

L3 Économie-Gestion 2013–2014

Mathématiques 5

Julia Matos

Université d'Evry Val-d'Essonne

Table des matières

1	Optimisation sans contraintes	2
1.1	Rappels de topologie de \mathbb{R}^n	2
1.2	Différentiabilité d'une fonction à plusieurs variables	3
1.3	Maximum et minimum d'une fonction à plusieurs variables	6
1.4	Cas des fonctions convexes ou concaves	11
2	Optimisation avec contraintes	13
2.1	Contraintes d'égalité	13
2.2	Contraintes d'inégalité	17
3	Équations et systèmes de récurrence	21
3.1	Équations de récurrence linéaires d'ordre p à coefficients réels constants	21
3.2	Systèmes de récurrence linéaires à coefficients réels constants et d'ordre 1	25
3.3	Équilibre et stabilité pour une équation de récurrence	29
3.4	Équilibre et stabilité pour un système de récurrence d'ordre 1	30
4	Équations différentielles et systèmes différentiels linéaires	33
4.1	Équations différentielles du premier ordre	33
4.1.1	Équations différentielles linéaires du premier ordre	34
4.1.2	Équations se ramenant à des équations linéaires	36
4.2	Équations différentielles linéaires du second ordre	37
4.2.1	Coefficients constants	37
4.3	Équations différentielles linéaires d'ordre p	39
4.3.1	Coefficients constants	40
4.4	Systèmes différentiels linéaires du premier ordre	41
4.4.1	Cas des coefficients constants	42
4.4.2	Équations différentielles d'ordre p	44
4.5	Équilibre et stabilité	45

1 Optimisation sans contraintes

Le problème mathématique dit d'optimisation joue un rôle majeur dans l'analyse économique pour les nombreuses applications qu'il trouve dans la théorie économique. Formellement un tel problème consiste à maximiser ou à minimiser une fonction donnée f de la variable $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ en sachant que x appartient à un sous-ensemble D de \mathbb{R}^n .

Dans ce chapitre, nous étudions les problèmes d'optimisation les plus simples, c'est-à-dire les problèmes d'optimisation libres ou sans contraintes. Dans ce cas, l'ensemble D est un ouvert de \mathbb{R}^n ou bien $D = \mathbb{R}^n$.

1.1 Rappels de topologie de \mathbb{R}^n

Nous énonçons ce qu'il faut savoir sur \mathbb{R}^n .

Soient $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ deux vecteurs de \mathbb{R}^n . On appelle *produit scalaire* de x et y le nombre

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x_1 y_1 + \dots + x_n y_n.$$

Soit $x \in \mathbb{R}^n$. On appelle la *norme* de x , le nombre réel positif ou nul,

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \dots + x_n^2}, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Cette norme $\|\cdot\|$ s'appelle la norme euclidienne. Dans le cas $n = 2$, c'est la norme correspondant au théorème de Pythagore. Quand $n = 1$, on a $\|x\| = |x|$.

Il y a d'autres normes sur \mathbb{R}^n , dites normes usuelles. Pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on pose

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Les applications $\|\cdot\|_1$ et $\|\cdot\|_\infty$ ainsi définies sont aussi des normes sur \mathbb{R}^n et s'appellent respectivement la norme ℓ^1 et la norme du max ou norme infinie.

Étant donné $a \in \mathbb{R}^n$ et $r > 0$, la boule ouverte de centre a et rayon r est définie par :

$$B(a, r) = \{a \in \mathbb{R}^n : \|x - a\| < r\}.$$

Si $n = 1$, $B(a, r)$ est l'intervalle $]a - r, a + r[$. Si $n = 2$, $B(a, r)$ est le disque de centre a et rayon r privé du cercle. Si $n = 3$, alors $B(a, r)$ est la sphère de centre a et rayon r privée de la surface.

La boule fermée de centre a et rayon $r \geq 0$ est définie par :

$$\bar{B}(a, r) = \{x \in \mathbb{R}^n : \|x - a\| \leq r\}.$$

On appelle *boule unité* de \mathbb{R}^n la boule de centre 0 et rayon 1 : $B(0, 1)$ et *sphère unité* à l'ensemble $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$.

Une partie $A \subset \mathbb{R}^n$ est dite *bornée* s'il existe une boule qui la contient (c'est-à-dire, il existe $a \in \mathbb{R}^n$ et $r > 0$ tels que $A \subset B(a, r)$).

Une partie O de \mathbb{R}^n est dite *ouverte* si, pour tout $x \in O$, il existe $r > 0$ tel que $B(x, r) \subset O$. En particulier, $B(a, r)$ est un ensemble ouvert de \mathbb{R}^n et \mathbb{R}^n est un ouvert. Dans \mathbb{R} , tout intervalle du type $]a, b[$ est un ouvert.

Un ensemble F de \mathbb{R}^n est dit un *fermé* si son complémentaire dans \mathbb{R}^n est un ouvert. En particulier, $\bar{B}(a, r)$ est un fermé dans \mathbb{R}^n . Dans \mathbb{R} , tout intervalle du type $[a, b]$ est un fermé.

1.2 Différentiabilité d'une fonction à plusieurs variables

Rappelons qu'une fonction f d'une variable réelle x est *dérivable* en $a \in \mathbb{R}$, s'il existe $f'(a) \in \mathbb{R}$ tel que

$$\begin{aligned}\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} &= \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h} = f'(a) \\ \iff \lim_{h \rightarrow 0} \frac{f(a + h) - f(a) - f'(a)h}{h} &= 0.\end{aligned}$$

La dérivée $f'(a)$ de f en a mesure l'effet de la variation infinitésimal $\Delta x = x - a$ de a sur $f(a)$:

$$f(x) - f(a) = f(a + \Delta x) - f(a) \simeq f'(a)\Delta x,$$

où l'approximation précédente devient une identité dans le cas d'une fonction affine $f(x) = mx + q$, m étant la dérivée de f en tout point x et, mesure la variation exacte de la fonction f pour chaque unité d'augmentation de x .

Exemple. On considère la *fonction de demande* $x = F(p)$ qui exprime la quantité x en fonction du prix p . Alors, la *fonction de demande marginale* $F'(p)$ décrit l'effet d'une augmentation d'une unité du prix sur le comportement d'achat des consommateurs. Il s'agit d'une mesure de sensibilité dépendante des unités choisies pour mesurer les quantités et les prix. Pour que la mesure de sensibilité soit complètement indépendante des unités choisies, il faut considérer la fonction *élasticité-prix* donnée par

$$E(p) = \frac{F'(p)p}{F(p)}.$$

La demande d'un bien qui est plutôt insensible aux variations de prix aura une élasticité proche de zéro.

Si la fonction f d'une variable réelle x est dérivable en tout point d'un intervalle ouvert I de \mathbb{R} , on définit sur I la fonction *dérivée première* $f'(x)$. Si cette fonction est elle-même dérivable en un point $a \in I$, on appelle *dérivée seconde* de f en a la dérivée $(f')'(a)$ de f' en a et on note $f''(a)$, et ainsi de suite.

Soit $f : D \rightarrow \mathbb{R}$ une fonction de n variables réelles où D est un ouvert de \mathbb{R}^n et $x = (x_1, \dots, x_n) \in D$. On dit que f est dérivable en $a = (a_1, \dots, a_n) \in D$ par rapport à la variable x_i , lorsque la limite

$$\lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a)}{h}$$

existe et est finie. Dans ce cas, on note $\frac{\partial f}{\partial x_i}(a)$ sa valeur et on l'appelle la dérivée partielle de f par rapport à la variable x_i en a .

L'interprétation précédente de la dérivée partielle première d'une fonction d'une variable reste valable pour les dérivées partielles premières d'une fonction de plusieurs variables, c'est-à-dire $\frac{\partial f}{\partial x_i}(a)$ mesure l'effet de la variation infinitésimal h de a dans la i -ème variable.

Exemple. On considère la *fonction de production* $Q = F(K, L)$ qui exprime la quantité de production Q en fonction du capital K et du travail L . Alors la fonction *productivité*

marginale du capital $\frac{\partial F}{\partial K}(K, L)$ décrit l'effet d'une augmentation d'une unité du capital K sur la quantité produite Q lorsque la quantité de travail L utilisée reste fixe.

Propriétés des dérivées partielles : Soient f et g deux fonctions définies sur D ouvert de \mathbb{R}^n admettant la dérivée partielle par rapport à la variable x_i en $a \in D$. Alors, on a

1. $\frac{\partial(f+g)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a)$.
2. $\frac{\partial(fg)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a)g(a) + f(a)\frac{\partial g}{\partial x_i}(a)$.
3. Si $g(a) \neq 0$,

$$\frac{\partial}{\partial x_i} \left(\frac{f}{g} \right) (a) = \frac{\frac{\partial f}{\partial x_i}(a)g(a) - f(a)\frac{\partial g}{\partial x_i}(a)}{g^2(a)}.$$

Si la fonction $f : D \rightarrow \mathbb{R}$ est dérivable en a par rapport à toutes les variables x_i , $1 \leq i \leq n$, alors on appelle *gradient* de f en a et on note $\nabla f(a)$ le vecteur

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right).$$

Exemple. Soit la fonction $f(x, y) = e^y \cos x$, $(x, y) \in \mathbb{R}^2$. Alors,

$$\nabla f(x, y) = (-e^y \sin x, e^y \cos x) \quad \text{pour tout } (x, y) \in \mathbb{R}^2.$$

Soit $f(x) = f(x_1, \dots, x_n)$ une fonction admettant toutes les dérivées partielles d'ordre 1 en $a \in D$. On dit que f est *différentiable* en a si

$$\lim_{h \rightarrow \vec{0}} \frac{f(a+h) - f(a) - \langle \nabla f(a), h \rangle}{\|h\|} = 0$$

Alors, on obtient la *Formule de Taylor d'ordre 1* de f en a :

$$\begin{aligned} f(a+h) &= f(a) + \langle \nabla f(a), h \rangle + \varepsilon(h)\|h\| \\ &= f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a)h_i + \varepsilon(h)\|h\|, \end{aligned}$$

où $h \in \mathbb{R}^n$ et $\lim_{h \rightarrow \vec{0}} \varepsilon(h) = 0$.

Enfin, on définit les dérivées partielles d'ordre supérieure d'une fonction à n variables f comme pour les fonctions d'une variable réelle. La *dérivée partielle seconde* de f par rapport à x_j et x_i (dans cet ordre) en $a \in U$ est la dérivée partielle de la fonction $\frac{\partial f}{\partial x_i}$ par rapport à x_j en a , et on note

$$\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Il est important de savoir qu'en général

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \neq \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

On utilise également la notation suivante

$$\frac{\partial^2 f}{\partial x_i^2}(a) = \frac{\partial^2 f}{\partial x_i \partial x_i}(a),$$

pour tout $1 \leq i \leq n$.

Si $f : D \rightarrow \mathbb{R}$, D ouvert de \mathbb{R}^n , admet toutes les dérivées partielles du second ordre en a , alors on appelle *matrice hessienne* de f en a la matrice carrée d'ordre n donnée par :

$$D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1,\dots,n} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{pmatrix}.$$

En particulier, pour une fonction de deux variables réelles $f(x, y)$, la matrice hessienne en un point (x_0, y_0) est

$$D^2 f(x_0, y_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{pmatrix}.$$

On dit que f est k fois continument différentiable en $a \in D$ ou de classe C^k en a si f est continue en a et toutes les dérivées partielles jusqu'à l'ordre k de f en a existent et sont continues en a . On dit que f est de classe C^k sur D si f est classe C^k en tout point $a \in D$.

Une fonction de classe C^1 sur D est différentiable sur D mais la réciproque est fausse.

Théorème de Schwarz.

Si f est de classe C^2 en a , alors

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a),$$

pour tout $i, j = 1, \dots, n$.

Par conséquent, sous les conditions du théorème de Scharwz, la matrice hessienne de f en a est symétrique ($D^2 f(a) = {}^t D^2 f(a)$).

Un autre résultat fondamental qui sera important dans la section suivante.

Formule de Taylor-Young d'ordre 2.

Si $f : D \rightarrow \mathbb{R}$ est de classe C^2 sur D et $a \in D$. Alors, pour h suffisamment petit,

$$\begin{aligned} f(a+h) &= f(a) + \langle \nabla f(a), h \rangle + \frac{1}{2} {}^t h D^2 f(a) h + \varepsilon(h) \|h\|^2 \\ &= f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i + \sum_{i=1}^n \sum_{j=1}^n h_i \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_j + \varepsilon(h) \|h\|^2, \end{aligned}$$

avec $\lim_{h \rightarrow \vec{0}} \varepsilon(h) = 0$.

1.3 Maximum et minimum d'une fonction à plusieurs variables

Soit D un ensemble de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$. On dit que a est un *maximum local* de f , s'il existe $r > 0$ tel que

$$\forall x \in B(a, r) \cap D, \quad f(x) \leq f(a).$$

On dit que a est un *minimum local* de f , s'il existe $r > 0$ tel que

$$\forall x \in B(a, r) \cap D, \quad f(x) \geq f(a).$$

On dit que a est un maximum ou minimum strict si les inégalités précédentes sont strictes pour $x \neq a$. Un point a est un *extremum local* de f si il est un maximum local ou un minimum local de f

De plus, on dit que a est un *maximum global* de f si

$$\forall x \in D, \quad f(x) \leq f(a).$$

Et on dit que a est un *minimum global* de f , si

$$\forall x \in D, \quad f(x) \geq f(a).$$

Pour déterminer les maxima et minima d'une fonction f sur un ensemble D , les propriétés topologiques de D sont déterminantes. On a le résultat fondamental suivant qui devient faux si une des deux propriétés topologiques de D n'est pas vérifiée.

Théorème 1.1 *Si D est un fermé borné de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$ est continue sur D , alors f atteint son maximum global et son minimum global sur D , c'est-à-dire, il existe $a, b \in D$ tels que $f(a) \leq f(x) \leq f(b)$, pour tout $x \in D$.*

Dans la pratique, pour déterminer les maxima et minima d'une fonction f à plusieurs variables, on utilise les théorèmes suivants qui généralisent les résultats connus dans le cas d'une fonction à une variable et qui deviennent faux si D n'est pas ouvert dans \mathbb{R}^n .

Théorème 1.2 (Condition nécessaire du premier ordre) *Soit D un ouvert de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$ différentiable sur D . Si $a \in D$ est un extremum de f , alors $\nabla f(a) = \vec{0}$.*

Preuve : Supposons que $a \in D$ est un extremum de f . Soit $h \in \mathbb{R}^n$ fixé. On considère $\varphi : \mathbb{R} \rightarrow \mathbb{R}^n$ définie par $\varphi(t) = a + th$. On a $\varphi(0) = a$ et $\varphi'(t) = h$. La fonction $g = f \circ \varphi$ à valeurs réelles est définie et différentiable sur l'ouvert de $\mathbb{R} : U = \{t \in \mathbb{R} : \varphi(t) \in D\}$. Alors, le point $t = 0$ est un extremum de g et donc $g'(0) = \langle \nabla f(a), h \rangle = 0$. Comme h est arbitraire dans \mathbb{R}^n , on a $\nabla f(a) = \vec{0}$.

La réciproque de ce théorème est évidemment fausse. Un point peut être critique sans être un extremum. Exemple : si $f : \mathbb{R} \rightarrow \mathbb{R}$ est définie par $f(t) = t^3$, alors $f'(0) = 0$ mais 0 n'est pas un extremum local.

Dans ce qui suit, on suppose $f : D \rightarrow \mathbb{R}$ au moins de classe C^2 .

On définit $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ par

$$Q(h) = {}^t h D^2 f(a) h = \sum_{i=1}^n \sum_{j=1}^n h_i \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_j,$$

où $h = (h_1, \dots, h_n) \in \mathbb{R}^n$. On appelle Q la *forme quadratique* associée à la matrice $D^2f(a)$. On suppose $\nabla f(a) = \vec{0}$. Alors, pour h suffisamment petit, la formule de Taylor-Young d'ordre 2 s'écrit :

$$\begin{aligned} f(a+h) - f(a) &= \langle \nabla f(a), h \rangle + \frac{1}{2} {}^t h D^2 f(a) h + \varepsilon(h) \|h\|^2 \\ &= \frac{1}{2} {}^t h D^2 f(a) h + \varepsilon(h) \|h\|^2. \end{aligned}$$

Théorème 1.3 (Condition nécessaire du second ordre) Soit $f : D \rightarrow \mathbb{R}$ de classe C^2 en $a \in D$.

1. Si a est un minimum local de f , alors $\nabla f(a) = \vec{0}$ et

$$Q(h) \geq 0, \quad \forall h \in \mathbb{R}^n,$$

c'est-à-dire Q est forme quadratique semi-définie positive.

2. Si a est un maximum local de f , alors $\nabla f(a) = \vec{0}$ et

$$Q(h) \leq 0, \quad \forall h \in \mathbb{R}^n,$$

c'est-à-dire Q est forme quadratique semi-définie négative.

On appelle *point critique* ou *point stationnaire* de f tout point $a \in D$ tel que $\nabla f(a) = \vec{0}$. Les éventuels maxima et minima de f dans D doivent être recherchés dans l'ensemble de points critiques de f . Si un point critique de f vérifie une des conditions nécessaires du théorème précédent, on ne peut rien conclure sur la nature de ce point. Pour pouvoir déterminer si un point critique est un maximum ou minimum de f , il faut vérifier des conditions plus strictes que les précédentes, portant toujours sur les dérivées partielles secondes de f .

On donne des conditions suffisantes pour qu'un point critique a soit un extremum de f .

Théorème 1.4 (Conditions suffisantes) Soit D un ouvert de \mathbb{R}^n , $f : D \rightarrow \mathbb{R}$ de classe C^2 et $a \in D$ un point critique de f .

1. Si Q est définie positive, c'est-à-dire

$$\forall h \in \mathbb{R}^n \setminus \{\vec{0}\}, \quad Q(h) > 0, \tag{1.1}$$

alors a est un minimum local strict de f .

2. Si Q est définie négative, c'est-à-dire

$$\forall h \in \mathbb{R}^n \setminus \{\vec{0}\}, \quad Q(h) < 0, \tag{1.2}$$

alors a est un maximum local strict de f .

Preuve :

1. Il existe $\alpha > 0$ tel que $Q(h) \geq \alpha \|h\|^2$, pour tout $h \in \mathbb{R}^n$. Par la formule de Taylor-Young, pour $h = x - a$, on a

$$f(x) - f(a) = \frac{1}{2} Q(x-a) + \varepsilon(x-a) \|x-a\|^2 \geq \left(\frac{\alpha}{2} + \varepsilon(x-a) \right) \|x-a\|^2.$$

Ainsi, pour x suffisamment proche de a et $x \neq a$, le signe de α prédomine et l'on a $f(x) > f(a)$.

