x_6 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 5 & 6 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \qquad X_B^* = \begin{pmatrix} 90 \\ 30 \\ 60 \\ 0 \end{pmatrix}$$

ÉTAPE 4

1. Le vecteur y vérifie :

$$\begin{cases} y_1 = 0 \\ 5y_1 + y_2 + y_4 = 39 \\ 6y_1 + 2y_2 + y_3 + y_4 = 52 \\ y_3 = 0 \end{cases}$$

On trouve $y = (0, 13, 0, 26)$.

2. Cherchons un indice i d'une variable hors-base telle que $c_i - ya_i > 0$:

$$c_1 - ya_1 = -10\ ;$$
$$c_5 - ya_5 = -13\ ;$$
$$c_7 - ya_7 = -26.$$

Il n'y a plus de colonne entrante ; la solution actuelle est donc optimale. Elle correspond, comme on le voit en consultant la base B et le vecteur X_B^*, aux valeurs suivantes des variables de choix :

$$x_1^* = 0\ ;\ x_2^* = 30\ ;\ x_3^* = 60.$$

Le profit maximum est donc obtenu en fabriquant 0 hm de fil F_1, 30 hm de fil F_2 et 60 hm de fil F_3. Le profit est alors 4290 centaines d'euros, soit 429 000 €. Il reste 90 kg de cuivre ; tout le cadmium, tout l'étain et les 90 jours de travail sont utilisés.

b. Pour exprimer le problème dual, écrivons le problème primal sous forme standard :

$$\text{Maximiser } 42x_1 + 39x_2 + 52x_3$$

avec les contraintes
$$\begin{cases} 9x_1 + 5x_2 + 6x_3 \leq 600 \\ 2x_1 + \ x_2 + 2x_3 \leq 150 \\ \qquad\qquad\quad x_3 \leq 60 \\ x_1 + \ x_2 + \ \ x_3 \leq 90 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{cases}$$

On en déduit l'expression du problème dual :

Minimiser $600y_1 + 150y_2 + 60y_3 + 90y_4$

avec les contraintes
$$\begin{cases} 9y_1 + 2y_2 \qquad + \ y_4 \geq 42 \\ 5y_1 + \ y_2 \qquad + \ y_4 \geq 39 \\ 6y_1 + 2y_2 + y_3 + y_4 \geq 52 \\ y_1 \geq 0, y_2 \geq 0, y_3 \geq 0, y_4 \geq 0 \end{cases}$$

ou encore, mis sous forme standard (sachant que le maximum obtenu ci-dessous sera égal à l'opposé du minimum obtenu ci-dessus) :

Maximiser $-600y_1 - 150y_2 - 60y_3 - 90y_4$

avec les contraintes
$$\begin{cases} -9y_1 - 2y_2 \qquad - y_4 \leq -42 \\ -5y_1 - \ \ y_2 \qquad - y_4 \leq -39 \\ -6y_1 - 2y_2 - y_3 - y_4 \leq -52 \\ y_1 \geq 0, y_2 \geq 0, y_3 \geq 0, y_4 \geq 0 \end{cases}$$

Le dernier vecteur y calculé dans la question a donne une solution optimale du problème dual :

$$y_1^* = 0,\ y_2^* = 13,\ y_3^* = 0,\ y_4^* = 26,$$

ce qui donne bien la valeur 4290 (centaines d'euros) à la fonction objectif du problème dual.

c. La question est celle de la décision à prendre si le profit associé au fil F_1 passe de 4200 € à 4800 € ou de 4200 € à 5400 €. Dans les termes de notre algorithme, on peut la reformuler comme suit : par rapport à la question a, si on modifie maintenant c_1 en lui donnant la valeur 48 (puis 54), la colonne a_1 devient-elle entrante ?
En reprenant les calculs de la fin de la question a, celui de $c_1 - y.a_1$ lors du point 2 de l'étape 4 conduit alors successivement à – 4 (dans le cas 48) puis à 2 (dans le cas 54). On voit que lorsque l'on augmente le bénéfice associé au fil F_1 de 4200 € à 4800 €, on ne change pas le maximum du profit espéré, alors que si on lui attribue la valeur 5400 €, il y a lieu de reconsidérer la politique de l'entreprise. Revenant au vocabulaire de l'algorithme, on voit dans ce dernier cas que la colonne a_1 est entrante et on reprend là les calculs, au troisième point de l'étape 4.

3. Le calcul du vecteur d tel que $Bd = a_1$ conduit à la résolution du système suivant :

$$\begin{cases} d_1 + 5d_2 + 6d_3 \qquad = 9 \\ \qquad d_2 + 2d_3 \qquad = 2 \\ \qquad\qquad d_3 + d_4 = 0 \\ \qquad d_2 \ + \ d_3 \qquad = 1 \end{cases} \text{ soit } \ d = \begin{pmatrix} 3 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

4. On a :

$$X_B^* - td = \begin{pmatrix} 90-3t \\ 30 \\ 60-t \\ 0+t \end{pmatrix}; \qquad \text{d'où} \qquad \begin{cases} t \le 30 \\ t \text{ quelconque} \\ t \le 60 \\ t \text{ quelconque} \end{cases}$$

On voit donc que l'on doit prendre $t = 30$ et que x_4 sort.

5. L'actualisation conduit à :

$$X_B = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_6 \end{pmatrix} \quad B = \begin{pmatrix} 9 & 5 & 6 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad X_B^* = \begin{pmatrix} 30 \\ 30 \\ 30 \\ 30 \end{pmatrix}$$

ÉTAPE 5

1. Le vecteur y vérifie :

$$\begin{cases} 9y_1 + 2y_2 + y_4 = 54 \\ 5y_1 + y_2 + y_4 = 39 \\ 6y_1 + 2y_2 + y_3 + y_4 = 52 \\ y_3 = 0 \end{cases}$$

On trouve $y = (2/3, 37/3, 0, 70/3)$.

2. Cherchons un indice i d'une variable hors-base telle que $c_i - ya_i > 0$.

$c_4 - ya_4 = -2/3$;
$c_5 - ya_5 = -37/3$;
$c_7 - ya_7 = -70/3$.

Aucune colonne n'est entrante ; on a ainsi déterminé la nouvelle solution optimale, donnée (en hectomètres) par : $x_1^* = 30$; $x_2^* = 30$; $x_3^* = 30$. Le nouveau profit est de 4350 centaines d'euros (on vérifie aisément que c'est aussi la valeur donnée à la fonction objectif du problème dual par $y = (2/3, 37/3, 0, 70/3)$, comme l'indique la théorie).

d. La question est proche de la précédente, mais concerne ce coup-ci une variable en base. Il convient donc, par rapport à la question a, de reprendre les calculs au début de l'étape 4. En effet, les fluctuations sur c_3 (exprimé en centaines d'euros) affectent le dernier calcul de y, mais pas le fait que la solution trouvée à la fin de l'étape 3 est réalisable et basique : elle peut donc nous servir de point de départ. L'étape 4 devient alors :

ÉTAPE 4

1. Le vecteur y vérifie :

$$\begin{cases} y_1 = 0 \\ 5y_1 + y_2 + y_4 = 39 \\ 6y_1 + 2y_2 + y_3 + y_4 = c_3 \\ y_3 = 0 \end{cases}$$

On trouve $y = (0, c_3 - 39, 0, 78 - c_3)$.

