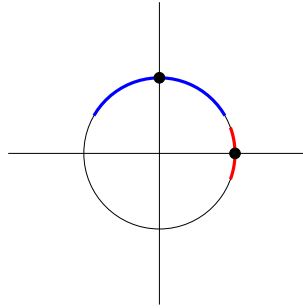


Analyse dans \mathbf{R}^n

Chapitre no. 4, Fonctions implicites, extrema liés

On considère la fonction $f(x, y) = x^2 + y^2 - 1$ définie dans le plan \mathbf{R}^2 . L'ensemble des points définie par $f(x, y) = 0$ correspond au cercle unité :



Autour du point $(0, 1)$, l'équation détermine y en fonction de x : $y = +\sqrt{1-x^2}$; plus précisément, donné le point $(0, 1)$ vérifiant $f(0, 1) = 0$, il existe un voisinage de $x = 0$, disons $]-\varepsilon, +\varepsilon[$ tel que pour $x \in]-\varepsilon, +\varepsilon[$, on a $y = +\sqrt{1-x^2}$. Ce n'est plus le cas si on choisit le point $(1, 0)$. En effet pour $x > 1$, y n'est pas défini, et pour $x < 1$, on a deux choix de racines, soit $+\sqrt{1-y^2}$, soit $-\sqrt{1-y^2}$. On remarque que le gradient de f :

$$\text{grad } f = 2 \begin{pmatrix} x \\ y \end{pmatrix}$$

est horizontal au point $(1, 0)$, c'est à dire le niveau $f(x, y) = 0$ est vertical.

En général, pour une fonction $f(x, y)$, l'équation $f(x, y) = 0$ détermine $y = y(x)$ dans un voisinage de x_0 , où (x_0, y_0) vérifie $f(x_0, y_0) = 0$, si $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

Exemple : Soit $f(x, y) = x^2y + xy^2 + xy + 1$. Le point $(-1, 1)$ vérifie $f(-1, 1) = 0$. La dérivée partielle $\partial f / \partial y = x^2 + 2xy + x$ vaut -2 au point $(-1, 1)$, donc l'équation $f(x, y) = 0$ détermine $y = y(x)$ implicitement en fonction de x dans un voisinage de $x = -1$. On peut effectivement résoudre l'équation quadratique en y : $xy^2 + (x^2 + x)y + 1$:

$$y = \frac{-(x^2 + x) \pm \sqrt{(x^2 + x)^2 - 4x}}{2x}$$

Pour savoir quelle racine il faut choisir, on doit substituer le point $x = -1$. Dans ce cas $y = \pm\sqrt{4}/(-2)$ et, afin que $y = 1$ en $x = -1$, on doit choisir la racine négative :

$$y = \frac{-(x^2 + x) - \sqrt{(x^2 + x)^2 - 4x}}{2x}$$

Le cadre général est donné par le théorème suivant.

Théorème des fonctions implicites : On suppose $n > p$ et on écrit $\mathbf{R}^n = \mathbf{R}^{n-p} \times \mathbf{R}^p$; soit $x = (x_1, \dots, x_{n-p})$ coordonnées sur \mathbf{R}^{n-p} , et $y = (y_1, \dots, y_p)$ coordonnées sur \mathbf{R}^p . Soit $U \times V$ un ouvert dans $\mathbf{R}^{n-p} \times \mathbf{R}^p = \mathbf{R}^n$ et soit $f : U \times V \rightarrow \mathbf{R}^p$ une application de classe C^1

telle que $f(a_1, \dots, a_{n-p}, b_1, \dots, b_p) = 0$ et le jacobien $\det \frac{\partial(f_1, \dots, f_p)}{\partial(y_1, \dots, y_p)}$ est non-nul au point $(a_1, \dots, a_{n-p}, b_1, \dots, b_p) \in U \times V$. Alors, il existe un ouvert connexe $W \subset U$ de (a_1, \dots, a_{n-p}) , et p fonctions scalaires uniques $g_i(x)$ ($1 \leq i \leq p$), définies et continues dans W tel que $g_i(a_1, \dots, a_{n-p}) = b_i$ ($1 \leq i \leq p$), et

$$f_i(x_1, \dots, x_{n-p}, g_1(x_1, \dots, x_{n-p}), \dots, g_p(x_1, \dots, x_{n-p})) = 0$$

pour $i = 1, \dots, p$ et pour tout $x = (x_1, \dots, x_{n-p}) \in W$. De plus, les fonctions g_i sont de classe C^1 dans W , et la matrice jacobienne $\frac{\partial(g_1, \dots, g_p)}{\partial(x_1, \dots, x_{n-p})}$ est donnée par $-B^{-1}A$, où A (resp. B) est obtenue en remplaçant y_i par $g_i(x_1, \dots, x_{n-p})$ dans la matrice jacobienne $\frac{\partial(f_1, \dots, f_p)}{\partial(x_1, \dots, x_{n-p})}$ (resp. $\frac{\partial(f_1, \dots, f_p)}{\partial(y_1, \dots, y_p)}$).