2. On raisonne de façon analogue. Il existe $\beta > 0$ tel que $Q(h) \leq -\beta\|h\|^2$, pour tout $h \in \mathbb{R}^n$. Alors, pour tout x suffisamment proche de a et $x \neq a$, on a $f(x) < f(a)$.

Si la forme quadratique $Q(h) = {}^t h D^2 f(a) h$ est indéfinie en un point critique a , c'est-à-dire qu'ils existent $h, k \in \mathbb{R}^n$ tels que $Q(h) > 0$ et $Q(k) < 0$, alors on dit que a est un *point selle* ou *point col* de f . Ce terme provient du fait que la fonction f au point a est minimisée dans certaines directions et maximisée dans d'autres directions, son graphe a la forme d'une selle ou d'un col d'une montagne.

Le signe de la forme quadratique $Q(h)$ dépend du signe des mineurs principaux de la matrice $D^2 f(a)$. Si A est une matrice carrée d'ordre n , la *sous-matrice principale d'ordre k* de A est la matrice extraite de A en éliminant les $n - k$ dernières lignes et les $n - k$ dernières colonnes de A . On appelle alors *mineur principal d'ordre k* de A le déterminant de la sous-matrice principale d'ordre k de A et on le note $M_k(A)$. Le Théorème 1.4 peut être alors énoncé de façon équivalente comme suit.

Théorème 1.5 (Conditions suffisantes) *Soit D un ouvert de \mathbb{R}^n , $f : D \rightarrow \mathbb{R}$ de classe C^2 et $a \in D$ un point critique de f .*

1. *Si les n mineurs principaux de $D^2 f(a)$ sont tous positifs, c'est-à-dire*

$$\frac{\partial^2 f}{\partial x_1^2}(a) > 0, \left| \begin{array}{cc} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) \end{array} \right| > 0, \left| \begin{array}{ccc} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_3}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_3}(a) \\ \frac{\partial^2 f}{\partial x_3 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_3 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_3^2}(a) \end{array} \right| > 0, \dots,$$

alors a est un minimum local strict de f .

2. *Si les n mineurs principaux de $D^2 f(a)$ alternent de signe, le premier étant négatif, c'est-à-dire*

$$\frac{\partial^2 f}{\partial x_1^2}(a) < 0, \left| \begin{array}{cc} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) \end{array} \right| > 0, \left| \begin{array}{ccc} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_3}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_3}(a) \\ \frac{\partial^2 f}{\partial x_3 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_3 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_3^2}(a) \end{array} \right| < 0, \dots,$$

alors a est un maximum local strict de f .

Remarque. Soit $a \in D$ un point critique de f . Si $\frac{\partial^2 f}{\partial x_1^2}(a) > 0$ et l'un des autres mineurs de $D^2 f(a)$ est négatif, alors a n'est pas un extremum de f .

Considérons le cas de la dimension 2. Soit $f : D \rightarrow \mathbb{R}$ une fonction C^2 sur un ouvert D de \mathbb{R}^2 et $a = (x_0, y_0) \in D$ un point critique de f , c'est-à-dire

$$\frac{\partial f}{\partial x}(a) = \frac{\partial f}{\partial y}(a) = 0.$$

La forme quadratique $Q(h)$ est définie par

$$Q(h) = \frac{\partial^2 f}{\partial x^2}(a)h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(a)h_1 h_2 + \frac{\partial^2 f}{\partial y^2}(a)h_2^2, \quad h = (h_1, h_2).$$

On note

$$r = \frac{\partial^2 f}{\partial x^2}(a), \quad s = \frac{\partial^2 f}{\partial x \partial y}(a), \quad t = \frac{\partial^2 f}{\partial y^2}(a)$$

(ce que l'on appelle les *notations de Monge*). Il est alors possible d'énoncer le Théorème 1.5 de la façon suivante.

Théorème 1.6 (Conditions suffisantes en dimension 2) *Soit D un ouvert de \mathbb{R}^2 , $f : D \rightarrow \mathbb{R}$ de classe C^2 et $a \in D$ un point critique de f .*

1. *Si $rt - s^2 > 0$ et $r > 0$, alors a un minimum local strict.*
2. *Si $rt - s^2 > 0$ et $r < 0$, alors a un maximum local strict.*
3. *Si $rt - s^2 < 0$, alors a n'est pas extremum de f . On dit que a est un point selle ou col.*
4. *Si $rt - s^2 = 0$, le comportement de f au voisinage de a dépend des termes suivants de son développement de Taylor.*

Finalement, sous les hypothèses des théorèmes précédents, $D^2f(a)$ est une matrice réelle symétrique et donc elle est diagonalisable. Plus précisément, il existe D matrice diagonale (des valeurs propres de $D^2f(a)$) et P matrice orthogonale ($P^{-1} = P^t$) telles que $A = P^{-1}DP$. Le résultat suivant est équivalent aux Théorèmes 1.4 et 1.5.

Théorème 1.7 *Soit D ouvert de \mathbb{R}^n , $f : D \rightarrow \mathbb{R}$ de classe C^2 et $a \in D$ un point critique. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres de $D^2f(a)$. On suppose que toutes les valeurs propres sont non nulles, c'est-à-dire le point critique a est régulier.*

1. *Si, pour tout $1 \leq i \leq n$, $\lambda_i > 0$, alors a est un minimum relatif strict de f .*
2. *Si, pour tout $1 \leq i \leq n$, $\lambda_i < 0$, alors a est un maximum relatif strict de f .*
3. *S'il existe $i, j \in \{1, \dots, n\}$ tels que $\lambda_i < 0$ et $\lambda_j > 0$, alors a n'est pas un extremum. On dit que a est un point selle ou col.*

Exemple 1. On considère la fonction $f(x, y) = x^3 - y^3 + 9xy$, $(x, y) \in \mathbb{R}^2$. On a

$$\frac{\partial f}{\partial x}(x, y) = 3x^2 + 9y, \quad \frac{\partial f}{\partial y}(x, y) = -3y^2 + 9x.$$

Alors, $\nabla f(x, y) = (0, 0)$ si et seulement si $(x, y) = (0, 0)$ ou $(x, y) = (3, -3)$. La fonction f admet deux points critiques : $(0, 0)$ et $(3, -3)$. On a

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 6x, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = 9, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = -6y.$$

Donc,

$$D^2f(0, 0) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix} \quad \text{et} \quad D^2f(3, -3) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}$$

La forme quadratique associée à $D^2f(0, 0)$ est donnée par

$$Q(h) = {}^t h D^2f(0, 0) h = 18h_1 h_2, \quad h = (h_1, h_2) \in \mathbb{R}^2,$$

et elle est indéfinie, donc $(0, 0)$ est un point selle. Par contre, les mineurs principaux de $D^2f(3, -3)$ sont tous positifs (respectivement, 18 et 243). Donc, $(3, -3)$ est un point de

minimum local de f dans \mathbb{R}^2 . Puisque, $f(x, 0) \rightarrow -\infty$ lorsque $x \rightarrow -\infty$, $(3, -3)$ n'est pas un point de minimum global de f dans \mathbb{R}^2 .

Exemple 2. On considère la fonction $f(x, y) = xy$, $(x, y) \in \mathbb{R}^2$. On a

$$\frac{\partial f}{\partial x}(x, y) = y, \quad \frac{\partial f}{\partial y}(x, y) = x.$$

Alors, $\nabla f(x, y) = (0, 0)$ si et seulement si $(x, y) = (0, 0)$. La fonction f admet un seul point critique : $(0, 0)$. On a

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 0, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = 1, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 0$$

et

$$D^2 f(0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Les deux valeurs propres de $D^2 f(0, 0)$: 1 et -1 ont des signes contraires, alors $(0, 0)$ est un point selle. En conclusion, f n'admet pas ni de maxima locaux ni de minima locaux dans \mathbb{R}^2 .

Exemple 3. On considère la fonction $f(x, y) = x^2 + y^2 - xy$, $(x, y) \in \mathbb{R}^2$. On va calculer les éventuels maxima et minima de f sur \mathbb{R}^2 et sur $B(0, 1)$. On a

$$\frac{\partial f}{\partial x}(x, y) = 2x - y, \quad \frac{\partial f}{\partial y}(x, y) = 2y - x.$$

Alors, $\nabla f(x, y) = (0, 0)$ si et seulement si $(x, y) = (0, 0)$. La fonction f admet un seul point critique : $(0, 0)$. On a

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 2, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = -1, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 2$$

et

$$D^2 f(0, 0) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Les mineurs principaux de $D^2 f(0, 0)$ d'ordre 1 et 2 sont respectivement 2 et 3 (deux nombres positifs). Donc, $(0, 0)$ est un point de minimum local de f dans \mathbb{R}^2 et aussi dans $B(0, 1)$. De plus, pour tout $(x, y) \in \mathbb{R}^2$,

$$f(x, y) = x^2 - xy + \frac{1}{4}y^2 + \frac{3}{4}y^2 = \left(x - \frac{y}{2}\right)^2 + \frac{3}{4}y^2 \geq 0 = f(0, 0).$$

D'où, $(0, 0)$ est point de minimum global de f dans \mathbb{R}^2 .

Exemple 4. On considère la fonction $f(x, y) = (x^2 + y^2)^2$, $(x, y) \in \mathbb{R}^2$. On va calculer les éventuels maxima et minima de f sur \mathbb{R}^2 et sur $B(0, 1)$. On a

$$\frac{\partial f}{\partial x}(x, y) = 4x(x^2 + y^2), \quad \frac{\partial f}{\partial y}(x, y) = 4y(x^2 + y^2).$$

Alors, $\nabla f(x, y) = (0, 0)$ si et seulement si $(x, y) = (0, 0)$. La fonction f admet un seul point critique : $(0, 0)$. On a

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 12x^2 + 4y^2, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = 8xy, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 12y^2 + 4x^2.$$

Alors, la matrice hessienne de f en $(0, 0)$ est nulle et on ne peut rien conclure à partir de sa forme quadratique associée. Mais on remarque que $f(x, y) \geq 0 = f(0, 0)$, pour tout $(x, y) \in \mathbb{R}^2$. Donc, $(0, 0)$ est le minimum global de f dans \mathbb{R}^2 et aussi dans $B(0, 1)$.

1.4 Cas des fonctions convexes ou concaves

On dit qu'un ensemble A de \mathbb{R}^n est convexe si et seulement si pour tout $\lambda \in [0, 1]$ et tout point $x, y \in A$ on a $\lambda x + (1 - \lambda)y \in A$. Cela signifie que le segment d'extrémités x et y est entièrement contenu dans A .

Une fonction définie sur un sous-ensemble convexe A de \mathbb{R}^n est *convexe* si, pour tout $\lambda \in [0, 1]$ et tout point $x, y \in A$ on a

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

De même, une fonction définie sur un sous-ensemble convexe A de \mathbb{R}^n est *concave* si, pour tout $\lambda \in [0, 1]$ et tout point $x, y \in A$ on a

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

Théorème 1.8 (Caractérisation des fonctions convexes) *Soit D un ouvert convexe de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$ une fonction de classe C^1 . Alors f est convexe si et seulement si, pour tout $x, y \in D$, on a*

$$\langle \nabla f(x), y - x \rangle \leq f(y) - f(x). \quad (1.3)$$

De plus, si f est de classe C^2 dans D , alors f est convexe si et seulement si, pour tout $x \in D$, tout $h \in \mathbb{R}^n$, on a

$${}^t h D^2 f(x) h \geq 0,$$

c'est-à-dire, la matrice hessienne $D^2 f(x)$ est semi-définie positive.

Théorème 1.9 (Caractérisation des fonctions concaves) *Soit D un ouvert convexe de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$ une fonction de classe C^1 . Alors f est concave si et seulement si, pour tout $x, y \in D$, on a*

$$\langle \nabla f(x), y - x \rangle \geq f(y) - f(x). \quad (1.4)$$

De plus, si f est de classe C^2 dans D , alors f est concave si et seulement si, pour tout $x \in D$, tout $h \in \mathbb{R}^n$, on a

$${}^t h D^2 f(x) h \leq 0,$$

c'est-à-dire, la matrice hessienne $D^2 f(x)$ est semi-définie négative.

La condition (1.3) nous dit qu'une fonction de classe C^1 est convexe si et seulement si le plan tangent au graphe de la fonction est toujours au-dessous du graphe. De façon analogue, la condition (1.4) nous dit qu'une fonction de classe C^1 est concave si et seulement si le plan tangent au graphe de la fonction est toujours au-dessus du graphe.

On peut démontrer facilement l'existence d'un minimum global pour une fonction convexe et l'existence d'un maximum global pour une fonction concave.

Théorème 1.10 Soit D un ouvert convexe de \mathbb{R}^n et $f : D \rightarrow \mathbb{R}$ une fonction de classe C^1 et $a \in D$ un point critique de f . Alors, si f est convexe, a est un point de minimum global de f sur D . Et, si f est concave, a est un point de maximum global de f sur D .

Exemple. On considère un monopole discriminant produisant un bien unique pour deux types de consommateurs, le consommateur 1 et le consommateur 2. Soient $p(q_1) = 50 - 5q_1$ et $p(q_2) = 100 - 10q_2$ le prix du bien en fonction de la quantité consommée par chacun des deux types de consommateurs. Soit $c(q) = 90 + 20q$ la fonction coût de la quantité totale $q = q_1 + q_2$ consommée. La fonction profit du monopole discriminant est alors

$$f(q_1, q_2) = q_1(50 - 5q_1) + q_2(100 - 10q_2) - (90 + 20q).$$

On a $\nabla f(q_1, q_2) = (0, 0)$ si et seulement si $(q_1, q_2) = (3, 4)$. La matrice hessienne de f en $(q_1, q_2) \in \mathbb{R}^2$ est

$$D^2 f(q_1, q_2) = \begin{pmatrix} -10 & 0 \\ 0 & -20 \end{pmatrix}$$

qui est une matrice semi-définie négative. Alors, f est concave sur \mathbb{R}^2 et le point critique $(3, 4)$ maximise le profit du monopole sur \mathbb{R}^2 et ainsi sur l'ensemble ouvert et convexe des quantités consommées admissibles $]0, +\infty[\times]0, +\infty[$.

2 Optimisation avec contraintes

Le problème de maximiser ou minimiser une fonction donnée sur un sous-ensemble de \mathbb{R}^n devient un problème *d'optimisation sous contraintes* quand l'ensemble D est un fermé. Un tel problème s'écrit souvent de la manière suivante : maximiser ou minimiser une fonction $f(x) = f(x_1, \dots, x_n)$ en sachant que $x \in \mathbb{R}^n$ doit satisfaire les contraintes d'égalité

$$h_1(x_1, \dots, x_n) = c_1, \quad h_2(x_1, \dots, x_n) = c_2, \quad \dots, \quad h_m(x_1, \dots, x_n) = c_m, \quad (2.1)$$

et les contraintes d'inégalité

$$g_1(x_1, \dots, x_n) \leq b_1, \quad g_2(x_1, \dots, x_n) \leq b_2, \quad \dots, \quad g_k(x_1, \dots, x_n) \leq b_k. \quad (2.2)$$

La fonction f est appelée *fonction objectif* et les fonctions h_i , $1 \leq i \leq m$, et g_j , $1 \leq j \leq k$, *fonctions contraintes*.

Exemple. (Le problème du consommateur). Soit $f(x) = f(x_1, \dots, x_n)$ la fonction d'utilité d'un consommateur, c'est-à-dire la fonction qui donne un indice de satisfaction d'un consommateur en fonction des quantités x_i , $i = 1, \dots, n$, consommées du bien i . Soit p_i , $i = 1, \dots, n$, le prix d'une unité du bien i et I le revenu du consommateur. Le problème du consommateur est alors celui de trouver les quantités x_i qui maximisent $f(x)$ en sachant que

$$p_1x_1 + p_2x_2 + \dots + p_nx_n \leq I$$

et

$$x_1 \geq 0, \dots, x_n \geq 0.$$

2.1 Contraintes d'égalité

On commence par considérer le problème d'optimisation sous les contraintes les plus simples, les contraintes d'égalité (2.1).

Soit $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une fonction vectorielle d'au moins classe C^1 sur \mathbb{R}^n . On a $h = (h_1, \dots, h_m)$ où $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$, pour tout $1 \leq i \leq m$. Soit $c = (c_1, \dots, c_m) \in \mathbb{R}^m$ et D l'ensemble des points \mathbb{R}^n respectant les m contraintes d'égalité (2.1),

$$D = \{x \in \mathbb{R}^n : h(x) = c\}. \quad (2.3)$$

L'ensemble D est un fermé de \mathbb{R}^n .

Il est utile d'introduire la *fonction de Lagrange* ou *Lagrangien* définie par :

$$\mathcal{L}(x_1, \dots, x_n, \mu_1, \dots, \mu_m) = f(x) - [\mu_1(h_1(x) - c_1) + \dots + \mu_m(h_m(x) - c_m)], \quad (2.4)$$

$x = (x_1, \dots, x_n) \in \mathbb{R}^n$ et où $\mu_1, \dots, \mu_m \in \mathbb{R}$ sont m nouvelles variables réelles appelées les *multiplicateurs de Lagrange*. Le Lagrangien \mathcal{L} est une fonction de $n + m$ variables réelles (x, μ_1, \dots, μ_m) . Il coïncide avec la fonction f sur l'ensemble D (défini par (2.3)) et il nous permet de transformer un problème sous contraintes à n variables (x_1, \dots, x_n) en un problème sans contraintes à $n + m$ variables.

Cas de la dimension 2 : Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $D = \{(x, y) \in \mathbb{R}^2 : h(x, y) = c\}$ (une unique contrainte d'égalité). La fonction de Lagrange associée est :

$$\mathcal{L}(x, y, \mu) = f(x, y) - \mu(h(x, y) - c). \quad (2.5)$$

Soit (\bar{x}, \bar{y}) un maximum ou minimum local de f sur l'ensemble D . On montre que la courbe de niveau $f(x, y) = f(\bar{x}, \bar{y})$ est tangente en (\bar{x}, \bar{y}) à la courbe représentant l'ensemble D . Donc, les vecteurs gradients $\nabla f(\bar{x}, \bar{y})$ et $\nabla h(\bar{x}, \bar{y})$ ont la même direction en (\bar{x}, \bar{y}) , c'est-à-dire qu'il existe $\bar{\mu} \in \mathbb{R}$ tel que

$$\nabla f(\bar{x}, \bar{y}) = \bar{\mu} \nabla h(\bar{x}, \bar{y}). \quad (2.6)$$

On obtient alors le théorème suivant.

