RAPPELS MATHÉMATIQUES POUR LA MICROÉCONOMIE

ALDRIC LABARTHE - PARIS 1 PANTHÉON SORBONNE

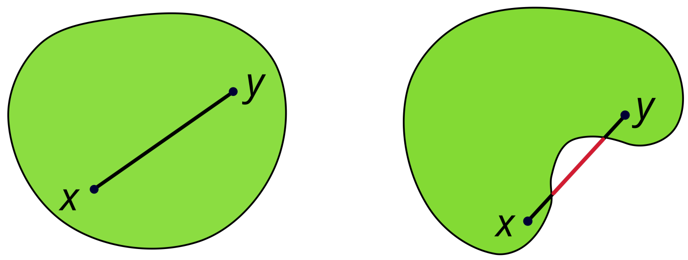
L'objectif de ce document est de proposer une synthèse des principales notions mathématiques qui seront nécessaires à la pratique de la microéconomie de niveau licence-master. Face à l'impossibilité d'être exhaustif, l'accent est placé ici sur les notions d'analyse et d'optimisation à visée pratique. A cet effet, on se bornera ci-après à une étude dans l'espace euclidien $\mathbb{R}^n$ *avec* $n \in \mathbb{N}$ *fini et fixé, muni de son produit scalaire usuel. L'attention du lecteur est attirée prioritairement sur les éléments marqués par* $\star$ *qui constituent des notions indispensables à la pratique de l'économie formalisée.*

1. Rappels théoriques

1.1. Convexité.

Définition 1 (Ensemble convexe $\star$)**.** *Un ensemble* $\mathcal{S}$ *est dit convexe ssi*

$$\forall (x, y) \in \mathcal{S}^2, \forall \lambda \in [0, 1], x \in \mathcal{S} \wedge y \in \mathcal{S} \Rightarrow \lambda x + (1 - \lambda) y \in \mathcal{S}$$



(A) Ensemble convexe (B) Ensemble non-convexe

Figure 1. **Illustration de la définition de la convexité**.

Définition 2 (Application convexe $\star$)**.** *Une application* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *(n et m finis) est:*

- *convexe, ssi* $\forall (x, y) \in (\mathbb{R}^n)^2, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda) y) \leqslant \lambda f(x) + (1 - \lambda) f(y)$
- *strictement convexe, ssi* $\forall (x, y) \in (\mathbb{R}^n)^2, \forall \lambda \in]0, 1[, f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y)$

Définition 3 (Epigraphe)**.** *Pour* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$*, on appelle épigraphe de f l'ensemble*

$$epi\, f = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}^m : f(x) \leqslant \alpha\}$$

Proposition 1. $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *est convexe* $\Leftrightarrow$ *epi* f *est un convexe de* $\mathbb{R}^n \times \mathbb{R}^m$

Définition 4 (Quasi-convexité $\star$)**.** *Une application* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *(n et m finis et fixés) est:*

- *quasi-convexe, ssi* $\forall (x, y) \in (\mathbb{R}^n)^2, \forall z \in [x, y], f(z) \leqslant \max\{f(x), f(y)\}$
- *strictement quasi-convexe, ssi* $\forall (x, y) \in (\mathbb{R}^n)^2, \forall z \in]x, y[, f(z) < \max\{f(x), f(y)\}$

Définition 5 (Ensemble de sous-niveau)**.** *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ une application, on appelle ensemble de sous-niveau α de f l'ensemble*

$$S_\alpha(f) = \{x \in \mathbb{R}^n : f(x) \leqslant \alpha\}$$

Proposition 2. *$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ est quasi-convexe $\Leftrightarrow \forall \alpha \in \mathbb{R}^m$, $S_\alpha(f)$ est un convexe de $\mathbb{R}^n$*

Démonstration. ($\Rightarrow$) Prenons $\alpha \in \mathbb{R}^m$ quelconque et considérons $S_\alpha(f) = \{x \in \mathbb{R}^n : f(x) \leqslant \alpha\}$. Prenons $(x, y) \in (S_\alpha(f))^2$ et $z = \lambda x + (1 - \lambda)y$ pour $\lambda \in [0, 1]$. On a $f(z) \leqslant \max\{f(x), f(y)\} \Rightarrow f(z) \leqslant \alpha$ d'où $f(z) \in S_\alpha(f)$. (On raisonne pour $S_\alpha(f)$ non-vide, car $\varnothing$ est convexe).