Remarque 1. On peut économiser la notation en écrivant $g = \begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix} : W \rightarrow \mathbf{R}^p$,

$a = (a_1, \dots, a_{n-p})$ et $b = (b_1, \dots, b_p)$, d'où on a $g(a) = b$ et $f(x, g(x)) = 0$ ($x \in W$), pour la conclusion du théorème, ainsi que la formule pour la matrice jacobienne : $J_g(x) = -J_2(g(x))^{-1}J_1(x)$, où $J_1(x)$ est la matrice jacobienne de f par rapport à la variable x , et $J_2(y)$ la matrice jacobienne de f par rapport à la variable y .

2. Lorsque $n = 2$ et $p = 1$, on obtient les conclusions de la discussion de l'introduction à ce chapitre : soit $f(x, y)$ une fonction de deux variables, et soit (a, b) un point vérifiant $f(a, b) = 0$ et $(\partial f / \partial y)(a, b) \neq 0$; alors il existe un voisinage W de a dans $U \subset \mathbf{R}$, et une fonction $g : W \rightarrow \mathbf{R}$ telle que $g(a) = b$ et $f(x, g(x)) = 0$. La conclusion pour la jacobienne :

$$\frac{\partial g}{\partial x}(x) = - \left(\frac{\partial f}{\partial y}(x, g(x)) \right)^{-1} \frac{\partial f}{\partial x}(x, g(x))$$

se déduit en dérivant l'équation $f(x, g(x)) = 0$ par rapport à x :

$$0 = \frac{\partial}{\partial x} f(x, g(x)) = \frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x)) \frac{\partial g}{\partial x}(x).$$

3. Pour démontrer le théorème des fonctions implicites, on peut associer à l'application f , une application $F : U \times V \rightarrow \mathbf{R}^n$ dont la dérivée est un isomorphisme au point (a, b) , puis on applique le théorème de l'application inverse $g(x)$ (voir la référence 1, pour plus de détails).

Pour simplicité, on suppose $n = 2$ et $p = 1$. Soit $F(x, y) = (x, f(x, y))$. La matrice jacobienne de F en (a, b) est donnée par

$$J_F(a, b) = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x}(a, b) & \frac{\partial f}{\partial y}(a, b) \end{pmatrix}$$

qui est de rang 2 par l'hypothèse que $\frac{\partial f}{\partial y}(a, b) \neq 0$. Par le théorème de l'application inverse, il existe une inverse locale $G : B \rightarrow A$ de classe C^1 , où B est un voisinage de $(a, 0)$

et A un voisinage de (a, b) . Alors $G(x, 0) = (x, g(x))$ détermine la fonction $y = g(x)$:
 $(x, 0) = F(G(x, 0)) = F(x, g(x)) = (x, f(x, g(x)))$.

Exemple : Soit $f : \mathbf{R}^4 \rightarrow \mathbf{R}^3$ définie par

$$f(x_1, y_1, y_2, y_3) = (x_1^2 + y_1^2 - y_2^2 - y_3^2, x_1 y_2 + y_1 y_3, -x_1 y_3 + y_1 y_2) - (2, 0, 0)$$

On voit que $f(1, 1, 0, 0) = 0$. La matrice jacobienne

$$\frac{\partial(f_1, f_2, f_3)}{\partial(y_1, y_2, y_3)} = \begin{pmatrix} 2y_1 & -2y_2 & -2y_3 \\ y_3 & x_1 & y_1 \\ y_2 & y_1 & -x_1 \end{pmatrix}$$

est de rang 3 au point $(1, 1, 0, 0)$. D'après le théorème de la fonction implicite, il existe trois fonctions $g_1(x_1), g_2(x_1), g_3(x_1)$ vérifiant