Théorème 2.1 (Condition nécessaire en dimension 2 et une contrainte) *Soient f, h deux fonctions de classe C^1 dans \mathbb{R}^2 et soit (\bar{x}, \bar{y}) un point de maximum ou minimum de f sur $D = \{(x, y) \in \mathbb{R}^2 : h(x, y) = c\}$ tel que $\nabla h(\bar{x}, \bar{y}) \neq (0, 0)$. Alors, il existe un unique $\bar{\mu} \in \mathbb{R}$ tel que le point $(\bar{x}, \bar{y}, \bar{\mu}) \in \mathbb{R}^3$ est un point critique de la fonction de Lagrange définie par (2.5), c'est-à-dire*

$$\frac{\partial \mathcal{L}}{\partial x}(\bar{x}, \bar{y}, \bar{\mu}) = 0, \quad \frac{\partial \mathcal{L}}{\partial y}(\bar{x}, \bar{y}, \bar{\mu}) = 0, \quad \frac{\partial \mathcal{L}}{\partial \mu}(\bar{x}, \bar{y}, \bar{\mu}) = 0.$$

La condition $\nabla h(\bar{x}, \bar{y}) \neq (0, 0)$ est appelée *condition de qualification* de la contrainte. Si elle est satisfaite, le système (2.6) admet une solution unique. De plus, la condition

$$\frac{\partial \mathcal{L}}{\partial \mu}(\bar{x}, \bar{y}, \bar{\mu}) = -(h(\bar{x}, \bar{y}) - c) = 0 \iff h(\bar{x}, \bar{y}) = c$$

traduit le fait que $(\bar{x}, \bar{y}) \in D$.

Le théorème précédent nous donne uniquement l'ensemble des points candidats à point d'extremum de f sur D (c'est-à-dire, sous la contrainte d'égalité $h(x, y) = c$). Comme dans le cas d'un problème d'extremums sans contrainte, pour déterminer la solution il faut établir des conditions suffisantes (du second ordre) permettant d'identifier les maxima et minima de f sur D parmi l'ensemble des points satisfaisant la condition nécessaire.

Théorème 2.2 (Conditions suffisantes en dimension 2 et une contrainte) *Soient f, h deux fonctions de classe C^2 dans \mathbb{R}^2 ,*

$$D = \{(x, y) \in \mathbb{R}^2 : h(x, y) = c\} \quad \text{et} \quad \mathcal{L}(x, y, \mu) = f(x, y) - \mu(h(x, y) - c).$$

Soit $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ tel que $\nabla h(\bar{x}, \bar{y}) \neq (0, 0)$. Alors :

1. *(\bar{x}, \bar{y}) est un point de maximum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}$ tel que $(\bar{x}, \bar{y}, \bar{\mu})$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, y, \mu)$ et*

$${}^t v D_{x,y}^2 \mathcal{L}(\bar{x}, \bar{y}, \bar{\mu}) v < 0, \quad (2.7)$$

pour tout $v \in \mathbb{R}^2$ non nul tel que $\langle \nabla h(\bar{x}, \bar{y}), v \rangle = 0$.

2. *(\bar{x}, \bar{y}) est un point de minimum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}$ tel que $(\bar{x}, \bar{y}, \bar{\mu})$ est un point critique de la fonction de Lagrange \mathcal{L} et*

$${}^t v D_{x,y}^2 \mathcal{L}(\bar{x}, \bar{y}, \bar{\mu}) v > 0, \quad (2.8)$$

pour tout $v \in \mathbb{R}^2$ non nul tel que $\langle \nabla h(\bar{x}, \bar{y}), v \rangle = 0$.

Les conditions (2.7) et (2.8) stipulent que la forme quadratique associée à la matrice hessienne par rapport aux variables (x, y) du Lagrangien en $(\bar{x}, \bar{y}, \bar{\mu})$ est respectivement définie négative et définie positive sur l'hyperplan tangent à D au point (\bar{x}, \bar{y}) . Le théorème suivant, donne deux conditions suffisantes du second ordre plus faciles à vérifier que (2.7) et (2.8) et qui impliquent celles-ci.

Théorème 2.3 (Conditions suffisantes en dimension 2 et une contrainte) Soient f , h deux fonctions de classe C^2 dans \mathbb{R}^2 , $D = \{(x, y) \in \mathbb{R}^2 : h(x, y) = c\}$ et $\mathcal{L}(x, y, \mu) = f(x, y) - \mu(h(x, y) - c)$. Alors :

1. $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ est un point de maximum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}$ tel que $(\bar{x}, \bar{y}, \bar{\mu}) \in \mathbb{R}^3$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, y, \mu)$ et

$$\det \begin{pmatrix} 0 & {}^t\nabla h(\bar{x}, \bar{y}) \\ \nabla h(\bar{x}, \bar{y}) & D_{x,y}^2 \mathcal{L}(\bar{x}, \bar{y}, \bar{\mu}) \end{pmatrix} > 0.$$

2. $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ est un point de minimum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}$ tel que $(\bar{x}, \bar{y}, \bar{\mu}) \in \mathbb{R}^3$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, y, \mu)$ et

$$\det \begin{pmatrix} 0 & {}^t\nabla h(\bar{x}, \bar{y}) \\ \nabla h(\bar{x}, \bar{y}) & D_{x,y}^2 \mathcal{L}(\bar{x}, \bar{y}, \bar{\mu}) \end{pmatrix} < 0.$$

Remarque. On appelle matrice hessienne bordée associée à $D^2\mathcal{L}(\bar{x}, \bar{y}, \bar{\mu})$ la matrice

$$H = \begin{pmatrix} 0 & {}^t\nabla h(\bar{x}, \bar{y}) \\ \nabla h(\bar{x}, \bar{y}) & D_{x,y}^2 \mathcal{L}(\bar{x}, \bar{y}, \bar{\mu}) \end{pmatrix} = \begin{pmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{pmatrix} (\bar{x}, \bar{y}, \bar{\mu}).$$

On remarque alors que la condition de qualification de la contrainte $\nabla h(\bar{x}, \bar{y}) \neq (0, 0)$ est une conséquence des conditions $\det H > 0$ et $\det H < 0$.

Exemple. Soit $f(x, y) = xy$ et $D = \{(x, y) \in \mathbb{R}^2 : x + 4y = 16\}$. La fonction contrainte h est définie par $h(x, y) = x + 4y$ et $\nabla h(x, y) = (1, 4) \neq (0, 0)$ donc la condition de qualification est vérifiée en tout point de \mathbb{R}^2 . On définit la fonction de Lagrange

$$\mathcal{L}(x, y, \mu) = f(x, y) - \mu(h(x, y) - 16) = xy - \mu(x + 4y - 16).$$

On vérifie facilement que $\nabla \mathcal{L}(x, y, \mu) = (0, 0, 0)$ si et seulement si $(x, y, \mu) = (8, 2, 2)$. Finalement,

$$H(8, 2, 2) = \begin{vmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{vmatrix} (8, 2, 2) = \begin{vmatrix} 0 & 1 & 4 \\ 1 & 0 & 1 \\ 4 & 1 & 0 \end{vmatrix} = 8 > 0.$$

Donc, $(8, 2)$ est un point de maximum local de f sur D .

Cas général : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $D = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : h(x) = c\}$ où $h = (h_1, \dots, h_m)$ et $c = (c_1, \dots, c_m) \in \mathbb{R}^m$ avec $1 \leq m < n$. Dans ce cas, la condition de qualification de la contrainte est que la matrice jacobienne de la fonction vectorielle h :

$$Dh(x) = \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x) & \dots & \frac{\partial h_1}{\partial x_n}(x) \\ \frac{\partial h_2}{\partial x_1}(x) & \dots & \frac{\partial h_2}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(x) & \dots & \frac{\partial h_m}{\partial x_n}(x) \end{pmatrix}$$

soit de rang maximal et donc égal à m . Cette condition implique que les m lignes de la matrice jacobienne $Dh(x)$ sont linéairement indépendantes et donc l'espace vectoriel engendré par les m lignes est de dimension m . Les trois théorèmes précédents se généralisent comme suit.

Théorème 2.4 (Condition nécessaire en dimension n et $1 \leq m < n$ contraintes)

Soient f et h_1, \dots, h_m $m + 1$ fonctions de classe C^1 dans \mathbb{R}^n avec $1 \leq m < n$ et soit $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$ un point de maximum ou minimum de f sur $D = \{x \in \mathbb{R}^n : h(x) = c\}$ tel que la matrice jacobienne $Dh(\bar{x})$ est de rang m . Alors, il existe un unique $\bar{\mu} \in \mathbb{R}^m$ tel que le point $(\bar{x}, \bar{\mu}) \in \mathbb{R}^{n+m}$ est un point critique de la fonction de Lagrange définie par (2.5), c'est-à-dire

$$\frac{\partial \mathcal{L}}{\partial x_i}(\bar{x}, \bar{\mu}) = 0, \quad \frac{\partial \mathcal{L}}{\partial \mu_j}(\bar{x}, \bar{\mu}) = 0, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m.$$

On remarque que les conditions

$$\frac{\partial \mathcal{L}}{\partial \mu_j}(\bar{x}, \bar{\mu}) = -(h_j(\bar{x}) - c_j) = 0 \iff h_j(\bar{x}) = c_j, \quad 1 \leq j \leq m,$$

traduisent le fait que $\bar{x} \in D$.

Théorème 2.5 (Conditions suffisantes en dimension n et $1 \leq m < n$ contraintes)

Soient f et h_1, \dots, h_m $m + 1$ fonctions de classe C^2 dans \mathbb{R}^n , avec $1 \leq m < n$, $D = \{x \in \mathbb{R}^n : h_1(x) = c_1, \dots, h_m(x) = c_m\}$ et

$$\mathcal{L}(x_1, \dots, x_n, \mu_1, \dots, \mu_m) = f(x_1, \dots, x_n) - \sum_{j=1}^m \mu_j (h_j(x_1, \dots, x_n) - c_j).$$

Soit $\bar{x} \in \mathbb{R}^n$ tel que $Dh(\bar{x})$ est de rang m . Alors :

1. $\bar{x} \in \mathbb{R}^n$ est un point de maximum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}^m$ tel que $(\bar{x}, \bar{\mu}) \in \mathbb{R}^{n+m}$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, \mu)$ et

$${}^t v D_x^2 \mathcal{L}(\bar{x}, \bar{\mu}) v < 0, \quad (2.9)$$

pour tout $v \in \mathbb{R}^n$ non nul tel que $Dh(\bar{x})v = 0$.

2. $\bar{x} \in \mathbb{R}^n$ est un point de minimum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}^m$ tel que $(\bar{x}, \bar{\mu}) \in \mathbb{R}^{n+m}$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, \mu)$ et

$${}^t v D_x^2 \mathcal{L}(\bar{x}, \bar{\mu}) v > 0, \quad (2.10)$$

pour tout $v \in \mathbb{R}^n$ non nul tel que $Dh(\bar{x})v = 0$.

Théorème 2.6 (Conditions suffisantes en dimension n et $1 \leq m < n$ contraintes)

Soient f et h_1, \dots, h_m $m + 1$ fonctions de classe C^2 dans \mathbb{R}^n , avec $1 \leq m < n$, $D = \{x \in \mathbb{R}^n : h_1(x) = c_1, \dots, h_m(x) = c_m\}$ et

$$\mathcal{L}(x_1, \dots, x_n, \mu_1, \dots, \mu_m) = f(x_1, \dots, x_n) - \sum_{j=1}^m \mu_j (h_j(x_1, \dots, x_n) - c_j).$$

Alors :

1. $\bar{x} \in \mathbb{R}^n$ est un point de maximum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}^m$ tel que $(\bar{x}, \bar{\mu}) \in \mathbb{R}^{n+m}$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, \mu)$ et si la matrice

$$H(\bar{x}, \bar{\mu}) = \begin{pmatrix} 0 & Dh(\bar{x}) \\ {}^t Dh(\bar{x}) & D_x^2 \mathcal{L}(\bar{x}, \bar{\mu}) \end{pmatrix}$$

est telle que les $n - m$ derniers mineurs principaux alternent de signe, le déterminant de H étant du même signe que $(-1)^n$.

2. $\bar{x} \in \mathbb{R}^n$ est un point de minimum local de f sur D s'il existe $\bar{\mu} \in \mathbb{R}^m$ tel que $(\bar{x}, \bar{\mu}) \in \mathbb{R}^{n+m}$ est un point critique de la fonction de Lagrange $\mathcal{L}(x, \mu)$ et si la matrice

$$H(\bar{x}, \bar{\mu}) = \begin{pmatrix} 0 & Dh(\bar{x}) \\ {}^t Dh(\bar{x}) & D_x^2 \mathcal{L}(\bar{x}, \bar{\mu}) \end{pmatrix}$$

est telle que les $n - m$ derniers mineurs principaux sont tous du signe $(-1)^m$.

Remarque. On appelle *matrice hessienne bordée* associée à la matrice carrée d'ordre $(m + n) \times (m + n)$ définie par :

$$H = \begin{pmatrix} 0 & \cdots & 0 & \frac{\partial h_1}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial h_1}{\partial x_n}(\bar{x}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\partial h_k}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial h_k}{\partial x_n}(\bar{x}) \\ \frac{\partial h_1}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial h_m}{\partial x_1}(\bar{x}) & \frac{\partial^2 \mathcal{L}}{\partial x_1^2}(\bar{x}, \bar{\mu}) & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n}(\bar{x}, \bar{\mu}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n}(\bar{x}) & \cdots & \frac{\partial h_m}{\partial x_n}(\bar{x}) & \frac{\partial^2 \mathcal{L}}{\partial x_n \partial x_1}(\bar{x}, \bar{\mu}) & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2}(\bar{x}, \bar{\mu}) \end{pmatrix}.$$

La condition de qualification de la contrainte vectorielle h est maintenant une conséquence de la condition sur le signe du déterminant de la matrice bordée H .

2.2 Contraintes d'inégalité

La majorité des problèmes d'optimisation sous contraintes en économie sont caractérisés par des contraintes d'inégalité du type (2.2).

Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ avec $i = 1, \dots, k$, des fonctions d'au moins classe C^1 sur \mathbb{R}^n . Soit $b = (b_1, \dots, b_k) \in \mathbb{R}^k$ et D l'ensemble de points respectant les k contraintes d'inégalité (2.2) :

$$D = \{x \in \mathbb{R}^n : g_1(x) \leq b_1, \dots, g_k(x) \leq b_k\}.$$

L'ensemble D est un fermé de \mathbb{R}^n .

On définit la *fonction de Lagrange* ou *Lagrangien* par

$$\mathcal{L}(x_1, \dots, x_n, \lambda_1, \lambda_k) = f(x) - [\lambda_1(g_1(x) - b_1) + \dots + \lambda_k(g_k(x) - b_k)], \quad (2.11)$$

$x = (x_1, \dots, x_n)$ et où $\lambda_1, \dots, \lambda_k$ sont k nouvelles variables appelées *multiplicateurs de Lagrange*. On remarque que le Lagrangien (2.11) ne coïncide plus avec la fonction f sur l'ensemble D comme dans le cas des contraintes d'égalité mais il nous permet toujours de transformer le problème sous contraintes à n variables en un problème sans contraintes à $n + k$ variables.

Cas de la dimension 2 : Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $D = \{(x, y) \in \mathbb{R}^2 : g(x, y) \leq b\}$ (une unique contrainte d'inégalité). La fonction de Lagrange associée est :

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \mu(g(x, y) - b). \quad (2.12)$$

Soit (\bar{x}, \bar{y}) un maximum ou minimum local de f sur l'ensemble D . Deux cas sont possibles : $g(\bar{x}, \bar{y}) = b$ ou $g(\bar{x}, \bar{y}) < b$.

Dans le premier cas, $g(\bar{x}, \bar{y}) = b$, on dit que la contrainte est *saturée* en (\bar{x}, \bar{y}) . Comme dans le cas d'une contrainte d'égalité, la courbe de niveau $f(x, y) = f(\bar{x}, \bar{y})$ est tangente en (\bar{x}, \bar{y}) à la courbe représentant l'ensemble $\{(x, y) \in \mathbb{R}^2 : g(x, y) = b\}$. Donc, les vecteurs gradients $\nabla f(\bar{x}, \bar{y})$ et $\nabla g(\bar{x}, \bar{y})$ ont la même direction en (\bar{x}, \bar{y}) , c'est-à-dire qu'il existe $\bar{\lambda} \in \mathbb{R}$ tel que

$$\nabla f(\bar{x}, \bar{y}) = \bar{\lambda} \nabla g(\bar{x}, \bar{y}). \quad (2.13)$$

De plus, puisque le vecteur gradient d'une fonction admet comme direction celle pour laquelle la fonction s'accroît le plus rapidement, les vecteurs gradients $\nabla f(\bar{x}, \bar{y})$ et $\nabla g(\bar{x}, \bar{y})$ doivent s'orienter dans le même sens si (\bar{x}, \bar{y}) est un point de maximum et dans le sens opposé si (\bar{x}, \bar{y}) est un point de minimum. Donc, le multiplicateur de Lagrange dans (2.13) doit vérifier $\bar{\lambda} \geq 0$ si (\bar{x}, \bar{y}) est un point de maximum et $\bar{\lambda} \leq 0$ si (\bar{x}, \bar{y}) est un point de minimum.

Dans le cas $g(\bar{x}, \bar{y}) < b$, on dit que la contrainte *n'est pas saturée* en (\bar{x}, \bar{y}) et le point (\bar{x}, \bar{y}) est un point de maximum ou minimum local sans contraintes car il appartient à l'ensemble ouvert $\{(x, y) \in \mathbb{R}^2 : g(x, y) < b\}$. Il doit donc vérifier

$$\nabla f(\bar{x}, \bar{y}) = (0, 0),$$

et les dérivées de g n'interviennent pas dans la caractérisation de (\bar{x}, \bar{y}) .

On obtient le théorème suivant.

