**UNIVERSITE DE LISALA**

**CENTRE INTERUNIVERSITAIRE DE RECHERCHE PLURIDISCIPLINAIRE (CIREP)**
**STATUT : UNIVERSITE PUBLIQUE**
**Web : www.cirep.ac.cd**
**Email : info@cirep.ac.cd**

# NOTES DE COURS DES MATHEMATIQUES APPLIQUEES

**OBJECTIFS DU COURS**

***Objectif général***

L’objectif général de ce cours et d’analyser et de modéliser les relations économiques à l'aide de données empiriques en économétrie et d'optimiser les processus et de prendre des décisions efficaces en utilisant des modèles mathématiques et des techniques d'optimisation en recherche opérationnelle.

***Objectifs spécifiques :***

- ✓ Identifier et quantifier les relations causales entre les variables économiques.
- ✓ Estimer les paramètres des modèles économétriques pour évaluer l'impact des variables explicatives sur la variable dépendante.
- ✓ Tester l'adéquation des modèles économétriques et évaluer leur capacité à prédire les comportements économiques futurs.
- ✓ Utiliser les résultats des analyses économétriques pour formuler des recommandations politiques ou stratégiques.
- ✓ Modéliser les problèmes de décision complexes sous forme de problèmes d'optimisation mathématique.
- ✓ Développer des méthodes et des algorithmes pour résoudre efficacement les problèmes d'optimisation.
- ✓ Analyser et évaluer les solutions optimales pour prendre des décisions éclairées.
- ✓ Appliquer les techniques d'optimisation à divers domaines tels que la logistique, la production, la finance, etc.

# 1. Qu'est-ce que l'économétrie ?

Ce premier chapitre est consacré à la présentation de l'économétrie et à sa liaison avec la théorie économique. Nous abordons tout d'abord la notion de modèle ainsi que les différentes étapes de la modélisation. L'apport de l'économétrie en tant qu'outil de validation est étudié en II. Enfin, la théorie de la corrélation – fondement de l'économétrie – fait l'objet de la section III.

## I. La notion de modèle

### A. Définition

Il est délicat de fournir une définition unique de la notion de modèle[1]. Dans le cadre de l'économétrie, nous pouvons considérer qu'un modèle consiste en une *présentation formalisée d'un phénomène* sous forme d'équations dont les variables sont des grandeurs économiques. L'objectif du modèle est de représenter les traits les plus marquants d'une réalité qu'il cherche à styliser. Le modèle est donc l'outil que le modélisateur utilise lorsqu'il cherche à comprendre et à expliquer des phénomènes. Pour ce faire, il émet des hypothèses et explicite des relations.

1. La notion de modèle est relative au point de vue auquel nous nous plaçons : la physique, l'épistémologie...

*Pourquoi des modèles ?*

*Nombreux sont ceux – sociologues, économistes ou physiciens – qui fondent leurs analyses ou leurs jugements sur des raisonnements construits et élaborés. Ces constructions refèrent implicitement à des modèles ; alors pourquoi ne pas expliciter clairement les hypothèses et les relations au sein d'un modèle ?*

Le modèle est donc une présentation schématique et partielle d'une réalité naturellement plus complexe. Toute la difficulté de la modélisation consiste à ne retenir que la ou les représentations intéressantes pour le problème que le modélisateur cherche à expliciter. Ce choix dépend de la nature du problème, du type de décision ou de l'étude à effectuer. La même réalité peut ainsi être formalisée de diverses manières en fonction des objectifs.

## B. La construction des modèles en économétrie

Dans les sciences sociales, et particulièrement en économie, les phénomènes étudiés concernent le plus souvent des comportements afin de mieux comprendre la nature et le fonctionnement des systèmes économiques. L'objectif du modélisateur est, dans le cadre de l'économétrie et au travers d'une mesure statistique, de permettre aux agents économiques (ménages, entreprises, État...) d'intervenir de manière plus efficace. La construction d'un modèle comporte un certain nombre d'étapes qui sont toutes importantes. En effet, en cas de faiblesse d'un des « maillons », le modèle peut se trouver invalidé pour cause d'hypothèses manquantes, de données non représentatives ou observées avec des erreurs, etc. Examinons les différentes étapes à suivre lors de la construction d'un modèle, ceci à partir de l'exemple du modèle keynésien simplifié.

### 1) Référence à une théorie

Une théorie s'exprime au travers d'hypothèses auxquelles le modèle fait référence. Dans la théorie keynésienne, quatre propositions sont fondamentales :

1. la consommation et le revenu sont liés ;
2. le niveau d'investissement privé et le taux d'intérêt sont également liés ;
3. il existe un investissement autonome public ;
4. enfin, le produit national est égal à la consommation plus l'investissement privé et public.

## 2) Formalisation des relations et choix de la forme des fonctions

À partir des propositions précédentes, nous pouvons construire des relations :

1. la consommation est fonction du revenu : $C = f(Y)$ avec $f' > 0$ ;
2. l'investissement privé dépend du taux d'intérêt : $I = g(r)$ avec $g' < 0$ ;
3. il existe un investissement autonome public : $\overline{I}$ ;
4. enfin, le produit national (ou le revenu national) est égal à la consommation plus l'investissement : $Y \equiv C + I + \overline{I}$.

À ce stade, nous n'avons postulé aucune forme particulière en ce qui concerne les fonctions $f$ et $g$. Ainsi, bien que des considérations d'ordre théorique nous renseignent sur le signe des dérivées, il existe une multitude de fonctions de formes très différentes et ayant des signes de dérivées identiques, par exemple $C = a_0 + a_1 Y$ et $C = a_0 Y^{a_1}$. Cependant ces deux relations ne reflètent pas le même comportement ; une augmentation du revenu provoque un accroissement proportionnel pour la première relation, alors que, dans la seconde, l'effet s'estompe avec l'augmentation du revenu (si $0 < a_1 < 1$). Nous appelons « forme fonctionnelle » ce choix (arbitraire ou fondé) de spécification précise du modèle. Dans notre exemple, le modèle explicité s'écrit :

$$
\begin{aligned}
&C = a_0 + a_1 Y && \text{avec } a_0 > 0 \text{ et } 0 < a_1 < 1 \\
& && a_1 = \text{propension marginale à consommer} \\
& && \text{et } a_0 = \text{consommation incompressible ;} \\
&I = b_0 + b_1 r && \text{avec } b_0 > 0 \text{ et } b_1 < 0 \text{ ;} \\
&Y \equiv C + I + \overline{I} &&
\end{aligned}
$$

Les deux premières équations reflètent des relations de comportements alors que la troisième est une identité (aucun paramètre n'est à estimer).

## 3) Sélection et mesure des variables

Le modèle étant spécifié, il convient de collecter les variables représentatives des phénomènes économiques. Ce choix n'est pas neutre et peut conduire à des résultats différents, les questions qu'il convient de se poser sont par exemple :

– *Faut-il raisonner en euros constants ou en euros courants ?*

– *Les données sont-elles brutes ou CVS*[1] ?

– *Quel taux d'intérêt faut-il retenir (taux au jour le jour, taux directeur de la Banque Centrale Européenne,...) ?* etc.

---

1. Corrigées des Variations Saisonnières.

Nous distinguons plusieurs types de données selon que le modèle est spécifié en :

– *série temporelle :* c'est le cas le plus fréquent en économétrie, il s'agit de variables observées à intervalles de temps réguliers (la consommation annuelle, totale France, exprimée en euros courants sur 20 ans) ;

– *coupe instantanée :* les données sont observées au même instant et concernent les valeurs prises par la variable pour un groupe d'individus[1] spécifiques (consommation observée des agriculteurs pour une année donnée) ;

– *panel :* la variable représente les valeurs prises par un échantillon d'individus à intervalles réguliers (la consommation d'un échantillon de ménages de la région parisienne sur 20 ans) ;

– *cohorte :* très proches des données de panel, les données de cohorte se distinguent de la précédente par la constance de l'échantillon, les individus sondés sont les mêmes d'une période sur l'autre.

## 4) Décalages temporels

Dans le cadre de modèle spécifié en séries temporelles, les relations entre les variables ne sont pas toujours synchrones mais peuvent être décalées dans le temps. Nous pouvons concevoir que la consommation de l'année $t$ est expliquée par le revenu de l'année $t-1$ et non celui de l'année $t$ . Pour lever cette ambiguïté, il est d'usage d'écrire le modèle en le spécifiant à l'aide d'un indice de temps : $C_t = a_0 + a_1 Y_{t-1}$. La variable $Y_{t-1}$ est appelée « variable endogène retardée ».

> *On appelle « variable exogène » une variable dont les valeurs sont prédéterminées, et « variable endogène » une variable dont les valeurs dépendent des variables exogènes.*

## 5) Validation du modèle

La dernière étape est celle de la validation[2] du modèle :

– *Les relations spécifiées sont-elles valides ?*

– *Peut-on estimer avec suffisamment de précision les coefficients ?*

– *Le modèle est-il vérifié sur la totalité de la période ?*

– *Les coefficients sont-ils stables ? Etc.*

À toutes ces questions, les techniques économétriques s'efforcent d'apporter des réponses.

1. Le terme d'individu est employé au sens statistique, c'est-à-dire comme un élément d'une population : une personne, une parcelle de terre...

2. Validation, c'est-à-dire en conformité avec les données disponibles.

# II. Le rôle de l'économétrie

## A. L'économétrie comme validation de la théorie

L'économétrie est un outil à la disposition de l'économiste qui lui permet d'infirmer ou de confirmer les théories qu'il construit. Le théoricien postule des relations ; l'application de méthodes économétriques fournit des estimations sur la valeur des coefficients ainsi que la précision attendue.

Une question se pose alors : pourquoi estimer ces relations, et les tester statistiquement ? Plusieurs raisons incitent à cette démarche : tout d'abord cela force l'individu à établir clairement et à estimer les interrelations sous-jacentes. Ensuite, la confiance aveugle dans l'intuition peut mener à l'ignorance de liaisons importantes ou à leur mauvaise utilisation. De plus, des relations marginales mais néanmoins explicatives, qui ne sont qu'un élément d'un modèle global, doivent être testées et validées afin de les mettre à leur véritable place. Enfin, il est nécessaire de fournir, en même temps que l'estimation des relations, une mesure de la confiance que l'économiste peut avoir en celles-ci, c'est-à-dire la précision que l'on peut en attendre. Là encore, l'utilisation de méthodes purement qualitatives exclut toute mesure quantitative de la fiabilité d'une relation.