$$f(x_1, g_1(x_1), g_2(x_1), g_3(x_1)) = 0$$

Pour les trouver, on résout

$$\begin{cases} x_1^2 + y_1^2 - y_2^2 - y_3^2 = 2 \\ x_1 y_2 + y_1 y_3 = 0 \\ -x_1 y_3 + y_1 y_2 = 0 \end{cases}$$

pour y_i en fonction de x_1 . On multiplie la 2ème équation par y_2 et la 3ème par y_3 , puis on soustrait pour en déduire que $y_2 = y_3 = 0$, puis $x_1^2 + y_1^2 = 2$ avec solution $y_1 = \sqrt{2 - x_1^2}$ dans un voisinage de $(x_1, y_1, y_2, y_3) = (1, 1, 0, 0)$. Il s'agit de l'arc d'un cercle de rayon $\sqrt{2}$. Donc $g(x_1) = (\sqrt{2 - x_1^2}, 0, 0)$.

Exercice : Vérifier la formule pour la matrice jacobienne de g .

Exemple : Soit $f(x, y, z) = x^2 + y^2 - z^2 - 1$. Alors $f(1, 1, 1) = 0$ et $\frac{\partial f}{\partial z} = -2z$ est non-nulle au point $(1, 1, 1)$. Par le théorème de la fonction implicite, l'équation $f(x, y, z) = 0$ détermine z en fonction de x, y dans un voisinage de $(1, 1, 1)$. Explicitement $z = \sqrt{x^2 + y^2 - 1}$, avec racine positive. La formule pour la matrice jacobienne :

$$\frac{\partial z}{\partial(x, y)} = - \left(\frac{\partial f}{\partial z} \right)^{-1} \frac{\partial f}{\partial(x, y)} = \frac{1}{2z} \begin{pmatrix} 2x & 2y \end{pmatrix}$$

est bien confirmée par l'expression pour $z(x, y)$.

Fibres d'une application. Soit $f : U \rightarrow \mathbf{R}^p$ ($n \geq p$) une application C^1 , et soit $y \in \mathbf{R}^p$. On appelle *fibres de f sur y* l'ensemble $f^{-1}(y) = \{x \in U : f(x) = y\}$. Bien évidemment, il faut que $y \in f(U)$ afin que $f^{-1}(y)$ soit non-vidé. On dit que $x \in U$ est un *point régulier* si le rang de $Df(x)$ est égale à p (sinon c'est un point critique, ou singulier). Un fibre est dit régulier si tous ses points sont réguliers. Le théorème des fonctions implicites entraînent qu'on peut paramétrer localement un fibre régulier par des ouverts de \mathbf{R}^{n-p} .

Remarque : Dans le théorème des fonctions implicites, il n'y a rien de spéciales en ce qui concerne les premières $n - p$ coordonnées. En effet, si le rang de Df est p , il existe p coordonnées (qu'on a appelées y_1, \dots, y_p dans le théorème), par rapport auxquelles le déterminant de la matrice jacobienne soit non-nul. Par exemple, si on prend l'exemple $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ donnée par $f(x, y) = x^2 + y^2 - 1$, alors dans un voisinage de $(0, 1)$, on peut paramétrer le fibre $f^{-1}(0)$ par $y = \sqrt{1 - x^2}$, dans un voisinage de $(1, 0)$ par $x = \sqrt{1 - y^2}$, dans un voisinage de $(0, -1)$ par $y = -\sqrt{1 - x^2}$, et dans un voisinage de $(-1, 0)$ par $x = -\sqrt{1 - y^2}$. En effet, tous les points de $f^{-1}(0)$ sont réguliers, et le fibre est une courbe localement paramétrée par des intervalles de \mathbf{R} . Par contre, si un fibre n'est pas régulier, ce n'est plus le cas qu'on a nécessairement une paramétrisation par des ouverts de \mathbf{R}^{n-p} .

Par exemple, soit $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ donnée par $f(x, y) = xy$. Pour $c \in \mathbf{R} \setminus \{0\}$, $f^{-1}(0)$ est un hyperbole, mais $f^{-1}(0)$ est la réunion des deux axes : $x = 0$ et $y = 0$. L'origine est un point critique et le fibre autour de ce point n'admet pas de paramétrisation par un intervalle.

Extréma liés. On cherche à maximiser ou minimiser une fonction $f(x_1, \dots, x_n)$ lorsque les variables x_1, \dots, x_n sont liées par certaines relations $g_1(x_1, \dots, x_n) = g_2(x_1, \dots, x_n) = \dots = g_p(x_1, \dots, x_n) = 0$ ($x \in U \subset \mathbf{R}^n$ ouvert). On remarque que

$$g := \begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix} : U \rightarrow \mathbf{R}^p$$

et que les relations imposent que les points appartiennent au fibre $g^{-1}(0)$.