Théorème 2.7 (Condition nécessaire en dimension 2 et une contrainte) *Soient f et g deux fonctions de classe C^1 dans \mathbb{R}^2 et soit (\bar{x}, \bar{y}) un point de maximum ou minimum local de f sur l'ensemble $D = \{(x, y) \in \mathbb{R}^2 : g(x, y) \leq b\}$ tel que $\nabla g(\bar{x}, \bar{y}) \neq (0, 0)$ si $g(\bar{x}, \bar{y}) = b$. Alors, il existe un unique $\bar{\lambda} \in \mathbb{R}$ tel que*

$$\frac{\partial \mathcal{L}}{\partial x}(\bar{x}, \bar{y}, \bar{\lambda}) = 0, \quad \frac{\partial \mathcal{L}}{\partial y}(\bar{x}, \bar{y}, \bar{\lambda}) = 0,$$

et

$$\bar{\lambda}(g(\bar{x}, \bar{y}) - b) = 0,$$

avec $\bar{\lambda} \geq 0$ si (\bar{x}, \bar{y}) est un point de maximum et $\bar{\lambda} \leq 0$ si (\bar{x}, \bar{y}) est un point de minimum.

La généralisation naturelle du théorème précédent au cas des fonctions de n variables sous k contraintes d'inégalité est la suivante.

Théorème 2.8 (Condition nécessaire en dimension n et k contraintes) *Soient f et g_1, \dots, g_k , $k+1$ fonctions de classe C^1 dans \mathbb{R}^n et soit $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$ un point de maximum ou minimum local de f sur l'ensemble*

$$D = \{x \in \mathbb{R}^n : g_1(x) \leq b_1, \dots, g_k(x) \leq b_k\}.$$

Sans perte de généralité, on suppose qu'il existe $k_0 \leq k$ tel que

$$g_1(\bar{x}) - b_1 = \dots = g_{k_0}(\bar{x}) - b_{k_0} = 0 \quad \text{et} \quad g_{k_0+1}(\bar{x}) < b_{k_0+1}, \dots, g_k(\bar{x}) < b_k,$$

c'est-à-dire les k_0 premières contraintes sont saturées en \bar{x} et les dernières $k - k_0$ ne le sont pas. Supposons enfin que le rang de la matrice jacobienne des k_0 contraintes saturées calculée en \bar{x}

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\bar{x}) & \dots & \frac{\partial g_1}{\partial x_n}(\bar{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{k_0}}{\partial x_1}(\bar{x}) & \dots & \frac{\partial g_{k_0}}{\partial x_n}(\bar{x}) \end{pmatrix}$$

est maximal et donc égal à k_0 . Alors, il existe un unique $\bar{\lambda} = (\bar{\lambda}_1, \dots, \bar{\lambda}_k) \in \mathbb{R}^k$ tel que

$$\frac{\partial \mathcal{L}}{\partial x_i}(\bar{x}, \bar{\lambda}) = 0, \quad i = 1, \dots, n,$$

et

$$\bar{\lambda}_j(g_j(\bar{x}) - b_j) = 0, \quad j = 1, \dots, k,$$

avec $\bar{\lambda}_j \geq 0$ pour tout $j = 1, \dots, k$, si \bar{x} est un point de maximum et $\bar{\lambda}_j \leq 0$ pour tout $j = 1, \dots, k$, si \bar{x} est un point de minimum.

Remarque. Dans les théorèmes précédents, le signe des multiplicateurs de Lagrange est lié aux inégalités définies par les k contraintes

$$g_i(x) \leq b_i \quad i = 1, \dots, k.$$

Si l'on change les inégalités en

$$g_i(x) \geq b_i \quad i = 1, \dots, k,$$

il faut changer aussi le signe des multiplicateurs de Lagrange.

Exemple. On considère à nouveau le problème du consommateur : trouver les quantités x_i qui maximisent $f(x) = f(x_1, \dots, x_n)$ en sachant que

$$p_1x_1 + p_2x_2 + \dots + p_nx_n \leq I$$

et

$$x_1 \geq 0, \dots, x_n \geq 0.$$

Sous une hypothèse de “non satiété” et en ignorant les contraintes de non négativité $x_1 \geq 0, \dots, x_n \geq 0$, le problème devient un problème d’optimisation sous contraintes d’égalité. En effet, l’hypothèse de “non satiété” peut se traduire par : $\frac{\partial f}{\partial x_i}(x) > 0$, pour au moins un indice i en chaque panier $x = (x_1, \dots, x_n)$. Donc, le multiplicateur de Lagrange λ , dans la fonction de Lagrange

$$\mathcal{L}(x, \lambda) = f(x) - \lambda(p_1x_1 + p_2x_2 + \dots + p_nx_n - I),$$

ne peut pas être nul au point de maximum, car cas contraire, on aurait

$$\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) - \lambda(p_1, \dots, p_n) = \vec{0} \implies \frac{\partial f}{\partial x_i}(x) = 0, \quad \text{pour tout } i = 1, \dots, n.$$

D’où $p_1x_1 + p_2x_2 + \dots + p_nx_n = I$: la contrainte de budget est saturée au point de maximum, le consommateur dépense tout son revenu disponible.

Exemple. Soit $f(x, y) = xy$ et $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$. On calcule les éventuels maxima et minima de f sur D . La fonction contrainte est $g(x, y) = x^2 + y^2$ et $\nabla g(x, y) = (2x, 2y) \neq (0, 0)$ pour tout $(x, y) \in D \setminus (0, 0)$. Donc, la condition de qualification est vérifiée en tout point où la contrainte est saturée. On définit la fonction de Lagrange

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - 1) = xy - (x^2 + y^2 - 1).$$

Les conditions nécessaires impliquent :

$$\frac{\partial \mathcal{L}}{\partial x}(x, y, \lambda) = y - 2\lambda x = 0, \quad \frac{\partial \mathcal{L}}{\partial y}(x, y, \lambda) = x - 2\lambda y = 0, \quad \lambda(x^2 + y^2 - 1) = 0.$$

Si $\lambda = 0$, alors $(x, y) = (0, 0)$. Mais la fonction f atteint des valeurs positives et négatives en toute boule centrée en $(0, 0)$. Donc, $(0, 0)$ est un point selle.

Si $\lambda \neq 0$, alors la contrainte est saturée, c'est-à-dire $x^2 + y^2 = 1$, et

$$\lambda = \frac{y}{2x} = \frac{x}{2y}.$$

Alors, les points critiques de \mathcal{L} sont :

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{1}{2}\right), \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{1}{2}\right), \quad \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, -\frac{1}{2}\right), \quad \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, -\frac{1}{2}\right).$$

Comme la contrainte est saturée en ces points, les conditions suffisantes du second ordre pour des contraintes d'égalité impliquent que les points $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ et $(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ sont deux points de maximum de f sur D et que $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ et $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ sont deux points de minimum de f sur D .

3 Équations et systèmes de récurrence

Dans ce chapitre, on considère un vecteur $u_n \in \mathbb{R}^k$ qui représente k états à la date n . Le but est de déterminer les états u_n pour toute date n à partir des états initiaux à l'aide d'une relation de récurrence. On est ramené à l'étude d'une suite numérique vectorielle et de son comportement (ou nature) lorsque n tend vers $+\infty$.

On rappelle qu'une suite numérique est une application de \mathbb{N} dans un espace vectoriel F , notée $(u_n)_{n \in \mathbb{N}}$ où $u_n = f(n)$ est appelé le *terme général* de la suite. Si $F = \mathbb{R}^k$ ou $F = \mathbb{C}^k$, on dit que la suite est réelle ou complexe, respectivement. On rappelle que l'espace des suites numériques $(u_n)_{n \in \mathbb{N}}$ muni de l'addition : $(u_n)_{n \in \mathbb{N}} + (v_n)_{n \in \mathbb{N}} = (u_n + v_n)_{n \in \mathbb{N}}$, et de la multiplication par un scalaire $\lambda : \lambda(u_n)_{n \in \mathbb{N}} = (\lambda u_n)_{n \in \mathbb{N}}$, est un espace vectoriel.

3.1 Équations de récurrence linéaires d'ordre p à coefficients réels constants

Soit $p \in \mathbb{N}$. On appelle *équation de récurrence linéaire d'ordre p à coefficients réels constants* une équation de la forme

$$a_0 u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = f_n, \quad \forall n \geq p, \quad (3.1)$$

où $a_0, a_1, \dots, a_p \in \mathbb{R}$ avec $a_0 \neq 0$ et $(f_n)_{n \in \mathbb{N}}$ est une suite réelle. La suite réelle $(u_n)_{n \in \mathbb{N}}$ est l'inconnue à déterminer.

L'équation (3.1) décrit l'évolution de l'état u_n à la date n en fonction de la valeur de l'état aux dates $n-p, \dots, n-1$. Si l'on connaît les p états u_0, \dots, u_{p-1} , alors on connaît u_n pour tout n . On appelle les p états u_0, \dots, u_{p-1} les *conditions initiales* de (3.1).

Toute suite réelle $(u_n)_{n \in \mathbb{N}}$ vérifiant l'équation (3.1) est appelée une *solution réelle* de (3.1). Résoudre (3.1) consiste à trouver toutes les solutions réelles de cette équation, c'est-à-dire la *solution générale* et réelle.

Si $f_n = 0$, pour tout $n \in \mathbb{N}$, on dit que l'équation de récurrence est *homogène*.

On considère l'équation homogène associée à (3.1) :

$$a_0 u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = 0, \quad \forall n \geq p, \quad (3.2)$$

L'espace des solutions de (3.2) est un sous-espace vectoriel de dimension p de l'espace des suites numériques. Pour le calculer, il suffit de déterminer une base $\{(u_n^1)_{n \in \mathbb{N}}, \dots, (u_n^p)_{n \in \mathbb{N}}\}$, c'est-à-dire p solutions linéairement indépendantes de (3.2). Alors, la solution générale $(u_n)_{n \in \mathbb{N}}$ de (3.2) est

$$u_n = \alpha_1 u_n^1 + \dots + \alpha_p u_n^p, \quad \forall n \in \mathbb{N},$$

avec $\alpha_1, \dots, \alpha_p \in \mathbb{R}$.

La solution générale $(z_n)_{n \in \mathbb{N}}$ de (3.1) est de la forme

$$z_n = u_n + v_n$$

où $(u_n)_{n \in \mathbb{N}}$ est la solution générale de l'équation homogène associée (3.2) et $(v_n)_{n \in \mathbb{N}}$ est une solution particulière de (3.1).

Remarque. L'équation de récurrence (3.1) s'écrit de façon équivalente comme

$$a_0 v_{n+p} + a_1 v_{n+p-1} + \dots + a_p v_n = f_{n+p}, \quad \forall n \geq 0.$$

Il suffit de poser $v_{n+p} = u_n$, pour tout $n \in \mathbb{N}$.

Sans perte de généralité, on suppose dans la suite $a_0 = 1$.

Équations homogènes d'ordre 1. L'équation de récurrence (3.2) s'écrit

$$u_n = au_{n-1}, \quad n \geq 1,$$

où $a \in \mathbb{R}$. Il s'agit d'une suite géométrique de raison a et premier terme u_0 . La solution générale est $u_n = a^n u_0$.

Équations homogènes d'ordre p , $p \geq 2$.

Les solutions de l'équation homogène (3.2) (avec $a_0 = 1$) :

$$u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = 0, \quad \forall n \geq p,$$

forment un espace vectoriel de dimension p . Nous considérons l'équation caractéristique associée à (3.2) :

$$\lambda^p + a_1 \lambda^{p-1} + a_2 \lambda^{p-2} + \dots + a_{p-1} \lambda + a_p = 0. \quad (3.3)$$

Cette équation admet exactement p racines dans \mathbb{C} . Nous remarquons que si λ est une racine (réelle ou complexe) de (3.3), alors la suite $(u_n)_{n \in \mathbb{N}}$ définie par $u_n = \lambda^n$ est solution (réelle ou complexe) de (3.2). Pour déterminer la solution générale de (3.2), nous avons les règles suivantes.

1. Si les p racines $\lambda_1, \dots, \lambda_p$ de l'équation caractéristique (3.3) sont réelles et distinctes, la solution générale $(u_n)_{n \in \mathbb{N}}$ réelle de (3.2) est

$$u_n = \alpha_1 \lambda_1^n + \dots + \alpha_p \lambda_p^n, \quad \alpha_1, \dots, \alpha_p \in \mathbb{R}.$$

2. Si λ est une racine réelle de multiplicité k de (3.3), on lui associe les k solutions réelles et linéairement indépendantes :

$$(\lambda^n)_{n \in \mathbb{N}}, (n\lambda^n)_{n \in \mathbb{N}}, \dots, (n^{k-1}\lambda^n)_{n \in \mathbb{N}}.$$

3. Si $\lambda = \rho e^{i\theta}$ et $\bar{\lambda} = \rho e^{-i\theta}$ sont deux racines complexes conjuguées de (3.3) de multiplicité k , on leur associe les $2k$ solutions réelles et linéairement indépendantes de (3.2) :

$$(n^j \rho^n \cos(n\theta))_{n \in \mathbb{N}}, (n^j \rho^n \sin(n\theta))_{n \in \mathbb{N}},$$

avec $0 \leq j \leq k - 1$. Remarquer que

$$n^j \rho^n \cos(n\theta) = \operatorname{Re}(n^j \lambda^n) \quad \text{et} \quad n^j \rho^n \sin(n\theta) = \operatorname{Im}(n^j \lambda^n).$$

Remarque. Supposant $a_p \neq 0$, $\lambda = 0$ n'est pas racine de l'équation caractéristique (3.3).

Exemple. Résoudre l'équation de récurrence homogène d'ordre 3 :

$$u_n - u_{n-1} + u_{n-2} - u_{n-3} = 0, \quad n \geq 3.$$

L'équation caractéristique associée est : $\lambda^3 - \lambda^2 + \lambda - 1 = 0$. On a $\lambda^3 - \lambda^2 + \lambda - 1 = (\lambda - 1)(\lambda^2 + 1)$. Alors, les racines de l'équation caractéristique sont : $\lambda_1 = 1$, $\lambda_2 = i = e^{i\frac{\pi}{2}}$ et $\lambda_3 = -i = e^{-i\frac{\pi}{2}}$. La solution générale $(u_n)_{n \in \mathbb{N}}$ est alors définie par

$$u_n = \alpha_1 + \alpha_2 \cos(n\frac{\pi}{2}) + \alpha_3 \sin(n\frac{\pi}{2}), \quad \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}.$$

Pour déterminer une solution particulière de l'équation non homogène

$$u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = f_n, \quad \forall n \geq p, \quad (3.4)$$

nous pouvons appliquer la *méthode des coefficients indéterminés*, pour des seconds membres $(f_n)_{n \in \mathbb{N}}$ particuliers.

Équations non homogènes d'ordre 1.

On considère l'équation

$$u_n = a u_{n-1} + f_n, \quad n \geq 1, \quad (3.5)$$

donc la solution général de l'équation homogène associée est $u_n = \alpha a^n$, $\alpha \in \mathbb{R}$. Il y a plusieurs cas.

1. $f_n = b$, pour tout $n \in \mathbb{N}$ (second membre constant), alors $(u_n)_{n \in \mathbb{N}}$ est une suite arithmético-géométrique et on cherche une solution particulière $(v_n)_{n \in \mathbb{N}}$ de (3.5) de la façon suivante :

- (a) Si $a \neq 1$, $v_n = c$ avec $c \in \mathbb{R}$. En effet, $c = \frac{b}{1-a}$ et la solution générale de (3.5) est la suite de terme général

$$u_n = \alpha a^n + \frac{b}{1-a}, \quad \alpha \in \mathbb{R}.$$

- (b) Si $a = 1$, $v_n = n\gamma$ avec $\gamma \in \mathbb{R}$. Alors, $\gamma = b$ et la solution générale de (3.5) est la suite de terme général $u_n = u_0 + nb$.

2. f_n est un polynôme en n de degré $k \geq 1$,

$$f_n = P_k(n) = \beta_k n^k + \beta_{k-1} n^{k-1} + \dots + \beta_0,$$

avec $\beta_k \neq 0$. On cherche alors une solution particulière $(v_n)_{n \in \mathbb{N}}$ de (3.5) de la façon suivante :

- (a) si $a \neq 1$, $v_n = Q_k(n)$ où Q_k est un polynôme de degré k en n .
 - (b) si $a = 1$, $v_n = nQ_k(n)$ où Q_k est un polynôme de degré k en n .
3. $f_n = n^p r^n$, $r \in \mathbb{R}$. Alors, (3.5) admet une solution particulière $(v_n)_{n \in \mathbb{N}}$ de la forme :
 - (a) si $a \neq r$, $v_n = Q_p(n) r^n$ où Q_p est un polynôme en n de degré p .
 - (b) si $a = r$, $v_n = nQ_p(n) r^n$ où Q_p est un polynôme en n de degré p .
 4. $(f_n)_{n \in \mathbb{N}}$ est quelconque. Alors, il est facile de vérifier que la solution réelle de (3.5) correspondant à $u_0 = 0$ est la suite $(u_n)_{n \in \mathbb{N}}$ définie par

$$u_n = \sum_{j=0}^{n-1} a^j f_{n-j}, \quad n \geq 1.$$

La solution générale de (3.5) est alors

$$u_n = a^n u_0 + \sum_{j=0}^{n-1} a^j f_{n-j}, \quad n \geq 1,$$

que l'on appelle *solution rétrograde*.

Équations non homogènes d'ordre p , $p \geq 2$.

On considère l'équation

$$u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = f_n, \quad n \geq p. \quad (3.6)$$

Supposons $f_n = r^n P(n)$ où $r \in \mathbb{R}$ et P est un polynôme de degré k en n . Alors, (3.6) admet une solution particulière $(v_n)_{n \in \mathbb{N}}$ de la forme :

1. $v_n = r^n Q(n)$, si r n'est pas racine de l'équation caractéristique associée,
2. $v_n = n^m r^n Q(n)$, si r est racine de multiplicité m de l'équation caractéristique associée (avec $1 \leq m \leq p$),

où $Q(n)$ est dans les deux cas un polynôme de degré k en n .

