UNIVERSITE ORAN 2
FACULTE DES SCIENCES ECONOMIQUES, DES SCIENCES DE GESTION ET DES SCIENCES COMMERCIALES

COMPLEMENTS DE MATHEMATIQUES

Notes de cours

Rafik BOUKLIA-HASSANE

P L A N

CHAP 1. RAPPEL SUR LES OPERATIONS SUR LES MATRICES

On note par $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ la matrice de format (n, m) suivante :

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

1. Quelques matrices particulières

- Une matrice carrée est dite ***diagonale*** si tous ses éléments en dehors de la diagonale sont nuls.
- Une matrice ***triangulaire*** supérieure (inférieure) est une matrice telle que tous les éléments situés en dessous (en dessus) de la diagonale sont nuls :

$$I_n = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nm} \end{pmatrix}$$

- La matrice ***unitaire*** I_n est une matrice diagonale d'ordre n dont tous les éléments de la diagonale sont égaux à 1 :

$$I_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

- La ***transposé*** de la matrice $A\,(n, m)$, notée ${}^tA$ (ou bien A'), est la matrice (m, n) obtenue en transposant les lignes de A avec ses colonnes.
- Une matrice carrée est ***symétrique*** si elle est égale à sa propre transposée : $(a_{ij} = a_{ji})$

2. Opérations sur les matrices

Somme de matrices

La somme des matrices est définie sur l'ensemble des matrices ayant le même format.

Soient deux matrices de même format $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ et $B = (b_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$

La matrice C somme de $A + B$ est obtenue en additionnant termes à termes les éléments de A et de B. Elle est définie par :

$$C = (c_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \quad \text{avec } c_{ij} = a_{ij} + b_{ij} \; \forall i,j$$

Exemple

Soit $A = \begin{pmatrix} 1 & -4 \\ 2 & 0 \\ -1 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 5 & 3 \\ -2 & 2 \\ 1 & 0 \end{pmatrix}$. Alors $A + B = \begin{pmatrix} 6 & -1 \\ 0 & 2 \\ 0 & 3 \end{pmatrix}$.

Propriété

On a ${}^t(A+B) = {}^tA + {}^tB$

Multiplication par un scalaire

La multiplication de la matrice A par un scalaire λ est la matrice $C = \lambda A$ obtenue en multipliant chaque terme de A par le scalaire λ. La matrice C est définie par :

$$C = (c_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \quad \text{avec } c_{ij} = \lambda a_{ij} \; \forall i,j$$

Exemple

Soit $A = \begin{pmatrix} 2 & 3 & 0 \\ 4 & 1 & -1 \end{pmatrix}$ et $\lambda = -2$. Alors $\lambda A = \begin{pmatrix} -4 & -6 & 0 \\ -8 & -2 & 2 \end{pmatrix}$.

Remarque

L'ensemble M_{nxm} des matrices de format (nxm) munie de l'addition et de la multiplication par un scalaire, telles que définies ci-dessus, forment un espace vectoriel. L'élément neutre pour l'addition est la matrice nulle d'ordre (nxm).

Particulièrement, la somme des matrices est associative et commutative.

Multiplication des matrices

L'opération de multiplication de deux matrices A et B n'est définie que si le nombre de colonnes de A est égal au nombre de lignes de B. La matrice C obtenue aura le nombre de ligne de A et le nombre de colonnes de B :

$$\underset{(n,m)}{C} = \underset{(n,p)}{A} \quad \underset{(p,m)}{B}$$

Le terme c_{ij} de la matrice C est obtenu en multipliant termes à termes la $i^{ème}$ ligne de A avec la $j^{ème}$ colonne de B.

Exemple

$$\underset{(2,3)}{A} = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & -1 \end{pmatrix} \quad ; \quad \underset{(3,2)}{B} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 3 \end{pmatrix}$$

Le format de la matrice C sera $(2,2)$ et :

$$C = \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}$$

Remarque

La multiplication des matrices est associative mais *n'est pas commutative*.

Propriété

On a ${}^{t}(AB) = {}^{t}B\ {}^{t}A$.

3. Calcul de déterminants.

Déterminant d'ordre 2

$$\begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} = (1 * 3) - (2 * -1) = 5$$

Déterminant d'ordre 3

Règle de Sarrus par exemple

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ -1 & 0 & 4 \end{vmatrix}$$
$$= [(1 * 1 * 4) + (2 * 3 * -1) + (2 * 0 * 1)] - [(-1 * 1 * 1) + (2 * 2 * 4) + (0 * 3 * 1)] = -17$$

Déterminants d'ordre supérieur à 3

On peut calculer le déterminant dans ce cas en développant suivant une ligne ou une colonne (qui possède le plus de 0) et en utilisant la règle des signes.

$$\begin{vmatrix} 1 & 2 & 1 & 0 \\ -2 & 3 & 2 & 3 \\ 0 & -1 & 0 & -2 \\ 0 & 1 & 3 & 1 \end{vmatrix}$$

$$= +1 * \begin{vmatrix} 3 & 2 & 3 \\ -1 & 0 & -2 \\ 1 & 3 & 1 \end{vmatrix} - (-2) * \begin{vmatrix} 2 & 1 & 0 \\ -1 & 0 & -2 \\ 1 & 3 & 1 \end{vmatrix} + 0 * \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix}$$

$$- 0 * \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} = \cdots$$

Propriétés

On a les propriétés suivantes :

- $|{}^{t}A| = |A|$
- $|A * B| = |A| * |B|$ (1)

4. Inversion de matrices

Définition

Une matrice est inversible si et seulement si il existe une matrice A^{-1} telle que :

$$A * A^{-1} = A^{-1} * A = I$$

où I est la matrice unité.

Propriété d'existence

Une matrice carrée A possède un inverse –noté A^{-1} -- si et seulement si son déterminant est différent de 0.

Etape de calcul de l'inverse d'une matrice *A*

- Calcul de $|A|$. Si $|A| = 0$, l'inverse de A n'existe pas. Sinon $(|A| \neq 0)$ on procède aux étapes suivantes :
- Calcul de la transposée ${}^{t}A$
- Calcul des cofacteurs de ${}^{t}A$
- Calcul de l'adjointe de A notée $Adj(A)$ obtenue en appliquant la règle des signes à la matrice des cofacteurs
- Calcul de l'inverse de A : $A^{-1} = \frac{Adj(A)}{|A|}$

Exemple

On calcule l'inverse de la matrice A suivante :

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

- $1^{\text{ère}}$ étape : Calcul du déterminant.

  $|A| = -7$. La matrice est donc inversible.

- $2^{\text{ème}}$ étape : Calcul de la matrice transposée :

$$^{t}A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

- $3^{\text{ème}}$ étape : Calcul de la matrice des cofacterus

$$\begin{pmatrix} \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 0 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & -1 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -1 & 2 & 2 \\ 3 & 1 & 1 \\ 2 & 3 & -4 \end{pmatrix}$$

- $4^{\text{ème}}$ étape : Calcul de la matrice adjointe :

$$Adj(A) = \begin{pmatrix} (+)-1 & (-)2 & (+)2 \\ (-)3 & (+)1 & (-)1 \\ (+)2 & (-)3 & (+)-4 \end{pmatrix} = \begin{pmatrix} -1 & -2 & 2 \\ -3 & 1 & -1 \\ 2 & -3 & -4 \end{pmatrix}$$

- $5^{\text{ème}}$ étape : Calcul de la matrice inverse :

$$A^{-1} = -\frac{1}{7}\begin{pmatrix} -1 & -2 & 2 \\ -3 & 1 & -1 \\ 2 & -3 & -4 \end{pmatrix} = \begin{pmatrix} 1/7 & 2/7 & -2/7 \\ 3/7 & -1/7 & 1/7 \\ -2/7 & 3/7 & -4/7 \end{pmatrix}$$

Exercice

Vérifier que $A * A^{-1} = A^{-1} * A = I$

Propriétés

On a les propriétés suivantes :

- $|A^{-1}| = \frac{1}{|A|}$ (2)
- $(AB)^{-1} = B^{-1}A^{-1}$

CHAP 2. VALEURS PROPRES ET VECTEURS PROPRES

Dans toute la suite, les matrices sont supposées carrées sauf mention contraire.

1- Définitions des valeurs propres et vecteurs propres

Définition

Soit une matrice carrée de n lignes et colonnes $A(n,n)$. On appelle valeur propre de A un nombre (réél ou complexe) λ tel qu'il existe un vecteur $V \neq 0$ vérifiant :

$$AV = \lambda V$$

Définition

Soit une matrice carrée $A(n,n)$. On appelle vecteur propre de A associé à la valeur propre λ un vecteur $V \neq 0$ vérifiant :

$$AV = \lambda V \qquad (3)$$

Remarque

La relation (1) peut également s'écrire (évident):

$$(A - \lambda I)V = 0 \qquad (4)$$

<u>*Question*</u> **:** Etant donnée une valeur propre λ d'une matrice carrée A, combien existe-t-il de vecteurs propres associées à cette matrice : un seul, une infinité ou aucun ?

Propriétés

- <u>*P1*</u> : Soit λ une valeur propre de la matrice A et V_1 et V_2 deux vecteurs propres associés à la valeur propre λ. Alors $V_1 + V_2$ est encore un vecteur propre associé à la valeur propre λ.
- <u>*P2*</u> : Soit V un vecteur propre associé à la valeur propre λ de A et α un scalaire. Alors αV est aussi un vecteur propre associé à λ.

Démonstration : à faire à titre d'exercice.

