Calcul Différentiel

Cours avec Exerices Corrigés

Département de Mathématiques

Université de Batna 2

Dr. El Amir DJEFFAL
l.djeffal@univ-batna2.dz

2020

Contents

1 Calcul Différentiel **2**
1.1 Dérivées. Matrice jacobienne. Gradient 2
1.2 Propriétés des dérivées partielles. 3
1.3 Derivées partielles d'ordre supérieur. Fonctions de classe C^k. Théorème de Schwarz 5
1.4 Différentielle 5

2 Propriétés géométriques des fonctions de plusieurs variables **9**
2.1 Dérivée directionnelle 9
2.2 Gradient. 10
2.2.1 **Propriété [a] : Le gradient est perpendiculaire à la ligne de niveau** 10
2.2.2 **Propriété [b] : Le gradient indique la ligne de plus grande pente** 11

2.3 Formule de Taylor 11
2.4 Vecteur normal et plan tangent à un graphe d'une fonction de 2 variables 13

3 Extrema **15**
3.1 Extrema locaux et globaux. Définition 15
3.2 Théorème des extrema sur un compact 16
3.3 Extrema de fonctions de 2 variables - critère par le déterminant de matrice Hessienne 17
3.4 Extrema liés 19
3.5 Extrema d'une fonction de $n > 2$ variables 21

4 Exercices corrigés **22**

1 Calcul Différentiel

1.1 Dérivées. Matrice jacobienne. Gradient

Rappel. Soit $f : I \to$ une fonction dérivable sur un intervalle $I \in$. La dérivée de f au point $a \in I$ est :

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} \tag{1}$$

Soit $f :\ D \subset R^p \to R^q$ et $A \in D$. Une expression du type " $\lim\limits_{X\to A} \dfrac{f(X) - f(A)}{X - A}$ " n'est pas bien définie parce que diviser par $X - A$, qui est un vecteur de R^p, n'a aucun sens! Néanmoins, si on fixe toutes les composantes de X sauf une, on peut définir des dérivées partielles. Soit $f :\ D \subset R^p \to R^q$ et $A \in D$. Pour $i = 1, \ldots, p$, on appelle dérivée partielle par rapport à x_i de f en $A = (a_1, \cdots a_p)$, et on note $\dfrac{\partial f}{\partial x_i}(A)$, ou bien $f'_{x_i}(A)$, la dérivée de la fonction partielle $f_{A,i}$ prise en a_i:

$$\frac{\partial f}{\partial x_i}(A) = f'_{A,i}(a_i) = \lim_{x_i\to a_i} \frac{f(a_1, \cdots, x_i, \cdots, a_p) - f(a_1, \cdots, a_i, \cdots, a_p)}{x_i - a_i}.$$

Pour une fonction de deux variables $f : D \in^2\to$ en point $A = (a, b) \in D$ les dérivées partielles de $f(x, y)$ en (a, b) sont les dérivées des fonctions partielles $f(x, b)f(a, y)$ qui se calculent alors :

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h} \frac{\partial f}{\partial y}(a, b) = \lim_{k\to 0} \frac{f(a, b+k) - f(a, b)}{k}.$$

Parfois, on les note aussi $f'_x(a, b)f'_y(a, b)$. Soit $f(x, y) = 2x^2 - 3xy + 4y^2$. Calculer les dérivées partielles au point $(1, 2)$. En considérant y constant et en dérivant par rapport à x on a :

$$\frac{\partial f}{\partial x}\Big|_{(x,y)=(1,2)} = (4x-3y)\Big|_{(x,y)=(1,2)} = -2$$

En considérant x constant et en dérivant par rapport à y on a :

$$\frac{\partial f}{\partial y}\Big|_{(x,y)=(1,2)} = (-3x+8y)\Big|_{(x,y)=(1,2)} = 13$$

La matrice des dérivées partielles de $f\ :^p\to^q$ s'appelle la <u>matrice jacobienne</u> ou la <u>Jacobienne</u> de f. La matrice jacobienne $Jac(f)(X_0)$ fait passer de R^p dans R^q: elle a p colonnes et q lignes.

$$Jac(f)(X_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(X_0) & \dots & \frac{\partial f_1}{\partial x_p}(X_0) \\ \vdots & & \vdots \\ \frac{\partial f_q}{\partial x_1}(X_0) & \dots & \frac{\partial f_q}{\partial x_p}(X_0) \end{pmatrix}. \tag{2}$$

Autrement dit, pour une fonction vectorielle $f(x_1, \cdot x_p)$ à valeurs dans ^{q} la matrice jacobienne a pour colonnes les vecteurs $\frac{\partial f}{\partial x_i}$. En particulier, pour une fonction de p variables à valeurs réelles, la matrice jacobienne est simplement une matrice-ligne :

$$Jac(f)(x_1, \cdots, x_p) = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_p}\right).$$

Sa matrice transposée - la matrice-colonne :

$$f(x_1, \cdots, x_p) = \ {}^{\dagger}\!\left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_p}\right)$$

s'appelle le <u>gradient</u> de f.