2. Calculons, pour les variables hors-base x_i, les différences $c_i - ya_i$:

$$c_1 - ya_1 = 42 - c_3 \; ; \; c_5 - ya_5 = 39 - c_3 \; ; \; c_7 - ya_7 = c_3 - 78.$$

Si c_3 est compris entre 42 et 78 inclus, on constate que les trois termes précédents sont négatifs ou nuls : les variables hors-base ne peuvent donc entrer en base et la solution trouvée à la première question reste optimale. Si c_3 sort de l'intervalle [42, 78], il convient de continuer les calculs comme précédemment pour savoir si la solution de la première question reste optimale ou non (car le fait d'avoir une variable hors-base qui entre dans la base n'entraîne pas nécessairement une croissance de la fonction objectif, à cause de la dégénérescence de la solution).

e. Considérons maintenant la question de savoir si on a intérêt à acheter de l'étain ou à payer des jours de travail supplémentaires. Pour cela, le profit associé au fil F_1 étant 5400 € ($c_1 = 54$), on repart des résultats obtenus à la question c. On peut appliquer le théorème sur la signification économique du dual puisque la solution n'est pas dégénérée pour $c_1 = 54$: acheter une quantité q d'étain et payer j jours supplémentaires revient à ajouter à b le vecteur $\delta b = (0, 0, q, j)^t$. Si q et j sont assez petits pour que la base optimale (du problème avec $c_3 = 54$) reste réalisable, alors le profit optimal vaut 4290 + (2/3, 37/3, 0, 70/3)δb, soit : 4290 + 70j/3, puisque le dernier vecteur y trouvé à la question c est (2/3, 37/3, 0, 70/3).
Acheter de l'étain ne permet donc pas d'espérer une augmentation du profit, quel que soit d'ailleurs son prix (positif). C'est ce qu'indique le théorème sur la signification économique du dual : une solution optimale du problème dual étant (2/3, 37/3, 0, 70/3), l'achat d'un décagramme d'étain rapporte $y_3^* = 0$. Il était d'ailleurs prévisible qu'il ne serait pas intéressant de s'en procurer une quantité supplémentaire, puisque la solution optimale obtenue à la question c ne consommait pas tout l'étain disponible. On a par conséquent $q = 0$.
En revanche, le coût d'une journée supplémentaire étant inférieure à ce que celle-ci rapporte ($y_4^* > 20$), il est intéressant de payer de tels jours supplémentaires. Pour que ce raisonnement reste valable, il faut que la solution $X_B^* + \delta X_B^*$ de l'équation : $B(X_B^* + \delta X_B^*) = b + \delta b$ ait toutes ses composantes positives ou nulles, B étant la matrice définie par la base trouvée à l'étape 4 de la question c. Or, X_B^* est solution de $BX_B^* = b$. Il vient donc $B.\delta X_B^* = \delta b$, c'est-à-dire :

$$\begin{cases} 9\delta x_1^* + 5\delta x_2^* + 6\delta x_3^* = 0 \\ 2\delta x_1^* + \delta x_2^* + 2\delta x_3^* = 0 \\ \delta x_3^* + \delta x_6^* = 0 \\ \delta x_1^* + \delta x_2^* + \delta x_3^* = j \end{cases}$$

qui admet pour solution : $\delta x_1^* = -4j/3$, $\delta x_2^* = 2j$, $\delta x_3^* = j/3$, $\delta x_6^* = -j/3$. La valeur maximum de j est donc 22,5 puisque, pour avoir $X_B^* + \delta X_B^* \geq 0$, j doit vérifier :

$$\begin{cases} 30 - 4j/3 \geq 0 \\ 30 + 2j \geq 0 \\ 30 + j/3 \geq 0 \\ 30 - j/3 \geq 0 \end{cases}$$

On peut conseiller de payer 22,5 jours supplémentaires ; le nouveau plan optimum est alors donné par : $x_1^* = 0$; $x_2^* = 75$; $x_3^* = 37,5$.
On peut vérifier que le profit correspondant est alors de 4875 centaines d'euros ; il a augmenté de $j \times y_4^* = 22,5 \times 70 / 3 = 525$ centaines d'euros.

f. Envisager un nouveau fil, que nous appellerons F_0, revient à ajouter une colonne à la matrice des contraintes A, colonne que nous appellerons a_0 et qui est $a_0 = (8, 2, 0, 1)^t$. Pour savoir si on a intérêt à fabriquer ce fil, il suffit de voir si, lors de la dernière phase de la question a., cette colonne est entrante. Rappelons que nous avons calculé y qui vaut, à moment-là, (0, 13, 0, 26). De sorte que :

$$c_0 - ya_0 = 55 - (26 + 26) = 3,$$

quantité strictement positive : la colonne a_0 est entrante. On reprend les calculs de la question a. à l'étape 4, point 3.

3. On cherche alors d tel que $Bd = a_0$, ou encore :

$$\begin{cases} d_1 + 5d_2 + 6d_3 = 8 \\ d_2 + 2d_3 = 2 \\ d_3 + d_4 = 0 \\ d_2 + d_3 = 1 \end{cases}, \qquad \text{soit} \quad d = \begin{pmatrix} 2 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

4. On a :

$$X_B{}^* - td = \begin{pmatrix} 90-2t \\ 30 \\ 60-t \\ t \end{pmatrix}; \qquad \text{d'où} \quad \begin{cases} t \le 45 \\ t \text{ quelconque} \\ t \le 60 \\ t \text{ quelconque} \end{cases}$$

On voit donc que l'on doit prendre $t = 45$ et que c'est la première variable qui sort, soit x_4.

5. L'actualisation conduit à :

$$X_B = \begin{pmatrix} x_0 \\ x_2 \\ x_3 \\ x_6 \end{pmatrix} \qquad B = \begin{pmatrix} 8 & 5 & 6 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad X_B{}^* = \begin{pmatrix} 45 \\ 30 \\ 15 \\ 45 \end{pmatrix}$$

ÉTAPE 5

1. Le vecteur y vérifie :

$$\begin{cases} 8y_1 + 2y_2 + y_4 = 55 \\ 5y_1 + y_2 + y_4 = 39 \\ 6y_1 + 2y_2 + y_3 + y_4 = 52 \\ y_3 = 0 \end{cases}$$

On trouve y = (1,5 ; 11,5 ; 0 ; 20).