Remarques

- <u>*R1*</u> : Les propriétés P1 et P2 s'expriment de façon condensée et de manière équivalente par :

Soient V_1 et V_2 deux vecteurs propres associés à une valeur propre λ de A et α et β deux scalaires (nombres réels ou complexes), alors $\alpha V_1 + \beta V_2$ est encore une vecteur propre associé à la valeur propre λ de A.

Cela signifie qu'étant donnée une valeur propre λ d'une matrice A, il existe une infinité de vecteurs propres associés à cette valeur propre λ .

- R2 : la remarque R1 permet de voir que dans l'ensemble E_λ des vecteurs propres associés à une valeur propre λ d'une matrice, l'opération d'addition et de multiplication par un scalaire est *stable* (dans le sens où le résultat de chacune de ces deux opérations appartient encore à l'ensemble des vecteurs E_λ). On conclut donc que *l'ensemble des vecteurs propres E_λ associés à une valeur propre λ d'une matrice constitue un sous espace vectoriel*.

2. Calcul des valeurs propres et des vecteurs propres

Définition

On appelle polynôme caractéristique d'une matrice A d'ordre n, le polynome $P(\lambda)$ en λ défini par :

$P(\lambda) = det(A - \lambda I)$ où I est la matrice unité d'ordre n.

Exemple

Soit la matrice $A = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}$. Alors :

$$P(\lambda) = |A - \lambda I| = det\left|\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| = \begin{vmatrix} 1-\lambda & 2 \\ 1 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) - 2.$$

Donc $P(\lambda) = \lambda^2 - 5\lambda + 2$

Remarque

L'ordre du polynôme caractéristique $P(\lambda)$ est égal à l'ordre de la matrice A.

Théorème

L'ensemble des valeurs propres d'une matrice A sont les solutions du polynôme caractéristique $P(\lambda) = 0$.

Démonstration (en cours).

Théorème de d'Alembert

Un polynôme de degré n possède n racines dans C. Les solutions complexes sont conjuguées deux à deux.

Conséquence : Une matrice d'ordre n possède n valeurs propres dans $\mathbb{C}$. Ce résultat naturellement n'est pas valable dans R.

Exemple

Calculer dans R les valeurs propres et vecteurs propres de la matrice $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$.

1- Calcul du polynôme caractéristique :

$P(\lambda) = |A - \lambda I| = det\left|\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| = \begin{vmatrix} 1-\lambda & 2 \\ 1 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) - 2.$

Donc $P(\lambda) = \lambda^2 - 3\lambda = \lambda(\lambda - 3)$

2- Calcul des valeurs propres :

$P(\lambda) = 0 \Longrightarrow \lambda(\lambda - 3) = 0.$ Donc $\lambda_1 = 0$ et $\lambda_2 = 3$

Le cas de matrices d'ordre supérieur à 2 sera traité en cours.

3- Calculs des vecteurs propres :

Système définissant les vecteurs propres

Soit $V = \begin{pmatrix} x \\ y \end{pmatrix}$ un vecteur propre associé à la valeur propre λ. Par définition (cf. Eq. 4), on a : $(\mathrm{A} - \lambda\mathrm{I})\mathrm{V} = 0$. En substituant à A son expression, on obtient l'équation matricielle :

$$\begin{pmatrix} 1-\lambda & 2 \\ 1 & 2-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

qui se déclinent en deux équations scalaires :

$$\begin{array}{l} (1-\lambda)x + 2y = 0 \\ x + (2-\lambda)y = 0 \end{array} \qquad (5)$$

Par construction, ce système n'est pas un système de Cramer. Il possède nécessairement une infinité de solutions (Cf. R1 et voir directement pourquoi).

Vecteurs propres associés à $\lambda_1 = 0$

On remplace dans le système (3) λ par sa valeur 0 :

$$\begin{array}{l} x + 2y = 0 \\ x + 2y = 0 \end{array}$$

On vient que la deuxième équation n'apporte pas d'information supplémentaire (étant la même que la première équation). On n'a donc qu'une seule équation avec deux inconnues. Pour sa résolution on fixe l'inconnue par exemple x à c puis on cherche y en fonction de c. On obtient :

$$x = c$$
$$y = -\frac{c}{2}$$

Donc les vecteurs propres associés à $\lambda_1 = 0$ sont :

$$V_1 = \begin{pmatrix} c \\ -\frac{c}{2} \end{pmatrix} = c \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}$$

Vecteurs propres associés à $\lambda_2 = 3$

On remplace dans le système (3) λ par sa valeur 3 :

$$-2x + 2y = 0$$
$$x - y = 0$$

La deuxième équation là aussi n'est qu'une forme différente de la première équation (qu'on obtient en divisant la première équation par -2). On pose là encore $x = c$; ce qui implique que $y = c$. Donc :

$$V_2 = \begin{pmatrix} c \\ c \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Exercice (calculatoire)

Soit la matrice A:

$$A = \begin{pmatrix} 3 & -4 & 2 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{pmatrix}$$

Trouver la matrice des valeurs propres Λ, la matrice des vecteurs propres P.

Résultats : Valeurs propres :$\lambda = 1,2,3$.

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix}$$

3. Trace d'une matrice carrée

Définition

On appelle trace d'une matrice carrée A et note $Tr(A)$ la somme des éléments de sa diagonale :

$$Tr(A) = \sum_{i=1}^{n} a_{ii}$$

La notion de trace sera abondamment utilisée en économétrie.

Propriété

On ne change pas la trace d'un produit de matrices si on opère une permutation circulaire :

$Tr(AB) = Tr(BA)$; $Tr(ABC) = Tr(CAB) = Tr(BCA); \quad etc.$

Propriété

La trace d'une matrice est égale à la somme de ses valeurs propres.

(Pour la démonstration : voir ci-dessous les propriétés des matrices semblables)

4. Matrice semblables

Définition

Deux matrices A et B sont dites semblables si et seulement si il existe une matrice P inversible telle que :

$$A = PBP^{-1}$$

Deux matrices semblables sont associées à la même application linéaire sauf que celle-ci est rapportée à deux bases différentes correspondant aux matrices de passage P et P^{-1}. D'où la qualificatif de matrices 'semblables'.

Propriété

Deux matrices A et B semblables ont :

- Le même déterminant
- le même polynôme caractéristique
- les mêmes valeurs propres
- la même trace.

Démonstration :

- $|A| \underbrace{=}_{def\ \ semblable} |PBP^{-1}| \underbrace{=}_{Eq.(1)\ et\ (2)} |P||B||P^{-1}| = |B|$

- $|A-\lambda I| \underbrace{=}_{définition} |PBP^{-1}-\lambda I| = |PBP^{-1}-\lambda PIP^{-1}| = |P(B-\lambda I)P^{-1}| =$
  $|P||B-\lambda I||P^{-1}| = |B-\lambda I|$
- $3^{ème}$: évident.
- $Tr(A) \underbrace{=}_{def\ semblable} Tr(PBP^{-1}) \underbrace{=}_{perm.circulare} Tr(P^{-1}PB) = Tr(B)$

5. Diagonalisation des matrices : théorème fondamental de diagonalisation

Remarque : Dans ce qui suit, on n'étudie que le cas où la matrice A considérée possède des valeurs propres distinctes.

Théorème fondamental de diagonalisation

Soit A une matrice carrée possédant n valeurs propres distinctes λ_1, λ_2, ..., λ_n. Alors la matrice A peut s'écrire comme :

$$A = P\Lambda P^{-1} \qquad (6)$$

où $\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ est la matrice diagonale des valeurs propres et $P = (V_1, \quad V_2, \dots, \quad V_n)$ est la matrice inversible des vecteurs propres.

Remarque

Ce théorème montre qu'une matrice carrée est semblable à la matrice diagonale de ses valeurs propres.

Propriété (conséquence)

- La somme des valeurs propres d'une matrice est égale à la trace de cette matrice :

$$\sum_{i=1}^{n} \lambda_i = Tr(A)$$

- Le produit des valeurs propres d'une matrice est égal à son déterminant :

$$\prod_{i=1}^{n} \lambda_i = Det(A)$$

Exercice : démontrer la propriété ci-dessus

Exercice (calculatoire)

Soit la matrice A:

$$A = \begin{pmatrix} 3 & -4 & 2 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{pmatrix}$$

Trouver la matrice des valeurs propres Λ, la matrice des vecteurs propres P (voir exercice ci-dessus) et vérifier que :

$A = P\Lambda P^{-1}$.

CHAP 3. APPLICATION A LA RESOLUTION DES EQUATIONS LINEAIRES AUX DIFFRENCES FINIES

1. Définition d'un système d'équations linéaires aux différences finies

Il s'agit d'un système d'équations décrivant l'évolution (linéaire) de l'état d'un système défini à un instant t par ses n coordonnées $(x_t^1 \quad x_t^2, \dots, \quad x_t^n)$. Dans le cas où le système est à trois dimensions $(x_t, \quad y_t, \quad z_t)$, les équations d'évolution seront :

$$\begin{aligned} x_t &= a_1 x_{t-1} + a_2 y_{t-1} + a_3 z_{t-1} \\ y_t &= b_1 x_{t-1} + b_2 y_{t-1} + b_3 z_{t-1} \\ z_t &= c_1 x_{t-1} + c_2 y_{t-1} + c_3 z_{t-1} \end{aligned} \qquad (7)$$

Remarque : ce système peut s'écrire sous forme matricielle :

$$X_t = AX_{t-1} \qquad (8)$$

où la matrice du système $A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ et le vecteur $X_t = \begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix}$

2. Résolution dans le cas où les conditions initiales du système X_0 sont connues.

Résoudre le système linéaire équivaut à trouver l'expression de l'état du système X_t en fonction du temps t (et non plus en fonction de son expression X_{t-1} en $t-1$).