Si $A := g^{-1}(0) = \{x \in U : g_1(x) = \dots = g_p(x) = 0\}$, et $a \in U$ est un extremum de $f|_A$ (la restriction de f à A), on dit que $f|_A$ admet un extremum relatif en $a \in A$.

Théorème : Soit U un ouvert non-vidé de \mathbf{R}^n , $p \geq 1$, et f, g_1, g_2, \dots, g_p des fonctions C^1 définies sur U . On pose $A = \{x \in U : g_1(x) = \dots = g_p(x) = 0\}$. Si $f|_A$ admet un extremum relatif en $a \in A$ et si les gradients $\text{grad } g_1, \dots, \text{grad } g_p$ sont linéairement indépendants en a , alors il existe des réels $\lambda_1, \dots, \lambda_p$ tels que

$$\text{grad } f(a) = \sum_{j=1}^p \lambda_j \text{grad } g_j(a)$$

Preuve : Comme pour le théorème des fonctions implicites, on écrit $(x, y) = (x_1, \dots, x_{n-p}, y_1, \dots, y_p)$.

Si le jacobien $\det \frac{\partial(g_1, \dots, g_p)}{\partial(y_1, \dots, y_p)}$ est non-nul en un point $a \in \mathbf{R}^{n-p} \times \mathbf{R}^p$, on peut trouver une fonction $\varphi(x)$ telle que $g(x, \varphi(x)) = 0$. Il s'ensuit que dans un voisinage de a , la restriction de f à l'ensemble $A = \{(x, y) : g(x, y) = 0\}$ est donnée par la fonction de $x = (x_1, \dots, x_{n-p})$: $f(x, \varphi(x))$. Afin que a soit un point critique de f restreinte aux points liés par les relations $g(x, y) = 0$, il faut que les dérivées partielles $\frac{\partial}{\partial x_i} f(x, \varphi(x))$ s'annulent :

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^p \frac{\partial f}{\partial y_j} \frac{\partial \varphi_j}{\partial x_i} = 0 \quad (1)$$

pour chaque $i = 1, \dots, n - p$ où on écrit $\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_p \end{pmatrix}$.

En dérivant $g(x, \varphi(x)) = 0$, on obtient

$$\frac{\partial g}{\partial x_i} + \sum_{j=1}^p \frac{\partial g}{\partial y_j} \frac{\partial \varphi_j}{\partial x_i} = 0 \quad (2)$$

pour chaque $i = 1, \dots, n - p$. Introduisons la matrice

$$B = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \cdots & \frac{\partial f}{\partial x_{n-p}}(a) & \frac{\partial f}{\partial y_1}(a) & \cdots & \frac{\partial f}{\partial y_p}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_{n-p}}(a) & \frac{\partial g_1}{\partial y_1}(a) & \cdots & \frac{\partial g_1}{\partial y_p}(a) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_p}{\partial x_1}(a) & \cdots & \frac{\partial g_p}{\partial x_{n-p}}(a) & \frac{\partial g_p}{\partial y_1}(a) & \cdots & \frac{\partial g_p}{\partial y_p}(a) \end{pmatrix}$$

Les équations (1) et (2)] montrent que les $n - p$ premières collones de B s'expriment linéairement en fonctions des p dernières, d'où le rang de B est inférieur ou égal à p . Il s'ensuit que les $p + 1$ lignes de B sont liées. Il existe donc des réels non tous nuls $\alpha_1, \dots, \alpha_{p+1}$ tels que

$$\alpha_1 \text{grad } g_1(a) + \cdots + \alpha_p \text{grad } g_p(a) + \alpha_{p+1} \text{grad } f(a) = 0. \quad (3)$$

Comme la famille $\{\text{grad } g_1(a), \dots, \text{grad } g_p(a)\}$ est indépendante, on a nécessairement $\alpha_{p+1} \neq 0$. Pour tout $i = 1, \dots, p$, posons alors $\lambda_i = -\alpha_i/\alpha_{p+1}$. Dans ce cas, l'égalité (3) conduit exactement à $\text{grad } f(a) = \sum_{i=1}^p \lambda_i \text{grad } g_i(a)$. \square

On appelle les constantes $\lambda_1, \dots, \lambda_p$ *multiplicateurs de Lagrange*.

