

L'optimisation revisitée

Ce cours vous est proposé par Odile Brandière, Université de Paris Sud 11, UFR Jean Monnet et AUNEGe, l'Université Numérique en Économie Gestion.

Table des matières

Introduction.....	2
Fonctions à plusieurs variables – optimisation libre.....	3
Maximum local – minimum local d'une fonction à deux variables	3
Condition du premier ordre.....	4
Condition du second ordre.....	4
Le cas des fonctions à plus de deux variables	6
Optimisation sous contrainte	8
Programmation linéaire	11
Références	13

Introduction

Objectif de la leçon : Nous résolvons les problèmes d'optimisation libre à plusieurs variables et nous abordons les problèmes d'optimisation liée. Nous traitons aussi de la programmation linéaire, outil très utile pour les gestionnaires.

Cette leçon est fondamentale pour des économistes et des gestionnaires qui sont très souvent confrontés à des problèmes d'optimisation.

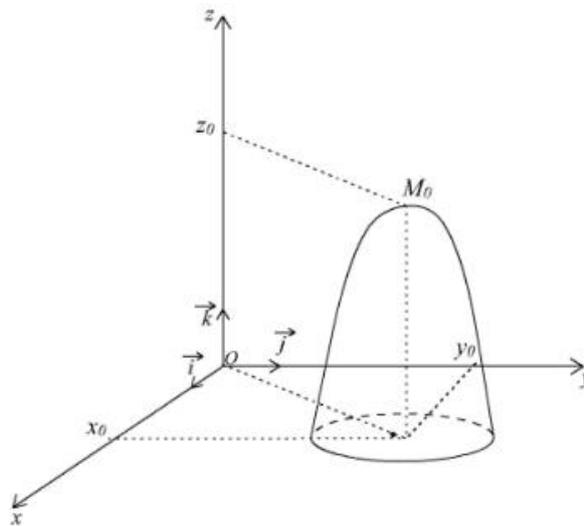
Fonctions à plusieurs variables – optimisation libre

Maximum local – minimum local d'une fonction à deux variables

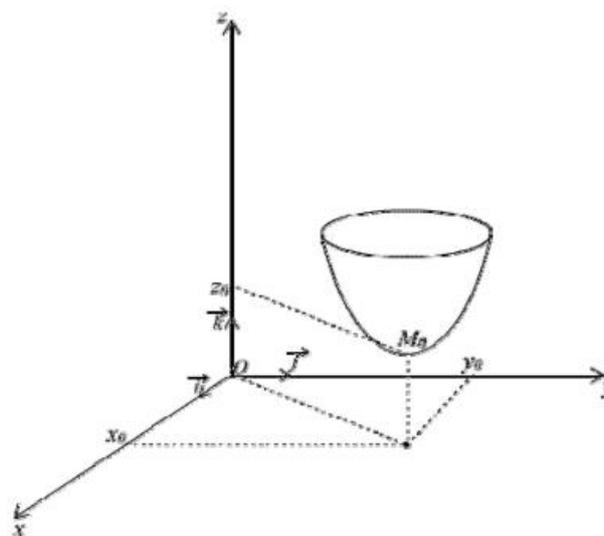
En généralisant les notions de maximum local et minimum local définis dans le cas d'une fonction à une variable au cas d'une fonction f à deux variables telle que $(x,y) \rightarrow f(x,y)$

On dira que f atteint un **maximum local** en (x_0, y_0) lorsque :

$f(x_0, y_0) \geq f(x, y)$ pour tout (x, y) appartenant à un voisinage de (x_0, y_0) (un voisinage de (x_0, y_0) , c'est par exemple un disque de centre (x_0, y_0) et de rayon non nul).



De même, f atteindra un minimum local en (x_0, y_0) si et seulement si pour tout (x, y) appartenant à un voisinage de (x_0, y_0) : $f(x_0, y_0) \leq f(x, y)$.



