OPTIMISATION LINÉAIRE

HÉDI NABLI

[1]

Ce cours d'optimisation linéaire est destiné à des étudiants en Mathématiques apliquées ou Informatique, niveau troisième année universitaire. Il est basé essentiellement sur la référence ci-après :

> Hédi Nabli, "Recherche Opérationnelle : Algorithme du Simplexe et ses Applications", *Centre de Publication Universitaire*, Tunisie (2006)

Ce livre propose entre autres

(1) une nouvelle méthode de recherche d'une base réalisable initiale pour l'algorithme du simplexe. Cette technique ne fait pas intervenir les variables artificielles. De plus il est souvent possible d'aboutir, en une seule itération, à une base réalisable.

(2) Grâce aux tableaux formels, une nouvelle alternative, autre que l'algorithme dual-simplexe, est proposée.

(3) Pour les problèmes de transport, un algorithme récursif pour la détermination des éléments d'un tableau de transport est mis en oeuvre.

Dans ce document, les résultats mathématiques sont donnés sans démonstration. Pour plus de détails, on peut consulter la référence ci-dessus.

1. PROGRAMMATION LINÉAIRE

Un **programme linéaire** est un problème dans lequel on est amené à maximiser (ou minimiser) une application linéaire, appelée *fonction d'objectif* ou *fonction économique*, sur un ensemble d'équations et/ou d'inéquations linéaires, dites *contraintes*. Autrement dit, la **programmation linéaire** est une branche des mathématiques qui a pour but de résoudre des problèmes d'optimisation linéaire de type

$$\begin{cases} \max(\text{ou } \min)[Z(x_1,\ldots,x_p) = \sum_{j=1}^{p} c_j x_j] \\ \sum_{j=1}^{p} a_{ij}x_j \leq (\text{et/ou } =)\ b_i, \text{ pour } i = 1,\ldots,m. \end{cases}$$

Les coefficients c_j, a_{ij} et b_i sont des réels fixés et les x_j sont des variables réelles. Les contraintes d'inégalités éventuelles sont toutes larges et non strictes. Il se peut qu'une contrainte d'inégalité soit de type " $\geq$ ". En multipliant chaque inégalité de type " $\geq$ " par (-1), on peut supposer que toutes les contraintes d'inégalité sont de type " $\leq$ ".

1.1. Généralités.

Dans le contexte de la programmation linéaire, le terme programmation désigne l'organisation des calculs et non la réalisation d'un programme informatique. Du point de vue des applications, l'optimisation linéaire est d'une grande portée. Elle s'applique à des problèmes très variés qui sont issus de l'économie, de l'ingénierie, de la physique ou encore des modèles probabilistes. Dans ce cadre, on peut citer par exemple, les problèmes de type **gestion de stock, gestion de production, transport de marchandise, affectation du personnel, systèmes industriels, réseaux de communication**, etc. Pour les modèles de programmation linéaire, on est souvent amené à maximiser un gain ou minimiser un coût. Ceci explique d'ailleurs pourquoi la fonction à maximiser s'appelle fonction d'objectif ou économique.

[1]Ce travail a été réalisé à l'occasion d'un projet Tempus, action JEP-31147-2003, impliquant d'une part l'université Paris-Sud, l'université de Lille (USTL) et l'université de Delft (TU Delft) et d'autre part l'université de Monastir (ISM et FSM) et l'université de Sousse (ISITC)

1.2. Exemple.

Exemple 1. Une usine fabrique deux produits (A) et (B) à l'aide des matières premières I, II et III. Le fonctionnement de l'usine est comme suit :

- 1 unité du produit (A) nécessite 2 unités de I et 1 unité de II.
- 1 unité du produit (B) nécessite 1 unité de I, 2 unités de II et 1 unité de III.

On suppose que l'usine dispose des matières premières I, II et III en quantités respectives 8, 7 et 3. Le profit dû à la fabrication d'une unité du produit (A) (resp. (B)) est égal à 4 (resp. 5) Dinars Tunisiens (DT). L'objectif étant de maximiser le profit tout en respectant les contraintes sur la matière première.

Formulation mathématique 1. Si on désigne respectivement par x_1 et x_2 les quantités vendues du produit (A) et (B), le gain total vaut

$$Z(x_1,x_2) = 4x_1 + 5x_2$$

D'autre part, la disponibilité en matières premières revient à exiger des quantités, utilisées pour la fabrication de x_1 et x_2, qui soient inférieures aux quantités disponibles. Ces contraintes s'expriment par les inégalités suivantes :

$$\begin{cases} 2x_1 + x_2 & \leq 8 \\ x_1 + 2x_2 & \leq 7 \\ x_2 & \leq 3 \end{cases}$$

Bien entendu, les variables x_1 et x_2 doivent être positives. En conclusion, le problème de maximisation du profit se traduit mathématiquement par un programme linéaire qui s'écrit sous la forme :

$$\begin{cases} \max[Z(x_1,x_2) = & 4x_1 + 5x_2] \\ 2x_1 + x_2 & \leq 8 \\ x_1 + 2x_2 & \leq 7 \\ x_2 & \leq 3 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

1.3. **Domaine réalisable.** Pour un problème d'optimisation linéaire, tout point qui vérifie l'ensemble des contraintes s'appelle point *réalisable* ou *admissible*. L'ensemble de tous les points réalisables s'appelle *domaine réalisable*. Il est facile de vérifier que le domaine réalisable d'un programme linéaire est convexe. On rappelle qu'un ensemble est dit *convexe* si à chaque fois qu'il contient deux points x et y, il contient le segment joignant x et y. Ce segment est souvent noté par $[x,y]$ et il est défini par

$$[x,y] = \{\lambda x + (1-\lambda)y \ / \ \lambda \in [0,1]\}$$

Un point est dit *optimal* s'il est admissible et s'il réalise l'optimum de la fonction d'objectif sur le domaine réalisable. Par exemple, pour le problème

$$\begin{cases} \max[Z(x_1,x_2) = x_1 - x_2] \\ x_1 + x_2 \leq 1 \\ x_2 \geq 0, \end{cases}$$

tout point réalisable (x_1,x_2) vérifie

$$Z(x_1,x_2) = x_1 - x_2 \leq x_1 \leq x_1 + x_2 \leq 1.$$

D'autre part, le point $(1,0)$ est réalisable et on a $Z(1,0) = 1$. Donc, la solution réalisable $(1,0)$ est optimale et la valeur maximale de Z sur le domaine réalisable est $Z(1,0) = 1$.

1.4. **Présentation des méthodes.** Nous allons présenter des techniques qui permettent de résoudre les programmes linéaires que l'on convient désormais de noter (PL). Diverses méthodes ont été proposées dans la littérature :

- **La méthode graphique** : l'utilisation de cette méthode est restreinte aux (PL) ayant un nombre de variables au plus égal à 3.
- **La méthode des sommets** : le nombre de sommets étant prohibitif, cette méthode est très coûteuse en temps de calcul.
- **La méthode du simplexe** : algorithme itératif mis au point par George Dantzig en 1951.

2. Méthode graphique

Nous allons décrire la méthode graphique à travers trois exemples représentatifs, ayant chacun deux variables structurelles. Le premier admet une et une seule solution optimale. Le deuxième admet une infinité de solutions optimales. Par contre, le troisième exemple n'admet aucune solution optimale.

La résolution d'un (PL) par la méthode graphique se fait selon une démarche commune, celle-ci peut être résumée en quatre directives :

(1) Schématiser le domaine réalisable dans un repère orthonormé.

(2) Dans le même repère, représenter l'hyperplan $\Delta_0\ : Z(x) = 0$. Noter que tout hyperplan Δ_r d'équation $Z(x) = r$, $r \in \mathbb{R}$, est parallèle à Δ_0.

(3) Si le (PL) considéré est de type maximisation (*resp.* minimisation), indiquer, par une flèche, le sens de déplacement parallèle de Δ_r qui fait augmenter (*resp.* diminuer) r.

(4) Déterminer la droite $\Delta_{r_{\max}}$ la plus éloignée (dans le sens de la flèche) de Δ_0 et qui coupe le domaine réalisable $\mathcal{R}$. Tout point de $\Delta_{r_{\max}} \cap \mathcal{R}$ est une solution optimale. Si une telle droite $\Delta_{r_{\max}}$ n'existe pas, cela veut dire que l'optimum est infini.

Appliquons ce principe général sur l'exemple 1 défini dans le paragraphe précédent.