2. Cherchons un indice i d'une variable hors-base telle que $c_i - ya_i > 0$.

$$c_1 - ya_1 = -14{,}5\ ;$$
$$c_4 - ya_4 = -1{,}5\ ;$$
$$c_5 - ya_5 = -11{,}5\ ;$$
$$c_7 - ya_7 = -20.$$

Il n'y a plus de colonne entrante ; la solution actuelle est donc optimale. Elle correspond aux valeurs suivantes des variables de choix (en hectomètres) :

$$x_0{}^* = 45\ ;\ x_1{}^* = 0\ ;\ x_2{}^* = 30\ ;\ x_3{}^* = 15.$$

g. Enfin, le gainage introduit une nouvelle contrainte primale : $x_2 + x_3 \le 80$, et donc une nouvelle ligne dans la matrice A des contraintes primales. Si on considère le problème dual, cela se traduit par l'introduction d'une colonne supplémentaire dans la matrice A^t des contraintes duales. Il suffit donc de procéder comme à la question précédente, mais par rapport au problème dual dont la forme standard est la suivante, en appelant y_0 la nouvelle variable duale :

$$\text{Maximiser} -80y_0 - 600y_1 - 150y_2 - 60y_3 - 90y_4$$

avec les contraintes
$$\begin{cases} \quad -9y_1 - 2y_2 \quad - y_4 \le -42 \\ -y_0 - 5y_1 - \quad y_2 \quad - y_4 \le -39 \\ -y_0 - 6y_1 - 2y_2 - y_3 - y_4 \le -52 \\ y_0 \ge 0, y_1 \ge 0, y_2 \ge 0, y_3 \ge 0, y_4 \ge 0 \end{cases}$$

Si on part de la solution optimale décrite à la question b, la base considérée correspond aux variables y_2, y_4 et y_5. Le vecteur Y_B, la matrice B associée et le vecteur Y_B^* sont donnés par :

$$Y_B = \begin{pmatrix} y_2 \\ y_4 \\ y_5 \end{pmatrix} \qquad B = \begin{pmatrix} -2 & -1 & 1 \\ -1 & -1 & 0 \\ -2 & -1 & 0 \end{pmatrix} \qquad Y_B^* = \begin{pmatrix} 13 \\ 26 \\ 10 \end{pmatrix}$$

Par ailleurs, les valeurs optimales des variables primales donnent le vecteur x défini par $xB = (-150, -90, 0)$ (ces valeurs correspondant aux coefficients de y_2, y_4 et y_5 dans la fonction objectif du problème dual sous forme standard) : $x = (0, 30, 60)$. On peut maintenant appliquer l'algorithme du simplexe à partir de cette base, en regardant si la colonne (duale) $(0, -1, -1)^t$ associée à la nouvelle contrainte primale est entrante :

$$-80 - [0 \times 0 + 30 \times (-1) + 60 \times (-1)] = 10,$$

quantité strictement positive : la colonne $\begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$ est entrante.

On cherche alors d tel que $Bd = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$:

$$\begin{cases} -2d_1 - d_2 + d_3 = 0 \\ -d_1 - d_2 = -1 \\ -2d_1 - d_2 = -1 \end{cases} \quad \text{d'où} \quad d = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Y_B^* devient $Y_B^* - td = \begin{pmatrix} 13 \\ 26-t \\ 10-t \end{pmatrix}$; d'où $\begin{cases} t \text{ quelconque} \\ t \le 26 \\ t \le 10 \end{cases}$

On voit donc que l'on doit prendre $t = 10$ et que c'est la troisième variable qui sort, soit y_5. L'actualisation conduit à :

$$Y_B = \begin{pmatrix} y_2 \\ y_4 \\ y_0 \end{pmatrix} \qquad B = \begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix} \qquad Y_B^* = \begin{pmatrix} 13 \\ 16 \\ 10 \end{pmatrix}$$

$$Y_B = \begin{pmatrix} y_2 \\ y_4 \\ y_0 \end{pmatrix} \qquad B = \begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix} \qquad Y_B^* = \begin{pmatrix} 13 \\ 16 \\ 10 \end{pmatrix}$$

On continue à partir de cette nouvelle solution. Le vecteur $x = (x_1, x_2, x_3)$ vérifie

$$\begin{cases} -2x_1 - x_2 - 2x_3 = -150 \\ -x_1 - x_2 - x_3 = -90 \\ -x_2 - x_3 = -80 \end{cases}$$

d'où $x = (10, 30, 50)$. Il est facile ensuite de voir qu'il n'y a plus de colonne entrante : la résolution du problème dual est terminée. Comme le dual du dual est le problème primal, on a aussi résolu celui-ci ; le dernier vecteur x calculé ci-dessus donne une solution optimale du nouveau problème primal : il conviendra de produire 1000 mètres de fil F_1, 3000 mètres de fil F_2 et 5000 mètres de fil F_3 ; le profit associé sera égal à 419 200 €, il restera 60 kilogrammes de cuivre et 10 décagrammes d'étain ; les autres ressources (cadmium, force de travail et gaine) sont entièrement consommées.

4. Chapitre IV

4. Corrigé de l'exercice

1. Définissons le problème dual obtenu en relâchant les contraintes. À chaque contrainte, mise sous la forme $d_v(x) - D_v \leq 0$, on associe un multiplicateur de Lagrange $\lambda_v \geq 0$. Notons Λ le vecteur dont les composantes sont les différents multiplicateurs de Lagrange. La fonction de Lagrange $L(x, \Lambda)$ vaut alors :

$$L(x, \Lambda) = \sum_{u \in U} c(u)x(u) + \sum_{v \in V} \lambda_v \left[d_v(x) - D_v\right].$$

On en déduit l'expression de la fonction duale $w(\Lambda)$:

$$w(\Lambda) = \operatorname*{Min}_{x \in S} \left\{ \sum_{u \in U} c(u)x(u) + \sum_{v \in V} \lambda_v \left[d_v(x) - D_v\right] \right\}$$

où S est l'ensemble des arbres.

Le problème dual consiste à maximiser $w(\Lambda)$ pour $\Lambda \geq 0$, c'est-à-dire à résoudre :

$$\operatorname*{Maximiser}_{\Lambda \geq 0} \quad \operatorname*{Min}_{x \in S} \left\{ \sum_{u \in U} c(u)x(u) + \sum_{v \in V} \lambda_v \left[d_v(x) - D_v\right] \right\}$$

2. L'inconvénient majeur de la forme précédente est qu'elle fait intervenir à la fois les arêtes et les sommets du graphe. Ceci nous empêche de reconnaître un problème d'arbre de poids minimum, problème que l'on sait bien résoudre. Transformons le second terme en éliminant les sommets.