## B. L'économétrie comme outil d'investigation

L'économétrie n'est pas seulement un système de validation, mais également un outil d'analyse. Nous pouvons citer quelques domaines où l'économétrie apporte une aide à la modélisation, à la réflexion théorique ou à l'action économique par :

– la mise en évidence de relations entre des variables économiques qui n'étaient pas *a priori* évidentes ou pressenties ;

– l'induction statistique ou l'inférence statistique consiste à inférer, à partir des caractéristiques d'un échantillon, les caractéristiques d'une population. Elle permet de déterminer des intervalles de confiance pour des paramètres du modèle ou de tester si un paramètre est significativement[1] inférieur, supérieur ou simplement différent d'une valeur fixée ;

1. Au sens statistique, c'est-à-dire avec un seuil (risque d'erreur à ne pas dépasser, souvent 5 %).

– la simulation qui mesure l'impact de la modification de la valeur d'une variable sur une autre $(\Sigma C_t = a_1 \Sigma Y_t)$ ;
– la prévision[1], par l'utilisation de modèles économétriques, qui est utilisée par les pouvoirs publics ou l'entreprise afin d'anticiper et éventuellement de réagir à l'environnement économique.

Dans cet ouvrage, nous nous efforcerons de montrer, à l'aide d'exemples, les différentes facettes de l'utilisation des techniques économétriques dans des contextes et pour des objectifs différents.

# III. La théorie de la corrélation

## A. Présentation générale

Lorsque deux phénomènes ont une évolution commune, nous disons qu'ils sont « corrélés ». La corrélation simple mesure le degré de liaison existant entre ces deux phénomènes représentés par des variables. Si nous cherchons une relation entre trois variables ou plus, nous ferons appel alors à la notion de corrélation multiple.

Nous pouvons distinguer la corrélation linéaire, lorsque tous les points du couple de valeurs $(x,y)$ des deux variables semblent alignés sur une droite, de la corrélation non linéaire lorsque le couple de valeurs se trouve sur une même courbe d'allure quelconque.

Deux variables peuvent être :

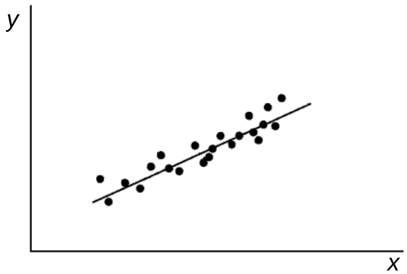
– en corrélation positive ; on constate alors une augmentation (ou diminution, ou constance) simultanée des valeurs des deux variables ;
– en corrélation négative, lorsque les valeurs de l'une augmentent, les valeurs de l'autre diminuent ;
– non corrélées, il n'y a aucune relation entre les variations des valeurs de l'une des variables et les valeurs de l'autre.

Le tableau 1, en croisant les critères de linéarité et de corrélation, renvoie à une représentation graphique.

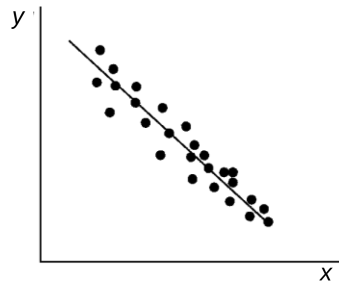
1. Pour découvrir l'utilisation de l'économétrie à des fins de prévision de ventes, voir Bourbonnais R. et Usunier J. C. (2013).

*Tableau 1 – Linéarité et corrélation*

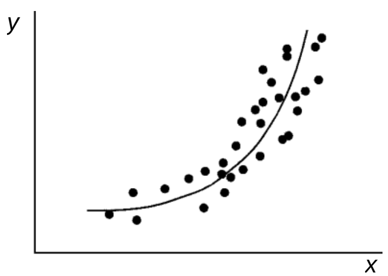
| | Corrélation positive | Corrélation négative | Absence de corrélation |
|---|---|---|---|
| Relation linéaire | Graphe 1 | Graphe 2 | Graphe 5 |
| Relation non linéaire | Graphe 3 | Graphe 4 | Graphe 5 |



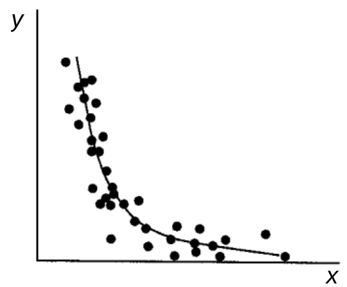

*Graphe 1*



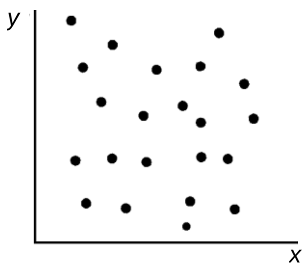

*Graphe 2*




*Graphe 3*




*Graphe 4*




*Graphe 5*

# B. Mesure et limite du coefficient de corrélation

## 1) Le coefficient de corrélation linéaire

La représentation graphique ne donne qu'une « impression » de la corrélation entre deux variables sans donner une idée précise de l'intensité de la liaison, c'est pourquoi nous calculons une statistique appelée *coefficient de corrélation linéaire simple,* noté $r_{x,y}$. Il est égal à :

$$r_{x,y} = \frac{\text{Cov}\,(x,y)}{\sigma_x\,\sigma_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad [1]$$

avec :

Cov $(x,y)$ = covariance entre $x$ et $y$ ;
$\sigma_x$ et $\sigma_y$ = écart type de $x$ et écart type de $y$ ;
$n$ = nombre d'observations.

En développant la formule [1], il vient :

$$r_{x,y} = \frac{n\sum_{i=1}^{n}x_i\,y_i - \sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{\sqrt{n\sum_{i=1}^{n}x_i^2 - \left(\sum_{i=1}^{n}x_i\right)^2}\sqrt{n\sum_{i=1}^{n}y_i^2 - \left(\sum_{i=1}^{n}y_i\right)^2}} \qquad [2]$$

On peut démontrer que, par construction ce coefficient reste compris entre $-1$ et 1 :

– proche de 1, les variables sont corrélées positivement ;

– proche de $-1$, les variables sont corrélées négativement ;

– proche de 0, les variables ne sont pas corrélées.

Dans la pratique, ce coefficient est rarement très proche de l'une de ces trois bornes et il est donc difficile de proposer une interprétation fiable à la simple lecture de ce coefficient. Ceci est surtout vrai en économie où les variables sont toutes plus au moins liées entre elles. De plus, il n'est calculé qu'à partir d'un échantillon d'observations et non pas sur l'ensemble des valeurs. On appelle $\rho_{x,y}$

ce coefficient empirique qui est une estimation du coefficient vrai $r_{x,y}$. La théorie des tests statistiques nous permet de lever cette indétermination.

Soit à tester l'hypothèse H0 : $r_{x,y} = 0$, contre l'hypothèse H1 : $r_{x,y} \neq 0$.

Sous l'hypothèse H0, nous pouvons démontrer que $\dfrac{\rho_{x,y}}{\sqrt{\dfrac{1-\rho_{x,y}^2}{n-2}}}$ suit une loi de Student à $n-2$ degrés de liberté[1]. Nous calculons alors une statistique, appelé le $t$ de Student empirique :

$$\boxed{t^* = \frac{|\rho_{x,y}|}{\sqrt{\dfrac{1-\rho_{x,y}^2}{n-2}}}} \qquad [3]$$

Si $t^* > t_{n-2}^{\alpha/2}$ valeur lue dans une table de Student[2] au seuil $\alpha = 0{,}05$ (5 %) à $n-2$ degrés de liberté[3], nous rejetons l'hypothèse H0, le coefficient de corrélation est donc significativement différent de 0 ; dans le cas contraire, l'hypothèse d'un coefficient de corrélation nul est acceptée. La loi de Student étant symétrique, nous calculons la valeur absolue du $t$ empirique et nous procédons au test par comparaison avec la valeur lue directement dans la table.

1. La notion de degrés de liberté est explicitée au chapitre 2.
2. Les lois de probabilité sont en fin d'ouvrage.
3. Si le nombre d'observations $n$ est supérieur à 30, on peut approximer la loi de Student par une loi normale, soit $t^{\alpha/2} \approx 1{,}96$.

## Exercice n° 1

### fichier C1EX1

### Calcul d'un coefficient de corrélation

Un agronome s'intéresse à la liaison pouvant exister entre le rendement de maïs $x$ (en quintal) d'une parcelle de terre et la quantité d'engrais $y$ (en kilo). Il relève 10 couples de données consignés dans le tableau 2

*Tableau 2 – Rendement de maïs et quantité d'engrais*

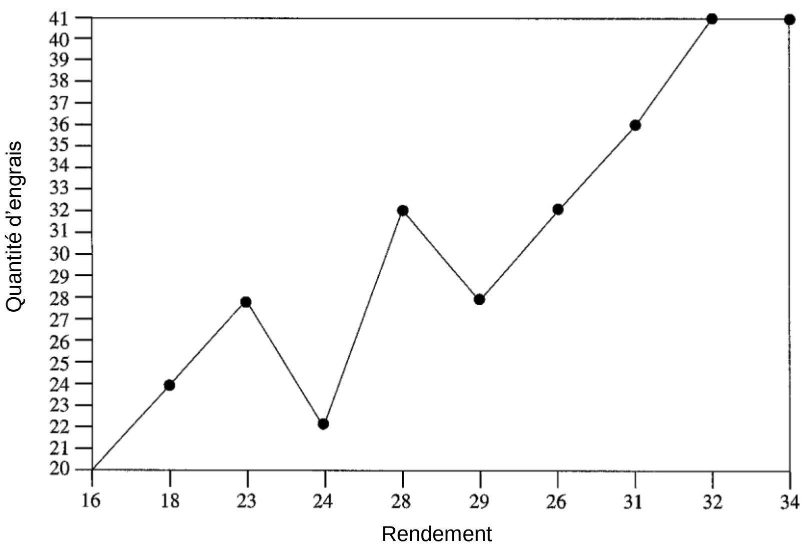
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Rendement $x$ | 16 | 18 | 23 | 24 | 28 | 29 | 26 | 31 | 32 | 34 |
| Engrais $y$ | 20 | 24 | 28 | 22 | 32 | 28 | 32 | 36 | 41 | 41 |

**1)** Tracer le nuage de points et le commenter.

**2)** Calculer le coefficient de corrélation simple et tester sa signification par rapport à 0 pour un seuil $\alpha = 0,05$.

**Solution**

**1)** Le nuage de points (graphique 6) indique que les couples de valeurs sont approximativement alignés : les deux variables semblent corrélées positivement.