L'équation (8) permet d'écrire, par des substitutions successives (à démontrer par récurrence):

$$X_1 = AX_0$$

$$X_2 = AX_1 = A(AX_0) = A^2X_0$$

$$X_3 = AX_2 = A(A^2X_0) = A^3X_0$$

$$X_t = AX_{t-1} = A(A^{t-1}X_0) = A^tX_0$$

Ainsi, la relation $X_t = A^tX_0$ constitue la solution recherchée puisqu'elle permet d'exprimer à tout moment la valeur de X_t (l'état du système) dès lors que l'on connait son état initial. Toutefois, la difficulté est que le calcul de A^t peut être fastidieux si la matrice A est grande ou si le temps est lointain (t grand). Une idée est de substituer à la matrice A une matrice plus simple, en l'occurrence une matrice diagonale Λ dont la puissance est facile à calculer. En utilisant le théorème fondamental de la diagonalisation (6), on peut écrire :

$$X_t = (P\,\Lambda\,P^{-1})^t X_0$$

Or $(P\,\Lambda\,P^{-1})^t = P\,\Lambda^t\,P^{-1}$. Donc:

$$\boxed{X_t = P\,\Lambda^t\,P^{-1}\,X_0} \qquad (9)$$

où Λ est la matrice des valeurs propres et P la matrice des vecteurs propres de A.

Remarque

Lorsque les conditions initiales sont données la solution de (7) est unique.

3. Résolution dans le cas où les conditions initiales du système X_0 ne sont pas connues.

Si X_0n'est pas connu, alors $P^{-1}\,X_0$ ne l'est pas non plus. On pose alors que $P^{-1}\,X_0$ (qui est une matrice (n,1) est un vecteur : $P^{-1}\,X_0 = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$ où les éléments k_i sont des constantes dépendant des conditions initiales.

En remplaçant dans (9), on aura :

$$X_t = P\,\Lambda^t\,P^{-1}\,X_0 = (V_1 \quad \ldots \quad V_n)\begin{pmatrix} \lambda_1^t & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^t \end{pmatrix}\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = (\lambda_1^t V_1 \quad \ldots \quad \lambda_n^t V_n)\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$$

La multiplication par blocs donne alors :

$$\boxed{X_t = k_1\lambda_1^t V_1 + \cdots + k_n\lambda_n^t V_n = \sum_{i=1}^{n} k_i\,\lambda_i^t V_i}$$

Remarque

Dans ce cas, il y a une infinité de solutions, lesquelles dépendent des conditions initiales à travers les paramètresk_i.

4. Eude du comportement de long terme de la solution

Définition

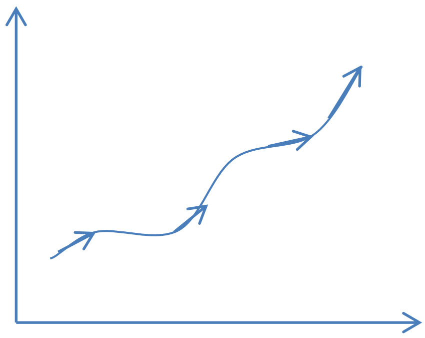
Le modèle $X_t = AX_{t-1}$ est dit stable (ou convergent) si ses solutions tendent vers 0 lorsque t tend vers l'infini. Dans le cas contraire, le modèle est dit explosif (ou divergent)=

$$\lim_{t\to+\infty} X_t \begin{cases} = 0 \iff modèle\ stable \\ \neq 0 \iff modèle\ divergent \end{cases}$$

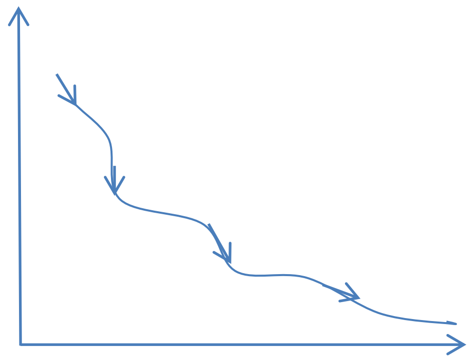
On s’attend à ce qu’un modèle économique soit stable pour qu’il soit raisonnable.

Définition

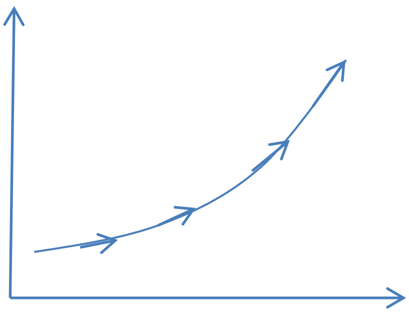
Le modèle $X_t = AX_{t-1}$ est dit monotone si ses solutions ne sont pas sujettes à des fluctuations. Dans le cas contraire, il est dit cyclique.



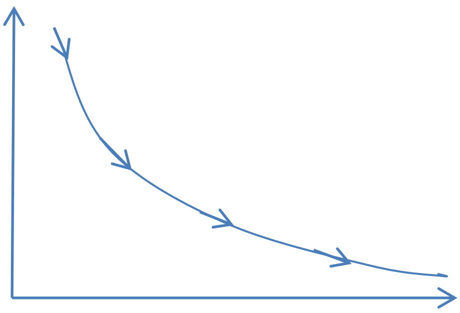
a- Système monotone divergent



b- Système cyclique convergent



c- Système monotone divergent



d- Système monotone convergent

Propriétés

Le modèle $X_t = AX_{t-1}$ est : (10)

- Stable si toutes ses valeurs propres sont en modules inférieures à 1 : $|\lambda_i| < 1 \ \forall i = 1, \dots, n$.

- Monotone si toutes ses valeurs propres sont réelles.

Récapitulatif

	$\lvert\lambda_i\rvert < 1$	$\lvert\lambda_i\rvert > 1$
λ_i réél	Convergent monotone	Divergent monotone
λ_i complexe	Convergent cyclique	Divergent cyclique

5. Cas des systèmes d'équations différentielles linéaires du premier ordre.

Définition

Un système d'équations différentielles linéaires de premier ordre s'écrit sous la forme (à trois variables):

$$\begin{aligned} x'(t) &= a_1x(t) + a_2y(t) + a_3z(t) \\ y'(t) &= b_1x(t) + b_2y(t) + b_3z(t) \\ z'(t) &= c_1x(t) + c_2y(t) + c_3z(t) \end{aligned} \qquad (11)$$

Ce système peut s'écrire sous forme matricielle :

$$X'(t) = AX(t) \qquad (12)$$

où la matrice du système $A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$ et le vecteur $X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$

NB : Attention, le prime désigne le symbole de dérivation et non le symbole de transposition.

Résolution

Pour résoudre ce système, on utilise le théorème de diagonalisation pour écrire la matrice A sous la forme $A = P\Lambda P^{-1}$ en supposant que les valeurs propres de A sont toutes distinctes. Dans ce cas, (12) se réécrit : $X'(t) = P\Lambda P^{-1}\mathrm{X(t)}$ ou encore :

$$P^{-1}X'(t) = \Lambda P^{-1}X(t) \qquad (13)$$

En posant

En posant $U(t) = P^{-1}X(t)$ (12) se réécrit :

$$U'(t) = \Lambda\, U(t).$$

Comme $\Lambda = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$ est une matrice diagonale, on a :

$$\begin{pmatrix} u_1'(t) \\ \vdots \\ u_n'(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix}.$$

On peut alors revenir à des équations différentielles scalaires simples :

$$\begin{gathered} u_1'(t) = \lambda_1 u_1(t) \\ \vdots \\ u_n'(t) = \lambda_n u_n(t) \end{gathered}$$

Considérons la première équation différentielle : $u_1'(t) = \lambda_1 u_1(t)$. Elle peut s'écrire $\frac{u_1'(t)}{u_1(t)} = \lambda_1$, dont la résolution est (à démontrer):

$$u_1(t) = k_1 e^{\lambda_1 t} \text{ avec } k_1 \in R$$

En répétant la même méthode de résolution pour les autres variables du système, on obtient :

$$U(t) = \begin{pmatrix} k_1 e^{\lambda_1 t} \\ \vdots \\ k_n e^{\lambda_n t} \end{pmatrix} = P^{-1} X(t)$$

On en déduit alors que :

$$X(t) = PU = (V_1 \quad \dots \quad V_n) \begin{pmatrix} k_1 e^{\lambda_1 t} \\ \vdots \\ k_n e^{\lambda_n t} \end{pmatrix}.$$

D'où les solutions :

$$\boxed{X(t) = \sum_{i=1}^{i=n} k_i e^{\lambda_i t} V_i}$$

Remarquer la similarité de la solution avec le cas d'un système d'équations aux différences finies.

Dans ce cas, le comportement asymptotique du système sera régi par la propriété :

Propriété

Le système d'équations différentielles linéaires de premier ordre sera :

- Stable si $\lambda_i < 0 \quad \forall i = 1, \dots, n$.
- Instable (divergent) si $\exists i \text{ } tel \text{ } que \text{ } \lambda_i < 0 \quad et \quad k_i \neq 0$.

Exercice

Etudier, sans résoudre, la dynamique de long terme (stabilité et monotonicité) des modèles suivants :

1 :

$$\begin{cases} x_t = -0.5x_{t-1} \\ y_t = -0.5x_{t-1} - 0.5z_{t-1} \\ z_t = x_{t-1} + 0.5z_{t-1} \end{cases}$$

2.

$$\begin{cases} x_t = 0.5x_{t-1} + z_{t-1} \\ y_t = z_{t-1} \\ z_t = -0.5y_{t-1} \end{cases}$$

3.

$$\begin{cases} x'(t) = -3x(t) - 3y(t) + 4z(t) \\ y'(t) = -2y(t) \\ z'(t) = -2x(t) - 3y(t) + 3z(t). \end{cases}$$

Approfondissement

Blanchard, O. J., & Kahn, C. M. (1980). The solution of linear difference models under rational expectations. *Econometrica*, 1305-1311.