($\Leftarrow$) Prenons $(x, y) \in (\mathbb{R}^n)^2$ quelconques et $\alpha = \max\{f(x), f(y)\}$. Comme $S_\alpha(f)$ est convexe on a, $\forall z \in [x, y], f(z) \in S_\alpha(f) \Rightarrow f(z) \leqslant \alpha = \max\{f(x), f(y)\}$. $\square$

Théoreme 1. *Pour $(f, g) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$,*

- *f (strictement) convexe $\Leftrightarrow -f$ (resp. strictement) concave*
- *f et g (strictement) convexes $\Rightarrow f + g$ (resp. strictement) convexe*
- *f strictement croissante, convexe et g convexe $\Rightarrow f \circ g$ est convexe*

Remarque 1. *Attention, les règles du Théorème 1 sur les fonctions convexes ne sont en général pas valables pour les fonctions quasi-convexes.*

1.2. **Différentiabilité.**

Définition 6 (Différentiabilité et différentielle)**.** *On dit que $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ est différentiable en un point $a \in \mathbb{R}^n$ s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$, représentée par une matrice $\mathbf{D}f(a) \in \mathbb{R}^{m\times n}$ telle que :*

$$\lim_{h\to 0} \frac{\|f(a + h) - f(a) - L(h)\|}{\|h\|} = 0$$

La matrice $\mathbf{D}f(a)$ est appelée la matrice jacobienne de f en a, et elle représente la meilleure approximation linéaire de f autour de a. Elle s'écrit comme:

$$\mathbf{D}f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

où f_i représente la i-ème composante de l'application f. $\frac{\partial f_i}{\partial x_j}(a)$ désigne la dérivée partielle de f_i par rapport à la variable x_j en a. Cette dérivée partielle mesure le taux de variation de f_i en fonction de x_j lorsque les autres variables sont maintenues fixes. Formellement:

$$\frac{\partial f}{\partial x_j}(a) = \lim_{h\to 0} \frac{f(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

Théoreme 2. *Soient $f : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^p \to \mathbb{R}^m$ deux applications différentiables. Si f est différentiable en un point $a \in \mathbb{R}^n$, et g est différentiable en $f(a) \in \mathbb{R}^p$, alors la fonction composée $h = g \circ f : \mathbb{R}^n \to \mathbb{R}^m$, définie par $h(x) = g(f(x))$, est différentiable en a, et sa dérivée est donnée par $\mathbf{D}h(a) = \mathbf{D}g(f(a)) \cdot \mathbf{D}f(a)$.*