1.2 Propriétés des dérivées partielles.

Les dérivées partielles d'une fonction qui est obtenue par des opérations algébriques sur d'autres fonctions (somme, produit, fraction) suivent les mêmes règles.

Si une fonction $f\ :\ D\ \subset^p\to^q$ est obtenue par des opérations algébriques (somme, produit, fraction) sur les fonctions $g, h\ :\ D\ \subset^p\to^q$, ses dérivées partielles peuvent être obtenues à partir des dérivées partielles de g et h par les formules de dérivée de somme, produit, fraction habituelles ($(u+v)' = u' + v'$, etc.)

Les dérivées partielles d'une composition de fonctions sont plus compliquées.

Rappel : règle de chaîne. Soit $g : I \subset\to J \subset, g : x \mapsto g(x),\ h : J \subset\to, h : y \mapsto h(y)$ et $f : I \subset\to, f : x \mapsto h(g(x))$. On a :

$$\Big|_{x=x_0} = \frac{\mathrm{d}h}{}\Big|_{y=g(x_0)} \cdot \ \Big|_{x=x_0}$$

Soient

$$\begin{array}{rl} g: & D \subset^p \rightarrow E \subset^m, g : X \mapsto g(X) = (g_1(X), \cdots, g_m(X)), \\ h: & E \subset^m \rightarrow^q, h : Y \mapsto h(Y) = (h_1(Y), \cdots, h_q(Y)), \\ f: & D \subset^p \rightarrow^q, f : X \mapsto h(g(X)) = f(X) = (f_1(X), \cdots f_q(X)) \end{array}$$

des fonctions telles que g en $X_0 \in D$ et h en $g(X_0) \in E$ sont des fonctions continument dérivables (i.e. les dérivées partielles existent et sont continues) alors pour tout $i \in \{1, \cdots, p\}, j \in \{1, \cdots, q\}$:

$$\begin{aligned} \frac{\partial f_j}{\partial x_i}(X_0) = & \frac{\partial (h \circ g)_j}{\partial x_i}(X_0) \\ = & \frac{\partial h_j}{\partial y_1}(g(X_0))\frac{\partial g_1}{\partial x_i}(X_0) + \cdots + \frac{\partial h_j}{\partial y_m}(g(X_0))\frac{\partial g_m}{\partial x_i}(X_0) \end{aligned} \tag{3}$$

ce qui nous donne les entrées d'une matrice jacobienne de f qui est un produit des matrices jacobiennes de h et g.

En particulier, si

$$h :^2 \rightarrow, \ (y_1, y_2) \mapsto h(y_1, y_2) g :^p \rightarrow^2, \ X \mapsto (g_1(X), g_2(X))$$

pour $f = h \circ g :^p \rightarrow$ on a

$$\frac{\partial (f)}{\partial x_i}(X_0) = \frac{\partial (h \circ g)}{\partial x_i}(X_0) = \frac{\partial h}{\partial y_1}(g(X_0))\frac{\partial (g_1)}{\partial x_i}(X_0) + \frac{\partial h}{\partial y_2}(g(X_0))\frac{\partial (g_2)}{\partial x_i}(X_0)$$

1. Soit $f(x) = e^x \sin^2 x$. On peut voir f comme une composition de deux fonctions $g :\rightarrow^2, g(x) = (e^x, \sin x) h :^2 \rightarrow, h(y_1, y_2) = y_1 \cdot (y_2)^2$.

   On a deux façons de calculer la dérivée de f - directement ou en utilisant la Proposition (1.2) :

   $$\begin{aligned} f'(x) &= \frac{\partial (y_1 \cdot (y_2)^2)}{\partial y_1} \cdot \frac{\partial (y_1)}{\partial x} + \frac{\partial (y_1 \cdot (y_2)^2)}{\partial y_2} \cdot \frac{\partial (y_2)}{\partial x} \\ &= (y_2)^2 e^x + 2 y_1 y_2 \cos x = \sin^2 x \cdot e^x + 2 e^x \sin x \cos x. \end{aligned}$$

2. On peut aussi résoudre des problèmes comme celui-là :

   Soient $f(x) = F(x, \phi(x)) = 0$, où $f(x)$ et $\phi(x)$ sont des fonctions d'une variable et $F(y_1, y_2)$ est une fonction de deux variables. Calculer $\phi'(x)$ en fonction des dérivées de F.