CHAP 4. APPLICATIONS ECONOMIQUES DES VALEURS ET VECTEURS PROPRES

1. ETUDE DE LA STABILITE DES MODELES ECONOMIQUES

L'oscillateur de Samuelson (Samuelson, P. A. (1939). A Synthesis of the Principle of Acceleration and the Multiplier. *The Journal of Political Economy*, 786-797.)

On considère le modèle économique suivant décrivant une économie fermée (avec des notations évidentes):

$$C_t = aY_{t-1}$$

$$I_t = b(C_t - C_{t-1})$$

$$Y_t = C_t + I_t + G_t$$

Les questions posées :

- Interpréter économiquement le modèle : Le système économique est décrit ici par (C_t, I_t, Y_t). Identifier le processus du multiplicateur et de l'accélérateur dans le cas où $G_t = G$.
- On veut également étudier la stabilité du modèle dans le long terme pour voir s'il est économiquement raisonnable.
- Il faut enfin résoudre le modèle.

Premier cas : $G_t = 0$ (pas de dépenses publiques).

Interprétation économique du modèle. Recherche personnelle puis discussion en cours.

Etude du comportement de long terme du modèle de Samuelson du multiplicateur-accélérateur :

Pour cela, on écrit tout d'abord le modèle sous la forme standard : $X_t = AX_{t-1}$. Il est alors nécessaire de transposer à gauche les variables contemporaines et à droite les variables retardées de sorte que :

$$C_t = aY_{t-1}$$

$$I_t = b(C_t - C_{t-1}) = b(aY_{t-1} - C_{t-1}) = -bC_{t-1} + abY_{t-1}$$

$$Y_t = -bC_{t-1} + (a + ab)Y_{t-1}$$

$$\begin{pmatrix} C_t \\ I_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & a \\ -b & 0 & ab \\ -b & 0 & a+ab \end{pmatrix} \begin{pmatrix} C_{t-1} \\ I_{t-1} \\ Y_{t-1} \end{pmatrix}$$

Pour étudier le comportement de long terme du modèle, On calcule alors les valeurs propres de la matrice du système en fonction des paramètres du modèle, particulièrement la propension à consommer le revenu ainsi que le coefficient de l'accélérateur.

Cas particulier : Soit une propension à consommer $a = 0.8$ et un coefficient de l'accélérateur $b = 0.3$. Montrer alors que ce modèle de fluctuation est stable et monotone.

[On aura dans ce cas :

$$\begin{pmatrix} C_t \\ I_t \\ Y_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0.8 \\ -0.3 & 0 & 0.24 \\ -0.3 & 0 & 1.04 \end{pmatrix} \begin{pmatrix} C_{t-1} \\ I_{t-1} \\ Y_{t-1} \end{pmatrix}$$

On trouve comme valeurs propres $\lambda_1 = 0.34$ et $\lambda_2 = 0.7$]

Deuxième cas : $G_t = 0$ (pas de dépenses publiques).

On introduit l'Etat avec des dépenses publiques constantes$G_t = G$. La résolution dans ce cas passe par la détermination de la 'solution stationnaire'

Définition

On appelle solution stationnaire du système matricielle $X_t = AX_{t-1} + B$ une solution qui vérifie la condition $X_t = X_{t-1} = X^*$.

Propriété

Dans le cas où la matrice $(I - A)$ est inversible, les propriétés asymptotiques (stabilité et monotonicité) du modèle $X_t = AX_{t-1} + B$ ne dépendent pas de B (elles sont identiques à celles du modèle $X_t = AX_{t-1}$.

Pour le voir, on détermine d'abord la solution stationnaire du système $X_t = AX_{t-1} + B$. Supposons que cette solution existe. Alors elle vérifie : $X^* = AX^* + B$. Donc :

$$(I - A)X^* = B.$$

Puisque $(I - A)$est inversible, alors : $X^* = (I - A)^{-1}B$

Maintenant que nous disposons de la solution stationnaire, on fait le changement de variable : $Y_t = X_t - X^*$. La variable Y_t s'interprète comme l'écart de X_t par rapport à sa valeur de long terme (ie. les fluctuations de X autour de son équilibre de long terme). Alors, Y vérifie (*à montrer*) :

$$Y_t = AY_{t-1}$$

On est donc ramené à l'étude d'un système homogène dont on connait les propriétés asymptotiques (cf propriété 10). On résout alors le modèle en Y à l'aide des propriétés établies ci-dessus puis on se ramène à la variable X par la transformation $X_t = Y_t + X^*$.

Exercice

En utilisant Excel pour le calcul matriciel, résoudre le modèle

$$C_t = aY_{t-1}$$

$$I_t = b(Y_t - Y_{t-1})$$

$$Y_t = C_t + I_t + G$$

Avec les valeurs $a = 0.8$, $b = 0.3$ et $G = 10$.

Exercice de statique comparative : On étudie le comportement de la solution stationnaire à la suite d'un choc sur les dépenses publiques. Calculer numériquement l'écart ΔX^* lorsque les dépenses publiques passent de $G_0 = 10$ à $G_1 = 12$.

2. ANALYSE DES DONNEES – ANALYSE EN COMPOSANTES PRINCIPALES

Position du problème

La croissance g d'une économie nationale dépend de son niveau initial de développement approché par le $PIB\ par\ habitant$ (plus celui-ci est faible et plus on s'attend à ce que la croissance sera élevée du fait d'un potentiel de rattrapage), de son taux d'investissement $I/_Y$ en capital physique et humain, de ses politiques économiques approximées par son degré d'ouverture $Open$.

$$g_t = \alpha_1 PIBcap_t + \alpha_2\, I/Y_t + \alpha_3 Open_t + \varepsilon_t$$

On peut également estimer que la croissance dépend également du niveau de développement des infrastructures. Cependant on dispose de plusieurs variables approximant les infrastructures (longueur du réseau routier, réseau téléphoniques, réseau de TIC, etc…)

Quelles variables choisir et quelles variables exclure ?

On ne peut pas inclure l'ensemble des variables d'infrastructure pour au moins deux raisons :

- Inclure l'ensemble des variables diminue le degré de liberté
- Ces variables sont souvent multicolinéaires (elles véhiculent les mêmes informations ou des informations relativement similaires) sachant que la multicolinéarité diminue la précision des résultats d'estimation.

Une solution est de représenter les infrastructures par une ou deux variables seulement sachant que celles-ci doivent reproduire au mieux les évolutions des variables d'infrastructures primaires (routes, téléphones, TIC, développement logistique,…)

L'analyse en composantes principales permet d' « agréger » 'au mieux' les multiples variables d'infrastructures dont on dispose.

L'identification des composantes principales.

Soit les variables primaires d'infrastructure :

	X_t^1 Longueur réseau routier	X_t^2 Téléphones par habitant	X_t^3 TIC (Ordinateur par habitant)	X_t^4 Indice de dévpt de la logistique du commerce
2008	1000	0.3	0.010	0.4
2009	1010	0.32	0.015	0.4
2010	1010	0.35	0.019	0.3
2011	1060	0.40	0.025	0.35
2012	1080	0.40	0.026	0.45
2013	1100	0.45	0.30	0.4

Il s'agit d'agréger ces données en une ou deux variables représentatives des données d'infrastructure de l'économie considérée. Les étapes de la méthode d'analyse en composantes principales sont les suivantes :

Soient X^j les variables (d'infrastructures initiales).

- On centre et réduit les variables X^j. Soit Z^j les nouvelles variables centrées-réduites ;
- On détermine la matrice de corrélation R des variables Z^j
- On calcule les matrices Λ et P des valeurs propres et des vecteurs propres de R.
- Les composantes principales C sont alors données par :

$$C = X\left[P\Lambda^{-1/2}\right]$$

où les valeurs propres sont classées par ordre décroissants.

Les valeurs propres (de la matrice de corrélation des variables centrées réduites) jouent un rôle en ce qu'ils représentent l'inertie expliquée par les différents axes factoriels. Ainsi, le premier facteur associé à la plus grande valeur propre (première colonne de la matrice C)

explique ${\lambda_1}/{\sum_{i=1}^{n} \lambda_i}$ de la variance totale. Si on prend les deux premiers facteurs, on explique ${(\lambda_1 + \lambda_2)}/{\sum_{i=1}^{n} \lambda_i}$, etc.

CHAP 5. LES FORMES QUADRATIQUES

L'objet de ce chapitre est de construire la notion de signe d'une matrice. Dans l'absolu, on ne peut pas parler du signe d'une matrice du fait que les éléments d'une matrice peuvent être de signe tout à fait différent.

L'idée est d'associer de façon bijective à une matrice symétrique une expression réelle (une forme quadratique) et de définir le signe d'une matrice par le signe de la forme quadratique qui lui est associée.

Le signe des matrice entrent notamment dans la formulation des conditions de maximisation de second ordre (conditions suffisantes) des fonctions à plusieurs variables avec ou sans contraintes. Il sera utilisé dans le cours d'optimisation statique.