Condition du premier ordre

Supposons que f ait des dérivées partielles du 1^{er} ordre en (x_0, y_0) .

Si f admet un maximum local en (x_0, y_0) , les fonctions d'une variable, $\phi : x \rightarrow f(x, y_0)$ (y_0 est fixé, la seule variable est x) et $\psi : y \rightarrow f(x_0, y)$ (x_0 est fixé, la variable est y) admettent un maximum local en x_0 pour ϕ et en y_0 pour ψ . Or D'après l'hypothèse, ϕ est dérivable en x_0 ($\phi'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0)$) et ψ en y_0 ($\psi'(y_0) = \frac{\partial f}{\partial y}(x_0, y_0)$).

Donc la dérivée première de ϕ est nulle en x_0 et celle de ψ est nulle en y_0 , ce qui s'écrit :

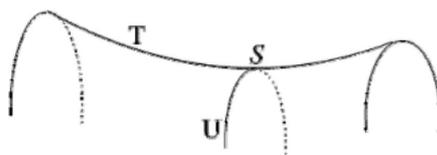
$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0 \quad (\text{Conditions du 1er ordre}).$$

Un point dont les coordonnées vérifient ces conditions s'appelle un point stationnaire.

Attention : Ces conditions ne sont pas suffisantes pour avoir un extremum (il suffit d'observer le dessin ci-dessous).

En effet S correspond à un maximum pour ϕ (correspond à la courbe U , y étant constant et égal à y_0) et à un minimum pour ψ (correspond à la courbe T , x étant constant et égal à x_0) donc $\frac{\partial f}{\partial x}(x_0, y_0) = 0$ et $\frac{\partial f}{\partial y}(x_0, y_0) = 0$.

Or S n'est pas un extremum pour f , on dit que c'est un **point-col** ou un **point-selle**.



Condition du second ordre

Si f a des dérivées partielles du second ordre continues sur un voisinage V de (x_0, y_0) , on a la formule de Taylor :

$$\forall (x, y) \in V: \quad f(x, y) = f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) + \frac{1}{2} \left[(x - x_0)^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + 2(x - x_0)(y - y_0) \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) + (y - y_0)^2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \right] + o((x - x_0)^2, (y - y_0)^2)$$

Et si les conditions du 1^{er} ordre sont vérifiées au point stationnaire (x_0, y_0) , on obtient

$$\Delta f = f(x, y) - f(x_0, y_0) = \frac{1}{2} \left[(x - x_0)^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + 2(x - x_0)(y - y_0) \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) + (y - y_0)^2 \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \right] + o((x - x_0)^2, (y - y_0)^2)$$

f admet effectivement un extremum local en (x_0, y_0) si et seulement si Δf est de signe constant sur un voisinage de (x_0, y_0) .

Il suffit donc que la quantité entre crochets soit de signe constant puisque $o((x - x_0)^2, (y - y_0)^2)$ est négligeable devant le crochet (sauf si le crochet est nul sur un voisinage de (x_0, y_0)).

Cela s'écrit : $\left(\frac{x-x_0}{y-y_0}\right)^2 \frac{\partial^2 f}{\partial x^2}(x_0, y_0) + 2\left(\frac{x-x_0}{y-y_0}\right) \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)$ de signe constant. C'est un trinôme du second degré en $\frac{x-x_0}{y-y_0}$ qui aura un signe constant si son discriminant est strictement négatif, soit $\frac{\partial^2 f}{\partial x^2}(x_0, y_0)^2 - \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) < 0$

Le trinôme sera positif si $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ (ou $\frac{\partial^2 f}{\partial y^2}(x_0, y_0) > 0$), et la différence $f(x, y) - f(x_0, y_0)$ sera positive, on aura donc un minimum local en (x_0, y_0) .

Par contre si $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ (ou $\frac{\partial^2 f}{\partial y^2}(x_0, y_0) < 0$) ce sera un maximum local.

