# Optimisation linéaire

L'optimisation linéaire est un domaine de la recherche opérationnelle. Elle consiste à modéliser des problèmes de recherche opérationnelle à l'aide d'inégalités linéaires et à résoudre de tels problèmes à l'aide d'une approche algorithmique spécifique.

## 1 Définitions et exemple

Un *problème (d'optimisation) linéaire (PL)* est un problème d'optimisation (minimisation ou maximisation) d'une fonction linéaire sous des contraintes linéaires, les variables étant continues. Voici un exemple de problème d'optimisation linéaire :

**PL 1**

$$\min 2x_1 - x_2 \tag{1}$$
$$x_1 + x_2 = 10 \tag{2}$$
$$5x_1 + x_2 \geq -2 \tag{3}$$
$$x_1 \leq 0 \tag{4}$$

Le problème **PL 1** possède deux variables continues nommées $x_1$ et $x_2$. La ligne (1) est la fonction linéaire (à minimiser), appelée *fonction objective* ou *fonction coût*. Les lignes (2)-(4) correspondent aux inégalités devant être respectées par les variables $x_1$ et $x_2$ et sont appelées *contraintes* du problème.

L'optimisation linéaire permet de modéliser beaucoup de problèmes de recherche opérationnelle. À titre d'exemple, considérons le problème suivant [1].

**Énoncé :**

Une entreprise de fabrication de yaourts souhaite produire des yaourts à la fraise. Il lui est possible de fabriquer deux types de yaourts à la fraise : allégé et normal. La fabrication d'un litre de chaque type de yaourt nécessite différentes matières premières : produire un litre de yaourt allégé nécessite 2 kilos de fraises et 1 litre de lait, alors que la production d'un litre de yaourt normal nécessite 1 kilo de fraises, 2 litres de lait et 1 kilo de sucre.

Un litre de yaourt allégé peut être vendu 40 euros alors qu'un litre de yaourt normal peut être vendu 50 euros.

Étant donné que l'entreprise possède 800 kilos de fraises, 700 litres de lait et 300 kilos de sucre, combien de litres de yaourt allégé et normal doit-elle fabriquer pour maximiser son revenu ?

**Modélisation :**

Dans ce problème, il faut déterminer le nombre de litres de yaourt de chaque type à produire. On a donc deux variables. On notera $x_A$ la quantité (en litres) de yaourt allégé et $x_N$ la quantité (en litres) de yaourt normal.

L'utilisation de chaque ressource (fraise, lait et sucre) donne lieu à une contrainte. Par exemple, la quantité de fraises nécessaire pour produire les deux types de yaourts ne doit pas dépasser la quantité en stock. Si l'on exprime la contrainte avec les variables $x_A$ et $x_N$, on obtient :

$$2x_A + x_N \leq 800$$

Comme une quantité ne peut être négative, nous avons aussi les contraintes $x_A \geq 0$ et $x_N \geq 0$. Finalement, l'objectif consiste à maximiser le revenu correspondant à la vente des quantités de yaourts produits. On obtient alors le problème d'optimisation linéaire **PL 2** :

1. Exemple issu du cours de P. Pesneau

**PL 2**

$$\max 40x_A + 50x_N \qquad \text{(revenu)} \qquad (5)$$
$$2x_A + x_N \leq 800 \qquad \text{(fraise)} \qquad (6)$$
$$x_A + 2x_N \leq 700 \qquad \text{(lait)} \qquad (7)$$
$$x_N \leq 300 \qquad \text{(sucre)} \qquad (8)$$
$$x_A \geq 0 \qquad \text{(quantité positive)} \qquad (9)$$
$$x_N \geq 0 \qquad \text{(quantité positive)} \qquad (10)$$

## 2 Résolution graphique

Si l'on a un PL avec deux variables (deux dimensions), on peut alors représenter les contraintes dans le plan. L'axe des abscisses correspond à la première variable et l'axe des ordonnées à la deuxième variable. Chaque inégalité peut être représentée par un demi-espace défini par la droite correspondant à l'équation associée à l'inégalité. L'intersection de ces demi-espaces définit l'ensemble des solutions réalisables du problème et est appelé *domaine réalisable* du problème linéaire.

Dans le problème **PL 2**, on a 5 inégalités, définissant 5 demi-plans (Le demi-plan associé à une inégalité est celui qui ne contient pas le trait bleu). Le domaine réalisable est dessiné en vert sur la figure 1.