1. Définition d'une forme quadratique

Une forme quadratique d'ordre deux est une expression de la forme :

$$Q(x,y) = a_{11}x^2 + a_{12}xy + a_{21}yx + a_{22}y^2 \quad \text{avec } a_{ij} \in R$$

Naturellement, comme on est dans R, $xy = yx$, néanmoins cette écriture a le mérite de 'standardiser' l'expression de la forme quadratique étant entendu que $x^2 = xx$ et $y^2 = yy$.

Exemple

$Q(x,y) = 2x^2 + 3xy - y^2$ est une forme quadratique d'ordre deux.

Par analogie, une forme quadratique d'ordre 3 est une expression qui s'écrit sous la forme :

$$Q(x,y,z) = (a_{11}x^2 + a_{12}xy + a_{13}xz) + (a_{21}yx + a_{22}y^2 + a_{23}yz) + (a_{31}zx + a_{32}zy + a_{33}zz)$$

Naturellement, certains coefficients a_{ij} peuvent être nuls et le terme correspondant ne pas apparaître dans l'expression de la forme quadratique.

Exemple

$Q(x,y,z) = 2x^2 + 3xy + 6xz - y^2 + 3z^2$ est une forme quadratique d'ordre trois.

Formulation générale

Une forme quadratique d'ordre n $Q(x_1, x_2, .., x_n)$ est une expression de la forme :

$$Q(x_1, x_2, .., x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

Exercice

Retrouver à partir du cas général l'expression des formes quadratiques d'ordre deux et trois ci-dessus.

2. Représentation matricielle d'une forme quadratique

On peut aisément vérifier qu'une forme quadratique d'ordre n peut s'écrire sous la forme matricielle suivante :

$$Q(x_1, x_2, .., x_n) = (x_1, x_2, .., x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Si on pose $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$, alors la forme quadratique s'écrit sous forme condensée:

$$Q(X) = X'AX$$

<u>*Question*</u> : étant donnée une forme quadratique Q, combien existe-t-il de matrice A associée à la forme Q ?

Pour rappel, on souhaite avoir une relation <u>*biunivoque*</u> entre les matrices et les formes quadratiques afin d'affecter le signe de la forme quadratique à la matrice qui lui est associée.

Est-ce le cas ? Contre exemple :

Soit la forme quadratique d'ordre deux vue plus haut : $Q(x, y) = 2x^2 + 3xy - y^2$

La matrice associée à cette forme quadratique est :

$$A = \begin{pmatrix} 2 & 3 \\ 0 & -1 \end{pmatrix}$$

Toutefois, la même forme quadratique peut également s'écrire :

$$Q(x, y) = 2x^2 + 1.5xy + 1.5yx - y^2$$

Dans ce cas, la matrice qui lui est associée sera :

$$A = \begin{pmatrix} 2 & 1.5 \\ 1.5 & -1 \end{pmatrix}$$

On voit que pour une même forme quadratique, il peut exister une infinité de matrice associée. Il est donc nécessaire de se donner une convention qui permette de limiter le choix de la matrice associée, plus précisément une convention qui permette l'unicité du choix de la matrice associée. Cette convention ne doit pas être arbitraire mais doit permettre la facilité de calcul.

La règle adoptée est d'opter pour une matrice associée symétrique.

Plusieurs raisons justifient ce choix :

- Une matrice symétrique a toutes ses valeurs propres réelles (non complexes)
- La matrice P des vecteurs propres d'une matrice symétrique a pour inverse sa transposée : $P^{-1} = P'$. La matrice symétrique peut donc se décomposer en $P\Lambda P'$

Par ce choix, la matrice associée à une forme quadratique donnée sera unique.

Exemple

Soit la forme quadratique $Q(x, y, z) = 2x^2 + 3xy + 6xz - y^2 + 3z^2$. Sa matrice associée au sens de la convention retenue est la matrice symétrique:

$$A = \begin{pmatrix} 2 & 1.5 & 3 \\ 1.5 & -1 & 0 \\ 3 & 0 & 3 \end{pmatrix}$$

Inversement, la forme quadratique associée à la matrice

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

est :

$$Q(x, y) = x^2 + 2y^2 - z^2 + 2xxy - 2xz$$

3. Signe des formes quadratiques (et des matrices)

Considérons la forme quadratique :

$$Q(x, y) = x^2 + 2xy + 2y^2$$

Son signe dépend à priori des valeurs de x et de y. Toutefois, on peut réécrire Q sous la forme :

$$Q(x, y) = (x + y)^2 + y^2$$

On peut voir que $Q(x, y) > 0$ quelque soit $(x, y) \neq (0,0)$.

En effet, si $x \neq 0$ et $y \neq 0$, il est évident que $Q > 0$

Par ailleurs, si $x \neq 0$ et $y = 0$, alors $Q(x, y) = x^2 > 0$.

Enfin, si $x = 0$ et $y \neq 0$, alors $Q(x, y) = 2y^2 > 0$.

Au total, $Q(x, y)$ a toujours un signe strictement positif quelque soit $(x, y) \neq (0,0)$. On dira dans ce cas que $Q(x, y)$ (et la matrice A qui lui est associée) est définie-positive.

Considérons le deuxième exemple suivant :

$$Q(x, y) = (x + y)^2$$

On peut facilement voir que le signe de Q dépend dans un certain sens des valeurs de x et de y. En effet :

$$Q(x, y) = \begin{cases} > 0 \; si \;\; x \neq -y \\ = 0 \; si \; x = -y \end{cases}$$

On dira dans ce cas que $Q(x, y)$ est semi définie-positive.

Enfin, dans l'exemple suivant, $Q(x, y)$ change de signe suivant les valeurs de x et y :

$$Q(x, y) = x^2 - y^2 = \begin{cases} > 0 \; si \; |x| > |y| \\ = 0 \; si \; |x| = |y| \\ < 0 \; si \; |x| < |y| \end{cases}$$

On dira dans ce cas que $Q(x, y)$ est non définie.

Définitions

- Une forme quadratique $Q(X)$ (et la matrice A qui lui est associée) est définie-positive ssi :

$$Q(X) > 0 \; \forall X \neq 0$$

- Une forme quadratique $Q(X)$ (et la matrice A qui lui est associée) est définie-négative ssi :

$$Q(X) < 0 \; \forall X \neq 0$$

- Une forme quadratique $Q(X)$ (et la matrice A qui lui est associée) est semi définie-positive ssi :

$$Q(X) \geq 0 \; \forall X \neq 0 \; et \; \exists X^* \neq 0 \; tel \, que \, Q(X^*) = 0$$

- Une forme quadratique $Q(X)$ (et la matrice A qui lui est associée) est semi définie-négative ssi :

$$Q(X) \leq 0 \; \forall X \neq 0 \; et \, \exists X^* \neq 0 \; tel \, que \, Q(X^*) = 0$$

- Une forme quadratique $Q(X)$ (et la matrice A qui lui est associée) est non définie ssi :

$$\exists X_1 \neq 0 \; tel \; que \; Q(X_1) > 0 \;\; et \;\; \exists X_2 \neq 0 \; tel \, que \, Q(X_2) < 0.$$

Le signe d'une matrice symétrique est le signe de la forme quadratique qui lui est associée.

4. Identification du signe d'une matrice symétrique

Théorème 1 sur le signe des matrices

- Une matrice symétrique est définie-positive (négative) ssi toutes ses valeurs propres sont strictement positives (négatives) :

  $Q(X) définie\ positive\ (négative) \Leftrightarrow\ \lambda_i > 0\ (< 0)\ \forall i = 1, \dots, n$.

- Une matrice symétrique est semi définie-positive (négative) ssi toutes ses valeurs propres sont positives (négatives) ou nulles et au moins une des valeurs propres est nulle :

  $$Q(X) semi\ définie\ positive\ (négative) \Leftrightarrow\ \lambda_i \geq 0\ (\leq 0)\ \forall i = 1, \dots, n\ et\ \exists j\ tel\ que\ \lambda_j = 0$$

- Une matrice symétrique est non définie ssi elle possède des valeurs propres strictement positives et des valeurs propres strictement négatives :

  $Q(X)\ non\ définie \Leftrightarrow\ \exists \lambda_i > 0\ et\ \exists \lambda_i < 0$

Démonstration de la première proposition :

$$Q(X) définie\ positive \Leftrightarrow \forall X \neq 0: Q(X) = X'AX > 0 \underbrace{\Leftrightarrow}_{décomposition} \forall X \neq 0: X'(\mathrm{P}\Lambda P^{-1})X > 0 \underbrace{\Leftrightarrow}_{car\ A\ sym.} \forall X \neq 0: X'P\Lambda P'X > 0$$

On pose $Y = P'X$. Alors :

- $Y' = X'P$
- $X \neq 0 \Leftrightarrow Y \neq 0$ (car P inversible)

Alors :

$$Q(X) définie\ positive \Leftrightarrow \forall Y \neq 0: Y'\Lambda Y > 0 \qquad (10)$$

Or Λ est la matrice *diagonale* des valeurs propres :

$\Lambda = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$ de sorte que :

$Y'\Lambda Y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$. Alors, la relation (10) peut se réécrite :

$$Q(X) définie\ positive \Leftrightarrow \forall Y \neq 0: \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 > 0 \qquad (11)$$

On va monter que $Q(X) définie\ positive \implies \lambda_i > 0$

Comme la relation (11) est vraie $\forall Y \neq 0$, elle l'est notamment pour le vecteur particulier Y dont tous les éléments sont nuls sauf la i[ème] composante qui est égale à 1 :

$$Y = (0,0, \dots ,1,0, \dots ,0)$$

En appliquant la relation (11) à ce vecteur particulier, on obtient :

$$Q(X) définie\ positive \Rightarrow \lambda_1 0 + \lambda_2 0 + \dots + \lambda_i 1^2 + \lambda_n 0 = \lambda_1 > 0$$

Il en est de même pour le signe de toutes les autres valeurs propres.