Si $\frac{\partial^2 f}{\partial x^2}(x_0, y_0)^2 - \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) > 0$, le trinôme change de signe et cela correspond à un point-col. On n'a pas d'extremum.

Si $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)^2 - \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) = 0$ on ne peut pas conclure et il faut pousser plus loin le développement limité.

Si f est une fonction des deux variables x et y , ayant des dérivées partielles du second ordre continues, pour que f admette un extremum local en (x_0, y_0) , il faut que

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0 \quad (\text{conditions du premier ordre})$$

Mais ces conditions ne sont pas suffisantes et un point dont les coordonnées les vérifient s'appelle un point stationnaire.

Si de plus $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)^2 - \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) < 0$ alors f admet effectivement un extremum local en (x_0, y_0) (condition du second ordre).

C'est un **maximum** si $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ (ou $f''_{yy}(x_0, y_0) < 0$).

C'est un **minimum** si $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ (ou $f''_{yy}(x_0, y_0) > 0$).

Si $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)^2 - \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) > 0$, f n'admet pas d'extremum en (x_0, y_0) mais un **point-col** (ou **point-selle**).

Si $\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)^2 - \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) = 0$ **on ne peut conclure**.

Le cas des fonctions à plus de deux variables

Soit f une fonction de 3 variables ou plus, différentiable.

Pour déterminer les extrema d'une fonction de 3 variables (ou plus), on commence par déterminer les points stationnaires (x_0, y_0, z_0) en écrivant que toutes les dérivées partielles sont nulles.

La condition du second ordre ne s'applique plus. On doit alors étudier directement la différence $\Delta f = f(x, y, z) - f(x_0, y_0, z_0)$. Il est souvent astucieux de faire le changement de variables $x = x_0 + h, y = y_0 + k$ et $z = z_0 + t$.

On se ramène ainsi à un voisinage de $(0, 0, 0)$.

En effet, (x, y, z) voisin de (x_0, y_0, z_0) équivaut à (h, k, t) voisin de $(0, 0, 0)$.

On pensera aussi à utiliser des développements limités (à l'ordre 2).

Si Δf est positif (respectivement négatif) pour (h, k, t) voisin de $(0, 0, 0)$, (x_0, y_0, z_0) correspond à un minimum (respectivement un maximum) de f .

Exemple : Déterminons les extrema de la fonction suivante : $f : (x, y, z) \rightarrow x \ln y + z \ln x - y$.

N.B. : On pourra remarquer que $-\frac{1}{2}k^2 + hk + zh = -\frac{1}{2}(k-h)^2 + \frac{1}{2}(h+z)^2 - \frac{1}{2}z^2$

Solution : **1^{ère} étape : ensemble de dérivation**

f est définie et admet des dérivées partielles d'ordre 2 continues sur $D = \{(x, y, z), x > 0 \text{ et } y > 0\}$.

2^{ème} étape : recherche des points stationnaires, conditions du premier ordre.

Sur $D : (x_0, y_0, z_0)$ est un point stationnaire si et seulement si : (I)
$$\begin{cases} \ln y_0 + \frac{z_0}{x_0} = 0 \\ \frac{x_0}{y_0} - 1 = 0 \\ \ln x_0 = 0 \end{cases}$$

La dernière équation donne $x_0 = 1$, la deuxième donne alors $y_0 = 1$ et la troisième $z_0 = 0$.

Et (I) $\Leftrightarrow \begin{cases} x_0 = 1 \\ y_0 = 1 \\ z_0 = 0 \end{cases}$. Ainsi $A(1,1,0)$ est le seul point stationnaire.

3^{ème} étape : nature des points stationnaires

Pour savoir si le point stationnaire $A(1,1,0)$ est un extremum, on commence par poser $x = 1 + h$ et $y = 1 + k$ (pas de changement de variable pour z car z est déjà voisin de 0).

