Jacques BAIR
Roland HINNION
Daniel JUSTENS

S. B. P. M.

Applications économiques au service de la mathématique

Société Belge des Professeurs de
Mathématique d'expression française

1989

Jacques BAIR
Université de Liège
Roland HINNION
Université Libre de Bruxelles
Daniel JUSTENS
Institut Cooremans

S. B. P. M.

Applications économiques au service de la mathématique

Société Belge des Professeurs de
Mathématique d'expression française

1989

Avant-propos

La Société Belge des Professeurs de Mathématique d'expression française a un rôle primordial : l'amélioration de l'enseignement de la mathématique. Cela l'a amenée à publier elle-même de nombreux ouvrages. Qu'il me suffise de rappeler les documents de la série *RENOVER*, « Les nombres transcendants » par C. Radoux et les deux tomes des « Olympiades mathématiques belges ».

Ce qui caractérisera l'enseignement de notre discipline pendant la seconde moitié du vingtième siècle n'est peut-être pas le flux des réformes et des contre-réformes qui se succèdent depuis bientôt trente ans, mais plutôt le fait que la mathématique, réservée auparavant aux physiciens et aux ingénieurs, est maintenant un outil essentiel pour les chimistes, biologistes, psychologues, agronomes, économistes,...

D'autre part, le nombre d'étudiants fréquentant les Facultés de sciences économiques et les Instituts supérieurs de commerce, a connu une progression spectaculaire. Il était donc opportun de publier un recueil qui montre différents aspects des applications de la mathématique à l'économie.

L'ouvrage comporte cinq chapitres. Les trois premiers sont l'œuvre de Jacques Bair, qui enseigne à l'Université de Liège. La matière traitée reprend pour l'essentiel les conférences faites par l'auteur aux congrès de la SBPMef et à d'autres tribunes.

Daniel Justens est spécialiste de la mathématique financière. Il enseigne à l'Institut d'Enseignement Supérieur Lucien Cooremans et a rédigé le chapitre IV.

Enfin, Roland Hinnion se consacre particulièrement à l'étude des points fixes. Il enseigne à l'Université Libre de Bruxelles et est l'auteur du dernier chapitre.

Si, à l'aide de cet ouvrage, les professeurs peuvent se familiariser avec de nouvelles applications de leur discipline et si, en outre, ils y trouvent matière à parfaire la préparation de leurs élèves à l'Enseignement Supérieur, le travail n'aura pas été inutile.

J. Wilmet

Président de la SBPMef

©SBPMef 1989

Table des matières

Chapitre 0

Historique sommaire des rapports entre Mathématiques et Economie.

par *Jacques BAIR*

L'utilisation des mathématiques au service d'autres disciplines évoque en premier lieu la longue et profonde coopération entre les mathématiques et les sciences physiques prises au sens large, incluant notamment les sciences de la nature et l'art de l'ingénieur.

Les rapports entre les mathématiques et les sciences humaines sont beaucoup plus récents et aussi moins bien établis. Cela résulte de la plus grande diversité (donc complexité) des situations rencontrées : par exemple, il est plus difficile de comprendre rigoureusement le comportement d'un groupe d'individus aux intérêts souvent divergents, plutôt que le mouvement d'un corps soumis à des forces bien connues. Le nombre de paramètres dont il faut tenir compte dans les sciences humaines est généralement très élevé, puisque les facteurs d'influence y sont fort nombreux ; cela oblige le praticien à manipuler un nombre considérable de données, souvent groupées en des tableaux imposants. Grâce notamment au développement du calcul matriciel, qui analyse précisément des tableaux de nombres, et à l'apparition des ordinateurs, qui permettent d'effectuer des calculs sur un nombre considérable de grandeurs, les mathématiques ont vraiment pu être exploitées dans les sciences humaines. Il est en effet intéressant de constater que les méthodes mathématiques ont fait leur percée en économie à l'époque où pratiquement naissait le calcul matriciel, sous l'impulsion des mathématiciens anglais A. Cayley (1821-1895) et J. Sylvester (1814-1897) [19]. Toutefois, ce n'est que depuis les progrès fulgurants de l'informatique au cours de ce siècle, que les mathématiques jouent un rôle fondamental dans les sciences humaines.

Reprenons l'histoire de l'économie mathématique à ses débuts. Les premiers théoriciens de l'économie politique moderne, avec à leur tête l'anglais D. Ricardo (1772-1823), n'employaient aucune technique scientifique ; ils faisaient, selon le mot de J.R. Hicks (prix Nobel d'Economie en 1972), des *« mathématiques dans la coulisse »*, c'est-à-dire sans le savoir. En effet, les lois économiques classiques reposent sur l'hypothèse de l'homme rationnel (ou *homo œconomicus*), à la recherche du plus grand gain et de la satisfaction maximale ; les mathématiciens ne pouvaient rester indifférents à cette conception de l'économie, puisqu'ils y reconnaissaient des problèmes qui leur étaient tout naturellement destinés : par exemple, la recherche du maximum des fonctions d'utilité d'un consommateur et du profit d'un entrepreneur.

Le philosophe et mathématicien français A. Cournot (1801-1877) apparaît aujourd'hui comme le véritable fondateur de l'économie mathématique. Ses mérites furent toutefois reconnus de façon posthume.

Les mathématiques entrèrent en économie, de façon définitive, à la suite de deux faits bien précis.

Au dix-neuvième siècle, certains économistes découvraient l'importance de la notion de fonction ; par exemple, l'italien V. Pareto (1848-1923) déclarait : *« les économistes littéraires perdent leur temps à chercher des relations de cause à effet là où il n'y a que des relations fonctionnelles réversibles entre des données qui se conditionnent mutuellement »*. Pratiquement au même moment, naissait, en 1871, le marginalisme, développé séparément par le français (qui a fait école à Lausanne) L. Walras (1834-1910) et l'australo-anglais S. Jevons (1835-1882) ; cette théorie attirait l'attention sur l'importance de la dernière unité dans l'influence de chaque variable et ouvrait de ce fait la porte à l'utilisation du calcul infinitésimal [3].

A cette époque, les mathématiques étaient principalement utilisées pour rendre l'économie politique moins littéraire, c'est-à-dire plus précise. On se contentait alors de faire appel à l'outillage existant, avec quelques innovations, comme l'introduction de la notion d'élasticité d'une fonction en remplacement de sa dérivée. Au cours du vingtième siècle, les économistes se rendent compte de la puissance de la démarche scientifique et sont même amenés à créer des théories mathématiques nouvelles pour résoudre leurs problèmes spécifiques [35]. Principalement depuis la fin de la deuxième guerre mondiale, les méthodes mathématiques sont devenues prépondérantes en économie [40].

En conséquence, les anciennes polémiques entre économistes « scientifiques » et « littéraires » ont pratiquement disparu [52]. La situation actuelle sur ce point est fort bien résumée par B. Russell (1872-1970), un des grands esprits de ce siècle, à la fois mathématicien, philosophe, psychologue et sociologue : *« Beaucoup de gens éprouvent une haine passionnée pour l'abstraction et cela, si nous ne nous trompons, à cause de sa difficulté intellectuelle. Mais comme ils ne veulent pas donner cette raison, il en inventent toutes sortes d'autres auxquelles ils attachent une grande importance et qu'ils croient décisives. Ils disent que la réalité est concrète et qu'en faisant des abstractions on laisse échapper tout l'essentiel. Ils disent encore que toute abstraction est falsification et qu'en omettant un des aspects du réel on court le risque de s'enfoncer dans le faux en jugeant seulement d'après les aspects qui restent. Ceux qui raisonnent ainsi demeurent tout à fait en dehors de la science.* [1]

Les économistes modernes utilisent de plus en plus des théories mathématiques fort variées, comme l'analyse, la topologie, l'algèbre, le calcul des probabilités, la statistique, etc. Au surplus, leurs recherches ont provoqué le développement de certaines disciplines purement mathématiques ; par exemple, la théorie de l'optimisation a fait, en ce vingtième siècle, des progrès considérables, tandis que la recherche opérationnelle connaît, depuis ces dernières décennies, un essor fulgurant, notamment en programmation mathématique (qu'elle soit linéraire, quadratique, convexe, quasi-convexe, multicritère ou dynamique), ou encore dans les théories des graphes et des jeux.

On peut distinguer actuellement plusieurs catégories d'applications des mathématiques à l'économie. Tout d'abord, pour résoudre des problèmes concrets posés par le commerce, les finances, l'industrie ou l'agriculture, les praticiens ont recours à l'arithmétique commerciale, à l'algèbre financière, ou encore à la recherche opérationnelle. Les mathématiques peuvent également être utilisées comme outil de logique formelle, par exemple pour établir une théorie des prix en microéconomie ou pour étudier les fluctuations dans un système macroéconomique. On les exploite également comme un moyen d'investigation statistique : par exemple, certaines fonctions abstraites, comme la célèbre courbe des revenus de PARETO, peuvent être vérifiées par des techniques statistiques. La mise en œuvre simultanée de ces trois premières applications a donné naissance à l'économétrie, qui est apparue, depuis 1930, comme une discipline autonome, recourant à la théorie économique, à la formulation mathématique et à l'analyse statistique [42]. En 1969, le premier prix Nobel d'Economie fut décerné conjointement au norvégien R. FRISCH et au néerlandais J. TINBERGEN, qui peuvent être tous deux considérés comme des fondateurs de l'économétrie. Par la suite, cette distinction suprême fut attribuée à de nombreux autres économètres ou à des économistes-mathématiciens : P. SAMUELSON (1970), célèbre pour ses travaux qui mettent en œuvre économie et mathématique, ainsi que pour ses fameux écrits sur l'économique ; K. ARROW (1972), qui généralisa notamment la théorie de KUHN-TUCKER sur la programmation convexe au cas quasi-convexe ; W. LEONTIEF (1973), qui se consacra principalement aux relations interindustrielles grâce à l'analyse input-output ; L. KANTOROVITCH (1975), un des pionniers de la programmation linéaire ; L. KLEIN (1980), auteur de contributions importantes à l'économétrie ; G. DEBREU (1983), mathématicien de formation, qui est l'auteur d'une théorie axiomatique de l'équilibre économique.

Comme les mathématiques prennent de plus en plus d'importance dans les sciences humaines, il est fort probable qu'à l'avenir les sciences économiques et sociales seront une source d'inspiration pour la création de nouvelles disciplines mathématiques, tout comme la physique l'a été par le passé [39]. Il convient toutefois de se souvenir que les mathématiques sont ici un instrument et non une fin en soi et que, pour influencer les responsables chargés d'élaborer les politiques, les idées doivent être traduites des équations en langage courant [48] ; citons à ce propos l'économiste français M. ALLAIS : « *certes, l'outil mathématique est indispensable pour analyser et comprendre les phénomènes économiques, mais le travail essentiel, ce n'est pas le maniement de cet outil logique, c'est le choix et la discussion des prémisses qui doivent être fondés sur l'observation des faits. L'erreur des adversaires de l'emploi de l'outil mathématique en économique, c'est de ne pas connaître les vastes possibilités qu'il offre ; l'erreur de certains mathématiciens, c'est parfois de prendre pour un but, ce qui n'est et ne peut être qu'un moyen.* » [1]

Chapitre 1

Aperçu de l'usage du calcul matriciel dans les sciences humaines

par *Jacques BAIR*

L'algèbre linéaire est un outil mathématique très puissant qui peut être appréhendé de plusieurs façons. Les mathématiciens représentent souvent les applications linéaires par des matrices et déduisent les règles du calcul matriciel des propriétés relatives aux applications linéaires. Pour un praticien, cette vision des choses n'est guère parlante : il doit en effet manipuler des tableaux de nombres qui sont soit les résultats d'observations empiriques, soit des tableaux obtenus à partir de données par des règles bien précises.

Nous nous proposons de montrer comment certaines de ces règles, parmi les plus importantes, peuvent être illustrées simplement par des problèmes concrets et facilement accessibles.

Nous traiterons sommairement quelques sujets classiques. Il s'agit de l'analyse input-output pour les tableaux d'échanges interindustriels, d'une introduction à l'étude des chaînes de Markov (mathématicien russe, 1856-1922), ainsi que de nouvelles disciplines en recherche opérationnelle, surtout la théorie des graphes (mise en évidence notamment par König, 1936, et l'école hongroise), la théorie des jeux (exposée dans un traité de von Neumann et Morgenstern, 1944) et la programmation linéaire (conçue initialement par Kantorovitch, 1940, et rendue célèbre par la méthode du simplexe due à Dantzig, 1949).

1.1 Quelques matrices usuelles

1.1.1 Matrices de commandes

Supposons, par exemple, que trois clients puissent acheter quatre produits. Pour fixer les idées, considérons une commande effectuée antérieurement, soit

$$C = \begin{pmatrix} 5 & 2 & 4 & 1 \\ 3 & 0 & 2 & 3 \\ 2 & 1 & 5 & 0 \end{pmatrix},$$

les lignes étant relatives aux personnes et les colonnes se rapportant aux biens.

Pour doubler la commande, il suffit de considérer la matrice $2C$. Plus généralement, on peut introduire aisément l'opération de multiplication d'une matrice par un scalaire.

De la même manière, l'addition de deux commandes, résumées par les matrices C_1 et C_2, est évidemment donnée par la somme C_1+C_2. L'addition de deux matrices apparaît dès lors comme une opération tout à fait naturelle.

1.1.2 Matrices de prix

Penchons-nous à présent sur les prix unitaires de ces quatre produits : ils peuvent être rassemblés dans une nouvelle matrice dont les lignes concernent les produits, la première colonne les prix unitaires d'achat et la seconde colonne les frais unitaires de transport ; à titre d'exemple soit

$$P = \begin{pmatrix} 4 & 0,2 \\ 2 & 0,1 \\ 4 & 0,3 \\ 5 & 0,1 \end{pmatrix}$$

la matrice des prix unitaires pour les quatre articles considérés. Il est aisé de constater que les factures globales à payer par les clients pour l'achat et le transport des biens commandés à l'aide de la matrice C sont réunies dans la matrice

$$F = CP = \begin{pmatrix} 45 & 2,5 \\ 35 & 1,5 \\ 30 & 2 \end{pmatrix},$$

tandis que les sommes totales à payer par chacun des trois clients sont données par le vecteur-colonne

$$T = F \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 47,5 \\ 36,5 \\ 32 \end{pmatrix}.$$

Ce vecteur T peut aussi être obtenu en multipliant la matrice C par le vecteur Q donnant, pour chaque produit, le prix unitaire total à payer (soit la somme du prix unitaire d'achat et du prix unitaire de transport) :

$$T = CQ, \quad \text{avec} \quad Q = P \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4,2 \\ 2,1 \\ 4,3 \\ 5,1 \end{pmatrix};$$

cet exemple illustre bien la règle d'associativité de la multiplication matricielle.

Ainsi, on constate que la multiplication de la matrice C par le vecteur Q agit comme une application d'un espace à 4 dimensions sur un espace à 3 dimensions ; cela signifie concrètement que les 3 comptes peuvent être obtenus à partir des 4 prix unitaires.

1.2 Matrices et théorie des graphes

1.2.1 Généralités

Les rapports entre les individus d'un univers sont souvent d'ordre qualitatif et non quantitatif ; c'est pourquoi, les mathématiques n'ont guère été utilisées dans les sciences humaines avant ce siècle. Depuis quelques années, une discipline mathématique, à savoir la théorie des graphes, permet d'étudier rigoureusement de telles relations qualitatives et fournit la solution à de nombreux problèmes concrets (notamment de gestion).

Par définition, un *graphe* $G = (S, V)$ est composé d'un ensemble fini S qui est l'univers donné, dont les éléments $X_1, X_2, \ldots, X_n$ sont appelés les *sommets* et figurés par des points, ainsi que d'un ensemble V dont les éléments sont des couples de deux sommets : la mise en relation d'un sommet X_i avec un autre X_j se nomme un *arc* de X_i vers X_j, se note (X_i, X_j) et est représenté par un trait, rectiligne ou incurvé, reliant X_i à X_j (sans passer par un autre sommet) et muni d'une flèche pointée vers X_j.

Mentionnons quelques exemples classiques et utiles.

- Le graphe de la confiance (ou de la domination) dont les sommets sont des personnes, un arc reliant deux individus si le premier a confiance en (ou domine) le second.
- Le graphe des sens de circulation, dont les sommets sont des nœuds de communications (des carrefours), les arcs indiquant la possibilité de se déplacer d'un sommet à un autre de façon directe (c'est-à-dire sans passer par un troisième sommet).
- Le graphe des affectations dont les sommets sont soit des emplois vacants, soit des candidats à certains de ces emplois, tout arc traduisant la candidature d'une personne à un poste (le problème consistant bien entendu à satisfaire le plus grand nombre possible de candidats, éventuellement au moindre coût).
- Le graphe des ordonnancements, rendu célèbre par la méthode PERT, qui est appliqué dans l'étude de tout projet complexe de construction ou de développement pour coordonner au mieux le travail au cours du temps : les sommets sont les états du projet au début et au terme de diverses tâches, tandis que les arcs sont précisément les opérations à effectuer.

1.2.2 Matrice booléenne associée à un graphe

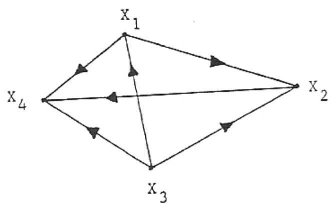
La situation figurée par un graphe peut aussi être décrite par un tableau carré de nombres, appelé la matrice M associée au graphe. Cette matrice offre d'autres ressources de manipulation que le graphe lui-même et se prête même à l'application du calcul matriciel classique. Les éléments a_{ij} de M sont affectés de deux indices entiers, dont le premier sert au repérage de la ligne, le second celui de la colonne ; chaque indice varie de 1 à n, où n désigne le nombre de sommets du graphe. On convient que a_{ij} vaut 1 si le graphe contient l'arc (X_i, X_j), 0 dans le cas contraire ; comme les éléments a_{ij} valent uniquement 0 et 1, la matrice M est souvent qualifiée de *booléenne* ou binaire.

1.2.3 Dénombrement des chemins de longueur donnée

Dans un graphe G, un *chemin* $(X_i, X_l, X_m, , \dots, X_s, X_j)$ reliant le départ X_i à l'arrivée X_j est une succession d'arcs $(X_i, X_l), (X_l, X_m), \dots, (X_s, X_j)$ tels que X_i soit le départ du premier, X_j soit l'arrivée du dernier, et l'arrivée de chaque arc autre que le dernier soit le départ du suivant. Le nombre d'arcs qui composent un chemin est la *longueur* de ce chemin.

Considérons l'élément b_{ij} de la matrice M^2, où M désigne la matrice booléenne associée à G ; il vaut $b_{ij} = \sum_{k=1}^{n} a_{ik}a_{kj}$, n représentant le nombre de sommets. Comme $a_{ik}a_{kj} = 1$ si et seulement si $a_{ik} = a_{kj} = 1$, on peut déclarer que $a_{ik}a_{kj} = 1$ si et seulement s'il existe un chemin (X_i, X_k, X_j), de longueur 2, reliant X_i à X_j (en passant par X_k). En conséquence, la valeur de b_{ij} dans M^2 est le nombre de chemins de longueur 2 de X_i à X_j. De façon plus générale, la valeur de p_{ij} dans $M^p = (p_{ij})$ est le nombre de chemins de longueur p joignant X_i à X_j.

Ces considérations peuvent être illustrées par cet exemple, facile à interpréter.



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

$$M^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ M^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ M^4 = 0.$$

1.3 Analyse input-output

1.3.1 Position du problème

Considérons une économie qui comporte n secteurs S_i produisant chacun un seul type de bien. Désignons par x_{ij} la quantité de ce bien livrée par le secteur S_i au vecteur S_j pour alimenter l'activité de celui-ci. Les nombres x_{ij}, parfois appelés *livraisons intermédiaires*, se rangent en une matrice dont les lignes correspondent aux productions (x_{ij} est l'*output* de S_i consommé par S_j) et les colonnes aux consommations (x_{ij} est l'*input* de S_j fourni par S_i) ; la matrice (x_{ij}) est souvent baptisée matrice d'input-output ou tableau d'échanges interindustriels.

De plus, les secteurs effectuent généralement des livraisons des produits finis dans le but de satisfaire une demande finale : soit d_i la valeur de la demande finale du secteur S_i. L'output du secteur S_i est dès lors donné par $x_i = \sum_{j=1}^{n} x_{ij} + d_i$, pour tout $i = 1, \ldots, n$.

Ainsi, la production totale du secteur S_i intervient pour satisfaire sa demande finale d_i, mais elle participe également aux autres livraisons finales via les livraisons intermédiaires. On est dès lors en présence d'un procédé cumulatif compliqué, qui peut être analysé par le calcul matriciel.

Des relations techniques lient la production du secteur S_j aux utilisations par S_j des différentes catégories de produit. En première approximation et à court terme, il est permis de supposer la grandeur x_{ij} proportionnelle à x_j soit $x_{ij} = a_{ij}x_j$. Cette hypothèse signifie que chaque secteur doit utiliser une proportion constante d'input pour la production de son output ; elle fait apparaître les constantes $a_{ij} = \frac{x_{ij}}{x_j}$, qui forment la matrice technologique $A = (a_{ij})$: a_{ij} représente la valeur de l'output du secteur S_i que S_j doit acquérir pour produire la quantité correspondant à une unité (monétaire) de son propre bien.

Dans ces conditions, le secteur S_i doit produire les outputs $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$ en vue d'alimenter tous les secteurs. Pour satisfaire en plus la demande finale, il faut que $x_i = \sum_{j=1}^{n} a_{ij}x_j + d_i$ (pour $i = 1, 2, \ldots, n$), soit matriciellement $X = AX + D$, ou encore $(I - A)X = D$, avec $X = (x_i)$,$A = (a_{ij})$ et $D = (d_i)$.

1.3.2 Solution globale

Le système $(I - A)X = D$ est résoluble et cramérien lorsque la matrice $I - A$ est régulière ; dans ce cas, l'unique solution est donnée par l'égalité

$$X = (I - A)^{-1}D.$$

Pour illustrer ce résultat, considérons une économie (fictive) de deux secteurs, répondant aux données suivantes (exprimées en une certaine unité monétaire) :

ventes / achats		secteurs utilisateurs: agriculture	secteurs utilisateurs: industrie	demandes finales	produit final
secteurs fournisseurs	agriculture	500	350	150	1000
	industrie	320	360	120	800

Le problème posé consiste à trouver le « vecteur output » de cette économie si la demande finale devient égale à 200 unités monétaires pour l'agriculture et 100 pour l'industrie. La matrice technologique est donnée par $A = \begin{pmatrix} \frac{1}{2} & \frac{7}{16} \\ \frac{8}{25} & \frac{9}{20} \end{pmatrix}$.

On vérifiera sans peine que $(I - A)^{-1}\begin{pmatrix} 150 \\ 200 \end{pmatrix} = \begin{pmatrix} 1000 \\ 800 \end{pmatrix}$, ce qui correspond bien aux données initiales. A présent, si $D = \begin{pmatrix} 200 \\ 100 \end{pmatrix}$, alors

$$X = (I - A)^{-1}\begin{pmatrix} 200 \\ 100 \end{pmatrix} = \frac{200}{27}\begin{pmatrix} \frac{11}{20} & \frac{7}{16} \\ \frac{8}{25} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 200 \\ 100 \end{pmatrix} = \begin{pmatrix} 1138,999 \\ 844,444 \end{pmatrix}.$$

1.3.3 Signification économique de la solution

La solution globale dans l'analyse input-output peut souvent être trouvée par l'étude spectrale (c'est-à-dire des valeurs et vecteurs propres) de la matrice technologique A. La matrice n-carrée A possède évidemment n valeurs propres (réelles ou complexes, comptées avec leur multiplicité). Dans la pratique, il est permis de supposer toutes les valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$ distinctes, quitte à modifier légèrement les données observées par plus de précision dans les mesures et à effectuer une approximation par le théorème de Bellman. De la sorte, A est diagonalisable : il existe une matrice régulière R telle que $R^{-1}AR = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. De plus, si du point de vue des entrées et des coûts, chaque secteur S_j consomme effectivement des facteurs primaires (travail, capital), on a $\sum_{i=1}^n a_{ij} < 1$ pour $j = 1, 2, \ldots, n$; il est alors aisé de vérifier que chaque valeur propre λ_j est, en module, inférieure à l'unité. Un calcul élémentaire montre que, pour tout entier p, $A^p = R\,\mathrm{diag}(\lambda_1^p, \lambda_2^p, \ldots, \lambda_n^p)R^{-1}$; un passage à la limite livre, avec des notations classiques, $\lim_{p\to+\infty} A^p = 0$ puisque $\lim_{p\to+\infty} \lambda_j^p = 0$ pour tout $j = 1, 2, \ldots, n$. Par ailleurs, l'égalité $I - A^p = (I-A)(I+A+A^2+\cdots+A^{p-1})$ entraîne

$$\begin{aligned}
\lim_{p\to+\infty}(I - A^p) &= I \\
&= \lim_{p\to+\infty}(I-A)(I+A+\cdots+A^{p-1}) \\
&= (I-A)\lim_{p\to+\infty}(I+A+\cdots+A^{p-1}) \\
&= (I-A)\sum_{k=0}^{\infty} A^k
\end{aligned}$$

En conclusion,

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

D'un point de vue économique, la formule $(I - A)^{-1} = I + A + A^2 + \cdots + A^p + \cdots$ admet une interprétation intéressante. De fait, pour satisfaire sa demande finale, chaque secteur S_i devra d'abord produire un output égal à d_i. Cette production exigera de tous les autres secteurs des outputs supplémentaires qui seront utilisés comme inputs ; ces quantités additionnelles dépendront des coefficients techniques et seront données par la matrice AD. Ces outputs AD vont, à leur tour, réclamer de nouveaux outputs qui seront exploités comme inputs, leur quantité provenant du produit matriciel A^2D. De même, les outputs A^2D vont exiger la production correspondant à A^3D, et ainsi de suite. Au total, pour satisfaire la demande finale, on aura besoin d'une production globale donnée par

$$D + AD + A^2D + A^3D + \cdots = (I + A + A^2 + A^3 + \cdots)D = \left(\sum_{k=0}^{\infty} A^k\right) D.$$

1.4 Quelques exemples simples de modèles linéaires

1.4.1 Equilibre du marché

Considérons tout d'abord le cas d'un marché isolé et déterminons le prix d'équilibre de l'unique bien en question. Le modèle comporte trois grandeurs : la demande D, l'offre S et le prix unitaire p ([1]) ; les quantités D et S sont exprimées, par exemple, en tonnes par mois et le prix p en francs. Nous supposerons que l'offre et la demande dépendent linéairement du prix, ce qui peut être considéré comme une bonne approximation dans un intervalle de temps limité. On a donc $D = ap + b$ et $S = cp + d$,où a,b,c et d sont des constantes ; généralement, la demande D est décroissante par rapport au prix, ce qui se traduit par le caractère négatif du coefficient angulaire a de la droite de demande, tandis que l'offre est une fonction croissante de p, d'où la pente c de la droite d'offre est positive.

Comme condition d'équilibre, nous allons adopter l'hypothèse classique selon laquelle l'offre est égale à la demande. Nous sommes donc en présence de trois équations : une condition d'équilibre $D = S$, et deux équations de comportement $D = ap + b$ et $S = cp + d$. Il existe alors un seul prix $\overline{p}$ vérifiant ces trois égalités, à savoir $\overline{p} = \frac{d-b}{a-c}$, la quantité écoulée étant donnée par $\overline{q} = \frac{ad-bc}{a-c}$, la réponse n'étant acceptable du point de vue économique que si le produit ad est inférieur à bc (car la quantité $\overline{q}$ doit être positive).

La discussion ci-dessus se rapporte à un marché isolé. Elle peut être étendue au cas d'un marché à n biens : pour chaque produit P_i, nous noterons D_i la demande, S_i l'offre et p_i le prix unitaire ; l'équilibre général sera atteint lorsque $D_i = S_i$ pour tout $i = 1, 2, \ldots, n$. En guise d'exemple (fictif), recherchons l'équilibre pour un marché à trois biens P_1,P_2 et P_3, dont la demande et l'offre sont données par les équations suivantes

$$\left\{\begin{array}{lcl} D_1 & = & 2 - p_1 + p_2 + p_3 \\ D_2 & = & 10 + p_1 - 2p_2 + p_3 \\ D_3 & = & 5 + p_1 + p_2 - p_3 \end{array}\right. \qquad \left\{\begin{array}{lcl} S_1 & = & -2 + 2p_1 \\ S_2 & = & -2 + p_2 \\ S_3 & = & -3 + 2p_3 \,. \end{array}\right.$$

Remarquons que l'offre S_i est une fonction linéaire croissante en l'unique variable p_i ; par contre, la demande D_i dépend linéairement de p_1,p_2 et p_3,

([1]) les symboles D (pour « demande ») et S (pour *« supply »* en anglais) sont quelquefois remplacés respectivement par q^d et q^o. Par ailleurs, le prix unitaire est parfois noté P au lieu de p.

le coefficient de p_i étant négatif (car la demande du produit P_i diminue si p_i augmente), tandis que le coefficient de p_j, pour $j \neq i$, est positif (suggérant par là que P_j peut être regardé comme un « substitut » de P_i). L'équilibre de ce modèle est atteint lorsque $3p_1 - p_2 - p_3 = 4$, $-p_1 + 3p_2 - p_3 = 12$ et $-p_1 - p_2 + 3p_3 = 8$. La solution de ce systéme cramérien est donnée par $\overline{p}_1 = 7$, $\overline{p}_2 = 9$, $\overline{p}_3 = 8$, avec $\overline{q}_1 = 12$, $\overline{q}_2 = 7$ et $\overline{q}_3 = 13$.

1.4.2 Ajustement linéaire

Dans la pratique, il est très fréquent de rechercher une liaison entre plusieurs grandeurs pour voir comment une variable y (qualifiée de *dépendante* ou *expliquée*) dépend de plusieurs autres $x_1, x_2, \ldots, x_n$ (dites *indépendantes* ou *explicatives*). Une situation favorable est celle où se dégage une relation linéaire du type $y = a + a_1x_1 + a_2x_2 + \cdots + a_nx_n$, le problème consistant alors à calculer les coefficients a et a_i $(i = 1, 2, \cdots, n)$ correspondant à des mesures expérimentales.

En guise d'application concrète, analysons le marché d'un produit P : sa quantité q commandée mensuellement dépend du prix unitaire p de ce bien, mais aussi des prix unitaires p_1 et p_2 de deux produits « concurrents » P_1 et P_2. Au cours des six premiers mois de l'année écoulée, les prix p, p_1 et p_2 ont été modifiés, ce qui a donné lieu aux résultats suivants (les prix étant exprimés en une certaine unité monétaire, et la quantité q, par exemple, en tonnes) :

mois	p	p_1	p_2	q
janvier	1	1	1	21
février	2	1	1	17
mars	1	2	1	24
avril	1	1	2	23
mai	2	3	3	27
juin	4	3	1	15

Nous allons voir que la quantité q dépend linéairement de p, p_1, p_2, et peut se mettre sous la forme $q = a + bp + cp_1 + dp_2$; nous allons également calculer les coefficients a, b, c, d qui correspondent à ces observations. Il s'agit de résoudre le système linéaire suivant :

$$\left\{ \begin{array}{rcl} a+b+c+d & = & 21 \\ a+2b+c+d & = & 17 \\ a+b+2c+d & = & 24 \\ a+b+c+2d & = & 23 \\ a+2b+3c+3d & = & 27 \\ a+4b+3c+d & = & 15 \end{array} \right.$$

Ce système (à six équations et quatre inconnues a, b, c et d) est résoluble et déterminé ; sa solution peut être obtenue par une méthode numérique (Crout, Banachiewicz, Gauss,...) : on trouve

$$\begin{array}{rclcrcl} a & = & 20 & \quad & c & = & 3 \\ b & = & -4 & & d & = & 2\,. \end{array}$$

D'où $q = 20 - 4p + 3p_1 + 2p_2$.

Signalons que, dans la réalité, le nombre d'observations dépasse généralement le nombre de paramètres à déterminer, ce qui conduit à la résolution d'un système qui renferme plus d'équations que d'inconnues. Souvent (et contrairement à l'exemple précédent qui est, à cet égard, « idéal »), un tel système n'est pas résoluble, mais admet une espèce de « solution qui est la meilleure possible » (au sens des moindres carrés) et qui peut être obtenue à l'aide de l'inverse généralisé de la matrice des coefficients.

1.4.3 Un modèle keynésien simple pour le revenu national

L'économiste anglais J.M.Keynes (1833-1946) préconisait de stimuler l'investissement public et d'augmenter l'intervention de l'Etat dans le but de maintenir l'activité à un niveau élevé de production et d'emploi.

Sous une forme très simple, le modèle keynésien du revenu national comporte deux équations. La première est une condition d'équilibre exprimant que le revenu national y est égal au total des dépenses, soit encore à la somme de la consommation c, de l'investissement i et de la dépense gouvernementale g. (2) La deuxième est une équation de comportement donnant c en fonction de y : nous supposerons que la consommation dépend linéairement du revenu national, ce qui se traduit par l'égalité $c = \bar{c} + ky$, où $\bar{c}$ et k sont des constantes positives, $\bar{c}$ représentant en quelque sorte la *« consommation vitale »* (égale à la consommation pour $y = 0$), k étant un facteur de proportionnalité (compris entre 0 et 1) qui désigne la propension marginale à consommer (en ce sens que la consommation augmente de k unités chaque fois que le revenu national s'accroît d'une unité).

L'investissement et la dépense gouvernementale sont des grandeurs exogènes, qui doivent donc être regardées comme fixées a priori et assimilées à des constantes ; par contre, le revenu national et la consommation sont des grandeurs endogènes, qui doivent dès lors être obtenues (en fonction de i et de g) en résolvant le système formé par les deux équations : $y = c + i + g$ et

(2) On désigne souvent le revenu national par le symbole y (pour *« yield »* en anglais), et non par la lettre R qui pourrait aussi indiquer la recette. Par ailleurs, on emploie indifféremment des majuscules ou des minuscules pour désigner les différentes grandeurs considérées.

$c = \overline{c} + ky$. Il s'agit en réalité de résoudre le système cramérien suivant :

$$\left\{\begin{array}{rcl} y - c & = & i + g \\ -ky + c & = & \overline{c} \end{array}\right.$$

On trouve aisément l'unique solution :

$$\left\{\begin{array}{rcl} y & = & m(i + g + \overline{c}) \\ c & = & m(\overline{c} + k(i + g)) \end{array}\right., \text{ avec } m = (1 - k)^{-1}.$$

On constate notamment que le revenu national y est égal au produit de la somme des trois grandeurs exogènes i, g et $\overline{c}$ par le *« multiplicateur d'impact »* m qui est strictement supérieur à 1.

1.5 Initiation à la théorie des jeux

1.5.1 Position du problème

Il s'agit essentiellement de présenter un modèle mathématique ([3]) dans lequel les intérêts, souvent contradictoires, de deux ou plusieurs décideurs (individus, firmes, pays, ...) sont pris en considération.

Nous traiterons uniquement le cas de deux *joueurs* A et B, dont les objectifs sont de gagner le plus possible ou de perdre le moins possible. Le gain de l'un est égal à la perte de l'autre, le jeu étant alors dit de somme nulle. La règle du jeu peut être résumée par un tableau, appelée la matrice du jeu, qui exprime les gains de A (un gain négatif devant être regardé comme une perte). De façon précise, A peut choisir entre m options ou stratégies $x_1, x_2, \ldots, x_m$, et B entre n stratégies $y_1, y_2, \ldots, y_n$; les choix sont faits simultanément. Lorsque A utilise la stratégie x_i et B la stratégie y_j, B paie à A le montant a_{ij} : de la sorte, la matrice $G = (a_{ij})$ du jeu représente les gains de A, ceux de B seraient donnés par la matrice $-G$.

Une ligne de G est relative à une stratégie du joueur A et une colonne à une stratégie de B. L'intérêt de A est de choisir la i-ème ligne de façon à rendre a_{ij} maximum, tandis que celui de B est de sélectionner la j-ème colonne pour laquelle a_{ij} est minimum.

([3]) La théorie mathématique des jeux de stratégie a été développée principalement par von Neumann dans plusieurs mémoires parus entre 1928 et 1942 ; elle a été exposée de façon systématique dans le traité de von Neumann-Morgenstern : *Theory of games and economic behavior*, Princeton New York, 1944

L'hypothèse de départ consiste à supposer les deux joueurs A et B intelligents (leurs décisions étant donc rationnelles) et prudents (en ce sens qu'ils ne prennent aucun risque).

1.5.2 Exemple d'un jeu strictement déterminé

Analysons un problème concret. Deux firmes concurrentes A et B se partagent un marché. Leurs directeurs doivent, à une époque déterminée, organiser leur campagne publicitaire, qui peut se faire soit à la radio, soit dans la presse écrite. Les informations suivantes ont été récoltées. Si la firme A fait uniquement sa publicité à la radio, elle aura un gain supplémentaire de 4 unités monétaires (u.m.) dans le cas où son concurrent prendrait la même décision, et un gain de 2 u.m. si l'adversaire fait sa publicité dans la presse écrite ; par contre si A choisit la presse, il perd 3 u.m. quand B opte pour la publicité à la radio, et gagne 1 u.m. lorsque B fait sa campagne dans la presse. La matrice du jeu, qui représente les gains de la firme A, se présente comme suit :

$$G = \begin{pmatrix} 4 & 2 \\ -3 & 1 \end{pmatrix}$$

Il est clair que la firme A choisira à coup sûr la radio pour y placer sa publicité, car cette stratégie est dominante en ce sens que les nombres de la première ligne sont supérieurs aux éléments correspondants de la deuxième ligne. D'un autre côté, l'intérêt de la firme B est alors de faire sa publicité dans la presse écrite, puisque sa perte sera dans ce cas minimale. Il existe donc une stratégie qui réalise un compromis satisfaisant les deux parties et le jeu est alors qualifié de strictement déterminé : cette stratégie optimale correspond au « *point-selle* » (nombre le plus petit dans sa ligne et le plus grand dans sa colonne) situé sur la première ligne et la deuxième colonne de la matrice G.