Exemple 1. On considère l'équation de récurrence

$$u_n + u_{n-1} - 2u_{n-2} = (-2)^n, \quad n \geq 2.$$

L'équation caractéristique associée est

$$\lambda^2 + \lambda - 2 = 0.$$

Ses racines sont : $\lambda_1 = 1$ et $\lambda_2 = -2$. Alors, la solution générale de l'équation homogène associée est

$$u_n = \alpha + \beta(-2)^n, \quad \alpha, \beta \in \mathbb{R}.$$

Le second membre de l'équation non homogène est $f_n = (-2)^n$ et -2 est une racine simple du polynôme caractéristique de l'équation homogène associée. Donc, on cherche une solution particulière de la forme

$$v_n = bn(-2)^n, \quad b \in \mathbb{R}.$$

On obtient $b = \frac{2}{3}$. Finalement, la solution générale de l'équation non homogène est

$$u_n = \alpha + \beta(-2)^n + \frac{2}{3}n(-2)^n, \quad \alpha, \beta \in \mathbb{R}.$$

Exemple 2. (Équation de mouvement du capital) Soit K_n le capital à la date n , I_n l'investissement à la date n et δ le taux de dépréciation. On a

$$K_n = (1 - \delta)K_{n-1} + I_n.$$

Alors,

$$K_n = K_0(1 - \delta)^n + \sum_{k=1}^n (1 - \delta)^{n-k} I_k.$$

3.2 Systèmes de récurrence linéaires à coefficients réels constants et d'ordre 1

Un système de récurrence linéaire à coefficients constants et d'ordre 1 est un système du type

$$U_n = AU_{n-1} + F_n, \quad n \geq 1, \quad (3.7)$$

où $U_n \in \mathbb{R}^p$ est le vecteur inconnu de p états à déterminer pour tout n , A est une matrice carrée à coefficients réels d'ordre p et $(F_n)_{n \in \mathbb{N}}$ est une suite de vecteurs de \mathbb{R}^p donnée.

Résoudre (3.7) consiste à trouver toutes les solutions réelles de (3.7), c'est-à-dire la solution générale. Le système homogène associée à (3.7) est

$$U_n = AU_{n-1}, \quad n \geq 1. \quad (3.8)$$

La solution générale $(Z_n)_{n \in \mathbb{N}}$ de (3.7) est de la forme $Z_n = U_n + V_n$ où $(U_n)_{n \in \mathbb{N}}$ est la solution générale de (3.8) et $(V_n)_{n \in \mathbb{N}}$ est une solution particulière de (3.7). Les solutions de (3.8) forment un sous-espace vectoriel de dimension p de l'espace des suites de vecteurs de \mathbb{R}^p .

La solution générale de (3.8) est la suite réelle $(U_n)_{n \in \mathbb{N}}$ donnée par $U_n = A^n C$, avec $C \in \mathbb{R}^p$ vecteur arbitraire. Déterminer la solution générale de (3.8) se ramène donc à déterminer la puissance n -ième de A qui se fait à l'aide de la diagonalisation ou de la réduction à la forme de Jourdan de la matrice A .

On remarque également que

1. Si $\lambda \in \mathbb{R}$ est une valeur propre de A et $V \in \mathbb{R}^p$ est un vecteur propre associé à λ , alors $(U_n)_{n \in \mathbb{N}}$, avec $U_n = \lambda^n V$, est solution de (3.8).
2. Si $\mu = \rho e^{i\theta}$ et $\bar{\mu} = \rho e^{-i\theta}$ sont deux valeurs propres complexes conjuguées de A et $V \in \mathbb{R}^p$ est un vecteur propre associé à μ (dans \mathbb{C}^p), alors les suites de termes généraux

$$\operatorname{Re}(\mu^n V) \quad \text{et} \quad \operatorname{Im}(\mu^n V)$$

sont des solutions réelles et linéairement indépendantes de (3.8). Si $V = W + iZ$ avec $W, Z \in \mathbb{R}^p$, on a

$$\operatorname{Re}(\mu^n V) = \rho^n (\cos(n\theta)W - \sin(n\theta)Z) \quad \text{et} \quad \operatorname{Im}(\mu^n V) = \rho^n (\cos(n\theta)Z + \sin(n\theta)W).$$

On a alors les deux cas suivants.

A est diagonalisable dans \mathbb{R} .

Il existe une base (V_1, \dots, V_p) de \mathbb{R}^p formé de vecteurs propres de A , de valeurs propres correspondants : $\lambda_1, \dots, \lambda_p$. Alors, $A = PDP^{-1}$ où P est la matrice inversible dont la j -ième colonne est V_j et

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots \\ 0 & \lambda_2 & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & \lambda_p \end{pmatrix}.$$

La solution générale de (3.8) est alors donnée par la combinaison linéaire des p solutions linéairement indépendantes de (3.8) :

$$(\lambda_1^n V_1)_{n \in \mathbb{N}}, \dots, (\lambda_p^n V_p)_{n \in \mathbb{N}},$$

c'est-à-dire

$$U_n = \alpha_1 \lambda_1^n V_1 + \dots + \alpha_p \lambda_p^n V_p, \quad \alpha_1, \dots, \alpha_p \in \mathbb{R}. \quad (3.9)$$

En effet, on a

$$A^n = PD^nP^{-1} \quad \text{avec} \quad D^n = \begin{pmatrix} \lambda_1^n & 0 & 0 & \dots \\ 0 & \lambda_2^n & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & \lambda_p^n \end{pmatrix}.$$

La solution générale du système est $(U_n)_{n \in \mathbb{N}}$ de terme général

$$U_n = A^n C = PD^nP^{-1}C$$

On note $\tilde{C} = P^{-1}C \in \mathbb{R}^p$. Alors,

$$U_n = PD^n\tilde{C} = \sum_{j=1}^p \tilde{C}_j \lambda_j^n V_j.$$

La solution générale de (3.8) est alors donnée par (3.9).

A n'est pas diagonalisable dans \mathbb{R} .

Si A est diagonalisable dans \mathbb{C} , il existe (V_1, \dots, V_p) base de \mathbb{C}^p formé de vecteurs propres de A . On peut supposer, sans perte de généralité, que $V_1, \dots, V_k \in \mathbb{R}^p$ de valeur propres correspondantes $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ et $V_{k+1}, \dots, V_p \in \mathbb{R}^p$ de valeur propres correspondantes $\mu_1, \bar{\mu}_1, \dots, \mu_m, \bar{\mu}_m \in \mathbb{C}$, avec $p = k + 2m$. La solution générale de (3.8) est alors donnée par

$$U_n = \alpha_1 \lambda_1^n V_1 + \dots + \alpha_k \lambda_k^n V_k + \beta_1 \operatorname{Re}(\mu_1^n V_{k+1}) + \beta_2 \operatorname{Im}(\mu_1^n V_{k+1}) + \dots + \beta_{2m-1} \operatorname{Re}(\mu_m^n V_{p-1}) + \beta_{2m} \operatorname{Im}(\mu_m^n V_{p-1}).$$

Si A n'est pas diagonalisable ni dans \mathbb{R} ni dans \mathbb{C} , on peut toujours trouver une matrice de Jourdan J carrée et triangulaire supérieure d'ordre p et une matrice inversible P telles que $A = PJP^{-1}$. Donc, la solution de (3.8) est de la forme $U_n = A^n C = PJ^nP^{-1}C = PJ^n\tilde{C}$ avec $\tilde{C} = P^{-1}C \in \mathbb{R}^p$. Enfin, en résolvant le système

$$\tilde{U}_n = P^{-1}U_n = J^n\tilde{C}$$

à partir de la dernière équation (car la matrice J^n est triangulaire supérieure), on obtient \tilde{U}_n et ensuite $U_n = P\tilde{U}_n$.

Dans le cas d'un système de deux équations ($p = 2$), si A n'est pas diagonalisable (ni dans \mathbb{R} ni dans \mathbb{C}) alors A admet une seule valeur propre λ_0 réelle telle que $\ker(A - \lambda_0 I)$ est un sous-espace propre de dimension 1.

On peut trouver une matrice inversible P telle que $A = PJP^{-1}$ où $J = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$. Les colonnes P_1 et P_2 de la matrice P sont respectivement un vecteur propre associé à λ_0 et un vecteur tel que

$$(A - \lambda_0 I)P_2 = P_1.$$

En effet,

$$A = PJP^{-1} \iff AP = PJ \iff \begin{cases} AP_1 = \lambda_0 P_1 \\ AP_2 = P_1 + \lambda_0 P_2 \end{cases}$$

Il est facile de vérifier (par récurrence sur n) que

$$J^n = \begin{pmatrix} \lambda_0^n & n\lambda_0^n \\ 0 & \lambda_0^n \end{pmatrix}.$$

La solution générale de (3.8) est alors

$$U_n = A^n C = PJ^n \tilde{C} = (\tilde{c}_1 \lambda_0^n + \tilde{c}_2 n \lambda_0^{n-1}) P_1 + \tilde{c}_2 \lambda_0^n P_2,$$

avec $\tilde{C} = (\tilde{c}_1, \tilde{c}_2)^t = P^{-1} C \in \mathbb{R}^2$.

Exemple 1. On considère deux populations dont la dynamique est décrite par le système suivant

$$\begin{cases} x_{n+1} = \frac{1}{3}x_n + \frac{1}{2}y_n \\ y_{n+1} = \frac{2}{3}x_n + \frac{1}{2}y_n \end{cases}$$

Étudier l'évolution des populations x_n, y_n lorsque n tend vers $+\infty$.

On pose $U_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$. Alors, on a $U_{n+1} = AU_n$, pour tout $n \geq 0$, avec

$$A = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

Les valeurs propres de A sont $\lambda_1 = -\frac{1}{6}$ et $\lambda_2 = 1$. Alors, A est diagonalisable dans \mathbb{R} . On a $\ker(A - \lambda_1) = \mathbb{R} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $\ker(A - \lambda_2) = \mathbb{R} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Donc, la solution générale du système de récurrence est

$$U_n = \alpha \left(-\frac{1}{6}\right)^n \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \alpha \left(-\frac{1}{6}\right)^n + 3\beta \\ -\alpha \left(-\frac{1}{6}\right)^n + 4\beta \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R}.$$

Finalement, on conclut que $(x_n, y_n) \rightarrow (3\beta, 4\beta)$ lorsque $n \rightarrow +\infty$.

Exemple 2. On considère le système de récurrence $U_{n+1} = AU_n$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Les valeurs propres de A sont i et $-i$. Un vecteur propre associé à i est $P = (1, -i) = W + iZ$, avec $W = (1, 0)$ et $Z = (0, -1)$. Alors, la solution générale est

$$U_n = \alpha(\cos(n\frac{\pi}{2})W - \sin(n\frac{\pi}{2})Z) + \beta(\cos(n\frac{\pi}{2})Z + \sin(n\frac{\pi}{2})W) = \alpha \begin{pmatrix} \cos(n\frac{\pi}{2}) \\ \sin(n\frac{\pi}{2}) \end{pmatrix} + \beta \begin{pmatrix} \sin(n\frac{\pi}{2}) \\ -\cos(n\frac{\pi}{2}) \end{pmatrix},$$

$\alpha, \beta \in \mathbb{R}$.

Exemple 3. On considère le système de récurrence $U_{n+1} = AU_n$ avec $A = \begin{pmatrix} 4 & 1 \\ -1 & 2 \end{pmatrix}$.

On a $\det(A - \lambda I) = (\lambda - 3)^2$. Alors, 3 est l'unique valeur propre de A , d'ordre de multiplicité

2. De plus, $\ker(A - 3I) = \mathbb{R} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et donc A n'est pas diagonalisable. On a $A = PJP^{-1}$ avec

$$P = (P_1, P_2), \quad P_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad J = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$$

et $P_2 = \begin{pmatrix} x \\ y \end{pmatrix}$ telle que

$$AP_2 = P_1 + 3P_2 \iff x + y = 1.$$

Soit $P_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Alors, la solution générale du système de récurrence est

$$U_n = (\alpha 3^n + \beta n 3^n) P_1 + \beta 3^n P_2 = \alpha 3^n \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \beta 3^{n-1} \begin{pmatrix} n \\ 3 - n \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{R}$.

Écriture matricielle. On remarque que toute équation de récurrence linéaire d'ordre p à coefficients réels constants

$$u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = f_n, \quad \forall n \geq p,$$

peut s'écrire sous la forme d'un système de récurrence linéaire à coefficients réels constants d'ordre 1.

Lorsque $p = 2$, on pose $U_n = \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}$. L'équation de récurrence

$$u_n + a_1 u_{n-1} + a_2 u_{n-2} = f_n, \quad \forall n \geq 2,$$

s'écrit alors

$$\begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix} = \begin{pmatrix} -a_1 & -a_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_{n-1} \\ u_{n-2} \end{pmatrix} + \begin{pmatrix} f_n \\ 0 \end{pmatrix},$$

soit $U_n = AU_{n-1} + F_n$ où

$$A = \begin{pmatrix} -a_1 & -a_2 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad F_n = \begin{pmatrix} f_n \\ 0 \end{pmatrix}.$$

Dans le cas général, on pose $U_n = \begin{pmatrix} u_n \\ u_{n-1} \\ \vdots \\ u_{n-p+1} \end{pmatrix}$. Le système de récurrence d'ordre 1 s'écrit

alors $U_n = AU_{n-1} + F_n$ avec

$$A = \begin{pmatrix} -a_1 & -a_2 & \dots & -a_p \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & \dots & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix} \quad \text{et} \quad F_n = \begin{pmatrix} f_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

3.3 Équilibre et stabilité pour une équation de récurrence

On considère l'équation de récurrence non homogène d'ordre p

$$u_n + a_1 u_{n-1} + \dots + a_p u_{n-p} = c, \quad \forall n \geq p, \quad (3.10)$$

où $c \in \mathbb{R}$. Dans cette équation le second membre est constant, c'est-à-dire $f_n = c$, pour tout $n \in \mathbb{N}$.

On appelle *état stationnaire* ou *équilibre* de l'équation (3.10) tout nombre $u_e \in \mathbb{R}$ tel que la suite constante $(u_n)_{n \in \mathbb{N}}$ définie par $u_n = u_e$, pour tout $n \in \mathbb{N}$, est solution de (3.10).

Il est clair que u_e est un équilibre de (3.10) si et seulement si

$$u_e(1 + a_1 + a_2 + \dots + a_p) = c.$$

Nous avons les cas suivants.

1. Si $1 + a_1 + a_2 + \dots + a_p \neq 0$, alors l'équation (3.10) admet un seul équilibre donné par

$$u_e = c(1 + a_1 + a_2 + \dots + a_p)^{-1}.$$

En particulier, si $c = 0$, alors $u_e = 0$.

2. Si $1 + a_1 + a_2 + \dots + a_p = 0$ et $c \neq 0$, alors il n'y a pas d'équilibre pour l'équation (3.10).
3. Si $1 + a_1 + a_2 + \dots + a_p = 0$ et $c = 0$ (l'équation est homogène), alors il y a une infinité d'équilibres, tout nombre réel étant un équilibre de l'équation (3.10).

On dit qu'un équilibre u_e de l'équation (3.10) est *globalement stable* si toute suite $(u_n)_{n \in \mathbb{N}}$ solution de (3.10) converge vers u_e lorsque $n \rightarrow +\infty$, quelles que soient les conditions initiales u_0, u_1, \dots, u_p de la solution.

Si un équilibre u_e n'est pas globalement stable, on appelle *ensemble de stabilité* de u_e , l'ensemble des conditions initiales assurant la convergence vers u_e lorsque $n \rightarrow +\infty$ des solutions correspondantes.

Enfin, comme un équilibre u_e est une solution particulière de (3.10), toute solution $(u_n)_{n \in \mathbb{N}}$ de (3.10) est de la forme $u_n = u_e + v_n$, où $(v_n)_{n \in \mathbb{N}}$ est une solution de l'équation homogène associée à (3.10). Nous avons le théorème suivant.

Théorème 3.1 *Un équilibre u_e de (3.10) est globalement stable si et seulement si toutes les racines (réelles ou complexes) de l'équation caractéristique associée à (3.10) sont de module inférieur à 1.*

Remarque. En général, si $(f_n)_{n \in \mathbb{N}}$ n'est pas constante, le problème d'étudier la stabilité d'une solution particulière non constante $(u_n^e)_{n \in \mathbb{N}}$ de (3.10) consiste à vérifier si $u_n - u_n^e$ converge vers 0 lorsque $n \rightarrow +\infty$. Les définitions d'équilibre globalement stable et d'ensemble de stabilité restent valables ainsi que le théorème précédent.

Équations de récurrence d'ordre 1. On considère l'équation $u_n = au_{n-1}$ qui admet la solution générale $u_n = u_0 a^n$, $u_0 \in \mathbb{R}$. Les éventuels équilibres sont les points fixes de l'équation, c'est-à-dire $u_e \in \mathbb{R}$ tels que $u_e = au_e$ et leur stabilité dépend de la valeur de a .

1. Si $a = 0$, l'équation admet seulement la solution identiquement nulle qui est aussi l'unique état stationnaire. De plus, il est globalement stable.

2. Si $a \in]0, 1[$, l'unique état stationnaire de l'équation est la suite identiquement nulle et, toute solution $(u_n)_{n \in \mathbb{N}}$ de l'équation converge vers 0 lorsque $n \rightarrow +\infty$, c'est-à-dire l'équilibre est globalement stable.
3. Si $a = 1$, toute solution de l'équation est une suite constante $u_n = u_0$, pour tout $n \in \mathbb{N}$, et toute solution de l'équation est un équilibre qui n'est pas stable.
4. Si $a \in]1, +\infty[$, l'équilibre de l'équation est à nouveau la suite identiquement nulle et toute autre solution explose : si $u_0 > 0$, la solution $(u_n)_{n \in \mathbb{N}}$ tend vers $+\infty$ lorsque $n \rightarrow +\infty$ et si $u_0 < 0$, la solution $(u_n)_{n \in \mathbb{N}}$ tend vers $-\infty$ lorsque $n \rightarrow +\infty$. En conclusion, l'état stationnaire n'est pas stable.
5. Si $a \in]-1, 0[$, l'équilibre de la solution est la suite identiquement nulle et toute solution $(u_n)_{n \in \mathbb{N}}$ converge vers 0 lorsque $n \rightarrow +\infty$, c'est-à-dire l'équilibre est globalement stable. De plus, la solution prend alternativement des valeurs positives et négatives.
6. Si $a \in]-\infty, -1[$, l'équilibre de la solution est la suite identiquement nulle et il n'est pas stable car toute solution de l'équation "explose" prenant alternativement des valeurs positives et négatives.
7. Si $a = -1$, l'équilibre de la solution est la suite identiquement nulle et il n'est pas stable car toute solution de l'équation prend alternativement les valeurs $-u_0$ et u_0 .

Quand le second membre de l'équation de l'ordre 1 est constant non nul, si $a \neq 0$, toute solution s'écrit $u_n = ca^n + \alpha$ où α est l'état stationnaire et c est une constante qui dépend de la condition initiale. Dans ce cas, α est globalement stable si et seulement si $|a| < 1$. Par contre, si $a = 1$ alors l'équation non homogène n'admet pas d'états stationnaires.