Inversement, on voit facilement que si tous les éléments λ_i sont strictement positifs, alors pour tout vecteur $Y \neq 0$, le nombre $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2 > 0$, ce qui par la relation (11) prouve que la forme quadratique est définie-positive.

Application

Trouver le signe de la forme quadratique : $Q(x, y, z) = 2x^2 - y^2 + z^2 + 2xy$

Exercice

Discuter suivant les valeurs de a, le signe de l'expression :

$$Q(x, y, z) = ax^2 + xy + y^2$$

Exercice

Démontrer en vous inspirant de la discussion ci-dessus les autres propositions.

La méthode d'indentification du signe d'une matrice est générale mais a l'inconvénient de demander le calcul des vecteurs (du moins de leur signe), ce qui peut être problématique dès que le format de la matrice est supérieur à deux car cela nécessite la résolution de polynôme de degré supérieur à deux. Une autre méthode utile, moins générale, mais qui ne nécessite pas le calcul des valeurs propres peut être utilisée. Son inconvénient est qu'elle ne fournit une équivalence logique que dans le cas de matrices définies-positives (négatives).

Cette limitation est cependant moins grave que cela peut être apparaître de prime abord car très souvent, on s'intéresse surtout au cas où les matrices doivent être définies positives (négatives). C'est le cas généralement des matrices de variance-covariance en économétrie. C'est aussi la condition de signe de la matrice hessienne dans les problèmes d'optimisation libre.

Avant d'énoncer cette propriété, on définit le mineur d'une matrice carrée.

Le mineur d'ordre i d'une matrice A et la matrice A_i extraite de A en ne retenant que les i premières lignes et les i premières colonnes de A. Formellement, on pose la définition suivante :

Définition

Soit une matrice carrée $A = \left(a_{pq}\right)_{p,q=1,\dots,n.}$. On appelle mineur d'ordre i de A, la matrice A_i définie par : $A = \left(a_{pq}\right)_{p,q=1,\dots i.}$.

Exemple

Soit la matrice $A = \begin{pmatrix} -1 & -4 & -1 \\ 1 & -3 & 0 \\ -1 & 4 & 2 \end{pmatrix}$. Alors :

$$A_1 = (1)$$

$$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & 0 \\ -1 & 4 & -2 \end{pmatrix} = A.$$

Théorème 2 sur le signe des matrices

- Une matrice symétrique est définie positive si et seulement si le déterminant de tous ses mineurs est positif :

  $$A\ d\acute{e}finie\ positive \Leftrightarrow |A_i| > 0\ \forall i = 1, \dots, n.$$

- Une matrice symétrique est définie négative si et seulement si le signe des déterminants de ses mineurs est alterné avec $|A_1| < 0$:

  $$A\ d\acute{e}finie\ positive \Leftrightarrow |A_1| < 0\ ;\ |A_2| > 0\ ;\ |A_3| < 0\ ; \dots$$

Exercice

Trouver le signe de la matrice :

$$A = \begin{pmatrix} -1 & -3 & -1 \\ 2 & +3 & 0 \\ 0 & 4 & 2 \end{pmatrix}$$

CHAP 6. LA DERIVATION CHAINE

1. Rappels

On considère la composition des fonctions à une variable suivante :

$$\begin{array}{ccccc} R & \rightarrow & R & \rightarrow & R \\ t \rightarrow & x(t) & \rightarrow & f[x(t)] = h(t) & \end{array}$$

Pour rappel, la dérivée de h est :

$$h'(t) = f'[x(t)]x'(t)$$

2. Première généralisation

On opère une première généralisation de ce résultat en considérant considère que la fonction x est à valeurs dans R^n :

$$\begin{array}{ccccc} R & \rightarrow & R^n & \rightarrow & R \\ t \rightarrow x(t) = & (x_1(t), x_2(t), \ldots, x_n(t)) & \rightarrow & f(x_1(t), x_2(t), \ldots, x_n(t)) = h(t) & \end{array}$$

La fonction h qui va toujours de R dans R est la composition des deux fonctions x et f

$$h(t) = (fox)(t)$$

A titre d'exemple, on peut considérer la variable t comme le temps et les x_i comme des fonctions variant dans le temps. La question posée est de savoir comment la fonction composée h se comporte dans le temps, autrement dit de pouvoir calculer $\frac{dh}{dt}$ dans ce cas plus général.

Théorème de la dérivation en chaîne

Si les fonctions x_i sont de classe C^1 au voisinage d'un point t_0 et la fonction f est également de classe C^1 au voisinage du point $x(t_0)$, alors la fonction h est de classe C^1 au point t_0 et on a :

$$\frac{dh}{dt}(t_0) = \frac{\partial f}{\partial x_1}[x(t_0)]\frac{dx_1}{dt}(t_0) + \frac{\partial f}{\partial x_2}[x(t_0)]\frac{dx_2}{dt}(t_0) + \cdots + \frac{\partial f}{\partial x_n}[x(t_0)]\frac{dx_n}{dt}(t_0)$$

Exemple.

Soit les fonctions $x(t) = (x_1(t), x_2(t)) = (2t, -t)$; $f(x_1, x_2) = x_1^2 + \frac{1}{x_2}$ et $h(t) = (fox)(t)$

Calculer la dérivée $h'(t)$.

On a : $h'(t) = f'_{x_1}\big(x_1(t), x_2(t)\big)x'_1(t) + f'_{x_2}\big(x_1(t), x_2(t)\big)x'_2(t)$

Or $x'_1(t) = 2$; $x'_2(t) = -1$; $f'_{x_1} = 2x_1$ et $f'_{x_2} = -\frac{1}{x_2^2}$. Donc :

$$h'(t) = 2(2t)2 + (-1)\frac{-1}{t^2} = 8t + \frac{1}{t^2}$$

Vérification : on a par substitution :

$$h(t) = (fox)(t) = f[x(t)] = f(2t, -t) = (2t)^2 + \frac{1}{(-t)} = 4t^2 - \frac{1}{t}$$

En dérivant par rapport à t, on vérifie que

$$h'(t) = 8t + \frac{1}{t^2}$$

3. Deuxième généralisation

On procède à une deuxième généralisation en considérant que la fonction x(t) est définie dans R^p au lieu de R : $x : t = (t_1, t_2, \ldots, t_p) \to x(t_1, t_2, \ldots, t_p)$.

Dans ce cas, la fonction composée $h(t) = (fox)(t)$ sera définie de R^p dans R et, par conséquent, on peut définir les dérivées suivantes :

$$h'_{t_1}(t), h'_{t_2}(t), \ldots h'_{t_p}(t)$$

Toutefois, cette généralisation est immédiate dès lors qu'il suffit de dériver successivement par rapport à chacune des variables t_i en considérant les autres variables comme des constantes et d'appliquer le résultat ci-dessus :

$$\frac{dh}{dt_i}(t_0) = \frac{\partial f}{\partial x_1}[x(t_0)]\frac{dx_1}{dt_i}(t_0) + \frac{\partial f}{\partial x_2}[x(t_0)]\frac{dx_2}{dt_i}(t_0) + \cdots + \frac{\partial f}{\partial x_n}[x(t_0)]\frac{dx_n}{dt_i}(t_0)$$

Exemple (Simon et Blume p. 294)

On considère une fonction de production de Cobb Douglas $Q(K, L) = K^{3/4}L^{1/4}$ dans laquelle le capital K et le travail L sont des fonctions du temps t et du taux d'intérêt r : $K = K(t, r) = \frac{10t^2}{r}$ et $L = L(t, r) = 6t^2 + 250r$. Calculer $\frac{\partial Q}{\partial t}$ et $\frac{\partial Q}{\partial r}$ au point $t = 10$ et $r = 0.1$.

Par application de la dérivation en chaîne, on a :

$$\frac{\partial Q}{\partial t}(10, 0.1) = \frac{\partial Q}{\partial K}(K, L)\frac{\partial K}{\partial t}(10{,}0.1) + \frac{\partial Q}{\partial L}(K, L)\frac{\partial L}{\partial t}(10{,}0.1)$$

Il faut évaluer les quantités K et L au point $(10, 0.1)$ puis les différentes dérivées partielles intervenant dans la relation ci-dessus :

$K(10, 0.1) = \frac{10.10^2}{0.1} = 10000\ et r(10{,}0.1) = 6.10^2 + 250.(0.1) = 625$

$$\frac{\partial Q}{\partial K}(10000{,}625) = \frac{3}{4}K^{-1/4}L^{1/4} = \frac{3}{4}10000^{-1/4}625^{1/4}$$

$$\frac{\partial Q}{\partial L}(10000{,}625) = \frac{1}{4}K^{3/4}L^{-3/4} = \frac{1}{4}10000^{3/4}625^{-3/4}$$

$$\frac{\partial K}{\partial t}(10{,}0.1) = \frac{20t}{r} = \frac{20.10}{0.1} = 2000$$

$$\frac{\partial L}{\partial t}(10{,}0.1) = 12t = 12.10 = 120.$$

En remplaçant dans l'expression de la dérivée $\frac{\partial Q}{\partial t}$, on obtient : $\frac{\partial Q}{\partial t} = 3960.$

Exercice

Calculer de la même manière $\frac{\partial Q}{\partial r}(10, 0.1)$.

Cas particulier important.