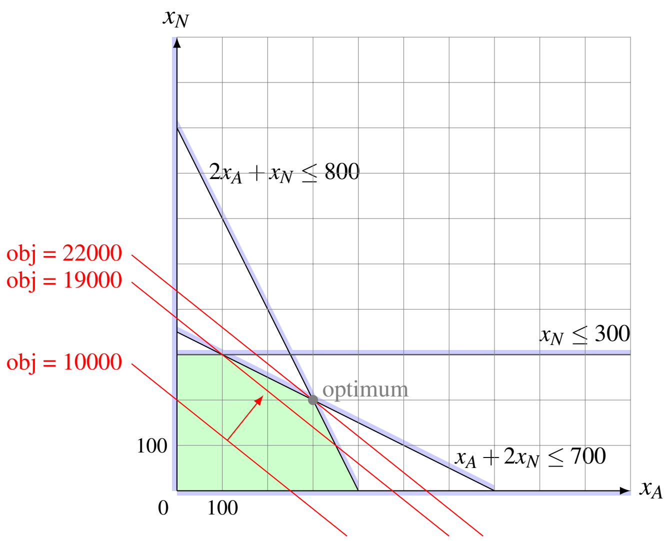


FIGURE 1 – Domaine réalisable et optimum

**Définition 1** Un *sommet* du domaine réalisable est une solution vérifiant deux inégalités du problème linéaire à l'égalité.

Dans cet exemple, le domaine réalisable possède 5 sommets :
— $(0,0)$ vérifiant les inégalités (9) et (10) à l'égalité,
— $(400,0)$ vérifiant les inégalités (6) et (10) à l'égalité,
— $(300,200)$ vérifiant les inégalités (6) et (7) à l'égalité,
— $(100,300)$ vérifiant les inégalités (7) et (8) à l'égalité,
— $(0,300)$ vérifiant les inégalités (8) et (9) à l'égalité.

**Définition 2** Deux sommets du domaine réalisable sont dits *voisins* s'ils vérifient à l'égalité une même inégalité du PL.

Les sommets $(0,0)$ et $(400,0)$ sont voisins, de même que $(400,0)$ et $(300,200)$.

**R** Les concepts de sommets et sommets voisins existent également lorsque l'on a plus de deux variables (pour des dimensions supérieures à deux).

À toute solution est associé le coût de cette solution, correspondant à la valeur de la fonction objective pour cette solution. Plusieurs solutions ont le même coût. Ceci est représenté dans la figure par des droites correspondant aux solutions de même coût. Le vecteur indique la direction dans laquelle la valeur de la fonction objective augmente. On obtient la solution optimale (c'est-à-dire la meilleure solution) en suivant la direction donnée par ce vecteur jusqu'à la frontière du domaine réalisable.

Dans l'exemple, la solution optimale est la solution $(300,200)$. Pour optimiser son profit, l'entreprise doit donc produire 300 litres de yaourt allégé et 200 litres de yaourt normal.

**R** S'il existe une solution optimale, alors (au moins) un des sommets du domaine réalisable est une solution optimale.

## 3 Formes générale, standard et canonique

Dans sa forme générale, un problème linéaire peut contenir des contraintes données par des inégalités ($\leq$ ou $\geq$) ou des égalités. De plus, les variables peuvent être positives ou nulles, négatives ou nulles ou sans signe. L'objectif peut être une maximisation ou une minimisation. Le **PL 1** donné en début de chapitre est un exemple de PL de forme générale.

**Définition 3** Un PL est dit sous forme *canonique* si :
— l'objectif est une maximisation,
— les contraintes sont toutes des inégalités de type $\leq$,
— les variables sont positives ou nulles.

**R** Tout PL (forme générale) peut être écrit sous forme canonique en appliquant quelques modifications :
— Si une variable $x$ est négative ou nulle, on remplace cette variable par la variable $x'$ de signe contraire ($x'$ vaut $-x$).
— Si une variable $x$ est *libre* (sans contrainte de signe), on peut alors l'écrire comme la différence de deux variables $x^+ - x^-$ satisfaisant $x^+ \geq 0$ et $x^- \geq 0$.
— Chaque égalité peut être remplacée par deux inégalités. On peut aussi multiplier une contrainte par -1 pour changer son signe.
— Minimiser $z$ se ramène à maximiser $-z$.