*Graphique 6 – Nuage du couple de valeurs : rendement-quantité d'engrais*

**2)** Afin d'appliquer la formule [2], nous dressons le tableau de calcul 3.

*Tableau 3 – Calcul d'un coefficient de corrélation*

| | $x$ | $y$ | $x^2$ | $y^2$ | $xy$ |
|---|---|---|---|---|---|
| | 16 | 20 | 256 | 400 | 320 |
| | 18 | 24 | 324 | 576 | 432 |
| | 23 | 28 | 529 | 784 | 644 |
| | 24 | 22 | 576 | 484 | 528 |
| | 28 | 32 | 784 | 1 024 | 896 |
| | 29 | 28 | 841 | 784 | 812 |
| | 26 | 32 | 676 | 1 024 | 832 |
| | 31 | 36 | 961 | 1 296 | 1 116 |
| | 32 | 41 | 1 024 | 1 681 | 1 312 |
| | 34 | 41 | 1 156 | 1 681 | 1 394 |
| Somme | 261 | 304 | 7 127 | 9 734 | 8 286 |

$$\rho_{x,y} = \frac{(10)(8\,286) - (261)(304)}{\sqrt{(10)(7\,127) - 261^2}\,\sqrt{(10)(9\,734) - 304^2}} = \frac{3\,516}{(56{,}11)(70{,}17)}$$

soit $\rho_{x,y} = 0{,}89$ et $\rho^2_{x,y} = 0{,}79$

Le $t$ de Student empirique (d'après [3]) est égal à :

$$t^* = \frac{|\rho_{x,y}|}{\sqrt{\dfrac{(1-\rho^2_{x,y})}{n-2}}} = \frac{0{,}89}{0{,}1\,620} = 5{,}49 > t_8^{0{,}025} = 2{,}306$$

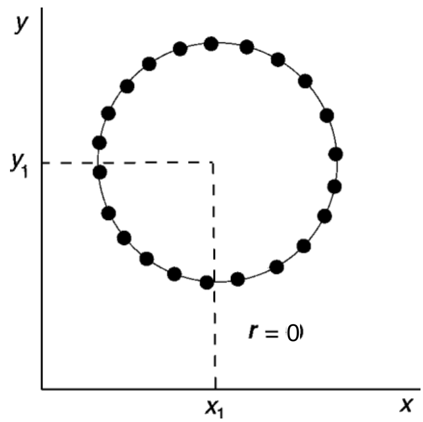
le coefficient de corrélation entre $x$ et $y$ est significativement différent de 0.

## 2) Limites de la notion de corrélation

### a) La relation testée est linéaire

L'application de la formule [1] ou [2] ne permet de déterminer que des corrélations linéaires entre variables. Un coefficient de corrélation nul indique que la covariance entre la variable $x$ et la variable $y$ est égale à 0. C'est ainsi que deux variables en totale dépendance peuvent avoir un coefficient de corrélation nul, comme l'illustre l'exemple suivant : l'équation d'un cercle nous est donnée par $(x - x_1)^2 + (y - y_1)^2 = R^2$ ; les variables $x$ et $y$ sont bien liées entre elles fonctionnellement (graphique 7) et pourtant leur covariance est nulle et donc leur coefficient de corrélation égal à 0.

Pour pallier cette limite, il convient éventuellement de transformer les variables, préalablement au calcul du coefficient de corrélation, afin de linéariser leur relation, par exemple au moyen d'une transformation de type logarithmique.




*Graphique 7 – Relation fonctionnelle n'est pas corrélation linéaire*

**b) Corrélation n'est pas causalité**

Le fait d'avoir un coefficient de corrélation élevé entre deux variables ne signifie pas qu'il existe un autre lien que statistique. En d'autres termes, une covariance significativement différente de 0 n'implique pas une liaison d'ordre économique, physique ou autre. Nous appelons *corrélation fortuite* ce type de corrélation que rien ne peut expliquer.

L'exemple le plus fameux concerne la forte corrélation existante entre le nombre de taches solaires observées et le taux de criminalité aux États-Unis. Cela ne signifie pas qu'il existe une relation entre les deux variables, mais qu'une troisième variable, l'évolution de long terme (la tendance) ici, explique conjointement les deux phénomènes. La théorie de la cointégration traite de ce problème (*cf.* chapitre 11).

# 2. Le modèle de régression simple

Nous commençons notre étude par le modèle le plus simple : une variable endogène est expliquée par une variable exogène. Après avoir étudié les conséquences probabilistes de l'erreur d'observation, nous présentons en I. les formules de base permettant d'estimer les paramètres du modèle. Les hypothèses stochastiques et leurs conséquences sont étudiées au paragraphe II.
En III. et IV., la qualité de l'estimation d'un modèle est examinée à l'aide des premiers tests statistiques (Student, Fisher).
Enfin, en V., le modèle de régression simple est étudié en tant qu'outil de prévision avec le degré de confiance que nous pouvons en attendre.

## I. Présentation du modèle

### A. Exemple introductif

Soit la fonction de consommation keynésienne :

$$C = a_0 + a_1 Y$$

où :
$C$ = consommation,
$Y$ = revenu,
$a_1$ = propension marginale à consommer,
$a_0$ = consommation autonome ou incompressible.

### 1) Vocabulaire

- La variable consommation est appelée « variable à expliquer » ou « variable endogène ».
- La variable revenu est appelée « variable explicative » ou « variable exogène » (c'est le revenu qui explique la consommation).
- $a_1$ et $a_0$ sont les paramètres du modèle ou encore les coefficients de régression.

### 2) Spécification

Nous pouvons distinguer deux types de spécifications :

- Les modèles en série temporelle, les variables représentent des phénomènes observés à intervalles de temps réguliers, par exemple la consommation et le revenu annuel sur 20 ans pour un pays donné. Le modèle s'écrit alors :

$$C_t = a_0 + a_1 Y_t \qquad t = 1, \ldots, 20$$

où :

$C_t =$ consommation au temps $t$,
$Y_t =$ revenu au temps $t$.

- Les modèles en coupe instantanée, les variables représentent des phénomènes observés au même instant mais concernant plusieurs individus, par exemple la consommation et le revenu observés sur un échantillon de 20 pays. Le modèle s'écrit alors :

$$C_i = a_0 + a_1 Y_i \qquad i = 1, \ldots, 20$$

où :

$C_i =$ consommation du pays $i$ pour une année donnée,

$Y_i =$ revenu du pays $i$ pour une année donnée.

## B. Rôle du terme aléatoire

Le modèle tel qu'il vient d'être spécifié n'est qu'une caricature de la réalité. En effet ne retenir que le revenu pour expliquer la consommation est à l'évidence même insuffisant ; il existe une multitude d'autres facteurs susceptibles d'expliquer la consommation. C'est pourquoi nous ajoutons un terme $(\varepsilon_t)$ qui synthétise l'ensemble de ces informations non explicitées dans le modèle : $C_t = a_0 + a_1 Y_t + \varepsilon_t$ si le modèle est spécifié en série temporelle $(C_i = a_0 + a_1 Y_i + \varepsilon_i$ si le modèle est spécifié en coupe instantanée), où $\varepsilon_t$ représente l'erreur de spécification du modèle, c'est-à-dire l'ensemble des phénomènes explicatifs de la consommation non liés au revenu. Le terme $\varepsilon_t$ mesure la différence entre les valeurs réellement observées de $C_t$ et les valeurs qui auraient été observées si la relation spécifiée avait été rigoureusement exacte. Le terme $\varepsilon_t$ regroupe donc trois erreurs :

– une erreur de spécification, c'est-à-dire le fait que la seule variable expli- cative n'est pas suffisante pour rendre compte de la totalité du phénomè- ne expliqué ;

– une erreur de mesure, les données ne représentent pas exactement le phé- nomène ;

– une erreur de fluctuation d'échantillonnage, d'un échantillon à l'autre les observations, et donc les estimations, sont légèrement différentes.

## Exercice n° 1

### fichier C2EX1

**Génération d'une consommation aléatoire**

Le tableau 1 présente le revenu moyen par habitant sur 10 ans exprimé en dollars pour un pays.

*Tableau 1 – Évolution du revenu moyen par habitant en dollars*

| Année | Revenu |
|---|---|
| 1 | 8 000 |
| 2 | 9 000 |
| 3 | 9 500 |
| 4 | 9 500 |
| 5 | 9 800 |
| 6 | 11 000 |
| 7 | 12 000 |
| 8 | 13 000 |
| 9 | 15 000 |
| 10 | 16 000 |

Sachant que la propension marginale à consommer est de 0,8 et que la consomma- tion incompressible est 1 000, on demande :

**1)** de calculer la consommation théorique sur les 10 ans ;

**2)** considérant que notre erreur d'observation suit une loi normale de moyenne 0 et de variance 20 000, de générer cette variable aléatoire et de calculer une consommationobservée tenant compte de cette erreur.

**Solution**

Les calculs des questions **1)** et **2)** sont présentés dans le tableau 2.

La consommation théorique (colonne 3) est calculée par application directe de la formule : $C_t = 1\,000 + 0{,}8\,Y_t$.

Recherche Opérationnelle et Optimisation

### Les origines de la recherche opérationnelle

Si la recherche opérationnelle, en abrégé RO, est aujourd'hui présente dans la plupart des domaines civils, ses racines sont habituellement attribués aux services militaires. La seconde guerre mondiale, de part son envergure, créa une besoin urgent d'allouer de manière efficace des ressources limitées aux différentes opérations militaires et aux activités au sein de chaque opération. En particulier, l'organisation militaire britannique, puis américaine, mis à contribution un grand nombre de scientifiques pour gérer ces allocations, et s'occuper d'autres problèmes stratégiques et tactiques. Ce faisant, ils furent appelés à poursuivre des recherches sur des opérations (militaires), et constituèrent les premières équipes de RO. Leurs efforts furent signifactifs dans la marche vers la victoire, par exemple en ce qui touche l'utilisation du radar, nouvellement développé. Ces succès encouragèrent la poursuite de l'utilisation de la RO dans d'autres domaines. La croissance importante de l'industrie d'après-guerre entraîna des problèmes, causés par la complexité croissante et la spécialisation dans les organisations, problèmes en fait proches de ceux présent lors du conflit. Au début des années 1950's, la RO avait pénétré une multitude d'organisations commerciales, industrielles, et gouvernementales. Et ce n'était que le début.
Au moins deux autres facteurs ont joué un rôle clé dans la croissance rapide de la RO. Tout d'abord, des progrès substantiels ont été obtenus très tôt afin d'améliorer les techniques de RO. Ces techniques, dans leur mise en pratique, furent soutenues par l'essor des outils informatiques [1].