Dans certains cas, la fonction $x(t) = (t, x_2(t))$ (autrement dit, on se situe dans le cas où la fonction x est définie dans R et prend ses valeurs dans R^2 (première généralisation) avec $x_1(t) = t$. Alors $\frac{dx_1}{dt} = 1$

Dans ce cas, $h'(t)$ s'exprime par :

$$h'(t) = \frac{\partial f}{\partial x_1}(t_0) + \frac{\partial f}{\partial x_2}[x(t_0)]\frac{dx_2}{dt}(t_0)$$

Ce résultat sera utilisé pour la démonstration du théorème des fonctions implicites qu'on va présenter ci-dessous.

4. Application : Le théorème des fonctions implicites

Théorème des fonctions implicites

Soit une fonction vérifiant $f(x, y)$ vérifiant les hypothèses suivantes :

- Les dérivées partielles (f'_x, f'_y) de f existent et sont continues
- Il existe un point (a, b) tel que $f(a, b) = 0$ et $f'_y(a, b) \neq 0$

Alors, si $f(x, y) = 0$ pour $x \in I(a)$ et $y \in I(b)$: il existe une fonction $\theta: I(a) \to I(b)$ telle que $y = \theta(x)$ vérifiant :

$$\theta'(x) = -\frac{f'_x(x,\ \theta(x))}{f'_y(x,\ \theta(x))}$$

Ce théorème signifie que bien qu'on ne connaisse pas la fonction θ (elle n'est définie qu'implicitement), on peut parfois être enseigné sur le signe de sa dérivée.

Démonstration.

Soit $y = \theta(x)$ sur l'intervalle $I(a)$. Comme $f(x, y) = 0$, alors : $f(x, \theta(x)) = h(x) = 0$ sur $I(a)$. On déduit que sur cet intervalle on a $h'(x) = 0$

Par ailleurs, on considère la composition des fonctions suivantes :

$$x \to \big(x, \theta(x)\big) \to f\big(x, \theta(x)\big) = h(x)$$

En appliquant le théorème de la dérivation en chaîne, on obtient :

$$h'(x) = f'_x\big(x, \theta(x)\big) + f'_y\big(x, \theta(x)\big)\theta'(x).$$

Comme $f'_y\big(x, \theta(x)\big) \neq 0$, on déduit que :

$$\theta'(x) = -\frac{f'_x\big(x, \theta(x)\big)}{f'_y\big(x, \theta(x)\big)}$$

Exemple 1

Les courbes d'iso-utilité $y = \theta(x)$ (autrement dit, les combinaisons entre x et y procurant la même utilité sont décroissantes.

On le montre sans à avoir à spécifier explicitement la fonction d'utilité $U(x, y)$ mais uniquement en retenant l'hypothèse que la fonction U est de classe C^1 et que $U'_x > 0$ et $U'_y > 0$.

En effet, la courbe d'indifférence est définie par $U(x, y) = c$ ou encore :

$$U(x, y) - c = 0$$

Cette équation définit une relation implicite entre y et x, soit $y = \theta(x)$. On a de plus,

$$\theta'(x) = -\frac{U'_x\big(x, \theta(x)\big)}{U'_y\big(x, \theta(x)\big)}$$

Ce qui montre d'après les seules hypothèses sur le signe des dérivées que $\theta'(x) < 0$.

Exemple 2

L'équilibre keynésien sur le marché des biens est défini par :

$$Y = C(Y - T) + I(Y, r) + G$$

On veut en déduire à l'équilibre le sens de la relation entre la production Y et le taux d'intérêt r. On ne spécifie pas la forme de la fonction de consommation ou celle de l'investissement mais on fait les hypothèses suivantes :

- la consommation est une fonction croissante du revenu (pour des taxes constantes)
- L'investissement est une fonction décroissante du taux d'intérêt et croissant par rapport à la production.
- L'accroissement marginal de la somme de la consommation et de l'investissement résultant d'une augmentation d'une unité de la production est inférieur à 1.

Ces hypothèses se traduisent par les relations $0 < C_Y' + I_Y' < 1$ et $I_r^{'} < 0$.

La relation d'équilibre peut se réécrire pour T et G fixé: $Y - C(Y - T) - I(r, Y) - G = f(r, Y) = 0$.

En appliquant le théorème des fonctions implicites, on obtient :

$$Y'(r) = -\frac{-I_r'(r, Y)}{1 - C_Y'(Y - T) - I_Y'(r, Y)}$$

Ce qui montre que la relation entre la production et le taux d'intérêt (relation IS) est négative dans le modèle keynésien.

CHAP 7. FONCTIONS CONVEXES, FONCTIONS CONCAVES

1. Définition d'un ensemble convexe

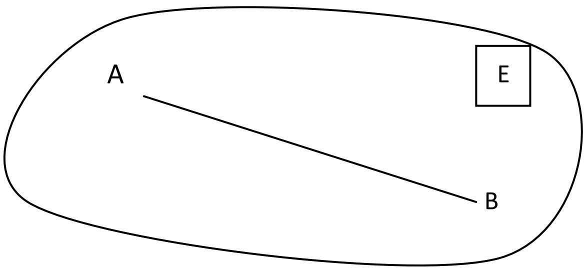
Définition d'un segment de droite

Soient $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ deux points de R^n. Le segment de droite $I = [A, B] \subset R^n$ est l'ensemble des points $X \in R^n$ qui s'écrivent sous la forme : $X = \lambda A + (1 - \lambda)B$ avec $\lambda \in [0,1]$:

$I = \{X \in R^n \colon \exists \lambda \in [0,1]\ tel\ que\ X = \lambda A + (1 - \lambda)B\}$

Définition d'un ensemble convexe

Soit A un ensemble de R^n. On dit que A est ensemble convexe si à chaque fois qu'il contient deux points A et B, alors il contient tout le segment $I = [A, B]$



Ainsi, l'ensemble E de la figure ci-dessus est un ensemble convexe.

Plus précisément,

$E \subset R^n$ convexe $\Leftrightarrow$ { $\forall A \in E, B \in E$ et $\forall \lambda \in [0, 1]$: $\lambda A + (1 - \lambda)B \in E$

2. Définition d'une fonction convexe / concave

Définition

Soit une fonction f définie par : $f \colon A \subset R^n \quad \rightarrow \ R$ où le sous ensemble A de R^n est un ensemble convexe (ouvert). Alors :

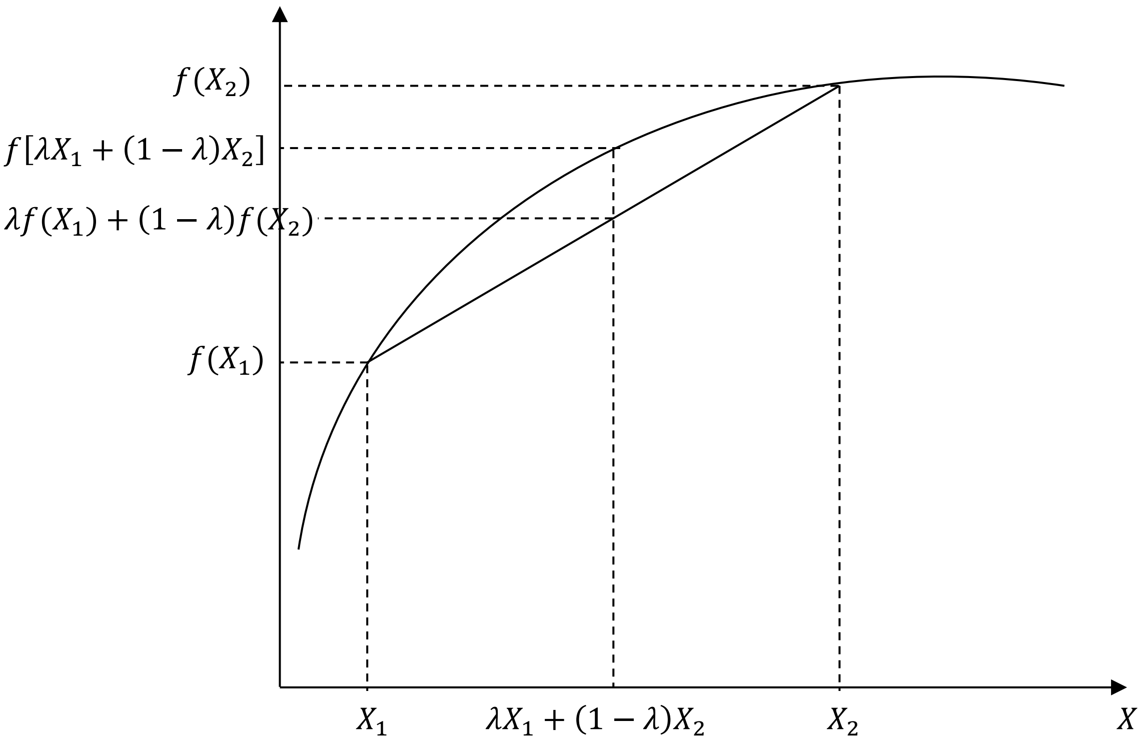
- f est concave si et seulement si :

$\lambda f(X_1) + (1-\lambda) f(X_2) \ \leq \ f[\lambda X_1 + (1-\lambda) X_2] \quad \forall X_1, X_2 \in R^n$ et $\forall \lambda \in [0, 1]$

- f est convexe si et seulement si :

$\lambda f(X_1 + (1-\lambda) f(X_2) \ \geq \ f[\lambda X_1 + (1-\lambda) X_2] \quad \forall X_1, X_2 \in R^n$ et $\forall \lambda \in [0, 1]$

La fonction f de la figure ci-dessous est une fonction concave.