Le problème linéaire **PL 1** sous forme générale peut s'écrire, sous forme canonique, de la manière suivante :

**PL 3**

$$
\begin{aligned}
&\max 2x'_1 + x_2^+ - x_2^- \\
&-x'_1 + x_2^+ - x_2^- \leq 10 \\
&x'_1 - x_2^+ + x_2^- \leq -10 \\
&5x'_1 - x_2^+ + x_2^- \leq 2 \\
&x'_1, x_2^+, x_2^- \geq 0
\end{aligned}
$$

**Définition 4** Un problème linéaire est dit sous forme *standard* si :
— l'objectif est une maximisation,
— les contraintes sont toutes des égalités,
— les variables sont positives ou nulles.

**R** On peut facilement passer d'une forme canonique à une forme standard en ajoutant, dans chaque contrainte, une *variable d'écart*, correspondant à la différence entre les valeurs du membre de droite et du membre de gauche.

Le **PL 3** s'écrit sous forme standard :

**PL 4**

$$
\begin{aligned}
&\max 2x'_1 + x_2^+ - x_2^- \\
&-x'_1 + x_2^+ - x_2^- + e_1 = 10 \\
&x'_1 - x_2^+ + x_2^- + e_2 = -10 \\
&5x'_1 - x_2^+ + x_2^- + e_3 = 2 \\
&x'_1, x_2^+, x_2^-, e_1, e_2, e_3 \geq 0
\end{aligned}
$$

## 4 Algorithme du simplexe

Pour résoudre un PL, on va utiliser un algorithme appelé *algorithme du simplexe* et introduit par George Dantzig en 1947. L'idée de cet algorithme est de partir d'une solution, puis de chercher à chaque itération une solution voisine qui améliore la fonction objective, jusqu'à ce que la solution obtenue soit optimale.

On supposera dans cette section que le PL à résoudre est sous forme canonique et que les membres de droite sont positifs ou nuls. La première étape consiste à écrire ce PL sous forme standard en ajoutant les variables d'écart. On a alors une première solution en fixant toutes les variables originales (*i.e.*, celles qui ne sont pas des variables d'écart) à 0 et chaque variable d'écart à la valeur du membre de droite de la contrainte.

À titre d'exemple, considérons le problème linéaire de l'exemple sur les yaourts. Celui-ci est sous forme canonique et les membres de droite sont tous positifs. La transformation de ce PL en forme standard donne :

**PL 5**

$$
\begin{aligned}
&\max 40x_A + 50x_N + 0e_1 + 0e_2 + 0e_3 \\
&2x_A + x_N + e_1 = 800 \\
&x_A + 2x_N + e_2 = 700 \\
&x_N + e_3 = 300 \\
&x_A, x_N, e_1, e_2, e_3 \geq 0
\end{aligned}
$$

Une solution du problème est $(0,0,800,700,300)$ avec un coût (de la fonction objective) de 0.

**R** Si un membre de droite est strictement négatif, alors cette solution n'est pas réalisable puisque les variables d'écart doivent être positives ou nulles. Nous verrons par la suite comment traiter ce cas.

On peut réécrire le PL en mettant à gauche les variables non nulles ($e_1$, $e_2$ et $e_3$). On obtient ainsi le problème linéaire, appelé *dictionnaire associé à $e_1$, $e_2$ et $e_3$* :

**Dictionnaire 1**

$$
\begin{aligned}
&\max 40x_A + 50x_N \\
&e_1 = 800 - 2x_A - x_N \\
&e_2 = 700 - x_A - 2x_N \\
&e_3 = 300 - x_N \\
&x_A, x_N, e_1, e_2, e_3 \geq 0
\end{aligned}
$$

Les variables à gauche sont appelées variables *de base* alors que celles à droite sont appelées variables *hors base*. Les variables de base ainsi que la fonction objective sont exprimées en fonction des variables hors base. La solution est donnée en fixant les variables hors base à 0. Les variables de base sont les variables ayant une valeur non nulle [2]. Cette solution est la *solution associée au dictionnaire*.