1.5.3 Exemple d'un jeu non strictement déterminé

Dans le cas d'un jeu non strictement déterminé, c'est-à-dire qui ne possède aucun point-selle, la réponse au problème posé est moins simple. On peut démontrer qu'elle est apportée par une stratégie mixte : dans une partie comportant un grand nombre de coups, les deux joueurs A et B doivent choisir un certain nombre de stratégies selon des fréquences appropriées.

A titre d'exemple, traitons le jeu caractérisé par la matrice

$$G = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}.$$

Il n'existe aucun point-selle. Chaque joueur devra adopter une stratégie « mixte » ; ainsi, A choisira la première stratégie avec la probabilité p_1, et la seconde avec la probabilité p_2 : on a forcément $p_1 + p_2 = 1$. Si B choisit sa première stratégie, le gain espéré de A sera donné par $E_1 = 1.p_1 + (-1).p_2$; par contre, si B opte pour sa deuxième stratégie, l'espérance du gain de A sera de $E_2 = 0.p_1 + 2.p_2$.

En vertu de nos hypothèses sur les deux joueurs, A s'efforcera de maximiser ses gains minimaux assurés, ce qui sera visiblement réalisé lorsque $E_1 = E_2$, soit $p_1 - p_2 = 2p_1$ avec $p_1 + p_2 = 1$, c'est-à-dire pour $p_1 = \frac{3}{4}$ et $p_2 = \frac{1}{4}$: la valeur théorique du jeu (qui représentera l'espérance de gain pour le joueur A) sera égale à $\frac{1}{2}$.

1.6 Exemples d'introduction aux chaînes de Markov

1.6.1 Définitions

Examinons une suite d'épreuves E_1, E_2,. . .,E_k,. . .dont les résultats appartiennent à un ensemble fini $\{S_1, S_2, \ldots, S_n\}$. Si le résultat de la k-ème épreuve E_k est S_i, on dit que le système est dans l'état S_i à la k-ème transition (ou encore au temps k).

On appelle chaîne de Markov une telle succession d'épreuves, dans la mesure où la condition suivante est satisfaite : pour tous indices i et j de $1, 2, \ldots, n$, la probabilité pour que le système passe, en une seule épreuve, de l'état S_i à l'état S_j ne varie pas au cours du temps et est donc indépendante du numéro k de l'épreuve au cours de laquelle s'effectue cette transition. Pour chaque couple d'états (S_i, S_j), on note alors p_{ij} la probabilité constante pour que S_j se produise immédiatement après S_i. Les nombres p_{ij} forment la matrice $P = (p_{ij})$, appelée la *matrice de transition* du processus stochastique considéré.

1.6.2 Premier exemple

En guise d'exemple très simple, considérons un employé qui se rend chaque jour à son bureau soit en voiture, soit en train. Il a décidé de ne jamais prendre la voiture deux jours consécutifs ; par contre, le jour où il fait le trajet en train, il prend le soir une pièce de monnaie parfaitement équilibrée et la lance : s'il obtient face (resp. pile), il partira le lendemain en train (resp. en voiture). Cette situation évoque un processus stochastique dont tous les résultats possibles sont les suivants :{v : il utilise la voiture} et {t : il part en train}. De plus, le résultat du jour dépend de ce qui s'est passé la veille. La matrice de transition décrivant le modèle peut donc se présenter comme suit : $P = \begin{pmatrix} 0 & 1 \\ 0,5 & 0,5 \end{pmatrix}$, les premières ligne et colonne se rapportant à la voiture, les secondes au train.

Admettons que, le lundi, notre personnage jette un dé, puis décide de partir en train si et seulement si le jet donne 1 pour résultat : il a donc alors 5 chances sur 6 de prendre sa voiture, pour 1 chance sur 6 de partir en train, ce qui peut être résumé par le vecteur de probabilité $X_0 = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \end{pmatrix}$. Le mardi de cette semaine, ses chances de partir en voiture ou en train sont respectivement données par le vecteur $X_1 = X_0 P = \begin{pmatrix} \frac{1}{12} & \frac{11}{12} \end{pmatrix}$; le mercredi, le vecteur de probabilité sera égal à $X_2 = X_1 P = X_0 P^2 = \begin{pmatrix} \frac{11}{24} & \frac{13}{24} \end{pmatrix}$; plus généralement, après k jours, son vecteur de probabilité sera $X_k = X_0 P^k$.

Pour savoir ce qui se passera dans le futur, diagonalisons la matrice de transition P. La matrice régulière $R = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$ donne lieu à l'égalité

$$R^{-1} P R = \operatorname{diag} \begin{pmatrix} 1 & -\frac{1}{2} \end{pmatrix}$$

d'où

$$P = R \operatorname{diag} \begin{pmatrix} 1 & \frac{-1}{2} \end{pmatrix} R^{-1}$$

puis, de proche en proche,

$$\begin{aligned} P^k &= R D^k R^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (-0,5)^k \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{-1}{3} & \frac{1}{3} \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \end{pmatrix} + (-0,5)^k \begin{pmatrix} -2 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{-1}{3} & \frac{1}{3} \end{pmatrix} \end{aligned}$$

Comme $(-0,5)^k$ converge vers 0 lorsque k tend vers l'infini, un passage à la limite livre

$$\lim_{k\to\infty} P^k = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Le vecteur de probabilité $Y = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \end{pmatrix}$ est un *point fixe* pour P et correspond à une distribution stationnaire de la chaîne de Markov (puisque $YP = Y$). En pratique, cela signifie qu'à très long terme, l'employé a une chance sur trois de partir en voiture.

1.6.3 Deuxième exemple

Traitons un autre cas concret. Supposons que des firmes soient classées en trois catégories. On pourrait, par exemple, s'intéresser à la production et mettre dans une première (resp. deuxième ; troisième) classe les entreprises dont la production, par rapport à l'année précédente, a augmenté de plus de 5 % (resp. a varié de 5 % ou moins ; a diminué de plus de 5 %). La matrice de transition est, par exemple, la suivante (les lignes se référant aux résultats de l'année passée, et les colonnes à ceux de cette année) :

$$P = \begin{pmatrix} 0,7 & 0,3 & 0 \\ 0,2 & 0,5 & 0,3 \\ 0,2 & 0,2 & 0,6 \end{pmatrix}$$

Pratiquement, on constate que si la production a augmenté l'an passé de plus de 5 %, elle a 7 chances sur 10 (probabilité égale à 0,7) d'encore augmenter de plus de 5 % cette année, 3 chances sur 10 de varier de 5 % ou moins ; par contre, il est « très probable » qu'elle ne diminuera pas de plus de 5 %. De même, on voit que si la production est demeurée pratiquement constante (resp. a eu tendance à diminuer) l'an passé, elle a plus de chances de rester stable (resp. de continuer à décroître) cette année. Pour connaître l'évolution à long terme d'une firme, il convient de rechercher si la matrice P ne possède pas un point fixe, c'est-à-dire un vecteur de probabilité Y tel que $YP = Y$; cela revient ici à résoudre le système linéaire que voici :

$$\left\{ \begin{array}{rcrcrcl} 0,7y_1 & + & 0,2y_2 & + & 0,3y_3 & = & y_1 \\ 0,3y_1 & + & 0,5y_2 & + & 0,2y_3 & = & y_2 \\ & & 0,3y_2 & + & 0,6y_3 & = & y_3 \\ y_1 & + & y_2 & + & y_3 & = & 1 \end{array} \right.$$

On trouve aisément $y_1 = \frac{2}{5}$, $y_2 = \frac{12}{35}$ et $y_3 = \frac{9}{35}$, de sorte que les puissances P^k de la matrice P possèdent, lorsque k tend vers l'infini, une limite $\overline{P}$ donnée par

$$\overline{P} = \lim_{k\to\infty} P^k = \begin{pmatrix} \frac{2}{5} & \frac{12}{35} & \frac{9}{35} \\ \frac{2}{5} & \frac{12}{35} & \frac{9}{35} \\ \frac{2}{5} & \frac{12}{35} & \frac{9}{35} \end{pmatrix} \approx \begin{pmatrix} 0,4 & 0,3428571 & 0,2571428 \\ 0,4 & 0,3428571 & 0,2571428 \\ 0,4 & 0,3428571 & 0,2571428 \end{pmatrix}.$$

On peut effectivement observer que les puissances de P convergent vers $\overline{P}$; ainsi :

$$P^2 = \begin{pmatrix} 0,55 & 0,36 & 0,09 \\ 0,3 & 0,37 & 0,33 \\ 0,3 & 0,28 & 0,42 \end{pmatrix}, \cdots, P^5 = \begin{pmatrix} 0,41875 & 0,35067 & 0,23058 \\ 0,3875 & 0,33869 & 0,27381 \\ 0,3875 & 0,33626 & 0,27624 \end{pmatrix}$$

$$P^{10} = \begin{pmatrix} 0,4005859375 & 0,3431463156 & 0,2562677469 \\ 0,399609375 & 0,3426668917 & 0,2577237333 \\ 0,399609375 & 0,3626609868 & 0,2577296382 \end{pmatrix}$$

...

$$P^{20} = \begin{pmatrix} 0,400000572204589 & 0,342857428937022 & 0,257141998858387 \\ 0,399999618530273 & 0,342856952152166 & 0,25714342931756 \\ 0,399999618530273 & 0,342856952117298 & 0,257143429352428 \end{pmatrix}$$

...

$$P^{24} = \begin{pmatrix} 0,400000035762786 & 0,342857160738354 & 0,257142803498858 \\ 0,399999976158141 & 0,342857130936455 & 0,257142892905402 \\ 0,399999976158141 & 0,342857130936173 & 0,257142892905684 \end{pmatrix}.$$

De ce qui précède, on peut notamment tirer cette conclusion : après un temps suffisamment long, la firme en question a 4 chances sur 10 (probabilité égale à $\frac{2}{5}$) de voir sa production augmenter de plus de 5 % sur une année, et cela indépendamment des résultats de l'année précédente.

1.7 Introduction élémentaire à la programmation linéaire

1.7.1 Position du problème

La programmation linéaire a pour objet la recherche du maximum ou du minimum d'une fonction linéaire $f(x) = \sum_{i=1}^{n} c_i x_i$, à coefficients c_i réels et à n variables $x_1, x_2, \ldots, x_n$, sur un ensemble E de l'espace numérique à n dimensions.

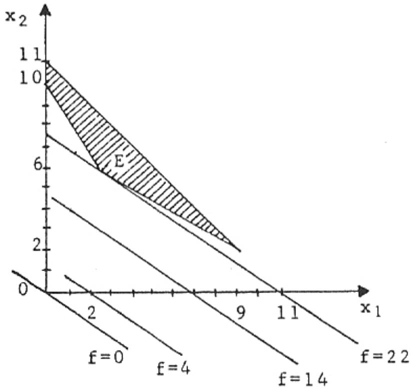
La fonction linéaire à optimiser reçoit souvent le nom (discutable) de fonction économique. Les variables x_i ont une signification concrète (quantités produites, durées, valeurs monétaires,...) qui les obligent presque toujours à être positives ou nulles ; en dehors de ces conditions générales de non-négativité, les vraies contraintes, c'est-à-dire celles qui sont spécifiques au problème traité et doivent être précisées dans chaque cas, peuvent toujours se mettre sous la forme $\sum_{i=1}^{n} a_i x_i \leqslant b$ quitte à multiplier les deux membres de l'inégalité par -1, ou à cumuler deux inégalités pour former une égalité). En résumé, il s'agit d'optimiser une fonction f qui peut se mettre sous la forme matricielle CX sachant que $X \geqslant 0$ et $AX \leqslant B$.

1.7.2 Exemple à deux variables

Illustrons la théorie par un exemple très simple (mais qui pourrait être aisément adapté pour rencontrer des situations plus complexes).

Deux ouvriers A et B travaillent à temps réduit dans un petit atelier et fabriquent tous les deux des pièces de type P_1 et de type P_2 ; A (resp. B) gagne 2 (resp. 3) unités monétaires par heure ; A (resp. B) produit 10 (resp. 5) pièces P_1 et 4 (resp. 7) pièces P_2 chaque heure. L'atelier reçoit une commande pressante de 50 pièces de chaque type. Pour des raisons budgétaires, il n'est pas possible de consacrer plus de 11 heures (au total) pour la fabrication des diverses pièces. Le problème posé à l'entrepreneur est de savoir combien d'heures A et B devront travailler pour que leur salaire total soit minimum.

Désignons par x_1 et x_2 le nombre d'heures de travail pour A et B respectivement. On doit avoir $x_1 \geqslant 0, x_2 \geqslant 0, x_1 + x_2 \leqslant 11, 4x_1 + 7x_2 \geqslant 50$ et $10x_1 + 5x_2 \geqslant 50$. Ces contraintes définissent l'ensemble E des programmes réalisables qui est, dans le plan euclidien, le quadrilatère convexe dont les sommets sont les points $(2, 6), (9, 2), (0, 11)$ et $(0, 10)$. La fonction économique, qu'il faut rendre minimum, est donnée par $f(x_1, x_2) = 2x_1 + 3x_2$. Il est visible que le minimum de la fonction f sur E a pour valeur 22 unités monétaires et est atteint au point extrême $(2, 6)$. Concrètement, l'ouvrier A devra travailler 2 heures, tandis que B sera à l'ouvrage pendant 6 heures.



Notons que cette solution n'est pas tellement évidente, puisqu'il faut employer plus longtemps l'ouvrier B qui gagne le plus ; bien sûr, A et B n'ont pas le même rendement, mais observons que A produit 5 pièces P_1 de plus que B, alors que B ne fabrique que 3 pièces P_2 de plus que A (par heure).

1.7.3 Méthode algébrique de résolution pour 3 variables et plus

Lorsque le nombre de variables est plus élevé que 2, un problème de programmation linéaire peut être résolu en recherchant tous les points extrêmes de l'ensemble E des programmes réalisables ; la plus grande (resp. la plus petite) valeur de la fonction économique en un de ces points donne le maximum (resp. le minimum) recherché, pour autant que ce dernier existe ([4]) ; la restriction énoncée en dernier lieu est rencontrée simplement dès qu'on connaît un majorant (resp. un minorant) de la fonction, plus particulièrement quand E est borné, c'est-à-dire compact.

Traitons par cette méthode un problème concret (et simple) de publicité. Une société fabrique de la poudre à lessiver ; elle désire commercialiser son produit dans une région pour laquelle une étude de marché a montré que l'impact d'une annonce publicitaire était donné par le tableau suivant :

médium	audience (en millions)	audience féminine (en millions)	coût par annonce (en unités monétaires)
télévision	10	7	500
radio	1	0,6	60
presse écrite	2	0,8	65

([4]) Nous reviendrons plus en détail sur ce théorème fondamental de la programmation linéaire au paragraphe 2.4.1.

Le directeur commercial cherche à déterminer le budget de publicité qui lui permette d'atteindre le public qu'il s'est fixé ; il voudrait toucher au moins 20 millions de personnes, dont au moins 14 millions de femmes. Comment peut-il rendre minimum le coût de sa campagne publicitaire compte tenu des contraintes créées par ses objectifs ?

Pour répondre à cette question, considérons les trois inconnues x_1, x_2 et x_3 qui représentent le nombre d'annonces respectivement à la télévision, à la radio et dans la presse écrite. On doit avoir $x_1 \geqslant 0$, $x_2 \geqslant 0$,$x_3 \geqslant 0$, ainsi que les vraies contraintes $10x_1 + x_2 + 2x_3 \geqslant 20$ et $7x_1 + 0,6x_2 + 0,8x_3 \geqslant 14$. La recherche des points extrêmes de l'ensemble E des programmes réalisables passe par la résolution des 10 systèmes linéaires et cramériens suivants :

	Solutions
1) $\begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{cases}$	$x_1 = x_2 = x_3 = 0$
2) $\begin{cases} x_1 = 0 \\ x_2 = 0 \\ 10x_1 + x_2 + 2x_3 = 20 \end{cases}$	$x_1 = x_2 = 0, x_3 = 10$
3) $\begin{cases} x_1 = 0 \\ x_2 = 0 \\ 7x_1 + 0,6x_2 + 0,8x_3 = 14 \end{cases}$	$x_1 = x_2 = 0, x_3 = 35/2$
4) $\begin{cases} x_1 = 0 \\ x_3 = 0 \\ 10x_1 + x_2 + 2x_3 = 20 \end{cases}$	$x_1 = x_3 = 0, x_2 = 20$
5) $\begin{cases} x_1 = 0 \\ x_3 = 0 \\ 7x_1 + 0,6x_2 + 0,8x_3 = 14 \end{cases}$	$x_1 = x_3 = 0, x_2 = 70/3$
6) $\begin{cases} x_2 = 0 \\ x_3 = 0 \\ 10x_1 + x_2 + 2x_3 = 20 \end{cases}$	$x_1 = 2, x_2 = x_3 = 0$
7) $\begin{cases} x_2 = 0 \\ x_3 = 0 \\ 7x_1 + 0,6x_2 + 0,8x_3 = 14 \end{cases}$	$x_1 = 2, x_2 = x_3 = 0$
8) $\begin{cases} x_1 = 0 \\ 10x_1 + x_2 + 2x_3 = 20 \\ 7x_1 + 0,6x_2 + 0,8x_3 = 14 \end{cases}$	$x_1 = 0, x_2 = 30, x_3 = -5$

$$9)\left\{\begin{array}{rcrcrcl} & & x_2 & & & = & 0 \\ 10x_1 & + & x_2 & + & 2x_3 & = & 20 \\ 7x_1 & + & 0,6x_2 & + & 0,8x_3 & = & 14 \end{array}\right. \qquad x_1 = 2, x_2 = x_3 = 0$$

$$10)\left\{\begin{array}{rcrcrcl} & & & & x_3 & = & 0 \\ 10x_1 & + & x_2 & + & 2x_3 & = & 20 \\ 7x_1 & + & 0,6x_2 & + & 0,8x_3 & = & 14 \end{array}\right. \qquad x_1 = 2, x_2 = x_3 = 0$$

Or, nous savons qu'une solution $x = (x_1, x_2, x_3)$ d'un tel système est un sommet de E si elle appartient précisément à E (c'est-à-dire si elle satisfait à toutes les inégalités initiales). Les points extrêmes de E sont donc $(0, 0, 35/2), (0, 70/3, 0)$ et $(2, 0, 0)$, où la fonction économique $f(x_1, x_2, x_3) = 500x_1 + 60x_2 + 65x_3$ prend respectivement les valeurs $1137, 5,\ 1400$ et 1000 unités monétaires. Comme la fonction f est visiblement non négative sur E (déjà d'après sa définition de coût de campagne publicitaire), l'existence du minimum est assurée, et c'est le point $(2, 0, 0)$ qui la procure. Pratiquement, il faut uniquement faire deux annonces à la télévision (moyen publicitaire le plus coûteux, mais le plus efficace).

La méthode qui vient d'être utilisée pour résoudre ce problème particulier est tout à fait générale, mais devient vite impraticable lorsque le nombre d'inconnues ou de variables augmente. En effet, le nombre de programmes extrêmes (donc le nombre de systèmes cramériens à résoudre) peut devenir très grand, puisqu'il peut atteindre la valeur $\binom{m}{n} = \frac{m!}{n!\,(m-n)!}$, où m désigne le nombre de contraintes définissant E dans l'espace à n dimensions. C'est pourquoi différentes techniques (dont l'algorithme du simplexe dû à Dantzig) ont été mises au point dans le but d'atteindre assez rapidement la solution optimale d'un programme linéaire.

Chapitre 2

Applications statiques de l'analyse mathématique

par *Jacques BAIR*

A la suite de J.R. HICKS, « j'appelle *statiques* des parties de l'économie théorique où nous ne nous soucions pas des dates » (dans *Value and Capital*, 1934).

Nous allons nous placer ici dans ce cadre statique. Cela nous permettra de mettre en évidence l'utilisation des calculs différentiel et intégral en économie.

Tout d'abord, nous donnerons de nombreuses applications de la dérivation (éventuellement partielle). A cet effet, nous passerons en revue quelques fonctions, explicites et implicites, classiques, puis nous aborderons l'important problème d'optimisation au travers quelques problèmes concrets typiques. Nous illustrerons également la notion d'intégration.

Précisons au préalable quelques particularités terminologiques.

La notion de dérivée intervient notamment dans les théories « marginalistes » par l'adjonction du qualificatif *marginal* accolé au nom de la fonction. Par exemple, la dérivée du coût de production $C = f(q)$ de la quantité q d'un bien Q porte le nom de *coût marginal*; celui-ci, qui est la limite pour Δq tendant vers 0 de $\frac{\Delta C}{\Delta q}$, vaut approximativement le coût de la dernière unité produite (ce que l'on peut voir en prenant $\Delta q = 1$)

Dans la pratique, il convient souvent de considérer non pas les variations absolues $\Delta x, \Delta y, \ldots$, des variables $x, y, \ldots$, mais les *variations relatives* $\frac{\Delta x}{x}, \frac{\Delta y}{y}, \ldots$ ou variations en pourcentage. Par exemple, on dira que la vente d'un article a augmenté (resp. diminué) de 5% lorsque la variation relative de la quantité vendue a été de 0,05 (resp. -0,05) pour cet article.

Soit une fonction $y = f(x)$ pour laquelle on s'intéresse aux variations relatives de x et de y. Au lieu de chercher la valeur marginale y' de y, c'est-à-dire la limite du rapport $\frac{\Delta y}{\Delta x}$ pour Δx tendant vers 0, on étudie la limite du rapport des variations relatives, c'est-à-dire de $\frac{\Delta y/y}{\Delta x/x} = \frac{x}{y}\frac{\Delta y}{\Delta x}$. Cette limite est appelée l'*élasticité* de y par rapport à x et peut se noter $E_x y$ (1) ; visiblement,

$$E_x y = \frac{x}{y} \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{x}{y} y' = \frac{d \ln y}{d \ln x} = \frac{(\ln y)'}{(\ln x)'}$$

Cette élasticité $E_x y$ vaut approximativement la grandeur relative (ou en pourcentage) de l'augmentation (si elle est positive) ou de la diminution (si elle est négative) de y lorsque x augmente (en valeur relative) de 1%.

La notion d'élasticité joue un rôle important en économie ; elle traduit la faculté de variation d'un phénomène en fonction d'un autre. Introduite par A. Marshall (1842-1924, économiste anglais considéré comme ayant été le maître de Keynes), l'idée de ce concept trouve son origine dans les travaux de A.R.J. Turgot (1727-1781, économiste français souvent désigné comme le pricipal fondateur de l'économie politique) qui avait constaté qu'un ressort s'allonge de moins en moins au fur et à mesure que la charge supportée s'alourdit.

Les fonctions $y = f(x)$ qui admettent une élasticité constante a sont de la forme $y = Kx^a$

Soit à présent une fonction f de n variables indépendantes $x_1, x_2, \ldots, x_n$. Pour dériver partiellement f par rapport à une variable x_i, on assimile en fait toutes les variables x_j, autres que x_i, à des constantes et on considère dès lors f comme une fonction de l'unique variable x_i : il s'agit de la clause *ceteris paribus* ou encore *toutes choses égales par ailleurs*. Pour une telle fonction f, on peut également introduire une *élasticité partielle* de f par rapport à chaque variable x_i : $E_i f = \frac{x_i}{f} f'_i$, où f'_i désigne la dérivée partielle de f par rapport à x_i.

(1) On utilise également la notation $\varepsilon_{y/x}$ à la place de $E_x y$.

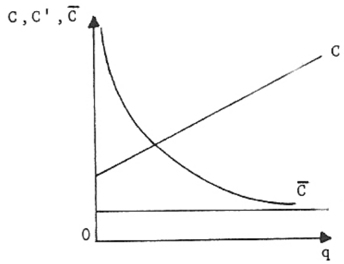
2.1 Exemples de fonctions explicites

2.1.1 Fonctions de coût

On appelle *coût total* la somme globale de toutes les dépenses engagées dans la production d'une certaine quantité de biens.

Supposons tout d'abord que le coût total C soit uniquement fonction de la quantité q produite. Le *coût moyen* $\overline{C}$ est le coût de l'unité du bien produit, soit $\overline{C} = \frac{C}{q}$; par ailleurs, le *coût marginal* C' est la dérivée de la fonction C par rapport à la variable q : la valeur de C' représente approximativement le nombre de pour-cent dont C augmente lorsque la quantité produite q croît de 1%.

En première approximation, on peut prendre $C = aq + b$, où b représente les frais fixes et a le coût de fabrication d'une unité. Le coût moyen $\overline{C} = a + \frac{b}{q}$ est généralement proportionnel au prix de vente unitaire p, avec un coefficient de proportionnalité supérieur à 1 ; dès lors, si la production est réglée sur les commandes, une diminution de la quantité achetée engendre une augmentation des prix.



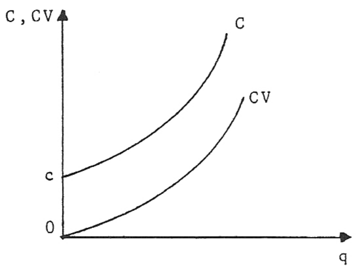
Ce phénomène économique, qui peut être illustré par de nombreux exemples concrets, peut aussi se retrouver pour des fonctions plus compliquées que la fonction linéaire.

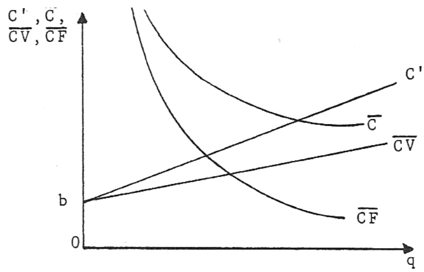
Les fonctions de coût peuvent avoir de nombreuses formes différentes ; par exemple, on fait quelquefois appel à l'une des fonctions suivantes pour définir le coût [2, p. 120] : $aq+b$, aq^2+bq+c, $\sqrt{aq+b}+c$, aq^3-bq^2+cq+d, $aq\frac{q+b}{q+c}+d$, $aq^2\frac{q+b}{q+c}+d$, ae^{bq}, $q^a e^{bq+c}+d$ (où a, b, c et d sont des constantes).

En général, une fonction de coût $f(q)$ possède un certain nombre de propriétés assez caractéristiques. Si aucune unité n'est produite, le coût total $f(0)$ est positif ou nul ; si $f(0) > 0$, $f(0)$ représente les frais fixes CF de production ; le coût variable CV s'obtient alors en retranchant ces frais fixes du coût total, soit $CV = C - CF$. La fonction $f(q)$ est croissante, d'où $f'(q) > 0$. Souvent, le coût marginal croît ; le graphique de la fonction $f(q)$ tourne alors sa concavité vers le haut ou, de façon équivalente, $f''(q) \geqslant 0$. Lorsque la production est peu élevée, le coût marginal est quelquefois décroissant.

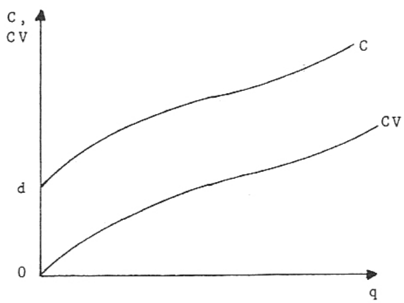
On admet souvent que la forme *normale* des fonctions de coût total est donnée par une fonction quadratique ou cubique.

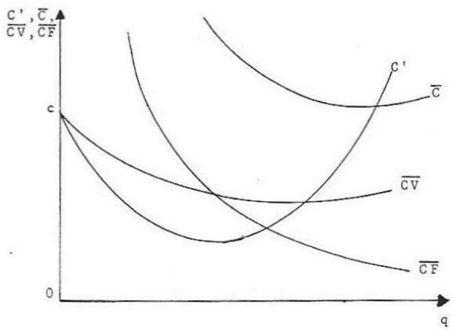
Soit une fonction de coût total quadratique, à savoir : $C = aq^2 + bq + c$, avec $a > 0$, $b \geqslant 0$, $c \geqslant 0$. Le coût total C, le coût marginal $C' = 2aq + b$ et le coût variable moyen $\overline{C}\,\overline{V} = aq + b$ sont toujours croissants, le deuxième C' étant supérieur au troisième $\overline{C}\,\overline{V}$. Par ailleurs, le coût fixe moyen $\overline{C}\,\overline{F} = \frac{CF}{q}$ diminue de façon régulière, la courbe représentant ce coût étant une hyperbole équilatère. Enfin, le coût total moyen $\overline{C}$ a pour sa part la forme dite en U.





Soit une fonction de coût total cubique, à savoir : $C = aq^3 - bq^2 + cq + d$, avec $a > 0, b \geqslant 0, c \geqslant 0, d \geqslant 0$ et $b^2 \leqslant 3ac$. On a $\overline{C} = aq^2 - bq + c + \frac{d}{q}$ et $C' = 3aq^2 - 2bq + c$. La fonction de coût total est toujours croissante, car $C'(q) \geqslant 0$ pour toute valeur de q, et admet un graphique qui tourne d'abord sa concavité vers le bas, car $C''(0) \leqslant 0$, puis vers le haut, car $C''(q) \geqslant 0$ dès que $q > \frac{b}{3a}$.





Les courbes de coût marginal C' et de coût variable moyen $\overline{C}\,\overline{V} = aq^2 - bq + c$ ont toutes les deux la forme d'une parabole tournant sa concavité vers le haut. On observera que le coût marginal C' atteint son minimum avant la fonction $\overline{C}\,\overline{V}$, et que $\overline{C}\,\overline{V}$ atteint son minimum avant $\overline{C}$.

Par ailleurs, la courbe C' passe par le minimum de la courbe $\overline{C}\,\overline{V}$ et par celui de la courbe $\overline{C}$ (puisque $\frac{d}{dq}\overline{C} = \frac{1}{q}(C' - \overline{C})$ et $\frac{d}{dq}\overline{C}\,\overline{V} = \frac{1}{q}(C' - \overline{C}\,\overline{V})$). La courbe $\overline{C}\,\overline{F}$ est une hyperbole équilatère indépendante de la forme des autres courbes de coût. La distance d'un point de la courbe $\overline{C}$ à son correspondant de la courbe $\overline{C}\,\overline{V}$, situé à la même abscisse, est égale à $\overline{C}\,\overline{F}$, et donc décroît lorsque la quantité q augmente.

Les considérations précédentes valent pour le court terme, en supposant bien déterminé l'équipement utilisé.

Admettons à présent que le coût à court terme varie avec le volume q de la production et la taille λ des installations ; il peut dès lors se traduire par une égalité du type $C(q, \lambda) = C$, où nous supposons continûment dérivable la fonction $C(q, \lambda)$. A long terme, l'entrepreneur rationnel va choisir l'équipement de taille λ pour lequel la dépense occasionnée pour la production de chaque volume q est la plus faible ; la condition du premier ordre relative à un tel minimum s'écrit alors $\frac{\partial C}{\partial \lambda}(q, \lambda) = 0$. Le coût à long terme est donc généralement obtenu en éliminant λ entre les équations $C(q, \lambda) = C$ et $\frac{\partial C}{\partial \lambda}(q, \lambda) = 0$; la courbe de coût à long terme est donc l'enveloppe des courbes de coût à court terme.

En guise d'illustration numérique, supposons connues les fonctions de coût à court terme pour quelques valeurs de la taille de l'équipement :

$$C_1 = 0,009q^3 - 0,5q^2 + 17q + 10 \text{ pour } \lambda = 1$$

$$C_2 = 0,009q^3 - 0,5q^2 + 17q + 40 \text{ pour } \lambda = 2$$

$$C_3 = 0,009q^3 - 0,5q^2 + 17q + 90 \text{ pour } \lambda = 3$$

Ces expressions sont des cas particuliers de la fonction

$$C(q, \lambda) = 0,009q^3 - 0,5q^2 + (19 - 2\lambda)q + 10\lambda^2$$

La fonction C de coût à long terme s'obtient en éliminant λ entre les équations

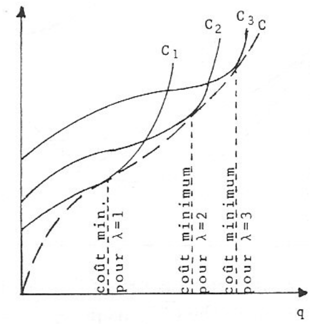
$$C(q, \lambda) = C$$

et

$$\frac{\partial C}{\partial \lambda} = -2q + 20\lambda = 0,$$

ce qui donne

$$C = 0,009q^3 - 0,6q^2 + 10q.$$



2.1.2 Fonctions de demande

Pour un bien déterminé, dont les quantités sont mesurées à l'aide d'une certaine unité, la demande q, c'est-à-dire la quantité vendue pendant une période de référence convenue, est fonction du prix de vente unitaire p. Elle dépend d'ailleurs d'autres éléments tels que la qualité du produit, les qualités et prix des biens concurrents ou complémentaires, les conditions économiques du moment, etc... Mais on peut, au moins en esprit, supposer fixes tous les éléments, autres que p, qui peuvent influencer la demande, de manière à considérer la dépendance $q = f(p)$ à l'égard de p seulement : f est alors la *fonction de demande.*

L'ensemble de définition de f est évidemment inclus dans l'ensemble des réels positifs ou nuls, et les valeurs de f sont elles-mêmes non négatives. Bien que seuls des prix rationnels soient pratiqués, on cherche à profiter au maximum de la théorie des fonctions en supposant continue (sur un intervalle) la fonction f.

En général, la fonction f est décroissante. Néanmoins, il existe au moins trois situations très classiques pour lesquelles la quantité demandée q croît en même temps que le prix p.

1. *Le paradoxe de* GIFFEN. Dans le passé, la consommation de biens de grande nécessité (par exemple, le pain ou la pomme de terre) augmentait en même temps que le prix de ce produit. En effet, la hausse

du prix d'un tel bien absorbait une part importante des revenus des travailleurs, de sorte que ceux-ci devaient renoncer à acheter d'autres denrées alimentaires (plus luxueuses, mais plus coûteuses) et les remplaçaient forcément par un bien de nécessité ; en revanche, une baisse du prix de ce produit permettait de réaliser une économie et ainsi d'acheter moins de ce bien au profit d'une plus grande consommation d'articles plus chers. Voici un exemple concret ([2]) illustrant cet effet ou paradoxe de GIFFEN (1837-1910, économiste britannique qui constata le premier ce phénomène). Une famille américaine dispose de 110 dollars par an pour sa consommation de pain et de viande. Lorsque le prix d'une miche de pain (resp. un kilo de viande) est fixé à 0,3 dollar (resp. 2 dollars), le ménage achète 300 miches de pain par an et 10 kilos de viande. Si le prix du pain passe à 0,1 dollar la miche, le prix de la viande restant inchangé, la famille peut s'offrir, avec ses 110 dollars, 42 kilos de viande et seulement 260 miches de pain : une diminution du prix du pain entraîne donc une consommation moindre de pain au profit d'une augmentation de la quantité de viande.

2. *L'effet* VEBLEN (économiste et sociologue américain, 1857-1929).

   Un bien peut être convoité pour son prix élevé, principalement en fonction de la signification sociale de son achat (par exemple, pour des raisons de « standing »). Il en est ainsi pour certains objets de luxe, comme le diamant ou le manteau de fourrure, qui seraient probablement moins demandés si leur prix était moins élevé.

3. *La spéculation*

   Ce phénomène bien connu se produit dans des conjonctures très particulières. Lorsqu'une hausse des prix s'annonce, les consommateurs augmentent leur demande (même au risque de faire des « provisions ») s'ils croient en la poursuite du mouvement des prix ; au contraire, lorsqu'une baisse des prix se produit, les clients ont tendance à réduire leur demande, car ils espèrent que les prix vont continuer à descendre et préfèrent dès lors attendre un moment plus propice pour effectuer leurs achats.

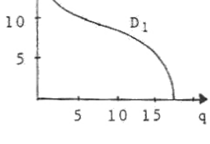
La rapidité de la décroissance de f est fonction du marché et dépend de l'élasticité $E_p f$ de la demande par rapport au prix. Le marché du bien considéré est *élastique* si $|E_p f| > 1$, *inélastique* si $|E_p f| < 1$. Par exemple,

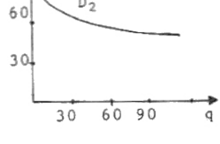
([2]) Cet exemple est tiré de l'ouvrage *Microéconomique, cours et problèmes par* D. Salvatore, Mc-Graw Hill, Série Schaum, 1985, p. 72

le marché d'une marque de poudre à lessiver est généralement élastique, car la demande de cette marque réagit fortement aux variations des prix en raison de la concurrence des autres marques et de l'existence de substituts valables (tels que les liquides pour la lessive) ; par contre, le marché du sucre, toutes marques confondues, est théoriquement inélastique, vu l'absence de concurrent sérieux pour cette denrée et la nécessité d'acheter ce produit quel qu'en soit le prix.

Avant d'examiner quelques exemples concrets de fonctions de demande, signalons que les économistes ont l'habitude de porter les quantités demandées en abscisses et les prix en ordonnées.