Pour cela, remarquons que le degré d'un sommet v est égal au nombre d'arêtes qui lui sont incidentes :

$$d_v(x) = \sum_{\substack{u \in U \text{ et} \\ \{i(u) = v \text{ ou } j(u) = v\}}} x(u)$$

D'où il vient :

$$\sum_{v\in V}\lambda_v\big(d_v(x)-D_v\big) = \sum_{v\in V}\lambda_v d_v(x) - \sum_{v\in V}\lambda_v D_v$$

$$= \sum_{v\in V}\ \sum_{\substack{u\in U \text{ et} \\ \{i(u)=v \text{ ou } j(u)=v\}}}\lambda_v x(u) - \sum_{v\in V}\lambda_v D_v$$

Considérons une arête u quelconque ; $x(u)$ apparaît à deux endroits dans le dernier membre de l'expression ci-dessus, correspondant aux deux extrémités de u. Par conséquent, si on regroupe les termes selon les arêtes u, x_u aura pour coefficient $\lambda_{i(u)} + \lambda_{j(u)}$. D'autre part, faire la somme sur tous les sommets possibles revient à envisager toutes les arêtes du graphe. On se retrouve donc avec l'expression :

$$\sum_{v\in V}\lambda_v\big(d_v(x)-D_v\big) = \sum_{u\in U}x(u)\big(\lambda_{i(u)}+\lambda_{j(u)}\big) - \sum_{v\in V}\lambda_v D_v$$

et $$L(x, \Lambda) = \sum_{u\in U}x(u)\big(c(u)+\lambda_{i(u)}+\lambda_{j(u)}\big) - \sum_{v\in V}\lambda_v D_v\ .$$

Le problème dual devient :

$$\underset{\Lambda\geq 0}{\text{Maximiser}}\ \ \underset{x\in S}{\text{Min}}\left\{\sum_{u\in U}x(u)\big(c(u)+\lambda_{i(u)}+\lambda_{j(u)}\big)-\sum_{v\in V}\lambda_v D_v\right\}.$$

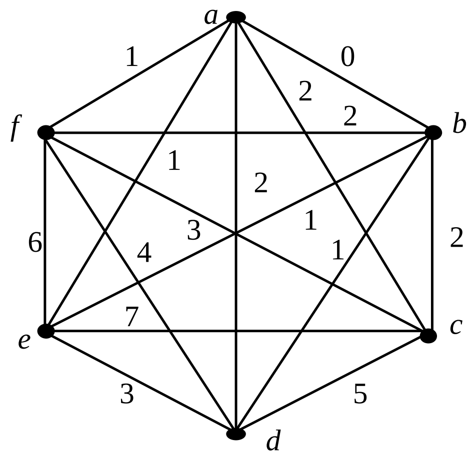
On constate que le terme : $\underset{x\in S}{\text{Min}}\left\{\sum_{u\in U}x(u)\big(c(u)+\lambda_{i(u)}+\lambda_{j(u)}\big)\right\}$ conduit à chercher un arbre de poids minimum pour les valuations $c(u)+\lambda_{i(u)}+\lambda_{j(u)}$, lesquelles changent à chaque itération en fonction des λ_v. On pourra par conséquent appliquer un algorithme calculant un arbre couvrant de poids minimum, par exemple celui de Kruskal, ou encore celui de Prim...

3. Pour traiter l'application, nous utiliserons les notations du chapitre IV. En particulier :

- $\Lambda = (\lambda_a, \lambda_b)$;
- z^* représente l'optimum du problème primal ;
- w^* représente le maximum de la fonction duale w ;
- si on considère $k - 1$ arbres $x_1, \ldots, x_{k-1}$, alors :
  $w_k(\Lambda) = \text{Min}\{L(x_1, \Lambda), \ \ldots, L(x_{k-1}, \Lambda)\}$;
- on note w_k^* le maximum de $w_k(\Lambda)$ et Λ_k la valeur de Λ telle que $w_k(\Lambda_k) = w_k^*$.

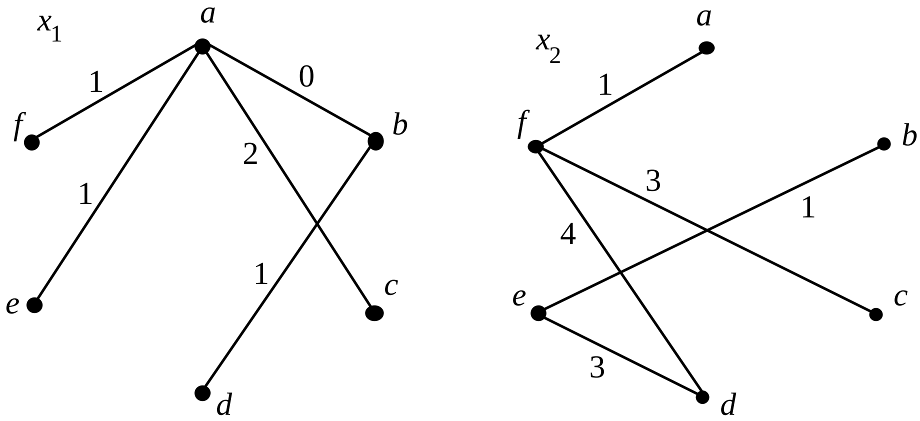
On rappelle l'encadrement : $w(\Lambda_k) \leq w^* \leq w_k^*$.

L'application consiste, au moyen de la relaxation lagrangienne, à chercher dans le graphe ci-contre un arbre couvrant de poids minimum (« APM ») parmi les arbres couvrants x respectant $d_a(x) \leq 2$ et $d_b(x) \leq 2$, ou au moins d'obtenir un minorant de l'optimum cherché.



INITIALISATION

On détermine un « petit » nombre d'arbres de façon à initialiser la relaxation. Nous choisissons ici un arbre de plus petit coût x_1 et un arbre x_2 dont les degrés en a et en b soient minimum, c'est-à-dire prennent la valeur 1.