   On considère $f(x)$ en tant qu'une fonction composée:

   $$f'(x) = \frac{\partial F}{\partial y_1} + \frac{\partial F}{\partial y_2}\frac{\mathrm{d}\phi(x)}{} = F_1'(x, \phi(x)) + F_2'(x, \phi(x))\phi'(x) = 0.$$

   D'où $\phi'(x) = -\dfrac{F_1'}{F_2'}(x, \phi(x))$.

1.3 Derivées partielles d'ordre supérieur. Fonctions de classe C^k. Théorème de Schwarz

Soit $f : D \subset^p \to$. Les dérivées partielles définissent p nouvelles fonctions

$$f'_{x_i}(x_1, \cdots, x_p) = \frac{\partial f}{\partial x_i}(x_1, \cdots, x_p).$$

On peut regarder les dérivées partielles de chacune de ces nouvelles fonctions. Cela nous donne les dérivées partielles d'ordre 2 (aussi appellées les dérivées partielles secondes) et à leur tour on peut regarder les dérivées partielles des dérivées partielles d'ordre 2, etc. Cela s'écrit par exemple :

$$f''_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

Une fonction $f : D \subset^p \to$ de <u>classe C^k</u> est une fonction dont toutes les dérivées partielles jusqu'à l'ordre k existent et sont continues. Une fonction est dite de <u>classe C^∞</u> si elle est de classe C^k pour tout $k \in$.

(Schwarz) Soit $f : D \subset^p \to$ une fonction de classe C^2 sur D. Les fonctions de dérivées partielles d'ordre 2, $\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$ sont égales en tout point de D.

Le théorème de Schwarz implique que les dérivées partielles d'ordre $k,\ k \geq 2$, d'une fonction de classe C^k, $f : D \subset^p \to$ ne dépendent pas de l'ordre dans lequel les dérivées partielles sont prises. Par exemple, pour une fonction de deux variables $f(x, y)$ de classe C^3, on a : $\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x^2 \partial y}$.

1.4 Différentielle

(Chapitre 2 de [**Rouv**].)

Lors de l'équation (1), en essayant de généraliser l'expression pour la dérivée d'une fonction d'une variable aux fonctions de plusieurs variables, nous avons introduit les fonctions de dérivées partielles, qui sont utiles et révèlent certaines informations sur le comportement de la fonction mais n'apportent pas toute l'information.

On considère à nouveau l'exemple **??**. La fonction $f :^2 \to$ définie de la façon suivante :

$$f : \ (x, y) \mapsto \begin{cases} \dfrac{xy}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0), \\ \\ 0 \text{ si } (x, y) = (0, 0). \end{cases}$$

On calcule sa dérivée partielle par rapport à x :

- $\forall (x_0, y_0) \neq (0, 0),\ \frac{\partial f}{\partial x}(x_0, y_0) = \left(\frac{x y_0}{x^2 + y_0^2}\right)'\Big|_{x = x_0} = \frac{y_0^3 - x^2 y_0}{(x^2 + y_0^2)^2}\Big|_{x = x_0} = \frac{y_0^3 - x_0^2 y_0}{(x_0^2 + y_0^2)^2}$.

- $\dfrac{\partial f}{\partial x}(0,0)$ est la dérivée de $x \mapsto \begin{cases} 0 \text{ si } x \neq 0, \\ 0 \text{ si } x = 0, \end{cases}$

  donc $\dfrac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \dfrac{f(0+h,0) - f(0,0)}{h} = 0.$

On voit que $\dfrac{\partial f}{\partial x}(0,0)$ existe, de même $\dfrac{\partial f}{\partial y}(0,0)$ existe et vaut 0, et pourtant f n'est même pas continue en $(0,0)$. Donc les dérivées partielles ne suffisent pas à décrire la régularité de la fonction.

Nous allons réécrire l'équation (1) sans division et la généraliser aux fonctions de plusieurs variables.

Soit $f : D \subset^p \to^q$, $A \in D$. La différentielle (A) de f au point A est une application linéaire de ^{p} dans ^{q} telle que au voisinage de A on a :

$$f(A+H) - f(A) = ((A))\,(H) + r(H), \text{ où } r(H) = o(\|H\|). \tag{4}$$

Ici, $H \in^p$, tel que $A+H$ est au voisinage de A. La fonction f est dite différentiable au point A si elle possède une différentielle en ce point. La fonction f est dite différentiable dans un domaine D si elle est différentiable en tout point de D.