Exemple : On suppose $(x, y, z) \in \mathbf{R}^3$ vérifie les deux équations $x^2 + y^2 = 1$ et $x - 2y + 3z = 0$. Calculer la distance maximal du point (x, y, z) au dessus du xy -plan.

Soit $A = \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 = 1 \text{ et } x - 2y + 3z = 0\}$. On veut trouver le maximum de z relatif. Soit $g_1(x, y, z) = x^2 + y^2 - 1$ et $g_2(x, y, z) = x - 2y + 3z$. En un point extremal $a \in A$, on a

$$\text{grad } z(a) = \lambda_1 \text{grad } g_1(a) + \lambda_2 \text{grad } g_2(a)$$

c'est à dire

$$(0, 0, 1) = \lambda_1(2x, 2y, 0) + \lambda_2(1, -2, 3)$$

D'où $\lambda_2 = \frac{1}{3}$ et $2\lambda_1(x, y, 0) = (-\frac{1}{3}, \frac{2}{3}, 0)$. C'est à dire $x/y = -1/2 \Rightarrow y = -2x$. Mais $g_1(x, y, z) = 0 \Rightarrow 5x^2 = 1$ et $x = \pm 1/\sqrt{5}$. D'après $g_2(x, y, z) = 0$, on en déduit que

$$3z = -x + 2y = -5x = \mp \frac{5}{\sqrt{5}} = \mp \sqrt{5}$$

Pour le maximum, on trouve alors $z = +\sqrt{5}/3$, qui est achevé au point $(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, \frac{\sqrt{5}}{3})$.

Exemple. Montrer l'inégalité arithmético-géométrique : pour tous réels x_1, \dots, x_n positifs :

$$(x_1 \cdots x_n)^{1/n} \leq \frac{x_1 + \cdots + x_n}{n}$$

6

Définissons $f : \mathbf{R}^n \rightarrow \mathbf{R}$ par $f(x_1, \dots, x_n) = x_1 \cdots x_n$. Soit $s > 0$ et considérons l'application $g : \mathbf{R}^n \rightarrow \mathbf{R}$ définie par $g(x_1, \dots, x_n) = x_1 + \cdots + x_n - s$. Posons $K_s := \{x \in (\mathbf{R}(\geq 0))^n : g = 0\}$ et soit U l'ouvert $(\mathbf{R}(> 0))^n$. Soit A l'ensemble $\{x \in U : g(x) = 0\} \subset K_s$.

L'ensemble K_s est fermé et borné, donc compact dans \mathbf{R}^n . Puisque f est continue, elle présente un maximum en un point $a \in K_s$. On remarque que $f = 0$ sur $K_s \setminus A$. Puisque $f(a) \geq f(\frac{s}{n}, \dots, \frac{s}{n}) > 0$, on voit que $a \in A$. Si $a = (a_1, \dots, a_n)$, il s'ensuit que $a_i \neq 0$ pour tout $i = 1, \dots, n$.

Or f et g sont de classe C^1 sur U , d'après le théorème des extrema liés, il existe $\lambda \in \mathbf{R}$ tel que $\text{grad } f(a) = \lambda \text{grad } g(a)$, c'est à dire pour chaque i on a $\frac{\partial f}{\partial x_i}(a) = \frac{\partial g}{\partial x_i}(a)$, ce qui entraîne

$$\forall i = 1, \dots, n, \frac{f(a)}{a_i} = \lambda$$

Mais $f(a) > 0$ et donc $\lambda \neq 0$. On en déduit que les a_i sont tous égaux à la constante $C = f(a)/\lambda$.

Mais $a \in A$, ce qui donne $\sum_{i=1}^n a_i - s = 0$, soit $nC - s = 0$.

Enfin, sur K_s , f atteint son maximum en $(\frac{s}{n}, \dots, \frac{s}{n})$. En d'autres termes :

$$\forall x \in K_s, f(x) \leq f\left(\frac{s}{n}, \dots, \frac{s}{n}\right) = \left(\frac{s}{n}\right)^n$$

En appliquant cette inégalité à $s = \sum_{i=1}^n x_i > 0$ et $x \in K_s$, il vient ainsi

$$(f(x))^{1/n} \leq \frac{1}{n} \sum_{i=1}^n x_i$$

Références 1. J. Dieudonné, Foundations of Modern Analysis, Academic Press 1969.

2. J. Lelong-Ferrand et J. M. Arnaudiès, Cours de Mathématiques Tome 2 : Analyse, 4ème édition, Dunod Université 1977.