2.1. Exemple 1.

Reprenons l'exemple 1

$$\left\{\begin{array}{rl} \max[Z(x_1,x_2) & = 4x_1 + 5x_2] \\ 2x_1 + x_2 & \leq 8 \\ x_1 + 2x_2 & \leq 7 \\ x_2 & \leq 3 \\ x_1 \geq 0 & x_2 \geq 0 \end{array}\right.$$

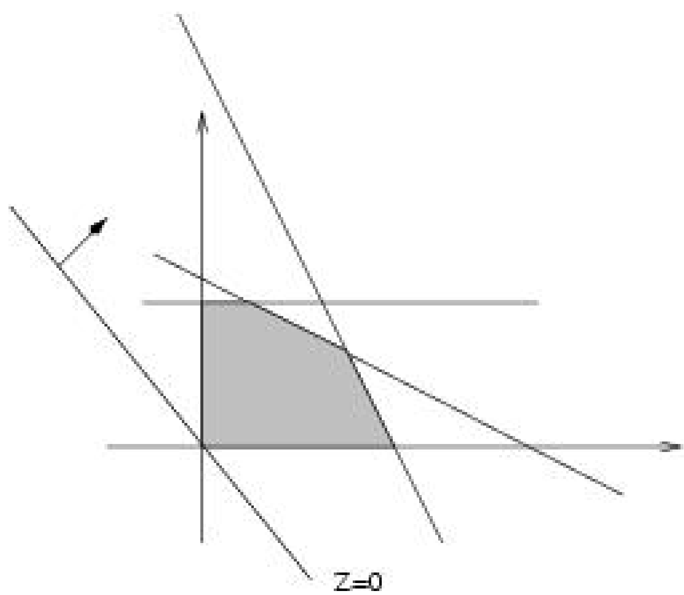

Fig. 1. Représentation graphique pour l'exemple 1

Le domaine réalisable est schématisé en gris sur la Figure. Toute droite Δ_r d'équation $Z(x_1,x_2) = r$, où r est un réel fixé, est parallèle à la droite $\Delta_0 : Z(x_1,x_2) = 0$. On constate aussi que tout déplacement parallèle de Δ_r dans le sens de la flèche (voir dessin ci-après) fait augmenter r. Donc maximiser la fonction Z, tout en satisfaisant aux contraintes du problème, revient à chercher la droite Δ_r la plus éloignée (dans le sens de la flèche) de Δ_0 et qui coupe le domaine réalisable. Du point de vue graphique, il s'agit de la droite $\Delta_{r_{\max}}$ passant par le point $(3,2)$ et parallèle à Δ_0. Ainsi, on peut conclure que le point $(3,2)$ est la seule solution optimale qui donne une valeur maximale de la fonction d'objectif égale à $Z(3,2) = 22$. Autrement dit, le meilleur profit vaut 22DT obtenu en fabriquant 3 unités du produit (A) et 2 unités du produit (B).

2.2. **Exemple 2.** On considère le (PL) suivant :

$$\begin{cases} \max[Z(x_1,x_2) & = -6x_1+15x_2] \\ 2x_1-x_2 & \geq 4 \\ -2x_1+5x_2 & \leq 4 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

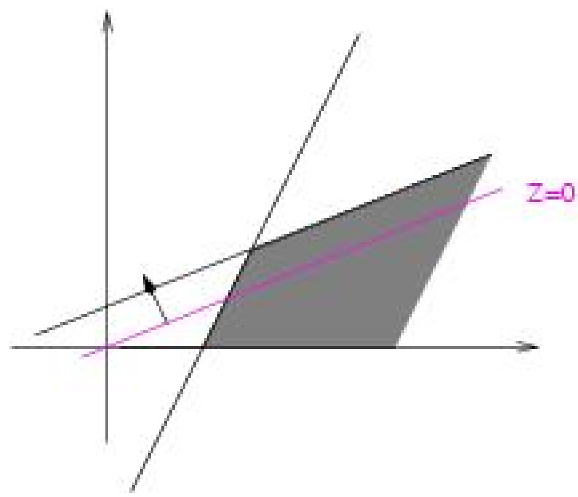

FIG. 2. Représentation graphique pour l'Exemple 2

Le domaine réalisable est la partie non bornée représentée partiellement en gris sur la Figure. Tout déplacement parallèle de $\Delta_r : Z(x_1,x_2) = r$ dans le sens de la flèche (voir dessin ci-après) fait augmenter r. Donc tout point de la demi-droite $[A,B)$ est optimal et la valeur maximale de Z vaut $Z(A) = Z(3,2) = 12$ (voir dessin ci-après).

2.3. **Exemple 3.** On considère maintenant le (PL) suivant :

$$\begin{cases} \max[Z(x_1,x_2) & = x_1+x_2] \\ -x_1+x_2 & \leq 1 \\ -x_1+2x_2 & \leq 4 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

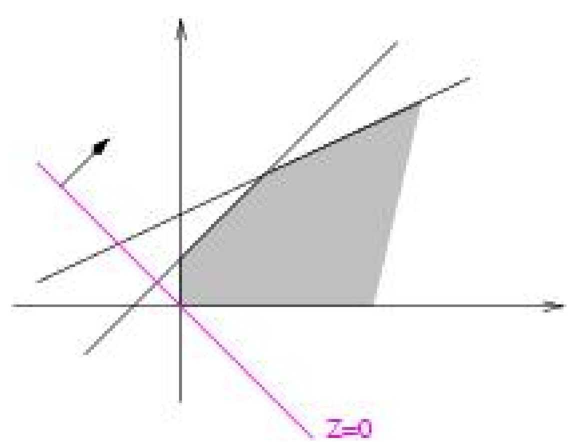

FIG. 3. Représentation graphique pour l'Exemple 3

Le domaine réalisable associé à ce (PL) est la partie non bornée représentée partiellement en gris dans la Figure. Dans le graphique, on observe que toute droite Δ_r d'équation $Z(x_1,x_2) = r$, avec $r \geq 0$, coupe le domaine réalisable. Donc, $\sup Z = +\infty$ et par conséquent, le problème n'admet aucune solution optimale (voir dessin ci-dessous).

2.4. **Remarque.** Tous les exemples traités ci-dessus possèdent deux variables structurelles x_1 et x_2. Pour des (PL) en dimension 3, le raisonnement pour la méthode graphique reste le même. La seule différence est que l'ensemble des points (x_1, x_2, x_3) vérifiant une contrainte

$$a_1x_1 + a_2x_2 + a_3x_3 \leq b$$

se représente par un demi-espace qui est délimité par le plan d'équation

$$a_1x_1 + a_2x_2 + a_3x_3 = b \; .$$

2.5. **Exercice.**

Exercice 1. A venir

3. Méthode des sommets

La méthode des sommetsest simple, elle se base sur la **résolution des systèmes linéaires de Cramer**. Cependant, elle est dans la pratique inexploitable car le nombre de systèmes à résoudre est en général trop important.

Néanmoins, la méthode des sommets est d'une utilité incontestable : la recherche de l'optimum éventuel sur tout le domaine réalisable se ramène au calcul de l'optimum de la fonction d'objectif restreinte à un sous ensemble fini. Ce résultat sera l'un des outils clé de la **méthode du simplexe**.

Pour toute contrainte d'inégalité $\sum a_{ij}x_j \leq b_i$ (*resp.* contrainte de positivité $x_i \geq 0$) d'un (PL) donné, l'hyperplan associé $\sum a_{ij}x_j \leq b_i$ (*resp.* $x_i = 0$) s'appelle *hyperplan frontière*.

3.1. **Résultat préliminaire.** Considérons un (PL) à p variables réelles. Un point A du domaine réalisable est appelé *sommet*, s'il existe p hyperplans frontières dont l'intersection est réduite à A.

Théorème 3.1. *Soit $Z : \mathbb{R}^p \to \mathbb{R}$ une application linéaire et $\mathcal{R}$ un polyèdre de $\mathbb{R}^p$. Si Z admet son optimum global en un point de $\mathcal{R}$, il est aussi atteint en un sommet de $\mathcal{R}$.*

Sans nul doute, l'intérêt de ce théorème est immédiat :

L'optimum (que ce soit maximum ou minimum) d'un (PL), lorsqu'il existe, peut toujours être réalisé sur un sommet.

La méthode des sommets pour la résolution d'un (PL), consiste donc à :
– déterminer tous les sommets du domaine réalisable,
– puis calculer la valeur de Z en chacun de ces sommets.

Le sommet doté de la meilleure valeur de Z serait une solution optimale. Une chose reste à vérifier pour l'application de cette méthode : s'assurer tout d'abord que le (PL) considéré admet bien une solution (*i.e.* l'optimum est fini). Cette condition est souvent difficile à vérifier. Néanmoins, lorsque le domaine réalisable est borné, et par suite compact, l'existence d'une solution optimale est certaine.

Exercice 2. WIMS : Application 1 : Problème de trains

Exercice 3. WIMS : Application 2 : Production optimale

3.2. **Mise en oeuvre.** Appliquons la méthode des sommets aux trois exemples précédents. Pour l'exemple 1, les sommets du domaine réalisable sont les points (voir Figure 1)

$$(0,0)\; ;\; (4,0)\; ;\; (1,3)\; ;\; (3,2) \text{ et } (0,3).$$

La valeur de la fonction d'objectif Z en chacun de ces sommets est

$$Z(0,0) = 0\; ;\; Z(4,0) = 16\; ;\; Z(1,3) = 19$$
$$Z(3,2) = 22\; ;\; Z(0,3) = 15.$$

Il est clair que la plus grande valeur de Z en ces sommets est atteinte au point $(3,2)$ et vaut $Z(3,2) = 22$. Pour l'exemple 2.2, d'après le graphique, les seuls sommets sont $(2,0)$ et $(3,2)$. De plus, on a

$$Z(2,0) = -12 < Z(3,2) = 12,$$

ce qui entraîne que $(3,2)$ est une solution optimale. Concernant l'exemple 2.3, les sommets du domaine réalisable sont les points $(0,0)$, $(0,1)$ et $(2,3)$. La valeur de Z en chacun de ces points vaut

$$Z(0,0)=0, Z(0,1)=1, Z(2,3)=5.$$

Cependant, on ne peut affirmer que $(2,3)$ est une solution optimale, puisqu'on a déjà vu que la fonction d'objectif n'est pas bornée sur son domaine réalisable.