On a $c(x_1) = 5$ et $c(x_2) = 12$.
Toute solution à notre problème aura un coût au moins égal à la solution de coût minimum : $5 \leq z^*$; l'arbre x_2 de coût 12 est réalisable : $z^* \leq 12$.
En ce qui concerne w^*, on peut remarquer que $w(0) = c(x_1) = 5$; comme on a $w^* \geq w(0)$, il vient $w^* \geq 5$. Par ailleurs, $w^* \leq z^*$: $w^* \leq 12$.
On peut en fait considérer que x_1 est un APM obtenu pour $\Lambda = 0$, et x_2 un arbre correspondant à des valeurs très grandes de λ_a et λ_b de telle sorte que le coût devienne négligeable par rapport aux contraintes dans la fonction de Lagrange. On obtient :

$$L(x_1, \Lambda) = \mathrm{c}(x_1) + \lambda_a[d_a(x_1) - D_a] + \lambda_b[d_b(x_1) - D_b] = 5 + 2\lambda_a$$

$$L(x_2, \Lambda) = \mathrm{c}(x_2) + \lambda_a[d_a(x_2) - D_a] + \lambda_b[d_b(x_2) - D_b] = 12 - \lambda_a - \lambda_b$$

ÉTAPE 1

a. On résout Maximiser Min$\{L(x_1, \Lambda), L(x_2, \Lambda)\}$ pour $\Lambda \geq 0$. On pose :

$$w_3(\Lambda) = \mathrm{Min}[L(x_1, \Lambda), L(x_2, \Lambda)]$$

On est donc amené à résoudre :

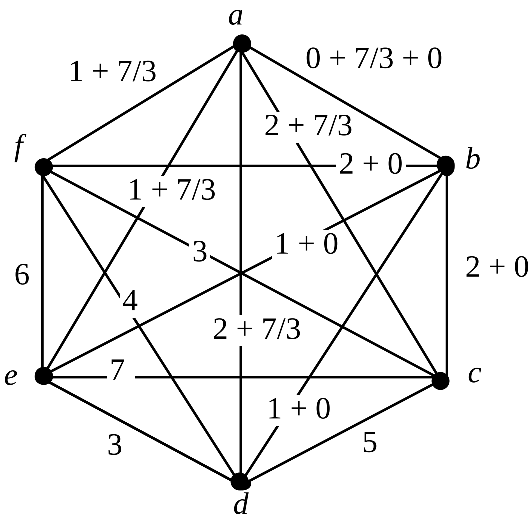
$$\text{Maximiser } w$$

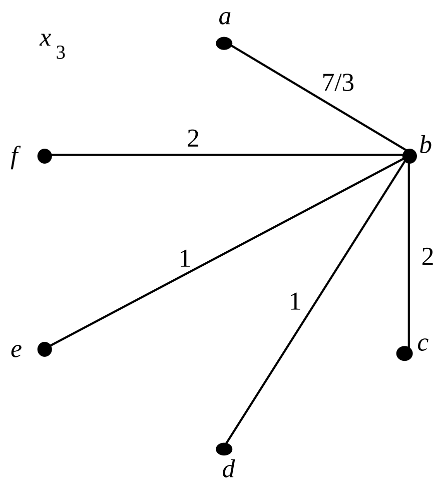
avec les contraintes $\begin{cases} w - 2\lambda_a \leq 5 \\ w + \lambda_a + \lambda_b \leq 12 \\ \lambda_a \geq 0, \lambda_b \geq 0 \end{cases}$

On peut résoudre ce problème à l'aide de l'algorithme du simplexe (voir les détails de calcul plus loin, à la fin de l'exercice). On trouve alors comme valeurs optimales de ce problème :

$$w_3^* = \frac{29}{3} \quad \text{et} \quad \Lambda_3 = \left(\frac{7}{3}, 0\right).$$

b. On calcule ensuite l'arbre x_3 minimisant $L(x, \Lambda_3) = L\left(x, \left(\frac{7}{3}, 0\right)\right)$. C'est l'APM du graphe suivant (à gauche), dont les poids valent $c(u) + \lambda_{i(u)} + \lambda_{j(u)}$; x_3 est représenté à droite :





Le poids de x_3 pour les nouvelles valuations vaut $\frac{25}{3}$; on a donc :

$$w(\Lambda_3) = L(x_3, \Lambda_3) = \frac{25}{3} - 2\lambda_a - 2\lambda_b = \frac{11}{3}$$

c. Pour le problème dual, on a actuellement : Max(5, $\frac{11}{3}$) = 5 $\leq w^* \leq \frac{29}{3}$.

Pour le problème primal, l'arbre x_3 n'étant pas réalisable, on n'a pas amélioré l'encadrement : $5 \leq z^* \leq 12$. Enfin, le coût de x_3 étant 6, les degrés des sommets a et b étant respectivement égaux à 1 et 5, la fonction de Lagrange associée à x_3 est $L(x_3, \Lambda) = 6 - \lambda_a + 3\lambda_b$.

ÉTAPE 2

a. On résout Maximiser Min$\{L(x_1, \Lambda), L(x_2, \Lambda), L(x_3, \Lambda)\}$ pour $\Lambda \geq 0$. On pose :

$$w_4(\Lambda) = \text{Min}\{L(x_1, \Lambda), L(x_2, \Lambda), L(x_3, \Lambda)\}$$

On est donc amené à résoudre :

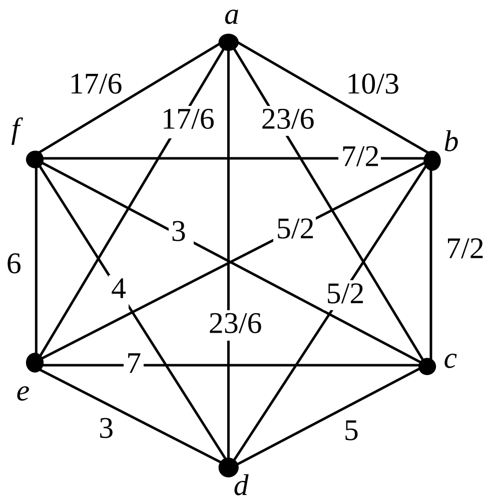
$$\text{Maximiser } w$$

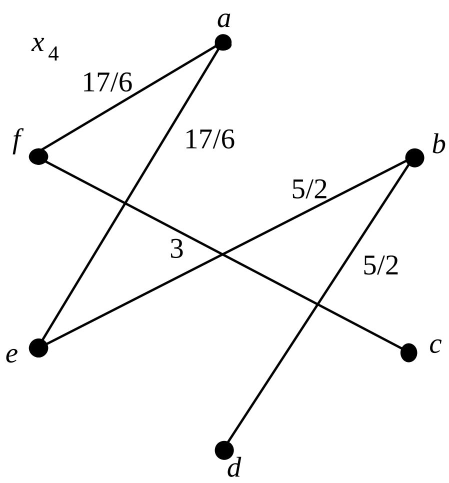
avec les contraintes
$$\begin{cases} w - 2\lambda_a \leq 5 \\ w + \lambda_a + \lambda_b \leq 12 \\ w + \lambda_a - 3\lambda_b \leq 6 \\ \lambda_a \geq 0, \lambda_b \geq 0 \end{cases}$$

L'algorithme du simplexe (voir détails plus loin) donne pour solution optimale :

$$w_4^* = \frac{26}{3} \quad \text{et} \quad \Lambda_4 = \left(\frac{11}{6}, \frac{3}{2}\right).$$

b. On calcule ensuite l'arbre x_4 minimisant $L(x, \Lambda_4) = L\left(x, \left(\frac{11}{6}, \frac{3}{2}\right)\right)$. C'est l'APM du graphe ci-dessous (à gauche), dont les poids sont donnés par $c(u) + \lambda_{i(u)} + \lambda_{j(u)}$; x_4 est représenté à droite.