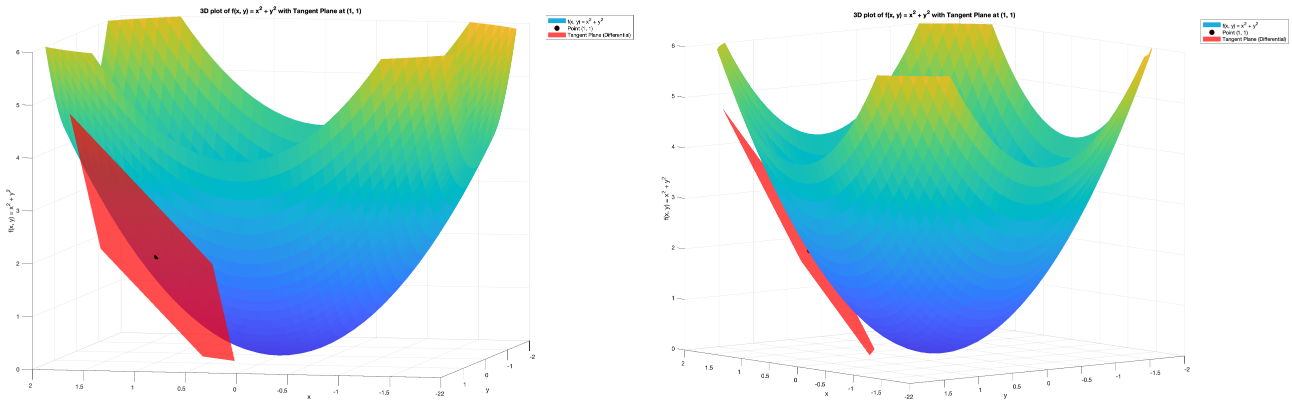


FIGURE 2. **Illustration de l'approximation induite par la différentielle d'une fonction en un point**. Avec $f : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad f(x, y) = x^2 + y^2$, différentiée en $(1, 1)$. Dans $\mathbb{R}^2$, une application linéaire est représentée par un hyperplan.

Remarque 2. *Ce théorème, qui peut apparaître difficile en première lecture, n'est ni plus ni moins qu'une transposition de la règle de dérivation en chaîne, dont l'énoncé dans l'espace des fonctions réelles $(f, u) \in (\mathcal{C}^1)^2$ est usuellement connu sous la forme, pour tout x du domaine de g, $(f \circ g)'(x) = g'(x) \cdot (f' \circ g)(x)$.*

Remarque 3. *Dans le cadre de ce cours, le problème de la différentiabilité d'une application ne sera étudié qu'à travers le théorème de composition. Ainsi, pour déterminer si une application dont on connaît une expression analytique est différentiable, on se reportera aux fonctions usuelles qui la composent, et pour lesquelles on admettra la différentiabilité. On se rappellera par exemple de la différentiabilité sur tout le domaine de définition de* $\ln, \exp$ *ou de toutes les fonctions polynomiales, ainsi que de la linéarité de l'opérateur de différentiabilité.*

Remarque 4. *Le calcul différentiel devient particulièrement lourd en notations lorsque l'on considère des applications de $\mathbb{R}^n$ dans $\mathbb{R}^m$. En microéconomie, il est davantage d'usage de se borner au cas où f est une fonction, c'est à dire $m = 1$, ce que nous supposerons donc désormais.*

Définition 7 (Gradient $\star$)**.** *Pour $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, on appelle gradient de f au point $a \in \mathbb{R}^n$ le vecteur:*

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_i}(a) \right)_{i \leqslant n}$$

On remarque que $\nabla f(a)$ est l'écriture quand $m = 1$ de $\mathbf{D}f(a)$.

Définition 8 (Hessienne $\star$)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction deux fois continûment différentiable. La matrice hessienne de f en un point $a \in \mathbb{R}^n$, notée $\mathbf{H_f}(a)$, est la matrice carrée d'ordre n dont les éléments sont les dérivées secondes partielles de f en a. Elle est donnée par :*

$$\mathbf{H_f}(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

Chaque élément $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ représente la dérivée seconde partielle croisée de f par rapport aux variables x_i et x_j.