Exercice 4. WIMS : Méthode graphique générale

3.3. **Champ d'application.** La méthode des sommets peut s'appliquer même pour des (PL) ayant un nombre de variables structurelles strictement supérieur à 3. Afin de déterminer un sommet, on choisit p hyperplans (p étant le nombre de variables) parmi les n contraintes du domaine réalisable (y compris les contraintes de positivité : $x_i \geq 0$). On obtient ainsi un système linéaire d'ordre p que l'on résoud. Si ce système admet une et une seule solution A, on vérifie si A satisfait les $(n-p)$ contraintes restantes. Dans ce cas, A est un sommet, sinon le choix considéré de ces p équations linéaires ne permet pas d'avoir un sommet et pour ce faire, on effectue un autre choix de p hyperplans frontières parmi les n contraintes. Dans le but d'obtenir tous les sommets, il va falloir balayer tous les choix possibles dont le nombre est égal à $\binom{n}{p}=\dfrac{n!}{p!(n-p)!}$. Chaque choix qui aboutit à un système de Cramer dont la solution est réalisable permet d'obtenir un sommet. Enfin, la solution du problème considéré est obtenue en prenant l'optimum de la fonction d'objectif sur tous les sommets.

Remarque 1. Dans la pratique, la méthode des sommets est en fait très coûteuse en temps de calcul. D'une part, la détermination d'un sommet exige la résolution d'un système linéaire de Cramer d'ordre p dont la solution doit satisfaire les contraintes restantes. D'autre part, le nombre $\binom{n}{p}$ est en général très important. A titre indicatif, on a $\binom{25}{10}=3.268.760$! Par ailleurs, cette méthode n'est applicable que lorsqu'on est certain que le (PL) admet une solution optimale, ce qui est a priori difficile à vérifier.

Exercice 5. Résolution par la méthode des sommets : A venir

4. Méthode du simplexe

Dans ce chapitre, nous allons étudier une méthode qui évite de parcourir tous les sommets. Plus précisément, le passage d'un sommet à un autre se fait en améliorant la valeur de la fonction à optimiser. De plus, elle fournit un test qui permet d'affirmer le cas échéant que le problème n'admette pas de solution. Nous verrons que les résultats concernant les problèmes de type maximisation, comparés à ceux relatifs aux problèmes de minimisation, sont très similaires.

4.1. **Forme canonique.** On convient de dire qu'*un vecteur u est inférieur à un vecteur v*, et on écrit $u \leq v$, si pour tout i, on a $u_i \leq v_i$. Ici, le terme u_i désigne la i-ième composante du vecteur u. Nous mettons en garde sur le fait que

la négation de $u \leq v$ n'est pas $u > v$.

En effet, l'inégalité $u > v$ traduit la propriété $u_i > v_i$ pour toute composante i. Par contre, la négation de $u \leq v$, que l'on convient de noter $u \not\leq v$, exprime que l'inégalité $u_i > v_i$ ait lieu pour au moins une composante i.

Une *forme canonique* (*resp. forme standard*) est un (PL) où toutes les contraintes sont des inégalités (*resp.* égalités) et les variables sont astreintes à être positives.

On a déjà expliqué qu'on peut supposer et sans perte de généralité, que les contraintes d'inégalité sont toutes de type " $\leq$ ". Par conséquent, on peut affirmer que tout (PL) sous forme canonique s'écrit sous la forme suivante :

$$\begin{cases} \max(\text{ou } \min)[Z(x)=c^*x] \\ Ax \leq b \\ x \geq 0_{\mathbb{R}^p} \end{cases}$$

où $c \in \mathbb{R}^p$, A, matrice (m,p) et $b \in \mathbb{R}^m$ sont fixés.

Le vecteur x a p composantes réelles x_i et * désigne l'opérateur transposé. Les inégalités

$$x_i \geq 0 \, ; i=1,\ldots,p$$

(*i.e.* $x \geq 0_{\mathbb{R}^p}$) s'appellent *contraintes de positivité*. Désormais, on décide de noter un (PL) sous forme canonique par (FC). Remarquer que les trois exemples précédents sont tous écrits sous forme canonique.

4.2. **Forme standard.** La mise sous *forme standard* d'un (PL) quelconque consiste à transformer les contraintes d'inégalités (mise à part les contraintes de positivité) en égalité tout en imposant aux variables d'être positives. Pour ce faire, on procède comme suit :

– A chaque contrainte de type " $\leq$ " (*resp.* " $\geq$ "), ajouter (*resp.* retrancher) une variable tout en lui imposant d'être positive. Celle-ci s'appelle *variable d'écart*.
– Remplacer chaque variable x_i sans restriction de signe par la différence de deux variables positives : $x_i = x_i^+ - x_i^-$ avec $x_i^+ \geq 0$ et $x_i^- \geq 0$.
– Si une variable x_i est négative, effectuer le changement de variable $x_i' = -x_i$.

Exemple 2. Mettre sous forme standard le (PL) suivant :

$$\begin{cases} \max[Z(x_1,x_2,x_3) & = 2x_1 - x_2 + 5x_3] \\ x_1 - 2x_2 & = 3 \\ x_1 - x_3 & \geq -1 \\ x_2 + x_3 & \leq 4 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

On peut dire à juste titre que ce (PL) n'est pas une forme canonique car la première contrainte est une égalité et la troisième variable x_3 n'est pas astreinte à être positive. Pour la mise sous forme standard, la première contrainte, qui est déjà une égalité, reste inchangée. Pour la deuxième, on lui retranche une variable positive x_4, elle devient

$$x_1 - x_3 - x_4 = -1.$$

Quant à la troisième contrainte, on lui ajoute une variable positive x_5 pour devenir

$$x_2 + x_3 + x_5 = 4.$$

Concernant la variable x_3 qui est sans restriction de signe, elle est remplacée par $x_3^+ - x_3^-$, avec $x_3^+ \geq 0$ et $x_3^- \geq 0$. En conclusion, la forme standard associée à ce (PL) s'écrit

$$\begin{cases} \max[Z(x) = 2x_1 - x_2 + 5x_3^+ - 5x_3^-] & \\ x_1 - 2x_2 = & 3 \\ x_1 - x_3^+ + x_3^- - x_4 = & -1 \\ x_2 + x_3^+ - x_3^- + x_5 = & 4 \\ x = (x_1, x_2, x_3^+, x_3^-, x_4, x_5) \geq 0_{\mathbb{R}^6}. & \end{cases}$$

Les variables du (PL) initial s'appellent *variables de décision*, celles de la forme standard associée s'appellent *variables structurelles*.

4.3. **Relation entre la frome canonique et standard.** L'écriture d'un (PL) sous forme standard est introduite tout simplement parce que l'algorithme du simplexe ne s'applique qu'aux formes standards. En d'autres termes,

pour résoudre un (PL) par le simplexe, il faut tout d'abord l'écrire sous forme standard.

Etant donné qu'on résoud le (PL) proprement dit et non la forme standard associée, la question naturelle qu'on se pose est de savoir **comment retrouver une solution optimale du (PL) de départ à partir d'une solution optimale de sa forme standard**. Le théorème suivant établit le lien entre une solution optimale d'un (PL) et celle de sa forme standard, et on se limite au cas où le (PL) considéré est une (FC).

Théorème 4.1. *v est une solution optimale de (FC) si et seulement si $\begin{pmatrix} v \\ b - Av \end{pmatrix}$ est une solution optimale de la forme standard associée. De plus, on a*

$$Z(v) = \tilde{Z}\begin{pmatrix} v \\ b - Av \end{pmatrix}$$

Ce théorème se généralise bien évidemment pour tout (PL) quelque soit son type d'écriture. A titre d'illustration, pour l'Exemple 2 qui n'est pas sous forme canonique, si $(v_1, v_2, v_3, v_4, v_5, v_6)$ est une solution optimale de la forme standard associée, compte tenu des variables d'écarts et du changement de variable considérés, on peut dire que $(v_1, v_2, v_3 - v_4)$ est une solution optimale du (PL) initial.

Exercice 6. Forme standard : à venir

4.4. **Notion de base réalisable.** Désormais, on note tout (PL) écrit sous forme standard par (FS). Par définition, l'écriture matricielle d'un problème (FS) possède la forme suivante :

$$(\mathrm{FS})\begin{cases}\max(\text{ou } \min)[Z(y)=c^*y]\\ My=b\\ y\geq 0_{\mathbb{R}^n}\end{cases}$$

où

$$c\in\mathbb{R}^n,\ M\in M_{m,n}\text{ et } b\in\mathbb{R}^m \text{ sont fixés.}$$

La matrice M contient toujours plus de colonnes que de lignes (*i.e.* $n>m$) vu que le nombre de variables structurelles n (variables de décision + variables d'écarts + variables dûes aux variables sans restriction de signe) dépasse le nombre de contraintes m. Sans perte de généralité, on suppose que la matrice M est de rang m. Sinon,

(1) ou bien il y a une (ou plusieurs) équation qui est combinaison linéaire des autres équations, qu'il faut alors enlever,

(2) ou bien le système est incompatible et dans ce cas le système linéaire $My=b$ n'admet aucune solution.