### 2. Définition de la Recherche Opérationnelle RO

La recherche Opérationnelle est une discipline qui permet de formuler des problèmes par des supports scientifiques, mathématiques et informatiques pour aider à mieux décider. La Recherche Opérationnelle (R.O) est avant tout un outil d'aide à la décision.
Autrement définie, la Recherche Opérationnelle (RO) traduit des énoncés ou des cahiers de charges liés à des problématiques spécifiques sous forme de méthodes et des démarches à base d'équation mathématique, des algorithmes et des outils statistiques.
La Recherche Opérationnelle (RO) permet d'assurer la communication entre le milieu industriel (les secteurs extérieurs) et les applications théoriques dans le but de proposer les solutions les mieux adaptées à une situation donnée [2].
La procédure utilisée par la recherche Opérationnelle RO peut être schématisée comme suit :

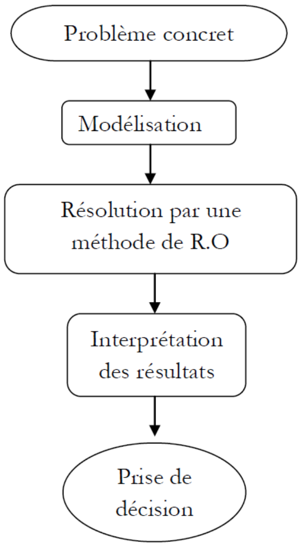



Figure 1. Schéma de procédure utilisée par la RO

## 3. Quelques problèmes de recherche opérationnelle

• Comment aller le plus vite de Batna à Oran, en voiture ?

• Comment ordonnancer les tâches d'un projet en fonction de la main d'oeuvre, tout en minimisant sa durée ?

• Comment investir ses 100000 DA d'économie de sorte à maximiser le profit obtenu après deux ans ?

• Trouver un (plus court) chemin entre deux villes : problème du plus court chemin dans les graphes

• Broadcast de coût minimum dans un réseau : problème des arbres recouvrant de poids minimum.

• Envoi d'un maximum d'information dans un réseau : problème du flot maximum [3].

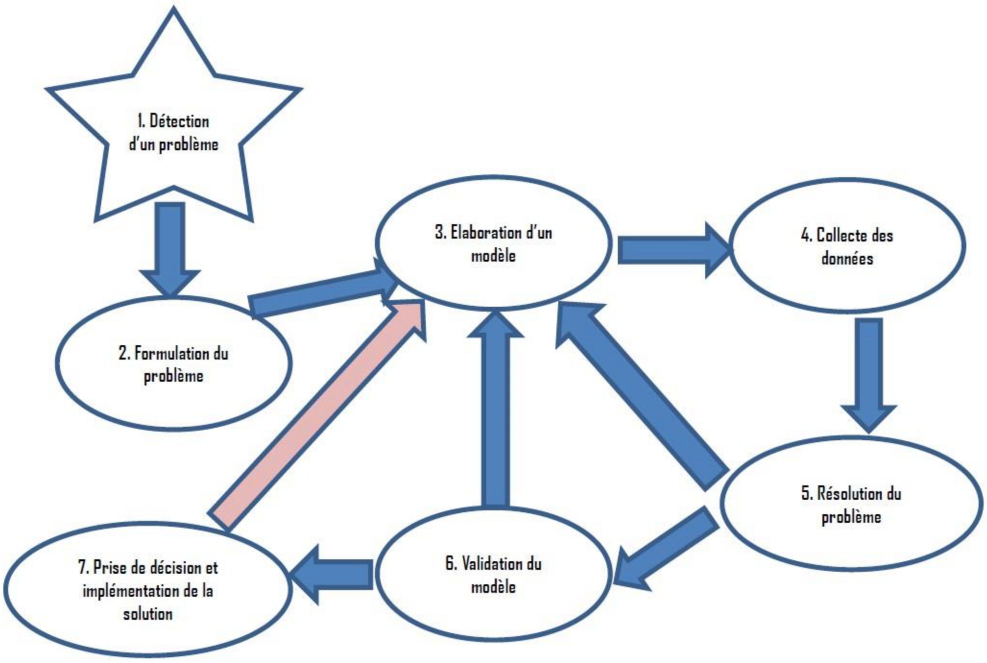



Figure 2. Vue d'ensemble de la modélisation d'un problème de recherche opérationnelle

1. **La détection d'un problème** : Les nécessités de l'action viennent des expériences vécues ; c'est la phase préscientifique.

2. **La formulation du problème** : Quel est le vrai problème à résoudre ? Quels critères permettent de juger si le problème est résolu de façon satisfaisante ?

3. **L'élaboration d'un modèle** : Il s'agit de représenter les principaux aspects de la réalité par un ensemble de formules, mathématiques le plus souvent qui mettent en jeu les variables de décisions concernées et leurs interactions. On lance des hypothèses, on élabore une théorie, on écrit un modèle. C'est la phase de conceptualisation, de construction théorique ; en un mot, c'est la phase de modélisation.

4. **La collecte des données** : Il faut préciser les paramètres du modèle en s'appuyant sur l'information recueillie dans l'environnement du problème à résoudre. L'élaboration du modèle s'éclaire à la lanterne des données. Le processus peut requérir plusieurs cycles impliquant les étapes 3, 4 et 5.

5. **La résolution du modèle** : C'est la phase où l'on souhaite recourir aux méthodes appropriées déjà disponible si on a réussi à classer le problème parmi ceux pour lesquels, on connait déjà une méthode d'approche. Sinon, il faut recourir à la simulation ou inventer une technique de résolution.

6. **La validation du modèle** : On confronte les conclusions obtenues du modèle aux opinions des personnes qui ont suffisamment d'expérience du problème traité pour apprécier ou critiquer la pertinence de la solution proposée. Si les avis reçus sont négatifs, on peut alors remettre en cause soit l'écriture du modèle retenu, soit la valeur de ses paramètres, soit les critères d'appréciation de la solution. On peut aller jusqu'à remettre en cause l'approche choisie pour résoudre le problème et partant le modèle retenu.

7. **La prise de décision et l'implémentation de la solution** : Comment implémenter la solution obtenue ? Doit-on s'arrêter là ? Il y a ici un retour possible vers le modèle initial pour le modifier ou l'enrichir des observations faites lors de la phase expérimentale. Une fois les révisions nécessaires apportées, le modèle enrichi permettra de tirer des conclusions mieux étayées.

## 4. Formulation d'un problème d'optimisation

La modélisation du problème ou le passage du texte vers les équations mathématiques ne se fait pas de manière aléatoire. Il doit vérifier certaines conditions et doit passer par plusieurs étapes
: l'analyse, la modélisation et le critère à optimiser décrit comme suit :

### 4.1 Analyse

Dans cette phase, on commence toujours par l'identification des données nécessaires à la résolution du problème, les valeurs de consommations ou les fiches techniques des articles, la quantité des ressources disponibles, les coûts de transport, la capacité de production, les gains engendrés par chaque article ou les dépenses établies pour la réalisation d'une action. Dans cette phase, on détermine les composantes, les enjeux et les limites du problème. Toutes ces informations sont primordiales pour la formulation des problèmes [2].

### 4.2 Modélisation

L'optimisation repose toujours sur des modèles mathématiques qui sont généralement simples et partiels. Pour définir le modèle du système, on définit :

**- les variables**

Une fois l'analyse effectuée, on détermine les variables du Système qui représente les décisions à prendre et qui doit répondre à un certain nombre de questions : que faut-il réaliser et en quelle quantité ? Combien doit-on acheter ? Combien de classes à faire dans une école ? Les variables représentent les inconnues du système.

**- Les contraintes**

Les objets mathématiques qui assurent l'interaction et la liaison des variables par rapport aux ressources disponibles et aux données du problème sont appelées contraintes.

### 4. 3 Application à un exemple

Une entreprise fabrique des tables et des chaises à partir de deux matières : le bois et la peinture, sachant que la réalisation d'une table nécessite **3 m** de bois et **4 kg** de peinture la réalisation d'une chaise nécessitent **2 m** de bois et **1 kg** de peinture. Les moyens financiers de l'entreprise acceptent un approvisionnement de **100 m** de bois et **120 kg** de peinture par semaine. Les produits ainsi fabriqués fournissent un bénéfice de **500 DA** par table et **300 DA** par chaise vendu.

***Question*** : formuler le problème

***Solution*** **:** reprenons les données de l'exemple sur un tableau :

| | Table | Chaise | Stock |
|---|---|---|---|
| Bois | 3 | 2 | 100 |
| Peinture | 4 | 1 | 120 |
| Gain | 500 | 300 | |

**-Les variables :**

x1 : représente la quantité des tables ; x2 :

représente la quantité des chaises.

**-Les contraintes :**

Pour le bois : $3x1 + 2x2 \leq 100$ (1)

Pour la peinture : $4x1 + 1x2 \leq 120$ (2)

Pour soulever le problème de la non-négativité des variables, puisqu'on ne peut pas produire moins deux tables ou moins vingt chaises nous supposons que les variables sont positives, en ajoutons une contrainte supplémentaire au problème :

$$x1, x2 \geq 0 \qquad (3)$$

$$\Rightarrow \text{Les contraintes :} \quad \begin{cases} 3x1 + 2x2 \leq 100 & (1) \\ 4x1 + 5x2 \leq 120 & (2) \\ x1, x2 \geq 0 & (3) \end{cases}$$

**- Fonction objective :**

Le gain : $z = 500x1 + 300x2$ $\Rightarrow$ L'objectif : max $z = 500x1 + 300x2$

### 4.4 Solutions d'un problème

**- Solution admissible**

On appelle solution admissible (ou solution) d'un problème tout vecteur $x \in S$ qui satisfait toutes les contraintes du problème.

**- Solution optimale**

On appelle solution optimale x*, la solution parmi toutes les solutions admissibles qui fournissent le meilleur résultat, c'est à dire de trouver le max pour un problème de maximisation ou trouver la min pour un problème de minimisation.

**Exemples d'applications**

**Problème de production.**

Une usine fabrique 2 produits P1 et P2 en utilisant un certain nombre de ressources : post opératoire (machine), main-d'oeuvre et emballage. Ces besoins sont indiqués dans le tableau ci- dessous. Par ailleurs, chaque ressource est disponible en quantités limitées (cf. tableau).

| | P1 | P2 | Ressource disponible |
|---|---|---|---|
| Heure /machine | 3 | 9 | 81 |
| main d'oeuvre | 4 | 5 | 55 |
| emballage | 2 | 1 | 20 |

Les deux produits P1 et P2 rapportent à la vente respectivement des bénéfices de 600 DA et 400 DA par unité. Quelles quantités de produits P1 et P2, doit produire l'usine afin de maximiser le bénéfice total venant de la vente des 2 produits ?