Définition 9 (Matrice définie positive). *Pour tout $n > 0$ fini, une matrice réelle $H \in \mathcal{M}(n,n)$ symétrique est:*

- *définie positive, si pour tout $x \in \mathbb{R}^n$ tel que $x \neq 0_{\mathbb{R}^n}$, $x^t H x > 0$*
- *positive, si pour tout $x \in \mathbb{R}^n$, $x^t H x \geqslant 0$*

Proposition 3. *Pour tout $n > 0$ fini, une matrice réelle $H \in \mathcal{M}(n,n)$ symétrique est (définie) positive, ssi toutes ses valeurs propres sont (resp. strictement) positives.*

Théoreme 3 ($\star$). *Si $f : \mathbb{R}^n \to \mathbb{R}$ une fonction deux fois continûment différentiable alors*

$$\begin{cases} \mathbf{H_f} \text{ définie positive} \Leftrightarrow f \text{ strictement convexe} \\ \mathbf{H_f} \text{ semi-définie positive} \Leftrightarrow f \text{ convexe} \end{cases}$$

1.3. **Optimisation.**

Remarque 5. *Dans cette section, nous nous concentrons sur les problèmes de minimisation, afin de rester cohérent avec notre caractérisation des fonctions convexes. Le lecteur attentif remarquera que minimiser une fonction f revient à maximiser -f.*

Définition 10 (Minimiseur $\star$). *Pour $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une application, $\hat{x}$ est un minimiseur de f sur $\mathcal{C} \subset \mathbb{R}^n$ ssi $\forall x \in \mathcal{C}, f(x) \geqslant f(\hat{x})$.*

Théoreme 4 ($\star$). *Si $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est deux fois continûment différentiable, alors l'ensemble des points qui minimisent f sur $\mathcal{C} \subset \mathbb{R}^n$ ($\mathcal{C}$ ouvert) est l'ensemble $\{\hat{x} \in \mathcal{C} : \nabla f(\hat{x}) = 0\}$.*

Remarque 6. *On dira que $\hat{x}$ est un minimiseur global de f ssi $\mathcal{C}$ est égal au domaine de f.*

Définition 11 (Programme d'optimisation $\star$). *On appelle problème d'optimisation contraint (en l'occurrence, minimisation) la recherche des points $\hat{x}$ tels que $\forall x \in \mathcal{C}, f(x) \geqslant f(\hat{x})$ avec $\mathcal{C} = \{x \in \mathbb{R}^n : g(x) = 0 \wedge h(x) \leqslant 0\}$. On appellera $\mathcal{C}$ l'ensemble réalisable (feasible set) et dirons que l'on minimise f sous les contraintes g et h (avec $f, g, h : \mathbb{R}^n \longrightarrow \mathbb{R}$). On écrira:*

$$\begin{cases} \textit{argmin } f(x) \\ s.c \begin{cases} g(x) = 0 \\ h(x) \leqslant 0 \end{cases} \end{cases}$$

Remarque 7 ($\star$). *Il est important de garder en mémoire la différence entre argmin et min (resp. max et argmax). Le premier consiste à trouver $\hat{x}$ tel que $f(\hat{x}) \leqslant f(x)$ pour tous les x du domaine considéré, là où le second consiste à trouver la valeur atteinte par f au point $\hat{x}$. En tant que microéconomistes, nous ne travaillons quasi-exclusivement avec argmin.*

Remarque 8 ($\star$). *La recherche de minimiseurs contraints de f est une tâche qu'il ne faut pas confondre avec la recherche de minimiseurs non contraints. Si effectivement, on peut considérer qu'une fois l'ensemble réalisable $\mathcal{C}$ défini, le problème peut parfois se réduire à la recherche d'un minimiseur local sur un ouvert, il est essentiel de garder en tête que* ***dans de très nombreux cas, en optimisation contrainte, il n'existe pas de** $x \in \mathcal{C}$ **tels que*** $\nabla f(x) = 0$*. Nous proposons donc la création d'une nouvelle fonction $\mathcal{L}$, pour laquelle les solutions du problème contraint de f vérifieront toutes $\nabla \mathcal{L}(x) = 0$.*

Définition 12 (Lagrangien $\star$)**.** *Pour $f, g, h : \mathbb{R}^n \longrightarrow \mathbb{R}$ formant le programme d'optimisation contraint,*

$$\begin{cases} \textit{argmin } f(x) \\ s.c \begin{cases} g(x) = 0 \\ h(x) \leqslant 0 \end{cases} \end{cases}$$

On appelle ***lagrangien*** *l'application $\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ définie par*

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda(g(x)) + \mu(h(x))$$

.