Il est clair que cette solution n'est pas optimale puisque la valeur de la fonction objective associée est 0. Comme le coefficient associé à la variable hors base $x_A$ dans la fonction objective est strictement positif, si l'on trouve une autre solution avec $x_A > 0$, alors cette solution sera meilleure (puisque son coût sera

2. Certaines variables de base peuvent être nulles dans le cas d'un PL dégénéré.

strictement plus grand que 0). On va donc chercher une solution avec $x_A > 0$ mais en laissant les autres variables hors base (ici $x_N$) à 0.

**R** On a choisi $x_A$ de manière arbitraire. On aurait aussi bien pu chercher une solution avec $x_N > 0$ (et $x_A$ à 0).

**Quelle valeur donner à $x_A$ ?**

— Si l'on fixe $x_A = 100$ (et $x_N = 0$), les valeurs des variables d'écart sont alors $e_1 = 800 - 200 = 600$, $e_2 = 700 - 100 = 600$, et $e_3 = 300$. On obtient alors la solution $(100, 0, 600, 600, 300)$ de coût $z = 4000$.

— Si l'on fixe $x_A = 500$, on obtient alors le point $(500, 0, -200, 200, 300)$ de coût $z = 20000$. Cependant, ce point n'est pas une solution puisque la contrainte $e_1 \geq 0$ n'est pas satisfaite !

Il faut donc trouver la valeur maximale que peut prendre $x_A$ tout en satisfaisant les contraintes. La contrainte $e_1 \geq 0$ implique que $800 - 2x_A \geq 0$, c'est-à-dire $x_A \leq 400$. La contrainte $e_2 \geq 0$ implique que $x_A \leq 700$.

Pour que le nouveau point soit une solution, il faut donc que $x_A$ soit au plus égal à 400. Comme on cherche à maximiser l'objectif, on pose $x_A = 400$. On obtient alors la solution $(400, 0, 0, 300, 300)$ de coût $z = 16000$.

Dans la nouvelle solution, la variable $x_A$ étant non nulle, elle devient une variable de base. On dit que $x_A$ *entre en base*. Comme $x_A = 400$, la variable $e_1$ passe à 0 et devient une variable hors base. On dit que $e_1$ *sort de la base*. Il faut maintenant écrire le dictionnaire associé aux variables de base $x_A$, $e_2$ et $e_3$. L'équation contenant $e_1$ permet d'exprimer la variable $x_A$ en fonction des variables hors base. De plus, on remplace dans les autres contraintes la variables $x_A$ par la valeur donnée par cette première équation. On obtient alors :

**Dictionnaire 2**

$$\begin{aligned}
&\max 16000 + 30x_N - 20e_1 \\
&x_A = 400 - \frac{1}{2}x_N - \frac{1}{2}e_1 \\
&e_2 = 300 - \frac{3}{2}x_N + \frac{1}{2}e_1 \\
&e_3 = 300 - x_N \\
&x_A, x_N, e_1, e_2, e_3 \geq 0
\end{aligned}$$

On recommence une nouvelle itération. Trouver une solution avec $x_N > 0$ améliorerait l'objectif. Pour obtenir une solution valide, on doit satisfaire :

$$\begin{aligned}
x_A &= 400 - \frac{1}{2}x_N \geq 0, \text{ ce qui implique } x_N \leq 800 \\
e_2 &= 300 - \frac{3}{2}x_N \geq 0, \text{ ce qui implique } x_N \leq 200 \\
e_3 &= 300 - x_N \geq 0, \quad \text{ce qui implique } x_N \leq 300
\end{aligned}$$

On pose donc $x_N = 200$. La variable $x_N$ entre en base et $e_2$ en sort. On obtient alors la solution $(300, 200, 0, 0, 100)$ de coût $z = 40 * 300 + 50 * 200 = 22000$.

**R** Si aucune borne supérieure n'est trouvée pour la variable entrante, alors le problème linéaire est non borné (la valeur de la fonction objective tend vers l'infini).

Le nouveau dictionnaire associé est :

**Dictionnaire 3**

$$\begin{aligned}
&\max 22000 - 10e_1 - 20e_2 \\
&x_A = 300 - \frac{2}{3}e_1 + \frac{1}{3}e_2 \\
&x_N = 200 + \frac{1}{3}e_1 - \frac{2}{3}e_2 \\
&e_3 = 100 - \frac{1}{3}e_1 + \frac{2}{3}e_2 \\
&x_A, x_N, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Toute solution voisine avec $e_1 > 0$ ou $e_2 > 0$ serait de moins bonne qualité puisque les coefficients de la fonction objective associés à $e_1$ et $e_2$ sont négatifs. La solution $(300, 200, 0, 0, 100)$ est donc optimale. Le dictionnaire associé à une solution optimale est dit *optimal*.