La forme de la loi de demande peut varier sensiblement d'un problème à l'autre, ainsi qu'en témoignent ces deux exemples typiques. La courbe D_1 représente la demande de sucre d'un ménage, le prix unitaire p et la quantité q étant exprimés en une certaine unité. Si le prix est inférieur à 15, aucune demande n'est enregistrée. Quand le prix diminue, la demande augmente lentement dans un premier temps, plus vite par la suite, puis à nouveau lentement pour de faibles prix. Cette courbe de demande, qui coupe les deux axes des prix et des quantités, est d'abord convexe, puis concave (voir la première figure ci-dessous). La courbe D_2 illustre la demande de cigarettes d'une marque X, en supposant les prix des marques concurrentes fixes, égaux à 60F le paquet. La demande de cigarettes X est évidemment très faible lorsque le prix est élevé, par exemple fixé à 90F le paquet ; mais, cette quantité augmente rapidement lorsque le prix baisse et devient même très importante (sur un marché assez large) lorsque le prix unitaire est inférieur ou égal à 50F. La courbe de demande rencontre donc l'axe vertical des prix, mais est pratiquement parallèle à une droite horizontale pour de faibles prix, de sorte que le point de rencontre de cette courbe avec l'axe des abscisses correspond à une demande très importante par rapport à celle relative au prix standard de 60F le paquet (voir la seconde figure ci-dessous).





En pratique, il est souvent intéressant de représenter la loi de la demande par une courbe d'équation connue, au moins dans un certain domaine de prix : par exemple, on fait quelquefois appel à une portion de ligne droite, de parabole, d'hyperbole ou de courbe exponentielle. Voici huit types courants de fonctions reliant la quantité demandée q au prix p (a, b, c étant des constantes) [2, p.144] :

(1) $q = \frac{a-p}{b}$	ou $p = a - bq$	(2) $q = \frac{a}{p+c} - b$	ou $p = \frac{a}{q+b} - c$
(3) $q = \frac{a-p^2}{b}$	ou $p = \sqrt{a - bq}$	(4) $q = \frac{a-\sqrt{p}}{b}$	ou $p = (a - bq)^2$
(5) $q = \sqrt{\frac{a-p}{b}}$	ou $p = a - bq^2$	(6) $q = bp^{-a} + c$	ou $p = (\frac{b}{q-c})^{\frac{1}{a}}$
(7) $q = ae^{-bp}$	ou $p = \frac{1}{b} \ln \frac{a}{q}$	(8) $q = p^a e^{-b(p+c)}$	

Notons que la dernière fonction est caractéristique d'une demande dont l'élasticité dépend linéairement du prix : en effet, l'égalité $E_p q = a - bp$ entraîne $q = Kp^a e^{-bp}$ (où k désigne une constate d'intégration) ; en particulier, la loi de la demande dont l'élasticité est une constante, égale à $-m$, est donnée par $q = Kp^{-m}$.

Une fonction de demande à une seule variable est généralement une première approximation, car la demande d'un produit dépend non seulement du prix unitaire de ce bien, mais aussi du prix de vente d'autres produits sur le marché ainsi que du revenu global disponible. Par exemple, la quantité q demandée d'un bien, vendu au prix unitaire p, est donnée par $q = kR^a p^{-b} \pi^c$, où a, b, c et k sont des constantes positives, R (resp. π) désigne le revenu total disponible (resp. le prix de vente moyen des autres produits). L'élasticité partielle *directe* (resp. *croisée*) de la demande vaut $E_p q = -b$ (resp. $E_\pi q = c$), tandis que l'élasticité partielle de la demande par rapport au revenu est égale à $E_R q = a$; dans la pratique, ces diverses élasticités ne sont généralement pas indépendantes. [7].

2.1.3 A propos du revenu brut

Nous allons considérer le marché d'un seul bien vendu au prix unitaire p et en quantité q ; le *revenu brut* (ou *recette totale*) est alors donné par la formule : $R = pq$.

Nous nous proposons d'analyser le comportement de cette recette R pour des variations du prix p.

Notons d'emblée que l'influence de p sur R est relativement complexe dans la mesure où le prix de vente intervient « doublement » dans la recette : directement sous forme du prix unitaire p, indirectement du fait qu'il influence la quantité q.

Le problème fondamental revient en somme à déterminer la forme précise de la relation entre la demande q et le prix p, ce qui peut être réalisé grâce à la notion d'élasticité de la demande, c'est-à-dire à l'élasticité $E_p q$ de la fonction « quantité demandée » par rapport à la variable « prix ».

Pour voir comment intervient l'élasticité de la demande dans notre problème, considérons le revenu brut R comme fonction de l'unique variable p, ce qui est permis puisque $q = f(p)$. Calculons ensuite la dérivée R' de R par rapport à p : nous obtenons aisément

$$R' = q + pq' = q(1 + \frac{p}{q}q') = q(1 + E_p q)$$

Comme, par hypothèse, q est toujours une quantité positive, le signe de R' dépend essentiellement de la valeur de $E_p q$. Plusieurs cas doivent être envisagés.

Lorsque la quantité marginale q', donc aussi $E_p q$, est positive, la dérivée R' est positive de sorte que la recette R est une fonction croissante du prix : une augmentation de prix entraînera un accroissement du revenu brut.

Pareille éventualité est peu fréquente, comme nous l'avons vu au paragraphe précédent. En général, bien sûr, l'élasticité de la demande est négative puisque la quantité demandée diminue le plus souvent lorsque le prix augmente. Le signe de R' dépend alors de la valeur absolue de $E_p q$.

Lorsque $-1 < E_p q < 0$, le marché est *inélastique* car la variation relative de la demande est, en valeur absolue, plus faible que celle du prix : la dérivée R' est positive et la recette R est à nouveau une fonction croissante du prix. Par contre, lorsque $E_p q < -1$, le marché est *élastique* car la valeur absolue de la variation relative de la demande dépasse celle du prix : $R' < 0$ et R décroît si p augmente.

Illustrons cette théorie par deux exemples concrets.

Une étude de marketing ([3]). En vue du lancement d'un nouveau rasoir électrique sur le marché français, un gestionnaire a émis ses prévisions pour trois prix de vente hypothétiques :

prix de vente (en francs)	100	120	140
volume prévu des ventes (en unités)	100000	85000	50000
recette totale	10000000	10200000	7200000

([3]) Ces données sont dues à D. Lindon, dans *Le marketing*, Editions F. Nathan, 1981, pp. 120-121

Ainsi, lorsque le prix passe de 100 à 120F, on doit s'attendre à une augmentation de la recette totale puisque la réduction du volume des ventes est modérée en raison de la faible élasticité (égale à -0,75). Par ailleurs, si le prix est porté de 100 à 140F, le revenu brut va diminuer à cause d'une forte élasticité de la demande (égale à -1,25). Le premier cas est typique d'un marché inélastique, le second se réfère à un marché élastique. Il convient encore de noter que, dans un tel marché, le coût total (égal à la somme des frais fixes et des coûts variables) est pratiquement le même dans les trois cas, de sorte que le bénéfice maximum sera obtenu pour un prix unitaire fixé à 120F.

Le marché des fruits en Belgique. La demande, supposée égale à l'offre, est ici limitée par l'importance de la production (c'est-à-dire de la récolte) ; quant au prix, il n'est pas fixé a priori, mais calculé de manière à écouler la récolte obtenue. Ce marché est (presque toujours) inélastique puisque la variation relative de la quantité récoltée est généralement inférieure en valeur absolue à celle du prix unitaire ; en effet, d'une récolte à l'autre, $\frac{\Delta q}{q}$ varie moins que $\frac{\Delta p}{p}$. Dans ces conditions, une récolte surabondante fait baisser les prix mais, aussi, la recette brute R puisque la fonction R en la variable p est croissante. Ce phénomène d'une récolte excellente et peu rémunératrice est bien connu : citons, pour le marché du blé, l'effet KING (secrétaire de la comptabilité publique dans le duché de Lancaster au 17e siècle) selon lequel les effets ne sont pas directement proportionnels aux causes ; il est aussi fréquent en Belgique, abstraction faite des complications comme un stockage spéculatif, des restrictions sur les exportations, etc. En guise d'exemple, rappelons que l'année 1982, qui fut très bonne dans la mesure où la production enregistra une hausse exceptionnelle de 7,4%, ne fit pas le bonheur des agriculteurs qui virent leur revenu brut, mais aussi leur revenu net car les coûts totaux restent pratiquement inchangés d'une année à l'autre, diminuer puisque le marché est inélastique (les prix unitaires baissant, dans ce cas, de plus de 10%).

2.1.4 Fonctions d'utilité

Un consommateur dispose d'un nombre fini n d'articles A_i en quantités respectives x_i (pour $i = 1, 2, \ldots, n$). Traditionnellement , on compte positivement les quantités x_i de biens effectivement consommés et négativement les quantités x_i de biens fournis (par exemple, le travail ou un service) par le consommateur. Le point $x = (x_1, x_2, \ldots, x_n)$ de $\mathbb{R}^n$ est appelé *programme* *de consommation* ou *complexe* (ou encore *panier) de biens*; l'ensemble X de tous les complexes possibles est l'*ensemble* (ou *espace) de consommation.*

Le consommateur est soumis à diverses contraintes.

Physiquement, tout panier x doit appartenir à un ensemble X donné *a priori* et dépendant éventuellement du consommateur considéré. Par exemple, si l'individu en question ne peut fournir aucune prestation, X sera simplement le sous-ensemble $\mathbb{R}^n_+$ de $\mathbb{R}^n$ composé par les points dont aucune coordonnée n'est négative.

Certaines hypothèses plus précises sur l'ensemble X sont quelquefois formulées; elles jouent le rôle de conditions suffisantes pour la validité des résultats, mais ne sont pas toujours nécessaires et peuvent être remplacées par d'autres dont le contenu est moins restrictif à certains points de vue.

L'ensemble X est le plus souvent supposé convexe, fermé, borné inférieurement et contenant l'origine O de $\mathbb{R}^n$; de plus, figure également dans X tout point y pour lequel existe un point x de X tel que $y_i \geqslant x_i$ pour $i = 1, 2, \ldots, n$.

Analysons un peu plus en détail ces conditions.

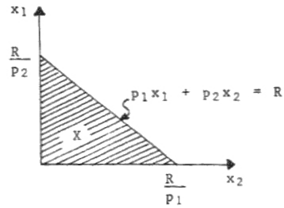
Les caractères convexe et fermé de X paraissent assez naturels et ne semblent pas restreindre la portée des résultats. Remarquons toutefois que la convexité n'est pas vérifiée notamment lorsque certains biens ne peuvent être consommés que par quantités entières et non de façon « continue ».

Le fait que X soit supposé borné inférieurement équivaut à l'existence d'un point x^* tel que $x_i \geqslant x_i^*$ pour tout $i = 1, 2, \ldots, n$, et pour tout x de X. Cette condition n'est pas restrictive puisque les quantités de travail fournies par le consommateur sont naturellement bornées supérieurement et que la consommation effective des autres biens ne peut pas être négative.

Supposer l'origine dans X signifie que l'individu peut se contenter d'une consommation nulle; cette hypothèse simplifie considérablement la théorie, mais est quelquefois peu réaliste puisqu'elle exclut l'existence d'un minimum vital (physique ou sociologique).

Enfin, la dernière condition traduit le fait que le consommateur a toujours la possibilité d'accepter un supplément de biens, même s'il ne doit rien en faire. En réalité, cette hypothèse, encore connue sous le nom de *libre disposition des excédents* , n'est pas toujours satisfaite. En effet, le consommateur dispose généralement d'un revenu R limité et doit acheter chaque produit A_i à un prix unitaire p_i bien défini. La valeur globale correspondant au complexe x ne peut pas dépasser R, ce qui se traduit par la *contrainte* *budgétaire* : $\sum_{i=1}^{n} p_i x_i \leqslant R$; cette inégalité délimite, dans l'espace $\mathbb{R}^n$ des consommations, un demi-espace fermé qui inclut X et qui admet l'hyperplan d'équation $\sum_{i=1}^{n} p_i x_i = R$ comme frontière.

Dans un modèle simple, X peut être regardé comme un polyèdre convexe qui est l'intersection de l'orthant non négatif $\mathbb{R}^n_+$ et du demi-espace fermé $\{x \in \mathbb{R}^n : \sum_{i=1}^{n} p_i x_i \leqslant R\}$. X est dans ce cas borné (donc compact), c'est-à-dire un polytope convexe, comme sur l'exemple de la figure ci-contre qui correspond à deux produits ($n = 2$).



Le consommateur doit choisir, parmi l'ensemble X, le complexe x qui lui semble le meilleur ; il doit dès lors être en mesure de décider quand il préfère un complexe à un autre.

Il convient donc d'établir un système de préférences du consommateur. A cet effet, il n'est pas nécessaire de se préoccuper des mobiles qui conduisent à ces préférences ; seule importe la vérification de certains axiomes qui traduisent une certaine cohérence interne des choix. Ces axiomes sont peu restrictifs et définissent en fait un préordre total noté $\succeq$, sur X : lorsque le complexe x sera préféré au complexe y, on écrira $x \succeq y$. Pour tout x de X, $x \succeq x$ (réflexivité) ; pour tous complexes x et y de X, on a $x \succeq y$ ou $y \succeq x$ (caractère total) ; enfin, si $x \succeq y$ et $y \succeq z$, alors $x \succeq z$ (transitivité).

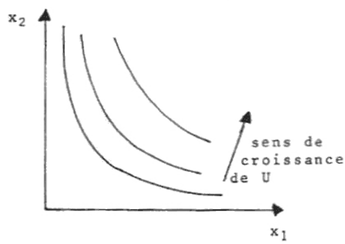
Ces conditions sont naturellement vérifiées si le système des préférences de l'individu est « logique » (en accord avec le concept d'*homo œconomicus*), mais peut se modifier au cours du temps en fonction de l'âge, de l'instruction et d'autres caractéristiques personnelles.

Enfin, une hypothèse supplémentaire et peu contraignante est parfois exigée : quel que soit le point x^* de X, l'ensemble $\{x \in X : x^* \succeq x\}$ de tous les x qui ne sont pas préférés à x^* et l'ensemble $\{x \in X : x \succeq x^*\}$ de tous les x auxquels x^* n'est pas préféré sont tous les deux fermés dans $\mathbb{R}^n$. Sous cette hypothèse, G. Debreu, économètre américain d'origine française qui reçut le prix Nobel en 1983, a démontré l'existence d'une fonction continue sur X qui mesure en quelque sorte l'utilité ou la satisfaction que retire le consommateur d'un complexe x [20] : il s'agit de la *fonction d'utilité* $U(x)$, à n variables $x_1, x_2, \ldots, x_n$, qui, à chaque programme x associe un nombre $U(x)$ (reflètant donc l'importance accordée par le consommateur au complexe x).

La fonction d'utilité U est généralement supposée continue, croissante et deux fois continûment dérivable. La continuité résulte des hypothèses formulées sur l'ensemble X de consommation, notamment de la condition énoncée par DEBREU. La croissance de U revient à dire que la valeur de $U(x)$ doit être au moins égale à $U(y)$ dès que le consommateur préfère le complexe x à y. Dès lors, chaque utilité marginale U_i' est habituellement positive : en effet, l'utilité du consommateur croît lorsque la quantité consommée de l'article A_i augmente, *ceteris paribus.* Par ailleurs, chaque utilité marginale U_i' est souvent décroissante, soit $U_{ii}'' < 0$: par exemple, la consommation du vingtième bonbon ne procure pas la même satisfaction que celle de la première friandise. Selon le théorème d'interversion de l'ordre des dérivées partielles, $U_{ij}'' = U_{ji}''$, ce qui signifie que la variation de l'utilité marginale du bien A_j provenant d'un petit changement dans la consommation de A_i est égale à la variation de l'utilité marginale de A_i pour une petite modification dans la consommation de A_j.

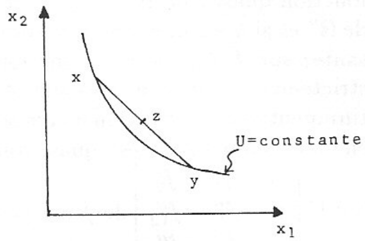
La fonction d'utilité U sert ainsi à classer les complexes suivant l'ordre choisi par le consommateur. En particulier, il peut exister plusieurs programmes différents qui donnent une même utilité : une plus grande quantité du produit A_i peut, par exemple, être associée à une plus faible quantité du produit A_j, de sorte que l'utilité globale reste inchangée ; l'ensemble des points x de X pour lesquels $U(x)$ prend une valeur constante est appelé *surface d'indifférence* (ou *surface d'égale utilité*) : il indique des assortiments de satisfaction égale ou des solutions offrant les mêmes avantages.

Il existe autant de surfaces d'indifférence que de valeurs atteintes par la fonction U ; deux complexes x et y appartiennent à la même surface d'indifférence si et seulement si le consommateur choisit indifféremment x ou y.



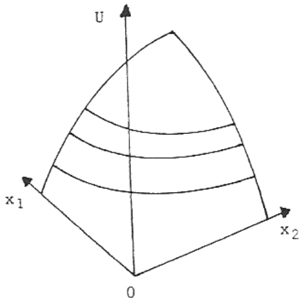
Pour $n = 2$, la représentation géométrique des surfaces (ici, des courbes) d'indifférence est facile à visualiser par un graphique sur lequel peut apparaître le sens de croissance de la fonction U. Souvent, la fonction U est prise strictement quasi-concave en ce sens que si $U(y) \geqslant U(x)$ pour deux complexes x et y différents, alors $U(z) > U(x)$ pour tout complexe z du segment $]x : y[$.

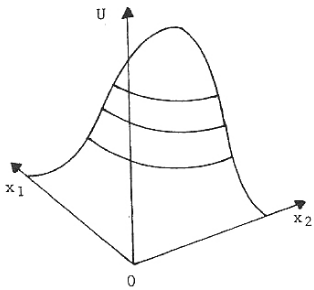
Cette hypothèse est assez naturelle, car un point z de $]x : y[$ présente, en quelque sorte, une composition intermédiaire entre celles de x et de y, donc, généralement, mieux équilibrée que celle de x ou celle de y.



Géométriquement, pour deux variables, cela signifie que les courbes d'indifférence ont leur concavité « tournée vers le haut » comme le montre la figure ci-contre.

Il est clair que toute fonction strictement concave est strictement quasi-concave, mais la réciproque de cette dernière propriété n'est pas toujours valable. Pour illustrer ces propos, analysons les graphes de ces deux fonctions U à deux variables x_1 et x_2 qui sont toutes les deux strictement quasi-concaves ; la première est strictement concave, tandis que le graphique de la deuxième fonction, en forme de cloche, possède des portions concaves et des portions convexes, ce qui signifie que les utilités marginales U_1' et U_2' peuvent être croissantes ou décroissantes.





Signalons encore que l'on emploie quelquefois l'hypothèse, moins forte, de *quasi-concavité* pour U, ce qui revient à supposer que si $U(y) \geqslant U(x)$ pour deux complexes x et y différents, alors $U(z) \geqslant U(x)$ pour tout complexe z appartenant au segment $]x : y[$. Cette hypothèse de quasi-concavité de l'utilité traduit concrètement le fait qu'un consommateur préfère généralement utiliser un peu de plusieurs biens plutôt que beaucoup d'un seul [45].

Donnons quelques exemples de fonctions quasi-concaves. Toute fonction concave (resp. strictement concave) sur un convexe A de $\mathbb{R}^n$ est quasi-concave (resp. strictement quasi-concave) sur A. Toute fonction monotone (resp. strictement monotone) d'une seule variable est quasi-concave (resp. strictement quasi-concave) sur tout intervalle où elle est définie. Si f est une fonction quasi-concave (resp. strictement quasi-concave) sur un convexe A de $\mathbb{R}^n$ et si h est une fonction numérique croissante (resp. strictement croissante) sur $f(A)$, alors la fonction composée hof est quasi-concave (resp. strictement quasi-concave) sur A. Une fonction $f(x_1, x_2)$, deux fois continûment dérivable sur un convexe ouvert A de $\mathbb{R}^2$ et telle que $f_1' > 0, f_2' > 0$ en tout point de A, est quasi-concave sur A si et seulement si le hessien bordé $\begin{vmatrix} 0 & f_1' & f_2' \\ f_1' & f_{11}'' & f_{12}'' \\ f_2' & f_{21}'' & f_{22}'' \end{vmatrix}$ de f est positif ou nul en tout point de A, ou encore si et seulement si toute courbe d'indifférence est convexe (c'est-à-dire tourne sa concavité vers le haut) [4, 7]. Pour illustrer ces résultats, mentionnons quelques expressions qui servent parfois à définir l'utilité pour deux variables (a, b, c étant des constantes positives adéquates) : $\frac{x_1+a}{b-\sqrt{x_2+c}}$, $(x_1+a)(x_2+b)$, $ax_1 + bx_2 + c\sqrt{x_1x_2}$, $ax_1^2 + 2bx_1x_2 + cx_2^2$ (avec $b^2 > ac$), $(x_1-a)^2 + (x_2-b)^2$ (avec $0 \leqslant x_1 \leqslant a$ et $0 \leqslant x_2 \leqslant b$), $a \ln x_1 + b \ln x_2$, $x_1^a x_2^b$ [2, pp.127, 293, 294].

Remarquons que cette dernière fonction, dite de Cobb-Douglas, est toujours strictement quasi-concave (pour $x_1 > 0$ et $x_2 > 0$) puisque $\ln(x_1^a x_2^b) = a \ln x_1 + b \ln x_2$ est visiblement strictement concave, mais qu'elle n'est concave que si et seulement si $a + b \leqslant 1$.

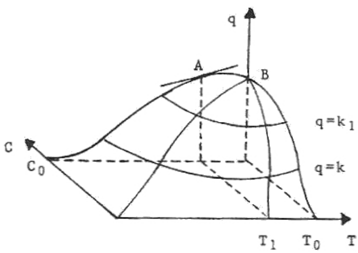
2.1.5 Fonctions de production

Le processus de fabrication d'un produit (ou *output*) à l'aide de plusieurs facteurs de production (ou *inputs*), tels que le travail, le capital, les matières premières, ..., décrit une fonction qui définit la quantité q de l'output la plus élevée qu'il est possible d'obtenir à partir des quantités données $x_1, x_2, \ldots, x_n$ des n inputs $P_1, P_2, \ldots, P_n$: $q = f(x_1, x_2, \ldots, x_n)$ porte le nom de *fonction de production* et sera supposée au moins deux fois continûment dérivable.

Il est aisé de voir que cette fonction doit être généralement croissante avec les quantités x_i des inputs ; cela entraîne le caractère positif de chaque productivité marginale f_i'. De plus, on admet souvent que la productivité marginale f_i' d'un facteur P_i est accrue si la quantité utilisée d'un autre facteur P_j est augmentée, ce qui rend positive chaque dérivée partielle seconde croisée f_{ij}'' : cette hypothèse est quelquefois connue sous le nom de *loi de* Wicksell [44]. Enfin, la fonction f est presque toujours prise concave, ou à tout le moins (strictement) quasi-concave. D'un point de vue formel, la fonction de production joue donc dans la théorie de la firme un rôle analogue à celui tenu par la fonction d'utilité dans la théorie du consommateur.

Un modèle simple ne retient que deux facteurs de production : le travail T et le capital C. La fonction de production s'écrit alors $q = f(T, C)$, où q désigne la quantité maximale d'output produit (à l'aide de T et C). Géométriquement, la fonction de production $q = f(T, C)$ peut être décrite par la *surface de production* qui est son graphique dans l'espace à trois dimensions.

La figure ci-contre est relative à une telle fonction représentée sur un sous-ensemble (le rectangle de sommets O, C_0, D, T_0) de son domaine de définition (le premier quadrant). Fixons le niveau du capital à une valeur C_o et considérons uniquement les variations de l'input T.



L'intersection de la surface de production par le plan d'équation $C = C_o$ détermine une courbe passant par les points C_o, A, B ; par exemple, le coefficient angulaire de la tangente au point A, de coordonnées T_1 et C_o, est donné par la dérivée partielle de f par rapport à T pour $C = C_o$ et $T = T_1$: il mesure le taux de variation de la quantité q en fonction d'un changement du travail T (autour de la valeur T_1, avec C constant, égal à C_o). Une même interprétation géométrique pourrait être donnée pour la dérivée partielle de f par rapport à C.

La fonction de production la plus célèbre est sans doute la *fonction de* COBB-DOUGLAS qui est de la forme suivante : $q = KC^aT^b$, où K, a et b sont des constantes positives. Cette fonction est homogène de degré $a + b$. Un passage aux logarithmes conduit à : $\ln q = \ln K + a. \ln C + b. \ln T$. Cette égalité montre qu'une variation de 1% de C (resp. de T), avec T (resp. C) constant, entraîne une variation de a% (resp. b%) de la quantité q ; de plus, une variation de 1% de T et de C entraîne une variation de $(a + b)$% de q. Dans certains cas est imposée la condition supplémentaire $a + b = 1$, ce qui signifie concrètement qu'il n'existe pas d'économies de dimension.

Signalons l'existence de nombreux autres modèles de production. Ainsi, certains économistes adjoignent à la fonction de COBB-DOUGLAS un facteur e^{ct} qui tient compte de l'évolution des techniques au cours du temps t : $q = KC^aT^be^{ct}$. L'introduction de ce facteur de *trend* (mot anglais signifiant mouvement de longue durée dans l'évolution d'un phénomène économique) est susceptible de deux interprétations : avec un stock de capital fixé et le même nombre de travailleurs, l'efficacité tend à augmenter avec le temps grâce à une meilleure connaissance de l'outillage (effet du *« learning by doing »*), tandis que les nouvelles techniques permettent de produire plus que les anciennes avec le même nombre de travailleurs et un équipement de même valeur (effet du « progrès de la technologie »).

Les fonctions de production ne sont pas toujours du type de COBB-DOUGLAS, mais sont néanmoins souvent supposées homogènes, les degrés d'homogénéité étant intimement reliés aux rendements à l'échelle (voir [44, pp.124-128]).

En raison de leur simplicité et de leurs propriétés remarquables, les fonctions *linéairement homogènes*, c'est-à-dire les fonctions homogènes de degré 1, jouent un rôle important dans la théorie de la production.

Donnons quelques exemples de fonctions de production linéairement homogènes du type $q = f(x_1, x_2)$, où q désigne la quantité d'output pour les quantités utilisées x_1 et x_2 des deux inputs, les coefficients $a, b, \ldots$ et les exposants α, β étant des constantes adéquates :

$$\begin{aligned}
q &= ax_1 + bx_2 \text{(fonction linéaire)}, \\
q &= ax_1^{\alpha} x_2^{1-\alpha} \text{(fonction de COBB-DOUGLAS)}, \\
q &= a[bx_1^{-\alpha} + (1-b)x_2^{-\alpha}]^{\frac{-1}{\alpha}} \text{ (fonction C.E.S.} \\
&\quad \text{ ou fonction à élasticité de substitution constante)}, \\
q &= \sqrt{ax_1^2 + bx_1x_2 + cx_2^2} \\
q &= \frac{ax_1^2 + bx_1x_2 + cx_2^2}{dx_1 + ex_2} \\
q &= \frac{ax_1^{\alpha} x_2^{1+\beta-\alpha}}{bx_1^{\beta} + cx_2^{\beta}} \\
q &= \frac{ax_1^{\alpha} x_2^{\beta} + bx_1^{\beta} x_2^{\alpha}}{cx_1^{\alpha+\beta-1} + dx_2^{\alpha+\beta-1}}
\end{aligned}$$

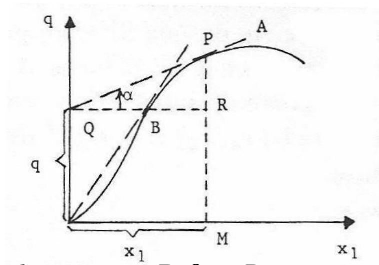
[44, p.98]

On vérifiera sans peine sur ces quelques cas particuliers la propriété fondamentale des fonctions linéairement homogènes, à savoir que la quantité produite q est multipliée par un coefficient t dès que chaque variable x_1 et x_2 est également multipliée par ce même facteur t.

Donnons d'autres propriétés intéressantes des fonctions de production linéairement homogènes (que nous supposerons au moins deux fois continûment dérivables).

Les valeurs moyennes $\frac{q}{x_1}$ et $\frac{q}{x_2}$, ainsi que les valeurs marginales f_1' et f_2' sont homogènes de degré 0, et peuvent être exprimées comme des fonctions du rapport $\frac{x_2}{x_1}$. On a de plus l'égalité suivante : $x_1 f_1' + x_2 f_2' = q$; ce résultat, connu sous le nom de *théorème d'Euler*, appelle des commentaires intéressants d'un point de vue économique. En effet, cette formule signifie que l'output global est égal à la somme des produits de la quantité de chaque facteur par sa productivité physique marginale. Dès lors, si l'entrepreneur paie les offreurs d'input suivant leur productivité marginale, la totalité de l'output est ainsi épuisée.

Le théorème d'EULER peut aussi être interprété géométriquement en considérant, par exemple, la courbe de produit total relative au premier facteur de production.



Soient P un point de cette courbe, M sa projection orthogonale sur l'axe des abscisses, Q l'intersection avec l'axe des ordonnées de la tangente menée par P à cette courbe et R le point situé à la même abscisse que P et à la même ordonnée que Q (voir la figure ci-contre).

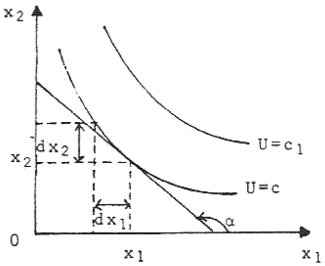
En travaillant dans le triangle rectangle de sommets P, Q et R, on trouve $|RP| = |QR| \operatorname{tg} \alpha = x_1 f_1'$, puisque $|QR| = x_1$ et $f_1' = \operatorname{tg} \alpha$. L'égalité $x_1 f_1' + x_2 f_2' = q$ entraîne $x_2 f_2' = q - x_1 f_1' = |MP| - |RP| = |MR|$; en conséquence, le produit total $|MP|$ est divisé en deux parties : $|RP| = x_1 f_1'$ et $|MR| = x_2 f_2'$. Au point A, la courbe présente un maximum, de sorte que $f_1'(A) = 0$, d'où $f_2'(A) = \frac{q}{x_2}$ en vertu de la formule d'EULER ; en ce point, la productivité physique moyenne du second facteur présente un maximum sur la section de la surface de production perpendiculairement à l'axe Ox_1 en effet, $\frac{\partial}{\partial x_2}(\frac{q}{x_2}) = 0$ équivaut à $\frac{x_2 f_2' - q}{x_2^2} = 0$, soit à $q = x_2 f_2'$. De même, au point B de la courbe, la productivité physique moyenne présente un maximum (car $f_1' = \frac{q}{x_1}$), d'où $f_2' = 0$, ce qui implique que le produit total est alors maximum sur la section de la surface de production perpendiculairement à l'axe $0x_1$. Ainsi, il existe une correspondance entre le maximum de la production totale pour la variation d'un facteur et le maximum de la productivité moyenne pour la variation de l'autre facteur.

2.2 Exemples de fonctions implicites

2.2.1 Courbe d'indifférence

Dans la théorie du consommateur, l'utilité n'est pas toujours mesurable de façon précise, mais est seulement une grandeur « *repérable* » ; par exemple, la fonction d'utilité permet de déclarer tel bien plus utile qu'un concurrent, mais pas de reconnaître une consommation deux fois plus utile qu'une autre. On parle dans ce cas de satisfaction (ou utilité) *ordinale* et la fonction U est alors définie à une fonction croissante près, puisque toute fonction obtenue en composant U avec une fonction monotone croissante représente les mêmes préférences du consommateur.

Plaçons-nous dans le cas où un consommateur peut acheter deux articles A_1 et A_2 en quantités respectives x_1 et x_2. L'ensemble de toutes les combinaisons des deux biens A_1 et A_2 pour lesquelles ce client obtient le même niveau de satisfaction constitue une *courbe d'indifférence* d'équation implicite $U(x_1, x_2) = c$, où U désigne la fonction d'utilité et c est une constante.



Par chaque point du premier quadrant, il passe une et une seule courbe d'indifférence : celle-ci correspond à des niveaux de satisfaction de plus en plus élevés au fur et à mesure que l'on progresse dans la direction du Nord-Est. Le long d'une courbe d'indifférence d'équation $U(x_1, x_2) = c$, on a bien entendu $dU = U_1' dx_1 + U_2' dx_2 = 0$.

En conséquence, à satisfaction égale, si la consommation A_1 décroît de dx_1 (avec $dx_1 < 0$), il en résultera une perte d'utilité, approximativement égale à $U_1' dx_1$, perte qui sera compensée par un gain, approximativement égal à $U_2' dx_2$, obtenu par une augmentation dx_2 de la quantité de A_2. La pente $m = \frac{dx_2}{dx_1} = \operatorname{tg} \alpha$ de la tangente en un point (x_1, x_2) de la courbe d'indifférence est le taux pour lequel le consommateur voudra substituer A_1 à A_2 ou A_2 à A_1, tout en restant à un même niveau donné de satisfaction ; l'opposé de cette pente m est le *taux de substitution des biens*, égal au rapport des utilités marginales : $m = -\frac{U_1'}{U_2'}$. De plus, chaque courbe d'indifférence est convexe, ce qui revient à dire que la dérivée seconde de x_2 par rapport à x_1 n'est pas négative ; de fait, on a $\frac{d^2x_2}{dx_1^2} = \frac{\overline{H}}{(U_2')^3}$, où $\overline{H}$ désigne le hessien bordé $\begin{vmatrix} 0 & U_1' & U_2' \\ U_1' & U_{11}'' & U_{12}'' \\ U_2' & U_{21}'' & U_{22}'' \end{vmatrix}$. La conclusion découle de la croissance de l'utilité, ainsi que de son caractère quasi-concave.

2.2.2 Isoquante

Une *isoquante* est l'ensemble de toutes les combinaisons possibles des inputs donnant un niveau fixé d'output. Formellement, cette notion joue le même rôle pour l'entreprise qu'une courbe d'indifférence dans la théorie du consommateur.

Considérons le cas de la fabrication d'un seul produit à l'aide de deux facteurs. Une isoquante est donc une courbe d'équation implicite $f(x_1, x_2) = c$, où f désigne la fonction de production, x_1 et x_2 sont les quantités des deux inputs et c est une constante. Géométriquement, cette courbe est la projection orthogonale dans le plan horizontal de base (d'équation $q = 0$) de la courbe qui est l'intersection de la surface de production et du plan horizontal d'équation $q = c$ (voir la figure dans le paragraphe 2.1.5). Bien entendu, plus l'isoquante est éloignée de l'origine, plus le niveau d'output qu'elle représente est élevé.

La pente de la tangente en un point d'une isoquante est le taux auquel un facteur de production doit être substitué à l'autre pour maintenir le même niveau d'output. L'opposé de cette pente est le *taux de substitution technique* : il est égal au rapport $\frac{f_1'}{f_2'}$ des productivités marginales des deux inputs.

Généralement, une isoquante est une courbe (strictement) convexe, la fonction de production étant le plus souvent (strictement) quasi-concave.

Par exemple, dans le cas d'une fonction de production du type Cobb-Douglas, toute isoquante d'équation implicite $Kx_1^a x_2^b = c$, est décroissante et strictement convexe (pour $x_1 > 0$ et $x_2 > 0$) ; en effet, le théorème de dérivation des fonctions implicites livre $\frac{dx_2}{dx_1} = -\frac{ax_2}{bx_1} < 0$, tandis que $\frac{d^2x_2}{dx_1^2} = -\frac{a}{bx_1^2}(x_1 \frac{dx_2}{dx_1} - x_2) > 0$, puisque $x_1 > 0$, $x_2 > 0$, $a > 0$, $b > 0$ et $\frac{dx_2}{dx_1} < 0$.

Le long d'une isoquante, l'intensité des facteurs (égale à $\frac{x_2}{x_1}$) se modifie au fur et à mesure que change le taux de substitution technique. L'importance d'une telle substitution est quelquefois mesurée par l'*élasticité de substitution*, introduite par HICKS et ROBINSON (4) et définie par

$$\sigma = \frac{d\ln\frac{x_2}{X_1}}{d\ln r_{21}}, \text{ avec } r_{21} = \frac{f_1'}{f_2'};$$

il est d'autant plus facile pour la firme de remplacer un facteur par l'autre (tout en conservant le même volume de production) que l'élasticité de substitution est élevée. On peut vérifier que l'élasticité de substitution pour une fonction de COBB-DOUGLAS est égale à 1, même si la somme $a+b$ des exposants est différente de 1. Plus généralement, la fonction de production à élasticité de substitution constante ou fonction C.E.S., qui a été étudiée notamment par R. M. SOLOW (prix Nobel en 1987) et peut s'écrire sous la forme

$$q = K[ax_1^{-b} + (1-a)x_2^{-b}]^{-\frac{c}{b}}$$

avec $c > 0$, se caractérise par une valeur constante (pas nécessairement égale à 1) de l'élasticité de substitution et généralise de ce fait la fonction de COBB-DOUGLAS (qui en est, en réalité, un cas limite).

2.2.3 Chemin d'expansion et classification des facteurs de production

Considérons une firme produisant une quantité q prédéterminée d'un output à l'aide de deux inputs P_1 et P_2 utilisés en quantités respectives x_1 et x_2.

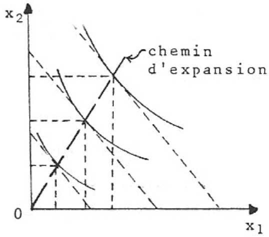
On appelle *chemin d'expansion* de la firme l'ensemble des points de tangence entre les isoquantes et les droites d'isocoût. Si p_1 et p_2 désignent les prix unitaires des deux inputs, une droite d'isocoût admet pour équation $p_1x_1 + p_2x_2 = k$, où k désigne une constante. En conséquence, le chemin d'expansion admet l'équation implicite $p_1f_1' + p_2f_2' = 0$, où f est la fonction de production.

On peut démontrer qu'un entrepreneur rationnel ne choisira que les quantités d'inputs se trouvant sur ce chemin d'expansion.