Remarque 9. *Malheureusement, nous ne pourrons dans le cadre de ce cours étendre le formalisme au delà de cette définition. A partir de maintenant, nous présentons quelques résultats non-rigoureux qui seront néanmoins utiles à l'analyse microéconomique. Le lecteur intéressé par les problèmes d'optimisation est encouragé à questionner l'existence d'une solution, et explorer le champ des conditions de qualification des contraintes. De même, nous supposerons à présent que toutes les contraintes sont de la forme $g(x) = 0$, c'est à dire qu'il n'y a pas de contraintes non-saturées. Nous nous bornons donc à*

$$\begin{cases} \textit{argmin } f(x) \\ s.c \ g(x) = 0 \end{cases} \qquad \mathcal{L}(x, \lambda) = f(x) + \lambda(g(x))$$

Proposition 4 ($\star$)**.** *Si f est continûment différentiable, tout minimiseur $\hat{x}$ de f sous la contrainte g vérifie $\nabla\mathcal{L}(\hat{x}) = 0$. Si f est convexe, alors on considèrera que tout point vérifiant $\nabla\mathcal{L}(x) = 0$ est un minimiseur local.*

Proposition 5 ($\star$)**.** *Si f est strictement convexe sur son domaine de définition, alors tout problème de minimisation de f n'admet qu'une seule solution.*

2. Exemple d'optimisation

On propose d'appliquer toutes ces notions à l'optimisation d'une fonction d'utilité CES (Constant elasticity of substitution):

$$\rho \leqslant 1, \quad U : \mathbb{R}_+^2 \longrightarrow \mathbb{R}, \quad U(x_1, x_2) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}}$$

Economiquement, il convient d'imposer que la consommation demeure dans l'ensemble économique réalisable, c'est-à-dire que la contrainte budgétaire $R \leqslant p_1x_1 + p_2x_2$ soit satisfaite. Par l'axiome d'insatiabilité des préférences, on a que le consommateur saturera la contrainte à l'équilibre, on peut donc la convertir en contrainte d'égalité et écrire le programme d'optimisation contraint:

$$\begin{cases} \text{argmax}_{x_1,x_2} U(x_1, x_2) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}} \\ \text{s.c. } R = p_1x_1 + p_2x_2 \end{cases}$$

On écrit alors le lagrangien associé:

$$\mathcal{L}(x_1, x_2, \lambda) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}} + \lambda(R - p_1x_1 - p_2x_2)$$

On observe que U est de classe $\mathcal{C}^2(\mathbb{R}_+^2)$ car elle est la composition de deux monomes, et d'une fonction puissance, elles mêmes de classe $\mathcal{C}^\infty$ sur le domaine. On peut donc rechercher les points critiques (condition du premier ordre) de $\mathcal{L}$:

$$\nabla\mathcal{L} = 0 \Leftrightarrow \begin{cases} \frac{\partial\mathcal{L}}{\partial x_1} = 0 \\ \frac{\partial\mathcal{L}}{\partial x_2} = 0 \\ \frac{\partial\mathcal{L}}{\partial \lambda} = 0 \end{cases} \Leftrightarrow \begin{cases} x_1^{\rho-1}\ (x_1^\rho + x_2^\rho)^{\frac{1-\rho}{\rho}} = \lambda p_1 \\ x_2^{\rho-1}\ (x_1^\rho + x_2^\rho)^{\frac{1-\rho}{\rho}} = \lambda p_2 \\ R = p_1x_1 + p_2x_2 \end{cases}$$