Puis on étudie le signe de $\Delta f =$ accroissement $f(x,y,z) - f(1,1,0) = f(1+h, 1+k, z) - f(1,1,0)$ pour (x,y,z) voisin de $(1,1,0)$ ou (h,k,z) voisin de $(0,0,0)$.

Si Δf a un signe constant pour (h,k,z) voisin de $(0,0,0)$, A est un extremum.

C'est un maximum si $\Delta f < 0$, et un minimum si $\Delta f > 0$.

Si le signe de Δ n'est pas constant, A est un point-col.

$$\Delta f = (1+h)\ln(1+k) + zt\ln(1+h) - (1+k) - (-1).$$

Pour étudier le signe de cette expression on en fait un développement limité à l'ordre 2 au voisinage de $(0,0,0)$ (c.f. rappel ci-dessous).

Or pour h et k voisins de zéro :

$$\ln(1+k) = k - \frac{k^2}{2} + k^2\varepsilon(k) \text{ avec } \lim_{k \rightarrow 0} \varepsilon(k) = 0 \text{ et } \ln(1+h) = h - \frac{h^2}{2} + h^2\varepsilon(h) \text{ avec } \lim_{h \rightarrow 0} \varepsilon(h) = 0.$$

$$\text{Donc } \Delta f = (1+h) \left[k - \frac{k^2}{2} + k^2\varepsilon(k) \right] + z \left[h - \frac{h^2}{2} + h^2\varepsilon(h) \right] - 1 - k + 1$$

$$\Delta f = -\frac{1}{2}k^2 + hk + zh + \left[-\frac{1}{2}hk^2 + k^2\varepsilon(h) + hk^2\varepsilon(k) - \frac{1}{2}zh^2 + zh^2\varepsilon(h) \right]$$

Or le crochet est négligeable devant les autres termes quand (h,k,z) est voisin de $(0,0,0)$, donc

$$\Delta f \sim -\frac{1}{2}(k-h)^2 + \frac{1}{2}(\sim h + z)^2 - \frac{1}{2}z^2 \text{ (ici on a utilisé le N.B. du début).}$$

Remarque : L'essentiel ici est de bien comprendre la conduite des calculs.

Dans le cadre de l'optimisation sans contrainte Δf est toujours équivalent aux termes de degré 2 (les termes constants et les termes de degré 1 s'annulent et les autres termes sont de degré supérieur à 2, donc négligeables).

Il faut comprendre que l'on cherche le signe de Δf au voisinage de $(0,0,0)$ et que les termes de degré 2, mis sous la forme d'une somme algébrique de carrés permettent de répondre à la question comme le prouve la suite.

Si $h = -z$ (et cela peut se produire dans tout voisinage de $(0,0,0)$) $\Delta\text{Tr}\{f\} < 0$ et si $z = 0$ et $h = k$ (et cela peut se produire dans tout voisinage de $(0,0,0)$) $\Delta\text{Tr}\{f\} > 0$. Δf n'a donc pas un signe constant au voisinage de $(0,0,0)$ et $A(1,1,0)$ est un point-col.

Rappel : pour x voisin de zéro (c.f. leçon 6 du cours de Mathématiques 1 (niveau L1)) et pour une fonction ε qui tend vers zéro en zéro :

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots + x^n + x^n + \varepsilon(x)$$

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \dots + \frac{m(m-1)\dots(m-n+1)}{n!}x^n + x^n\varepsilon(x) \quad (m \in \mathbb{R})$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \dots + (-1)^{n-1}\frac{x^n}{n} + x^n\varepsilon(x)$$