4.4.1. *Système de base.* Notons par M_i la i^{e} colonne de M. Le vecteur colonne My s'écrit alors

$$My=\sum_{i=1}^{n} y_iM_i$$

où y_i est i^{e} composante de y. On appelle *système de variables de base* tout ensemble de m variables distinctes y_i, $i\in J$, prises parmi $y_1,\ldots,y_n$, telles que le déterminant des vecteurs M_i, $i\in J$ est différent de zéro. On rappelle que l'entier m désigne le nombre de contraintes d'égalités qui apparaissent dans (FS). La matrice carrée formée par les m colonnes M_i, $i\in J$, s'appelle *base* et les variables y_i, $i\notin J$, s'appellent *variables hors base*. Autrement dit,

une base B du problème (FS) n'est autre qu'une sous matrice carrée de M d'ordre m et inversible.

Pour mieux visualiser une telle base B, on effectue des permutations de colonnes de sorte que la base B apparaisse sur les m premières colonnes de M. Ainsi, on met M sous la forme suivante :

$$M=[B\ \ N],$$

où $B=(M_i,i\in J)\in GL_m(\mathbb{R})$ et $N=(M_i,i\notin J)\in M_{m,n-m}$.

Les variables y_i, $i=1,\ldots,n$, peuvent aussi être classées selon B et N :

$$y=\left(\begin{smallmatrix}y_B\\ y_N\end{smallmatrix}\right),\ \text{où } y_B=(y_i,i\in J) \text{ et } y_N=(y_i,i\notin J).$$

Désormais, l'ensemble J des indices de base sera noté par J_B et son complémentaire par J_N. En respectant cette partition, le système $My=b$ vérifie :

$$\begin{aligned}My=b &\Leftrightarrow [B\,N]\left(y_B\right)\Leftrightarrow By_B+Ny_N=b\\ &\Leftrightarrow y_B=B^{-1}b-B^{-1}Ny_N.\end{aligned}$$

Par suite, l'ensemble des solutions dans $\mathbb{R}^n$ du système linéaire $My=b$ est égal à

$$\{y\in\mathbb{R}^n\ /\ My=b\}=\begin{pmatrix}B^{-1}b-B^{-1}Ny_N\\ y_N\end{pmatrix}\ /\ y_N\in\mathbb{R}^{n-m}\}. \tag{1}$$

Parmi les solutions de ce système, nous distinguons les solutions dites *de base* pour lesquelles $y_N=0_N$ (on écrit 0_N à la place de $0_{R^{n-m}}$ pour dire que les variables hors base sont nulles). Une telle solution existe et elle vaut

$$y=\begin{pmatrix}B^{-1}b\\ 0_N\end{pmatrix}$$

Cette solution de base est réalisable (*i.e.* elle vérifie les contraintes $My=b$ et $y\geq 0_{\mathbb{R}^n}$) si et seulement si

$$B^{-1}b\geq 0_B.$$

Dans ce cas, on dit que la *base B est réalisable*.

Noter que toute solution de base réalisable admet $(n-m)$ composantes nulles relatives aux variables hors base et que les autres composantes vérifient un système de Cramer d'ordre m.

4.4.2. *Optimalité et base réalisable.* Par analogie au Théorème 3.1, nous allons établir que lorsque (FS) admet une solution, on peut toujours chercher une solution optimale parmi les solutions de base réalisable.

Théorème 4.2. *Si un problème (FS) admet un optimum global, il existe toujours une solution optimale qui est une solution de base réalisable.*

Remarque 2. Tout sommet du domaine réalisable

$$\mathcal{Y} = \{y/My = b,\ y \geq 0_{\mathbb{R}^n}\}$$

est une solution de base réalisable, la réciproque est également vraie. Les sommets de $\mathcal{Y}$ correspondent donc aux solutions de base réalisable. De ce fait, l'algorithme du simplexe ne s'intéresse qu'aux solutions de type

$$y = \begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix}$$

avec $B^{-1}b \geq 0_B$. Ces solutions de base réalisable sont bien évidemment en nombre fini.

4.4.3. *Base dégénérée.* Une solution de base réalisable $y = \begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix}$ est dite *dégénérée*, si en plus des $(n-m)$ composantes nulles relatives à 0_N, il existe une composante du vecteur $B^{-1}b$ qui est nulle. Elle est *non dégénérée* dans le cas où toutes les composantes de $B^{-1}b$ sont non nulles et par suite strictement positives. De même, on dit qu'une base réalisable est dégénérée si la solution de base associée est dégénérée. Il est intéressant de noter qu'un point dégénéré est solution de plusieurs bases réalisables. En effet, tout vecteur colonne $u = B^{-1}b$ relatif à une base dégénérée B admet au moins une coordonnée u_{i_0}, $i_0 \in J_B$, qui est nulle. En considérant que $B = (M_i,\ i \in J_B)$, on obtient

$$u = B^{-1}b \Leftrightarrow Bu = b \Leftrightarrow \sum_{i \in J_B} u_i M_i = b$$
$$\Leftrightarrow \sum_{i \in J_B \setminus \{i_0\}} u_i M_i = b$$

car $u_{i_0} = 0$. Dans la matrice B, nous allons remplacer le vecteur colonne M_{i_0} par un autre vecteur M_{j_0} pour $j_0 \in J_N$ de sorte que la matrice B' ainsi obtenue soit inversible. Il est clair que u est solution du système linéaire $B'x = b$, étant donné que

$$\sum_{i \in J_B \setminus \{i_0\}} u_i M_i + 0 M_{j_0} = b.$$

Cette solution est unique en raison de la régularité de la matrice B'. Donc, les deux bases réalisables B et B' donnent toutes les deux la même solution de base $\begin{pmatrix} u \\ 0_N \end{pmatrix}$.

4.5. **Algorithme du simplexe. L'algorithme du simplexe** est un procédé itératif qui progresse dans un sens évolutif : il passe d'une solution de base réalisable non optimale à une autre solution ayant une meilleure valeur d'objectif. De cette façon, on évite de parcourir toutes les solutions de base réalisable dont le nombre est en général prohibitif. Pour vérifier la non optimalité d'une solution, un simple test sera effectué. De plus, grâce à l'algorithme du simplexe, on sera capable de détecter, le cas échéant, que l'optimum est infini.

4.5.1. *Condition d'optimalité.* La fonction d'objectif $Z(y) = c^*y$ est entièrement définie par la donnée du vecteur $c = (c_i, i = 1, \ldots, n)$. A chaque base $B = (M_i, i \in J_B)$, on note $c_B = (c_i, i \in J_B)$ et $c_N = (c_i, i \in J_N)$. D'une façon équivalente, le vecteur c_B (*resp.* c_N) s'obtient à partir de c en prenant les coordonnée c_i relatives aux variables de base (*resp.* hors base) associée à la matrice B (*resp.* N). Le classement du vecteur c selon J_B et J_N donne $c = \binom{c_B}{c_N}$. Naturellement, on dit qu'une *base B est optimale* si la solution associée $y = \binom{B^{-1}b}{0_N}$ est optimale.

Théorème 4.3. *On se donne une base réalisable B du problème (FS). Si la forme standard (FS) est de type maximisation (resp. minimisation), on considère le vecteur ligne*

$$w_N^* = c_N^* - c_B^* B^{-1} N$$

(resp. $w_N^* = c_B^* B^{-1} N - c_N^*$*). Alors, on a :*
- *(i)* $w_N^* \leq 0_N^* \Longrightarrow$ *B est optimale.*
- *(ii) Si de plus la base B est non dégénérée (i.e.* $B^{-1}b > 0_B$*), alors*

$$B \text{ est optimale } \Longrightarrow w_N^* \leq 0_N^*.$$

La condition d'optimalité $w_N^* \leq 0_N^*$ peut être interprétée comme un test d'arrêt des itérations pour l'algorithme du simplexe. Il reste à décrire la technique qui permet de passer d'une solution de base réalisable non optimale à une autre solution de base réalisable ayant une meilleure valeur économique. Cette partie fera l'objet du paragraphe suivant.

Interprétation 1. Donner une interprétation économique du vecteur w_N^*. A partir de la solution réalisable $y^{(0)} = \binom{B^{-1}b}{0_N}$, on construit le point y' obtenu en augmentant d'une unité la s-ième variable hors base. D'après l'égalité (1), ce point vaut $y' = \binom{B^{-1}b - B^{-1}Ne_s}{e_s}$, où e_s désigne le s-ième vecteur de la base canonique de $\mathbb{R}^{n-m}$. Si le point ainsi obtenu est réalisable, on obtient :

$$Z(y') = Z(y^{(0)}) \pm w_N^* e_s = Z(y^{(0)}) \pm w_s, \text{ car } w_N^* e_s = w_s.$$

Pour un problème de type maximisation, cette dernière égalité s'écrit $Z(y') = Z(y^{(0)}) + w_s$. Donc, l'élément w_s représente la quantité s'ajoutant à la valeur économique Z lorsqu'on incrémente d'une unité la s^{e} coordonnée hors base de $y^{(0)}$, pour autant que le nouveau point obtenu soit réalisable. Semblablement, lorsqu'il s'agit d'un problème de minimisation, w_s correspond à la valeur retranchée de la valeur économique. C'est la raison pour laquelle w_N^* s'appelle vecteur des *coûts réduits* (ou *coûts fictifs* ou encore *coûts marginaux*).

Exercice 7. Sommet-Base-Optimalité : A venir

4.5.2. *Amélioration de la fonction d'objectif.* On se donne une base réalisable $B = (B_1, \ldots, B_m)$. Le vecteur B_i désigne la i^{e} colonne de B qui est à fortiori une colonne de M. On note aussi N_j la colonne d'indice j de la matrice N. D'ailleurs, N_j n'est autre Ne_j où e_j est le j-ième vecteur de la base canonique de $\mathbb{R}^{n-m}$.