(4) Hicks J.R., *Theory of wages*, Londres, 1963, pp. 117, 244-245. Robinson J., *The Economics of Imperfect Competition*, Londres, 1969, pp. 256-260

Par exemple, si la production est donnée par la fonction de COBB-DOUGLAS $q = Kx_1^a x_2^b$, chaque point du chemin d'expansion doit satisfaire aux égalités suivantes :

$$\frac{p_1}{p_2} = \frac{f_1'}{f_2'} = \frac{Kax_1^{a-1}x_2^b}{Kbx_1^a x_2^{b-1}} = \frac{ax_2}{bx_1},$$



d'où $x_2 = \frac{b}{a}\frac{p_1}{p_2}x_1$. En conséquence, le chemin d'expansion est alors une droite passant par l'origine et de coefficient angulaire $\frac{bp_1}{ap_2}$ (voir la figure ci-contre). Notons encore que le chemin d'expansion est une droite pour toute fonction de production homogène.

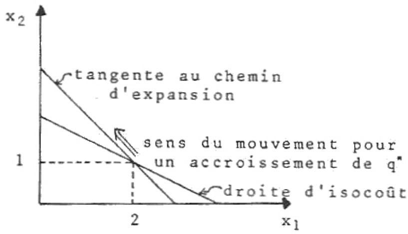
L'étude du chemin d'expansion permet une classification des facteurs de production.

Dans la théorie de la firme, un facteur est dit *normal* ou *supérieur* (resp. *inférieur*) si l'augmentation du volume de production provoque un accroissement (resp. une diminution) de l'utilisation du facteur.

Examinons le cas d'une entreprise qui produit un seul output (en quantité q) à l'aide de deux inputs P_1 et P_2. On sait que les quantités utilisées x_1 et x_2 de chaque facteur P_1 et P_2 déterminent un point situé sur le chemin d'expansion de la firme. Soient $q = f(x_1, x_2)$ la fonction de production considérée et $p_1x_1 + p_2x_2 = k$ l'équation d'une droite d'isocoût, k désignant le coût global de production, p_1 et p_2 les prix unitaires respectifs des deux facteurs P_1 et P_2. Nous admettrons que la productivité marginale de chaque facteur est positive, ce qui est généralement le cas.

Considérons un point $x^* = (x_1^*, x_2^*)$ du chemin d'expansion, point situé à l'intersection de l'isoquante d'équation implicite $f(x_1, x_2) = q^* = f(x_1^*, x_2^*)$ et de la droite d'isocoût d'équation $p_1x_1 + p_2x_2 = k^* = p_1x_1^* + p_2x_2^*$. En vertu du théorème d'existence des fonctions implicites, il existe, au voisinage du point x^*, une et une seule courbe d'équation explicite $x_2 = h(x_1)$ qui passe par x^* et coïncide avec le chemin d'expansion. Le théorème de dérivation des fonctions implicites permet de calculer le coefficient angulaire m^* de la tangente à cette courbe $x_2 = h(x_1)$ au point x^*, à savoir $m = \frac{p_1 f_{21}'' - p_2 f_{11}''}{p_2 f_{12}'' - p_1 f_{22}''}$, les dérivées partielles de f étant toutes prises au point x^*, convention que nous maintiendrons ultérieurement ; puisque $\frac{f_1'}{p_1} = \frac{f_2'}{p_2}$, on peut encore écrire $m = \frac{f_1' f_{21}'' - f_2' f_{11}''}{f_2' f_{12}'' - f_1' f_{22}''}$.

Comme exemple numérique ([5]), prenons la fonction de production $q = f(x_1, x_2) = \frac{x_1 - 1}{(x_2-2)^2}$ pour $x_1 > 1$ et $x_2 < 2$. Au point $x^* = (2, 1)$, on a $\frac{\partial f}{\partial x_1}(x^*) = 1$, $\frac{\partial f}{\partial x_2}(x^*) = 2$, $\frac{\partial^2 f}{\partial x_1^2}(x^*) = 0$, $\frac{\partial^2 f}{\partial x_1 \partial x_2}(x^*) = 2$ et $\frac{\partial^2 f}{\partial x_2^2}(x^*) = 6$, d'où $m^* = -1$, tandis que la droite d'isocoût passant par x^* est d'équation $x_2 = -\frac{1}{2}x_1 + 2$.

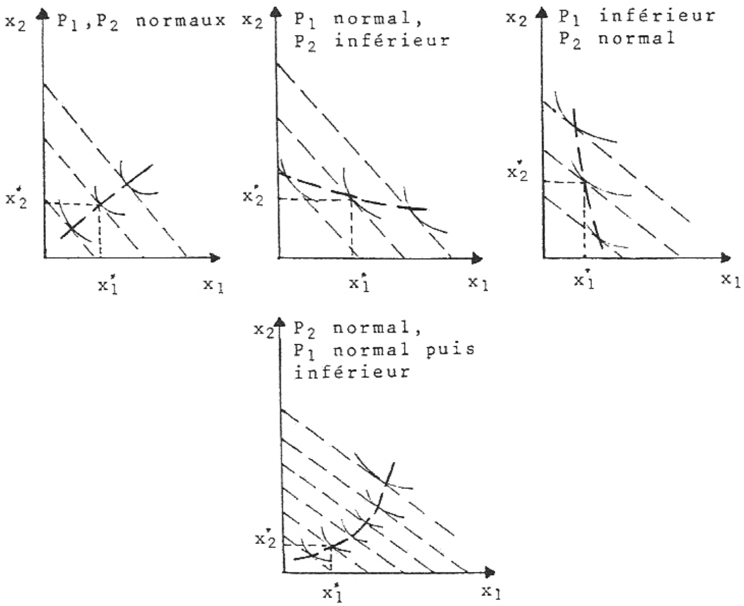


Un accroissement de la quantité produite autour de q^* va déplacer le point x^* le long du chemin d'expansion. En première approximation, ce mouvement peut se faire en partant du point x^* et en suivant la tangente au chemin d'expansion (c'est-à-dire la droite d'équation $x_1 + x_2 = 3$) de bas en haut et de la droite vers la gauche. Ainsi, une petite augmentation du volume de production autour de q^* provoque visiblement une diminution du facteur P_1, ce qui montre que P_1 est alors inférieur ; ce résultat est d'ailleurs confirmé par le caractère négatif de la dérivée partielle par rapport à x_2 du taux de substitution technique en x^* ; de fait, $\frac{\partial}{\partial x_2}(\frac{f_1'}{f_2'})(x^*) = -\frac{1}{2}$.

D'une façon générale, pour de petites variations du volume de production autour de la valeur q^*, chaque facteur est normal si m^* est strictement positif ; par contre, si m^* est strictement négatif, un facteur est normal et l'autre est inférieur : de façon plus précise, le facteur P_1 est normal et P_2 est inférieur si $0 > m^* > -\frac{p_1}{p_2}$, tandis que P_1 est inférieur et P_2 est normal si $m^* < -\frac{p_1}{p_2}$. Il est également possible de vérifier que le facteur P_1 (resp. P_2) est normal si et seulement si $\frac{\partial}{\partial x_2}(\frac{f_1'}{f_2'})(x^*) > 0$ (resp. $\frac{\partial}{\partial x_1}(\frac{f_1'}{f_2'})(x^*) < 0$).

Les trois premières figures ci-dessous illustrent les différents cas possibles, tandis que la quatrième rappelle le caractère local de tous ces résultats (en effet, les deux facteurs sont normaux au voisinage de x^* car m^* est positif, mais, par la suite, P_1 devient inférieur tandis que P_2 reste normal).

([5]) Cet exemple est dû à Ferguson C.E, *The Neoclassical Theory of Production and Distribution*, Cambridge University Press, 1969, p. 199.



2.3 Applications de l'intégration

2.3.1 Impôt sur le revenu

L'impôt belge sur le revenu des personnes physiques est une imposition dont le taux est progressif et varie par tranches.

En 1988, pour un revenu net imposable globalement de 113 000 F (obtenu en 1987), l'impôt de base s'élève à 300 F. Pour un revenu net (imposable globalement en 1987) supérieur à 113 000 F, l'impôt se calcule selon le tableau ci-dessous.

plus de	mais pas plus de	l'impôt de base dû s'élève à	plus sur l'excédent
113 000 F	206 500 F	300 F	24,1 %
206 500 F	258 000 F	22 834 F	27,8 %
258 000 F	309 500F	37 151F	35,8 %
309 500 F	413 000 F	55 588 F	39,595%
413 000 F	516 500 F	96 568 F	42,8 %
516 500 F	775 000 F	140 866 F	45,3 %
775 000 F	1 033 000 F	257 967 F	46,9 %
1 033 000 F	1 550 000 F	378 969 F	51,9 %
1 550 000 F	2 067 000 F	647 292 F	56,8 %
2 067 000 F	3 100 000 F	940 948 F	62,2 %
3 100 000 F	4 133 000 F	1 583 474 F	68,2 %
4 133 000 F	14 549 247 F	2 287 980 F	71,2 %
14 549 247 F	—	9 704 348 F	66,7 %

Taux d'imposition 1988

L'impôt à payer s'obtient en ajoutant à l'impôt de base, fixé par la limite inférieure de la tranche considérée, l'impôt sur l'excédent (qui est égal au produit du taux relatif à la tranche par la différence entre le revenu réel et la limite inférieure de la tranche).

Par exemple, pour un revenu imposable de 1 265 500 F, l'impôt de base sur 1 033 000 F est de 378 969 F, et sur l'excédent (c'est-à-dire sur 229 500 F), il est de $0{,}519 \times 229\,500$ F$=$ 119 500,5 F, arrondi à 119 111 F ; au total, l'impôt est donc de 378 969 F + 119 111 F = 498 080 F [6].

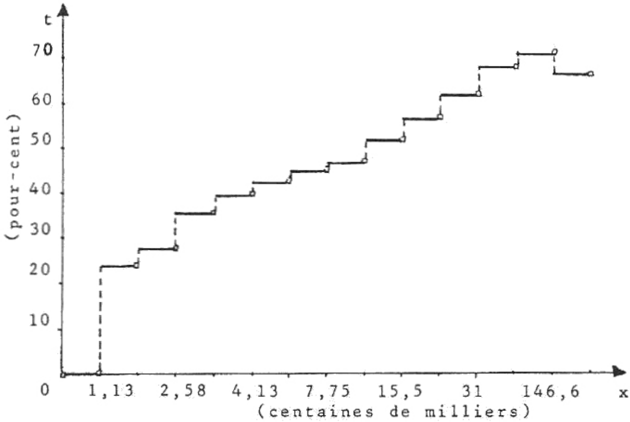
Il convient de remarquer que le taux d'imposition est marginal et non global : en effet, il s'applique uniquement à la dernière tranche du revenu et non au revenu tout entier.

Le taux t d'imposition est une fonction du revenu x ; il s'agit d'une fonction étagée (ou en escalier) dont le graphique, pour les revenus de 1987, est donné par la figure ci-dessous.

---

[6] Ces données et cet exemple sont tirés du livre « *L'almanach 1988 du contribuable* », Elsevier, RTL Edition.

LE TAUX D'IMPOSITION
en fonction de la tranche de revenu



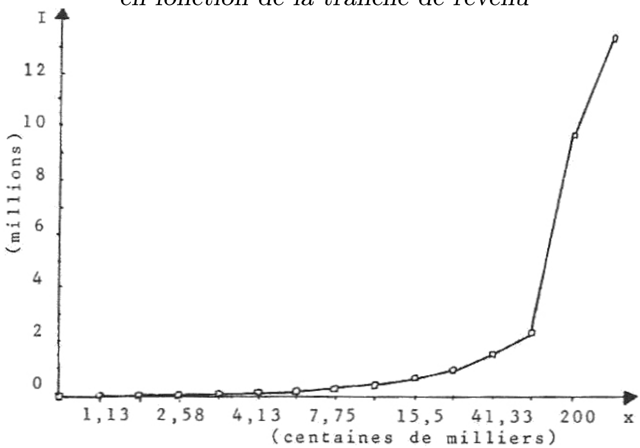
Il convient de constater que l'impôt de base coïncide (parfois à un léger arrondi près) à l'impôt global relatif à la borne supérieure de la tranche située immédiatement en-dessous. Calculons le taux théorique τ d'imposition pour que l'impôt global soit une fonction continue du revenu x.

Tranche	Taux théorique en %
$x \leqslant 113\ 000$	0,265486...
$113\ 000 < x \leqslant 206\ 500$	24,1
$206\ 500 < x \leqslant 258\ 000$	27,8
$258\ 000 < x \leqslant 309\ 500$	35,8
$309\ 500 < x \leqslant 413\ 000$	39,594202...
$413\ 000 < x \leqslant 516\ 500$	42,8
$516\ 500 < x \leqslant 775\ 000$	45,3001934...
$775\ 000 < x \leqslant 1\ 033\ 000$	46,9
$1\ 033\ 000 < x \leqslant 1\ 550\ 000$	51,9
$1\ 550\ 000 < x \leqslant 2\ 067\ 000$	56,8
$2\ 067\ 000 < x \leqslant 3\ 100\ 000$	62,2
$3\ 100\ 000 < x \leqslant 4\ 133\ 000$	68,2
$4\ 133\ 000 < x \leqslant 14\ 549\ 247$	71,2
$x > 14\ 549\ 247$	66,7

Ces taux théoriques sont très proches des valeurs correspondantes de la fonction t.

En travaillant avec ces nouvelles données, l'impôt est une fonction continue du revenu ; plus précisément, l'impôt $I(x)$ à payer pour un revenu x est de $\int_0^x \tau(u)\, du$ (à un éventuel arrondi près). Cette fonction $I(x)$ est affine par morceaux ; en voici le graphique :

L'IMPOT
en fonction de la tranche de revenu

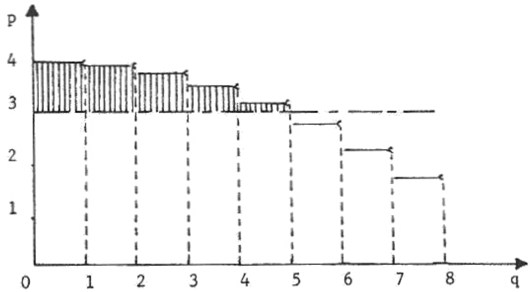


Il est intéressant de comparer les deux figures ci-dessus. Pour une même abscisse x, l'ordonnée $I(x)$ sur la deuxième figure représente le même nombre que l'aire délimitée, sur la première figure, par le graphique de la fonction, l'axe horizontal et les verticales d'abscisses 0 et x.

2.3.2 Surplus du consommateur et du producteur

Considérons en premier lieu le marché (fictif) d'un bien pour lequel est connue la loi de la demande : lorsque la quantité q demandée augmente, le prix unitaire p diminue (et réciproquement) selon la formule $25p = 100 - [E(q) + 0,5]^2$, où $E(q)$ désigne le plus grand entier contenu dans q. Les différentes possibilités de prix, pour une commande inférieure à 8 unités, sont représentées dans le tableau et sur la figure ci-dessous.

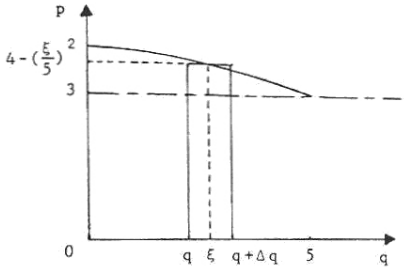
q		p
de	à moins de	
0	1	3,99
1	2	3,91
2	3	3,75
3	4	3,51
4	5	3,19
4	6	2,79
6	7	2,31
7	8	1,75



Soit un consommateur qui désire acheter 5 unités du produit au prix unitaire fixé à 3 u.m., prix qui est inférieur à celui dicté par la loi de la demande (pour $q \leqslant 5$) ; le client réalisera de la sorte un gain, égal respectivement à 0,99 (resp. 0,91 ; 0,75 ; 0,51 et 0,19) u.m. sur la première (resp. deuxième ; troisième ; quatrième et cinquième) unité acquise, soit un bénéfice net de 0,99+0,91+0,75+0,51+0,19 = 3,35 u.m. ; ce nombre 3,35 peut être mesuré par l'aire de la région hachurée sur la figure ci-dessus.

D'habitude, la demande n'est pas donnée par une fonction étagée, mais bien par une fonction continue. Modifions légèrement notre exemple introductif et envisageons à présent le cas plus « réaliste » pour lequel la demande obéit à la loi $25p = 100 - q^2$: nous allons analyser le comportement d'un consommateur qui souhaite encore acheter 5 unités de ce bien. Admettons que le client commande une petite quantité additionnelle Δq après avoir effectivement acheté une quantité q (avec $q + \Delta q < 5$). D'après la loi de la demande, le consommateur doit s'attendre à payer, pour cet achat supplémentaire, une somme approximativement égale à $[4 - (\frac{\xi}{5})^2]\Delta q$, où ξ désigne un nombre compris entre q et $q + \Delta q$ (voir figure ci-dessous). Par ailleurs, si le prix unitaire reste fixé à 3 u.m., le consommateur enregistre alors un gain donné par $[4 - (\frac{\xi}{5})^2 - 3]\Delta q = [1 - (\frac{\xi}{5})^2]\Delta q$. La somme de tous ces petits profits (maximaux), lorsque les quantités additionnelles Δq tendent vers 0 tandis que leur nombre croît indéfiniment, est appelée le *surplus du consommateur*.

Il s'agit en fait de la différence entre le montant global que le consommateur est prêt à payer pour obtenir $q_0 = 5$ unités du bien dans un régime monopolistique (7) et le prix total qu'il donnerait pour recevoir cette même quantité q_0 dans un marché parfaitement concurrentiel. (8)



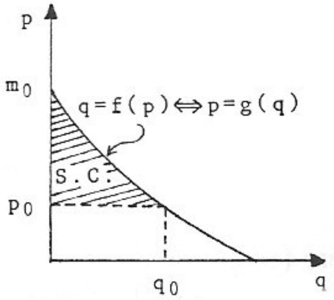
Ce surplus est mesuré par l'aire comprise entre le graphique de la fonction $p = 4 - \left(\frac{q}{5}\right)^2$, la droite horizontale d'équation $p = 3$, l'axe vertical $q = 0$ et la droite verticale d'abscisse $q = 5$, soit $\int_0^S [1 - (\frac{q}{5})^2] dq = \frac{10}{3}$ u.m.. Il convient de noter que cette dernière valeur est assez proche de la réponse obtenue dans le cas où la fonction de demande présentait un graphique composé de « paliers horizontaux ».

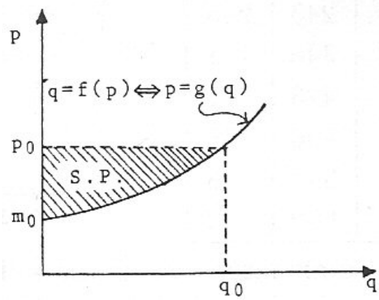
D'une manière tout à fait générale, la fonction de demande (resp. d'offre) détermine les quantités d'un produit qui doivent être achetées (resp. mises sur le marché) lorsque le prix varie.

Dans le marché d'un certain bien, supposons le prix unitaire parfaitement déterminé, de valeur p_0, et la quantité correspondante achetée (resp. offerte) égale à q_0 ; par exemple, dans un régime de concurrence parfaite, p_0 et q_0 peuvent respectivement coïncider avec le prix et la quantité assurant l'équilibre du marché, c'est-à-dire avec les coordonnées du point d'intersection des courbes d'offre et de demande. Toute personne qui compte acheter (resp. offrir) sa marchandise à un prix supérieur (resp. inférieur) à p_0 réalisera évidemment un gain. Sous certaines conditions économiques, le gain total de cette personne est représenté par l'aire de la région située en dessous (resp. au-dessus) de la courbe de demande (resp. de la courbe d'offre) et au-dessus (resp. en dessous) de la droite d'équation $p = p_0$. Cette aire a été appelée par MARSHALL le *surplus du consommateur* S.C. (resp. le *surplus du producteur* S.P.).

(7) Dans un *marché de monopole*, un seul offreur est confronté à une multitude de demandeurs. Le prix est fixé par le monopoleur et est univoquement déterminé par la quantité demandée (grâce à la loi de la demande).

(8) Dans un *régime de concurrence parfaite*, le marché comprend un grand nombre d'offreurs vendant un produit homogène ; chaque unité du bien est mise sur le marché au même prix, car un producteur sait que les prix de tous les biens ne seront pas influencés par les actions d'une entreprise considérée isolément ; ce prix fixe est déterminé par la loi de la demande appliquée à la dernière unité considérée.





Avec les notations indiquées sur les figures ci-dessus, on obtient les formules suivantes :

$$S.C. = \int_0^{q_0} g(q)dq - p_0q_0 = \int_{p_0}^{m_0} f(p)dp$$

et

$$S.P. = p_0q_0 - \int_0^{q_0} g(q)dq = \int_{m_0}^{p_0} f(p)dp$$

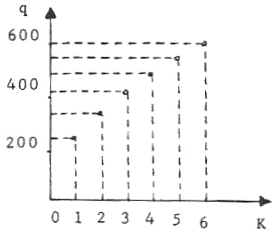
2.3.3 Productivité physique marginale du capital

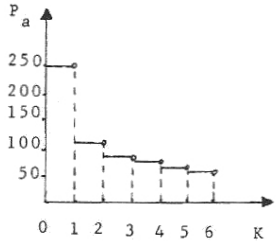
Nous allons étudier la variation du volume q de production d'un seul bien (ou output) en fonction, par exemple, du capital K, toutes choses égales par ailleurs.

A partir d'un exemple concret (9), nous nous proposons de relier le volume q de production à la productivité physique marginale, qui sera notée P_a et qui représente, rappelons-le, le produit additionnel résultant d'une unité supplémentaire de capital, *ceteris paribus*. Le tableau ci-dessous comprend les quantités q produites en fonction du nombre K des unités du capital, toutes choses égales par ailleurs, ainsi que les productivités physiques marginales P_a correspondantes ; ces mesures peuvent être visualisées sur les grpahiques donnant respectivement q et P_a en fonction de K.

(9) Ces données sont dues à P. Samuelson dans son livre « *L'économique, techniques modernes d'analyse économique* », Colin, Paris, 1957, pp. 545-549

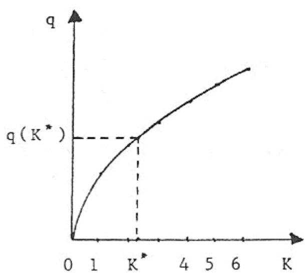
K	q	P_a
1	245	245
2	346	101
3	423	77
4	490	67
5	548	58
6	600	52

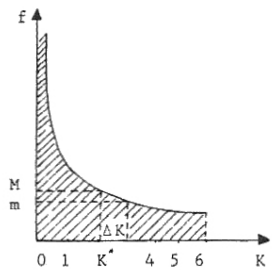




Commentons ces données. Pour la première (resp. deuxième, troisième, ...) unité de capital, la quantité totale produite augmente de 245 (resp. 346-245=101, 423-346=77, ...) unités ; ainsi, la productivité physique marginale P_a diminue lorsque le capital K augmente, et ceci conformément à la règle des rendements décroissants. Par ailleurs, le produit total résultant de l'adjonction de la première unité de capital est mesuré par l'aire de la première « tranche rectangulaire » dont la base horizontale est l'intervalle [0,1] sur l'axe des abscisses et la hauteur est égale à 245. Pour obtenir le produit total provenant de l'application de 2 unités de capital, il faut ajouter les surfaces des deux premières tranches, la seconde étant un rectangle dont la base est l'intervalle [1,2] et la hauteur est égale à 101 : en effet, la somme de ces deux aires vaut 101+245=346 ; et ainsi de suite pour les autres valeurs de K. Finalement, la somme des aires de toutes les 6 tranches donne le produit total pour 6 unités de capital, à savoir $245+101+77+67+58+52 = 600$.

Supposons désormais que le volume de production soit donné par une fonction continûment dérivable $q\,(K)$, de l'unique variable K, sur l'intervalle ouvert $I =]0,6[$. Recherchons une fonction $f(K)$ qui soit continue sur I et telle que l'aire comprise entre son graphique, l'axe horizontal, l'axe vertical et la droite verticale d'abscisse K^* soit précisément égale à la valeur $q\,(K^*)$ du volume de production pour K^* unités de capital, quel que soit le nombre K^* de I.





Pour une valeur K^* de I, considérons un petit accroissement ΔK du capital, avec toutefois $K^* + \Delta K < 6$. Le minimum m (resp. le maximum M) de la fonction f sur l'intervalle $[K^*, K^* + \Delta K]$ existe en vertu du théorème de Weierstrass ; de plus, par comparaison des aires, on a

$$m\Delta K \leqslant q\left(K^* + \Delta K\right) - q\left(K^*\right) \leqslant M\Delta K$$

soit

$$m \leqslant \frac{q\left(K^* + \Delta K\right) - q\left(K^*\right)}{\Delta K} \leqslant M.$$

Ces inégalités sont conservées lorsque ΔK tend vers 0, auquel cas m et M s'approchent de $f(K^*)$ par continuité. Un passage à la limite et un recours à la définition de la dérivée livrent dès lors $f(K^*) = q'(K^*)$, ce qui équivaut à dire que la fonction continue de productivité marginale f coïncide, sur l'intervalle I, avec la dérivée de la fonction q.

Par l'interprétation géométrique de l'intégrale définie, il apparaît que, pour tout nombre K de l'intervalle I, la quantité produite $q(K)$ est donnée par l'égalité suivante :

$$q(K) = \int_0^K f(x)dx$$

Il convient de remarquer que, sur notre exemple, l'expression considérée est une intégrale généralisée convergente (puisque $\lim_{K \to 0^+} f(K) = +\infty$, tandis que $q(K)$ est un nombre connu pour tout K dans I). Par ailleurs, il ne s'agit pas ici de la forme « normale » d'une courbe de produit total car, dans ce cas, la productivité physique marginale augmente en premier lieu, ce qui signifie que les rendements ne sont alors décroissants qu'après une phase initiale (correspondant en quelque sorte au « rodage ») à rendements croissants.

2.4 Quelques problèmes d'optimisation

Comme le faisait remarquer L.R. Klein (qui reçut le prix Nobel en 1980), *« economists try to select among alternative uses of scarce ressources in such a way as to make the most efficient (or least wasteful) employment of ressources to achieve stated ends. Stated in this way, we see clearly that* *economics involves optimization, and this is the engine that produces principles of economic analysis »* (10).

Effectivement, des problèmes de programmation mathématique surgissent d'une façon spontanée dans le monde des affaires, puisque la programmation mathématique a précisément pour objet l'étude des extrema (maximum et/ou minimum) d'une fonction soumise ou non à certaines contraintes.

Grâce à un exemple très simple, nous allons revenir sur le théorème fondamental exploité en programmation linéaire.

Ensuite, nous aborderons les deux problèmes fondamentaux d'optimisation rencontrés en microéconomie, et ceci toujours par le biais d'exemples sélectionnés. Nous traiterons d'abord le cas d'un consommateur qui essaie de maximiser son utilité tout en respectant une contrainte budgétaire, ou encore de minimiser son budget pour autant qu'il obtienne un niveau de satisfaction fixé. Ensuite, nous nous pencherons sur le cas d'un producteur qui cherche à maximiser son profit en respectant éventuellement une contrainte technologique, ou qui veut minimiser ses coûts de production pour un niveau d'output donné.

Enfin, nous donnerons un petît aperçu du problème important de la gestion des stocks.

Nous avons délibérément choisi de présenter des exemples très simples, qui peuvent être résolus à l'aide de théories mathématiques fort classiques, et pour lesquels la *fonction économique* (c'est-à-dire la fonction à optimiser) ne comporte que deux variables tout au plus. Nous avons éliminé les problèmes (plus complexes, mais néanmoins fort importants) *dynamiques*, c'est-à-dire ceux dont l'extremum varie au cours du temps, et *multicritères*, c'est-à-dire ceux pour lesquels existent plusieurs objectifs (parfois contradictoires) à optimaliser.

Signalons que nous développerons certains cas typiques pour lesquels la seule intuition pourrait vite conduire à de mauvais résultats : nous espérons de la sorte montrer l'utilité du raisonnement mathématique dans l'administration des affaires.

---

(10) Klein, The Role of Mathematics in Economics, in *The Mathematical Sciences : a Collection of Essays*, the M.J.T. Press, 1969, p. 162

2.4.1 Distribution des programmes optimaux en programmation linéaire

D'une façon générale, le problème posé en programmation linéaire consiste en la recherche du maximum ou du minimum d'une fonction linéaire

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} c_i x_i$$

sur un ensemble E (de $\mathbb{R}^n$) défini par des contraintes linéaires du type

$$\sum_{i=1}^{n} a_i x_i \leqslant b, \ \sum_{i=1}^{n} a_i x_i \geqslant b$$

ou encore, en combinant ces deux inégalités,

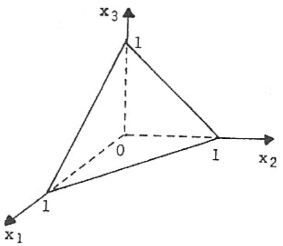
$$\sum_{i=1}^{n} a_i x_i = b$$

En particulier, on suppose souvent $x_i \geqslant 0$ pour $i = 1, 2, \ldots, n$; cet ensemble E, dont les éléments sont baptisés des *programmes réalisables*, est un polyèdre convexe de $\mathbb{R}^n$.

La théorie repose sur le fait que l'optimum éventuel de f sur E est nécessairement atteint (s'il existe) en un point extrême (appelé alors *programe optimal*) de E.

Pour démontrer ce théorème fondamental de la programmation linéaire, admettons que l'optimum recherché soit atteint en un point x^*. Quitte à travailler dans la variété affine engendrée par E, nous pouvons supposer non vide l'intérieur de E. Si la fonction f n'est pas constante sur E, elle ne possède aucun point stationnaire et est de ce fait dénuée d'extrémants locaux. Le point x^* ne peut appartenir qu'à la frontière $\dot{E}$, qui est la réunion d'intersections de E avec un hyperplan construit en remplaçant une inégalité de contrainte (conditions de non-négativité incluses) par l'égalité correspondante. Il existe donc un polyèdre convexe E_1 contenant x^* et inclus dans $\dot{E}$ (donc dans E), la dimension de E_1 étant strictement inférieure à celle de E. Dans la variété engendrée par E_1, la fonction économique reste affine et l'on peut recommencer le raisonnement ci-dessus à propos de E_1 au lieu de E. Après un nombre fini de telles opérations, on aboutit à la conclusion que x^* est un point extrême d'un polyèdre convexe inclus dans E (donc x^* est également un point extrême de E), à moins que la fonction économique ne reste constante sur un polyèdre convexe E^* inclus dans E, auquel cas tout point extrême de E^* (qui est aussi un point extrême pour E) procure à f l'optimum cherché, à savoir $f(x^*)$.

Illustrons ce résultat par un exemple simple. Recherchons les extrema de la fonction économique $f(x_1, x_2, x_3) = 2x_1 + x_2 + 3x_3$ sur le polyèdre convexe E défini par les conditions de non-négativité $x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0$ et la vraie contrainte $x_1 + x_2 + x_3 \leqslant 1$.



L'existence d'un maximum et d'un minimum de f sur E est garantie par la compacité de E (théorème de Weierstrass). Les dérivées partielles premières de f, à savoir $f_1' = 2, f_2' = 1$ et $f_3' = 3$, sont constantes et différentes de 0, de sorte qu'il n'existe aucun extremum local à l'intérieur de E et que tout extremum de f sur E est atteint en un point situé sur la frontière $\dot{E}$.
Cette dernière se décompose en quatre parties, à savoir les quatre triangles E_1 (resp. E_2, E_3, E_4) caractérisés par $x_1 = 0$ et $x_2 + x_3 \leqslant 1$ (resp. $x_2 = 0$ et $x_1 + x_3 \leqslant 1, x_3 = 0$ et $x_1 + x_2 \leqslant 1, x_1 + x_2 + x_3 = 1$ et $x_i \geqslant 0$ pour $i = 1, 2, 3$).

Considérons, par exemple, le triangle E_1, dont les sommets sont les points $(0, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$; comme $x_1 = 0$, la fonction ne dépend plus que des variables x_2 et x_3, et le problème consiste alors à rechercher les extrema de $f_1 = x_2 + 3x_3$ sous les contraintes $x_2 \geqslant 0, x_3 \geqslant 0$ et $x_2 + x_3 \leqslant 1$; à nouveau, les dérivées partielles de f_1 par rapport à x_2 et à x_3 sont non nulles, et l'extremum de f_1 sur E_1 est atteint sur la « frontière relative » de E_1, c'est-à-dire sur un des trois segments $[(0, 0, 0) : (0, 1, 0)]$, $[(0, 0, 0) : (0, 0, 1)]$ ou $[(0, 1, 0) : (0, 0, 1)]$; sur chacun de ceux-ci, on peut reprendre le même raisonnement, à savoir éliminer une variable, puis rechercher l'optimum d'une fonction affine de l'unique variable restante sur un intervalle, ce qui ne peut admettre comme solution qu'une des extrémités d'un des segments, c'est-à-dire un des sommets $(0, 0, 0)$, $(0, 1, 0)$ ou $(0, 0, 1)$. La même démarche doit ensuite être entreprise (de la même manière) sur les autres triangles E_2, E_3 et E_4. Notons toutefois que, sur la dernière partie E_4 de $\dot{E}$, on élimine une variable, par exemple x_3, grâce à l'égalité $x_3 = 1 - x_1 - x_2$, ce qui transforme la fonction linéaire f en $f_4(x_1, x_2) = 2x_1 + x_2 + 3(1 - x_1 - x_2) = 3 - x_1 - 2x_2$, fonction qui est également affine en x_1 et x_2. A la fin des opérations, on s'aperçoit que le maximum et le minimum de f sur E ne peuvent être atteints qu'en un des sommets de E, à savoir en un des quatre points $(0,0,0),(1,0,0),(0,1,0)$ et $(0,0,1)$. Ici, bien entendu, le maximum (resp. minimum) 3 (resp.0) est atteint au point $(0,0,1)$ (resp. $(0,0,0)$).

Observons encore que la fonction $f(x_1,x_2,x_3)=x_1+x_2+x_3$ admet, sur le même ensemble E, le minimum 0, atteint en l'origine, et le maximum 1, atteint aux trois autres sommets $(1,0,0),(0,1,0),(0,0,1)$, mais aussi en tous les points du triangle E_4 (qui est l'enveloppe convexe de ces trois sommets) ; en réalité, cette fonction f est constante sur E_4.

2.4.2 Problèmes d'optimisation dans la théorie du consommateur

Nous allons d'abord traiter en détail un exemple classique et simple : il mettra bien en évidence les types de problèmes d'optimisation posés dans la théorie microéconomique du consommateur, ainsi que les principales techniques de programmation utilisées.

Un consommateur peut disposer de deux biens P_1 et P_2. La fonction d'utilité est donnée par $U=x_1x_2$, ou x_1 (resp.x_2) désigne la quantité de P_1 (resp. de P_2) qu'il achète. Le prix unitaire p_1 (resp. p_2) de chaque produit vaut, pour fixer les idées, une unité monétaire. Il s'agit tout d'abord de rechercher le maximum de l'utilité du consommateur, sachant que celui-ci dispose d'un budget fixe b pour l'acquisition des deux biens : par exemple, $b=100$ unités monétaires.

Pour trouver le maximum de $U=x_1x_2$ sur l'ensemble $E=\{x\in\mathbb{R}^2 : x_1>0,x_2>0,x_1+x_2=100\}$, nous pourrions bien sûr éliminer la variable x_2, via la contrainte $x_2=100-x_1$, et chercher le maximum de la fonction $x_1(100-x_1)$ sur l'intervalle $[0,100]$. Nous allons toutefois utiliser la technique des multiplicateurs de LAGRANGE qui est bien adaptée à ce problème et admet une interprétation économique intéressante. Le lagrangien du problème est la fonction $L(x;\lambda)=x_1x_2+\lambda(100-x_1-x_2)$; il possède un seul point stationnaire, à savoir $x^*=(50,50)$, associé au multiplicateur de LAGRANGE $\lambda^*=50$. L'ensemble défini par la contrainte budgétaire ne possède aucun point singulier. Comme la matrice hessienne bordée par les coefficients de la contrainte, à savoir

$$\overline{\mathcal{H}}L(x;\lambda^*)=\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$$

est constante et de déterminant positif, le point x^* est un maximum local de $L(x;\lambda^*)$, donc de U sur E ; comme, de plus, la fonction U est strictement quasi-concave sur l'ouvert $\{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$, ce maximant local x^* est aussi maximant absolu de U sur E : le maximum de l'utilité sur E vaut alors $U(x^*) = 2500$.

Supposons à présent que le budget disponible b puisse être modifié ; l'utilité maximale $U_{max} = L(x^*;\lambda^*)$ va changer, de même que les quantités optimales x_1^* et x_2^* : ces grandeurs vont dépendre du budget b et l'on aura $U_{max} = L[x_1^*(b), x_2^*(b);\lambda^*(b)]$.

La variation du maximum en fonction du budget peut être estimée par le calcul de la dérivée de cette fonction U_{max} par rapport à b ; en utilisant le théorème de la dérivation des fonctions composées, on trouve

$$\begin{aligned}\frac{dU_{max}}{db} &= \frac{\partial L}{\partial x_1}[x^*(b);\lambda^*(b)]\frac{dx_1^*}{db} + \frac{\partial L}{\partial x_2}[x^*(b);\lambda^*(b)]\frac{dx_2^*}{db} \\ &\quad + \frac{\partial L}{\partial \lambda}[x^*(b);\lambda^*(b)]\frac{d\lambda^*}{db} + \frac{\partial L}{\partial b}[x^*(b);\lambda^*(b)] \\ &= \lambda^*\end{aligned}$$

puisque $\frac{\partial L}{\partial \lambda}[x^*(b);\lambda^*(b)] = 0$, $\frac{\partial L}{\partial b}[x^*(b);\lambda^*(b)] = \lambda^*$, et $\frac{\partial L}{\partial x_i}[x^*(b);\lambda^*(b)] = 0$ pour $i = 1, 2$ (car x^* reste un point stationnaire du lagrangien et la contrainte $x_1^* + x_2^* = b$ est toujours vérifiée). En conclusion, le multiplicateur de Lagrange λ^* mesure la sensibilité du maximum de l'utilité pour une variation de budget ([11]). Plus concrètement, λ^* donne une approximation de la variation de la valeur optimale U_{max} lorsque le budget s'accroît d'une unité monétaire. Pour illustrer ces propos, reprenons le cas où b vaut 100 : la valeur $U_{max} = 2500$ est atteinte pour $x_1^* = x_2^* = 50$, avec $\lambda^* = 50$; si le budget est porté à 101 unités monétaires, la nouvelle valeur optimale vaut $U(50,5; 50,5) = 2550,25$, de sorte que la différence entre ces deux valeurs extrémales est égale à $2550,25 - 2500 = 50,25$ et est très proche de $\lambda^* = 50$.