Le coût de x_4 pour les nouvelles valuations est 41/3, d'où :

$$w(\Lambda_4) = L(x_4, \Lambda_4) = \frac{41}{3} - 2\frac{11}{6} - 2\frac{3}{2} = 7.$$

c. Pour le problème dual, on a donc : $\text{Max}(5, \frac{11}{3}, 7) = 7 \leq w^* \leq \frac{26}{3}$.

Pour le problème primal, l'arbre x_4 étant réalisable, on a $z^* \leq 7$; d'où il vient :

$$7 \leq w^* \leq z^* \leq 7 : z^* = w^* = 7.$$

La résolution des deux problèmes, primal et dual, est terminée et x_4 est un arbre optimum du problème primal.

Clairement, il en sera de même dès que, pour une valeur donnée de Λ, on découvre en calculant $w(\Lambda)$ que l'égalité $w(\Lambda) = L(x, \Lambda)$ est vraie, où x est une solution qui sature exactement les contraintes. En effet, on a alors :

$$w(\Lambda) \leq w^* \leq z^* \leq c(x) = L(x, \Lambda) = w(\Lambda)$$

(puisque les contraintes sont exactement saturées) ; d'où il s'ensuit : $w^* = z^* = w(\Lambda)$.

DETAILS DES CALCULS POUR LES ETAPES 1 ET 2

On trouvera ici les calculs du simplexe des étapes 1 et 2. Afin de varier les plaisirs, la présentation du dictionnaire a été utilisée pour la première itération, et la forme matricielle pour la seconde.

Pour la première étape, on doit résoudre :

Maximiser w

avec les contraintes
$$\begin{cases} w - 2\lambda_a \leq 5 \\ w + \lambda_a + \lambda_b \leq 12 \\ \lambda_a \geq 0, \lambda_b \geq 0 \end{cases}$$

Ce problème n'est pas sous forme standard : il n'y a pas de contrainte sur le signe de w, mais il est facile de voir que $w = 0$ peut être atteint. On peut alors ajouter au problème la contrainte $w \geq 0$ sans modifier le maximum de w. On obtient ainsi un problème sous forme standard dont le premier dictionnaire s'écrit de la façon suivante, en appelant t la fonction objectif, et x_1 et x_2 les variables d'écart :

$$\begin{array}{l} x_1 = 5 - \quad w \; + 2\lambda_a \\ \underline{x_2 = 12 - w \quad - \lambda_a - \lambda_b} \\ t = \qquad w \end{array}$$

w entre en base et x_1 en sort. On obtient le nouveau dictionnaire :

$$\begin{array}{l} w = 5 - x_1 + 2\lambda_a \\ \underline{x_2 = 7 + x_1 - 3\lambda_a - \lambda_b} \\ t = 5 - x_1 + 2\lambda_a \end{array}$$

C'est maintenant au tour de λ_a d'entrer en base, que x_2 quitte pour donner :

$$\begin{array}{l} w = \frac{29}{3} - \frac{1}{3}x_1 - \frac{2}{3}x_2 - \frac{2}{3}\lambda_b \\ \lambda_a = \frac{7}{3} + \frac{1}{3}x_1 - \frac{1}{3}x_2 - \frac{1}{3}\lambda_b \\ \hline t = \frac{29}{3} - \frac{1}{3}x_1 - \frac{2}{3}x_2 - \frac{2}{3}\lambda_b \end{array}$$

Tous les coefficients de la dernière ligne étant négatifs, on a atteint l'optimum. Celui-ci correspond à : $w = \frac{29}{3}$, $\lambda_a = \frac{7}{3}$, $\lambda_b = 0$.

Pour la seconde étape, on doit résoudre :

Maximiser w

avec les contraintes $\begin{cases} w - 2\lambda_a \le 5 \\ w + \lambda_a + \lambda_b \le 12 \\ w + \lambda_a - 3\lambda_b \le 6 \\ \lambda_a \ge 0, \lambda_b \ge 0 \end{cases}$

problème auquel on peut ajouter, comme précédemment, la contrainte $w \ge 0$ afin qu'il prenne une forme standard. En posant $X = (w, \lambda_a, \lambda_b, x_1, x_2, x_3)^t$ où les x_i $(1 \le i \le 3)$ désignent les variables d'écart, $\beta = (5, 12, 6)^t$, $k = (1, 0, 0, 0, 0, 0)$ et

$$M = \begin{pmatrix} 1 & -2 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & -3 & 0 & 0 & 1 \end{pmatrix}$$

notre problème peut désormais s'écrire :

Maximiser $k.X$

avec les contraintes $\begin{cases} MX \le \beta \\ X \ge 0 \end{cases}$

1. On peut choisir la matrice identité (correspondant aux trois variables d'écart) comme matrice B inversible extraite de la matrice des contraintes M. En adaptant les notations du chapitre II, on obtient alors $X_B = \beta$.

2. On cherche un vecteur-ligne y tel que $yB = k_B = (0\ 0\ 0)$; on trouve $y = (0\ 0\ 0)$. Puis on cherche une colonne m_j de M telle que $ym_j < k_j$: $m_j = (1\ 1\ 1)^t$ convient : w est variable entrante. On cherche ensuite δ (vecteur-colonne) tel que $B\delta = m_j$: $\delta = m_j = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. X_B devient alors $X_B - t\delta$, où t est la valeur que va prendre w :

$$X_B - t\delta = \begin{pmatrix} 5 \\ 12 \\ 6 \end{pmatrix} - t\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5-t \\ 12-t \\ 6-t \end{pmatrix}$$

La plus grande valeur que peut prendre t pour que toutes les variables restent positives ou nulles est 5 : c'est alors x_1 qui sort de la base. On obtient donc :

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ et } X_B{}^* = \begin{pmatrix} 5 \\ 7 \\ 1 \end{pmatrix}$$

3. On recommence : recherche de y tel que $yB = k_B = (1\ \ 0\ \ 0)$; on trouve $y = (1\ \ 0\ \ 0)$. Recherche d'une colonne entrante m_j telle que $ym_j < k_j$: $m_j = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ convient : λ_a entre en base. Recherche de δ tel que $B\delta = m_j$: on trouve $\delta = \begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix}$. X_B devient :

$$X_B - t\delta = \begin{pmatrix} 5 \\ 7 \\ 1 \end{pmatrix} - t\begin{pmatrix} -2 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 5+2t \\ 7-3t \\ 1-3t \end{pmatrix}$$