Théorème 4.4. *On suppose que la base réalisable B vérifie $w_N^* \nleq 0_N^*$. Il existe donc une composante w_s du vecteur w_N^* qui est strictement positive. Considérons le s^{e} vecteur colonne $\gamma = B^{-1}N_s$ de la matrice $B^{-1}N$. Alors de deux choses l'une*
- *(i) ou bien $\gamma \leq 0_B$, et dans ce cas l'optimum est infini,*
- *(ii) ou bien $\gamma \nleq 0_B$. Dans ce cas, on calcule le réel*

$$\lambda = \min\{\frac{(B^{-1}b)_i}{\gamma_i} \;/\; \gamma_i > 0\},$$

et on détermine un indice $r \in \{1, \ldots, m\}$ pour lequel ce minimum λ est atteint (i.e. $\lambda = \frac{(B^{-1}b)_r}{\gamma_r}$). Alors, la matrice B', obtenue à partir de $B = (B_1, \ldots, B_m)$ en remplaçant le vecteur colonne B_r par N_s, est une base réalisable. La solution de base $y^{(1)} = \binom{B'^{-1}b}{0_{N'}}$ associée à B' vérifie

$$(B'^{-1}b)_i = \begin{cases} (B^{-1}b)_i - \lambda\gamma_i & \text{si } i \neq r \\ \lambda & \text{si } i = r. \end{cases}$$

De plus, on a $Z(y^{(1)}) = Z(y^{(0)}) \pm \lambda w_s$, selon que (FS) soit de tupe maximisation ou minilisation. Ici, le point $y^{(0)}$ représente la solution de base réalisable associée à B : $y^{(0)} = \binom{B^{-1}b}{0_N}$.

Il est intéressant de remarquer que dans le cas où la base réalisable B est non dégénérée, le paramètre λ est strictement positif. Sans aucun doute, le point construit $y^{(1)}$ possède, en cas de non dégénérescence, une valeur économique strictement meilleure que celle de $y^{(0)}$.

4.5.3. *Algorithme.*

/* *On suppose que l'on dispose d'une base réalisable B* */
$y_B = B^{-1}b$
$Z = c_B^* y_B$
$w_N^* = \begin{cases} c_N^* - c_B^* B^{-1}N & \text{si "max"} \\ c_B^* B^{-1}N - c_N^* & \text{si "min"} \end{cases}$
while $w_N^* \not\leq 0_N^*$ **do**
  choisir s telle que $w_s = w_N^* e_s > 0$
  $\gamma = B^{-1}N_s$
  if $\gamma \leq 0_B$ **then**
    optimum infini /* $\sup = +\infty$ *ou* $\inf = -\infty$ */
    exit
  else
    calculer $\lambda = \min\{\frac{(B^{-1}b)_i}{\gamma_i} \;/\; \gamma_i > 0\}$
    déterminer r pour lequel λ est atteint
  end if
  $Z \leftarrow \begin{cases} Z + \lambda w_s & \text{si "max"} \\ Z - \lambda w_s & \text{si "min"} \end{cases}$
  $B \leftarrow B'$ /* *B_r est remplacée par N_s* */
  $N \leftarrow N'$ /* *N_s est remplacée par B_r* */
  $(y_B)_i \leftarrow \begin{cases} (y_B)_i - \lambda\gamma_i & \text{pour } i \neq r \\ \lambda & \text{pour } i = r \end{cases}$
  Calculer w_N^*
end while
Optimum atteint en $y = \binom{y_B}{0_N}$ et il vaut Z /* *Fin de l'algorithme* */

4.5.4. *Règle de pivotage.* Dans la pratique, le choix de w_s est souvent multiple, il se peut que plusieurs composantes w_N^* soient strictement positives. Etant donnée que la nouvelle solution $y^{(1)}$ vérifie :

$$Z(y^{(1)}) = Z(y^{(0)}) \pm \lambda w_s$$

il serait commode de choisir la plus grande composante w_s de w_N^* strictement positive.

Il arrive aussi que le minimum λ soit atteint pour deux ou plusieurs coordonnées différentes. Dans ce cas de figure, on fixe comme critère de choix pour l'indice r, la première coordonnée trouvée j qui réalise le minimum λ. Pour des raisons de stabilité numérique et par analogie avec la méthode de Gauss avec pivot partiel, un autre critère de choix est souvent considéré. Il consiste à choisir r de façon à ce que γ_r soit le plus élevé parmi les éléments γ_i strictement positifs. Le critère qui fixe le choix de w_s et de l'indice r s'appelle **règle de pivotage**. La règle qui prend systématique la plus grande valeur positive de w_N^* s'appelle **règle du meilleur coût marginal**, c'est cette règle que l'on adopte dans ce cours.

5. ALGORITHME DU SIMPLEXE STANDARD

Dans ce chapitre, les données mise en oeuvre dans l'algorithme du simplexe seront représentées dans un tableau. A chaque itération correspond un tableau appelé **tableau simplexe**. Des formules itératives liant les éléments d'un tableau au tableau subséquent seront également établies.

5.1. **Tableau simplexe.** Les éléments qui interviennent dans l'algorithme du simplexe peuvent être récapitulés dans ce qui suit.

- Les variables de base et les variables hors base. La base réalisable B et la matrice N seront immédiatement identifiées à partir des variables de base et hors base.
- La solution $y = \begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix}$ associée à la base réalisable B. Puisque $y_N = 0_N$, on ne retient que le vecteur $y_B = B^{-1}b$.
- Le vecteur $w_N^* = \pm[c_N^* - c_B^* B^{-1}N]$ pour tester l'optimalité.

– La matrice $B^{-1}N$ afin de calculer w_N^* et le coefficient λ.
– La valeur de la fonction Z au point $y = \begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix}$ qui est égale à

$$Z(y) = c_B^* y_B.$$

A chaque base réalisable B correspond une itération de l'algorithme du simplexe. Pour rendre pratique l'application manuelle de l'algorithme du simplexe, les éléments décrits ci-dessus sont regroupés dans un tableau de la manière suivante :

$$\begin{array}{|ccc|c|c|} y_{i_{m+1}} & \cdots & y_{i_n} & & \\ & & & & y_{i_1} \\ & B^{-1}N & & B^{-1}b & \vdots \\ & & & & y_{i_m} \\ & w_N^* & & Z & \end{array}$$

Les coordonnées $y_{i_1},\dots,y_{i_m}$ correspondent aux variables de base, et les coordonnées $y_{i_{m+1}},\dots,y_{i_n}$ correspondent aux variables hors base. Le terme Z désigne la valeur de la fonction d'objectif Z au point réalisable $\begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix}$. D'après l'indexation des variables de base, on a :

$$Z = \sum_{j=1}^{m} c_{i_j} y_{i_j}, \text{ avec } y_{i_j} = (B^{-1}b)_j.$$

5.2. **Application.** Comme application de l'algorithme du simplexe, nous allons résoudre les trois exemples du premier chapitre en utilisant les tableaux simplexe.

5.2.1. *Exemple 1.* Reprenons l'Exemple 1 :

$$\begin{cases} \max[Z(x_1,x_2) & = 4x_1 + 5x_2] \\ 2x_1 + x_2 & \leq 8 \\ x_1 + 2x_2 & \leq 7 \\ x_2 & \leq 3 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

On écrit ce problème sous forme standard.: En ajoutant les variables d'écart, on obtient :

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

$$c^* = \begin{pmatrix} 4 & 5 & 0 & 0 & 0 \end{pmatrix} \text{ et } b = \begin{pmatrix} 8 \\ 7 \\ 3 \end{pmatrix}.$$

On cherche un système de variables de base réalisable: Il est clair que $\{y_3, y_4, y_5\}$ forme un système de variables de base réalisable. La base associée est $B = I_3$, elle vérifie bien

$$B^{-1}b = I_3 b = b \geq 0_{\mathbb{R}^3}.$$

Pour cette base, on obtient :

$$B^{-1}N = N = \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 0 & 1 \end{pmatrix},$$

$$w_N^* = c_N^* = \begin{pmatrix} 4 & 5 \end{pmatrix}$$

car $c_B^* = 0_B^*$ et

$$Z \leftarrow Z(0,0,8,7,3) = 0.$$

Rappelons que les coordonnées de chaque solution de base réalisable $\begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix}$ sont classées selon B et N. Pour la base $B = I_3$, si on écrit les coordonnées de la solution réalisable $\begin{pmatrix} B^{-1}b \\ 0_N \end{pmatrix} = \begin{pmatrix} b \\ 0_N \end{pmatrix}$ dans l'ordre y_i, $i = 1, \ldots, 5$, on obtient exactement le point $(0,0,8,7,3)$.