Cette application agit sur les vecteurs de ^{p} et les envoie vers ^{q}, en particulier $((A))(H) \in^q$. Le reste, $r(H) = o(\|H\|)$, dit "petit o" de $\|H\|$, est une fonction $r : D \subset^p \to^q$, négligeable devant $\|H\|$. On peut comparer leurs normes :

$$\lim_{\|H\| \to 0} \frac{\|r(H)\|_q}{\|H\|_p} = 0.$$

On peut réécrire la condition de différentiablité

$$\lim_{\|H\| \to 0} \frac{f(A+H) - f(A) - ((A))\,(H)}{\|H\|} = 0$$

Si elle existe, la différentielle (A) est unique. On la note selon les auteurs ou les circonstances :

$$L = Df(A) \text{ ou } (A) \text{ ou } D_A f \text{ ou } \mathrm{d}_A f.$$

La différentielle (A), si elle existe, est donnée par une matrice de taille $p \times q$ (une application linéaire de ^{p} vers ^{q} écrite dans des bases des espaces vectoriels ^{pq}). Cette matrice est appellée la matrice jacobienne.

La différentiabilité entraîne l'existence des dérivées partielles. On peut le voir sur un exemple d'une fonction f à p variables à valeurs réelles ($q = 1$). Par définition :

$$\frac{\partial f}{\partial x_i}(A) = \lim_{h_i \to 0} \frac{f(a_1, \cdots, a_i + h_i, \cdots, a_p) - f(a_1, \cdots, a_i, \cdots, a_p)}{h_i}.$$

Par définition de la différentielle on a aussi

$$\frac{f(a_1, \cdots, a_i + h_i, \cdots, a_p) - f(a_1, \cdots, a_i, \cdots, a_p)}{h_i} = \frac{(A)(H) + r(H)}{h_i}$$

Ici H est le vecteur transposé de $(0, \cdots, h_i, \cdots, 0)$. Donc $r(H) = o(\|H\|) = o(h_i)$ et

$$\lim_{h_i \to 0} \frac{(A)(H) + r(H)}{h_i} = \lim_{h_i \to 0} \frac{(A)(H)}{h_i}.$$

Donc ici

$$(A)\ \ ^{\dagger}(0, \cdots, 0, h_i, 0, \cdots, 0) = \frac{\partial f}{\partial x_i}(A) \cdot h_i$$

et par linéarité

$$(A)\ \ ^{\dagger}(h_1, \cdots, h_i, \cdots, h_p) = \sum_{i=0}^{p} \frac{\partial f}{\partial x_i}(A) \cdot h_i$$

Finalement, on remarque que

$$(A)(H) = Jac(f)(A)H(A) = \sum \frac{\partial f}{\partial x_i}(A)_i$$

car les différentielles de fonctions x_i, notées $_i$, satisfont $_i(H) = h_i$.

Dans les exercices de nature théorique, la différentiabilité est souvent établie en montrant directement par des majorations que le reste $r(H)$ est un $o(\|H\|)$. Mais si f est donnée explicitement au moyen des fonctions usuelles, on va plus vite en constatant simplement l'existence et la continuité de ses dérivées partielles. Si une fonction est de classe C^1 elle est différentiable. Soit $f : D \subset R^p \to R^q$ et $X_0 \in D$. Si $\forall i = 1, \ldots, p$, $\forall j = 1, \ldots, q$, $X \mapsto \dfrac{\partial f_j}{\partial x_i}(X)$ *existe au voisinage de* X_0 *et est continue en* X_0, alors f est différentiable en X_0.

En termes moins précis, que j'ai prise dans le livre [**Rouv**] et que je pense essentielle pour la compréhension du cours, la GRANDE IDÉE DU CALCUL DIFFÉRENTIEL :

$$\begin{pmatrix} \text{accroissement} \\ \text{de la fonction} \end{pmatrix} = \begin{pmatrix} \text{terme linéaire par rapport à} \\ \text{l'accroissement de la variable} \end{pmatrix} + \begin{pmatrix} \text{petit terme} \\ \text{correctif} \end{pmatrix} \quad (5)$$

Propriétés de la différentielle.

1. **Continuité.** Une fonction différentiable en un point est continue en ce point.

2. **Linéarité.** Soient $f, g : D \to^q$ deux fonctions définies sur une partie D de ^{p}. Si fg sont différentiables en $A \in D$, $\lambda \in$, alors $\mathrm{d}(f+g)(A) = (A) + (A)$ et $\mathrm{d}(\lambda f)(A) = \lambda(A)$.