Cet exemple illustre la propriété générale selon laquelle les fonctions de demande du consommateur sont homogènes de degré 0 par rapport aux prix et au budget : en effet, si tous les prix et le budget sont multipliés par un même facteur positif, les quantités demandées x_1^* et x_2^*, ainsi que l'utilité optimale restent inchangées.

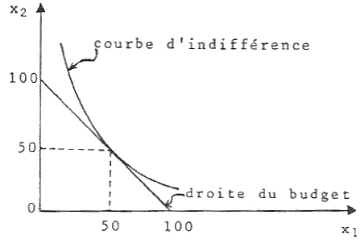
([11]) En général, un multiplicateur de Lagrange peut être interprété comme un taux de variation de l'objectif pour une modification infinitésimale du second membre de la contrainte ; la comparaison du nouvel équilibre (obtenu après cette modification) à l'ancien porte le nom de *statique comparative*.

Supposons désormais que le budget ne soit plus prédéterminé et désignons par p_1,p_2 les prix unitaires de P_1,P_2 respectivement. Il s'agit de rechercher les quantités des produits que le consommateur achètera en fonction de ces prix, sachant qu'il désire minimiser ses dépenses tout en exigeant que son utilité soit fixe, égale à U_0. Mathématiquement, il convient de minimiser la fonction coût $C = p_1x_1 + p_2x_2$ sous la contrainte $x_1x_2 = U_0$ (avec toujours $x_1 > 0$, $x_2 > 0$). Comme il n'existe aucun point singulier de l'ensemble E_1 défini par la contrainte, il faut construire le lagrangien $L_1(x, \mu) = p_1x_1 + p_2x_2 + \mu(U_0 - x_1x_2)$; un point stationnaire de L_1 est donné par $x^* = \left(\sqrt{U_0\frac{p_2}{p_1}}, \sqrt{U_0\frac{p_1}{p_2}}\right)$, associé au multiplicateur de Lagrange $\mu^* = \sqrt{\frac{p_1p_2}{U_0}}$; en particulier, si $p_1 = p_2 = 1$ et $U_0 = 2500$, on trouve $x_1^* = x_2^* = 50 = \frac{1}{\mu^*}$. Comme le déterminant de la matrice hessienne bordée en x^* est négatif, ce point x^* procure à $L_1(x; \mu^*)$ un minimum local ; $L_1(x^*; \mu^*)$ est aussi minimum local de C sur E_1, et dès lors le minimum absolu de C sur E_1 (en raison du caractère linéaire de la fonction C).

A la suite de cet exemple, il est facile de vérifier que les quantités x_1 et x_2 retenues (qui sont appelées les *fonctions de demande compensées* du consommateur) sont homogènes de degré 0 par rapport aux prix.

D'une manière générale, on peut affirmer que, dans les mêmes conditions, le problème de la recherche du maximum de l'utilité sous la contrainte budgétaire et celui de la recherche du minimum sous la contrainte d'une utilité prédéterminée sont tout à fait identiques (avec les mêmes données). Lorsque l'utilité est quasi-concave, c'est-à-dire lorsque les courbes d'indifférence sont convexes, la solution de ces problèmes est donnée par les conditions du premier ordre. Ces dernières expriment que le taux de substitution des biens $TSB = \frac{U_1'}{U_2'}$ est égal au rapport des prix $\frac{p_1}{p_2}$; elles se traduisent géométriquement par la tangence de la courbe d'indifférence et de la droite du budget adéquates.

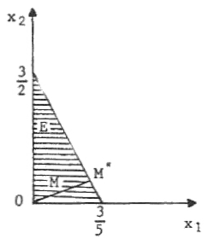
Ainsi, pour notre application numérique où $p_1 = p_2 = 1, b = 100$ et $U_0 = 2500$, la droite d'équation $x_1 + x_2 = 100$ est tangente à l'hyperbole équilatère d'équation $x_1x_2 = 2500$ au point $x^* = (50, 50)$.



Observons enfin que les multiplicateurs de Lagrange λ^* et μ^* des deux problèmes soulevés mesurent tous les deux la sensibilité de l'optimum pour une variation de la contrainte et sont l'inverse l'un de l'autre, à savoir $\lambda^* = \frac{1}{\mu^*}$ (= 50 pour nos données numériques $p_1 = p_2 = 1, b = 100$ et $U_0 = 2500$).

Pour introduire le problème général posé en programmation non linéaire, analysons le cas d'un (autre) consommateur qui peut disposer de deux biens et pour lequel la fonction d'utilité est donnée par $f(x_1, x_2) = (x_1 + 1)x_2$, où x_1 et x_2 désignent les quantités respectives des deux produits en question ; remarquons que cette fonction f est quasi-concave, mais non concave. L'individu cherche à maximiser son utilité, sachant que les variables x_1 et x_2 ne peuvent évidemment pas être négatives et qu'il ne peut dépenser plus qu'il ne possède, de sorte qu'il doit respecter une contrainte budgétaire du type $p_1x_1 + p_2x_2 \leqslant b$ (où p_1 et p_2 sont les prix unitaires des deux produits, et b le budget maximal prévu pour l'acquisition de ces deux biens) ; pour fixer les idées, nous prendrons $p_1 = 5, p_2 = 2$ et $b = 3$.

Mathématiquement, le problème à résoudre est le suivant : chercher le maximum de $(x_1+1)x_2$, sachant que $x_1 \geqslant 0, x_2 \geqslant 0$ et $5x_1 + 2x_2 \leqslant 3$. L'ensemble E des programmes réalisables, c'est-à-dire l'ensemble des points $x = (x_1, x_2)$ du plan qui vérifient toutes les contraintes, est le triangle de sommets $(0, 0)$, $(\frac{3}{5}, 0)$ et $(0, \frac{3}{2})$ (voir figure ci-contre).



Le caractère croissant de la fonction d'utilité oblige un maximum de f sur E (qui existe en vertu du théorème de Weierstrass) à être situé sur le segment $S = [(\frac{3}{5}, 0) : (0, \frac{3}{2})]$; en effet, un maximant ne peut pas se trouver, par exemple, en un point M intérieur à E, car le point M^* situé à l'intersection de la demi-droite $[0 : M)$ et du segment S procurerait une utilité supérieure ; de la même manière, un maximant de f sur E ne peut pas appartenir au segment $[(0, 0) : (0, \frac{3}{2})[$, ni à $[(0, 0) : (\frac{3}{5})[$. De la sorte, on se trouve apparemment en présence d'un problème déjà rencontré, puisqu'il s'agit de maximiser f sous la contrainte d'égalité $5x_1 + 2x_2 = 3$; la méthode de Lagrange conduit aux conditions du premier ordre suivantes : $\frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial \lambda} = 0$ pour $L = (x_1+1)x_2 + \lambda(3 - 5x_1 - 2x_2)$, soit $x_1^* = -\frac{1}{5}, x_2^* = 2$ et $\lambda^* = \frac{2}{5}$; cette solution est inacceptable en raison du caractère négatif d'une variable. En fait, dans un cas aussi simple, la solution peut être obtenue par simple élimination : l'égalité $5x_1 + 2x_2 = 3$ entraîne $x_2 = \frac{3}{2} - \frac{5}{2}x_1$, ce qui nous amène à rechercher le maximum de la fonction $(x_1 + 1)(\frac{3}{2} - \frac{5}{2}x_1) =$

$-\frac{5}{2}x_1^2 - x_1 + \frac{3}{2}$ sur l'intervalle $[0, \frac{3}{5}]$, car $x_1 \geqslant 0$ et $x_2 = \frac{3}{2} - \frac{5}{2}x_1 \geqslant 0$; il est clair que le maximum vaut $\frac{3}{2}$ et est atteint au point $(0, \frac{3}{2})$.

Cet exemple introductif nous amène à poser le problème général de programmation non linéaire, tel qu'il se rencontre effectivement dans de nombreuses situations économiques.

On recherche le maximum (éventuel) d'une fonction $f(x_1, \ldots, x_n)$ lorsque les n variables $x_1, x_2, \ldots, x_n$ sont positives ou nulles (*conditions de non-négativité*) et vérifient m *vraies contraintes* du type $g_j(x_1, \ldots, x_n) \leqslant v_j$ pour $j = 1, 2, \ldots, m$, où les b_j sont des constantes. Nous supposerons que les $m+1$ fonctions $f, g_1, \ldots, g_m$ sont autant de fois continûment dérivables que nécessaire sur un ouvert contenant l'ensemble E des programmes réalisables, à savoir $E = \bigcap_{j=1}^{m} \{x \in \mathbb{R}_+^n : g_j(x) \leqslant b_j\}$ où $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x = (x_1, x_2, \ldots, x_n), x_i \geqslant 0 \text{ pour } i = 1, 2, \ldots, n\}$. Les n variables $x_1, x_2, \ldots, x_n$ sont souvent appelées les instruments, f est la *fonction économique* ou *l'objectif*, les g_j sont les *fonctions de contraintes* et les b_j sont les *constantes de contraintes.*

Ce problème général peut être résolu par l'introduction des multiplicateurs de Kuhn-Tucker-Lagrange,[4, 7, 9, 16, 33].

2.4.3 Elaboration d'un plan optimal de production

Toute entreprise se doit de poursuivre des objectifs qui sont en accord avec les critères d'amélioration sélectionnés. Or, chaque objectif peut souvent se traduire quantitativement par la recherche du maximum ou du minimum d'une fonction économique. Par exemple, on s'efforcera de maximiser le volume de la production, la part de marché, le revenu brut, le bénéfice, la satisfaction du client, etc ; de même, on pourra tenter de minimiser le coût total de fabrication, les dépenses en matières premières ou en main-d'oeuvre, le taux des investissements, la quantité de déchets (ou de « pollution ») dans le processus de fabrication, etc.

Nous allons examiner quelques exemples concrets montrant que l'élaboration d'un plan de production fait appel à la programmation mathématique.

Un monopoleur reçoit une commande linéaire donnée par l'égalité $p = 50 - q$ (où p désigne le prix de vente unitaire du produit, q la quantité commandée et supposée vendue), tandis que son coût total de production dépend de la quantité produite et vaut $C = 30 + 10q$. Pour rendre maximum le profit du monopoleur, à savoir la fonction $P = pq - C = (50 - q)q - (30 + 10q)$, il convient d'annuler la dérivée de P par rapport à q, ce qui donne $q = 20$, d'où $p = 30$ et $P = 370$: cette valeur de 370 représente bien le maximum souhaité. Supposons à présent que cet entrepreneur soit capable de répartir ses clients dans deux marchés distincts pour lesquels les lois de demande sont données respectivement par $p_1 = \frac{10}{9}(30 - q_1)$ et $p_2 = 200 - 10q_2$, ce qui laisse inchangée la loi de demande globale pour un même prix p, puisque les égalités $q_1 = 30 - 0,9p_1$ et $q_2 = 20 - 0,1p_2$ donnent, pour $p = p_1 = p_2, q = q_1 + q_2 = 30 - 0,9p + 20 - 0,1p = 50 - p$. Le profit dépend maintenant des deux variables q_1 et q_2, soit $P = p_1q_1 + p_2q_2 - C = \frac{10}{9}(30 - q_1)q_1 + (200 - 10q_2)q_2 - 30 - 10(q_1 + q_2)$. Les conditions du premier ordre pour l'obtention du maximum de P, à savoir $P'_1 = 0$ et $P'_2 = 0$, conduisent aux résultats suivants : $q_1 = 10,5$ et $q_2 = 9,5$ avec $p_1 = 21,266, p_2 = 105$ et $P = 995$. On peut montrer que cette valeur de 995 est bien le maximum cherché. Ainsi, grâce à la discrimination du marché, le profit du monopoleur a augmenté sensiblement en passant de 370 à 995, tandis que les prix (surtout p_2) s'éloignent assez fortement de $p = 30$. Remarquons que le prix est le plus bas sur le marché où l'élasticité de la demande est, en valeur absolue, la plus grande ; en effet, $e_1 = -1,857$ tandis que $e_2 = -1,105$, l'élasticité de la demande étant donnée par la formule générale $e = \frac{p}{q}\frac{dp}{dq}$. Cet exemple illustre le phénomène bien connu, mais paradoxal, du *dumping* : une firme vend parfois un de ses produits plus cher dans son propre pays qu'à l'étranger ; cela provient du fait que l'élasticité de la demande est généralement plus forte sur un marché extérieur (en raison de la concurrence importante) que sur le marché intérieur (où l'entreprise dispose plus facilement d'un véritable monopole).

Voici un autre exemple qui met en évidence la puissance de l'outil mathématique.

Un industriel fabrique et vend deux produits P_1 et P_2. Il se propose bien entendu de rendre son revenu net maximum et dispose des données suivantes :

produit	prix unitaire	prix de revient	quantité vendue
P_1	p_1	$2F$	$x_1 = 11 - 2p_1 - 2p_2$
P_2	p_2	$1F$	$x_2 = 16 - 2p_1 - 3p_2$

La fonction à maximiser est le bénéfice : $B = x_1p_1 + x_2p_2 - 2x_1 - x_2 = (p_1 - 2)(11 - 2p_1 - 2p_2) + (16 - 2p_1 - 3p_2)(p_2 - 1)$. Les conditions du premier ordre qui reviennent à annuler les deux dérivées partielles premières de B

(par rapport à p_1 et à p_2) conduisent à la résolution du système formé par les deux équations linéaires : $4p_1+4p_2 = 17$ et $4p_1+6p_2 = 23$. La solution unique de ce système est donnée par $p_1 = \frac{5}{4}$ et $p_2 = 3$. Il est possible de montrer que le point $p = (\frac{5}{4}, 3)$ procure bien à B le maximum recherché. Constatons que, contrairement à ce que l'intuition laisserait supposer, l'industriel n'a pas intérêt à augmenter démesurément ses prix et doit même vendre son produit P_1 à perte (de 0,75 F par unité) ; ce produit P_1 joue en réalité un rôle « d'appel » et est vendu dans le seul but d'inciter les clients à acheter le produit P_2 qui, au contraire de P_1, est tout à fait rentable.

A l'aide d'un exemple simple, illustrons à présent les problèmes d'optimisation classiquement rencontrés dans la théorie de la firme.

Un entrepreneur utilise deux inputs variables X_1 et X_2 pour produire un seul output Q. La fonction de production, qui donne la quantité q de l'output en fonction des quantités x_1 et x_2 des inputs X_1 et X_2, est la suivante : $f(x) = \sqrt{x_1x_2} + x_1 + x_2$. Il achète ses inputs à des prix unitaires p_1 et p_2, par exemple $p_1 = p_2 = 1$ unité monétaire ; de plus, ses frais fixes sont négligeables. Il se pose deux questions qui, nous le verrons, admettent une même réponse :

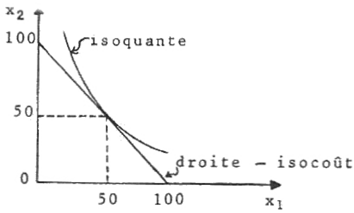
- a) comment maximiser la quantité produite de Q tout en respectant une contrainte budgétaire estimée à 100 unités monétaires pour l'acquisition des deux inputs ?
- b) comment minimiser le coût de production tout en exigeant d'atteindre un niveau donné d'output, par exemple 150 unités ?

Soit à rechercher d'abord le maximum de $f(x)$ sur l'ensemble $E = \{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0, x_1 + x_2 = 100\}$ qui ne possède aucun point singulier. Comme dans la théorie du consommateur, nous allons utiliser la technique des multiplicateurs de Lagrange, de préférence à la méthode d'élimination d'une variable grâce à la contrainte . Le lagrangien de ce problème vaut $L(x; \lambda) = \sqrt{x_1x_2} + x_1 + x_2 + \lambda(100 - x_1 - x_2)$; il admet pour point stationnaire x^*, avec $x_1^* = x_2^* = 50$, associé au multiplicateur de Lagrange $\lambda^* = \frac{3}{2}$. Comme $L(x; \lambda^*)$ est concave sur $\{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$, x^* est maximant absolu de $L(x; \lambda^*)$, donc maximant absolu de f sur E ; le maximum de la quantité produite vaut alors $f(x^*) = 150$. La valeur $\lambda^* = \frac{3}{2}$ du multiplicateur de Lagrange mesure la sensibilité de la valeur optimale de q pour une variation du budget. Par exemple, si le budget est porté à 101 unités monétaires, la quantité optimale de Q vaut visiblement $f(50,5; 50,5) = 151,5$, de sorte que $f(50,5; 50,5) - f(50; 50) = 1,5 = \lambda^*$.

Il s'agit à présent de minimiser la fonction coût $C = x_1 + x_2$ sur l'ensemble $E = \{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0, \sqrt{x_1 x_2} + x_1 + x_2 = 150\}$ qui ne possède aucun point singulier. Le lagrangien vaut $L_1(x;\mu) = x_1 + x_2 + \mu(150 - \sqrt{x_1 x_2} - x_1 - x_2)$; il admet pour point stationnaire $x_1^* = x_2^* = 50$, avec $\mu^* = \frac{2}{3}$. Ce point x^* procure bien le minimum de $L_1(x;\mu^*)$ sur l'ouvert $\{x \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\}$, puisque $L_1(x;\mu^*)$ y est convexe, donc le minimum de C sur E_1.

Constatons que les conditions du premier ordre pour la minimisation du coût sous une contrainte d'output sont semblables à celles pour la maximisation de la quantité produite sous une contrainte de coût. Elles expriment que le taux de substitution technique $TST = \frac{f_1'}{f_2'}$ est égal au rapport $\frac{p_1}{p_2}$ des prix des inputs (à savoir ici 1) ; elles se traduisent géométriquement par la tangence d'une droite d'isocoût (ici $x_1 + x_2 = 100$) et d'une isoquante (ici $\sqrt{x_1 x_2} + x_1 + x_2 = 150$).

En vertu du caractère concave de la fonction de production, ces conditions du premier ordre sont suffisantes pour atteindre l'extremun cherché. Observons enfin que le multiplicateur de Lagrange $\lambda^* = \frac{3}{2}$ du premier problème est l'inverse du multiplicateur du deuxième problème.

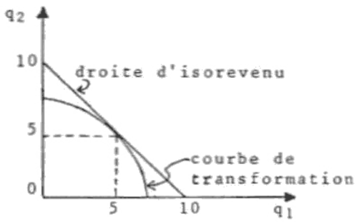


Dans la pratique, la contrainte est plus souvent une inégalité qu'une égalité. En guise d'exemple, supposons que la fonction de production d'une firme soit du type de COBB-DOUGLAS : $q = x_1^a x_2^b$ (avec $a > 0, b > 0$) ; les prix unitaires p_1 et p_2 des deux inputs sont des grandeurs constantes et considérées comme exogènes. Le problème posé à l'entrepreneur est de minimiser le coût global $C = p_1 x_1 + p_2 x_2$ pour produire au moins une quantité donnée (strictement positive) Q_0, c'est-à-dire sous la condition $x_1^a x_2^b \geqslant Q_0$, avec $x_1 \geqslant 0$ et $x_2 \geqslant 0$. Les contraintes obligent les deux variables à être strictement positives et la règle des multiplicateurs de KUHN-TUCKER-LAGRANGE peut ainsi être appliquée. Bien plus, la fonction coût est linéaire, donc quasi-convexe, tandis que la fonction de contrainte est à tout le moins quasi-concave. Toutes les hypothèses du théorème d'ARROW-ENTHOVEN sont visiblement satisfaites ; il en résulte que les conditions de KUHN-TUCKER sont suffisantes pour atteindre le minimum cherché. Il s'agit dès lors de trouver un point x^* et un réel λ^* non négatif qui obéissent aux relations suivantes : pour $L = p_1 x_1 + p_2 x_2 + \lambda(Q_0 - x_1^a x_2^b)$, $\frac{\partial L}{\partial x_i}(x^*;\lambda^*) \geqslant 0, x_i^* \geqslant 0$ et $x_i^* \frac{\partial L}{\partial x_i}(x^*;\lambda^*) = 0$ pour $i = 1, 2$, $\frac{\partial L}{\partial \lambda}(x^*;\lambda^*) \leqslant 0$ et $\lambda^* \frac{\partial L}{\partial \lambda}(x^*;\lambda^*) = 0$. Des calculs élémentaires livrent : $x_1^* = \left[Q_0\left(\frac{ap_2}{bp_1}\right)^b\right]^{\frac{1}{a+b}}$, $x_2^* = \left[Q_0\left(\frac{bp_1}{ap_2}\right)^a\right]^{\frac{1}{a+b}}$, ce qui établit le plan de production réclamant le coût global minimum, égal à $C^* = p_1x_1^* + p_2x_2^*$.

Des problèmes d'optimisation se posent également dans la théorie de *productions jointes*, c'est-à-dire dans l'étude de procédés de production qui fournissent plusieurs produits.

Examinons le cas d'un entrepreneur qui utilise un seul input X pour la production de deux outputs Q_1 et Q_2 ; un exemple classique est donné par l'élevage des moutons qui permet d'obtenir deux produits, à savoir la laine et la viande. La relation de production est ici donnée par $x = f(q_1, q_2) = q_1^2 + q_2^2$, où q_1, q_2 et x désignent les quantités respectives de Q_1, Q_2 et X. Les outputs sont vendus à des prix p_1 et p_2 déterminés, par exemple à une unité monétaire pour une unité de chaque output ($p_1 = p_2 = 1$). La première question à traiter est de savoir comment l'entrepreneur peut maximiser son revenu en utilisant une quantité prédéterminée, par exemple 50 unités, de son input X. Le problème consiste alors à maximiser la fonction de revenu, à savoir $R = q_1 + q_2$, sur l'ensemble $E = \{q \in \mathbb{R}^2 : q_1 > 0, q_2 > 0, q_1^2 + q_2^2 = 50\}$ qui n'admet aucun point singulier. Le lagrangien $L(q;\lambda) = q_1 + q_2 + \lambda(50 - q_1^2 - q_2^2)$ possède un seul point stationnaire, soit $q^* = (5, 5)$, associé au multiplicateur de Lagrange $\lambda^* = 0,1$. Comme la fonction $L(q;\lambda^*)$ est (partout) strictement concave, q^* est bien un maximant strict de $L(q;\lambda^*)$, donc aussi un maximant strict du revenu R sur l'ensemble E : le maximum cherché vaut $R^* = 10$. Si la quantité de l'input est portée, par exemple, à 51, le revenu optimal vaut visiblement $2\sqrt{25,5} \simeq 10,0995$ unités monétaires, de sorte que l'accroissement du maximum, soit 10,0995 - 10 = 0,0995, est assez bien estimé par le multiplicateur de Lagrange $\lambda^* = 0,1$. Le caractère linéaire du revenu et concave de la contrainte rend concave le lagrangien ; de la sorte, les conditions du premier ordre suffisent pour résoudre ce problème ; elles signifient que le taux de transformation des produits, à savoir $TTP = \frac{f_1'}{f_2'}$ (égal ici à $\frac{2q_1}{2q_2} = \frac{q_1}{q_2}$) doit être égal au rapport $\frac{p_1}{p_2}$ des prix (ici $\frac{p_1}{p_2} = 1$).

Géométriquement, cette égalité $TTP = \frac{p_1}{p_2}$ se visualise par la tangence d'une courbe de transformation des produits (ici $q_1^2 + q_2^2 = 40$) et d'une droite d'iso-revenu (ici $q_1 + q_2 = 10$).



Supposons à présent que l'entrepreneur désire maximiser son profit lorsque le prix unitaire de son input est connu, fixé par exemple à $p = 0,25$ unité monétaire. Il importe donc de rechercher le maximum de la fonction $\pi = q_1 + q_2 - 0,25(q_1^2 + q_2^2)$, sans aucune contrainte. La théorie générale des extrema libres conduit à la valeur extrémale $\pi^* = 2$, ce maximum étant atteint pour $q_1^* = q_2^* = 2$. Ainsi, la productivité marginale de X pour la production de chaque output est égale au prix de x : on a en effet, dans le cas général, $p = p_1 \frac{\partial q_1}{\partial x} = p_2 \frac{\partial q_2}{\partial x}$ (ici, $0,25 = 1.\frac{1}{2\times 2} = \frac{1}{4}$).

2.4.4 Un modèle simple pour la gestion d'un stock

Nous allons présenter, sur un exemple, un modèle très élémentaire (dû à Wilson) pour la gestion d'un stock.

La demande annuelle d'un produit de consommation courante, fabriqué et stocké par une entreprise, est supposée constante et s'élève à 500 000 unités. Le coût de stockage par unité est évalué à 0,05 F par jour, tandis que le coût par unité manquante (ou coût de pénurie qui représente notamment une évaluation des indemnités à verser aux clients à cause du retard de livraison) est, par jour, estimé au double du coût de stockage. Par ailleurs, le coût de lancement (qui provient des frais administratifs et publicitaires pour chaque opération de réapprovisionnement) est forfaitairement évalué à 648 F pour chaque commande, indépendamment du volume livré.

Le problème consiste à rendre minimum le coût annuel de la gestion du stock, d'en déduire la durée idéale séparant deux réapprovisionnements successifs ainsi que le volume optimal de chaque commande.

Nous supposerons que le réapprovisionnement a lieu instantanément par quantités constantes de n pièces et à intervalles réguliers de T jours ; le nombre de « rafales » (c'est-à-dire de commandes) est alors égal à $r = \frac{500\ 000}{n} = \frac{360}{T}$, en admettant qu'une année comporte 360 jours (soit 12 mois de 30 jours).

Dans l'hypothèse où le stock est toujours approvisionné, le coût de stockage par période de T jours est égal à $\frac{0{,}05}{2} nT$ (puisque la quantité moyenne stockée pendant ce laps de temps vaut $\frac{n}{2}$). Sur une année entière, le coût de stockage est égal à $C_s = \frac{0{,}05}{2} nTr = 9n$, tandis que le coût de lancement est de $C_l = 648r = \frac{324\,000\,000}{n}$. Le coût annuel global pour la gestion du stock vaut $C = C_s + C_l = 9n + \frac{324\,000\,000}{n}$. Le minimun de cette fonction C est atteint lorsque $9n^2 = 324000000$, égalité provenant de l'annulation première de C, d'où $n = 6000$ unités ; la durée optimale séparant deux commandes successives est alors égale à $T = \frac{360}{500\,000}6000 = 43,2$ jours et le coût annuel minimal de gestion du stock est de 108 000 F.

Admettons à présent la possibilité d'avoir une rupture du stock. La période T est divisée en une période T_1 pendant laquelle le stock est approvisionné et une période de T_2 dite de rupture de stock. L'approvisionnement, intervenant à la fin de T_2, a pour premier effet de satisfaire les demandes laissées en suspens pendant T_2 ; il ramène ensuite le stock à son niveau maximal s. Si n désigne encore la quantité commandée lors de chaque réapprovisionnement, on a visiblement $T_1 = \frac{s}{n}T$ et $T_2 = \frac{n-s}{n}T$. Comme la quantité moyenne stockée pendant T_1 vaut $\frac{s}{2}$, le coût global de stockage par période est de $\frac{0{,}05}{2}sT_1$. Par ailleurs, la pénurie moyenne pendant T_2 est de $\frac{1}{2}(n-s)$, d'où le coût total de pénurie par période s'élève à $\frac{0{,}1}{2}(n-s)T_2$. En conséquence, le coût global de gestion pendant une année s'exprime en fonction des variables n et s, à savoir $C = \left(\frac{0{,}05}{2}sT_1 + \frac{0{,}1}{2}(n-s)T_2 + 648\right)r$ ou encore $C = 18n + \frac{27}{n}s^2 + \frac{324\,000\,000}{n} - 36s$. En annulant les deux dérivées partielles premières de C par rapport à n et à s, on obtient $s = \frac{0{,}1}{0{,}1+0{,}05}n = \frac{2}{3}n$, le coefficient $\frac{2}{3}$ étant appelé le taux de pénurie. On trouve de plus $n = 6000\sqrt{\frac{3}{2}} = 7348$ unités, avec $s = \frac{2}{3}7348 = 4899$ et $T = 43,2\sqrt{\frac{3}{2}} = 53$ jours : ces valeurs assurent bien le minimum de la fonction C, à savoir $108000\sqrt{\frac{2}{3}} = 88181$ F.

Comme le taux de pénurie est inférieur à l'unité, l'acceptation de la pénurie a pour effet d'augmenter la période entre deux réapprovisionnements successifs et la rafale ; ce dernier point peut sembler surprenant, mais il s'explique par le fait qu'un modèle sans pénurie est un cas particulier du modèle correspondant avec pénurie, lorsque le coût de pénurie « devient » infini (ce qui équivaut somme toute à interdire la rupture du stock) : cette contrainte nécessite un stock moyen plus important pour une même consommation. Il est aussi étonnant de constater que le coût global de gestion diminue lorsque la pénurie est admise : la raison en est que le coût de stockage supplémentaire, destiné à éviter la rupture du stock, est supérieur aux frais qu'entraînerait cette rupture. Dans notre exemple, l'économie réalisée est de plus de 18%.

Chapitre 3

Modèles dynamiques en économie

par *Jacques Bair*

On qualifie de *dynamique* tout modèle pour lequel le temps fait partie de l'analyse et qui permet d'expliquer directement la variation temporelle des grandeurs ou de déterminer quand ces variables convergent, à long terme, vers des valeurs d'équilibre.

Le temps peut être regardé comme une variable *continue* ou *discrète*. Dans le premier cas, la variable temporelle peut prendre toute valeur comprise dans un intervalle de $\mathbb{R}$, souvent dans l'ensemble des réels non négatifs : cette situation se rencontre, par exemple, dans le calcul d'un intérêt composé continu, qui peut être déterminé instantanément à tout moment. Dans le second cas, les grandeurs considérées sont mesurées seulement à intervalles de temps réguliers : la variable temporelle ne prend dès lors que des valeurs entières (exprimées en une unité adéquate) ; un exemple est donné par la formation d'un capital placé à un taux d'intérêts composés constant et dont les intérêts sont calculés uniquement à la fin de chaque mois (ou de chaque année).

Quand on étudie des processus continus qui sont décrits à l'aide de fonctions (de $\mathbb{R}$ dans $\mathbb{R}$) dérivables, on peut être amené à résoudre des équations différentielles. Par exemple, la vitesse de croissance d'une grandeur au cours du temps permet d'obtenir, par simple intégration, la valeur de cette grandeur à tout instant, pour autant que l'on en connaisse la valeur initiale. Par contre, lorsque les processus sont discrets, on utilise souvent des *différences finies* définies par $\Delta y(t) = y(t+1) - y(t)$ [1], et l'on doit alors résoudre des équations récurrentes ou aux différences finies. Par exemple, si la différence $\Delta y(t)$ est connue à tout instant t, on trouve de proche en proche chaque valeur $y(t)$ en fonction de la valeur initiale $y(0)$.

Nous allons donner quelques modèles économiques simples et souvent classiques, discrets ou continus. Délibérément, nous n'avons retenu que des modèles comportant des équations (différentielles ou récurrentes) d'ordre 1 ou 2, linéaires et à coefficients constants.

3.1 Quelques modèles discrets

3.1.1 Valeur d'un capital placé à intérêts composés

La valeur y_t, à l'époque t, d'un capital placé à intérêts composés au taux annuel de $i\%$ vérifie l'équation : $y_t = (1 + 0,0i)y_{t-1}$. Si l'on effectue un versement (ou un retrait annuel) a_t, cette valeur obéit alors à l'équation $y_t = a_t + (1 + 0,0i)y_{t-1}$.

En guise d'exemple concret, considérons le cas d'une personne qui gagne 100 000 F à la loterie nationale et qui décide de placer ce gain à un taux annuel de 5%. Son capital (exprimé en francs) va s'élever à 100 000 F+ 100 000 F $\times$ 0,05 $=$ 100 000 F $\times$ 1,05 $=$ 105 000 F au terme de la première année, à 105 000 F $\times$ 1,05 $=$ 110 250 F au terme de la deuxième année, à 110 250 F $\times$ 1,05 $=$ 115 762,5 F au terme de la troisième année, etc. Au bout de 40 ans, son avoir sera de 100 000 F $\times 1,05^{40} = 703\ 999$ F : pour s'en convaincre, il suffit de résoudre l'équation récurrente, linéaire, d'ordre 1 et homogène : $y_t = 1,05 y_{t-1}$, avec la condition initiale $y_0 = 100\ 000$.

Supposons à présent que notre heureux gagnant soit moins économe et décide de retirer 3000 F à la fin de chaque année. Le capital dont il disposera après 40 ans ne se calcule plus de façon aussi immédiate que dans le premier cas : il s'agit en fait de résoudre l'équation récurrente linéaire, d'ordre 1 et

[1] Dans les équations récurrentes, la fonction y à découvrir n'est définie que pour des valeurs entières de la variable t, c'est pourquoi, l'on utilise assez souvent la notation y_t au lieu de $y(t)$.

non homogène : $y_t = 1,05 y_{t-1} - 3000$; la solution de cette équation est donnée par $y_t = 40000 \times 1,05^t + 60000$, d'où y_{40} =341 600 F.

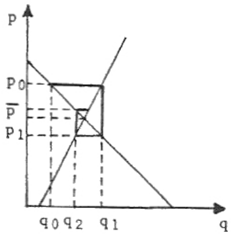
3.1.2 Modèle de Harrod pour l'analyse du revenu national

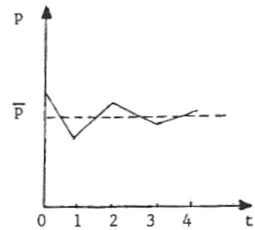
Dans ce modèle macroéconomique, on suppose, qu'à tout instant t, l'épargne e_t est proportionnelle au revenu national y_t, que l'investissement i_t est proportionnel à l'accroissement du revenu national $y_t - y_{t-1}$, et enfin que l'épargne est toujours entièrement investie ; par ailleurs, la valeur initiale $y_0 = y(0)$ du revenu est connue et positive. Ces hypothèses se traduisent par les égalités suivantes : $e_t = \alpha y_t, i_t = \beta(y_t - y_{t-1})$ et $e_t = i_t$, où α et β sont deux coefficients de proportionnalité positifs. On en déduit l'équation récurrente (linéaire, d'ordre 1 et homogène) : $\alpha y_t = \beta(y_t - y_{t-1})$, ou encore $y_t = \frac{\beta}{\beta-\alpha} y_{t-1}$. La solution vaut $y_t = y_0 \left(\frac{\beta}{\beta-\alpha}\right)^t$, d'où $e_t = i_t = \alpha \left(\frac{\beta}{\beta-\alpha}\right)^t$ tend, si $\beta > \alpha$, de façon explosive vers l'infini quant t tend vers l'infini (puisque $\frac{\beta}{\beta-\alpha} > 1$ si $\beta > \alpha$). Ce modèle représente une économie en expansion.

3.1.3 Modèle du cobweb

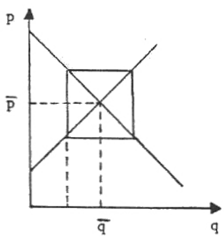
Une application économique classique des équations récurrentes est fournie par le modèle du « *cobweb* » (ou « toile d'araignée ») qui décrit un marché à réaction de l'offre retardée, par exemple celui du blé d'hiver : les plans de production sont élaborés après la moisson, mais la production correspondant à ces plans apparaît sur le marché un an plus tard ; la quantité demandée à une période quelconque dépend du prix à cette époque, tandis que la quantité offerte est fonction du prix de la période précédente. Si l'on suppose que ces dépendances sont linéaires et que la quantité offerte S_t à la période t est toujours égale à la quantité demandée D_t à cette période, $D_t = ap_t + b$ (a et b étant des constantes, avec $a < 0$ car D_t diminue lorsque le prix p_t augmente), $S_t = cp_{t-1} + d$ (c et d étant des constantes, avec $c > 0$ car S_t et p_t croissent simultanément) et $S_t = D_t$, d'où $p_t = \frac{c}{a} p_{t-1} + \frac{d-b}{a}$. La solution générale de cette équation récurrente est donnée par $p_t = (p_0 - \overline{p}) \left(\frac{c}{a}\right)^t + \overline{p}$, avec $\overline{p} = \frac{d-b}{a-c}$; notons que $\overline{p}$ représente l'abscisse du point d'intersection des droites $q = ap + b$ (pour la demande) et $q = cp + d$ (pour l'offre). La stabilité du système va donc dépendre du rapport $\frac{c}{a}$ (c'est-à-dire du rapport des coefficients angulaires des « droites » d'offre et de demande), sachant que $\frac{c}{a}$ est négatif.