D'où $t = \dfrac{1}{3}$ et x_3 sort de la base : et $B = \begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ et $X_B{}^* = \begin{pmatrix} 17/3 \\ 6 \\ 1/3 \end{pmatrix}$

4. Recherche de y tel que $yB = k_B = (1\ \ 0\ \ 0)$; on trouve $y = (\frac{1}{3}\ \ 0\ \ \frac{2}{3})$. Puis recherche d'une colonne entrante m_j telle que $ym_j < k_j$: $m_j = \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix}$ convient : λ_b entre en base. Recherche de δ tel que $B\delta = m_j$: on trouve $\delta = \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}$. X_B devient :

$$X_B - t\delta = \begin{pmatrix} 17/3 \\ 6 \\ 1/3 \end{pmatrix} - t\begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}$$

D'où $t = \dfrac{3}{2}$ et x_2 sort de la base : $B = \begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & -3 & 1 \end{pmatrix}$ et $X_B{}^* = \begin{pmatrix} 26/3 \\ 3/2 \\ 11/6 \end{pmatrix}$.

5. Recherche de y tel que $y.B = k_B = (1\ \ 0\ \ 0)$; on trouve $y = (\frac{1}{3}\ \ \frac{1}{2}\ \ \frac{1}{6})$. On ne peut plus alors trouver de colonne entrante : $y.m_j \geq k_j$ pour tout j. L'optimum est atteint : la matrice B étant constituée des colonnes 1, 3 et 2 de M, les composantes de X_B donnent les valeurs de w, de λ_b et de λ_a à l'optimum : $w = \dfrac{26}{3}$, $\lambda_b = \dfrac{3}{2}$ et $\lambda_a = \dfrac{11}{6}$.

5. Chapitre V

5. Corrigé de l'exercice

$$\nabla f(x,y) = \begin{pmatrix} e^{x+y} + 2x \\ e^{x+y} + 4y \end{pmatrix},\ \nabla^2 f(x,y) = \begin{pmatrix} e^{x+y} + 2 & e^{x+y} \\ e^{x+y} & e^{x+y} + 4 \end{pmatrix}$$

Le déterminant de la hessienne ainsi que sa trace étant positives, les valeurs propres de $\nabla^2 f(x, y)$ sont positives, $\nabla^2 f(x, y)$ est définie positive : f est donc convexe.
On en déduit que tout minimum local est global, or une condition nécessaire et suffisante pour que x* soit un minimum local 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 $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$; $\nabla f(x_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; $d_0 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

On cherche λ tel que $g(\lambda) = f(x_0 + \lambda d_0)$ 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 $\lambda = 0{,}216$.

$x_1 = \begin{pmatrix} -0{,}216 \\ -0{,}216 \end{pmatrix}$; $\nabla f(x_1) = \begin{pmatrix} 0{,}216 \\ -0{,}216 \end{pmatrix}$; $d_1 = \begin{pmatrix} -0{,}216 \\ 0{,}216 \end{pmatrix}$.

On constate que d_1 est orthogonal à d_0.

$g(\lambda) = f(x_1 + \lambda d_1) = f(-0{,}216(1 + \lambda), -0{,}216(1 - \lambda))$
$\quad = e^{-2\times 0{,}216} + (0{,}216)^2\,[(1 + \lambda)^2 + 2(1 - \lambda)^2]$

$h(\lambda) = (1+\lambda)^2 + 2(1 - \lambda)^2 = 3\lambda^2 - 2\lambda + 3$, est minimum pour $\lambda = \frac{1}{3}$

$$x_2 = \begin{pmatrix} -0{,}216(1+\frac{1}{3}) \\ -0{,}21628(1-\frac{1}{3}) \end{pmatrix} = \begin{pmatrix} -0{,}288 \\ -0{,}144 \end{pmatrix};\ \nabla f(x_2) = \begin{pmatrix} 0{,}0732 \\ 0{,}0732 \end{pmatrix};\ d_2 = \begin{pmatrix} -0{,}0732 \\ -0{,}0732 \end{pmatrix}$$

$g(\lambda) = f(-0{,}288 - \lambda, -0.144 - \lambda) = e^{-0.432-2\lambda} + (0{,}288 + \lambda)^2 + 2(0{,}144 + \lambda)^2$. On minimise par Newton, et on trouve $\lambda = 0{,}07119$, d'où :

$x_3 = \begin{pmatrix} -0{,}305 \\ -0{,}161 \end{pmatrix}$.

On peut continuer ainsi pour avoir plus de précision.

6. Chapitre VI

6.1. Corrigé de l'exercice 1

1. Soit $f(x, y) = (2x^2 + y^4)$, $h_1(x, y) = x - 1$ et $h_2(x, y) = x + ay - a - 1$.
2. Le problème s'écrit

Minimiser $f(x, y)$
avec les contraintes :
$$\begin{cases} h_1(x,y) \geq 0 \\ h_2(x,y) \geq 0 \end{cases}$$

$$\nabla^2 f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 \end{pmatrix}$$ qui est définie positive : la fonction f est convexe sur ℝ.

Par ailleurs, une fonction linéaire est concave (et d'ailleurs aussi convexe) : h_1 et h_2 sont concaves.

Il faut et il suffit donc que la condition de 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 Kuhn et Tucker est vérifiée, c'est montrer qu'il existe deux réels positifs ou nuls μ_1 et μ_2 tels que :

$$\nabla f(1, 1) = \mu_1 \nabla h_1(1, 1) + \mu_2 \nabla h_2(1, 1)$$

$$\nabla f(x,y) = \begin{pmatrix} 4x \\ 4y^3 \end{pmatrix} \qquad \nabla f(1,1) = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

$$\nabla h_1(1,\ 1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \nabla h_2(1,\ 1) = \begin{pmatrix} 1 \\ a \end{pmatrix}$$

Cherchons des coefficients μ_1 et μ_2 tels que :

$(4, 4) = \mu_1(1, 0) + \mu_2(1, a)$: $\begin{cases} \mu_1 + \mu_2 = 4 \\ a\mu_2 = 4 \end{cases}$

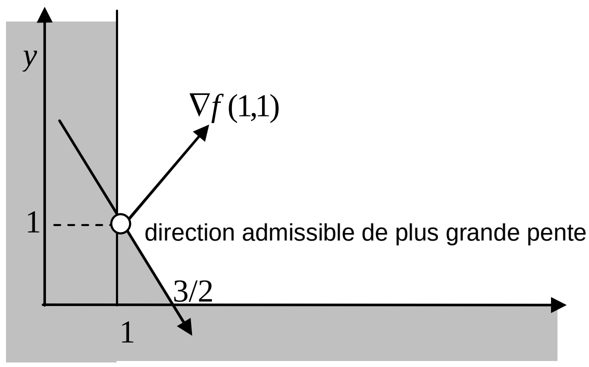
Pour qu'il y ait une solution, il faut et il suffit que $a \neq 0$ et alors :

$$\mu_1 = 4(1 - \frac{1}{a}),\ \mu_2 = \frac{4}{a}.$$

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 $\boxed{a \geq 1}$.