Itération 1:

y_1	y_2		
2	1	8	y_3
1	2	7	y_4
0	**1**	3	y_5
4	**5**	0	

Le vecteur ligne $w_N^* = (4\ 5)$ n'est pas négatif, il contient deux composantes strictement positives. Il a été convenu de choisir la plus grande qui est 5. Pour distinguer notre choix, on décide de l'écrire en gras. Faisant référence aux notations utilisées dans le Théorème 4.4, on a :

$$w_s = 5 \text{ et } \gamma = B^{-1}N_s = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Donc, au cours de l'itération suivante, la variable hors base qui se situe au niveau de w_s va devenir une variable de base. D'après le tableau ci-dessus, il s'agit de la variable y_2. Pour déterminer la variable qui prend place, il va falloir calculer le paramètre λ et en déduire l'indice r. D'après le même tableau, on a :

$$\lambda = \min\{\frac{8}{1}, \frac{7}{2}, \frac{3}{1}\} = 3$$

et donc $r = 3$. L'indice r correspond alors à la variable y_5. On décide de mettre en gras l'élément γ_r qui a permis de réaliser le minimum

$$\lambda = \frac{(B^{-1}b)_r}{\gamma_r}.$$

En résumé, lorsqu'on écrit en gras une composante w_s de w_N^*, cela signifie que la variable, figurant sur la même colonne que w_s du tableau simplexe, entre dans la base. Inversement, l'identification de γ_r permet de décider que la variable, située sur la même ligne que γ_r sort de la base. Naturellement, ces deux variables s'appellent respectivement **variable entrante** et **variable sortante**.

Itération 2: A ce stade de calcul, la base réalisable B' est associée aux variables $\{y_3, y_4, y_2\}$. Pour la prochaine itération, on aura besoin de B'' à la place de B'. Afin de ne pas encombrer le texte par les notations, il est naturel de garder les mêmes symboles B, N et λ pour toutes les itérations du simplexe. En tenant compte de l'ordre des indices $J_B = \{3,4,2\}$ des variables de base et de $J_N = \{1,5\}$, on obtient

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Par un simple calcul, on vérifie les égalités suivantes :

$$B^{-1}N = \begin{pmatrix} 2 & -1 \\ 1 & -2 \\ 0 & 1 \end{pmatrix}, B^{-1}b = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} \text{ et } w_N^* = \begin{pmatrix} 4 & -5 \end{pmatrix}.$$

Le tableau simplexe relatif à cette itération est

y_1	y_5		
2	-1	5	y_3
1	-2	1	y_4
0	1	3	y_2
4	-5	15	

Noter que le vecteur $B^{-1}b$ peut être calculé aussi via la formule

$$\begin{aligned}(B^{-1}b)_i &\leftarrow \begin{cases} (B^{-1}b)_i - \lambda\gamma_i & \text{pour } i \neq r \\ \lambda & \text{pour } i = r \end{cases} \\ &= \begin{pmatrix} 8-3\times 1 \\ 7-3\times 2 \\ \lambda \end{pmatrix} \\ &= \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}\end{aligned}$$

De même, la valeur $Z = 15$ qui correspond à $Z(0,3,5,1,0)$ peut être obtenue à travers le résultat $Z \leftarrow Z + \lambda w_s = 0 + 3 \times 5 = 15$. Ces deux constatations restent valables pour les itérations qui suivent et pour n'importe quel exemple.

Pour ce tableau, le vecteur w_N^* admet une seule composante strictement positive qui est $w_s = 4$. Donc, la variable y_1 entre dans la base au cours de l'itération 3. En ce qui concerne le nouveau paramètre λ, on a :

$$\gamma = B^{-1}N_s = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \text{ et par suite } \lambda = \min\{\frac{5}{2}, \frac{1}{1}\} = 1.$$

Ainsi, la variable y_4 devient hors base. Il est intéressant de noter que les matrices B et N peuvent être déduites du tableau simplexe à partir des variables de base et des variables hors base. Dans toute la suite et lors de l'exécution d'une itération simplexe, seul le tableau simplexe sera donné. Bien entendu, on précisera en gras les éléments w_s et γ_r et on calculera le coefficient λ, à chaque fois que cela s'avère nécessaire.

Itération 3:
$$\begin{array}{|rr|r|l} y_4 & y_5 & & \\ \hline -2 & \mathbf{3} & 3 & y_3 \\ 1 & -2 & 1 & y_1 \\ 0 & 1 & 3 & y_2 \\ \hline -4 & \mathbf{3} & 19 & \\ \hline \end{array}$$

Conclusion: On a

$$w_N^* = (-2\ -1) \leq (0\ 0)$$

donc le Théorème 4.3 assure que $(3,2,0,0,1)$ est une solution optimale pour (FS). D'après le Théorème 4.1, le point $(3,2)$ est optimal pour notre (PL) de départ. Il réalise une valeur maximale de Z égale à 22.

5.2.2. *Exemple 2.* Pour l'Exemple 2, le (PL) considéré est

$$\begin{cases} \max[Z(x_1,x_2) & = -6x_1 + 15x_2] \\ 2x_1 - x_2 & \geq 4 \\ -2x_1 + 5x_2 & \leq 4 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

On écrit ce problème sous forme standard.: En retranchant une variable d'écart à la première contrainte, et en ajoutant une seconde à la deuxième, on obtient :

$$M = \begin{pmatrix} 2 & -1 & -1 & 0 \\ -2 & 5 & 0 & 1 \end{pmatrix}, \ c^* = \begin{pmatrix} -6 & 15 & 0 & 0 \end{pmatrix} \text{ et } b = \begin{pmatrix} 4 \\ 4 \end{pmatrix}.$$

On cherche un système de variables de base réalisable: Prenons par exemple $\{y_1, y_4\}$ et montrons qu'il forme un système de variables de base réalisable. La matrice B associée à ce système vérifie :

$$B = \begin{pmatrix} 2 & 0 \\ -2 & 1 \end{pmatrix} \Rightarrow B^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & 1 \end{pmatrix} \Rightarrow B^{-1}b = \begin{pmatrix} 2 \\ 8 \end{pmatrix}.$$

Les composantes de $B^{-1}b$ étant positives, la matrice B est une base réalisable. On calcule

$$B^{-1}N = \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ 4 & -1 \end{pmatrix},$$

et

$$w_N^* = \begin{pmatrix} 15 & 0 \end{pmatrix} - \begin{pmatrix} -6 & 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ 4 & -1 \end{pmatrix} = \begin{pmatrix} 12 & -3 \end{pmatrix}.$$

La solution de base réalisable relative à B est $(2,0,0,8)$, elle donne une valeur d'objectif égale à $Z(2,0,0,8) = -12$.

Itération 1:

y_2	y_3		
$-\frac{1}{2}$	$-\frac{1}{2}$	2	y_1
4	-1	8	y_4
12	-3	-12	

$$\lambda = \min\{\frac{8}{4}\} = 2.$$

Itération 2:

y_4	y_3		
$\frac{1}{8}$	$-\frac{5}{8}$	3	y_1
$\frac{1}{4}$	$-\frac{1}{4}$	2	y_2
-3	0	12	

Le vecteur ligne $w_N^* = (-3\ 0)$ étant négatif, la solution associée $(3,2,0,0)$ est optimale pour (FS). Compte tenu du Théorème 4.1, on peut affirmer que $(3,2)$ est une solution optimale de ce (PL) et que la valeur maximale de Z sur le domaine réalisable est 12.

Remarque: On avait montré à l'aide du graphique que ce problème admettait une infinité de solutions optimales. Ce résultat peut être retrouvé via le dernier tableau simplexe. En effet, on remarque sur ce tableau une composante w_s nulle. Le vecteur $\gamma = B^{-1}N_s$ est égal à $\gamma = -\begin{pmatrix} 5/8 \\ 1/4 \end{pmatrix}$, qui est négatif. Donc, tout point de la forme $z_\lambda = \begin{pmatrix} B^{-1}b - \lambda\gamma \\ \lambda e_s \end{pmatrix}$ est réalisable pour (FS) et ceci pour tout réel λ positif. Du fait que $w_s = 0$, la valeur de $Z(z_\lambda)$, qui est égale à $12 + \lambda w_s$, reste inchangée. Par suite, le point z_λ est optimal pour (FS) quelque soit le réel positif λ.

Conclusion: En ce qui concerne le (PL) associé, nous allons utiliser comme d'habitude le Théorème 4.1. Si on classe les composantes de z_λ dans l'ordre y_i, $i = 1,\dots,4$, on obtient $z_\lambda = (3+\lambda\frac{5}{8}, 2+\lambda\frac{1}{4}, \lambda, 0)$. Par conséquent,

$$x_\lambda = (3+\lambda\frac{5}{8}, 2+\lambda\frac{1}{4}) = (3,2) + \lambda(\frac{5}{8}, \frac{1}{4})$$

est un point optimal pour (PL). Notons que $(\frac{5}{8}, \frac{1}{4})$ est un vecteur directeur de la droite (AB) (voir Figure 2). D'ailleurs, la demi-droite $[AB)$ coïncide avec l'ensemble des points $(3,2) + \lambda(\frac{5}{8}, \frac{1}{4})$ lorsque λ décrit $\mathbb{R}_+$.

5.2.3. *Exemple 3.* Pour l'Exemple 3 :

$$\begin{cases} \max[Z(x_1,x_2) & = x_1 + x_2] \\ -x_1 + x_2 & \leq 1 \\ -x_1 + 2x_2 & \leq 4 \\ x_1 \geq 0,\ x_2 \geq 0 & \end{cases}$$

On écrit ce problème sous forme standard: On a :

$$M = \begin{pmatrix} -1 & 1 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix},\ c^* = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \text{ et } b = \begin{pmatrix} 1 \\ 4 \end{pmatrix}.$$

On cherche un système de variables de base réalisable: Il est facile de vérifier que $B = I_2$ est une base réalisable.