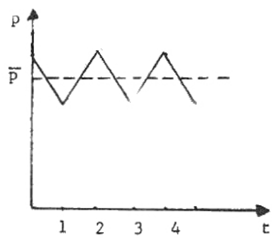
Si $|\frac{c}{a}| < 1$, il y a stabilité avec alternances amorties . Illustrons ce cas par les équations suivantes : $D_t = 10 - p_t$ et $S_t = 1 + 0,5p_{t-1}$; pour le prix initial $p_0 = 8$, l'offre initiale vaut $q_0 = 2$; comme p_0 diffère de la position d'équilibre $\overline{p} = 6$, l'entrepreneur est amené à offrir la quantité $q_1 = 5$ à la prochaine période ; le prix tombe instantanément à $p_1 = 5$; ce prix p_1 entraîne une offre de $q_2 = 3,5$, qui conduit au prix $p_2 = 6,5$; ce processus continue indéfiniment, formant un modèle dit « en toile d'araignée » (ou « cobweb ») dans le plan (Oq, Op) ; le niveau du prix prend respectivement les valeurs $p_0 = 8, p_1 = 5, p_2 = 6,5, p_3 = 5,75, p_4 = 6,125, \ldots$; le prix fluctue donc autour de la valeur $\overline{p} = 6$ vers laquelle il converge visiblement ; les deux figures ci-dessous permettent de bien comprendre un tel comportement du prix p par rapport à la quantité q (figure de gauche), ainsi qu'au cours du temps t (figure de droite).



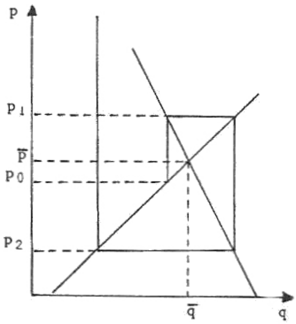


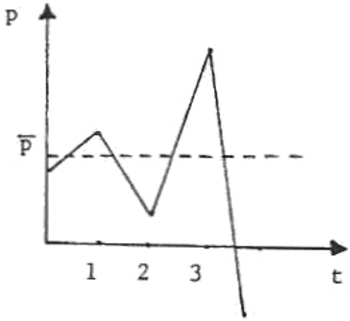
Si $\frac{c}{a} = -1$, c'est-à-dire si les deux droites $q = ap + b$ et $q = cp + d$ ont un coefficient angulaire de même valeur absolue et de signe contraire, on est en présence d'alternances entretenues ; cette situation est évoquée par les deux figures ci-dessous.





Enfin, si $| \frac{c}{a} | > 1$, le mouvement du prix est fait d'alternances explosives ; un exemple numérique (représentation graphique ci-dessous) est donné par $D_t = 10 - 0,5p_t$ et $S_t = 1 + p_{t-1}$.





3.1.4 La suite de Fibonacci

La suite de Fibonacci [2] se définit ainsi : ses deux premiers termes valent respectivement 0 et 1 ; tout autre terme est la somme des deux termes qui le précèdent. Il s'agit donc de la suite suivante : 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, ... ; cette suite est la solution particulière de l'équation récurrente (homogène, linéaire à coefficients constants, d'ordre 2) $y_t - y_{t-1} - y_{t-2} = 0$ qui obéit aux conditions initiales $y_0 = 0$ et $y_1 = 1$.

L'équation caractéristique $\lambda^2 - \lambda - 1 = 0$ admet les racines $\frac{1}{2}(1-\sqrt{5})$ et $\frac{1}{2}(1+\sqrt{5})$, de sorte que la solution générale s'écrit $y_t = C_1\left(\frac{1-\sqrt{5}}{2}\right)^t + C_2\left(\frac{1+\sqrt{5}}{2}\right)^t$.

En prenant successivement $t=0$ et $t=1$, on trouve $0 = C_1 + C_2; 2 = C_1(1-\sqrt{5}) + C_2(1+\sqrt{5})$; on en déduit $2 = \sqrt{5}(C_2 - C_1)$, puis $C_1 = -C_2 = -\frac{1}{\sqrt{5}}$.

Le terme général de la suite de Fibonacci est donc

$$y_t = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^t - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^t.$$

De façon inattendue, les termes (tous entiers) de cette suite s'écrivent ainsi à l'aide de nombres irrationnels.

La suite de Fibonacci trouve des applications variées et, parfois, surprenantes.

---

[2] Géomètre italien, encore connu sous le nom de Léonard de Pise, né à Pise vers 1175 et mort à une époque incertaine.

Ainsi, il existe dans la nature certaines plantes en forme de spirales ; ces courbes se présentent parfois en deux ensembles formant une révolution en sens opposé, chaque jeu étant constitué par un nombre fixe de spirales. Par exemple, chez la plupart des marguerites, il y a 21 spirales dans un sens et 34 dans l'autre ; dans les écailles de pommes de pin, on dénombre 5 spirales dans un sens et 8 dans l'autre ; dans les protubérances des ananas, on compte respectivement 8 et 13 spirales. Ces différents nombres (21 et 34, 5 et 8, 8 et 13) sont en fait constitués par deux termes consécutifs de la suite de FIBONACCI.

Par ailleurs, les nombres de FIBONACCI jouent aparemment un rôle important dans les arts. En effet, le nombre d'or , qui vaut approximativement 1,6 ou le rapport entre deux nombres consécutifs et supérieurs à 3 de la série de FIBONACCI, possède une valeur esthétique certaine : pour preuve, le rapport de 1 à 1,618 se retrouve dans de nombreux chefs-d'oeuvre de l'architecture ancienne et moderne.

La suite de FIBONACCI est également exploitée dans la construction d'algorithmes en programmation non linéaire [3]

Voici une autre application concrète et simple de cette théorie.

Soit un système de signalisation qui comprend seulement deux signes S_1 et S_2 (par exemple, les points et les traits en morse). On appelle message toute succession de ces deux signes. Sachant qu'un signe S_1 (resp. S_2) réclame exactement 1 (resp. 2) unité(s) de temps pour être transmis d'un endroit A vers un endroit B, et que le premier signe du message est toujours S_1, on peut calculer la capacité c de B par rapport à A, à savoir (d'après SHANNON) $c = \lim_{t\to\infty} \frac{\log_2 N(t)}{t}$, où $N(t)$ désigne le nombre total de messages qui peuvent être transmis de A vers B après t unités de temps. En partant du fait que le dernier signe de tout message d'une durée t est S_1 ou S_2, on établit l'équation récurrente $N_t = N_{t-1} + N_{t-2}$ avec $N_0 = 0$ et $N_1 = 1$. On trouve, comme ci-dessus, $N_t = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^t - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^t$, d'où $c = \lim_{t\to\infty} \frac{1}{t}\log_2\left[\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^t\left(1-\left(\frac{1-\sqrt{5}}{1+\sqrt{5}}\right)^t\right)\right] = \log_2\left(\frac{1+\sqrt{5}}{2}\right) \approx 0,7.$

---

(3) Voir, par exemple, la description d'une méthode du gradient pour l'obtention du maximum (ou minimum) d'une fonction d'une variable dans le livre de Mital K.V., *Optimization Methods in Operations Research and Systems Analysis*, Wiley, New Delhi, 1977, pp. 226-234.

3.1.5 L'oscillateur de Samuelson

Ce modèle économique célèbre fut présenté en 1939 par SAMUELSON ([4]), qui reçut en 1970 le prix Nobel de Sciences Economiques ; il fut ensuite repris par de nombreux auteurs, dont HICKS (autre prix Nobel, 1972) ; il est mieux connu sous le nom anglais de « multiplier-accelerator Model », dénomination qui sera justifiée ultérieurement. Nous allons le présenter dans une version fort simple, pour en étudier la stabilité.

Le revenu national y_t au cours de l'année t se décompose en la consommation c_t, l'investissement i_t et la dépense gouvernementale g_t à l'époque t : $y_t = c_t + i_t + g_t$; la consommation c_t est proportionnelle au revenu national y_{t-1} de la période précédente : $c_t = \gamma y_{t-1}$; enfin, l'investissement i_t est proportionnel à l'accroissement de consommation : $i_t = \alpha(c_t - c_{t-1})$; γ est un multiplicateur qui représente la propension marginale à consommer : il s'agit d'une constante strictement comprise entre 0 et 1 ; α, multiplicateur d'un accroissement, est appelé accélérateur par analogie avec la mécanique, car il multiplie en quelque sorte la vitesse d'évolution de la consommation : il s'agit d'une constante positive. Nous nous placerons dans le cas simple où la dépense gouvernementale exogène est constante, égale à g_0. En rassemblant toutes ces informations, on obtient l'équation récurrente d'ordre 2 : $y_t - \gamma(1+\alpha)y_{t-1} + \alpha\gamma y_{t-2} = g_0$.

Une solution particulière de cette équation est visiblement donnée par la constante $\frac{g_0}{1-\gamma}$.

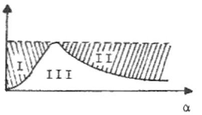
Recherchons la solution générale de l'équation homogène associée : $y_t - \gamma(1+\alpha)y_{t-1}+\alpha\gamma y_{t-2} = 0$. L'équation caractéristique $\lambda^2-\gamma(1+\alpha)\lambda+\alpha\gamma = 0$ possède deux racines λ_1, λ_2 (réelles ou complexes, éventuellement comptées deux fois) telles que $\lambda^2 - \gamma(1+\alpha)\lambda + \alpha\gamma = (\lambda - \lambda_1)(\lambda - \lambda_2) = 0, \lambda_1 + \lambda_2 = \gamma(1+\alpha), \lambda_1\lambda_2 = \alpha\gamma$ et $(1-\lambda_1)(1-\lambda_2) = 1-\gamma$.

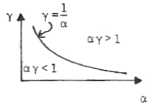
Nous devons considérer trois cas selon la nature des racines λ_1, λ_2.

- Premier cas : λ_1 et λ_2 sont deux racines réelles et distinctes, ce qui se produit lorsque $\gamma^2(1+\alpha)^2 > 4\alpha\gamma$; alors, $\lambda_1 = \frac{\gamma(1+\alpha)-\sqrt{\gamma^2(1+\alpha)^2-4\alpha\gamma}}{2}$, $\lambda_2 = \frac{\gamma(1+\alpha)+\sqrt{\gamma^2(1+\alpha)^2-4\alpha\gamma}}{2}$; ces deux racines sont de même signe (car $\lambda_1\lambda_2 = \alpha\gamma > 0$), et même toutes les deux positives (car $\lambda_1 + \lambda_2 = \gamma(1+\alpha) > 0$) ; de plus, elles sont toutes les deux inférieures à 1, ou bien toutes les deux supérieures à 1 (car $(1-\lambda_1)(1-\lambda_2) =$

([4]) Samuelson P., Interactions between the Multiplier Analysis and the Principle of Acceleration, *Review of Economic Statistics*, 1939, pp. 75-78.

$1-\gamma>0$) : la première (resp. seconde) éventualité se produit si et seulement si $\alpha\gamma<1$ (resp. $\alpha\gamma>1$). Ces considérations peuvent être interprétées graphiquement dans un repère orthonormé $(0\alpha,0\gamma)$. Le graphe de $\gamma=\frac{4\alpha}{(1+\alpha)^2}$ présente un maximum au point (1,1) et possède l'axe 0α comme asymptote ; la condition $\gamma^2(1+\alpha)^2-4\alpha\gamma>0$, équivalente à $\gamma>\frac{4\alpha}{(1+\alpha)^2}$, oblige le point (α,γ) d'appartenir à une des deux régions hachurées I ou II représentées ci-dessous.



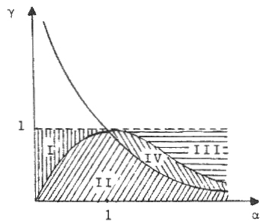


Par ailleurs, la courbe d'équation $\alpha\gamma=1$ (avec $\alpha>0,\gamma>0$) est une branche d'hyperbole équilatère : tout point (α,γ) situé en dessous (resp. au-dessus) de cette hyperbole obéit à l'inégalité $\alpha\gamma<1$ (resp. $\alpha\gamma>1$).
La solution générale du système est dès lors : $y_t=C_1\lambda_1^t+C_2\lambda_2^t+\frac{g_0}{1-\gamma}$. Le système est stable si et seulement si $\alpha\gamma<1$, puisque $0<\lambda_1<1$ et $0<\lambda_2<1$ entraînent $\lim_{t\to\infty}\lambda_1^t=\lim_{t\to\infty}\lambda_2^t=0$; lorsque $\alpha\gamma>1$, on est en présence d'un cas « explosif », car $\lambda_1>1$ et $\lambda_2>1$ donnent $\lim_{t\to\infty}\lambda_1^t=+\infty$.

- Deuxième cas : λ_1 et λ_2 sont deux racines réelles confondues, égales à $\lambda_1=\lambda_2=\frac{\gamma(1+\alpha)}{2}$, ce qui exige $\gamma^2(1+\alpha)^2=4\alpha\gamma$; l'équation caractéristique est alors le carré parfait de $\lambda-\frac{\gamma(1+\alpha)}{2}$ et la solution générale de l'équation récurrente vaut $y_t=(C_1+C_2t)\left[\frac{\gamma(1+\alpha)}{2}\right]^t+\frac{g_0}{1-\gamma}$. Géométriquement, le point (α,γ) se trouve sur la courbe d'équation $\gamma=\frac{4\alpha}{(1+\alpha)^2}$; selon qu'il se trouve sur la portion croissante $(0<\alpha<1)$ ou décroissante $(\alpha>1)$, le modèle sera stable (cas amorti, sans oscillations) ou instable (cas explosif, sans oscillations).
- Troisième cas : λ_1 et λ_2 sont deux racines complexes conjuguées, pour autant que $\gamma^2(1+\alpha)^2<4\alpha\gamma$, ce qui place le point (α,γ) en dessous de la courbe d'équation $\gamma=\frac{4\alpha}{(1+\alpha)^2}$ (c'est-à-dire dans la région III de la première figure ci-dessus). Comme l'équation caractéristique peut se mettre sous la forme $[\lambda-(x+iy)][\lambda-(x-iy)]=\lambda^2-2x\lambda+x^2+y^2$, avec $\lambda_1=x+iy$ et $\lambda_2=\overline{\lambda_1}=x-iy$, le module ρ de λ_1 et de λ_2 est visiblement égal à $\sqrt{\alpha\gamma}$, de sorte que l'on peut écrire $\lambda_1=\sqrt{\alpha\gamma}(\cos\theta+i\sin\theta)$ et $\lambda_2=\sqrt{\alpha\gamma}(\cos\theta-i\sin\theta)$, où θ est l'argument de λ_1. La solution générale de l'équation récurrente est : $y_t=(\alpha\gamma)^{\frac{t}{2}}(C_1\cos\theta t+$

$C_2 \sin\theta t) + \frac{g_0}{1-\gamma}$; on est dans le cas d'oscillations qui peuvent être explosives ($\alpha\gamma > 1$), entretenues ($\alpha\gamma = 1$) ou amorties ($\alpha\gamma < 1$).

La figure ci-dessous rassemble tous ces résultats.



Légendes : I stable : cas amorti sans oscillations
stable : cas amorti sans oscillations
II stable : cas amorti sans oscillations
III instable : cas explosif sans oscillations
instable : cas explosif sans oscillations
IV instable : cas explosif avec oscillations
instable : oscillations entretenues

3.2 Quelques modèles continus

3.2.1 Courbes de croissance

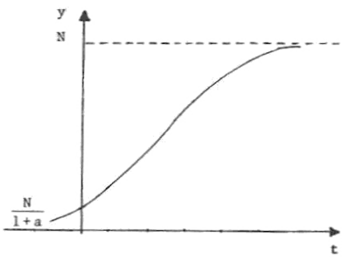
On doit souvent étudier l'évolution d'une grandeur y au cours du temps t ; par exemple, pour prévoir la demande de certains produits, on analyse le chiffre d'affaires d'une entreprise, le rythme de croissance d'un marché, etc. Les renseignements obtenus empiriquement sur cette fonction $y = f(t)$ peuvent être exploités en portant en abscisses les valeurs de la variable temporelle t et en ordonnées les valeurs correspondantes de $f(t)$: si l'on joint par une courbe continue les différents points ainsi tracés, on peut avoir une idée du type de liaison fonctionnelle reliant y à t. Par exemple, supposons les valeurs de la grandeur considérée mesurées à intervalles de temps réguliers. Si les valeurs sont en progression arithmétique, les points représentatifs de la fonction se disposent en ligne droite et la fonction $f(t)$ est linéaire. Par contre, si les valeurs de y sont en progression géométrique, il s'agit d'une fonction exponentielle ab^t qui se reconnaît par l'alignement des points relatifs aux données sur du papier semi-logarithmique. Enfin si, en portant en abscisses les $\log t$ et en ordonnées les $\log y$ correspondants, on obtient des points situés sur une même droite, la fonction est du type puissance at^b.

Une autre façon de reconnaître ces types de courbes, ainsi que d'autres également courants, consiste en l'examen de l'évolution de la pente de la courbe, ou du rapport de la pente à la grandeur elle-même, ou encore de l'élasticité. En effet, si le graphe de y' (resp. $\frac{y'}{y}; E_t y$) se réduit à une droite horizontale, la fonction $f(t)$ est linéaire (resp. exponentielle ; du type puissance). Dans le même ordre d'idées, si le graphe de y' (resp. $\frac{y'}{y}$; $E_t y$) est une droite oblique, celui de $f(t)$ est une parabole (resp. une parabole logarithmique ; une courbe de puissance modifiée) d'équation $y = at^2 + bt + c$ (resp. $y = ab^t c^{t^2}$; $y = at^b c^t$). Pour obtenir l'expression analytique de ces fonctions, il suffit de résoudre une équation différentielle élémentaire.

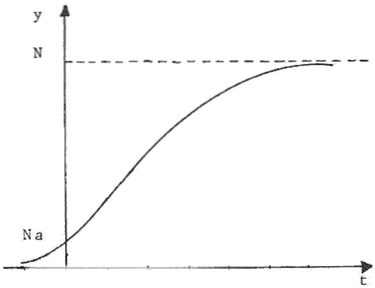
Les fonctions précédentes ont la particularité d'être constamment croissantes (pour autant que la dérivée reste positive) et possèdent alors l'infini comme limite lorsque t tend vers l'infini. Elles ne sont donc pas toujours adéquates pour décrire la réalité. En effet, beaucoup de grandeurs partent d'une valeur initiale (souvent nulle) et augmentent sans cesse, sans toutefois dépasser une valeur de *saturation* déterminée. Pour tenir compte de ce phénomène, on fait appel à des fonctions dont le graphe admet une asymptote horizontale (lorsque t tend vers l'infini), d'équation $y = N$ où N est la valeur de saturation en question. Plusieurs fonctions répondent à ces exigences.

La *courbe logistique*, définie par l'égalité $y = N(1 + ae^{-bt})^{-1}$, où N, a et b sont des constantes positives, est obtenue en supposant la vitesse de croissance y' de y proportionnelle à la grandeur y elle-même, mais aussi à la distance entre y et la valeur de saturation N (c'est-à-dire à la différence $N - y$, qui est appelée la *tension*).

Le graphique de cette fonction est une courbe en forme de « S incliné », débutant par une allure pratiquement exponentielle (car la croissance est proportionnelle à la grandeur), puis s'infléchissant et tendant finalement vers un *seuil* (car la croissance est proportionnelle à la tension) ; ce phénomène se traduit par une diminution progressive du coefficient angulaire de la tangente.

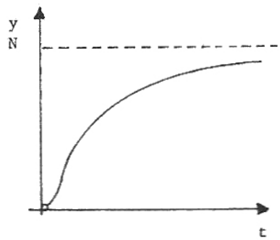


On peut reconnaître assez facilement cette courbe dans la pratique : les points doivent être alignés lorsque l'on porte en abscisses le temps t et en ordonnées le logarithme népérien du rapport de la pente y' au carré y^2 de la grandeur.

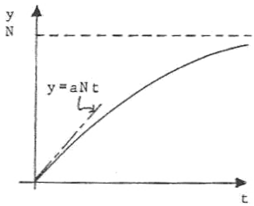


La *courbe de Gompertz*, qui obéit à l'équation $y = Na^{b^t}$, où a, b et N sont des constantes positives avec toutefois $b < 1$, est telle que le logarithme népérien du rapport de la pente y' à la grandeur y est une fonction linéaire du temps, le graphe de $\ln \frac{y'}{y}$ étant en réalité une droite de coefficient angulaire négatif. Une courbe de Gompertz ressemble fortement à une courbe logistique.

Une hyperbole logarithmique, d'équation $y = Ne^{\frac{-a}{t}}$, s'obtient en remarquant que le graphe de l'élasticité $E_t y$ est une hyperbole ayant les axes de coordonnées comme asymptotes ; cette fonction n'est évidemment pas définie en $t = 0$, mais la limite de y, pour t tendant vers 0 (par valeurs supérieures) est nulle.



Enfin, lorsque la vitesse de croissance y' de y est proportionnelle à l'écart entre le seuil N et la grandeur y elle-même, on obtient une exponentielle modifiée d'équation $y = N(1 - e^{-at})$.



Au voisinage de $t = 0$, cette courbe se comporte pratiquement comme la droite d'équation $y = aNt$; elle ne possède pas de point d'inflexion, mais tend également asymptotiquement vers l'horizontale $y = N$ lorsque t tend vers l'infini.

Voici un tableau résumant les paragraphes précédents.

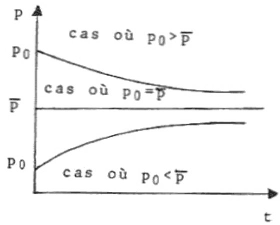
Si le graphique de	est (proche d')une	le type de courbe suggérée est une	d'expression analytique
y'	droite horizontale	ligne droite	$y = at + b$
y'	droite oblique	parabole	$y = at^2 + bt + c$
$\frac{y'}{y}$	droite horizontale	courbe exponentielle	$y = ab^t$
$\frac{y'}{y}$	droite oblique	parabole logarithmique	$y = ab^t c^{t^2}$
$E_t y$	droite horizontale	courbe de puissance	$y = at^b$
$E_t y$	droite oblique	courbe de puissance modifiée	$y = at^b c^t$
$E_t y$	hyperbole	hyperbole logarithmique	$y = Ne^{\frac{-a}{t}}$
$\ln y'$	droite descendante	courbe exponentielle modifiée	$y = N(1 - e^{-at})$
$\ln(\frac{y'}{y})$	droite descendante	courbe de Gompertz	$y = Na^{b^t}$
$\ln(\frac{y'}{y^2})$	droite descendante	courbe logistique	$y = \frac{N}{1+ae^{-bt}}$

3.2.2 Initiation à l'analyse dynamique de la demande

Dans l'étude statique, l'influence du temps sur les variables est ignorée. L'analyse dynamique est plus élaborée puisqu'elle tient compte des variations des grandeurs considérées au cours du temps.

En guise d'exemple, considérons le modèle simple où les équations de la demande et de l'offre sont respectivement données par $D = ap + b$ et $S = cp + d$ (où a, b, c et d sont des constantes). Ici, contrairement au cas statique, nous considérerons les prix p variables au cours du temps t et nous supposerons que le taux de variation du prix (ou vitesse de changement de prix) est proportionnelle à la différence entre la demande et l'offre. Cette hypothèse se traduit par l'égalité $\frac{dp}{dt} = \alpha(D - S)$, où α est la constante de proportionnalité. On est ainsi en présence d'une équation différentielle du premier ordre, à savoir $\frac{dp}{dt} = \alpha[(a-c)p+b-d]$ ou $\frac{dp}{dt}+\alpha(c-a)p = \alpha(b-d)$. La formule de Leibnitz donne $p(t) = Ce^{\alpha(a-c)t}+\frac{d-b}{a-c}$; la constante d'intégration C peut être calculée en faisant $t = 0$: $p(0) = p_0 = C + \frac{d-b}{a-c}$, soit $C = p_0 - \frac{d-b}{a-c}$: C représente donc la différence entre le prix p_0 à l'instant initial et la valeur constante $\overline{p} = \frac{d-b}{a-c}$ qui assure l'équilibre du marché. Au total, $p(t) = (p_0-\overline{p})e^{\alpha(a-c)t}+\overline{p}$. Généralement, les constantes α et c sont positives, tandis que a est négative ; dans ces conditions, le prix $p(t)$ converge vers $\overline{p}$ lorsque t tend vers l'infini ; le niveau $\overline{p}$ peut alors être interprété comme un équilibre, qui est ici dynamiquement stable. Ce terme $\overline{p}$ est en réalité une solution particulière de l'équation différentielle proposée. Nous avons donc une interprétation économique des solutions particulière ($p_2 = \overline{p}$) et générale ($p_1 = (p_0 - \overline{p})e^{\alpha(a-c)t}$) : p_2 représente le niveau d'équilibre du système, tandis que p_1 est l'écart depuis cet équilibre.

La convergence de la solution générale $p(t) = p_1 + p_2$ résulte du fait que p_1 possède $\overline{p}$ pour limite ; il y a lieu de distinguer trois cas selon le signe de la différence entre p_0 et $\overline{p}$: si $p_0 - \overline{p}$ est positif (resp. négatif), la courbe $p(t)$ se rapproche asymptotiquement de la droite horizontale d'équation $p(t) = \overline{p}$ en étant toujours située au-dessus (resp. en dessous) de celle-ci ; par contre, si $p_0 = \overline{p}$, le prix $p(t)$ reste bien sûr constant, égal à $\overline{p}$.



3.2.3 Modèle du fardeau de la dette publique

Soit un gouvernement qui a recours à des emprunts répétés.

Si l'impôt est le seul moyen utilisé pour couvrir les intérêts de ces emprunts, il est clair que les taxes ainsi imposées augmenteront sans cesse (pour autant que le taux d'intérêt ne baisse pas), ce qui constituera un « fardeau » pour le peuple. D'après l'économiste Domar [5], ce fardeau ne doit pas être mesuré par le montant global des impôts, mais bien par le rapport entre l'impôt supplémentaire et le revenu national ; en effet, si le revenu national croît, une augmentation absolue de l'impôt peut effectivement ne pas être pénible pour la population.

Ainsi, à tout instant t, le fardeau $F(t)$ s'exprime en fonction de l'impôt $I(t)$ et du revenu $R(t)$ à l'aide de la relation $F(t) = \frac{I(t)}{R(t)}$.

Domar a étudié le fardeau de la dette publique sous diverses hypothèses de croissance du revenu national.

Le taux de l'intérêt est supposé constant, égal à τ, de sorte que les intérêts nécessitent un impôt proportionnel à la dette publique $D(t)$ à l'instant t ; de façon précise, $I(t) = \tau D(t)$.

Le revenu national augmente avec un taux relatif constant α (avec $0 < \alpha < 1$), ce qui fournit une équation différentielle du premier ordre $R'(t) =$

---

[5] E.D. Domar, « The Burden of the Debt' and National Income », *American Economic Review*, 1944, pp. 798-827

$\alpha R(t)$, dont la solution générale est de la forme $R(t) = R_0 e^{\alpha t}$ (où R_0 désigne le revenu national à l'instant initial $t = 0$).

De plus, une fraction constante du revenu national a toujours tendance à devenir une nouvelle dette publique ; cette condition se traduit mathématiquement par l'égalité $D'(t) = \beta R(t)$, avec $0 < \beta < 1$. Dès lors, on obtient la relation suivante : $D'' - \alpha D' = 0$; il s'agit d'une équation différentielle linéaire à coefficients constants d'ordre 2, dont la solution générale s'écrit comme suit : $D(t) = C_1 + C_2 e^{\alpha t}$, avec $D_0 = C_1 + C_2$ et $\beta R_0 = \alpha C_2$, d'où $C_2 = \frac{\beta}{\alpha} R_0$ et $C_1 = D_0 - \frac{\beta}{\alpha} R_0$ (D_0 désignant la dette à l'instant $t = 0$). Au total,

$$D(t) = \frac{\beta}{\alpha}\ R_0 e^{\alpha t} + D_0 - \frac{\beta}{\alpha}\ R_0$$

En conséquence, le fardeau vaut $F(t) = \tau \frac{D(t)}{R(t)} = \tau \frac{\beta R_0 e^{\alpha t} + \alpha D_0 - \beta R_0}{\alpha R_0 e^{\alpha t}}$.

Il est possible de voir ce que deviendrait le fardeau si le gouvernement continuait indéfiniment à emprunter (à ce même taux τ) ; il suffit de calculer la limite de $F(t)$ lorsque t tend vers l'infini. Par la règle de l'Hospital, on trouve $\lim_{t \to +\infty} F(t) = \frac{\tau\beta}{\alpha}$; cette valeur est constante et dépend des paramètres α, β et τ.

3.2.4 Croissance d'un niveau par un taux

Au cours du temps, le niveau d'un stock se constitue généralement par intégration d'un flux, ce dernier étant souvent issu de la variation d'un stock. De la sorte, on peut affirmer que les niveaux sont alimentés par des taux et se vident également selon des taux.

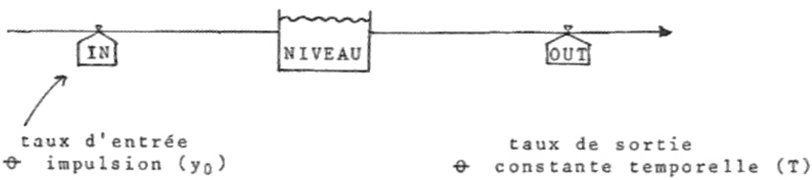
En guise d'exemples, un réservoir hydrologique dépend de la quantité d'eau qui lui est déversée par unité de temps, un capital se forme grâce au taux d'accumulation du capital, le stock d'un marché agricole varie, notamment, en fonction des taux de plantation et de récolte, une population résulte du taux de naissance net (nombre de naissances par an moins le nombre de décès par an), etc.

Le problème consiste à étudier comment varie le niveau d'un stock en fonction du temps. ([6])

([6]) Ce sujet est étudié de façon détaillée dans le cours *Dynamique des systèmes*,par C. De Bruyn, 3e cycle F.N.R.S. d'aide à la décision et recherche opérationnelle, Université de Liège, 1985, pp. 13-24.

Nous allons nous contenter d'examiner le cas d'un flot (monétaire ou autre) alimenté uniquement au temps initial ($t = 0$) par une quantité donnée y_0 d'un input du type « impulsion », la quantité y_0 étant donc livrée instantanément en $t = 0$. Ce flot est toutefois soumis à un délai en ce sens que l'input est « retenu » en s'accumulant dans un niveau, puis est entièrement redistribué en output par la suite. L'*ordre n du délai* est le nombre de fois que le flot est freiné avant d'être complètement délivré. Par ailleurs, le temps moyen pendant lequel le flot sera retenu dans le délai est supposé constant, égal à T : le taux de sortie du niveau $y(t)$ à l'instant t vaut donc $\frac{y(t)}{T}$.

Considérons tout d'abord un délai d'ordre 1 qui peut être schématisé comme suit : [7]



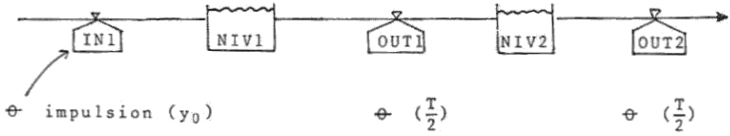
Graphe de flux pour un délai d'ordre 1

En désignant par $IN1(t)$ (resp. $OUT1(t)$) le taux d'entrée (resp. le taux de sortie) et par $y1(t)$ le niveau à l'instant t pour un délai d'ordre 1, on obtient : $y1(t + \Delta t) = y1(t) + \Delta t[IN1(t) - OUT1(t)]$, ou $\frac{y1(t+\Delta t)-y1(t)}{\Delta t} = -OUT1(t) = -\frac{y1(t)}{T}$ car $IN1(t) = 0$ du fait que le niveau initial a été constitué par une impulsion à l'instant $t = 0$. En passant à la limite pour Δt tendant vers 0, on tombe sur l'équation différentielle : $\frac{dy1(t)}{dt} = -\frac{1}{T}y1(t)$, dont la solution est donnée par $y1(t) = y_0 e^{-\frac{t}{T}}$, de sorte que le niveau décroît de façon exponentielle.

Un délai d'ordre 2 est une séquence de deux délais d'ordre 1 pour lesquels la constante de délai T est affectée pour moitié à chacun.

---

[7] Le symbole $\ominus$ désigne une grandeur exogène c'est-à-dire une grandeur qui n'est pas expliquée par le système, mais est déterminée en dehors de celui-ci.



Graphe de flux pour un délai d'ordre 2

La réponse à un délai d'ordre 2 s'obtient par l'étude des deux niveaux implicites $y1(t)$ (pour NIV1) et $y2(t)$ (pour NIV2), l'input d'impulsion (en $t = 0$) étant encore égal à y_0. Avec des notations évidentes, on trouve (grâce aux résultats précédents) :

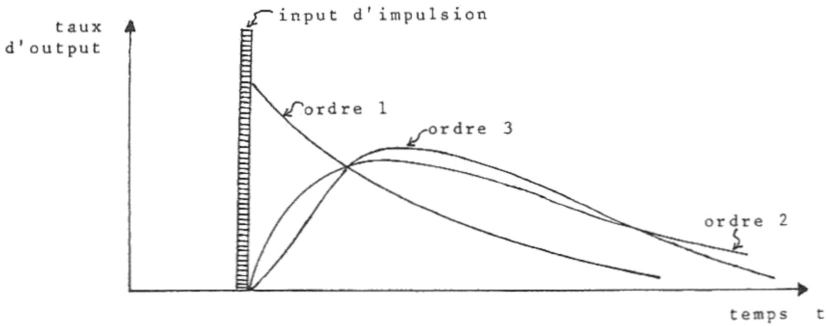
$$OUT1(t) = \frac{y1(t)}{\frac{T}{2}} = \frac{y_0 e^{-2\frac{t}{T}}}{\frac{T}{2}}, \quad OUT2(t) = \frac{y2(t)}{\frac{T}{2}},$$

$$y2(t + \Delta t) = y2(t) + \Delta t[OUT1(t) - OUT2(t)],$$

ce qui conduit à l'équation différentielle $\frac{dy2(t)}{dt} = \frac{y1(t)}{\frac{T}{2}}$ ou $\frac{dy2(t)}{dt} + \frac{2y2}{T} = \frac{2}{T}e^{\frac{-2t}{T}}$. La solution particulière de cette équation, qui devient nulle lorsque t tend vers 0, est la forme $y2(t) = \frac{2t}{T}e^{\frac{-2t}{T}}$.

De la même façon, un délai d'ordre 3 conduit à l'équation différentielle $\frac{dy3}{dt} + 3\frac{y3}{T} = \frac{9t}{T^2}y_0 e^{\frac{-3t}{T}}$, dont on retient la solution $y3(t) = 4,5\frac{t^2}{T^2}e^{-\frac{3t}{T}}$.

Les réponses des délais d'ordre 1, 2 et 3 à un input « impulsion » sont représentées ci-dessous. On notera que l'ordre 3 est intéressant dans la mesure où il correspond le plus souvent à la réalité, avec des rendements d'abord croissants, puis décroissants.



Chapitre 4

Fonctions et mathématique financière

par *Daniel JUSTENS*

L'univers économique ne se peut concevoir sans un marché de capitaux. L'érosion monétaire jointe à la nécessité de rémunérer les capitaux prêtés contraignent le mathématicien à proposer des modèles mathématiques d'évolution de capitaux au cours du temps. Le propre de la modélisation dans ce cas particulier est que son utilisation est uniquement question de convention. Le physicien peut interroger la nature pour tester son modèle. Le mathématicien de la finance peut concevoir un modèle parfait, cohérent et efficace et lui voir préférer, dans la réalité des choses, un modèle absurde, contradictoire et insuffisant.

Cette particularité de la mathématique financière nous contraint en permanence à préciser le contexte, théorique ou pratique, dans lequel se situe un problème à traiter ; les conventions particulières adoptées - plus ou moins arbitrairement - doivent également être connues.

La description de l'évolution d'un capital devrait être fonction de paramètres économiques. On peut la concevoir :

1. en fonction du taux d'inflation.
2. en fonction du pouvoir d'achat réel.

3. relativement à l'évolution d'autres monnaies.
4. indépendamment des paramètres économiques réels.
5. dans un univers aléatoire modélisé au moyen du calcul des probabilités.

Comme nous venons de l'exposer, tout n'est que question de convention. Celle qui est presque universellement acceptée est la suivante :

- On considère que l'évolution d'un capital doit être une fonction non décroissante du temps.
- La seule dépendance vis-à-vis du contexte économique doit se traduire par un seul nombre constant : le taux d'intérêt. Celui-ci doit être strictement positif.

Cette convention arbitraire est très loin d'être réaliste. La traduction d'un univers économique mouvant et instable par un seul nombre constant paraît particulièrement absurde. Nous verrons comment on peut améliorer cette contrainte.

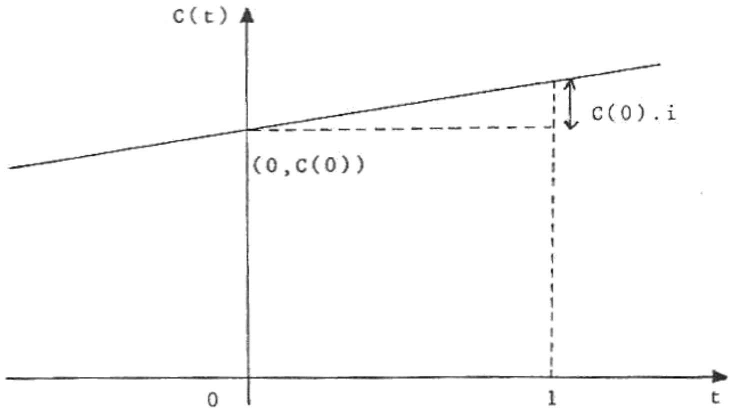
4.1 Modélisation linéaire et intérêt simple

Il semble relativement admis que la valeur d'un capital est fonction du moment de son échéance. Un montant de 100 000 aujourd'hui est « évidemment » préférable à un montant identique échéant dans deux ans. Cette évidence conduit à la première tentative de modélisation. Dans un but de simplification louable, on se tourne vers la représentation linéaire. Soit un capital échéant aujourd'hui. Nous le notons $C(0)$. Pour toute valeur t du temps, on aura :

$$C(t) = C(0).(1 + i.t) \tag{4.1}$$

Dans cette expression : $C(t)$ désigne la valeur du capital au temps t ; i est un coefficient de proportionnalité appelé taux d'intérêt. En toute cohérence, le taux doit être exprimé dans une unité compatible avec l'unité de temps. Pour évidente qu'elle puisse sembler, cette dernière remarque n'est pas vide de sens : beaucoup de contrats sont proposés avec remboursements mensuels ou semestriels alors que le taux d'intérêt donné est annuel. En toute généralité, i est supposé positif.