3.4 Équilibre et stabilité pour un système de récurrence d'ordre 1

Dans le cas d'un système homogène d'ordre 1

$$U_n = AU_{n-1}$$

où A est une matrice carrée d'ordre p non nulle, les états stationnaires sont les points fixes, c'est-à-dire les vecteurs $U_e \in \mathbb{R}^p$ tels que $U_e = AU_e$. Donc, si $A - I_d$ est inversible, il y a un seul équilibre $U_e = \vec{0}$. Sinon, il y en a une infinité, tous les vecteurs de $\ker(A - I_d)$.

Supposons que $A - I_d$ est inversible. On dit $U_e = \vec{0}$ est *globalement stable* si la solution générale $U_n = A^n U_0$ tend vers $\vec{0}$ lorsque $n \rightarrow +\infty$, pour tout vecteur des conditions initiales U_0 . Si $U_e = \vec{0}$ n'est pas globalement stable, on appelle *ensemble de stabilité* de U_e , l'ensemble des conditions initiales assurant la convergence vers U_e lorsque $n \rightarrow +\infty$ des solutions correspondantes.

Si A est diagonalisable, la solution générale $(U_n)_{n \in \mathbb{N}}$ du système s'écrit

$$U_n = \sum_{i=1}^p \alpha_i \lambda_i^n P_i = \alpha_1 \lambda_1^n P_1 + \dots + \alpha_p \lambda_p^n P_p,$$

où les λ_i sont les valeurs propres de A et P_i sont des vecteurs propres correspondants et linéairement indépendants. Alors, $U_e = \vec{0}$ est globalement stable si et seulement si les valeurs propres de A sont de module strictement inférieur à 1. Si A n'est pas diagonalisable, la condition nécessaire et suffisante précédente reste encore valable. Par exemple, si A est une

matrice carrée d'ordre 2 qui admet une unique valeur propre λ et est non diagonalisable ($\dim \ker(A - \lambda I) = 1$), nous avons vu que A peut s'écrire sous la forme

$$A = P \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} P^{-1}$$

et que la solution générale du système est donnée par

$$U_n = (\tilde{c}_1 \lambda^n + \tilde{c}_2 n \lambda^{n-1}) P_1 + \tilde{c}_2 \lambda^n P_2, \quad \tilde{c}_1, \tilde{c}_2 \in \mathbb{R},$$

où P_1 et P_2 sont les deux colonnes de la matrice P . Donc, à nouveau, $U_e = \vec{0}$ est globalement stable si et seulement si la valeur propre λ de A est de module strictement inférieur à 1.

Dans le cas d'un système non homogène

$$U_n = AU_{n-1} + B, \quad (3.11)$$

où $B \in \mathbb{R}^p$, on appelle *état stationnaire* du système toute suite $(U_n)_{n \in \mathbb{N}}$ constante et solution de (3.11).

Soit $U_n = U_e$, pour tout $n \in \mathbb{N}$, un état stationnaire de (3.11). Le vecteur U_e doit vérifier

$$U_e = AU_e + B \iff (I_d - A)U_e = B.$$

Si la matrice $A - I_d$ est inversible, il existe un seul état stationnaire donné par

$$U_e = (I_d - A)^{-1} B.$$

Si la matrice $A - I_d$ n'est pas inversible, il y a une infinité états stationnaires.

Supposons que $A - I_d$ est inversible. On dit que l'état stationnaire U_e est *globalement stable* si la solution générale U_n de (3.11) tend vers U_e lorsque $n \rightarrow +\infty$, pour tout vecteur des conditions initiales U_0 . Si U_e n'est pas globalement stable, on appelle *ensemble de stabilité* de U_e , l'ensemble des conditions initiales assurant la convergence vers U_e lorsque $n \rightarrow +\infty$ des solutions correspondantes.

On remarque que si $(U_n)_{n \in \mathbb{N}}$ est solution de (3.11), alors $Z_n = U_n - U_e$ est solution du système homogène associé $Z_n = AZ_{n-1}$. L'état stationnaire U_e de (3.11) est alors globalement stable si et seulement si l'équilibre $Z_e = \vec{0}$ du système homogène associé est globalement stable, donc si et seulement si les valeurs propres de A sont de module strictement inférieur à 1.

Exemple. On considère le système homogène d'ordre 1

$$U_n = AU_{n-1} \quad \text{avec} \quad A = \begin{pmatrix} \alpha & 1/2 \\ 1/2 & \alpha \end{pmatrix}.$$

Les valeurs propres de A sont $\lambda_1 = \alpha - \frac{1}{2}$ et $\lambda_2 = \alpha + \frac{1}{2}$ de vecteurs propres correspondants, respectivement, $P_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $P_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. La matrice de changement de coordonnées est

$$P = (P_1, P_2) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{et son inverse} \quad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

L'équilibre $U_e = \vec{0}$ est globalement stable si et seulement si $|\lambda_1| < 1$ et $|\lambda_2| < 1$, c'est-à-dire $-\frac{1}{2} < \alpha < \frac{1}{2}$. En effet, dans la base de vecteurs propres, le système s'écrit

$$\tilde{U}_n = \begin{pmatrix} \alpha - 1/2 & 0 \\ 0 & \alpha + 1/2 \end{pmatrix} \tilde{U}_{n-1},$$

avec $\tilde{U}_n = P^{-1}U_n$ et la solution générale est

$$\tilde{U}_n = \begin{pmatrix} (\alpha - 1/2)^n & 0 \\ 0 & (\alpha + 1/2)^n \end{pmatrix} \tilde{U}_0.$$

Donc, la solution générale du système du départ est

$$\begin{aligned} U_n &= P\tilde{U}_n = P \begin{pmatrix} (\alpha - 1/2)^n & 0 \\ 0 & (\alpha + 1/2)^n \end{pmatrix} \tilde{U}_0 \\ &= P \begin{pmatrix} (\alpha - 1/2)^n & 0 \\ 0 & (\alpha + 1/2)^n \end{pmatrix} P^{-1}U_0 \\ &= c_1(\alpha - \frac{1}{2})^n P_1 + c_2(\alpha + \frac{1}{2})^n P_2, \end{aligned}$$

où $(c_1, c_2)^t = P^{-1}U_0$. On voit bien alors que $U_n \rightarrow \vec{0}$ si et seulement si $|\alpha - \frac{1}{2}| < 1$ et $|\alpha + \frac{1}{2}| < 1$.

On considère maintenant le cas particulier d'une donnée initiale valeur propre : $U_0 = cP_1$, $c \in \mathbb{R}$. La solution correspondante est

$$U_n = c(\alpha - 1/2)^n P_1 = (\alpha - 1/2)^n U_0.$$

De même, si $U_0 = cP_2$ avec $c \in \mathbb{R}$, la solution $(U_n)_{n \in \mathbb{N}}$ correspondante est donnée par $U_n = (\alpha + 1/2)^n U_0$. En conclusion, si l'on part d'un vecteur propre la solution reste sur la droite déterminée par la valeur propre.

Si $U_0 = cP_1$, la solution $(U_n)_{n \in \mathbb{N}}$ tend vers $\vec{0}$ lorsque $n \rightarrow +\infty$ si et seulement si $|\alpha - 1/2| < 1$, tandis que si $U_0 = cP_2$, la solution $(U_n)_{n \in \mathbb{N}}$ tend vers $\vec{0}$ lorsque $n \rightarrow +\infty$ si et seulement si $|\alpha + 1/2| < 1$.

Par exemple, si $\alpha = 1$, le système est

$$U_n = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} U_{n-1},$$

et dans la base de valeurs propres on a

$$\tilde{U}_n = \begin{pmatrix} (1/2)^n & 0 \\ 0 & (3/2)^n \end{pmatrix} \tilde{U}_0.$$

Si l'on part d'un point initial sur la droite $y = -x$ déterminée par le vecteur P_1 , on reste sur cette droite et la solution $(U_n)_{n \in \mathbb{N}}$ tend vers $\vec{0}$ lorsque $n \rightarrow +\infty$. Mais si l'on part d'un point initial sur la droite $y = -x$ déterminée par le vecteur P_2 , on reste sur cette droite et la solution $(U_n)_{n \in \mathbb{N}}$ tend vers $+\infty$ (en norme) lorsque $n \rightarrow +\infty$.

4 Équations différentielles et systèmes différentiels linéaires

La théorie des équations différentielles ordinaires constitue l'un des outils fondamentaux des mathématiques. Elle permet de décrire des phénomènes d'évolution déterminants qui relèvent de l'économie, de la mécanique, de la physique et de la biologie.

L'objectif de ce chapitre est d'exposer des notions de base et d'apprendre à résoudre certains types d'équations et systèmes différentiels.

4.1 Équations différentielles du premier ordre

D'abord, on se propose d'étudier un certain nombre de types classiques d'équations différentielles du premier ordre pour lesquelles on sait ramener le calcul des solutions à des calculs de primitives.

On considère $U \subset \mathbb{R}^2$ et $f : U \rightarrow \mathbb{R}$. On appelle équation différentielle du premier ordre (résolue par rapport à la fonction inconnue y de la variable x) une équation de la forme

$$y'(x) = f(x, y(x)). \quad (4.1)$$

Une solution de (4.1) est une fonction dérivable y d'un intervalle I de \mathbb{R} à valeurs réelles telle que

$$(x, y(x)) \in U \text{ et } y'(x) = f(x, y(x)), \text{ pour tout } x \in I.$$

Résoudre (ou intégrer) l'équation différentielle (4.1) consiste à déterminer toutes les solutions de (4.1), c'est-à-dire la *solution générale*.

L'équation (4.1) avec la donnée initiale (x_0, y_0) , c'est-à-dire le problème

$$\begin{cases} y'(x) = f(x, y(x)) \\ y(x_0) = y_0 \end{cases} \quad (4.2)$$

est appelé *problème de Cauchy*.

Exemple : Le problème de Cauchy le plus simple est le suivant

$$\begin{cases} y'(x) = f(x) \\ y(x_0) = y_0 \end{cases}$$

Ce problème admet une solution et une seule donnée par

$$y(x) = y_0 + \int_{x_0}^x f(t) dt.$$

Pour $f(x) = \frac{1}{x}$, on a le problème de Cauchy suivant

$$\begin{cases} y'(x) = \frac{1}{x} \\ y(1) = 0 \end{cases}$$

dont l'unique solution est $y(x) = \ln x$, $x \in \mathbb{R}_+^* =]0, +\infty[$.

4.1.1 Équations différentielles linéaires du premier ordre

On appelle *équation différentielle linéaire du premier ordre*, une équation de la forme

$$a(x)y' + b(x)y = c(x) \quad (4.3)$$

où $a, b, c : I \rightarrow \mathbb{R}$ sont des fonctions continues sur un intervalle I de \mathbb{R} . On appelle a et b les *coefficients de l'équation* et c le *second membre*.

Si $c = 0$, l'équation est dite *sans second membre ou homogène*.

On considère l'*équation homogène (sans second membre)* associée à (4.3) :

$$a(x)y' + b(x)y = 0. \quad (4.4)$$

La solution générale de (4.3) s'écrit :

$$y = y_p + y_h,$$

où y_p est une solution particulière de (4.3) et où y_h est la solution générale de (4.4). On exprime ce résultat en disant que la solution générale de (4.3) s'obtient en additionnant une solution particulière de (4.3) avec la solution générale de l'équation homogène associée (4.4).

On suppose dans la suite que $a(x) \neq 0$ pour tout $x \in I$. Alors, la solution générale de (4.4) est donnée par

$$y(x) = Ke^{-\int \frac{b(x)}{a(x)} dx}, \quad K \in \mathbb{R}.$$

En effet, (4.4) est équivalente à

$$\frac{y'(x)}{y(x)} = -\frac{b(x)}{a(x)}, \quad \forall x \in I.$$

Alors,

$$\begin{aligned} \frac{d}{dx}(\ln(|y(x)|)) &= -\frac{b(x)}{a(x)}, \quad \forall x \in I \iff \ln(|y(x)|) = -\int \frac{b(x)}{a(x)} + C, \quad \forall x \in I, \quad C \in \mathbb{R} \\ &\iff |y(x)| = e^{-\int \frac{b(x)}{a(x)} dx + C}, \quad \forall x \in I, \quad C \in \mathbb{R}, \end{aligned}$$

où $\int \frac{b(x)}{a(x)} dx$ indique une primitive de $\frac{b(x)}{a(x)}$. D'où,

$$y(x) = Ke^{-\int \frac{b(x)}{a(x)} dx}, \quad K \in \mathbb{R}.$$

Exemple 1. L'équation linéaire du premier ordre : $y'(x) = 3y(x)$ admet comme solution générale sur \mathbb{R} :

$$y(x) = Ke^{\int 3 dx} = Ke^{3x}, \quad K \in \mathbb{R}.$$

Exemple 2. L'équation linéaire du premier ordre : $2y'(x) + \frac{1}{\sqrt{x}}y(x) = 0$ admet comme solution générale sur $]0, +\infty[$:

$$y(x) = Ke^{-\int \frac{1}{2\sqrt{x}} dx} = Ke^{-\sqrt{x}}, \quad K \in \mathbb{R}.$$

Exemple 3. L'équation linéaire du premier ordre : $xy'(x) + 2y(x) = 0$ admet comme solution générale sur $]0, +\infty[$:

$$y(x) = Ke^{-\int \frac{2}{x} dx} = Ke^{-2 \ln |x|} = \frac{K}{x^2}, \quad K \in \mathbb{R}.$$

Pour déterminer une solution particulière de (4.3), on écrit

$$y'(x) + \frac{b(x)}{a(x)}y(x) = \frac{c(x)}{a(x)}$$

et on multiplie l'équation par $e^{\int \frac{b(x)}{a(x)} dx}$ pour obtenir

$$\frac{d}{dx} \left(e^{\int \frac{b(x)}{a(x)} dx} y(x) \right) = \frac{c(x)}{a(x)} e^{\int \frac{b(x)}{a(x)} dx}.$$

Ainsi,

$$e^{\int \frac{b(x)}{a(x)} dx} y(x) = \int \frac{c(x)}{a(x)} e^{\int \frac{b(x)}{a(x)} dx} dx \iff y(x) = e^{-\int \frac{b(x)}{a(x)} dx} \int \frac{c(x)}{a(x)} e^{\int \frac{b(x)}{a(x)} dx} dx.$$

On conclut, finalement, que la solution générale de (4.3) est

$$y(x) = Ke^{-\int \frac{b(x)}{a(x)} dx} + e^{-\int \frac{b(x)}{a(x)} dx} \int \frac{c(x)}{a(x)} e^{\int \frac{b(x)}{a(x)} dx} dx, \quad K \in \mathbb{R}. \quad (4.5)$$

Dans le cas des coefficients constants, c'est-à-dire $a(x) = a$ et $b(x) = b$ pour tout $x \in I$, avec $a, b \in \mathbb{R}$, la solution générale (4.5) devient

$$y(x) = Ke^{-\frac{b}{a}x} + \frac{1}{a}e^{-\frac{b}{a}x} \int c(x)e^{\frac{b}{a}x} dx, \quad K \in \mathbb{R}.$$

Remarque : On détermine (4.5) en utilisant la *méthode de la variation de la constante*.

Méthode de la variation de la constante :

On cherche la solution générale de (4.3) de la forme

$$y(x) = K(x)e^{-\int \frac{b(x)}{a(x)} dx}.$$

Alors $K(x)$ satisfait :

$$K'(x) = \frac{c(x)}{a(x)} e^{\int \frac{b(x)}{a(x)} dx}.$$

La solution générale de (4.3) est donnée par (4.5).

Exemple : La solution générale de l'équation linéaire du premier ordre :

$$xy'(x) + 2y(x) + x = 0$$

est

$$y(x) = \frac{K}{x^2} - \frac{x}{3}, \quad K \in \mathbb{R}.$$

La solution du problème de Cauchy (4.3) et donnée initiale $y(x_0) = y_0$ est :

$$y(x) = e^{-\int_{x_0}^x \frac{b(t)}{a(t)} dt} \left(y_0 + \int_{x_0}^x \frac{c(t)}{a(t)} e^{\int_{x_0}^t \frac{b(s)}{a(s)} ds} dt \right).$$

4.1.2 Équations se ramenant à des équations linéaires

a) Équations de Bernoulli

Ce sont les équations de la forme

$$y'(x) = p(x)y + q(x)y^\alpha, \quad \alpha \in \mathbb{R} \setminus \{1\}, \quad (4.6)$$

avec $p, q : I \rightarrow \mathbb{R}$ continues. (Pour $\alpha = 1$, l'équation (4.6) est linéaire).

On se place dans le demi-plan supérieur $U = \{(x, y) \in \mathbb{R}^2 : y > 0\}$.

On pose $z(x) = y^{1-\alpha}(x)$. Alors,

$$(4.6) \iff \frac{1}{1-\alpha} z'(x) = p(x)z(x) + q(x).$$

On est donc ramené à une équation linéaire en z .

Exemple : La solution générale de l'équation de Bernoulli $y' + 2y - (x+2)\sqrt{y} = 0$ est

$$y(x) = \left(\frac{x+1}{2} + Ke^{-x}\right)^2, \quad K \in \mathbb{R},$$

pour $x \in I = \{x \in \mathbb{R} : \frac{x+1}{2} + Ke^{-x} > 0\}$.

b) Équations de Riccati

Ce sont les équations de la forme

$$y'(x) = a(x)y^2(x) + b(x)y(x) + c(x) \quad (4.7)$$

avec $a, b, c : I \rightarrow \mathbb{R}$ continues, c'est-à-dire $f(x, y)$ est un polynôme de degré ≤ 2 en y .

On sait résoudre (4.7) dès que l'on connaît une solution particulière y_p .

On pose $y = y_p + u$, c'est-à-dire $u = y - y_p$. On obtient

$$u'(x) = (2a(x)y_p(x) + b(x))u(x) + a(x)u^2(x).$$

C'est une équation de Bernoulli avec $\alpha = 2$. On la ramène à une équation linéaire en posant

$$z(x) = u^{1-\alpha}(x) = \frac{1}{u(x)}.$$

Remarquer que $y = y_p + \frac{1}{z}$ où z est solution d'une équation linéaire.

Exemple : L'équation de Riccati $y' + 4y + y^2 - 5 = 0$ admet comme solution particulière $y_p = 1$. On considère de changement de variable $y = 1 + \frac{1}{z}$. Alors, z est solution de l'équation linéaire

$$z' - 6z - 1 = 0.$$

La solution générale de cette équation linéaire est $z(x) = -\frac{1}{6} + Ke^{6x}$, $K \in \mathbb{R}$, $x \in \mathbb{R}$. Donc, la solution générale de l'équation de Riccati est

$$y(x) = \frac{\lambda e^{6x} + 5}{\lambda e^{6x} - 1}, \quad \lambda \in \mathbb{R},$$

$x \in I$ où l'intervalle I dépend de la constante λ .