- **Choix des variables (les inconnues)**

x1 et x2 sont respectivement les quantités des produits P1 et P2 à fabriquer ($x1, x2 \in N$).

- **Choix de la fonction objectif à maximiser**

La fonction objective F correspond au bénéfice total. Elle vaut : $F(x1, x2) = 600x1 + 400x2$.

Le problème se traduit donc par :

Max $F = 600\,x1 + 400\,x2$.

- **Détermination des contraintes**

a) La disponibilité de chacune des ressources s'écrit sous la forme suivante :

$$3x1 + 9x2 \leq 81$$

$$4x1 + 5x2 \leq 55$$

$$2x1 + x2 \leq 20$$

b) Positivité des variables : $x1, x2 \geq 0$.

En résumé, le problème de production se modélise sous la forme Max $F = 600\,x1 + 400\,x2$.

$$\begin{cases} 3x1 + 9x2 \leq 81 \\ 4x1 + 5x2 \leq 55 \\ 2x1 + x2 \leq 20 \quad x1, \\ x2 \geq 0. \end{cases}$$

# TD N° 1

## Formulation mathématique

Pour chaque exercice, formuler le problème de programmation linéaire

**Exercice 1 :** Un atelier de couture fabrique en série deux modèles de tablier. Le premier modèle nécessite 1 mètre de tissu, 4 heures de travail et rapporte 240 DA. Un tablier du deuxième modèle exige 2 mètres de tissu, 2 heures de travail et rapporte 160 DA. Sachant que l'atelier dispose mensuel de 150 mètres de tissu et de 400 heures de travail, et qu'il peut vendre toute sa fabrication. Combien de tablier faut-il fabriquer pour obtenir un bénéfice maximal ?

**Exercice 2 :** Un fabricant de gravier pour cours de maison en produit deux catégories : du gravier grossier et du gravier fin. Le gravier grossier nécessite 2h de broyage, 5h de criblage et de 8h de séchage, tandis que le gravier fin nécessite six heures de broyage, trois heures de criblage et deux heures de séchage. Le fabricant dispose de 36 heures pour le broyage, de 30 heures pour le criblage et de 40 heures pour le séchage. La marge de profit est de 4000 DA par unité de gravier grossier et de 5000 DA par unité de gravier fin. Le fabricant désire maximiser sa marge bénéficiaire.

**Exercice 3 :** Une association culturelle organise une exposition, pendant cette exposition des tasses de café au lait et des tasses de chocolat au lait sont vendues pour apporter une aide aux orphelins. Un sponsor a permis de procurer 40 litres de lait, 4 kg de sucre et assez de café et de chocolat pour faire 150 tasses de chaque boisson.
On prévoit de servir 2 sucres par tasse en moyenne ; chaque paquet d'un kilogramme de sucre contient 120 morceaux ; il faut 1/4 de litre de lait pour une tasse de chocolat et
1/12 de litre de lait pour une tasse de café.
Le trésorier du club propose de vendre 50 DA chaque tasse de chocolat au lait et 40 DA chaque tasse de café au lait.
Déterminez le nombre de tasses de chaque sorte à servir et calculez la recette maximale collectée

**Exercice 4 :** Une entreprise disposant de 10 000 m2 de bois en réserve, fabrique et commercialise 2 types de boîtes en bois. La fabrication d’une boîte en bois de type 1 et de type
2 nécessite, respectivement, 1 et 2 m2 de carton ainsi que 20 et 30 minutes de temps d’assemblage. Seules 200 heures de travail sont disponibles pendant la semaine à venir.
Les boites sont clouer et il faut quatre fois plus de clou pour une boîte du second type que pour une du premier. Le stock de clous disponible permet d’assembler au maximum 15000 boîtes du premier type. Les boites sont vendues, respectivement, 300 DA et 500 DA. Déterminez le nombre de boite de chaque type pour avoir la recette maximale.

# II : Programmation linéaire

## 2.1 Introduction

En mathématiques, les problèmes de **programmation linéaire** (PL) sont des problèmes d'optimisation où la fonction objective et les contraintes sont toutes linéaires.

## 2.2 Forme d’un programme linéaire

### 2.2.1 Forme générale

La forme générale d’un problème d’optimisation est la suivante :
Maximiser f (x) {ou Minimiser f(x)}
sous
gi(x) = 0 i ∈ I $^0$ contraintes égalité gi(x) ≤ 0 i ∈
I − contraintes inégalité gi(x) ≥ 0 i ∈ I +
contraintes inégalité x = [x1, ..., xn] T ≥ 0
D'une manière plus détaillée, un programme linéaire (PL) est dit sous forme générale s’il s'écrit de la façon suivante :

$$\max F = c_1 x_1 + c_2 x_2 + \ldots\ldots + c_n x_n \qquad (2.1)$$

La fonction 2.1 peut s'exprimer par une sommation de n terme (variables) comme suit :

$$\max F = \sum_j^n C_j X_j \qquad (2.2)$$

Sous les contraintes :

$$\left.\begin{array}{ll} & \sum_j^n a_{rj} X_j = br \\ \text{Ou} & \sum_{j=1}^n a_{ij} X_j \geq bi \\ \text{Ou} & \sum_{j=1}^n a_{kj} X_j \leq bk \end{array}\right\} \qquad (2.3)$$

### 2.2.2 Forme standard ou forme canonique d’un programme linéaire

Un problème linéaire est sous forme standard si toutes les contraintes sont des contraintes égalité. Pour transformer une contrainte d'inégalité en contrainte d'égalité, il faut ajouter vaux membres de gauche d’une contrainte, une quantité (une mesure) appelée variable d'écart qui sert à absorber l'écart entre le membre gauche et le membre droit d'une contrainte si l'écart existe, bien sûr, sinon cette variable vaut zéro.

La forme standard peut être donnée par la forme suivante :

$$\max z=CX$$

sous

$$\begin{cases} AX=b & \alpha \\ X \geq 0 & \beta \end{cases} \qquad (2.4)$$

Cette forme standard est obtenue en introduisant des variables d'écart dans toutes les contraintes d'inégalité(à) et que ces variables soient non négatives(β).

La forme canonique d'un programme linéaire peut être exprimée par un ensemble de vecteurs comme suit :

$$X = (x_1, x_2, \ldots\ldots x_n) \in R^n$$
$$c = (c_1, c_2, \ldots\ldots, c_n) \in R^n \qquad (2.5)$$
$$b = (b_1, b_2, \ldots., b_n) \in R^n$$

Et la matrice A de taille m x n

$$\begin{bmatrix} a_{11} & a_{12} \ldots. & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{mn} \end{bmatrix}$$

avec:

n = nombre de variables,

m = nombre de contraintes,

X = vecteur des variables de décision

A = matrice réelle m × n (matrice des contraintes) c= [c1, ,

cn] =vecteur-ligne des couts,

b=[b1, ............., bm]T=vecteur-colonne des seconds membres

Z= CX est la fonction objective, ou critère à minimiser.

**Exemple :**

Reprenons l'exemple du problème de production de l'introduction. Sous forme standard (en introduisant des variables d'écart), le PL s'écrit :

$$\max F = 600x_1 + 400x_2$$

Sous contrainte

$$4x_1 + 5\,x_2 + x_3 = 55$$

$$x_1 + 3\,x_2 + x_4 = 27$$

$$2x_1 + x_2 + x_5 = 20$$

$\mathbf{x_1, x_2, x_3, x_4, x_5 \geq 0}$

**Pour cet exemple :**

- Le vecteur des variables dans le système est :

-Le vecteur des coefficients $c$ est donné par $c = (600, 400, 0, 0, 0)$ puisque max F peut s'écrire comme suit :

$$\max F = 600x_1 + 400x_2 + 0x_3 + 0x_4 + 0x_5$$

- Le vecteur des coefficients $b$ est donné par $b = (55, 27, 20)$

- La matrice A :

$$A = \begin{bmatrix} 4 & 5 & 1 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 \end{bmatrix}$$

## 2.3 Résolution de programmes linéaires

Il existe plusieurs techniques de résolution pour les programmes linéaires. Cela dit nous présenterons dans cette section : la résolution graphique et la résolution analytique en détaillant deux procédures : méthode Simplexe et tableaux Simplexe.

### 2.3.1 Résolution graphique

La résolution graphique d'un problème linéaire consiste à tracer la droite qui sépare les demi- plans pour chaque contrainte tout en conservant le demi-plan acceptable, c'est à dire le demi- plan des solutions réalisables pour la contrainte. L'intersection des différents demi-plans de toutes les contraintes sans oublier les contraintes de positivité forme le polygone des solutions, appelé aussi "région des solutions admissibles".

Exemples : soit le problème suivant :

$$\max Z = 600x_1 + 400x_2$$

Sous contrainte :

$4x_1 + 5\,x_2 \leq\ 55$

$x_1 + 3\,x_2 \leq 27$

$2x_1 + x_2 \leq 20$

$x_1, x_2 \geq\ 0$ (2.8)

Le problème possède trois contraintes plus la contrainte de positivité. On commence à tracer chaque contrainte séparément.

$4x_1 + 5\,x_2 \leq\ 55$

On prend la première contrainte du système et on remplace l'inégalité par une égalité sans l'ajout de variable d'écart, (cette façon de faire est applicable seulement pour la résolution graphique). L'équation résultante correspond à une droite. Pour tracer une droite, il faudrait déterminer deux points, ce qui donne pour la première contrainte notre exemple :

$4x_1 + 5\,x_2 =\ 55$

-Si x1=0 $\Rightarrow$ $x$2=11 le premier point est ($x$1, $x$ 2) =(0 , 11)

-Si x2=0 $\Rightarrow$ $x$1=55/4 = 13.75 le deuxième point est ($x$1, $x$ 2) =(13.75 , 0)

À partir de ces points on trace la première droite et on conserve ce qui est en dessus de la droite. On élimine ensuite la partie supérieure puisque la contrainte est une contrainte d'infériorité.