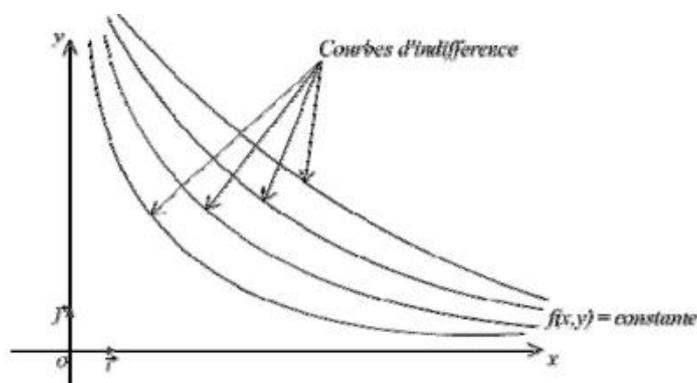
$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + x^n\varepsilon(x)$$

Optimisation sous contrainte

L'optimisation sous contrainte consiste à optimiser une fonction de plusieurs variables, ces variables étant liées entre elle par une relation (contrainte).

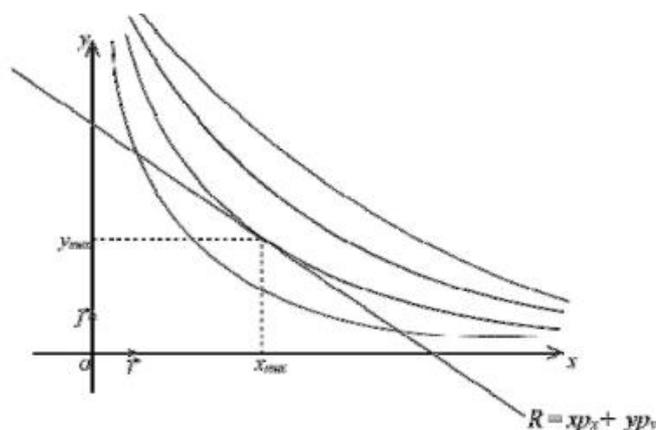
Nous considérerons que toutes les fonctions qui interviennent dans ce paragraphe ont des dérivées partielles secondes continues. Mais pour mieux comprendre cette nouvelle notion traitons un exemple simple.

Exemple économique : On considère un consommateur ayant un budget R et devant acheter 2 biens en quantité x et y , sachant que les prix respectifs de ces biens sont p_x et p_y . On sait d'autre part que ce consommateur est indifférent à certaines combinaisons des biens, d'où l'existence de courbes d'indifférence d'équation $f(x,y) = \text{constante}$



Le problème qui se pose est alors d'accroître au maximum la satisfaction du consommateur (la satisfaction du consommateur est d'autant plus grande qu'on s'éloigne de l'origine), tout en respectant la contrainte budgétaire

Ici on peut résoudre le problème graphiquement. Le point correspondant au maximum doit se trouver sur la droite de budget et sur la courbe d'indifférence le plus loin de l'origine possible. C'est le point de tangence du dessin.



Par le calcul, on est amené à maximiser f sous la contrainte (C) : $xp_x + yp_y = R$.

Or ici le problème se résout facilement puisque $y = \frac{R - xp_x}{p_y}$ et le problème revient à maximiser la fonction de la variable x , $\varphi(x) = f\left(x, \frac{R - xp_x}{p_y}\right)$.

Plus généralement, dans le cas simple où l'une des variables peut s'exprimer en fonction des autres, on pourra se ramener à une optimisation libre (sans contrainte) avec une variable de moins.

Dans le cas où une variable ne peut s'exprimer en fonction des autres, on utilise une fonction auxiliaire L , appelée Lagrangien et qui permet de se ramener à une optimisation sans contrainte.

Pour simplifier, limitons-nous au cas de deux variables, mais la généralisation à n variables est simple et naturelle.

Soit $f : (x,y) \rightarrow f(x,y)$ la fonction à optimiser sous la contrainte $g(x,y) = 0$.

Le Lagrangien L est la fonction des trois variables x, y et λ définie par :

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$$

λ est appelé multiplicateur de Lagrange.

f et g ont des dérivées partielles secondes continues donc L aussi.