4.2 Équations différentielles linéaires du second ordre

On appelle *équation différentielle linéaire du second ordre*, une équation de la forme

$$a(x)y'' + b(x)y' + c(x)y = d(x) \quad (4.8)$$

où $a, b, c, d : I \rightarrow \mathbb{R}$ sont des fonctions continues sur un intervalle I de \mathbb{R} . On appelle a, b et c les *coefficients de l'équation* et d le *second membre*.

Si $d = 0$, l'équation est dite *sans second membre ou homogène*.

Une solution de (4.8) est une fonction y de la variable x deux fois dérivable de I dans \mathbb{R} vérifiant, pour tout $x \in I$,

$$a(x)y''(x) + b(x)y'(x) + c(x)y(x) = d(x)$$

On considère l'équation homogène (sans second membre) associée à (4.8) :

$$a(x)y'' + b(x)y' + c(x)y = 0. \quad (4.9)$$

L'ensemble des solutions de (4.9) est un espace vectoriel de dimension 2. Pour déterminer la solution générale de (4.9), il suffit de déterminer deux solutions réelles linéairement indépendantes de (4.9) .

La solution générale de (4.8) s'écrit :

$$y = y_p + y_h,$$

où y_p est une solution particulière de (4.8) et où y_h est la solution générale de (4.9).

Le problème de Cauchy associé à (4.8) et la donnée initiale (x_0, y_0, y_1) est le problème

$$\begin{cases} a(x)y'' + b(x)y' + c(x)y = d(x) \\ y(x_0) = y_0, \quad y'(x_0) = y_1. \end{cases}$$

Pour une équation du second degré, la donnée initiale d'un problème de Cauchy revient à donner les valeurs de la fonction et de sa dérivée première en un point x_0 de I .

4.2.1 Coefficients constants

On s'intéresse aux équations du type (4.8) et (4.9) dont les coefficients sont constants, c'est-à-dire $a(x) = a$, $b(x) = b$ et $c(x) = c$, pour tout $x \in \mathbb{R}$, avec $a, b, c \in \mathbb{R}$ et $a \neq 0$. Alors, (4.8) et (4.9) s'écrivent

$$ay'' + by' + cy = d(x). \quad (4.10)$$

et

$$ay'' + by' + cy = 0. \quad (4.11)$$

On appelle *polynôme caractéristique* de l'équation (4.10) le polynôme :

$$P(\lambda) = a\lambda^2 + b\lambda + c, \quad (4.12)$$

et *équation caractéristique* de l'équation (4.10) :

$$P(\lambda) = a\lambda^2 + b\lambda + c = 0. \quad (4.13)$$

Il est facile de vérifier qu'une fonction du type $y(x) = e^{\alpha x}$ est solution de (4.10) si et seulement si α est solution de l'équation caractéristique (4.13). On en déduit les règles suivantes.

1. Si l'équation (4.13) admet deux solutions réelles distinctes λ_1 et λ_2 , alors la solution générale de (4.10) est

$$y(x) = Ae^{\lambda_1 x} + Be^{\lambda_2 x}, \quad A, B \in \mathbb{R}.$$

2. Si l'équation (4.13) admet une solution réelle double λ , alors la solution générale de (4.10) est

$$y(x) = Ae^{\lambda x} + Bxe^{\lambda x}, \quad A, B \in \mathbb{R}.$$

3. Si l'équation (4.13) admet une racine complexe non réelle $\lambda = \alpha + i\beta$ ($\beta \neq 0$), alors elle admet aussi la racine $\bar{\lambda}$ et la solution générale de (4.10) est

$$y(x) = Ae^{\alpha x} \cos(\beta x) + Be^{\alpha x} \sin(\beta x), \quad A, B \in \mathbb{R}.$$

Exemples :

1. La solution générale de $y'' - 4y = 0$ est

$$y(x) = Ae^{2x} + Be^{-2x}, \quad A, B \in \mathbb{R}.$$

2. La solution générale de $y'' - 2y' + y = 0$ est

$$y(x) = Ae^x + Bxe^x, \quad A, B \in \mathbb{R}.$$

3. La solution générale de $y'' + y = 0$ est

$$y(x) = A \cos(x) + B \sin(x), \quad A, B \in \mathbb{R}.$$

Revenons à l'équation avec second membre (4.10). Pour des seconds membres particuliers, nous pouvons appliquer la *méthode des coefficients indéterminés* pour déterminer une solution particulière de (4.10). Plus précisément, nous avons les règles suivantes.

1. Si $d(x) = e^{\alpha x} Q(x)$, où Q est un polynôme de degré k et $\alpha \in \mathbb{R}$, alors (4.10) admet une solution particulière de la forme

$$y_p(x) = e^{\alpha x} R(x)$$

avec R polynôme de degré m où

- (a) $m = k$, si α n'est pas racine de P ,
- (b) $m = k + r$, si α est racine de P de multiplicité r .

2. Si $d(x) = e^{\alpha x} \cos(\beta x) Q(x)$ ou $d(x) = e^{\alpha x} \sin(\beta x) Q(x)$ où Q est un polynôme de degré k et $\alpha \in \mathbb{R}$, alors (4.10) admet une solution particulière de la forme

$$y_p(x) = e^{\alpha x} (\cos(\beta x) R_1(x) + \sin(\beta x) R_2(x))$$

avec R_1 et R_2 polynômes de degré m où

- (a) $m = k$, si $\alpha + i\beta$ n'est pas racine de P ,
- (b) $m = k + 1$, si $\alpha + i\beta$ est racine de P .

3. Si $d(x)$ est la somme d'expressions du type précédent, on cherche une solution pour chacune des expressions et on fait la somme.

Exemple. On considère l'équation du second degré à coefficients constants non homogène $y'' - 2y' + y = x$. On a $b(x) = x$ et $\alpha = 0$ n'est pas racine du polynôme caractéristique $P(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Alors, on cherche une solution particulière de la forme $y_p(x) = ax + b$. On obtient $a = 1$ et $b = 2$. Donc, la solution générale de cette équation est

$$y(x) = x + 2 + Ae^x + Bxe^x, \quad A, B \in \mathbb{R}.$$

Remarque. Pour résoudre l'équation non homogène (4.10), on peut aussi utiliser la *méthode de la variation des constantes*. Supposons que la solution générale de l'équation homogène (4.11) est

$$y_h(x) = A\varphi_1(x) + B\varphi_2(x), \quad A, B \in \mathbb{R}.$$

Alors, la solution générale de l'équation non homogène (4.10) s'écrit

$$y_h(x) = A(x)\varphi_1(x) + B(x)\varphi_2(x),$$

avec

$$\begin{cases} A'(x)\varphi_1(x) + B'(x)\varphi_2(x) = 0 \\ A'(x)\varphi_1'(x) + B'(x)\varphi_2'(x) = \frac{d(x)}{a}. \end{cases}$$

4.3 Équations différentielles linéaires d'ordre p

On appelle *équation différentielle linéaire d'ordre p* , une équation de la forme

$$a_p(x)y^{(p)} + a_{p-1}(x)y^{(p-1)} + \dots + a_1(x)y' + a_0(x)y = d(x) \quad (4.14)$$

où $a_j : I \rightarrow \mathbb{R}$, $0 \leq j \leq p$, et d sont des fonctions continues sur un intervalle I de \mathbb{R} . On appelle a_j , $0 \leq j \leq p$, les *coefficients de l'équation* et d le *second membre*.

Si $f = 0$, l'équation est dite *sans second membre ou homogène*.

Une solution de (4.14) est une fonction y de la variable x , p fois dérivable de I dans \mathbb{R} vérifiant, pour tout $x \in I$,

$$a_p(x)y^{(p)}(x) + a_{p-1}(x)y^{(p-1)}(x) + \dots + a_1(x)y'(x) + a_0(x)y(x) = d(x)$$

On considère l'équation homogène (sans second membre) associée à (4.14) :

$$a_p(x)y^{(p)} + a_{p-1}(x)y^{(p-1)} + \dots + a_1(x)y' + a_0(x)y = 0. \quad (4.15)$$

L'ensemble des solutions de (4.15) est un espace vectoriel de dimension p . Pour déterminer la solution générale de (4.15), il suffit de trouver p solutions réelles linéairement indépendantes de (4.15).

La solution générale de (4.14) s'écrit :

$$y = y_p + y_h,$$

où y_p est une solution particulière de (4.14) et où y_h est la solution générale de (4.15).

Le problème de Cauchy associé à (4.14) et la donnée initiale $(x_0, y_0, \dots, y_{p-1})$ est le problème

$$\begin{cases} a_p(x)y^{(p)} + a_{p-1}(x)y^{(p-1)} + \dots + a_1(x)y' + a_0(x)y = d(x) \\ y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(p-1)}(x_0) = y_{p-1}. \end{cases}$$

4.3.1 Coefficients constants

On s'intéresse aux équations du type (4.14) et (4.15) dont les coefficients sont constants, c'est-à-dire $a_j(x) = a_j$, pour tout $x \in \mathbb{R}$, avec $a_j \in \mathbb{R}$ et $a_p \neq 0$, $0 \leq j \leq p$. Alors, (4.14) et (4.15) s'écrivent

$$a_p y^{(p)} + a_{p-1} y^{(p-1)} + \dots + a_1 y' + a_0 y = d(x) \quad (4.16)$$

et

$$a_p y^{(p)} + a_{p-1} y^{(p-1)} + \dots + a_1 y' + a_0 y = 0. \quad (4.17)$$

On appelle *polynôme caractéristique* de l'équation (4.16) le polynôme :

$$P(\lambda) = a_p \lambda^p + a_{p-1} \lambda^{(p-1)} + \dots + a_1 \lambda + a_0, \quad (4.18)$$

et *équation caractéristique* de l'équation (4.16) :

$$P(\lambda) = a_p \lambda^p + a_{p-1} \lambda^{(p-1)} + \dots + a_1 \lambda + a_0 = 0. \quad (4.19)$$

Pour déterminer la solution générale de (4.17), nous avons les règles suivantes.

1. Si les p racines de l'équation (4.19) sont toutes distinctes et réelles, alors la solution générale de (4.17) est

$$y(x) = A_1 e^{\lambda_1 x} + \dots + A_p e^{\lambda_p x}, \quad A_j \in \mathbb{R}, \quad 1 \leq j \leq p.$$

2. Si λ est une racine réelle de (4.19) de multiplicité k , on lui associe les k solutions réelles et linéairement indépendantes de (4.17)

$$y_j(x) = x^j e^{\lambda x}, \quad 0 \leq j \leq k - 1.$$

3. Si $\lambda = \alpha + i\beta$ et $\bar{\lambda} = \alpha - i\beta$ sont deux racines complexes conjuguées de (4.19) de multiplicité k , on leur associe les $2k$ solutions réelles et linéairement indépendantes de (4.17)

$$y_j(x) = x^j e^{\alpha x} \cos(\beta x), \quad z_j(x) = x^j e^{\alpha x} \sin(\beta x), \quad 0 \leq j \leq k - 1.$$

Pour déterminer une solution particulière de l'équation avec second membre (4.16), nous procédons comme pour les équations du second ordre. En particulier, pour des seconds membres particuliers, nous pouvons appliquer la *méthode des coefficients indéterminés* pour déterminer une solution particulière de (4.10). Plus précisément :

1. Si $d(x) = e^{\alpha x} Q(x)$, où Q est un polynôme de degré k et $\alpha \in \mathbb{R}$, alors (4.16) admet une solution particulière de la forme

$$y_p(x) = e^{\alpha x} R(x)$$

avec R polynôme de degré m où

(a) $m = k$, si α n'est pas racine de P ,

(b) $m = k + r$, si α est racine de P de multiplicité r .

2. Si $d(x) = e^{\alpha x} \cos(\beta x) Q(x)$ ou $d(x) = e^{\alpha x} \sin(\beta x) Q(x)$ où Q est un polynôme de degré k et $\alpha \in \mathbb{R}$, alors (4.10) admet une solution particulière de la forme

$$y_p(x) = e^{\alpha x} (\cos(\beta x) R_1(x) + \sin(\beta x) R_2(x))$$

avec R_1 et R_2 polynômes de degré m où

- (a) $m \leq k$, si $\alpha + i\beta$ n'est pas racine de P ,
 - (b) $m \leq k + r$, si $\alpha + i\beta$ est racine de P de multiplicité r .
3. Si $d(x)$ est la somme d'expressions du type précédent, on cherche une solution pour chacune de expressions et on fait la somme.

4.4 Systèmes différentiels linéaires du premier ordre

On appelle système différentiel linéaire du premier ordre dans \mathbb{R}^n un système de la forme

$$Y'(x) = A(x)Y(x) + B(x) \quad (4.20)$$

où A est une application continue de J intervalle ouvert de \mathbb{R} dans $M_n(\mathbb{R})$, l'espace des matrices carrées de dimension n à coefficients réels, c'est-à-dire

$$A(x) = (a_{ij}(x))_{1 \leq i, j \leq n},$$

$B(x) = (B_i(x))_{1 \leq i \leq n}$ est une application continue de J dans \mathbb{R}^n et $Y(x) = (y_i(x))_{1 \leq i \leq n}$ est le vecteur des n fonctions inconnues.

L'équation (4.20) est dite homogène ou sans second membre si $B(x) = 0$, pour tout $x \in J$.

Sous forme développée, (4.20) s'écrit :

$$\begin{cases} y_1'(x) = a_{11}(x)y_1(x) + \dots + a_{1n}(x)y_n(x) + b_1(x) \\ \vdots \\ y_n'(x) = a_{n1}(x)y_1(x) + \dots + a_{nn}(x)y_n(x) + b_n(x) \end{cases}$$

c'est-à-dire

$$y_i'(x) = \sum_{j=1}^n a_{ij}(x)y_j(x) + b_i(x), \quad \forall 1 \leq i \leq n.$$

Résoudre ou intégrer le système ci-dessus consiste à déterminer la *solution générale* $Y : J \rightarrow \mathbb{R}^n$, c'est-à-dire toutes les solutions. Le problème de déterminer la solution de (4.20) qui satisfait la donnée initiale $Y(x_0) = Y_0$, avec $(x_0, Y_0) \in J \times \mathbb{R}^n$ donné, s'appelle le problème de Cauchy associé à (x_0, Y_0) .

On considère le système homogène associé à (4.20) :

$$Y'(x) = A(x)Y(x). \quad (4.21)$$

Comme pour les équations linéaires d'ordre 1 et 2, la solution générale de (4.20) s'obtient en additionnant une solution particulière de (4.20) et la solution générale du système homogène associé (4.21). De plus, l'ensemble des solutions du système homogène (4.21) est un espace vectoriel de dimension n .

On dit qu'un ensemble de n fonctions vectorielles réelles $(Y_1(x), \dots, Y_n(x))$ est une *solution fondamentale* du système homogène (4.21) si c'est un ensemble de solutions linéairement indépendantes de (4.21). La solution générale de (4.21) s'écrit alors

$$Y(x) = \alpha_1 Y_1(x) + \dots + \alpha_n Y_n(x), \quad \alpha_j \in \mathbb{R}, 1 \leq j \leq n.$$

4.4.1 Cas des coefficients constants

On considère le système homogène

$$Y' = AY, \quad (4.22)$$

où $A \in M_n(\mathbb{R})$. Autrement dit, on considère un système de la forme (4.21) avec $J = \mathbb{R}$ et $A(x) = A$ pour tout $x \in J$. On peut montrer que si $\lambda \in \mathbb{R}$ est une valeur propre de A et V est un vecteur propre correspondant à λ , alors $Y(x) = e^{\lambda x}V$ est une solution de (4.22).

Rappels.

1. Soit $A \in M_n(\mathbb{R})$. On dit que $\lambda \in \mathbb{R}$ est une valeur propre de A s'il existe $V \in \mathbb{R}^n$, $V \neq 0$ tel que

$$AV = \lambda V.$$

On appelle V le vecteur propre de A correspondant à la valeur propre λ .

Autrement dit, λ est valeur propre de A si $\ker(A - \lambda I) \neq \{0\}$ où

$$\ker(A - \lambda I) = \{V \in \mathbb{R}^n : (A - \lambda I)V = 0\},$$

et I denote la matrice identité de $M_n(\mathbb{R})$. Si λ est une valeur propre de A , on appelle $\ker(A - \lambda I)$ l'espace propre de A associé à λ .

2. On appelle $P(\lambda) = \det(A - \lambda I)$ le polynôme caractéristique de A . Alors, λ est valeur propre de A si et seulement si λ est une racine de P .
3. Si A est diagonalisable sur \mathbb{R} , il existe une matrice $P \in M_n(\mathbb{R})$ inversible et D une matrice diagonale telles que $A = PDP^{-1}$. De plus,

$$D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}, \quad P = (V_1 \quad \dots \quad V_n),$$

où $\lambda_1, \dots, \lambda_n$ sont les valeurs propres réelles de A et V_j est un vecteur propre de A associé à λ_j , $1 \leq j \leq n$.

Cas simple : A est diagonalisable.

Il existe alors une base (V_1, \dots, V_n) de \mathbb{R}^n constituée de vecteurs propres de A , de valeurs propres correspondantes $\lambda_1, \dots, \lambda_n$. On obtient donc n solutions linéairement indépendantes

$$Y_j(x) = e^{\lambda_j x}V_j, \quad 1 \leq j \leq n.$$

La solution générale est alors donnée par

$$Y(x) = \alpha_1 e^{\lambda_1 x}V_1 + \dots + \alpha_n e^{\lambda_n x}V_n, \quad \alpha_j \in \mathbb{R}.$$

Cas de la dimension 2. On suppose $n = 2$.