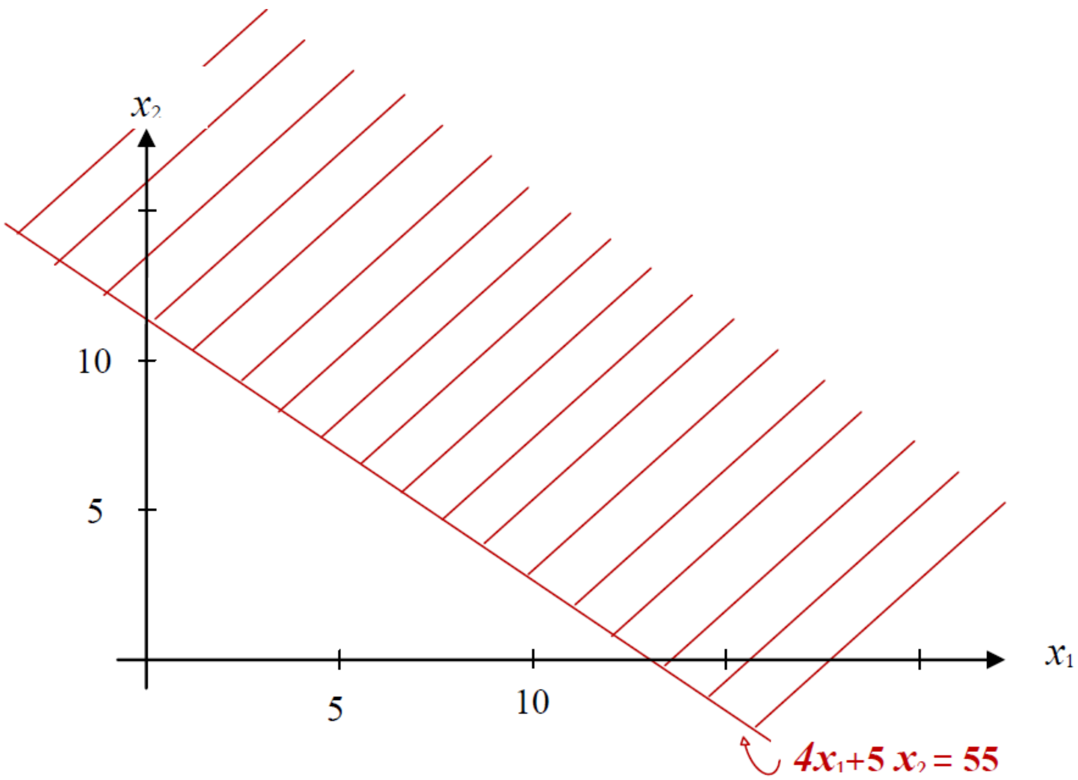



Figure. 3. Représentation de la première contrainte

Ainsi, on applique le même principe pour toutes les contraintes du système :

$$x_1 + 3\,x_2 = 27$$

-Si $x_1$=0 $\Longrightarrow$ $x_2$=9 le premier point est ($x_1$, $x_2$) = (0, 9)

-Si $x_2$=0 $\Longrightarrow$ $x_1$=27 le deuxième point est ($x_1$, $x_2$) = (27, 0)

$$2x_1 + x_2 = 20$$

-Si $x_1$=0 $\Longrightarrow$ $x_2$=20 le premier point est ($x_1$, $x_2$) = (0, 20)

-Si $x_2$=0 $\Longrightarrow$ $x_1$=10 le deuxième point est ($x_1$, $x_2$) = (10, 0)

Bien évidemment, il faut tracer aussi les contraintes de positivité. L'intersection de toutes les contraintes forme le polygone de solutions tel qu'il est donné par la figure 4.

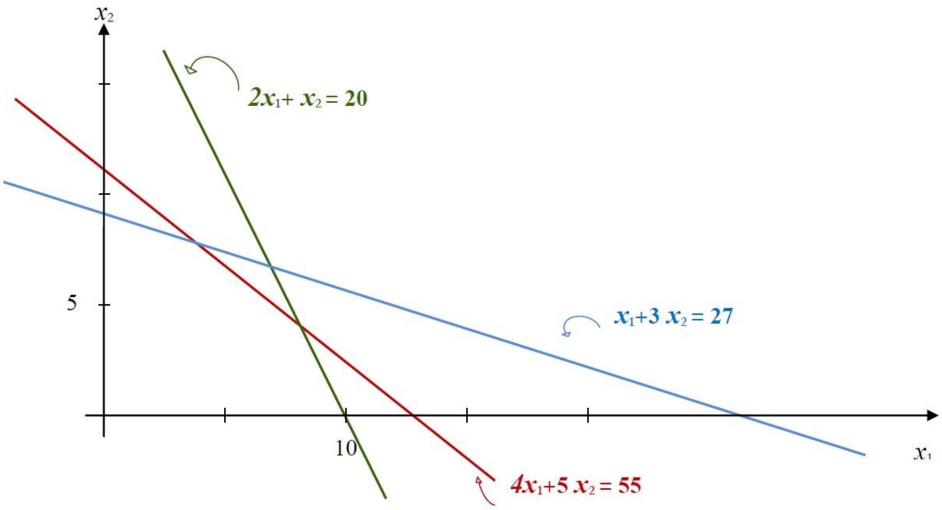



Figure. 4. Représentation des différentes contraintes

La question qui se pose maintenant est : quel est le point qui donne la valeur optimale pour ce problème ?
Pour répondre à cette question nous expliquerons la procédure à suivre pour déterminer la valeur maximale pour ce problème qui consiste à tracer la droite qui correspond à la fonction objective.
Tout d'abord, on prend l'équation de la fonction objective et on lui attribue la valeur de zéro puisque c'est la plus petite valeur pour la fonction objective. Ce qui donne pour cet exemple :

$$400x_1 + 600\,x_2 = 0,$$

$$\Longrightarrow 4\,x_1 = -\,6\,x_2 \Longrightarrow x_1 = \frac{-6}{4}\,x_2 \Longrightarrow x_1 = -1.5x$$

Le coefficient directeur de cette fonction est (−1, 1.5). Il passe par l'origine (0, 0).
Une fois la droite tracée on effectue une translation parallèle à la direction de la droite du bas vers le haut jusqu’à rencontrer le dernier point du polygone satisfaisant les contraintes.

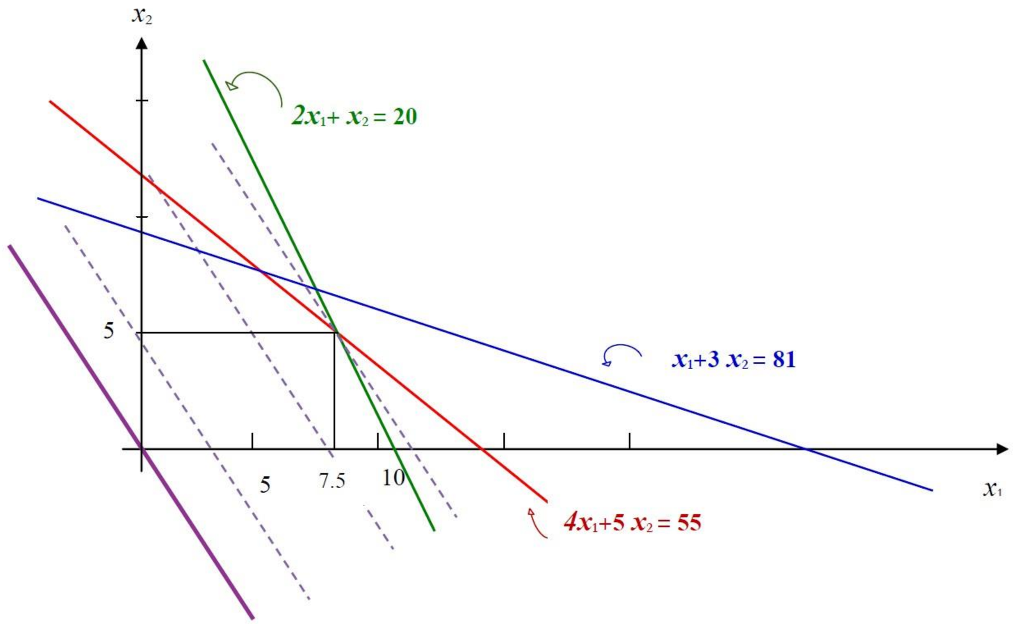


Figure. 5 Résolution graphique du problème de production

Le maximum de Z sur cet ensemble de contraintes est alors atteint. On obtient ainsi le point qui correspond à la solution optimale. Par la projection de ce point sur les axes x1 et x2 on obtient

: x1=7.5 et x2 = 5, ce qui donne une valeur maximale Z = 6000.

La méthode graphique pour la résolution d'un problème linéaire à deux variables est très facile à appliquer. Sa difficulté augmente par l'augmentation des variables dans le système.

Elle devient difficile pour trois variables et, voire impossible, au-delà de trois inconnus dans le système étudié.

Afin d'ôter cette difficulté, une méthode appelée méthode du simplexe a été proposée et développée par Dantzig afin de résoudre des problèmes de programmation linéaire avec plusieurs variables.

**Exemple 2**

Une entreprise fabrique des vêtements pour femmes, les profits réalisés sont de 50 DA par robe, et 30 DA par chemisier vendu, combien de robes et de chemisiers dit-elle vendre vendre pour réaliser un profit maximum en respectant les conditions suivantes ?

- On ne peut fabriquer plus de 80 items vêtement par mois.
- Il faut deux heures pour coudre une robe et une heure pour coudre un chemisier.

Or, la machine à coudre n'est disponible que pendant cent heures par mois.