On peut dès lors concevoir la première représentation graphique.



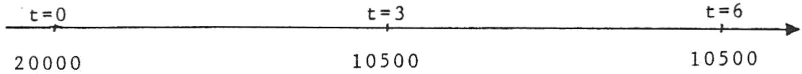
4.1.1 Critique de la modélisation linéaire

L'accroissement de capital après une unité de temps est de $C(0).i$. L'interprétation du taux d'intérêt comme coefficient de proportionnalité est donc cohérente. En est-il de même du modèle linéaire ? La réponse à cette question est non. En effet, la pente de la droite étant positive (i supposé positif), on peut trouver dans tous les cas une valeur de t pour laquelle le capital est nul. Pour toute valeur de t antérieure, le capital est négatif. On obtient donc, en interprétant ce résultat, une dette qui s'éteint d'elle-même, ce qui est déjà surprenant, et qui se transforme petit à petit en un capital de plus en plus confortable. A titre d'exercice, on peut se demander ce qui se passe avec un capital initial $C(0)$ négatif, donc une dette.

Dans tous les cas, on est tenté de se détourner du modèle linéaire en ce qui concerne le passé. Il faut cependant constater que cette modélisation est utilisée quotidiennement : elle porte le nom d'escompte simple !

Une autre critique va nous montrer l'incohérence de la modélisation linéaire : un même problème posé correctement à des moments différents conduit à des résultats significativement différents.

Un exemple élémentaire va nous en convaincre. Considérons un emprunt de 20 000 remboursable en deux versements de 10 500 échéant respectivement à l'issue des troisième et sixième mois. On peut traduire cette situation par le graphique suivant :



La mise en équation doit se faire en égalant les valeurs perçues et dépensées en tenant compte de leur évolution au cours du temps. Notons i le taux mensuel à estimer.

a. Mise en équation en début de contrat : $t = 0$.

On obtient facilement :

$$20\,000 = 10\,500(1 - 3i) + 10\,500(1 - 6i)$$

Il faut évidemment tenir compte des positions relatives des dates d'échéance. On trouve :

$$i = 0,010582$$

b. Mise en équation en fin de contrat : $t = 6$.

On obtient cette fois :

$$20\,000(1 + 6i) = 10\,500(1 + 3i) + 10\,500.$$

On en tire :

$$i = 0,012994.$$

On remarque que tous les problèmes financiers se traduisent par des relations du premier degré. Comme toujours, la linéarité est utilisable pour des problèmes à très court terme et en première approximation uniquement.

Il y a plus amusant encore : un cas réel va nous conduire à une équation parfaitement intégrée au système défini, équation dont la solution est totalement absurde. Voici cet exemple précis. Une société pratiquant la vente à tempérament propose un micro-ordinateur avec écran, imprimante et traitement de texte pour un acompte de 7 499 et 24 versements de 2 165. La valeur comptant affichée du bien est 49 999. L'ordinateur en question peut se trouver couramment dans le commerce au prix de 29 999 ! Cette dernière information modifie notablement la mise en équation. En effet, la quantité véritablement prêtée dans ce cas-ci n'est pas 49 999 - 7 499 mais bien 29 999 - 7 499 (valeur réelle du bien moins acompte).

Lorsque la mise en équation se fait en $t = 0$, la solution donnant le taux est admissible (encore qu'inexacte). On a l'équation :

$$22\,500 = 2\,165(1-i) + 2\,165(1-2i) + \cdots + 2\,165(1-24i)$$

On en tire :

$$22\,500 = 2\,165 \times 24 - 2\,165i(1+2+\cdots+24)$$

Et finalement :

$$\begin{aligned} 2\,165i \times \frac{24 \times 25}{2} &= 2\,165 \times 24 - 22\,500 \\ i &= 0,04536. \end{aligned}$$

Ce résultat est surtout amusant si l'on se rappelle que ce taux correspond aux interversements et est donc mensuel. Le taux annuel correspondant s'obtient en théorie de l'intérêt simple par simple proportionnalité (voir plus loin).

On obtient : $i_{annuel} = 0,5443$.

Ce dernier résultat peut sembler excessif. Et pourtant ...

Abordons à présent la mise en équation en $t = 24$, en fin de contrat. On obtient l'équation :

$$22\,500(1+24i) = 2\,165(1+23i) + \cdots + 2\,165(1+i) + 2\,165$$

On peut écrire :

$$22\,500 + 22\,500 \times 24i = 2\,165 \times 24 + 2\,165i(23+22+\cdots+1)$$

ou encore :

$$\begin{aligned} 22\,500 \times 24i - \frac{2\,165 \times 23 \times 24}{2} i &= 2\,165 \times 24 - 22\,500 \\ -57\,540i &= 29\,460 \\ i &= -0,51199 \end{aligned}$$

Malgré l'absurdité du résultat (un taux négatif alors que le montant des remboursements est nettement supérieur au montant emprunté !), on peut en donner une interprétation. Le taux trouvé est sur-infini : en ce sens que, quel que soit le taux fini choisi pour calculer les remboursements, il est impossible d'arriver au montant exigé par la société. A ce stade, il n'est pas totalement inutile de repréciser que toutes les données du problème sont réelles, que cette proposition de crédit existe.

4.1.2 Vers une certaine amélioration

On peut aisément éliminer la première objection. Voyons comment.

Dans ce qui suit, il est nécessaire de distinguer le passé du futur. En supposant t strictement positif, t peut désigner le futur, $-t$ le passé. Considérons à présent un capital $C(0)$ échéant aujourd'hui. Quelle était sa valeur il y a t unités de temps ? Autrement dit : que vaut $C(-t)$? Cette valeur peut évoluer dans le temps selon notre modèle linéaire : en effet, par rapport à l'instant $-t$, l'instant 0 représente le futur. On a donc en utilisant (4.1) :

$C(0) = C(-t).(1 + it)$

On en tire :

$$\boxed{C(-t) = \tfrac{C(0)}{1+it}}$$

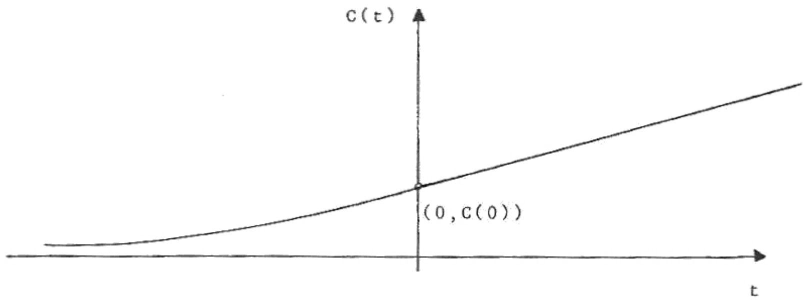
On peut se demander à quoi va ressembler notre modèle linéaire modifié. Une façon simple de décrire l'évolution du capital est la suivante. Soit t strictement positif :

$$\boxed{\begin{array}{c} C(t) = C(0).(1 + it) \\ \\ C(-t) = \dfrac{C(0)}{1 + it} \end{array}} \tag{4.2}$$

La représentation analytique (4.2) n'est pas mathématiquement souhaitable. Il est préférable de définir $C(t)$ par morceaux. On obtient alors :

$$\boxed{\begin{aligned} C(t) &= C(0).(1+it) & t \geqslant 0 \\ C(t) &= \frac{C(0)}{1-it} & t < 0 \end{aligned}} \tag{4.3}$$

On remarque que la fonction $C(t)$ ainsi obtenue est continue. A titre d'exercice, on peut se demander si elle est dérivable. La réponse à cette dernière question est oui. Graphiquement, on obtient donc :



La partie intéressante à étudier est évidemment le voisinage du point $(0, C(0))$.

Le modèle modifié est-il cohérent ? Encore une fois la réponse est non. Lorsque plusieurs capitaux sont à considérer avec des échéances différentes dans le temps, des mises en équations également cohérentes vis-à-vis du système ainsi défini conduisent à des résultats d'autant plus différents que la période considérée est grande.

Reprenons l'exemple qui a déjà été traité. L'équation obtenue est du deuxième degré. Lorsqu'il y a n remboursements, l'équation à traiter est de n-ième degré.

En $t = 0$:

$$20\,000 = \frac{10\,500}{1+3i} + \frac{10\,500}{1+6i}$$

En $t = 6$, l'équation est inchangée : tous les montants sont à calculer dans le futur.

4.1.3 Lien avec la réalité pratique

Les relations (4.1) et (4.3) sont cependant largement utilisées. L'utilisation de (4.1) pour le passé porte le nom d'actualisation commerciale, celle de (4.3) d'actualisation rationnelle.

Ces relations, quand elles sont utilisées, ne le sont pas dans un contexte continu. On admet que le temps ne peut se diviser en intervalles de longueur quelconque. Seuls certains rationnels seront admis, le choix de ceux-ci étant fonction du contexte. Ainsi lors d'un prêt, l'organisme financier demandera une rémunération en considérant le jour comme l'unité « naturelle » à utiliser. Lors d'un dépôt, ce même organisme distribuera la rémunération en fonction du nombre de quinzaines. . .une unité étrange qui peut valoir 13, 14, 15 ou 16 jours. La droite théorique va donc se transformer en pratique en fonction en escalier. Lorsque le jour est l'unité choisie, on peut considérer la droite comme une excellente approximation de l'escalier réel. Lorsque c'est la quinzaine, l'approximation devient d'autant plus mauvaise que les dates de dépôt et de retrait s'éloignent des points de subdivision choisis.

Un exemple peut nous en convaincre.

Considérons un capital de 100 000 placé au taux annuel de 6% (il est totalement inutile de compliquer les données expérimentales : le raisonnement étant élémentaire, pourquoi les exemples ne le seraient-ils pas ?).

Les quinzaines constituent en fait des « demi-mois » commençant le 1er et le 16 de chacun d'eux. Un dépôt se faisant le 1er pour se terminer le 1er d'un des mois suivants va donner une valeur réelle proche de la valeur théorique linéaire, la seule différence étant due au manque de cohérence de la notion temporelle de quinzaine.

Par contre si l'on considère un montant de 100 000 déposé le 10 janvier et retiré le 10 juin de la même année, on obtient :

Valeur linéaire théorique :

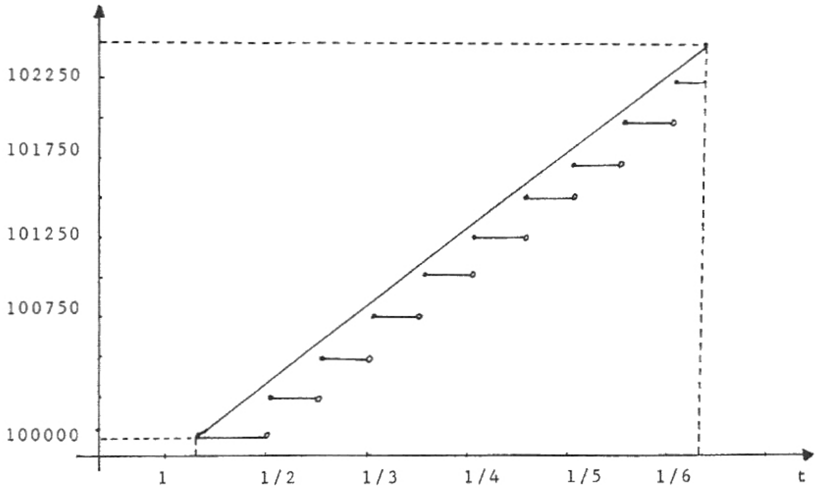
$$100\,000.(1 + \frac{151}{365}.0,06) = 102\,482$$

Valeur réelle :

$$100\,000.(1 + \frac{9}{24}.0,06) = 102\,250 \text{ (}^1\text{)}$$

([1]) Pour le calcul des durées de manière simple : utiliser la table située dans [36, p. 68]

Graphiquement, la situation est la suivante : (pour rendre les différences visibles, on ne représente que la partie intéressante de l'ordonnée).



Pour simplifier la représentation graphique, nous avons considéré les quinzaines comme de « vrais » 1/24. La différence après 5 mois est loin d'être négligeable : 232. Une règle de trois amusante consiste à calculer l'économie réalisée par an par l'organisme financier si l'on sait que pour les grandes banques belges, le montant global moyen des dépôts est de l'ordre de 100 000 000 000.

Un autre exercice consiste à introduire la notion de prime de fidélité qui augmente le taux d'intérêt lorsque le dépôt est resté constant pendant une durée déterminée. Dans ce cas, l'un des sauts est nettement plus important que les autres.

4.1.4 La notion de taux de chargement

La théorie de l'intérêt simple s'enrichit de la notion de taux de chargement, une notion pratique fréquemment confondue avec le taux d'intérêt. Elle est cependant totalement différente. En effet considérons un emprunt de montant 100 000 remboursable en 20 mensualités. On va considérer que chaque mensualité comportera une partie « remboursable proprement dit », appelée amortissement et une partie « rémunération du capital emprunté », les frais financiers.

La pratique des taux de chargement se base sur le raisonnement suivant : chaque remboursement est composé d'un vingtième du capital emprunté, ici 5000 et d'une charge financière constante proportionnelle au montant emprunté. Le coefficient de proportionnalité est notre taux de chargement. Pour un taux de $0,4\%$, le remboursement mensuel sera de $a = \frac{100\,000}{20} + 100\,000 \times 0,004 = 5400$.

Ce taux ne peut pas être le taux d'intérêt. En effet, les remboursements diminuent progressivement la dette restant due et donc les frais financiers ne peuvent être constants. Une charge financière constante ne se peut d'ailleurs JAMAIS concevoir dès lors que l'on rembourse progressivement le capital.

Alors qu'en est-il du taux de ce contrat ? Ce calcul est abordé dans la suite de manière correcte. Il est nécessaire d'utiliser la modélisation exponentielle pour obtenir des résultats satisfaisants. On trouve néanmoins dans la littérature et dans les textes de loi la relation :

$$i = \frac{24.r.n}{(n+1)}$$

Dans cette relation, n représente le nombre de remboursements et r le taux de chargement.

Cette relation n'est pas correcte. Outre le fait qu'elle se base exclusivement sur l'intérêt simple, elle est calculée à partir de deux hypothèses simplificatrices contradictoires. On suppose pour commencer que la charge financière est *constante* par remboursement et donc la partie amortissement. Des amortissements constants induisent des soldes restant dus en progression arithmétique et partant des frais financiers de même. Ce sont ces frais financiers *variables* qui servent au calcul de i. Une étude complète de la pratique des taux de chargement est faite dans [36, p. 148], étude qui montre que la relation approximative donne de bons résultats lorsque la durée du contrat est comprise entre 30 et 36 mois. En dehors de cet intervalle, les différences entre taux réel et approximation deviennent rapidement significatives.

4.2 Modélisation exponentielle et intérêt composé

Face à l'insuffisance évidente du modèle linéaire (4.1) et du modèle linéaire corrige (4.3), une deuxième tentative est la modélisation exponentielle :

$$\boxed{C(t) = C(0).(1+i)^t}. \tag{4.4}$$

Cette deuxième tentative est nettement plus fructueuse que la précédente et ce pour les raisons suivantes :

- La fonction exponentielle étant toujours définie positive, des absurdités telles celles rencontrées dans le modèle linéaire sont évitées.
- Quelle que soit la mise en équation envisagée, elle conduit à une solution unique indépendante des conventions choisies (voir [17, 36, 50]).

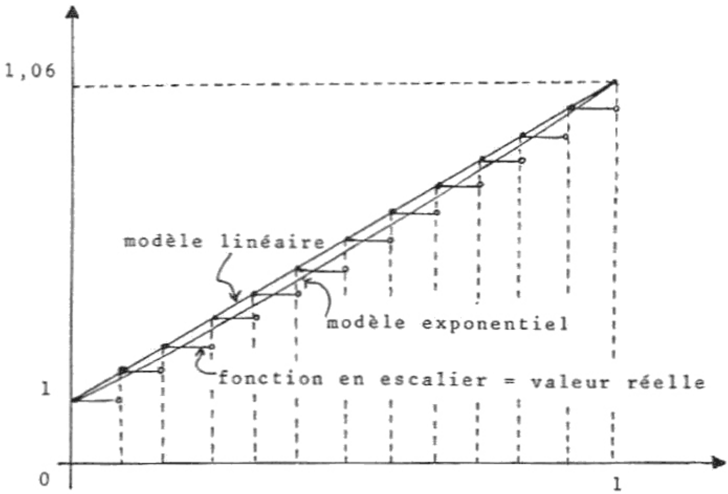
Des objections persistent néanmoins :

- Le changement d'unité de temps induit des formules non triviales de changement de taux. Cette difficulté est un prétexte facile pour les organismes financiers à utiliser des relations approximatives ...dont ils sont bien évidemment les seuls bénéficiaires.
- Pour le court terme et notamment pour les ventes à tempérament, l'usage et aussi la législation optent pour le calcul par intérêt simple et donc pour les relations (4.1) ou (4.3). (pour la partie légale, voir [36, §8.2, p. 88]).
- Enfin l'usage de la relation exponentielle n'est jamais envisagé directement : les organismes financiers se contentent de répéter des capitalisations à intérêt simple. Pour des valeurs naturelles du temps, cette façon de procéder donne lieu à des valeurs qui coïncident avec celles données par le modèle exponentiel. Pour des valeurs intermédiaires, cette façon de procéder conduit à la notion de capitalisation mixte, mélange malheureux d'intérêt simple et composé. (Voir [36] et [49]). Pour la capitalisation mixte, il importe de décomposer la partie temporelle t en partie entière + partie décimale. On pose : $t = n + d$ où $n = E(t)$, $d = D(t)$. La modélisation est alors :

  $$\boxed{C(t) = C(0).(1+i)^n.(1+d.i)} \tag{4.5}$$

  Graphiquement cette modélisation coïncide avec la modélisation exponentielle, pour les valeurs entières et prend la corde de celle-ci pour toute valeur intermédiaire. Un graphique est proposé dans la suite.
- Pour des valeurs inférieures à l'unité, la modélisation exponentielle inférieure à la modélisation linéaire, conduit à la situation suivante :

une valeur réelle supérieure à la valeur théorique. Cette situation est décrite par le graphique :



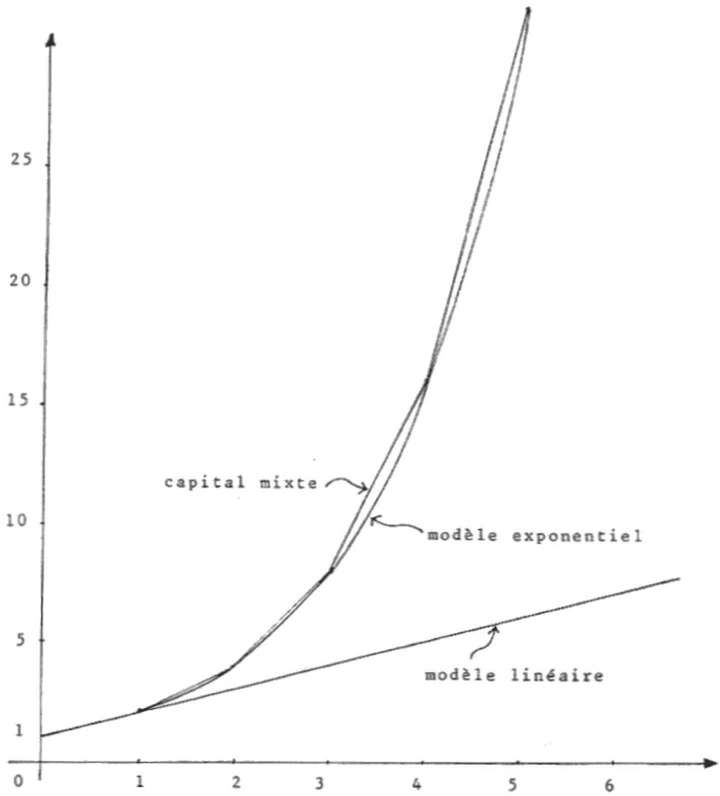
Envisageons à présent des durées supérieures à l'unité. Le graphique suivant décrit :

1. la modélisation linéaire (4.1)

2. la modélisation exponentielle (4.4)

3. la capitalisation mixte (4.5)

Par souci de simplification, on ne représente pas les fonctions en escaliers réelles mais leur lissage théorique. En effet, chaque unité doit être décomposée en 24 sous-périodes. Le dernier saut avant l'unité doit être plus important que les autres pour traduire la prime de fidélité. On a : ($i = 1; 100\%$ pour la visualisation).



4.3 Les formules de changement de taux

4.3.1 L'intérêt simple

Considérons à nouveau la relation (4.1) : $C(t) = C(0).(1 + it)$. Lorsque l'unité de temps est multipliée par k, la valeur de t est divisée par k : en effet 12 mois correspondent à une année. Pour établir un résultat cohérent, il suffit de considérer un taux lui aussi k fois plus grand. La notion ainsi introduite porte le nom de « taux proportionnel ».

4.3.2 L'intérêt composé

Ceci n'est plus le cas si on utilise la relation (4.4) :

$$C(t) = C(0).(1+i)^t$$

Notons i le taux pour l'unité choisie et i_k le taux correspondant à la nouvelle unité choisie, supposée k fois plus grande.

Dans le nouveau système, (4.4) devient :

$$C(\frac{t}{k}) = C(0).(1+i_k)^{(\frac{t}{k})}$$

Pour conserver la cohérence au système, on doit avoir : $C(t) = C(\frac{t}{k})$.

Pour être tout à fait rigoureux, des notations « primes » devraient être introduites qui nuiraient à la clarté du propos.

On obtient :

$$(1+i)^t = (1+i_k)^{(\frac{t}{k})}$$

On en tire :

$$\boxed{\begin{array}{c} i = (1+i_k)^{(\frac{1}{k})} - 1 \\ i_k = (1+i)^k - 1 \end{array}} \tag{4.6}$$

4.3.3 Lien entre l'intérêt simple et l'intérêt composé

On montre de manière évidente que l'intérêt simple est la réduction au premier degré du développement en série de MAC-LAURIN de l'intérêt composé. L'extension (4.3) de la modélisation linéaire satisfait également cette propriété. (Attention au rayon de convergence !). Les relations (4.6) de changement de taux se réduisent également à la relation de proportionnalité.

L'intérêt simple peut donc se concevoir comme la première approximation de l'intérêt composé. En tant que tel, on peut justifier son utilisation dans un certain voisinage de l'origine des temps. Pratiquement, on limite à un an la durée de ce voisinage. Des différences significatives sont néanmoins observées entre intérêt simple et intérêt composé pour des durées inférieures à l'année.

4.3.4 Comparaison entre les différents taux

Pour comparer les relations (4.6) et les taux proportionnels, il suffit d'étudier la fonction : (voir [49, p. 20])

$$F(i) = i_k - k.i = (1+i)^k - 1 - k.i$$

On dérive facilement cette fonction :

$$F'(i) = k.(1+i)^{k-1} - k = k.[(1+1)^{k-1} - 1]$$

Dès lors :

- Lorsque $0 < k < 1$, on a toujours : $F'(i) < 0$, ce qui implique $F(i)$ décroissante. Or $F(0) = 0$! (il suffit de remplacer i par 0 dans la première équation). Lorsque $i > 0$, on a donc obligatoirement : $F(i) < 0$, ou encore $\boxed{i_k < k.i.}$ Cette relation est facile à interpréter. En effet, le taux i_k représente le taux équivalent au taux i, le taux donnant mêmes valeurs (équivalent). Le taux $k.i$ est le taux réellement utilisé. La relation qui vient d'être démontrée signifie tout simplement que lorsque les remboursements s'effectuent plusieurs fois par an, le taux réellement utilisé est toujours supérieur au taux équivalent au taux annoncé.
- Lorsque $k > 1$, on montre exactement de la même manière que $i_k > k.i$. Cette relation peut s'interpréter comme suit : lorsque la période interversement est supérieure à l'année, il faut tenir compte de la (ou des) capitalisation(s) intermédiaire(s) qui modifie(nt) le taux.

4.4 L'exponentielle de base e et la capitalisation continue

4.4.1 Nécessité de l'extension du modèle exponentiel

La modification de l'unité de temps induit pour les relations (4.4) des modifications de taux non triviales (4.6). Néanmoins, par souci de simplification, la plupart des organismes financiers utilisent des relations de proportionnalité ...lorsque celles-ci sont avantageuses.

Pour contourner cette petite subtilité, on peut définir une capitalisation à des taux proportionnels pour des périodes infiniment courtes. On introduit ainsi la capitalisation continue. L'intérêt intrinsèque de cette notion est l'apparition « naturelle » du nombre e. En effet : la relation (4.4) était : $C(t) = C(0).(1+i)^t$. En décomposant l'intervalle $[0,t]$ en kt intervalles de longueur $(\frac{1}{k})$ pour lesquels on applique des taux proportionnels $\frac{i}{k}$, on obtient : $C'(t) = C(0).(1+\frac{i}{k})^{k.t}$.

Ici la notation « prime » est indispensable, la nouvelle valeur différant de $C(t)$. Il faudrait également noter $C'(kt)$ et non $C'(t)$, l'unité de temps ayant été modifiée. De telles subtilités ont parfois le tort de brouiller le propos véritable d'une interprétation.

En faisant tendre k vers l'infini, on obtient finalement la relation :

$$\boxed{C^*(t) = C(0).e^{it}}. \tag{4.7}$$

(démonstration : voir [36, p. 115])

La relation (4.7) a toutes les qualités de l'intérêt composé. De plus, les modifications d'unité de temps induisent des modifications de taux suivant la simple proportionnalité. Ce n'est pas tout, l'utilisation de la capitalisation continue permet de placer la théorie de l'intérêt dans un cadre plus large et la fait bénéficier de façon directe de résultats intéressants.

4.4.2 Les nouveaux horizons de la théorie de l'intérêt

La relation (4.7) peut se définir immédiatement par une équation différentielle. En effet si $C(t) = C(0).e^{it}$, on a : $C'(t) = C(0).i.e^{it} = i.C(t)$. ([2]).

Sous forme différentielle, on écrit :

$$\boxed{dC(t) = C(t).i.dt.} \tag{4.8}$$

L'énorme avantage de la relation (4.8) est qu'elle permet d'introduire des variations de taux d'intérêt. Rien ne nous empêche d'écrire :

$$dC(t) = C(t).i(t).dt$$

([2]) La notation « prime » désigne ici la dérivée première

Dans un souci de description réaliste de la réalité économique, on peut également faire du taux une variable aléatoire. On décompose alors le terme $i(t)dt$ en :

$$\boxed{i(t)dt = d(t)dt + \sigma(t)dw(t)} \tag{4.9}$$

1. $d(t)dt$ représente la partie déterministe décrivant la tendance évolutive du taux.
2. $\sigma(t)dw(t)$ est en fait la perturbation aléatoire introduite sous la forme d'un processus wienérien de coefficient de diffusion σ.

Un tel modèle a été développé récemment par P. Devolder dans [22].

4.4.3 Lien entre capitalisation continue et intérêt composé

Etant donné l'utilisation de taux proportionnels pour la capitalisation continue, on n'a pas : $C(t) = C^*(t)$.

On peut se demander pour quel taux j on obtiendrait des résultats « intérêt composé » équivalents aux résultats taux i « capitalisation continue ». On obtient : $C(t,j) = C^*(t,i)$, $C(0).(1+j)^t = C(0)e^{it}$. On en tire : $j = e^i - 1$. En utilisant la relation qui donne e^x en série de Mac-Laurin :

$$j = (1 + i + \frac{i^2}{2} + \frac{i^3}{6} \ + \cdots) - 1 = i + \frac{i^2}{2} + \frac{i^3}{6} + \cdots$$

La relation inverse est également intéressante : on a bien sûr :

$$i = \ln(1 + j).$$

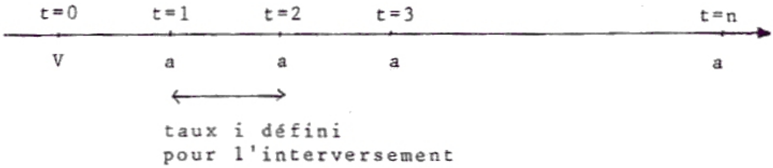
En passant par Mac-Laurin : $i = j - (\frac{j^2}{2} + (\frac{j^3}{3}) - \cdots$

La sous-estimation du taux en travaillant par capitalisation continue est donc de l'ordre de $\frac{j^2}{2}$. Elle est donc d'autant plus grande que le taux est élevé. Lorsque l'on considère non plus le taux d'intérêt d'un contrat mais par exemple le taux de rendement d'un investissement, celui-ci est une estimation, non une constante. Un taux de rendement élevé correspond généralement à un investissement de risque plus grand. L'utilisation de la capitalisation continue est alors une atténuation du rendement d'un investissement, atténuation qui est fonction du rendement lui-même et donc de son risque.

4.5 Successions de capitaux équidistants

4.5.1 Description du problème

Considérons un montant V remboursé par n versements de valeur a échéant à intervalles réguliers. Graphiquement le problème est symbolisé par :



Le principe de base de la mathématique financière est l'égalité des capitaux perçus et déboursés, en tenant compte de la modélisation choisie pour décrire leur évolution.

4.5.2 Modélisation linéaire

L'équation est :

$$V = a.(1-i) + a.(1-2i) + \ldots + a.(1-n.i)$$

On reconnaît immédiatement une progression arithmétique et l'on obtient : (voir [36, p. 78])

$$V = \frac{a.n.(2-i.(n+1))}{2}.$$

Cette relation est du premier degré en V, a, i et du deuxième degré en n. Tous les paramètres sont donc calculables. Nous savons qu'une équation non équivalente peut être proposée en $t = n$. Nous avons déjà abordé les problèmes que cette mise en équation pouvait introduire.

4.5.3 Modélisation exponentielle

On obtient immédiatemment :

$$V = a.(1+i)^{-1} + a.(1+i)^{-2} + \ldots + a.(1+i)^{-n}$$

La progression est géométrique cette fois et l'on trouve : (voir [36, p. 126])

$$\boxed{V = \tfrac{a}{i}.(1-(1+i)^{-n})} \tag{4.10}$$

La relation est du premier degré en a et V. Pour le calcul de n, l'utilisation des logarithmes s'avère nécessaire. On trouve :

$$\begin{aligned} n &= -\log_{(1+i)}(1 - V.\frac{i}{a}) \\ &= \frac{-\log(1 - V.\frac{i}{a})}{\log(1+i)} \end{aligned}$$

Le calcul de i est un peu plus amusant. On peut écrire (4.10) sous la forme :

$$i = \frac{a}{V}.(1-(1+i)^{-n})$$

Notons cette expression : $i = f(i)$.

A partir d'une valeur arbitraire i_0 (i_0 strictement positif), on peut construire la suite des valeurs :

$$\begin{aligned} i_1 &= f(i_0) \\ i_2 &= f(i_1) \\ &\vdots \\ i_k &= f(i_{k-1}). \end{aligned}$$

On peut montrer que cette suite converge toujours vers la solution du problème.

En effet :

$$|i_k - i_{k-1}| = |f(i_{k-1}) - f(i_{k-2})|.$$

En utilisant le théorème de la moyenne, on sait que l'on peut trouver un point intermédiaire entre i_{k-1} et i_{k-2} tel que l'on puisse écrire :

$$f(i_{k-1}) - f(i_{k-2}) = f'(i^*).(i_{k-1} - i_{k-2}).$$

On peut majorer facilement l'expression $f'(i^*)$ sur l'intervalle $]0, +\infty[$ [36, p. 141].

En notant β le majorant, ($\beta < 1$), on obtient :

$$|i_k - i_{k-1}| < \beta.|i_{k-1} - i_{k-2}|.$$

Les différences entre approximations successives sont de plus en plus petites, ce qui assure la convergence.

4.5.4 La résolution de nos cas particuliers

Emprunt remboursable en deux versements

Considérons à nouveau notre emprunt de 20 000 remboursable en 2 versements de 10 500, à l'issue des 3ème et 6ème mois. Plusieurs attitudes sont possibles. On peut travailler avec un taux trimestriel : i_t. Dans ce cas, la relation est :

$$20\,000 = \frac{10\,500}{i_t}.(1 - (1 + i_t)^{-2})$$

Nous utilisons évidemment les formules (4.10). La relation est du troisième degré. La suite construite comme expliqué précédemment converge vers la valeur $i_t = 0,033135133$. Le taux annuel équivalent est : $i = 0,13935428$. On peut travailler avec un taux mensuel : i_m. L'équation est alors :

$$20\,000 = 10\,500(1 + i_m)^{-3} + 10\,500(1 + i_m)^{-6}$$

Cette équation est du deuxième degré. En posant $Y = (1+i_m)^{-3}$, on obtient tout simplement :

$$20\,000 = 10\,500Y + 10\,500Y^2 \text{ ou } 10\,500Y^2 + 10\,500Y - 20\,000 = 0$$

On trouve :

$$Y = \frac{-10\,500 \pm \sqrt{(10\,500)^2 + 4 \times 10\,500 \times 20\,000}}{21\,000}$$

Soit : $Y_1 = 0,9679107276$, Y_2 est négatif et ne peut nous intéresser. On calcule i_m de la façon qui suit :

$$\begin{aligned}(1+i_m)^{-3} &= Y_1 \\ i_m &= Y_1^{-\frac{1}{3}} - 1 = 0,01093111934\end{aligned}$$

Quant au taux annuel équivalent : $i = (1+i_m)^{12} - 1 = 0,13935428$, il est (fort heureusement) égal au taux trouvé précédemment.

Un ordinateur vendu à tempérament

La mise en équation dérive immédiatemment des relations (4.10). On trouve :

$$22\,500 = \frac{2165}{i_m}(1 - (1+i_m)^{-24})$$

La relation qui en découle converge vers la valeur : $i_m = 0,08156696477$. Cette valeur est très nettement différente de la valeur trouvée par intérêt simple. En calculant le taux annuel équivalent, on trouve : $i_{annuel} = 1,562365$, soit plus de 150% ! Ce résultat sans être infini est néanmoins représentatif du risque que l'on court en acceptant n'importe quel crédit sans aucune vérification et sans faire jouer la concurrence.

L'application aux taux de chargement

Le problème est tout simplement de calculer le taux d'un emprunt de 100 000 remboursé par 20 versements mensuels de 5 400. L'équation est donc :

$$100\,000 = \frac{5400}{i_m}(1 - (1+i_m)^{-20})$$

La relation qui en découle converge vers la valeur : $i_m = 0,0074442749$. Le taux annuel équivalent est :

$$i_{annuel} = (1+i_m)^{12} - 1 = 0,09308113468$$

Il n'est pas rare d'entendre qu'il suffit de multiplier le taux de chargement par 12 pour obtenir le taux. On voit qu'il n'en est rien ! La relation légale donne dans ce cas particulier :

$$i = \frac{24 \times 0,004 \times 20}{21} = 0,09142857$$

L'approximation n'est valable qu'au pourcent près.

En guise de conclusion

La méthode que nous venons d'évoquer permet enfin la résolution des problèmes financiers réels, problèmes jusqu'ici difficiles à aborder numériquement par l'amateur intéressé. La micro-informatique à petits prix lève définitivement cet obstacle et permet la résolution immédiate de problèmes complexes. On peut par exemple travailler en univers aléatoire et simuler les événements futurs et atteindre ainsi une représentation beaucoup plus réaliste de l'univers économique [37, p.255]. Cette extension et bien d'autres a permis l'élévation de la traditionnelle algèbre financière au titre envié de « mathématique ».

Chapitre 5

Aspects théoriques et pratiques de la recherche de points fixes.

par *Roland HINNION*

5.1 A propos du théorème de Brouwer ...

Rappelons que :
- un *point fixe* pour une fonction f est un élément x (du domaine de f) tel que : $f(x) = x$.
- $\mathbb{R}$ étant l'ensemble des nombres réels, un *compact* de $\mathbb{R}^n$ est exactement un ensemble borné et fermé (fermé pour la topologie habituelle de $\mathbb{R}^n$).

Une formulation classique du théorème de Brouwer est alors :

Toute fonction continue d'un compact convexe K de $\mathbb{R}^n$ dans K admet au moins un point fixe.

Ce théorème reste bien entendu valable lorsqu'on remplace la condition « K compact convexe » par « K homéomorphe à un ensemble compact convexe » ; il s'ensuit que l'on peut (sans affaiblir le théorème), remplacer dans l'énoncé ci-dessus « K compact convexe » par (au choix) :

- K boule fermée
  ou
- K cube fermé
  ou
- K rectangle fermé
  etc

(avec l'interprétation habituelle de la notion de *rectangle* (par exemple) :

K du type $\{(x_1, x_2, \ldots, x_n) \,|\, \forall i \;\; a_i \leqslant x_i \leqslant b_i\}$, avec a_i et b_i donnés (dans $\mathbb{R}$)).

Petit détail : on suppose bien entendu K non vide !

L'utilité du théorème de Brouwer saute aux yeux si l'on remarque que $G(X) = 0$ est équivalent à $F(X) = X$, pour $F(X) \equiv G(X) + X$; ici G et F sont des fonctions du type $\mathbb{R}^n \to \mathbb{R}^n$ et X est une variable vectorielle. On voit donc que la résolution d'un système (non linéaire en général !) du type $G(X) = 0$ est équivalente à la recherche d'un point fixe pour F.

La résolution de systèmes est essentielle dans tous les domaines concernés par le quantitatif ; il est bien connu (par exemple) qu'un type classique de problème économique est la minimisation d'une fonction *coût* : $C(x, y)$ (par exemple à deux variables) ; on sait que cela mène souvent à la recherche des solutions du système exprimant que les dérivées partielles de C sont nulles, donc à un système non linéaire de deux équations à deux inconnues.