3. **Composition.** Soient $g : D \to E \subset^m$ définie sur une partie D de ^{p} et différentiable en $A \in D$, et $h : E \to^q$ différentiable en $g(A)$, alors $h \circ g$ est différentiable en A et la différentielle $\mathrm{d}(h \circ g)(A) = (g(A)) \times (A)$

La composition suit de la formule (4) :

$$\begin{aligned} h(g(A+H)) - h(g(A)) = &\quad \mathrm{d}h(g(A))(g(A+H)-g(A)) + \text{ petit reste} \\ = &\quad \mathrm{d}h(g(A))(A)(H) + \text{ un autre petit reste} \end{aligned}$$

En pratique c'est donné par le produit des matrices jacobiennes (comparer avec l'équation (3)).

Regardons maintenant une fonction $f : D \subset^2 \to$. Soit $(x, y) \in D$. On remarque que la différentielle d'une fonction $(x, y) \to f(x, y)$ au point (x, y) est égale à :

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

1. On reprend : soit une fonction $f :^2 \to$ définie de la façon suivante

$$f : \ (x, y) \mapsto \begin{cases} \dfrac{xy}{x^2+y^2} \text{ si } (x, y) \neq (0, 0), \\ \\ 0 \text{ si } (x, y) = (0, 0). \end{cases}$$

   On sait que f n'est pas différentiable en $(0, 0)$ parce qu'elle n'est même pas continue. Comment se comportent ses dérivées partielles au voisinage de $(0, 0)$?

   On a vu que si $(x_0, y_0) \neq (0, 0)$, $\dfrac{\partial f}{\partial x}(x_0, y_0) = \dfrac{y_0(y_0^2 - x_0^2)}{(x_0^2 + y_0^2)^2}$ et que $\dfrac{\partial f}{\partial x}(0, 0) = 0$. Donc $\dfrac{\partial f}{\partial x}$ est bien définie au voisinage de $(0, 0)$, mais elle n'est pas continue: si $x_n = 1/n$ et $y_n = 2/n$, on a $\lim_{n\to+\infty} x_n = 0$, $\lim_{n\to+\infty} y_n = 0$, et $\dfrac{\partial f}{\partial x}(x_n, y_n) = \dfrac{2/n^2}{(1/n^2 + 4/n^2)^2} = \dfrac{2/n^2}{25/n^4}$. Donc, $\lim_{n\to+\infty} \dfrac{\partial f}{\partial x}(x_n, y_n) \neq 0$.

2. Soit

$$f : \begin{cases} R^2 \to R^2 \\ \\ (x, y) \mapsto (x^2y^2, x + y) \end{cases}$$

   Est-elle différentiable en $(2, 3)$?

   Soit $(x_0, y_0) \in R^2$, $\dfrac{\partial f^1}{\partial x}(x_0, y_0) = 2x_0y_0^2$, $\dfrac{\partial f^1}{\partial y}(x_0, y_0) = 2y_0x_0^2$, $\dfrac{\partial f^2}{\partial x}(x_0, y_0) = 1$, $\dfrac{\partial f^2}{\partial y}(x_0, y_0) = 1$.

   Toutes ces dérivées partielles sont continues en $(2, 3)$ donc f est différentiable en $(2, 3)$. On a $Jac(f)(2, 3) = \begin{pmatrix} 36 & 24 \\ \\ 1 & 1 \end{pmatrix}$.

3. On considère :

$$f: \begin{cases} R \to R \\ x \mapsto x^2 \end{cases}$$

est-elle différentiable en 2? Soit $x_0 \in R$, $\dfrac{\partial f}{\partial x}(x_0) = 2x_0$. Elle est continue en 2 donc f est différentiable en 2 et $Jac(f)(2) = \dfrac{\partial f}{\partial x}(2) = f'(2) = 4$.

2 Propriétés géométriques des fonctions de plusieurs variables

2.1 Dérivée directionnelle

Soit $f : D \subset^p \to^q$, $A \in D$ et $\overrightarrow{V}$ un vecteur de ^{p}. On dit que f a une dérivée au point A en suivant le vecteur $\overrightarrow{V}$ si l'expression :

$$D_{\overrightarrow{V}} f(A) := \lim_{t \to 0} \frac{f(A + t\overrightarrow{V}) - f(A)}{t}$$