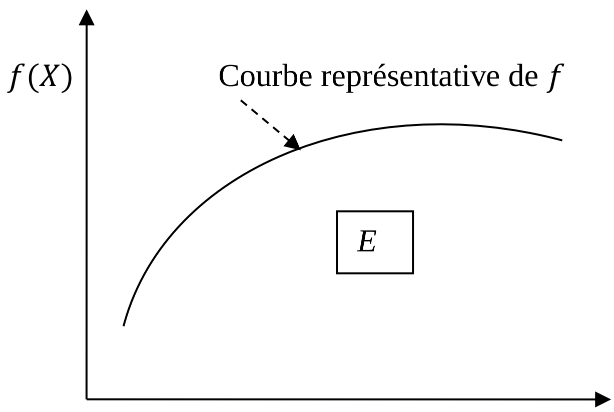


<u>Remarque</u> : si f est concave, alors la fonction $-f$ est convexe. C'est évident.

3. Propriétés

Propriété 1

f est concave $\Leftrightarrow$ {l'ensemble des points situés sur ou sous la courbe de f est convexe}



L'ensemble E des points situés sous la courbe représentative de la fonction concave f est convexe.

De même, l'ensemble des points situés sur ou sur la courbe d'une fonction convexe est convexe.

Démonstration :

Soient deux points $A\,(X_1, Y_1)$ et $B\,(X_2, Y_2)$ appartenant à E. Il faut montrer que le point $\lambda A + (1-\lambda)B \in E$, c'est-à-dire que : $\lambda Y_1 + (1-\lambda)Y \leq f[\lambda X_1 + (1-\lambda)X_2]$

Comme les points de E sont tous situés sous la courbe de f, alors :

$Y_1 \leq f(X_1)$ et $Y_2 \leq f(X_2)$.

Pour tout $\lambda \in [0, 1]$, on aura :

$\lambda Y_1 + (1-\lambda)Y_2 \leq \lambda f(X_1) + (1-\lambda)f(X_2)$ (*)

Comme f est concave, on a également :

$\lambda f(X_1) + (1-\lambda)f(X_2) \leq f[\lambda X_1 + (1-\lambda)X_2]$

Ce qui avec la relation (*) montre que :

$\lambda Y_1 + (1-\lambda)Y_2 \leq f[\lambda X_1 + (1-\lambda)X_2]$

Propriété 2

Soit A l'ensemble des points définis par

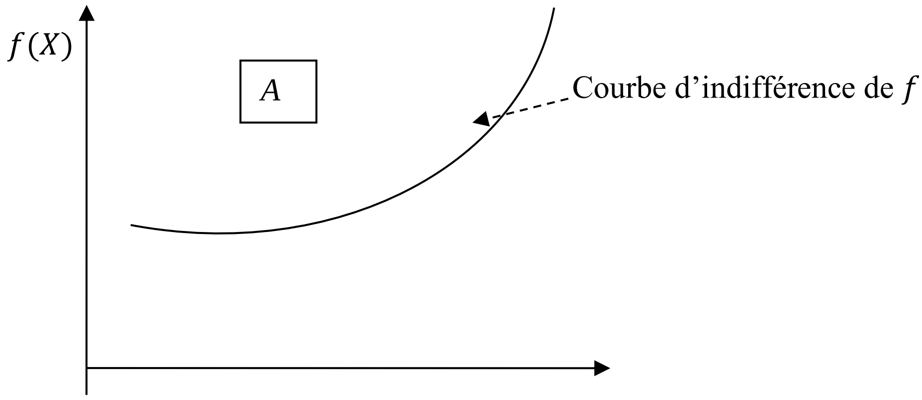
$$A = \{X \in R^n : f(X) \geq c\}$$

Ce sont les points situés au dessus de la courbe d'indifférence $(X) = c$. Alors :

$$f \text{ est concave} \Longrightarrow A \text{ est convexe}$$

La réciproque n'est pas vraie.

Cette propriété signifie que si f est concave, alors l'ensemble des points situés au dessus de la courbe d'indifférence $f(X) = c$ est convexe.:



A est convexe.

Démonstration

Soient deux points X_1 et X_2 appartenant à A. Il faut montrer que le point $\lambda X_1 + (1-\lambda)X_2 \in A$, c'est-à-dire que : $f[\lambda X_1 + (1-\lambda)X_2] \geq c$

Or :

$$\left.\begin{array}{l} X_1 \in A \Rightarrow f(X_1 \geq c \\ X_2 \in A \Rightarrow f(X_2 \geq c \end{array}\right\} \Rightarrow \lambda f(X_1) + (1-\lambda)f(X_2) \geq c \quad \text{pour tout} \quad \lambda \in [0,1] \quad (*)$$

Par ailleurs,

f concave $\Rightarrow f[\lambda X_1 + (1-\lambda)X_2] \geq \lambda f(X_1) + (1-\lambda)f(X_2)$

Ce qui implique d'après (*) que :

$f[\lambda X_1 + (1-\lambda)X_2] \geq c$

4. Opérations sur les fonctions concaves

Propriété 1

Si f et g sont deux fonctions concaves (convexes), alors la somme $f+g$ est encore une fonction concave (convexes).

(Démonstration facile : à faire comme exercice).

Propriété 2

$$\left.\begin{array}{c} f\ concave \\ g\ concave\ et\ \boldsymbol{croissante} \end{array}\right\} \Rightarrow fog \text{ est concave}$$

Démonstration :

Soient $X_1\ et\ X_2$ deux points de R^n.

On a $fog[\lambda X_1 + (1-\lambda)X_2] = \mathrm{f}\{\mathrm{g}[\lambda X_1 + (1-\lambda)X_2]\}$

Or g concave $\Longrightarrow \mathrm{g}[\lambda X_1 + (1-\lambda)X_2] \geq \lambda g(X_1) + (1-\lambda)g(X_2)$

Comme f est croissante, alors :

$$\mathrm{f}\{\mathrm{g}[\lambda X_1 + (1-\lambda)X_2]\} \geq \mathrm{f}\{\lambda g(X_1) + (1-\lambda)g(X_2)\}$$

Et comme f est concave (en posant $g(X_1) = \mathrm{X}_1$ et $g(X_2) = \mathrm{X}_2$:

$$\mathrm{f}\{\lambda g(X_1) + (1-\lambda)g(X_2)\} \geq \lambda f\left[g(X_1) + (1-\lambda)f[g(X_2)]\right]$$

ce qui montre que la fonction composée fog est concave. Noter que cette propriété n'est vrai que si la fonction g est croissante.

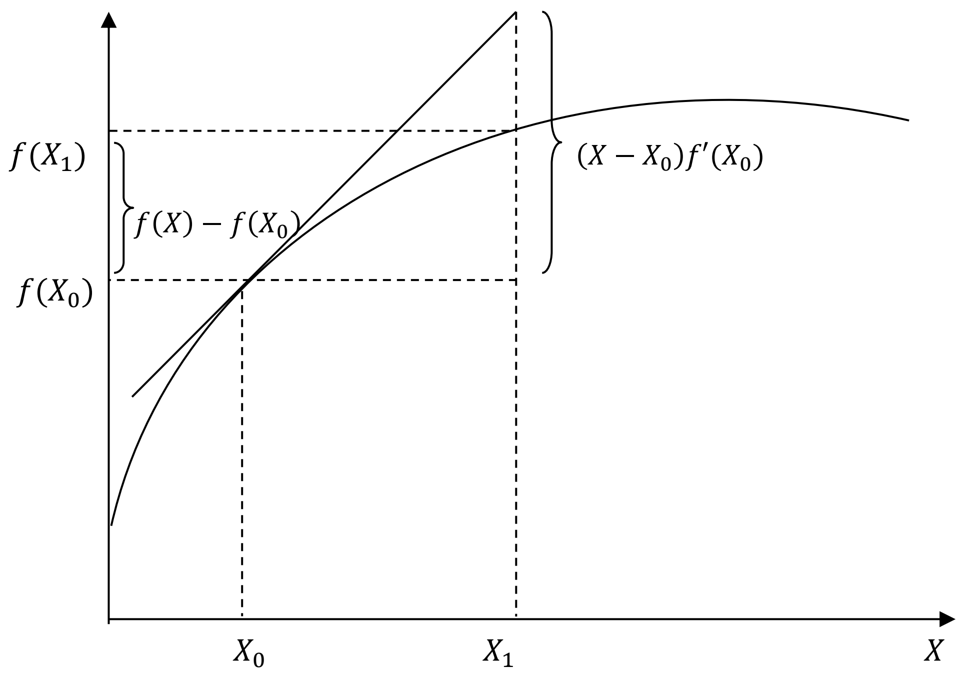
5. Propriétés des fonctions concaves dérivables.

Propriété 1 : Cas d'une fonction concave appartenant à $\mathcal{C}^1$

Soit une fonction $f: A \subset R^n \rightarrow R$ où A est convexe et dont les dérivées partielles existent et sont continues. Alors :

$$f\,concave \Longrightarrow f(X) - f(X_0) \leq (X - X_0)gradf(X_0) \quad \forall\, X\; et\; X_0 \in A$$

Dans le cas particulier où la fonction *f* est définie dans *R*, on a la représentation géométrique suivante qui donne une interprétation intuitive de cette inégalité.



Cette relation signifie également que si la fonction f est concave, sa courbe représentative est située nécessairement sous la tangente.

Propriété 2 : Cas d'une fonction concave appartenant à $\mathcal{C}^2$

Soit f une fonction deux fois continument dérivables définie sur un convexe A.

Alors, si $H_f(X)$ est la Hessienne de f au point X (sa matrice des dérivées secondes)

- f concave $\Leftrightarrow H_f(X)$ est définie négative pour tout $X \in A$
- f convexe $\Leftrightarrow H_f(X)$ est définie positive pour tout $X \in A$.