Itération 1:

y_1	y_2		
-1	1	1	y_3
-1	2	4	y_4
1	1	0	

Conclusion: Pour $w_s = 1$ écrit en gras, on a $\gamma = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Donc, d'après le Théorème 4.4, le problème vérifie :

$$\sup_{y \in \mathcal{Y}} Z(y) = +\infty.$$

Plus précisément, pour tout réel $\lambda \geq 0$, le point réalisable $y_\lambda = \begin{pmatrix} B^{-1}b - \lambda\gamma \\ \lambda e_s \end{pmatrix}$ possède une valeur d'objectif égale à :

$$Z(y_\lambda) = 0 + \lambda w_s = \lambda$$

dont la limite en $+\infty$ vaut $+\infty$. Le point y_λ est égal à

$$(\lambda, 0, 1+\lambda, 4+\lambda)$$

et par conséquent $x_\lambda = (\lambda, 0)$ est un point réalisable du (PL) initial et ceci pour tout réel $\lambda \geq 0$. Noter que x_λ décrit le demi axe $[Ox_1)$ lorsque λ parcourt l'ensemble des réels positifs. D'après la Figure 3, tous ces points sont effectivement réalisables.

5.3. **Sommets et solutions de base.** Tous les (PL) résolus jusqu'ici sont des formes canoniques puisque le domaine réalisable de chacun peut s'écrire sous la forme

$$\mathcal{R} = \{x \in \mathbb{R}^p \;/\; Ax \leq b \text{ et } x \geq 0_{\mathbb{R}^p}\}$$

Comme conséquence immédiate du Théorème 4.1, on peut affirmer que si $\begin{pmatrix} v \\ b - Av \end{pmatrix}$ est une solution d'une base réalisable de $\mathcal{Y}$, alors v est un sommet de $\mathcal{R}$. Essayons donc de voir le lien entre les différentes solutions de base réalisable parcourues par l'algorithme du simplexe et les sommets correspondants. Pour l'Exemple 1, à partir des quatre itérations effectuées, on voit que les sommets de $\mathcal{R}$ mis en jeu ont été parcourus dans l'ordre suivant :

$$\begin{cases} (0,0) \\ Z(0,0) = 0 \end{cases} \rightarrow \begin{cases} (0,3) \\ Z(0,3) = 15 \end{cases} \rightarrow \begin{cases} (1,3) \\ Z(1,3) = 19 \end{cases} \rightarrow \begin{cases} (3,2) \\ Z = 22. \end{cases}$$

En se référant à la Figure 1, on remarque que deux sommets relatifs à deux itérations consécutives sont toujours *adjacents* : ils appartiennent à un même segment de droite frontière de $\mathcal{R}$. Cela s'explique algébriquement par le fait que les bases B et B' ne différent que d'une seule colonne.

5.4. **Formules itératives.** A chaque itération de l'algorithme du simplexe, on est contraint à calculer la matrice $B^{-1}N$ et le vecteur ligne $w_N^* = \pm[c_N^* - c_B^* B^{-1}N]$. Ces calculs exigent préalablement la détermination de la matrice inverse B^{-1}. Dans cette section, nous allons voir comment surmonter cette difficulté. Tout simplement, on évitera le calcul de B^{-1}. Les éléments d'un tableau vont être déterminés itérativement à partir des éléments du tableau précédent.

5.4.1. *Matrice du pivot.* Posons $\beta = B^{-1}N$ la matrice réelle rectangulaire (m, p) avec $p = n - m$. Aussi, on pose $B = (B_1, \ldots, B_m)$ et $N = (N_1, \ldots, N_p)$ où B_i (*resp.* N_i) désigne la i^e colonne de B (*resp.* N). Supposons qu'à l'itération relative à la base B, on ait permuté le vecteur colonne B_r avec N_s. Au niveau du tableau, cela revient à écrire en gras l'élément (r, s) de la matrice β, que l'on note communément β_{rs}. Par conséquent, au cours de la prochaine itération, on aura

$$B' = (B_1, \ldots, B_{r-1}, N_s, B_{r+1}, \ldots, B_m)$$

comme base réalisable et

$$N' = (N_1, \ldots, N_{s-1}, B_r, N_{s+1}, \ldots, N_p).$$

Théorème 5.1. *Les relations qui lient* $\beta' = B'^{-1}N'$ *avec* $\beta = B^{-1}N$ *sont données par :*

$$\beta'_{is} = \begin{cases} \frac{1}{\beta_{rs}} & si\ i = r \\ -\frac{\beta_{is}}{\beta_{rs}} & si\ i \neq r \end{cases}$$

et

$$\begin{cases} \beta'_{rj} = \frac{\beta_{rj}}{\beta_{rs}} & si\ j \neq s \\ \beta'_{ij} = \beta_{ij} - \frac{\beta_{is}\beta_{rj}}{\beta_{rs}} & si\ j \neq s\ et\ i \neq r. \end{cases}$$

Afin de mieux visualiser les formules établies dans le Théorème 5.1, nous allons illustrer tout d'abord les éléments du tableau simplexe dans le schéma suivant. Le réel β_{rs} correspond à l'élément écrit en gras au niveau de la matrice $\beta = B^{-1}N$.

$$\begin{array}{|ccccc|c|}
\hline
\diamond & \cdots & \triangleleft & & & \\
\vdots & & \vdots & & & \\
\circ & \cdots & \beta_{rs} & \cdots & \circ & (B^{-1}b)_r \\
 & & \vdots & & \vdots & \\
 & & \triangleleft & \cdots & \diamond & \\
\hline
 & & w_s & & & \\
\hline
\end{array}$$

Ensuite, le tableau simplexe qui suit sera rempli moyennant des transformations qui peuvent être résumées en ces termes :

– au niveau de β_{rs}, on aura $\frac{1}{\beta_{rs}}$.
– Tout élément $\triangleleft$ de la colonne relative à β_{rs} devient $-\frac{\triangleleft}{\beta_{rs}}$.
– Tout élément $\circ$ de la ligne relative à β_{rs} devient $\frac{\circ}{\beta_{rs}}$.
– Tout élément $\diamond$ en dehors de la ligne et de la colonne relative à β_{rs} sera transformé en $\diamond - \frac{\circ \cdot \triangleleft}{\beta_{rs}}$. Dans cette dernière expression, on tient à préciser que $\circ$ (*resp.* $\triangleleft$) est l'élément de la matrice $\beta = B^{-1}N$ qui appartient à la même colonne (*resp.* ligne) que $\diamond$ et la même ligne (*resp.* colonne) que β_{rs}.

Les relations données dans le Théorème 5.1 sont identiques aux formules de Gauss-Jordan pour l'inversion d'une matrice. Pour cette raison, l'élément β_{rs} s'appelle *pivot*, faisant allusion au pivot de Gauss. A chaque itération du simplexe, le pivot est donc l'intersection de la variable sortante avec la variable entrante. Noter également que toute ligne, en dehors de la ligne du pivot, qui possède un zéro sur la colonne du pivot reste inchangée ; idem pour les colonnes possédant un zéro sur la ligne du pivot.

5.4.2. *Vecteur des coûts réduits.* On garde ici le même pivot β_{rs}. Cela veut dire que le vecteur colonne B_r a été échangé par N_s et vice versa. Les formules établies précédemment pour β' vont conduire aux relations liant les composantes de $w^*_{N'}$ avec celles de w^*_N. Pour cela, on pose :

$$w^*_{N'} = (w'_1 \ \dots \ w'_p) \quad \text{et} \quad w^*_N = (w_1 \ \dots \ w_p).$$

Théorème 5.2. *Les composantes de* $w^*_{N'}$ *sont données par*

$$w'_j = \begin{cases} -\frac{w_s}{\beta_{rs}} & si\ j = s \\ \\ w_j - \frac{w_s\beta_{rj}}{\beta_{rs}} & si\ j \neq s. \end{cases}$$

D'ores et déjà, on peut observer que les relation liant $w^*_{N'}$ à w^*_N sont les mêmes que les relations entre une ligne $\beta'_{i\bullet}$ et une ligne en dehors de la ligne du pivot $\beta_{i\bullet}$. Le terme w_s, qu'on a convenu d'écrire en gras, se situe sur la même colonne du pivot. Il est remplacé par $-\frac{w_s}{\beta_{rs}}$. Les opérations qu'il faut appliquer sur w_j, $j \neq k$, en vue d'obtenir w'_j sont identiques à celles effectuées sur les termes de symbole $\diamond$ dans le schéma du paragraphe précédent.