existe. $D_{\overrightarrow{V}} f(A)$ s'appelle la <u>dérivée directionnelle</u> de f en A en direction de vecteur $\overrightarrow{V}$. Les dérivées partielles $\dfrac{\partial f}{\partial x_i}$ sont des dérivées directionnelles de f en A en direction de vecteurs de base $e_i = \; ^t(0, \cdots, \underbrace{1}_{i}, \cdots, 0)$. Soit $f : D \subset^p \to$, de classe C^1, en $A \in D$ et $\overrightarrow{V}$ un vecteur de ^{p}. Alors, la dérivée directionnelle de f en A en direction de vecteur $\overrightarrow{V}$ est égale au produit scalaire du gradient de f au point A et du vecteur $\overrightarrow{V}$:

$$D_{\overrightarrow{V}} f(A) = f(A) \cdot \overrightarrow{V} \tag{6}$$

On va démontrer cette proposition pour le cas $p = 2$. La généralisation au cas $p > 2$ est assez directe. Soit $\{\overrightarrow{i}, \overrightarrow{j}\}$ une base orthonormale de ^{2} et $\overrightarrow{V} = \lambda \overrightarrow{i} + \mu \overrightarrow{j}$ $A = (x_0, y_0) \in D \in^2$. Soit une fonction $u :\to^2$, $u(t) = (x_0, y_0) + t\overrightarrow{V} = \left(x_0 + \lambda \overrightarrow{i}, y_0 + \mu \overrightarrow{j}\right) := (x(t), y(t))$. On considère une fonction d'une variable à valeurs réelles : $F(t) = f(u(t))$. C'est une fonction composée. Sa dérivée en 0 :

$$\begin{aligned} \frac{\mathrm{d}F}{}(0) &= \frac{\mathrm{d}(f \circ u)}{}(0) = \frac{\partial f}{\partial x}\Big|_{(x,y)=(x_0,y_0)} \cdot \frac{\partial x(t)}{\partial t}(0) + \frac{\partial f}{\partial y}\Big|_{(x,y)=(x_0,y_0)} \cdot \frac{\partial y(t)}{\partial t}(0) \\ &= \frac{\partial f}{\partial x}\Big|_{(x,y)=(x_0,y_0)} \cdot \lambda + \frac{\partial f}{\partial y}\Big|_{(x,y)=(x_0,y_0)} \cdot \mu = f \cdot \overrightarrow{V} \end{aligned}$$

car $\dfrac{\mathrm{d}x(t)}{}\Big|_{t=0} = \dfrac{\mathrm{d}(x_0 + \lambda t)}{}\Big|_{t=0} = \lambda$ et $\dfrac{\mathrm{d}y(t)}{}\Big|_{t=0} = \dfrac{\mathrm{d}(y_0 + \mu t)}{}\Big|_{t=0} = \mu$. De l'autre coté

$$\begin{aligned}\frac{\mathrm{d}F}{}(0) &= \lim_{t\to 0}\frac{F(t)-F(0)}{t} = \lim_{t\to 0}\frac{f(u(t))-f(u(0))}{t}\\ &= \lim_{t\to 0}\frac{f(x_0+\lambda t, y_0+\mu t)-f(x_0,y_0)}{t} := D_{\overrightarrow{V}} f(x_0,y_0)\end{aligned}$$

D'où la relation (6).

2.2 Gradient.

Soit $f : D \subset^2 \to$ une fonction de classe C^1. Son gradient, pris en tout point de D définit une fonction à valeurs vectorielles $f : D \subset^2 \to^2$, noté aussi :

$$\overrightarrow{\nabla} f(x,y) := \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)(x,y).$$

2.2.1 Propriété [a] : Le gradient est perpendiculaire à la ligne de niveau

Soit X un point d'une courbe $\Gamma \in^p$ et T une droite tangente à Γ au point X. On dit qu'un vecteur $\overrightarrow{V}$ est <u>perpendiculaire à la courbe</u> Γ au point X si $\overrightarrow{V}$ est perpendiculaire à T. Dans ce cas on dit aussi que $\overrightarrow{V}$ est <u>normal à la courbe</u> Γ au point X. En particulier, cela signifie que le produit scalaire de V et du vecteur directeur de T est égal à 0.

Soient $D \subset^2$, $f : D \to (x,y) \in D$, alors si $f(x,y) = a$, (x,y) appartient à la ligne de niveau $L_a(f)$. Le vecteur gradient $\overrightarrow{\nabla} f(x,y)$ est normal à la courbe $L_a(f)$ au point (x,y). Soit $(x+h, y+k) \in L_a(f)$ un point au voisinage de (x,y), qui appartient à la même courbe de niveau que (x,y).