$$\Rightarrow \left(\frac{x_i}{x_j}\right)^{\rho-1} = \frac{p_i}{p_j} \Leftrightarrow x_i = x_j\left(\frac{p_i}{p_j}\right)^{\frac{1}{\rho-1}} \qquad (i,j) \in \{(1,2),(2,1)\}$$

$$\Rightarrow R = p_ix_j\left(\frac{p_i}{p_j}\right)^{\frac{1}{\rho-1}} + p_jx_j = x_j\left(p_i^{\frac{\rho}{\rho-1}}p_j^{\frac{-1}{\rho-1}} + p_j\right)$$

$$\Rightarrow x_j = \frac{R}{p_i^{\frac{\rho}{\rho-1}}p_j^{\frac{-1}{\rho-1}} + p_j} \qquad (\rho \neq 1)$$

Nous obtenons donc un candidat maximiseur. Pour appliquer la Proposition 4, il convient de vérifier la concavité (il s'agit d'un problème de maximisation) de la fonction objectif U. **En pratique, dans ce cours, on écrira fréquemment que l'on *"suppose la condition du second ordre vérifiée"*, et nous arrêterons à cette étape.** Toutefois, le lecteur intéressé remarquera que cet argument est pleinement incomplet car le point candidat n'est en l'état aucunement garanti d'être réellement un maximiseur. On propose donc un exemple de justification, pour le lecteur intéressé dans la poursuite de l'économie formalisée.

On peut montrer que $U(x_1, x_2) = h(g(x_1), g(x_2))$ avec $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$ et $g : \mathbb{R} \longrightarrow \mathbb{R}$ telles que $h(h_1, h_2) = (h_1 + h_2)^{\frac{1}{\rho}}$, et $g(x) = x^\rho$. Il est très aisé de voir que g est concave pour $\rho < 1$, $g'(x) = \rho x^{\rho-1}$, $g''(x) = \rho(\rho-1)x^{\rho-2} < 0$ (pour $x \geqslant 0$, ce qui est le cas ici). La concavité de h est un peu plus longue à établir:

$$\frac{\partial h}{\partial h_i} = \frac{1}{\rho}(h_i + h_j)^{\frac{1-\rho}{\rho}} \geqslant 0$$

$$\frac{\partial^2 h}{\partial h_i^2} = \frac{\partial^2 h}{\partial h_i \partial h_j} = \frac{1-\rho}{\rho^2}(h_i + h_j)^{\frac{1-2\rho}{\rho}}$$

$$\mathbf{H}_g = \begin{pmatrix} \frac{\partial^2 h}{\partial h_1^2} & \frac{\partial^2 h}{\partial h_1 \partial h_2} \\ \frac{\partial^2 h}{\partial h_2 \partial h_1} & \frac{\partial^2 h}{\partial h_2^2} \end{pmatrix} = \frac{1-\rho}{\rho^2}(h_i + h_j)^{\frac{1-2\rho}{\rho}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

La matrice composée que de 1 n'a pour valeur propre que 1. D'où, $\mathbf{H}_g$ n'a comme valeur propre que $\frac{1-\rho}{\rho^2}(h_i + h_j)^{\frac{1-2\rho}{\rho}}$ qui est négative pour $\rho < 1$. D'où h est concave et croissante. Comme g est aussi concave, on peut affirmer que U est bien concave par le théorème de composition, confirmant que la solution trouvée est bien un maximum de U pour $\rho < 1$. Quand $\rho = 1$, U est linéaire.

On peut aussi s'intéresser au caractère substituable des biens considérés. On dérive pour cela x_j par rapport à p_i:

$$\frac{\partial x_j^*}{\partial p_i} = \frac{-\rho R p_i^{\frac{1}{\rho-1}} p_j^{\frac{-1}{\rho-1}}}{(\rho-1)\left(p_i^{\frac{\rho}{\rho-1}} p_j^{\frac{-1}{\rho-1}} + p_j\right)^2} < 0 \quad (p_i > 0, p_j > 0, R > 0)$$