On résout le système $Y' = AY$ de la façon suivante :

1. Si la matrice A admet deux valeurs propres réelles distinctes λ_1 et λ_2 de vecteurs propres correspondants V_1 et V_2 , alors A est diagonalisable dans \mathbb{R}^2 et la solution générale de (4.22) est

$$Y(x) = \alpha_1 e^{\lambda_1 x} V_1 + \alpha_2 e^{\lambda_2 x} V_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

2. Si la matrice A admet une valeur propre réelle double λ et $\dim \ker(A - \lambda I) = 2$, alors A est diagonalisable dans \mathbb{R}^2 . Soient V_1 et V_2 deux vecteurs linéairement indépendants de $\ker(A - \lambda I)$. La solution générale de (4.22) est donnée par

$$Y(x) = \alpha_1 e^{\lambda x} V_1 + \alpha_2 e^{\lambda x} V_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

3. Si la matrice A admet une valeur propre réelle double λ et $\dim \ker(A - \lambda I) = 1$, alors A n'est pas diagonalisable dans \mathbb{R}^2 . On prend $V_1 \in \ker(A - \lambda I)$ (c'est-à-dire, V_1 est un vecteur propre correspondant à λ) et on prend $V_2 \in \ker((A - \lambda I)^2)$ tel que (V_1, V_2) est une base de $\ker((A - \lambda I)^2)$. Alors, la solution générale de (4.22) est

$$Y(x) = \alpha_1 e^{\lambda x} V_1 + \alpha_2 e^{\lambda x} (I + x(A - \lambda I)) V_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

4. Si la matrice A admet deux valeurs propres complexes conjuguées $\lambda = \alpha + i\beta$ et $\bar{\lambda} = \alpha - i\beta$, A est diagonalisable dans \mathbb{C}^2 . Soit $V \in \mathbb{C}^2$ un vecteur propre correspondant à λ . Alors, la solution générale de (4.22) est

$$Y(x) = \alpha_1 \operatorname{Re}(e^{\lambda x} V) + \alpha_2 \operatorname{Im}(e^{\lambda x} V), \quad \alpha_1, \alpha_2 \in \mathbb{R}.$$

Exemples.

1. On considère le système

$$\begin{cases} x'(t) = 5x(t) \\ y'(t) = 3y(t) \end{cases}$$

Il peut s'écrire sous la forme : $Y' = AY$ avec

$$A = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix},$$

matrice diagonale. On peut alors résoudre séparément chacune des équations du système et on obtient $x(t) = \alpha e^{5t}$ et $y(t) = \beta e^{3t}$ avec $\alpha, \beta \in \mathbb{R}$.

2. On considère le système

$$\begin{cases} x'(t) = 5x(t) - 3y(t) \\ y'(t) = 2x(t) - 2y(t) \end{cases}$$

Les valeurs propres de la matrice A associée au système sont les racines de $\lambda^2 - 3\lambda - 4 = 0$, soit 4 et -1 . La matrice A est diagonalisable. Un vecteur propre associé à 4 est $(3, 1)$ et un vecteur propre associé à -1 est $(1, 2)$. La solution générale du système est

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha e^{4t} \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \beta e^{-t} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3\alpha e^{4t} + \beta e^{-t} \\ \alpha e^{4t} + 2\beta e^{-t} \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{R}$.

3. On considère le système

$$\begin{cases} x'(t) = 5x(t) - 3y(t) \\ y'(t) = 5y(t) \end{cases}$$

On peut résoudre séparément chacune des équations car la matrice A du système est triangulaire. On a $y(t) = \beta e^{5t}$ et x satisfait une équation linéaire non-homogène du premier ordre : $x'(t) = 5x(t) - 3\beta e^{5t}$. Donc $x(t) = \alpha e^{5t} - 3\beta t e^{5t}$, $\alpha, \beta \in \mathbb{R}$.

On peut aussi appliquer les résultats généraux dans le cas d'une matrice non diagonalisable (valeur propre $\lambda = 5$ de multiplicité 2 et sous-espace propre associé de dimension 1).

4. On considère le système

$$\begin{cases} x'(t) = y(t) \\ y'(t) = -x(t) \end{cases}$$

Les valeurs propres de la matrice A associée au système sont les racines de $\lambda^2 + 1 = 0$, donc $\lambda = \pm i$. Un vecteur propre de A associé à i est $(1, i)$. La solution générale est

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha \operatorname{Re} \left(e^{it} \begin{pmatrix} 1 \\ i \end{pmatrix} \right) + \beta \operatorname{Im} \left(e^{it} \begin{pmatrix} 1 \\ i \end{pmatrix} \right) = \begin{pmatrix} \alpha \cos t + \beta \sin t \\ -\alpha \sin t + \beta \cos t \end{pmatrix},$$

avec $\alpha, \beta \in \mathbb{R}$.

Cas général. On note P le polynôme caractéristique de A et on considère ses racines distinctes dans \mathbb{C} : $\lambda_1, \dots, \lambda_s$, d'ordre de multiplicité respectives r_1, \dots, r_s .

On obtient une base de l'espace des solutions de (4.22) dans \mathbb{R}^n de la façon suivante :

1. Si λ est une racine réelle de P de multiplicité r , on lui associe les r solutions :

$$Y_j(x) = e^{\lambda x} \sum_{k=0}^{r-1} \frac{x^k}{k!} (A - \lambda I)^k V_j, \quad 1 \leq j \leq r,$$

où (V_1, \dots, V_r) est une base de $\ker((A - \lambda I)^r)$.

2. Si $\mu, \bar{\mu}$ est un couple de racines complexes conjuguées de partie imaginaire non nulle et de multiplicité r , on lui associe $2r$ solutions :

$$Y_j(x) = \operatorname{Re} \left(e^{\mu x} \sum_{k=0}^{r-1} \frac{x^k}{k!} (A - \mu I)^k V_j \right) \quad \text{et} \quad Z_j(x) = \operatorname{Im} \left(e^{\mu x} \sum_{k=0}^{r-1} \frac{x^k}{k!} (A - \mu I)^k V_j \right),$$

$1 \leq j \leq r$, où (V_1, \dots, V_r) est une base de $\ker((A - \mu I)^r)$ (sous-espace vectoriel à valeurs dans \mathbb{C}^n).

4.4.2 Équations différentielles d'ordre p

Une équation différentielle linéaire d'ordre p , $p \geq 2$, peut s'écrire comme un système linéaire du premier ordre dans \mathbb{R}^p . On considère l'équation différentielle linéaire d'ordre p (4.14) :

$$a_p(x)y^{(p)} + a_{p-1}(x)y^{(p-1)} + \dots + a_1(x)y' + a_0(x)y = f(x).$$

Supposons $a_p(x) \neq 0$, pour tout $x \in J$. Alors, on a, pour tout $x \in J$,

$$y^{(p)}(x) + \frac{a_{n-1}(x)}{a_p(x)}y^{(p-1)}(x) + \dots + \frac{a_1(x)}{a_p(x)}y'(x) + \frac{a_0(x)}{a_p(x)}y(x) = \frac{f(x)}{a_p(x)}.$$

On associe à (4.14) le système d'ordre 1 dans \mathbb{R}^p

$$Y'(x) = A(x)Y + B(x), \quad (4.23)$$

où $A(x) \in M_p(\mathbb{R})$ et $B(x) \in \mathbb{R}^p$ se représentent de la façon suivante :

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\frac{a_0(x)}{a_p(x)} & -\frac{a_1(x)}{a_p(x)} & -\frac{a_2(x)}{a_p(x)} & \dots & -\frac{a_{n-1}(x)}{a_p(x)} \end{pmatrix}, \quad B(x) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{f(x)}{a_p(x)} \end{pmatrix}.$$

L'application qui à y solution de (4.14) associe $Y(x) = \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix}$, réalise une bijection de

l'ensemble de solutions de (4.14) sur l'ensemble de solutions de (4.23), la bijection réciproque étant l'application $Y(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_p(x) \end{pmatrix} \mapsto y(x) = y_1(x)$. Ces bijections sont linéaires si $f(x) = 0$ (cas homogène).

Exemple. Résoudre l'équation d'ordre 3 : $y''' - y'' - y' + y = e^{-x}$ équivaut à résoudre le système d'ordre 1 : $Y' = AY + B(x)$ avec

$$A(x) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}, \quad B(x) = \begin{pmatrix} 0 \\ 0 \\ e^{-x} \end{pmatrix}.$$

4.5 Équilibre et stabilité

On considère le système différentiel autonome du premier ordre n équations

$$\begin{cases} y_1'(x) = f_1(y_1(x), \dots, y_n(x)) \\ \vdots \\ y_n'(x) = f_n(y_1(x), \dots, y_n(x)) \end{cases} \quad (4.24)$$

ou, en notation vectorielle, $Y'(x) = F(Y(x))$.

On appelle *état stationnaire* du système (4.24) toute solution constante de (4.24)

$$Y(x) \equiv Y^e = (y_1^e, \dots, y_n^e) \in \mathbb{R}^n.$$

On remarque que $Y^e \in \mathbb{R}^n$ est un état stationnaire du système (4.24) si et seulement si

$$F(Y^e) = 0 \iff \begin{cases} f_1(y_1^e, \dots, y_n^e) = 0 \\ \vdots \\ f_n(y_1^e, \dots, y_n^e) = 0 \end{cases}$$

On dit qu'un état stationnaire est un *équilibre globalement (asymptotiquement) stable* si toute solution $Y(x)$ de (4.24) tend vers Y^e lorsque x tend vers $+\infty$, c'est-à-dire lorsque $\lim_{x \rightarrow +\infty} y_i(x) = y_i^e$ pour tout $i = 1, \dots, n$, quel que soit la condition initiale (x_0, Y_0) de la solution.

Si un état stationnaire $Y^e \in \mathbb{R}^n$ de (4.24) n'est pas globalement (asymptotiquement) stable, alors on appelle *ensemble de stabilité* de Y^e l'ensemble des conditions initiales $(x_0, Y_0) \in \mathbb{R}^{n+1}$ assurant la convergence vers Y^e lorsque $x \rightarrow +\infty$ des solutions associées. Si l'ensemble de stabilité de Y^e n'est pas vide, alors on dit que l'équilibre Y^e est *localement (asymptotiquement) stable*.

On peut observer aussi des cas de systèmes ayant un état stationnaire Y^e et dont la solution générale ne tend pas vers Y^e lorsque x tend vers $+\infty$ ni s'éloigne de Y^e . On dit alors que Y^e est un *équilibre neutre*.

Enfin, si un état stationnaire $Y^e \in \mathbb{R}^n$ de (4.24) n'est ni globalement (asymptotiquement) stable ni localement (asymptotiquement) stable ni neutre, on dit qu'il est un état stationnaire *instable*.

Remarque. Si le système (4.24) se réduit à une équation $y' = f(y)$, alors les éventuels états stationnaires sont les zéros de la fonction f . La nature de la stabilité des états stationnaires dépend alors du signe de f , de la dérivée première f' et éventuellement aussi du signe de la dérivée seconde f'' , dans un voisinage des états stationnaires.

Si le système (4.24) est linéaire, c'est-à-dire $Y' = AY$, alors si A est inversible, il n'y a qu'un état stationnaire : $Y^e = \vec{0} \in \mathbb{R}^n$ et la nature de la stabilité dépend uniquement des valeurs propres de la matrice A . Si A n'est pas inversible, il y a une infinité d'états stationnaires donnés par le noyau de A .

Supposons $n = 2$ et A inversible. On note λ_1 et λ_2 les deux valeurs propres de la matrice A . On a les cas suivants :

1. Si les valeurs propres sont réelles (distinctes ou égales), alors $Y^e = \vec{0}$ est globalement stable si et seulement si λ_1 et λ_2 sont négatives ; si par contre, λ_1 ou λ_2 est positive alors Y^e est instable ou localement stable.
2. Si les valeurs propres λ_1 et λ_2 sont complexes et conjuguées, alors $Y^e = \vec{0}$ est globalement stable si et seulement si $\text{Re}(\lambda_1) = \text{Re}(\lambda_2) < 0$; si $\text{Re}(\lambda_1) = \text{Re}(\lambda_2) > 0$, alors $Y^e = \vec{0}$ est instable.
3. Si les valeurs propres λ_1 et λ_2 sont complexes et conjuguées et $\text{Re}(\lambda_1) = \text{Re}(\lambda_2) = 0$, alors $Y^e = \vec{0}$ est neutre.

Exemple 1 (matrice à valeurs propres réelles et distinctes). On considère le système différentiel

$$\begin{cases} x'(t) = \alpha x(t) + \frac{1}{2}y(t) \\ y'(t) = \frac{1}{2}x(t) + \alpha y(t) \end{cases}$$

où $\alpha \in \mathbb{R}$. Les valeurs propres de la matrice associée à ce système sont : $\lambda_1 = \alpha - \frac{1}{2}$ et $\lambda_2 = \alpha + \frac{1}{2}$, les vecteurs propres correspondants étant : $v_1 = (1, -1)$ et $v_2 = (1, 1)$. La solution générale du système est alors

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \alpha_1 e^{(\alpha - \frac{1}{2})t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \alpha_2 e^{(\alpha + \frac{1}{2})t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 e^{(\alpha - \frac{1}{2})t} + \alpha_2 e^{(\alpha + \frac{1}{2})t} \\ -\alpha_1 e^{(\alpha - \frac{1}{2})t} + \alpha_2 e^{(\alpha + \frac{1}{2})t} \end{pmatrix},$$

avec $\alpha_1, \alpha_2 \in \mathbb{R}$. Le point d'équilibre $Y^e = (0, 0)$ est globalement (asymptotiquement) stable si et seulement si $\lambda_1 < 0$ et $\lambda_2 < 0$, c'est-à-dire $\alpha < -\frac{1}{2}$. Remarquons que

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Alors, si la condition initiale est un vecteur propre associé à λ_1 :

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \beta v_1,$$

avec $\beta \in \mathbb{R}$, on a $\alpha_2 = 0$ et la solution correspondante est

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \beta e^{(\alpha - \frac{1}{2})t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

De même, si la condition initiale est un vecteur propre associé à λ_2 :

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \beta v_2,$$

avec $\beta \in \mathbb{R}$, on a $\alpha_1 = 0$ et la solution correspondante est

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \beta e^{(\alpha + \frac{1}{2})t} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

On conclut que si l'on part d'un vecteur propre, la solution reste sur la droite déterminée par la direction du vecteur propre.

Dans le cas $\alpha = 1$, le système est

$$\begin{cases} x'(t) = x(t) + \frac{1}{2}y(t) \\ y'(t) = \frac{1}{2}x(t) + y(t) \end{cases}$$

et l'équilibre $Y^e = (0, 0)$ est instable car les deux valeurs propres sont positives ($\lambda_1 = \frac{1}{2}$ et $\lambda_2 = \frac{3}{2}$). De plus, $x(t)$ est croissante dans la région $x + \frac{1}{2}y \geq 0$ et $y(t)$ est croissante dans la région $\frac{1}{2}x + y \geq 0$.

Dans le cas $\alpha = 0$, le système est

$$\begin{cases} x'(t) = \frac{1}{2}y(t) \\ y'(t) = \frac{1}{2}x(t) \end{cases}$$

et l'équilibre $Y^e = (0, 0)$ n'est pas globalement stable car la valeur propre λ_2 est positive ($\lambda_1 = -\frac{1}{2}$ et $\lambda_2 = \frac{1}{2}$). Par contre, si la condition initiale (x_0, y_0) se trouve sur la droite $y = -x$

déterminée par la direction de $v_1 = (1, -1)$ (vecteur propre associée à λ_1), alors la solution reste sur cette droite et comme $\lambda_1 < 0$, elle converge vers $(0, 0)$ lorsque $t \rightarrow +\infty$. Donc, l'équilibre $Y^e = (0, 0)$ est localement stable et son ensemble de stabilité est $\{(x, -x) : x \in \mathbb{R}\}$. Enfin, $x(t)$ est croissante dans la région $y \geq 0$ et $y(t)$ est croissante dans la région $x \geq 0$.

Dans le cas $\alpha = -1$, le système est

$$\begin{cases} x'(t) = -x(t) + \frac{1}{2}y(t) \\ y'(t) = \frac{1}{2}x(t) - y(t) \end{cases}$$

et l'équilibre $Y^e = (0, 0)$ est globalement stable car les deux valeurs propres sont négatives ($\lambda_1 = -\frac{3}{2}$ et $\lambda_2 = -\frac{1}{2}$). De plus, $x(t)$ est croissante dans la région $-x + \frac{1}{2}y \geq 0$ et $y(t)$ est croissante dans la région $\frac{1}{2}x - y \geq 0$.

Exemple 2 (matrice à valeurs propres complexes conjuguées). On considère le système différentiel

$$\begin{cases} x'(t) = \alpha x(t) + \frac{1}{2}y(t) \\ y'(t) = -\frac{1}{2}x(t) + \alpha y(t) \end{cases}$$

où $\alpha \in \mathbb{R}$. Les valeurs propres de la matrice associée à ce système sont : $\lambda_1 = \alpha - \frac{i}{2}$ et $\lambda_2 = \alpha + \frac{i}{2}$, les vecteurs propres correspondants étant : $v_1 = (1, -i)$ et $v_2 = (1, i)$. La solution générale du système est alors

$$\begin{aligned} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \alpha_1 \operatorname{Re} \left(e^{(\alpha + \frac{i}{2})t} \begin{pmatrix} 1 \\ i \end{pmatrix} \right) + \alpha_2 \operatorname{Im} \left(e^{(\alpha + \frac{i}{2})t} \begin{pmatrix} 1 \\ i \end{pmatrix} \right) \\ &= \alpha_1 e^{\alpha t} \begin{pmatrix} \cos \frac{t}{2} \\ -\sin \frac{t}{2} \end{pmatrix} + \alpha_2 e^{\alpha t} \begin{pmatrix} \sin \frac{t}{2} \\ \cos \frac{t}{2} \end{pmatrix}, \end{aligned}$$

avec $\alpha_1, \alpha_2 \in \mathbb{R}$. L'équilibre $Y^e = (0, 0)$ est globalement stable si et seulement si $\alpha < 0$. Par contre, si $\alpha = 0$, la solution générale $(x(t), y(t))$ se trouve sur le cercle de centre $(0, 0)$ et rayon $\sqrt{\alpha_1^2 + \alpha_2^2}$, car $x(t)^2 + y(t)^2 = \alpha_1^2 + \alpha_2^2$, c'est-à-dire l'équilibre Y^e est neutre. Enfin, si $\alpha > 0$, l'équilibre est instable car $(x(t), y(t))$ se trouve sur le cercle de centre $(0, 0)$ et rayon $r(t) = e^{\alpha t} \sqrt{\alpha_1^2 + \alpha_2^2}$ et $r(t) \rightarrow +\infty$ lorsque $t \rightarrow +\infty$.

Références

- [1] J.-P. Demailly, *Analyse numérique et équations différentielles*, Presses Universitaires de Grenoble, 1996.
- [2] B. Guerrien, *Initiation aux mathématiques*, Economica, 1991.
- [3] P. Michel, *Cours de mathématiques pour économistes*, Economica, 1989.
- [4] G. Poulalion, G. Pupion, *Les mathématiques de l'économiste*, Vuibert, 2002.
- [5] C. Simon, L. Blume, *Mathématiques pour économistes*, De Boeck Université, 1998.