X= nombres de robes

Y= nombres de chemisiers $X \geq 0$

$Y \geq 0$

P= 50X + 30Y X + Y
≤ 80  2X + Y ≤ 100

| | | |
|---|---|---|
| X | 0 | 80 |
| Y | 80 | 0 |

| | | |
|---|---|---|
| X | 0 | 50 |
| Y | 100 | 0 |

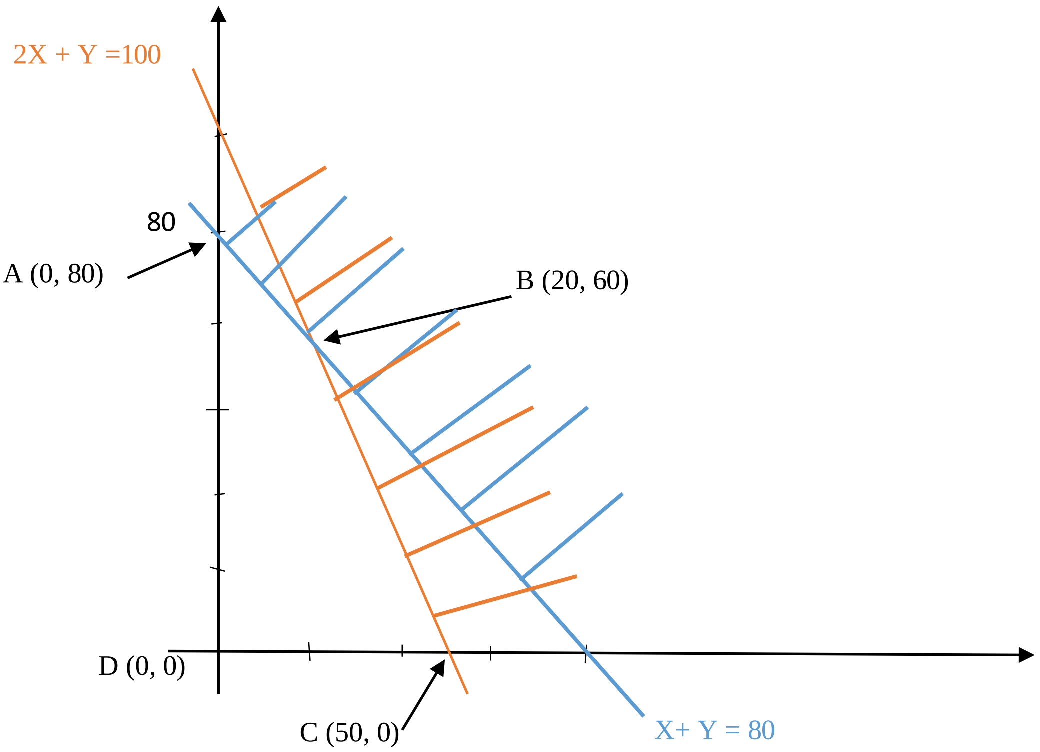



Figure. 6 Résolution graphique du problème de production

On va vérifier P= 50X
+ 30Y

P(A), (0,80) ⇒ P= 50(50) + 30 (80) =  2400

P(B), (20, 60) ⇒ P= 50(20) + 30 (60) = 2800

P(C), (50, 0) ⇒ P= 50(50, 0) + 30 (0) = 2500

P(D), (0,0) ⇒ P= 50(0) + 30(0) = 0

La solution optimale est 20 Robes et 60 Chemisiers

# TD N° 2

# Programmation Linéaire résolution graphique

Pour chaque exercice, formuler le problème de programmation linéaire et le résoudre graphiquement.

**Exercice 1 :** Un fabricant de basket fait un bénéfice de 800 DA sur chaque basket ordinaire et de 1500 DA sur chaque basket professionnel. Pour satisfaire a la demande des vendeurs, la production journalière de basket ordinaires devrait se situer entre 30 et 80, et la production journalière de basket professionnel 10 et 30. Pour maintenir une bonne qualité, le nombre de baskets produites ne devrait dépasser 80 par jour.
Combien de baskets de chaque type faudrait-il fabriquer pour réaliser un bénéfice maximum ?

**Exercice 2 :** Une association culturelle organise une exposition, pendant cette exposition des tasses de café au lait et des tasses de chocolat au lait sont vendues pour apporter une aide aux orphelins. Un sponsor a permis de procurer 20 litres de lait, 2 kg de sucre et assez de café et de chocolat pour faire 150 tasses de chaque boisson.
On prévoit de servir 2 sucres par tasse en moyenne ; chaque paquet de sucre contient 120 morceaux et le poids du paquet indiqué sur la boite est de 500 g ; il faut 1/4 de litre de lait pour une tasse de chocolat et 1/12 de litre de lait pour une tasse de café. Le trésorier du club propose de vendre 50 DA chaque tasse de chocolat au lait et 40 DA chaque tasse de café au lait.
Déterminez le nombre de tasses de chaque sorte à servir et calculez la recette maximale collectée ?

**Exercice 3 :** (Production de peinture). Une société produit de la peinture d'intérieur et d'extérieur à partir de deux produits de base M1 et M2.

**Données**

| | Quantité utilisée par tonne | | Quantité disponible par jour |
|---|---|---|---|
| | Extérieure | Intérieure | |
| M1 | 6 | 4 | 24 |
| M2 | 1 | 2 | 6 |
| Profit par tonne | 5 | 4 | |

**Contraintes supplémentaires**

– Demande maximum en peinture d'intérieur : 2 tonnes / jour.
– La production en peinture d'intérieur ne dépasser que d'une tonne celle d'extérieur.

## III. La méthode du simplexe :

La méthodologie proposée pour cette technique consiste à visiter tous les états possibles dans un système en partant d'un sommet vers un sommet adjacent de manière à réviser et améliorer la fonction objective [2]. Pour ce faire, nous procédons à l'exploration des différentes démarches selon l'ordre de priorité donné ci-dessous :

**Démarche 1 :** On démarre l'application de l'approche (méthode du simplexe) par la transformation des contraintes d'inégalité en contraintes d'égalité en ajoutant les variables d'écart.

**Démarche 2 :** Dans un second temps nous sélectionnons les variables originales comme variables hors-base et les variables d'écarts comme variable basique. puis nous effectuerons une permutation entre une variable hors-base de notre choix qui sera remplacée par une variable de base (entrante). Le choix de la variable entrante repose sur la variable dont le coefficient est le plus élevé dans la fonction objective.

**Démarche 3 :** La variable sortante est la première à s'annuler. On répète le processus jusqu'à ce que tous les coefficients de fonction objective soient négatifs ou nuls. Dans ce cas, on arrête et la solution optimale est trouvée.

**Exemple** : Considérons le problème suivant :

Max F = 40 y1 + 60 y2     3.1

Sous les contraintes

$2y1 + y2 \leq 70$     3.2

$y1 + y2 \leq 40$     3.3

$y1 + 3y2 \leq 90$     3.4

$y1 \geq 0$

$y2 \geq 0$

**Solution :**

Soit y3, y4 et y5, les variables d'écart relatif respectivement aux contraintes 3.2, 3.3 et 3.4, qui permettent de transformer les contraintes d'inégalité pour obtenir les égalités suivantes :

$2y1 + y2 + y3 = 70$     3.5

$y1 + y2 + y4 = 40$     3.6

$y1 + 3y2 + y5 = 90$     3.7

On commence à partir du point y1=0, y2=0 et on vérifie s'il est possible d'augmenter la fonction objective sachant que notre système est un système de trois équations avec cinq inconnues. Pour trouver la solution on impose deux des variables à zéro et on déduit les valeurs des trois variables restantes. Ce qui donne :

$y1 = 0, y2 =$
$0$ ;
Pour les variables de base et :
$y3 = 70$
$y4 = 40$
$y5 = 90$
Pour les variables hors base.
Et la valeur de la fonction objective $F = 40\ y1 + 60\ y2 = 0$, La question
qui se pose est : Est-il possible d'augmenter F ?
Pour répondre à cette question, on vérifie la fonction objective. Si les coefficients sont positifs, ce qui est le cas, on a donc la possibilité d'augmenter F. On favorise l'augmentation de la variable qui a le plus grand coefficient dans la fonction objective.

Donc on choisit y2 en maintenant $y1 = 0$.

On remplace $y1=0$ dans équation 3.5, 3.6, et 3.7

| | | | |
|---|---|---|---|
| **2y1** + y2 + y3 = 70 | → | y2 + y3 = 70 | 3.8 |
| **y1** + y2 + y4 = 40 | → | y2 + y4 = 40 | 3.9 |
| **y1** + 3y2 + y5 = 90 | → | 3y2 + y5 = 90 | 3.10 |

On réécrit les variables restantes en fonction d'y2 pour chaque équation. Ce qui nous donne : $y3 = 70 - y2 \geq 0$

$\rightarrow y2 \leq 70$ 3.11

$y4 = 40 - y2 \geq 0 \rightarrow y2 \leq 40$ 3.12

$y5 = 90 - 3y2 \geq 0 \rightarrow y2 \leq 30$ 3.13

On peut augmenter la valeur de y2 au maximum à 30 puisque c'est la valeur qui permet de satisfaire les équations 3.11, 3.12 et 3.13.
Nouvelle solution admissible :
$y1 = 0$ ; $y2 = 30$
; $y3 = 40$ ; $y4 =$
$10$ ; $y5 = 0$

On remarque bien qu'y1=0 est toujours maintenu à zéro. Par ailleurs, y5 est passé de 90 à 0 et y2 est passé de 0 à 30 .Donc on déduit que pour cette étape la variable sortante de la base est y2 et la variable entrante à la base est y5.
La nouvelle valeur de la fonction objective est : $F = 40\ y1 + 60y2 = 1800$,

Les coefficients de la fonction objective sont positifs. Ce qui permet améliorer d'avantage la solution obtenue. Dans ce cas on s'arrange pour exprimer toutes les équations en fonction de y1 et y5. Pour ce faire on retire la valeur de y2 de la troisième équation et on remplace son expression dans les deux premières. Il en est de même pour F, ce qui nous donne dans cette étape les équations 3.14, 3.15 et 3.16

$$5/3\ y1 + y3 - 1/3\ y5 = 40 \tag{3.14}$$

$$2/3\ y1 + y4 - 1/3\ y5 = 10 \tag{3.15}$$

$$1/3\ y1 + y2 + 1/3\ y5 = 30 \tag{3.16}$$

Nouvelle expression de la fonction objective F = 1800 + 20y1 - 20y5,

D'après l'expression de la fonction objective de l'étape précédente on choisit d'augmenter y1 en gardant y5 =

$$5/3\ y1 + y3 - \mathbf{1/3\ y5} = 40 \qquad y3 = 40 - 5/3\ y1 \geq 0 \rightarrow y1 \leq 24 \tag{3.16}$$

$$2/3\ y1 + y4 - \mathbf{1/3\ y5} = 10 \qquad y4 = 10 - 2/3\ y1 \geq 0 \rightarrow y1 \leq 15 \tag{3.17}$$

$$1/3\ y1 + y2 + \mathbf{1/3\ y5} = 30 \qquad y5 = 30 - 1/3\ y1 \geq 0 \rightarrow y1 \leq 90 \tag{3.18}$$

La valeur maximale admissible pour y1 = 15. Nouvelle solution admissible :

y1 = 15;

y2 = 25;

y3 = 15;

y4 = 0;

y5 = 0

Pour cette étape, la variable sortante de la base est y2 et la variable entrante à la base est y4. L'expression de la nouvelle valeur du critère est : F = 2100 - 30y4 - 10y5,

Les coefficients de y4 et y5 dans la fonction objective de la dernière étape sont négatifs, quelle que soient la valeur de y4 et y5 qui implique une diminution de la valeur du critère F. Il n'existe donc plus d'augmentation et d'amélioration possible et la dernière solution obtenue représente la solution optimale qui est égale à 2100.