**R** Un dictionnaire est optimal si tous les coefficients des variables hors-base dans la fonction objective sont négatifs ou nuls.

L'algorithme du simplexe peut donc être écrit de la manière suivante.

**Algorithme 1 :** Algorithme du simplexe

**Entrées :** PL canonique avec membres de droite positifs ou nuls
**Sorties :** Solution optimale ou problème non borné
**début**
- Écrire le PL sous forme standard
- $D \leftarrow$ dictionnaire associé aux variables d'écart
- **tant que** *$D$ n'est pas optimal* **faire**
  - $x_j \leftarrow$ variable entrante de $D$
  - $x_i \leftarrow$ variable sortante de $D$ et $x_j$
  - **si** *$x_i$ n'existe pas* **alors**
    - **retourner** Problème non borné
  - Mettre à jour $D$ ($x_j$ entre, $x_i$ sort)
- **retourner** Solution optimale de $D$

### Choix des variables entrante et sortante

À chaque itération, on doit faire entrer en base une variable dont le coefficient dans la fonction objective du dictionnaire est positif. Dans le cas où plusieurs variables peuvent entrer en base, on choisira la première dans l'ordre "lexicographique" (l'ordre dans lequel apparaissent les variables dans le PL). De la même manière, si plusieurs variables peuvent sortir de base, on choisira la première dans l'ordre lexicographique. Ce choix permet de s'assurer que l'algorithme du simplexe s'arrête dans tous les cas [3].

### Interprétation géométrique

Les solutions associées aux différents dictionnaires obtenus avec l'algorithme du simplexe correspondent à des sommets du domaine réalisable. Ainsi, les solutions obtenues par l'algorithme du simplexe dans l'exemple sont $(0, 0, 800, 700, 300)$, $(400, 0, 0, 300, 300)$ et $(300, 200, 0, 0, 100)$. Ces solutions correspondent aux sommets du domaine réalisable $(0, 0)$, $(400, 0)$ et $(300, 200)$.

L'algorithme du simplexe part d'un sommet du domaine réalisable. À chaque étape, il se déplace vers un sommet voisin améliorant l'objectif.

3. En effet, si le PL est dégénéré, il est possible, en prenant d'autres règles pour le choix des variables entrante et sortante, que l'algorithme du simplexe ne s'arrête jamais. Ceci est cependant très rare et l'explication de la dégénérescence des problèmes linéaires dépasse le cadre de ce cours.

## 5 Recherche d'une solution initiale

L'algorithme du simplexe a besoin d'une solution initiale (un sommet) pour commencer. Si tous les membres de droite sont positifs ou nuls, alors une solution peut facilement être trouvée (*c.f.* partie précédente). Dans le cas contraire, la recherche d'une solution initiale n'est pas forcément facile.

Une méthode pour trouver une telle solution initiale consiste à résoudre une version modifiée du PL. On considère une nouvelle variable, disons $x_0$, et on ajoute $-x_0$ dans chaque contrainte où le membre de droite est strictement négatif. On ajoute également la contrainte $x_0 \geq 0$ et on pose comme objectif $\min x_0$.

À titre d'exemple, supposons que l'on ait le PL (sous forme standard) :

**PL 6**

$$\begin{aligned}
&\max -x_1 - 2x_2 \\
&-x_1 - x_2 + e_1 = -1 \\
&x_1 + x_2 + e_2 = 2 \\
&-4x_1 - x_2 + e_3 = -4 \\
&x_1, x_2, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Le PL auxiliaire de ce problème est alors :

**PL 7**

$$\begin{aligned}
&\max -x_0 \\
&-x_1 - x_2 + e_1 - x_0 = -1 \\
&x_1 + x_2 + e_2 = 2 \\
&-4x_1 - x_2 + e_3 - x_0 = -4 \\
&x_1, x_2, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Une solution peut facilement être trouvée. On fixe $x_0$ égal à la valeur absolue du membre de droite le plus petit (ici, 4). On fixe les variables originales à 0. Les valeurs des variables d'écart se calculent alors en fonction des membres de droite et de $x_0$. Dans l'exemple, on obtient la solution $(4,0,0,3,2,0)$. Les variables de base associées à cette solution sont $x_0$, $e_1$ et $e_2$. Comme on a une solution initiale, on peut alors appliquer l'algorithme du simplexe.