5.2 Brouwer contre Brouwer ...

Brouwer énonça son théorème dans un article : Über Abbildung von Mannigfaltigkeiten, *Mathematische Annalen* 71, p.97-115. L'article se termine par l'indication : *Amsterdam, juli 1910 ...*

En 1951, Brouwer fournit un contre-exemple montrant que le théorème est faux ! Ajoutons cependant que Brouwer se place désormais dans le système intuitionniste dont il est le fondateur, et que le théorème du point fixe est faux dans ce système seulement ...Nous ne définirons pas l'intuitionnisme ici ; il suffira de savoir que l'intuitionnisme se base sur une logique différente de la logique classique, et que, en particulier, il refuse les démonstrations par l'absurde.

Il faut savoir que depuis la preuve initiale de BROUWER, beaucoup d'autres preuves ont été données, par des scientifiques d'horizons divers (analystes, économistes, géomètres,etc) ; certaines preuves ne nécessitent même plus qu'une connaissance assez élémentaire de l'analyse classique (voir par exemple KANNAÏ) ; malheureusement ce type de preuve (toujours par l'absurde !) ne permet pas de comprendre réellement ce qui est en jeu. D'autres preuves par contre font clairement apparaître le schéma suivant :

1. Des arguments combinatoires donnent une version finie du théorème du point fixe (en fait du point presque fixe, mais nous y reviendrons).
2. Des arguments purement techniques permettent de passer au théorème du point presque fixe.
3. Le théorème de BOLZANO-WEIERSTRASS permet de déduire du théorème précité le (vrai) théorème du point fixe.

Dans l'intuitionnisme, on peut reproduire les étapes 1 et 2, mais pas 3 car le théorème de BOLZANO-WEIERSTRASS est faux en intuitionnisme. Dans son article de 1951, BROUWER démontre d'ailleurs (en intuitionnisme) le théorème du point presque fixe.

De ceci on peut déduire que :

– Il ne peut exister aucun algorithme général permettant de déterminer un (vrai) point fixe ; en effet, si un tel algorithme existait, il fournirait une preuve intuitionniste du théorème du point fixe.
– Toutes les preuves (en logique classique) des théorèmes de BOLZANO-WEIERSTRASS et BROUWER (version pré-intuitionniste !) doivent contenir une étape où l'on utilise le principe de démonstration par l'absurde.

Remarquons que ce qui précède n'empêche pas que dans ces cas particuliers (où les hypothèses ne sont plus celles du théorème de BROUWER, c'est-à-dire f continue de K dans K, avec K compact convexe de $\mathbb{R}^n$) il puisse exister des preuves constructives (c'est-à-dire non par l'absurde, et fournissant un algorithme). L'exemple suivant (bien connu) illustre cette possibilité : soit f une contraction $K \to K$ (K borné, $K \subseteq \mathbb{R}^n$), c'est-à-dire une fonction continue telle qu'il existe un α tel que $0 < \alpha < 1$ et

$$\forall P, Q \in K \qquad d(f(P), f(Q)) < \alpha d(P, Q)$$

où d est la distance standard dans $\mathbb{R}^n$.

Si P_j est la suite définie par

$$P_0 \text{ quelconque et } P_{k+1} = f(P_k),$$

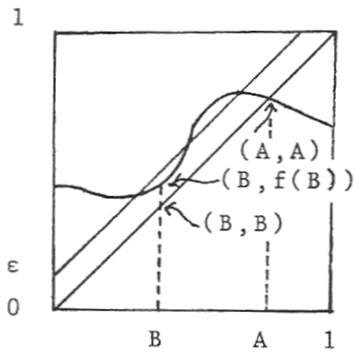
on montre facilement que $d(P_{j+k}, P_j) < \alpha^j d(P_k, P_0)$; comme K est borné, il existe $M \in \mathbb{R}$ tel que $d(P_k, P_0) < M$; la suite P_j est donc manifestement de CAUCHY, donc convergente. Sa limite P ne peut être qu'un point fixe, car en passant à la limite (sur $j \to \infty$) dans $f(P_j) = P_{j+1}$ on obtient $f(P) = P$.

5.3 Points presque fixes ...

Si ϵ est un réel strictement positif, on dit, par définition, que P est un point ϵ-fixe de f ssi $d(f(P), P) < \epsilon$.

La figure ci-dessous montre un point ϵ-fixe P et un vrai point fixe P' (dans le cas d'une fonction $f : [0,1] \to [0,1]$).

Remarquez que le théorème de BROUWER est trivial dans ce contexte (dimension 1) et qu'un point presque fixe (même pour ϵ très petit) n'est pas nécessairement proche d'un vrai point fixe.



A est fixe
B est ϵ-fixe

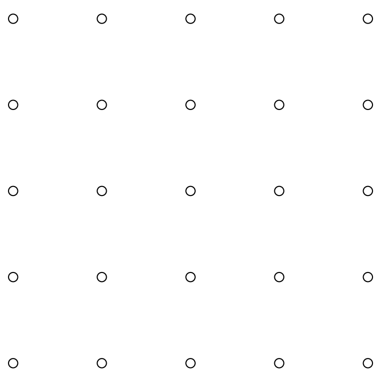
Pour le théoricien, les notions de *point fixe* et de *point presque fixe* sont donc bien distinctes ! D'un point de vue pratique, par contre, la recherche d'un point ϵ-fixe est, dans bien des cas, aussi intéressante que celle d'un vrai point fixe (pour ϵ assez petit) ; pour l'économiste par exemple, une production de 0,003 litres de lait équivaut à une production nulle !

L'intérêt des algorithmes permettant de déterminer des points ϵ-fixes est dès lors évident. Un de ces algorithmes est contenu dans la très élégante preuve géométrique suivante (du théorème de BROUWER) :

5.4 Une preuve géométrique du théorème de Brouwer . . .

Cette preuve est due à N. NIZETTE, et se déroule en trois étapes (cf le paragraphe 5.2) :

– Etape 1 : une version finie du théorème du point presque fixe.
  Pour pouvoir énoncer cette version, il nous faut introduire quelques notions et définitions ; nous nous limiterons à la dimension 2, pour des questions de représentabilité, mais la méthode se généralise aisément en dimension n (n quelconque).
  Un *réseau* est un ensemble du type



  De façon précise :

$$R = \left\{ (x, y) \mid \exists i, j \in \{1, 2, \ldots, k\} \text{ tels que } x = \frac{i}{k} \text{ et } y = \frac{j}{k} \right\}$$

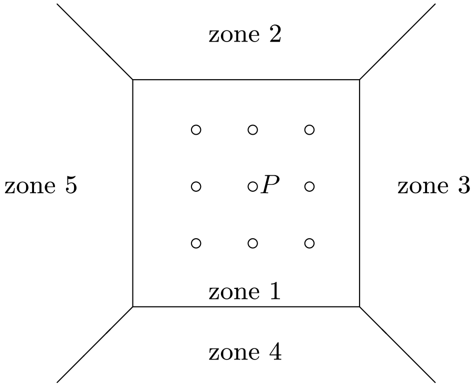
  k est la *taille* du réseau ; le réseau de la figure ci-dessus est de taille 5.
  Sur tout réseau on peut définir une notion naturelle de voisinage : dans la figure suivante, nous indiquons les points voisins de P par le signe « $\star$ » :

$$\begin{array}{ccccc}
\circ & \circ & \circ & \circ & \circ \\
\circ & \star & \star & \star & \circ \\
\circ & \star & \star P & \star & \circ \\
\circ & \star & \star & \star & \circ \\
\circ & \circ & \circ & \circ & \circ
\end{array}$$

Si P et P' sont voisins, on écrit $P \mathbb{V} P'$. Remarquez que $\mathbb{V}$ est une relation réflexive et symétrique (mais pas transitive !).
Une fonction f d'un réseau R dans R sera dite *continue* ssi

$$\forall P, P' \in R \qquad (P \mathbb{V} P' \Longrightarrow f(P) \mathbb{V} f(P'));$$

en clair : f est continue si elle respecte la relation de voisinage.
Un point P presque fixe est tel que $f(P) \mathbb{V} P$.
La version finie du point presque fixe peut s'énoncer maintenant :
Toute fonction continue d'un réseau fini dans lui-même admet un point presque fixe.
Preuve :
Soit f continue : $R \to R$, où R est un réseau fini. Munissons chaque point de R de l'une des cinq *flèches* suivantes :
– type 1 : flèche nulle, indiquée par $\odot$
– type 2 : flèche vers le haut :$\uparrow$
– type 3 : flèche vers la droite : $\rightarrow$
– type 4 : flèche vers le bas : $\downarrow$
– type 5 : flèche vers la gauche : $\leftarrow$
L'attribution des flèches indique que $f(P)$ est voisin de P (auquel cas P est un point presque fixe), ou que $f(P)$ « tombe trop loin de P », vers le haut (flèche de type 2), ou vers la droite (flèche de type 3), etc.
De façon précise :
P sera muni d'une flèche de type i si $f(P)$ se trouve dans la zone i :



Si $f(P)$ tombe sur la frontière entre deux zones, on choisit librement l'une des deux attributions possibles.
On se trouve maintenant dans la situation suivante :

- chaque point de R est muni d'une flèche
- comme f est une fonction $R \to R$, il n'y a pas de flèches « sortantes »
- comme f est continue, deux points voisins ne peuvent jamais porter de flèches « opposées » (par exemple de type 2 et de type 4)

Ces circonstances suffisent à garantir l'existence d'un point portant une flèche nulle (type 1). Pour le voir clairement, on ajoute au réseau R un bord artificiel muni de flèches « rentrantes » ; cela ne change pas les caractéristiques décrites ci-dessus, et n'introduit pas de nouveaux points presque fixes ; les points presque fixes du nouveau réseau fléché R' seront les mêmes que ceux du réseau R.

Exemple : R de taille 3 :

$$\begin{array}{ccc}
\circ & \circ & \circ \\
\downarrow & \downarrow & \downarrow \\
\circ \rightarrow & \circ \rightarrow & \circ \\
\uparrow & & \\
\circ & \circ \rightarrow & \odot
\end{array}$$

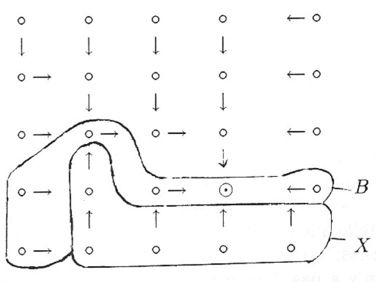
On passe à R' de taille 5 :

$$
\begin{array}{ccccccccc}
\circ & & \circ & & \circ & & \circ & & \circ \\
 & & & & & & & & \\
\circ & & \circ & & \circ & & \circ & & \circ \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
\circ & & \circ & \rightarrow & \circ & \rightarrow & \circ & & \circ \\
 & & \uparrow & & & & & & \\
\circ & & \circ & & \circ & \rightarrow & \odot & & \circ \\
 & & & & & & & & \\
\circ & & \circ & & \circ & & \circ & & \circ
\end{array}
$$

et l'on munit le bord de flèches « rentrantes » (il faut remarquer qu'à chaque coin on a le choix entre deux types de flèches : ici nous les disposerons dans le sens trigonométrique) :

$$
\begin{array}{ccccccccc}
\circ & & \circ & & \circ & & \circ & \leftarrow & \circ \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
\circ & \rightarrow & \circ & & \circ & & \circ & \leftarrow & \circ \\
 & & & & & & & & \\
\circ & \rightarrow & \circ & & \circ & & \circ & \leftarrow & \circ \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
\circ & \rightarrow & \circ & \rightarrow & \circ & \rightarrow & \circ & \leftarrow & \circ \\
 & & \uparrow & & & & \downarrow & & \\
\circ & \rightarrow & \circ & & \circ & \rightarrow & \odot & \leftarrow & \circ \\
 & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
\circ & \rightarrow & \circ & & \circ & & \circ & & \circ
\end{array}
$$

Considérons maintenant $X =$ la partie connexe maximale de R', rencontrant le bord inférieur de R', et dont tous les points sont munis de flèches ascendantes (type 2) :



Rappelons qu'un sous-ensemble X d'un réseau est dit connexe si deux points quelconques de X sont toujours reliés par un chemin situé dans X, un chemin étant une suite finie de points $P_0, P_1, \ldots, P_k$ telle que $\forall i \qquad P_{i-1} \mathbb{V} P_i$.

Le bord de X, dans R', est un certain chemin B. Deux points voisins n'étant jamais munis de flèches opposées, et X ne présentant que des flèches de type 2, il est clair que les points de B sont munis de flèches de type 1,3 ou 5 (jamais 2, par le caractère maximal de X). Comme une extrémité de B porte une flèche de type 3, et l'autre une flèche de type 5, il y aura nécessairement une flèche nulle dans B (vu la définition d'un chemin), donc un point presque fixe.

– Etape 2 : passage au théorème du point presque fixe (pour K compact convexe de $\mathbb{R}^n$).

Soit f une fonction continue : $K \to K$, où K est (par exemple) un carré de $\mathbb{R}^2$. Soit ϵ un réel > 0. On sait (puisque f est continue sur K compact) que f est uniformément continue sur K, et donc :

$$\exists \eta > 0 \text{ tel que } \forall P, P' \in K \ (d(P,P') < \eta \Rightarrow d(f(P), f(P')) < \frac{\epsilon}{\sqrt{2}}).$$

On subdivise alors le carré K en un réseau de taille $j+1$, de sorte que $\frac{1}{j} <$ minimum de $\{\frac{\epsilon}{2}, \frac{\eta}{2}\}$. Un petit calcul élémentaire permet de vérifier que dans ce cas, la fonction g construite ci-dessous, avec $g : R \to R$ (où R est le réseau décrit ci-dessus), est bien une fonction continue (au sens des réseaux) ; g admet donc un point presque fixe (au sens des réseaux) ; il est facile de voir que ce point presque fixe est un point ϵ-fixe pour f. Voici l'approximation g : $g(P) =$ un point (au choix) du réseau R, voisin de $f(P)$ (au sens de la relation $\mathbb{V}$ sur R).

Ce qui précède fournit clairement un algorithme pour la recherche d'un point ϵ-fixe.

– Etape 3 : Cette étape est tout à fait standard ; elle consiste à déduire le théorème du point fixe du théorème des points presque fixes. Soit donc f continue : $K \to K$, avec K compact convexe de $\mathbb{R}^n$. Puisque pour tout $\epsilon > 0$, il existe un point ϵ-fixe, considérons une suite P_i (dans K), telle que : $d(P_i, f(P_i)) < \frac{1}{i}$. Par le théorème de Bolzano-Weierstrass, cette suite admet une sous-suite convergente (pour laquelle il n'y a pas d'algorithme ! Ceci est donc la seule étape non-constructive de toute la preuve). Soit $(P_{i_k})_{k \in \mathbb{N}}$ une telle sous-suite (de limite P). En passant à la limite dans la relation : $d(P_{i_k}, f(P_{i_k})) < \frac{1}{i_k}$ on obtient : $d(P, f(P)) \leqslant 0$, d'où $f(P) = P$.

5.5 Un autre algorithme pour la recherche de points presque fixes

Les économistes ont mis au point d'autres algorithmes pour la recherche de points ϵ-fixes. Comme dans (5.4), ces algorithmes sont une partie d'une preuve (aussi constructive que possible) du théorème de Brouwer.

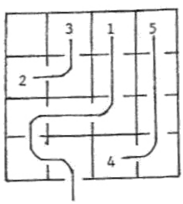
A titre d'exemple, nous décrivons ici l'algorithme de Scarf ; d'autres algorithmes, plus raffinés et plus performants, ont été mis au point par d'autres chercheurs.

L'argument de type fini (étape 1) est ici le « principe des deux portes ».

Considérons une maison dont toutes les pièces (en nombre fini !) se situent au rez-de-chaussée et comportent 0, 1 ou 2 portes. De plus, il y a exactement une porte donnant sur l'extérieur. Dans ce cas, le nombre de pièces à 1 porte est impair et, en partant de l'unique porte donnant sur l'extérieur, on aboutit nécessairement à une pièce à une porte. Ceci est évident si l'on remarque que :

- au départ de la porte d'entrée, il n'y a qu'un seul trajet possible, et il ne peut se terminer que dans une pièce à 1 porte.
- au départ d'une pièce à une porte qui n'est pas celle dont on vient de parler, on ne peut qu'aboutir dans une autre pièce à une porte ; excepté dans la pièce à une porte où l'on aboutit en venant de l'extérieur, les autres pièces à une porte sont couplées naturellement par l'unique chemin qui les relie ; il y a donc un nombre impair de pièces à une porte.

Exemple :



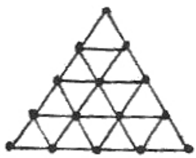
Décrivons maintenant l'algorithme de Scarf, en dimension 2. Soit S le triangle dont les sommets sont $(1,0,0),(0,1,0)$ et $(0,0,1)$, dans $\mathbb{R}^3$; l'ensemble S est donc la fermeture convexe de $\{(1,0,0),(0,1,0),(0,0,1)\}$, c'est-à-dire : $S = \{\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 | \forall i\ 0 \leqslant \alpha_i \leqslant 1 \text{ et } \alpha_1 + \alpha_2 + \alpha_3 = 1\}$ avec $e_1 = (1, 0, 0)$, etc.

Tout point de S se caractérise donc par les nombres α_i, appelés les coordonnées barycentriques de ce point. Nous écrirons : $P\langle 1\rangle$ pour α_i. Exemple : si P est $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ (NB : $P \in S$) $P\langle 1\rangle = \frac{1}{2}$ et $P\langle 2\rangle = P\langle 3\rangle = \frac{1}{4}$.

S sera le compact convexe dans lequel on recherche un point presque fixe (pour f continue : $S \to S$).

La triangulation régulière de S, de taille k, est le sous-ensemble (de S) des points dont les coordonnées barycentriques sont toutes du type $\frac{i}{k-1}$ (avec $0 \leqslant i \leqslant k-1, i$ entier).

Exemple : Triangulation régulière de taille 5 : T_5 est l'ensemble des sommets dans le dessin suivant :



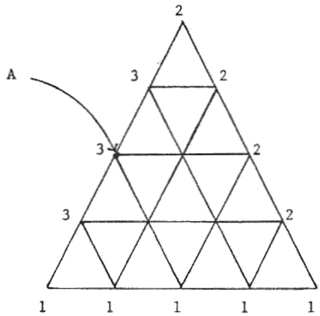
Soit f une fonction continue : $S \to S$. Soit T_k la triangulation de taille k.

L'indexation I de Scarf attribue un entier à chaque point de T_k. Voici comment :

Pour P un point de T_k : $I(P) =$
- le plus petit j tel que $P\langle j\rangle = 0$ s'il existe un tel j.
- le plus petit i tel que $(f(P))\langle i\rangle \geqslant P\langle i\rangle$ si aucune coordonnée barycentrique de P n'est nulle.

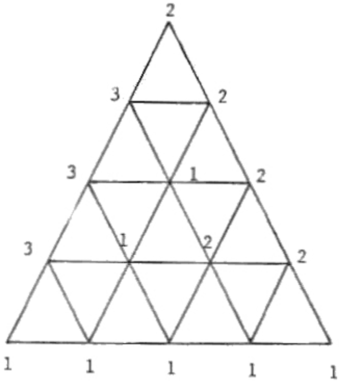
Pour les sommets e_i de S, l'indexation de Scarf donne (quelle que soit f) : $I(e_1) = 2; I(e_2) = 1; I(e_3) = 1$.

Voici ce qui se passe au bord de S pour une triangulation de taille 5 :



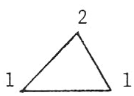
Le point A (par exemple) est $\frac{1}{2}e_1 + \frac{1}{2}e_2$, de coordonnées barycentriques : $\frac{1}{2}, \frac{1}{2}, 0$; donc $I(A) = 3$.

L'indexation des points de T_5 situés à l'intérieur de S dépend de f, puisque pour ces points aucune coordonnée barycentrique n'est nulle. Cela pourrait donner par exemple :

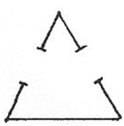


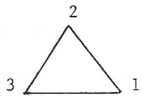
S peut maintenant être assimilé à une maison dont les pièces sont les triangles déterminés par la triangulation et les portes, les côtés (de ces triangles) dont les extrémités sont des sommets indexés par 1 et 2.

Ainsi :

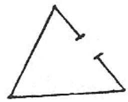


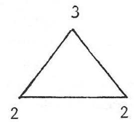
est une pièce à 2 portes :



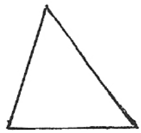


est une pièce à une porte :

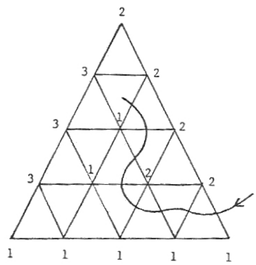




est une pièce sans porte :



L'indexation de Scarf garantit qu'il n'y a qu'une seule porte sur le bord de S. Les hypothèses du « principe des deux portes » sont satisfaites et il suffit de pénétrer par l'unique porte d'entrée et de suivre l'unique trajet possible pour découvrir une pièce à une porte, c'est-à-dire un triangle « complet » (qui porte les indices 1, 2 et 3) :



L'intérêt d'un tel triangle complet réside dans le fait que si A, B, C sont ses sommets (avec par exemple : $I(A) = 1, I(B) = 2$ et $I(C) = 3$), on a :

$$\begin{aligned} (f(A))\langle 1\rangle &\geqslant A\langle 1\rangle \\ (f(B))\langle 2\rangle &\geqslant B\langle 2\rangle \\ (f(C))\langle 3\rangle &\geqslant C\langle 3\rangle \end{aligned} \qquad \text{(Inégalités *)}$$

Ceci est dû au fait que pour tout point P de la triangulation régulière, $(f(P))\langle I(P)\rangle \geqslant P\langle I(P)\rangle$, par la définition même de I.

Si l'on avait $A = B = C$, on aurait $f(A) = A$, en vertu des relations d'inégalités ci-dessus et du fait que A et $f(A)$ se trouvent dans S (puisque $f : S \to S$), ce qui entraîne que : $A\langle 1\rangle + A\langle 2\rangle + A\langle 3\rangle = 1$ et $(f(A))\langle 1\rangle + (f(A))\langle 2\rangle + (f(A))\langle 3\rangle = 1$, avec $A\langle i\rangle$ et $(f(A))\langle i\rangle$ tous positifs ou nuls. Ceci entraîne que lorsque A, B, C (sommets d'un triangle complet) sont assez proches l'un de l'autre, chacun est un point presque fixe de f. Précisons ceci : considérons la suite $(T_k)_{k\in\mathbb{N}}$ des triangulations régulières de S.

Pour chaque T_k, on considère l'indexation de SCARF et on obtient un triangle complet correspondant, de sommets A_k, B_k, C_k. Comme S est compact, il existe des sous-suites $(A_{k_j})_{j\in\mathbb{N}}, (B_{k_j})_{j\in\mathbb{N}}, (C_{k_j})_{j\in\mathbb{N}}$ de limites respectives A, B, C (dans S). Les triangles de T_k deviennent plus petits lorsque k augmente, et à la limite, on aura : $A = B = C$. Les inégalités* (cf ci-dessus) s'appliquent pour chaque triple A_k, B_k, C_k, et sont conservées par passage à la limite. On se retrouve donc dans le cas examiné ci dessus (lorsque $A = B = C$) et A est un point fixe de f. La suite $(A_k)_{k\in\mathbb{N}}$ admet donc un vrai point fixe comme point d'accumulation. On en déduit facilement que : $\forall\epsilon > 0\, \exists k \in \mathbb{N}$ tel que $d(f(A_k), A_k) < \epsilon$.

On dispose donc d'un algorithme permettant de trouver un point ϵ-fixe ; cet algorithme consiste à déterminer un triangle complet dans T_k, et à tester (pour le sommet A par exemple) si $d(A, f(A)) < \epsilon$; si le test n'est pas concluant, on recommence avec (par exemple) T_{2k-1} (qui a l'avantage de contenir T_k, et nécessite donc moins de calculs pour la détermination de I).

5.6 Un exemple économique

Dans l'exemple classique suivant, le théorème de BROUWER sert à prouver l'existence d'un équilibre dans une économie d'échange. Une économie d'échange se caractérise par :

1. n biens de consommation
2. N consommateurs, définis par :
   - un espace de consommation $X^i \subseteq (\mathbb{R}^+)^n$, qui est l'ensemble des vecteurs de consommation possibles $x^i = (x_1^i, x_2^i, \ldots, x_n^i)$ pour le consommateur i ; $\mathbb{R}^+$ est l'ensemble des nombres réels $\geqslant 0$; x_j^i est la quantité de bien de type j consommé par le consommateur i ; i varie de 1 à N ; j varie de 1 à n.
   - une fonction d'utilité $U^i : X^i \to \mathbb{R}$, qui décrit les préférences du consommateur i ; x et y étant deux vecteurs (dans X^i), $U^i(x) > U^i(y)$ signifie que le consommateur i préfère x à y, et $U^i(x) = U^i(y)$ que i n'a pas de préférence entre x et y.

   – un vecteur de biens initiaux $W^i = (w_1^i, w_2^i, \ldots, wi_n)$ (NB : $W = W^1 + W^2 + \ldots + W^N$ est appelé le vecteur total de biens initiaux).

3. Des vecteurs de prix $p = (p_1, p_2, \ldots, p_n)$ appartenant à $(\mathbb{R}^+)^n$ (p_j est le prix d'une unité de bien de type j).

Pour chaque consommateur i, on définit une fonction « budget » B^i par : $B^i(p) = \{x \in X^i | p.x \leqslant p.W^i\}$; p varie bien entendu dans l'espace des prix et $p.x$ représente le produit scalaire usuel entre les vecteurs p et x, donc : $p.x = \sum_i p_i x_i$. Par la suite tous les produits scalaires seront simplement indiqués par « . ». Le souhait de chaque consommateur est de maximiser sa fonction d'utilité, et il aura donc pour tendance de choisir dans $B^i(p)$ le vecteur de consommation z^i tel que :

$$U^i(z^i) = max\{U^i(x) | x \in B^i(p)\}$$

(pour un vecteur de prix p fixé).

Un *équilibre* dans ce modèle est alors constitué par un vecteur de prix $p^\star$ et des vecteurs z^i tels que :

$(i) z^i \in B^i(p^\star)$ et $U^i(z^i) = max\{U^i(x) | x \in B^i(p^\star)\}$

$(ii) \sum_{i=1}^N z^i \leqslant W$

Notons $z^i(p)$ le vecteur qui réalise le maximum de U^i dans $B^i(p)$ (NB : nous supposons qu'un tel vecteur existe ; nous verrons plus loin des hypothèses raisonnables pour qu'il en soit bien ainsi).

Dans ces conditions, en définit la « demande excédentaire du marché » par :

$$E(p) = \left(\sum_{i=1}^N z^i(p) \right) - W.$$

A l'équilibre, on doit donc avoir : $E(p^\star) \leqslant 0$, c'est-à-dire que l'on doit pouvoir satisfaire toutes les demandes du marché.

Pour pouvoir assurer l'existence d'un équilibre, on fait les hypothèses suivantes :

1. Hypothèse sur les fonctions d'utilité :
   On suppose que toutes les fonctions U^i sont continues et strictement concaves (sur l'ensemble X^i, supposé bien entendu convexe). Cette

hypothèse garantit que chaque $B^i(p)$ non vide admet exactement un élément qui maximise U^i (existence et unicité de $z^i(p)$ pour chaque i tel que $B^i(p)$ est non vide).

De plus $z^i(p)$ et $E(p)$ seront des fonctions continues de p.

2. Hypothèse d'homogénéité de $z(p)$ et $E(p)$:

   pour tout prix p et tout nombre $a > 0, z(a.p) = z(p)$ et $E(a.p) = E(p)$ (homogénéité de degré 0) avec $z(p) = \sum_{i=1}^{N} z^i(p)$.

3. La troisième hypothèse est la (bien connue) « loi de WALRAS » : pour tout prix p : $p.E(p) = 0$. Cette hypothèse garantit que, pour un équilibre ($p^\star \geqslant 0$ et $E(p^\star) \leqslant 0$), on a nécessairement égalité entre l'offre et la demande pour les biens non gratuits : en effet, si $p^\star.E(p^\star) = 0$, les $p^\star_j.E_j(p^\star)$ doivent tous être nuls et donc $E_j(p^\star)$ est nul lorsque $p^\star_j$ ne l'est pas !

Existence d'un équilibre

L'hypothèse (2) permet d'effectuer une normalisation des prix : on peut se restreindre aux prix p tels que : $p \geqslant 0$ et $\sum_{j=1}^{n} p_j = 1$. La variable p varie donc maintenant dans un ensemble compact et convexe (du type considéré dans l'algorithme de SCARF). Considérons maintenant la fonction f suivante :

$$f(p) = \frac{(p + E(p))^+}{\sum_{j=1}^{n}(p_j + E_j(p))^+}$$

avec la convention : $b^+ = max\{b, 0\}$. Il suffit maintenant de montrer que f satisfait les hypothèses du théorème de BROUWER et que son point fixe correspond à un équilibre :

1. f est bien une fonction $K \to K$, avec $K = \{p \in \mathbb{R}^n | p \geqslant 0 \text{ et } \sum_{j=1}^{n} p_j = 1\}$ un ensemble compact convexe de $\mathbb{R}^n$; la vérification est facile ; le seul point délicat est la vérification du fait que f est bien définie : supposons que pour un p de K, on ait $\sum_{j=1}^{n}(p_j + E_j(p))^+ = 0$; alors pour chaque j : $(p_j + E_j(p))^+ = 0$ et donc : $p_j + E_j(p) \leqslant 0$; pour les j tels que $p_j > 0$, on a donc $E_j(p) < 0$ et donc $p_j.E_j(p) < 0$; en sommant, on obtient : $0 > \sum_{j=1}^{n} p_j.E_j(p) = p.E(p)$, ce qui contredit la loi de WALRAS. f est donc bien définie.
2. f est manifestement continue, puisque $E(p)$ l'est.

3. Un point fixe de f correspond bien à un équilibre : en effet, soit p un prix tel que $f(p) = p$.

Soit u le vecteur de $\mathbb{R}^n$ dont toutes les composantes valent 1 . On a donc : $p = f(p) = \frac{(p+E(p))^+}{(p+E(p))^+.u}$; en posant k ce dénominateur ($k > 0$), on a donc : $k.p = (p + E(p))^+$.

Considérons les indices j tels que $p_j > 0$: comme $k.p_j = p_j + E_j(p)$, on a : $k.p_j^2 = p_j^2 + p_j.E_j(p)$. En sommant, on obtient : $\sum_j k.p_j^2 = \sum_j p_j^2 + \sum_j p_j.E_j(p)$ (la somme étant prise indifféremment sur les j tels que $p_j > 0$ ou sur les j de 1 à n). Donc : $\sum_j k.p_j^2 = \sum_j p_j^2 + p.E(p)$; comme $p.E(p) = 0$ (loi de Walras) ceci entraîne : $k = 1$. Donc : $p = (p + E(p))^+$. Si p_j est $> 0, E_j(p)$ doit être nul ; si p_j est nul, on a : $0 = (p_j + E_j(p))^+ = (E_j(p))^+$ et donc $E_j(p)$ est $\leqslant 0$.

p est donc bien un équilibre.

Bibliographie

[1] Allais M., Puissance et dangers de l'utilisation de l'outil mathématique en économique, *Econometrica*, 22, 1954.

[2] Allen R.G.D., *Mathematical Analysis for Economists*, Mac Millan, London, 1947.

[3] Als M.G., *Histoire de l'économie mathématique*, Presses Universitaires de Bruxelles, 1961.

[4] Arrow K. J., Enthoven A.C., Quasi-concave Programming, *Econometrica*, 1961, 29, 779-800.

[5] Ayres F., *Mathematics of finance*, New York, Mc Graw-Hill, Schaum's outiline series in business, 19..

[6] Bair J., *Applications et compléments théoriques d'algèbre linéaire*, Presses Universitaires de Liège, 1987.

[7] Bair J., *Applications et compléments théoriques d'analyse mathématique*, Presses Universitaires de Liège, 1987-1988.

[8] Bair J., La notion d'élasticité explique l'influence des variations de prix sur le revenu brut de la vente d'un bien, *Mathématique et Pédagogie*, 1987, 64, 21-24.

[9] Bair J., Un problème de dualité en programmation quasi-concave, *Optimization*, 1988, 19, pp...

[10] Bair J., *Mathématiques générales : exercices et solutions*, De Boeck-Wesmael, Bruxelles, 1988.

[11] Bair J., *La programmation mathématique au service de la gestion*, en voie de publication.

[12] Balestra P.*Calcul matriciel pour économistes*, Castella, Albeuve, 1972.

[13] BAUMOL W.J., *Economic Dynamics*, Macmillan, New York, 1959.

[14] BOOT J.C., *Mathematical Reasoning in Economics and Management Science*, Prentice-Hall, Englewood Cliffs (New Jersey), 1967.

[15] BRADLEY I., MEEK R., *Matrices and Society*, Penguin, Harmondsworth, 1986.

[16] CHIANG A.C., *Fundamental Methods of Mathematical Economics*, 3e Edit., Mc Graw-Hill, Tokyo, 1984.

[17] CNAPELINCKX G., NYSTEN L., MEURICE P., *Algèbre financière*, De Boeck, Bruxelles, 1986, 21e éd.

[18] COHEN V., *Mathématiques pour les sciences sociales, tome 3 : processus discrets*, Presses Universitaires de France, Paris, 1971.

[19] COULSON A., *An Introduction to Matrices*, Longman, London, 1965.

[20] DEBREU G., *Théorie de la valeur : analyse axiomatique de l'équilibre économique*, Dunod, Paris, 1966.

[21] DESPLAS M., *Algèbre linéaire 2 et applications économiques,* Dalloz, 197..

[22] DEVOLDER, P, Opérations stochastiques de capitalisation, *Astin Bulletin* Vol. 16, 1986.

[23] FERRET, LANGLOIS, *Mathématiques appliquées* Foucher, Paris, 1979.

[24] FLETCHER T.J., *L'algèbre linéaire par ses applications*, CEDIC, Paris, 1972.

[25] FREDON D., *Mathématique, économie et gestion*, Cédic, Paris, 1976.

[26] GINSBURGH, WAELBROECK, *Fixed-Point Algorithms for Computation of Economic Equilibria*, Librairie Droz, Economie Appliquée Archives de l'ISMEA, Tome 33, n° 4, Genève, 1979.

[27] GOLDBERG S., *Introduction to Difference Equations*, Wiley, New York, 1958.

[28] HADLEY G., KEMP M.C., *Finite Mathematics in Business and Economics*, North-Holland, Amsterdam, 1972.

[29] HENDERSON J.M., QUANDT R.E., *Microéconomie : formulation mathématique élémentaire*, Dunod, Paris, 1982.

[30] HENRY S.G.B., *Elementary Mathematical Economics*, Sage Publications, Beverly Hills (California), 1969.

[31] HINNION R., Points (presque) fixes, Cahiers du Centre de Logique, vol 6, Logique et Informatique, Cabay, Louvain-la-Neuve, 1986.

[32] Houx, *Théorème de Brouwer : théorie et applications*, Mémoire ULB, Faculté des Sciences, Bruxelles, 1982-1983.

[33] Intriligator M.D.*Mathematical Optimization and Economic Theory*, Prentice-Hall, New Jersey, 1971.

[34] Jongmans F., Varlet J., *Notions de mathématique à l'usage des sciences humaines*, Université de Liège, 1968.

[35] Jongmans F., *Les mathématiques au dix-neuvième siècle*, APPS éditions, Bruxelles, 1987.

[36] Justens D., *Introduction à la mathématique financière*, De Boeck, Bruxelles, 1987.

[37] Justens D., *Statistique pour décideurs*, De Boeck, Bruxelles, 1988.

[38] Kemeny J., Snell J., Thompson G., *Algèbre moderne et activités humaines*, Dunod, Paris, 1960.

[39] Kemeny J., The Social Sciences Call on Mathematics, in *The Mathematical Sciences : a Collection of Essays*, The M.J.T. Press, 1969.

[40] Klein L., The Role of Mathematics in Economics, in *The Mathematical Sciences : a Collection of Essays*, The M.J.T. Press, 1969.

[41] Langaskens Y., Gazon J., *L'élaboration de modèles statistiques et dynamiques en économie*, Droz, Genève, 1978.

[42] Maillet P., *L'économétrie*, Presses Universitaires de France, Paris, 1976.

[43] Malinvaud E., *Leçons de théorie microéconomique*, Dunod, Paris, 1969.

[44] Minguet A., *Introduction à l'analyse microéconomique*, tome A, Presses Universitaires de Liège, 1983.

[45] Moulin H., Fogelman F., *La convexité dans les mathématiques de la décision*, Hermann, Paris, 1979.

[46] Nizette, *A propos du théorème de Brouwer*, Mélanges Paul Libois, Service de Géométrie, Université Libre de Bruxelles, 1964.

[47] Peston M.H., *Elementary Matrices for Economics*, Sage Publications, Beverly Hills, 1969.

[48] Quaden G., *Politique économique*, Labor, Bruxelles, 1985.

[49] Saada M., *Mathématiques financières*, Presses Universitaires de France, Coll. Que sais-je ? 19..

[50] SAADA M., *Traité pratique de mathématique financière, 1.Procédures fondamentales ; 2.Emprunts indivis ; 3.Investissements en avenir certain*, Vuibert, Paris, 3 volumes.

[51] SEARLE S.R., HAUSMAN W.H., *Matrix Algebra for Business and Economics*, Wiley, New York, 1970.

[52] VERMAAT A.J., Over de wiskunde in de economie, *De Economist*, 118, 1970.

[53] WILLIAMS J.D., *The Compleat Strategyst being a primer on the theory of games of strategy*, McGraw-Hill, New York, 1954.

Index

Index des Noms