## 3.1 Résolution par Tableaux de simplexe

On commence par la transformation des contraintes d'inégalité en contraintes d'égalité en introduisant des variables d'écart.

**Étape 1**

On construit un tableau à deux dimensions rxs où le nombre de colonnes r est égal au nombre de variables (les variables de décision plus les variables d'écart) dans le système plus une colonne de solution. Le nombre de lignes est égal au nombre d'équation dans le système sans considération des contraintes de positivité [2].

A la première itération, on sélectionne la variable qui a le coefficient le **plus élevé** dans la ligne objective (fonction objective). On encadre la colonne de la variable entrante que l'on appelle " **colonne pivot".**

**Étape 2**

On calcule le minimum du rapport du coefficient du membre de droite de chaque contrainte sur le coefficient correspondant à la colonne. Dans le cas ou le coefficient dans la colonne entrante est négatif ou infini, on ne le compte pas dans le calcul du minimum.
On encadre alors la ligne ou le minimum se produit. Cette ligne reçoit le nom de "**ligne**

**pivot".**
Le coefficient qui se trouve à l'intersection de la colonne pivot et de la ligne pivot est appelé
**"élément pivot".**

### Étape 3

On reconstruit le tableau du simplexe (il faut conserver la même dimension du tableau).
On commence, d'abord, par construire la nouvelle ligne pivot qui se calcule de la manière suivante :

$$Nouvelle\ ligne\ pivot = \frac{Ancienne\ ligne\ pivot}{Element\ pivot}$$

Puis, on calcule les autres lignes par la formule suivante :

**Toutes les autres lignes y compris z= (Ligne actuelle) - (l'élément de sa colonne pivot) × (nouvelle ligne pivot)**

**Exemple :** Une entreprise disposant de 10 000 m2 de planches de bois en réserve fabrique et commercialise 2 types de boîtes en bois. La fabrication d'une boîte en bois de type 1 et de type 2 requiert, respectivement, 1 et 2 m2 de bois ainsi que 2 et 3 minutes de temps d'assemblage. Seules 200 heures de travail sont disponibles pendant la semaine à venir. Les boites sont clouées et, il faut quatre fois plus de clou pour une boîte du second type que pour une du premier. Le stock de clous disponible permet d'assembler au maximum 15 000 boîtes du premier type. Les boites sont vendues, respectivement, 3 et 5 du chiffre d'affaires [2].

**Formulation du problème :**

$x1$ : quantité de boîtes en bois de type 1 $x2$ :

quantité de boîtes en bois de type 2 Max z = 3$x$1 +

5$x$2

Sous les contraintes :

$x1 + 2x2 \leq 10\ 000$

$2x1 + 3x2 \leq 12\ 000$

$x1+4x2 \leq 15\ 000$ 3.19

$x1 \geq 0$ et $x2 \geq 0$

Après avoir mis le programme linéaire sous sa forme standard Max z = 3*x1* +
5 *x2*

S.c. $x1 + 2x2 + x3 = 10\ 000$

$2x1 + 3x2 + x4 = 12\ 000$ 3.20

$x1 + 4x2 + x5 = 15\ 000$

$x1 \geq 0$ et $x2 \geq 0$

On peut réécrire le programme linéaire en fonction de toutes les variables du système (variables de décision et variables d'écarts)

Max z = $3x1 + 5\ x2 + 0\ x3 + 0\ x4 + 0\ x5$

$x1 + 2x2 + x3 + 0\ x4 + 0x5 = 10\ 000$

$2x1 + 3x2 + 0\ x3 + x4 + 0\ x5 = 12\ 000$ 2.30

$x1 + 4x2 + 0\ x3 + 0\ x4 + x5 = 15\ 000$

$x1 \geq 0$ , $x2 \geq 0$, $x3 \geq 0$ , $x4 \geq 0$ et $x5 \geq 0$

Puis, nous faisons juste une translation vers un tableau comme suit :

**Itération 1**

| | z | x1 | x2 | x3 | x4 | x5 | sol |
|---|---|---|---|---|---|---|---|
| z | -1 | 3 | 5 | 0 | 0 | 0 | 0 |
| x3 | 0 | 1 | 2 | 1 | 0 | 0 | 10000 |
| x4 | 0 | 2 | 3 | 0 | 1 | 0 | 12000 |
| x5 | 0 | 1 | 4 | 0 | 0 | 1 | 15000 |

Dans ce tableau, nous allons définir la colonne pivot, la ligne pivot et l'élément pivot.
Nous commençons par la colonne pivot. Pour ce faire, nous sélectionnerons la variable qui a le coefficient le **plus élevé** dans la ligne de la fonction objective qui est le 5 de la variable x2. Donc la colonne pivot est la colonne du x2.
Ensuite, nous définirons la ligne pivot en divisant chaque solution par les éléments correspondants dans la colonne pivot.

10000/2 = 5000

12000/3 = 4000

15000/4 = 3750

On choisit la plus petite valeur. Dans ce cas c'est la valeur de : 3750. Ce qui correspond à la dernière ligne du tableau qui sera la ligne pivot.

## Itération 1

Colonne pivot

| | z | x1 | x2 | x3 | x4 | x5 | sol | |
|---|---|---|---|---|---|---|---|---|
| z | -1 | 3 | 5 | 0 | 0 | 0 | 0 | |
| x3 | 0 | 1 | 2 | 1 | 0 | 0 | 10000 | 5000 |
| x4 | 0 | 2 | 3 | 0 | 1 | 0 | 12000 | 4000 |
| x5 | 0 | 1 | 4 | 0 | 0 | 1 | 15000 | 3750 |

Ligne Pivot

Élément Pivot

L'intersection entre la colonne pivot et la ligne pivot, nous donne l'élément pivot qui est égal à 4

**Itération 2**

Une fois l'élément pivot déterminé, on calcule la nouvelle ligne pivot en divisant l'ancienne ligne par l'élément pivot, ce qui donne :

$$\frac{0}{4} \quad \frac{1}{4} \quad \frac{4}{4} \quad \frac{0}{4} \quad \frac{0}{4} \quad \frac{1}{4} \quad \frac{15000}{4}$$

$$0 \quad 0.25 \quad 1 \quad 0 \quad 0 \quad 0.25 \quad 3750$$

On la place dans le nouveau tableau en conservant la même position que dans le premier tableau sauf que la variable dans la colonne pivot prend la place de la variable de la ligne pivot comme suit :

| | z | x1 | x2 | x3 | x4 | x5 | sol |
|---|---|---|---|---|---|---|---|
| z | | | | | | | |
| x3 | | | | | | | |
| x4 | | | | | | | |
| **x2** | 0 | 0,25 | 1 | 0 | 0 | 0,25 | 3750 |

Ensuite, on remplit toutes les autres lignes par la formule suivante :

**(Ligne actuelle)-(l'élément de sa colonne pivot) × (nouvelle ligne pivot)**

On applique le principe sur la ligne 3 (variable x4)

**Ligne actuelle** $\quad 0 \quad 2 \quad 3 \quad 0 \quad 1 \quad 0 \quad 12000$

**L'élément de sa colonne pivot :** 3

**Nouvelle ligne pivot :** $\quad 0 \quad 0.25 \quad 1 \quad 0 \quad 0 \quad 0.25 \quad 3750$

**Résultats obtenus pour cette ligne :**

$0 \quad 1.25 \quad 0 \quad 0 \quad 1 \quad -0.75 \quad 750$

Et de la même manière, on applique le même calcule pour les lignes de x3 et de z. On les pace dans les zones appropriées dans le tableau de la deuxième itération.

| | z | x1 | x2 | x3 | x4 | x5 | sol | |
|---|---|---|---|---|---|---|---|---|
| z | -1 | 1,75 | 0 | 0 | 0 | -1,25 | -18750 | |
| x3 | 0 | 0,5 | 0 | 1 | 0 | -0,5 | 2500 | 5000 |
| x4 | 0 | 1,25 | 0 | 0 | 1 | -0,75 | 750 | 600 |
| x2 | 0 | 0,25 | 1 | 0 | 0 | 0,25 | 3750 | 15000 |

À la fin de cette itération, on vérifie tous les coefficients de ligne z. S'ils sont négatifs ou nuls on s'arrête. On a trouvé la solution optimale. Ce n'est pas le cas ici, la valeur de x1 est positive. On construit donc une nouvelle itération et un nouveau tableau de la même manière que pour l'itération2.

### Itération 3

| | z | x1 | x2 | x3 | x4 | x5 | sol |
|---|---|---|---|---|---|---|---|
| z | -1 | 0 | 0 | 0 | -1,4 | -0,2 | -19800 |
| x3 | 0 | 0 | 0 | 1 | -0,4 | -0,2 | 2200 |
| x1 | 0 | 1 | 0 | 0 | 0,8 | -0,6 | 600 |
| x2 | 0 | 0 | 1 | 0 | -0,2 | 0,4 | 3600 |

D'après le tableau de l'itération 3 on remarque que les coefficients de la ligne z, sont négatifs ou nuls. Donc on s'arrête. On a trouvé la solution optimale avec :

$x1 = 600$

$x2 = 3600$

$z = 19800$

# TD N° 3

## Algorithme des tableaux de simplexe

Dans les exercices suivants, appliquer l'algorithme tableaux de simplexe pour trouver la solution optimale de programmation linéaire.

**Exercice 1 :** Soit le problème d'optimisation suivant : Maximiser $z = 240\ x1$

$+160\ x2$

Sous $1\ x1 + 2\ x2 \leq 150$

$4\ x1 + 2\ x2 \leq 400$

avec $x1\ ;\ x2 \geq 0$

**Exercice 2 :** Soit le problème d'optimisation suivant : Maximiser $z = 3x1 +$

$5x2$

Sous $x1 + 2\ x2 \leq 10000$ $2x1 + 3$

$x2 \leq 12000$

$x1 + 4\ x2 \leq 15000$

avec $x1\ ;\ x2 \geq 0$

**Exercice 3 :** Soit le problème d'optimisation suivant : Maximiser $z = 3x1 +$

$4x2 + 2x3$

Sous $x1 + x2 - x3 \leq 10$ $x1 - 2x2$

$+ 3x3 \leq 14$ avec $x1; x2; x3 \geq 0$

**Exercice 4 :** On veut résoudre le problème d'optimisation suivant : Maximiser $z = 4x1 + 6x2$

$+ 3x3$

Sous $x1 + 6x2 + 2x3 \leq 24$ $x1 + 3x2 -$

$2x3 \leq 9$

$x1; x2; x3 \geq 0$