5.5. **Algorithme du simplexe standard.** En ce qui concerne le vecteur $B'^{-1}b$, il a été indiqué d'avance qu'il peut être déterminé par les relations

$$(B'^{-1}b)_i = \begin{cases} (B^{-1}b)_i - \lambda\beta_{is} & \text{pour } i \neq r \\ \lambda & \text{pour } i = r. \end{cases}$$

Si on remplace λ par sa valeur $(B^{-1}b)_r/\beta_{rs}$, on obtient les mêmes formules qui régissent une colonne de β en dehors de la colonne du pivot. Maintenant, l'élément en bas à droite du tableau simplexe, à savoir Z, se calcule itérativement via la relation $Z \leftarrow Z + \lambda w_s$ ou $Z \leftarrow Z - \lambda w_s$ selon que le problème étudié soit de type maximisation ou minimisation. En tenant compte de l'expression de λ, il est facile de constater que la formule $Z \leftarrow Z - \lambda w_s$ s'identifie à la formule relative à un élément de symbole "$\diamond$". Afin de garder les mêmes formules, il est souhaitable de remplacer Z par $-Z$ lorsqu'il s'agit d'un problème de type maximisation. Au niveau de l'implémentation sur machine, il est commode de définir une matrice, que l'on note H, qui contient les éléments β, $B^{-1}b$, w_N^* et Z :

$$\begin{cases} H(1:m, 1:p) = \beta = B^{-1}N, \text{ avec } p = n - m \\ H(1:m, p+1) = B^{-1}b \\ H(m+1, 1:p) = w_N^* = \pm[c_N^* - c_B^* B^{-1}N] \\ H(m+1, p+1) = \pm Z(y^{(0)}) \end{cases}$$

La matrice H est de taille $(m+1, p+1)$. D'une itération à une autre, elle est calculée itérativement de la même façon que l'on détermine les éléments symbolisés par "$\diamond$", "$\triangleleft$" et "$\circ$". Afin de balayer tous les paramètres du tableau simplexe, on ajoute à H deux vecteurs entiers, notés *base* et *hbase* et de taille respective m et p. Pour chaque coordonnée i, $base(i)$ fournit l'indice de la i^{e} variable de base du tableau simplexe en cours. De même, $hbase(j)$ correspond à l'indice de la j^{e} variable hors base. La mise en oeuvre des résultats de ce paragraphe dans l'algorithme du simplexe aboutit à une formulation plus simplifiée au niveau de l'implémentation sur machine. Ici, on suppose que l'on prend systématiquement l'élément $w_s > 0$ le plus élevé. Naturellement, le plus grand parmi les composantes strictement positives est le maximum sur toutes les composantes de w_N^*.

Exercice 8. WIMS : Résolution par les tableaux simplexe
WIMS : Méthode du simplexe

6. Dualité en programmation linéaire

Dans le cadre de la programmation linéaire, on peut associer à chaque (PL) un autre (PL), de type opposé, nommé **programme dual**. Leurs propriétés sont étroitement liées :

Par souci d'optimiser les calculs lors de la résolution d'un (PL), il est plus avantageux, dans certaines circonstances, de **résoudre le problème dual.** Ensuite, la solution du (PL) initial sera déduite facilement par l'intermédiaire de la solution du dual.

Une des conséquences les plus fructueuses de la notion de dualité est **l'algorithme dual-simplexe** qui a été conçu, dans les années 50, par Lemke. Cet algorithme a un intérêt inéluctable en programmation linéaire.

6.1. **Primal - Dual.**

6.1.1. *Cas d'une forme canonique.* On considère un programme linéaire écrit sous forme canonique que l'on suppose de type maximisation :

$$\begin{cases} \max[Z(x) = c^* x] \\ Ax \leq b \\ x \geq 0_{\mathbb{R}^p}. \end{cases}$$

Von Neumann, un Mathématicien Allemand du début du 20-ième siècle, a défini le *dual* de ce problème par le programme linéaire de type minimisation suivant :

$$\begin{cases} \min[\psi(v) = b^* v] \\ A^* v \geq c \\ v \geq 0_{\mathbb{R}^m}. \end{cases}$$

Le problème initial à partir duquel on a défini le dual s'appelle *primal*. Pour une forme canonique, le second membre du dual est formé par les coefficients de la fonction économique du primal ; de même, les coefficients de la fonction d'objectif du dual ne sont autres que les composantes du second membre du dual ; aussi, la matrice

Algorithm 1 Algorithme du simplexe standard

Require: base réalisable
  Initialiser H, *base* et *hbase*
  Calculer $ws = \max[H(m+1, 1:p)]$
  while $ws > 0$ **do**
    déterminer l'indice s pour lequel le maximum ws est atteint
    if $H(1:m,s) \leq 0_{\mathbb{R}^m}$ **then**
      optimum infini /* sup $= +\infty$ ou inf $= -\infty$ */
      exit
    else
      calculer $\lambda = \min\{\frac{H(i,p+1)}{H(i,s)}$, pour $i = 1, \ldots, m \ / \ H(i,s) > 0\}$
      déterminer l'indice r pour lequel λ est atteint
    end if
    $base(r) \leftrightarrow hbase(s)$
    $piv = H(r,s)$
    $H(r,s) \leftarrow \frac{1}{piv}$
    $H(r,j) \leftarrow \frac{H(r,j)}{piv}$, pour $j = 1, \ldots, p+1$ et $j \neq s$
    $H(i,j) \leftarrow H(i,j) - H(r,j) * H(i,s)$, pour $i \neq r$ et $j \neq s$
    $H(i,s) \leftarrow -\frac{H(i,s)}{piv}$, pour $i = 1, \ldots, m+1$ et $i \neq r$
    Calculer $ws = \max[H(m+1, 1:p)]$
  end while
  Optimum vaut $Z = \pm H(m+1, p+1)$
  Une solution optimale, que l'on note $opty$, est donnée par :

$$\begin{cases} opty(base(1:m)) = H(1:m, p+1) \\ opty(hbase(1:p)) = 0. \end{cases}$$

régissant les contraintes du dual est la matrice transposée de la matrice du primal ; et enfin, les inégalités (mise à part les inégalités de positivité) du primal et du dual sont en sens inverse.

6.1.2. *Cas général.* Tout programme linéaire ne s'écrit pas forcément sous forme canonique. Il peut admettre en effet des contraintes d'égalités et/ou des variables **sans restriction de signe** (s.r.s. en abrégé). De ce fait, la forme générale d'un primal de type maximisation s'écrit de la façon suivante :

$$(\mathcal{P}) \begin{cases} \max[Z(x) = \sum_{j=1}^{p} c_j x_j] \\ \sum_{j=1}^{p} a_{ij} x_j \leq b_i, \text{ pour } i = 1, \ldots, l \\ \sum_{j=1}^{p} a_{ij} x_j = b_i, \text{ pour } i = l+1, \ldots, m \\ x_j \geq 0, \text{ pour } j = 1, \ldots, k \\ x_j \text{ s.r.s., pour } j = k+1, \ldots, p. \end{cases}$$

Après un certain nombre d'opérations algébriques, on montre que le dual de $(\mathcal{P})$ est :

$$(\mathcal{D})\begin{cases}\min[\psi(v)=\sum_{i=1}^{m} b_i v_i]\\ \sum_{i=1}^{m} a_{ij}v_i \geq c_j, \text{ pour } j=1,\dots,k\\ \sum_{i=1}^{m} a_{ij}v_i = c_j, \text{ pour } j=k+1,\dots,p\\ v_i\geq 0, \text{ pour } i=1,\dots,l\\ v_i \text{ s.r.s., pour } i=l+1,\dots,m.\end{cases}$$

6.1.3. *Tableau récapitulatif.* On note que lorsque le primal est écrit sous une forme générale, son dual l'est aussi. Le tableau ci-dessous récapitule les règles de correspondance entre un primal de type maximisation et son dual.

Primal	**Dual**
- Maximisation	- Minimisation
- Coefficients d'objectif	- Second membre
- Second membre	- Coefficients d'objectif
- Matrice régissant les contraintes	- Matrice transposée
- Relation liant les contraintes	- Nature des variables
– i-ième inégalité : $\leq$	– $v_i \geq 0$
– i-ième équation : =	– v_i s.r.s.
- Nature des variables	- Relation liant les contraintes
– $x_j \geq 0$	– j-ième inégalité : $\geq$
– x_j s.r.s.	– j-ième équation : =

Remarque 3. On aurait pu choisir comme primal un problème de minimisation, le dual aurait été alors un programme linéaire de type maximisation. Tout simplement parce que le dual du dual est le primal (facile à prouver). Autrement dit, si on permutait dans le tableau précédent les deux mots **Primal** et **Dual**, on obtiendrait les règles de correspondance entre un primal de minimisation et son dual.

6.2. **Optimalité pour le primal-dual.**

Théorème 6.1. *Si x_{op} et v_{op} sont deux solutions réalisables de $(\mathcal{P})$ et $(\mathcal{D})$ respectivement, alors $Z(x_{op})\leq \psi(v_{op})$. Si de plus on a $Z(x_{op})=\psi(v_{op})$, alors x_{op} et v_{op} sont deux solutions optimales de $(\mathcal{P})$ et $(\mathcal{D})$ respectivement.*

Dans la pratique, il est souvent difficile de trouver deux points réalisables x_{op} et v_{op} de $(\mathcal{P})$ et $(\mathcal{D})$ tels que $Z(x_{op})=\psi(v_{op})$. La situation la plus envisageable consiste à déterminer une solution optimale du primal via par exemple le simplexe, puis essayer d'en dégager une solution optimale du dual.

Théorème 6.2. *Soit B une base réalisable de la forme standard relative au primal $(\mathcal{P})$. Si la base B vérifie $w_N^*\leq 0_N^*$, alors $v_{op}^*=c_B^*B^{-1}$ est une solution optimale du dual $(\mathcal{D})$.*

Remarque 4. Afin de déterminer la solution optimale $v_{op}^*=c_B^*B^{-1}$, il est inutile de calculer la matrice inverse B^{-1}. Il suffit, en effet, de résoudre le système de Cramer $v^*B=c_B^*$. Par ailleurs, le Théorème précédent n'exige aucune condition sur la nature du primal (max ou min).

Index

Je remercie vivement Sophie Lemaire et Bernadette Perrin-Riou, enseignants-chercheurs à l'université de Paris-Sud, pour leur aide précieuse à réaliser ce document interactif.