Alors, $f(x+h,y+k) - f(x,y) = 0$ car les valeurs de f en ces deux points sont égales. De la grande idée du calcul différentiel (5) on a :

$$\begin{aligned} f(x+h,y+k)-f(x,y) = &\quad df(x,y)\cdot\begin{pmatrix} h \\ k\end{pmatrix} + o(\|(h,k)\|)\\ = &\quad \frac{\partial f}{\partial x}(x,y)\cdot h + \frac{\partial f}{\partial y}(x,y)\cdot k + o(\|(h,k)\|).\end{aligned}$$

On a $\lim_{(h,k)\to 0} o(\|(h,k)\|) = 0$, donc $\frac{\partial f}{\partial x}(x,y)\cdot h + \frac{\partial f}{\partial y}(x,y)\cdot k \to 0$ quand $(x+h, y+k) \to (x,y)$. Quand $(x+h,y+k) \to (x,y)$ tout en restant sur $L_a(f)$, le vecteur (h,k) est un vecteur tangent à $L_a(f)$. On a alors trouvé que le produit scalaire de $\left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right)(h,k)$ égale 0, on en déduit que ces deux vecteurs sont orthogonaux. A. $f(x,y) = x^2+y^2$. $L_a(f) = C((0,0),\sqrt{a})$ - cercle de centre $(0,0)$ et de rayon $\sqrt{a}$. $\frac{\partial f}{\partial x}(x,y) = 2x$, $\frac{\partial f}{\partial y}(x,y) = 2y$. On remarque que $(2x,2y) = 2(x,y)$ est 2 fois le vecteur radial qui est en effet orthogonal au cercle.

B. Soit la courbe d'équation $x^2 - y = 0$. Pour calculer la normale en chaque point de cette courbe, on la voit comme une ligne de niveau 0 de la fonction $f(x,y) = x^2 - y$. La normale est donc donnée par son gradient : $\overrightarrow{\nabla} f(x,y) = \begin{pmatrix} 2x \\ -1 \end{pmatrix}$.

2.2.2 Propriété [b] : Le gradient indique la ligne de plus grande pente

Sur le graphe de la fonction f on prend un point $(x, y, f(x, y))$, alors (x, y) est sur la ligne de niveau $a = f(x, y)$.

Le gradient en (x, y) indique la direction de plus grande pente ≥ 0 sur Γ_f à partir d'un point en question.

$$f((x,y) + \overrightarrow{v}) - f(x,y) = \overrightarrow{\nabla} f(x,y) \cdot \overrightarrow{v} + o(\|\overrightarrow{v}\|)$$

Le produit scalaire $\overrightarrow{\nabla} f(x,y) \cdot \overrightarrow{v}$ vaut $\|\overrightarrow{\nabla} f(x,y)\| \cdot \|\overrightarrow{v}\| \cos\theta$, où θ est l'angle entre les deux vecteurs. L'accroissement de la fonction atteint le maximum quand $\cos\theta = 1$, alors $\overrightarrow{v}$ doit être parallèle à $\overrightarrow{\nabla} f(x,y)$.

En suivant la ligne de plus grande pente dans D on a, sur le graphe, le chemin le plus court à parcourir pour obtenir une variation donnée de f. Autrement dit, si on veut passer le plus vite possible du niveau a au niveau b à partir d'un point (x, y) donné de niveau $a = f(x, y)$, il faut suivre le gradient.

2.3 Formule de Taylor

Rappel : petit o. Soient fg deux fonctions d'une variable à valeurs réelles. On dit que $g = o(f)$ au point a si :

$$\lim_{x \to a} \frac{g(x)}{f(x)} = 0$$

Exemple : $u(x) = x^3, v(x) = x^2 + 2x$. En $a = 0$ on a $u(x) = o(v(x))$ et en $a = +\infty$ on a $v(x) = o(u(x))$.

Rappel : La formule de Taylor avec le reste en forme de Lagrange. Si f est $n+1$ fois différentiable en a, on a une approximation de f par un polynôme :

$$f(a+t) = f(a) + f'(a)t + \cdots \frac{f^{(n)}(a)}{n!} t^n + r_n(a,t)$$

où il existe $\theta \in [a, a+t]$ tel que $r_n(a,t) = \dfrac{f^{(n+1)}(\theta)}{(n+1)!} t^{n+1}$. C'est une conséquence du théorème des accroissements finis : si f est continue et dérivable sur l'intervalle $[a, b]$, $a < b$ alors $\exists x_0 \in [a, b]$ tel que $f(b) = f(a) + f'(x_0)(b - a)$.