Montrons que tout extremum de f sous la contrainte $g(x,y) = 0$ correspond à un extremum de L .

Ecrivons les conditions du premier ordre pour L :

$$\frac{\partial L}{\partial x} = 0, \frac{\partial L}{\partial y} = 0 \text{ et } \frac{\partial L}{\partial \lambda} = 0.$$

Mais remarquons que : $\frac{\partial L}{\partial \lambda} = 0 \Leftrightarrow g(x,y) = 0$.

Donc un extremum de L est atteint lorsque x et y sont soumis à la contrainte $g(x,y) = 0$.

Or si $g(x,y) = 0$ on a $L(x,y,\lambda) = f(x,y)$, et f et L sont minimales ou maximales en même temps. Un extremum lié de f sous la contrainte $g(x,y) = 0$ correspond bien à un extremum libre de L .

Les conditions du premier ordre s'écrivent donc :

$$\begin{cases} \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 \\ g(x,y) = 0 \end{cases}$$

Les points stationnaires du Lagrangien ne sont pas nécessairement des extrema, sauf dans le cadre de la convexité étudié dans la leçon 7. Pour déterminer si un point stationnaire est un extrema il est nécessaire de vérifier les conditions du second ordre. Il existe pour cela un outil très utile, la matrice hessienne bordée. Nous ne définirons pas cette notion dans le cadre de ce cours car elle nécessite des notions mathématiques qui n'y sont pas abordées. Nous nous contenterons donc de la recherche des points stationnaires, tout en sachant que l'étude n'est pas terminée.

Synthèse des résultats à retenir en ce qui concerne l'optimisation sous contrainte :

Pour trouver les extrema de $f : (x,y) \rightarrow f(x,y)$ sous la contrainte $g(x,y) = 0$, on regarde d'abord si on peut exprimer l'une des variables en fonction de l'autre dans l'expression $g(x,y) = 0$.

Si c'est le cas, on est ramené à chercher un extremum libre de contrainte.

Sinon, on forme le Lagrangien $L : (x,y,\lambda) \rightarrow L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$.

Les extrema liés de f sous la contrainte $g(x,y) = 0$, sont les extrema libres de L .

Les conditions du premier ordre pour L sont : $\frac{\partial L}{\partial x} = 0, \frac{\partial L}{\partial y} = 0$ et $g(x,y) = 0$, elles permettent de déterminer les points stationnaires (x_0, y_0) .

Pour savoir si ce sont effectivement des extrema, il faut résoudre les conditions du second ordre (hors du cadre de ce cours s'il n'y a pas convexité).

Programmation linéaire

La programmation linéaire a pour objectif de déterminer graphiquement l'extremum d'une fonction reliant deux variables sous plusieurs contraintes prenant la forme d'inégalités.

Exemple :

- Un fabricant produit des tables et des bureaux. Chaque table nécessite 2,5 heures pour l'assemblage (A), 3 heures pour le polissage (P) et une heure pour la mise en caisse (C). Chaque bureau exige 1 heure pour l'assemblage, 3 heures pour le polissage et 2 heures pour la mise en caisse. L'entreprise ne peut disposer, chaque semaine que de 20 heures pour l'assemblage, 30 heures pour le polissage et de 16 heures pour la mise en caisse. Sa marge de profit est de 3€ par table et de 4€ par bureau.

Déterminer le nombre x de tables et le nombre y de bureaux que l'entreprise doit fabriquer pour obtenir le maximum de profit Π .