On part du dictionnaire donné par les variables en base de la solution initiale trouvée. Dans l'exemple, on part du dictionnaire suivant :

**Dictionnaire 4**

$$\begin{aligned}
&\max -4 + 4x_1 + x_2 - e_3 \\
&e_1 = 3 - 3x_1 + e_3 \\
&e_2 = 2 - x_1 - x_2 \\
&x_0 = 4 - 4x_1 - x_2 + e_3 \\
&x_A, x_N, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Si la solution optimale du problème auxiliaire a pour coût 0 (*i.e.*, $x_0$ vaut 0 dans cette solution), alors elle est solution du PL original et peut être utilisée comme solution initiale. Dans le cas contraire, le PL original n'a pas de solution.

Cette phase de recherche d'une solution initiale à l'aide de l'algorithme du simplexe est appelée *phase I du simplexe*. La résolution d'un PL à l'aide du simplexe avec recherche d'une solution initiale est appelée dans ce cas *simplexe en deux phases*.

## 5 Recherche d'une solution initiale

L'algorithme du simplexe a besoin d'une solution initiale (un sommet) pour commencer. Si tous les membres de droite sont positifs ou nuls, alors une solution peut facilement être trouvée (*c.f.* partie précédente). Dans le cas contraire, la recherche d'une solution initiale n'est pas forcément facile.

Une méthode pour trouver une telle solution initiale consiste à résoudre une version modifiée du PL. On considère une nouvelle variable, disons $x_0$, et on ajoute $-x_0$ dans chaque contrainte où le membre de droite est strictement négatif. On ajoute également la contrainte $x_0 \geq 0$ et on pose comme objectif $\min x_0$.

À titre d'exemple, supposons que l'on ait le PL (sous forme standard) :

**PL 6**

$$\begin{aligned}
&\max -x_1 - 2x_2 \\
&-x_1 - x_2 + e_1 = -1 \\
&x_1 + x_2 + e_2 = 2 \\
&-4x_1 - x_2 + e_3 = -4 \\
&x_1, x_2, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Le PL auxiliaire de ce problème est alors :

**PL 7**

$$\begin{aligned}
&\max -x_0 \\
&-x_1 - x_2 + e_1 - x_0 = -1 \\
&x_1 + x_2 + e_2 = 2 \\
&-4x_1 - x_2 + e_3 - x_0 = -4 \\
&x_1, x_2, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Une solution peut facilement être trouvée. On fixe $x_0$ égal à la valeur absolue du membre de droite le plus petit (ici, 4). On fixe les variables originales à 0. Les valeurs des variables d'écart se calculent alors en fonction des membres de droite et de $x_0$. Dans l'exemple, on obtient la solution $(4,0,0,3,2,0)$. Les variables de base associées à cette solution sont $x_0$, $e_1$ et $e_2$. Comme on a une solution initiale, on peut alors appliquer l'algorithme du simplexe.

On part du dictionnaire donné par les variables en base de la solution initiale trouvée. Dans l'exemple, on part du dictionnaire suivant :

**Dictionnaire 4**

$$\begin{aligned}
&\max -4 + 4x_1 + x_2 - e_3 \\
&e_1 = 3 - 3x_1 + e_3 \\
&e_2 = 2 - x_1 - x_2 \\
&x_0 = 4 - 4x_1 - x_2 + e_3 \\
&x_A, x_N, e_1, e_2, e_3 \geq 0
\end{aligned}$$

Si la solution optimale du problème auxiliaire a pour coût 0 (*i.e.*, $x_0$ vaut 0 dans cette solution), alors elle est solution du PL original et peut être utilisée comme solution initiale. Dans le cas contraire, le PL original n'a pas de solution.

Cette phase de recherche d'une solution initiale à l'aide de l'algorithme du simplexe est appelée *phase I du simplexe*. La résolution d'un PL à l'aide du simplexe avec recherche d'une solution initiale est appelée dans ce cas *simplexe en deux phases*.