2. C'est le problème de la question 1. avec $a = \frac{1}{2}$: 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 descente pour f en ce point.



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 $g'' = 4 + 48(1 - 2t)^2 > 0$ donc $g'(t)$ est croissante.

$g'(0) = -4 < 0$; $g'(\frac{1}{2}) = 6 > 0$. 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.

$\nabla f((1{,}0914 ; 0{,}8172)) = \begin{pmatrix} 4{,}3656 \\ 2{,}18296 \end{pmatrix}$ et $\nabla h_2((1{,}0914 ; 0{,}8172)) = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = \dfrac{1}{4{,}3656}\begin{pmatrix} 4{,}3656 \\ 2{,}18296 \end{pmatrix}$.

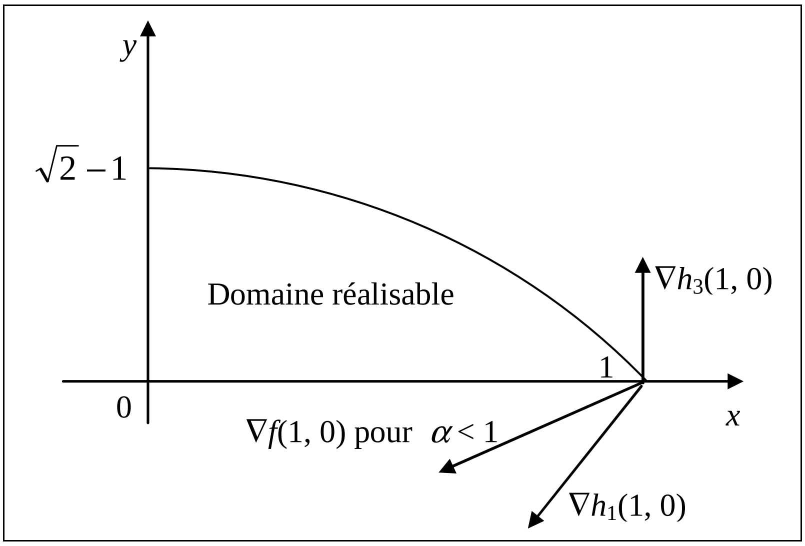
6.2. Corrigé de l'exercice 2

1. 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 : $h_1(x, y) = 2 - x^2 - (y + 1)^2$, $h_2(x, y) = x$, $h_3(x, y) = y$.

Le problème s'écrit :
$$\text{Minimiser } f_\alpha(x, y) \quad \text{avec} \begin{cases} h_1(x, y) \geq 0 \\ h_2(x, y) \geq 0 \\ h_3(x, y) \geq 0 \end{cases}$$

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 négatives. La fonction h_1 est concave, les fonctions h_2 et h_3 sont concaves car linéaires. La condition de 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 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 $\nabla f(1, 0) = \mu_1 \nabla h_1(1, 0) + \mu_3 \nabla h_3(1, 0)$. 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 $\nabla f(1, 0) = \mu_1 \nabla h_1(1, 0) + \mu_3 \nabla h_3(1, 0)$ s'écrit :

$$\begin{cases} -2 = -2\mu_1 \\ -2\alpha = -2\mu_1 + \mu_3 \end{cases} \text{ ou encore : } \begin{cases} \mu_1 = 1 \\ \mu_3 = 2(1-\alpha) \end{cases}.$$

Le point (1, 0) est un minimum local (et même un minimum global) si et seulement si : $\alpha \leq 1$ (en gardant l'hypothèse $\alpha \geq 0$).

.

2. Avant de relâcher la contrainte $x^2 + (y + 1)^2 \leq 2$, on l'écrit :

$$x^2 + (y + 1)^2 - 2 \leq 0.$$

La fonction de Lagrange est définie, pour $\lambda \geq 0$, par :

$$L(x, y, \lambda) = (x - 2)^2 + 2(y - 1)^2 + \lambda(x^2 + (y + 1)^2 - 2)$$

ou encore : $L(x, y, \lambda) = [(x - 2)^2 + \lambda x^2] + [2(y - 1)^2 + \lambda(y + 1)^2] - 2\lambda$

La fonction duale est définie, pour $\lambda \geq 0$, par :

$w(\lambda) = \text{minimum}_{x\geq 0,\, y\geq 0} L(x, y, \lambda)$.

Le problème dual est : maximiser $w(\lambda)$, avec $\lambda \geq 0$.

3. $w(1) = \text{minimum}_{x\geq 0,\, y\geq 0} L(x, y, 1)$.

$L(x, y, 1) = [(x - 2)^2 + x^2] + [2(y - 1)^2 + (y + 1)^2] - 2$.

La fonction $g(x) = (x - 2)^2 + x^2$ a pour dérivée : $g'(x) = 4x - 4$, et est donc minimum pour $x = 1$; son minimum pour $x \geq 0$ vaut 2.

La fonction $k(y) = 2(y - 1)^2 + (y + 1)^2$ a pour dérivée : $k'(x) = 6y - 2$, et est donc minimum pour $y = 1/3$; son minimum pour $y \geq 0$ vaut 25/9.

D'où : $w(1) = 2 + 24/9 - 2$: $w(1) = 24/9 = 8/3$.

$w(3) = \text{minimum}_{x\geq 0,\, y\geq 0} L(x, y, 3)$.

$L(x, y, 3) = [(x - 2)^2 + 3x^2] + [2(y - 1)^2 + 3(y + 1)^2] - 6$.

La fonction $g(x) = (x - 2)^2 + 3x^2$ a pour dérivée : $g'(x) = 8x - 4$, et est donc minimum pour $x = 1/2$; son minimum pour $x \geq 0$ vaut 3.

La fonction $k(y) = 2(y - 1)^2 + 3(y + 1)^2$ a pour dérivée : $k'(x) = 10y + 2$, et est donc minimum pour $y = -1/5$; l'étude des variations de k montre que son minimum pour $y \geq 0$ est atteint pour $y = 0$ et vaut 5.

D'où : $w(3) = 3 + 5 - 6$: $w(3) = 2$.

Comme le maximum de w minore le minimum du problème (P_2), le minimum de (P_2) est au moins égal à $w(1)$ et $w(3)$; le minimum du problème (P_2) vaut au moins 8/3.

Index