1) Exprimons les données sous formes d'équations et d'inéquations, la fonction à optimiser étant

$$\text{ici : } \Pi = 3x + 4y \text{ avec les contraintes suivantes : } \begin{cases} 2,5x + y \leq 20(\text{pour } A) \\ 3x + 3y \leq 30(\text{pour } P) \\ x + 2y \leq 16(\text{pour } C), x \in N, y \in N \end{cases}$$

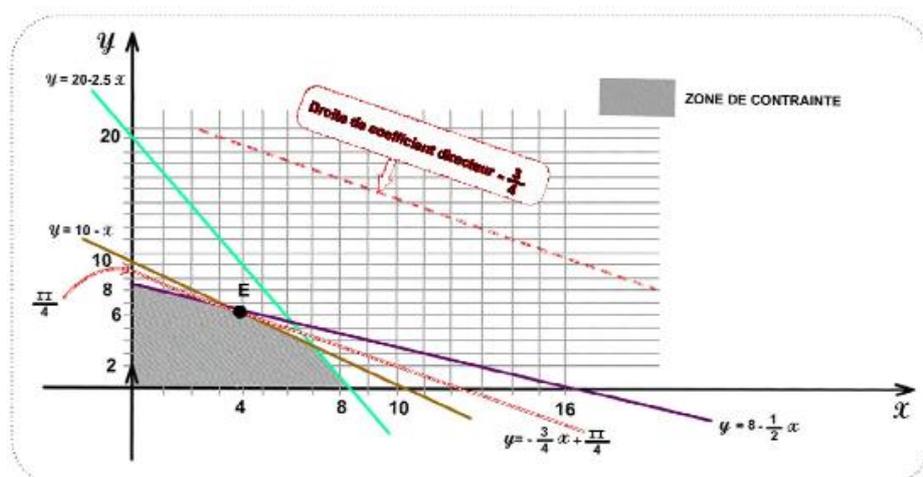
Ces 3 inégalités s'écrivent : (1) $y \leq 20 - 2,5x$

$$(2) y \leq 10 - x$$

$$(3) y \leq 8 - \frac{x}{2}$$

Graphiquement ces trois inégalités sont représentées par 3 demi-plans délimités par les droites d'équations : $y = 20 - 2,5x$, $y = 10 - x$ et $y = 8 - 0,5x$

D'où la représentation graphique suivante :



Le profit Π vérifie : $y = -\frac{3}{4}x + \frac{\Pi}{4}$, graphiquement, cette égalité est l'équation de droites toutes parallèles (coefficient directeur $-\frac{3}{4}$) et d'ordonnée à l'origine $\frac{\Pi}{4}$. Donc plus la droite s'écarte de 0 plus Π est grand. Pour obtenir le profit maximum, il suffit donc de tracer la droite de coefficient directeur $-\frac{3}{4}$, se situant dans la zone de contrainte et la plus éloignée possible de 0.

Le point E limite de la représentation graphique, se trouvant encore dans la zone de contrainte et sur la droite de profit maximum, permet de déterminer les valeurs x et y correspondant à ce profit maximum.

Ici E a pour coordonnées (4,6) et le profit correspondant est $\Pi = 36$ ($\frac{\Pi}{4} = 9$).

Ainsi pour obtenir le profit maximum, compte tenu des contraintes sur (A), (P) et (C), l'entreprise devra produire 4 chaises et 6 bureaux par semaine.

Remarquons que le profit est maximisé à l'intersection de deux droites de contrainte, en un point qualifié de point extrême.

Ceci est un résultat général :

Lorsque la fonction à maximiser peut avoir une valeur optimale dans la région accessible, celle-ci se situera à un point extrême.

Sur la figure précédente, il y a 4 points extrêmes possibles : (0,8), (7,3), (4,6), (8,0).

Références

Comment citer ce cours ?

Mathématiques 2, Odile Brandière, AUNEGe (<http://auneg.fr>), CC – BY NC ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Cette œuvre est mise à disposition dans le respect de la législation française protégeant le droit d'auteur, selon les termes du contrat de licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). En cas de conflit entre la législation française et les termes de ce contrat de licence, la clause non conforme à la législation française est réputée non écrite. Si la clause constitue un élément déterminant de l'engagement des parties ou de l'une d'elles, sa nullité emporte celle du contrat de licence tout entier.