Finalement, on a aussi **la formule de Taylor-Young** avec $r_n(a,t) = o(t^n)$:

$$f(a+t) - f(a) = \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} t^k + o(t^n)$$

C'est cette formule qu'on va généraliser au cas de plusieurs variables.

(Formule de Taylor) Soit $f : D \subset^p \to$ de classe C^n au voisinage du point $A(a_1, a_2, \cdots a_p) \in D$. Soient $H(h_1, \cdots, h_p) \in^p$ l'intervalle $[A, A+H] \subset D$. Alors,

$$f(A+H) - f(A) = \sum_{k=1}^{n} \frac{1}{k!} \left((h_1\partial_1 + \cdots h_p\partial_p)^k (f) \right)(A) + o(\|H\|^n)$$

Ici $\partial_i := \dfrac{\partial}{\partial x_i}$. Soit $F(t) = f(A+tH)$ une fonction composée d'une variable à valeurs réelles. On va utiliser la formule de Taylor-Lagrange pour cette fonction. Pour cela on remarque que :

$$F'(t) = \sum_{i=1}^{p} \frac{\partial f(A+tH)}{\partial x_i} \cdot \frac{\mathrm{d}(a_i + th_i)}{}.$$

Pour la k-ème dérivée de la fonction composée $F(t)$ on a :

$$F^{(k)}(t) = \sum (\partial_{i_k} \cdots (\partial_{i_1} f(A+tH) \underbrace{) \cdots)}_{k \text{ fois}} \frac{\mathrm{d}(a_{i_1} + th_{i_1})}{} \cdots \frac{\mathrm{d}(a_{i_k} + th_{i_k})}{}$$

où on prend la somme sur tout $i_1 \in \{1, \cdots, p\}, \cdots, i_k \in \{1, \cdots, p\}$. On remarque que $\dfrac{\mathrm{d}(a_{i_k} + th_{i_k})}{} = h_i, \forall i \in \{1, \cdots, p\}$. Par le binôme de Newton cette formule se réécrit :

$$F^{(k)}(t) = (h_1\partial_1 + \cdots + h_p\partial_p)^k f(A+tH)$$

On écrit a formule de Taylor-Lagrange pour la fonction $F(0+t)$ au voisinage de 0 :

$$F(t) = F(0) + F'(0)t + \cdots + \frac{1}{n!} F^{(n)}(0)t^n + \frac{1}{(n+1)!} F^{(n+1)}(\theta)t^{n+1}, \ \theta \in [0,t].$$

Pour $t = 1$ on a :

$$F(1) = F(0) + F'(0)t + \cdots \frac{1}{n!} F^{(n)}(0) + \frac{1}{(n+1)!} F^{(n+1)}(\theta), \ \theta \in [0,1]$$

D'où :

$$f(a_1+h_1, \cdots, a_p+h_p) = f(a_1, \cdots, a_p) + \sum_{k=1}^{n} \frac{1}{k!} (h_1\partial_1 + \cdots + h_p\partial_p)^k f(A) + r_n(A,H)$$

Le dernier terme est le reste :

$$r_n(A,H) = \frac{1}{(n+1)!} (h_1\partial_1 + \cdots + h_p\partial_p)^{n+1} f(A+\theta H) \equiv o(\|H\|^n).$$

En particulier, la formule de Taylor à l'ordre 2 est la suivante :

$$f(A+H)=f(A)+\sum_{i=1}^{p}\partial_i f(a)h_i+\frac{1}{2}\sum_{i,j=1}^{p}\partial_i\partial_j f(a)h_ih_j+o(\|H\|^2) \qquad (7)$$

La matrice-colonne des entrées $\partial_i f$ est la matrice Jacobienne. La matrice $p\times p$ des dérivées secondes

$$Hess_f(A):=[\alpha_{ij}]=[\partial_i\partial_j f(A)]$$

s'appelle la <u>matrice Hessienne</u> de f en A. Par le théorème de Schwarz cette matrice est symétrique si f est de classe C^2. La forme quadratique $\alpha(u)=\sum_{i,j=1}^{p}\alpha_{ij}u_iu_j$ s'appelle la <u>forme hessienne</u> de f en A.

L'idée de la formule de Taylor c'est de trouver une approximation de la fonction par un polynôme dans un voisinage d'un point donné.

En particulier, pour $p=2,\ A=(a,b),\ H=(h,k),\ (A+H)=(a+h,b+k)$ on a les formules de Taylor suivantes :

- $n=0$

$$f(A+H)-f(A)=o((\sqrt{h^2+k^2})^0)\Leftrightarrow \lim_{H\to 0}\frac{f(A+H)-f(A)}{1}=0$$

  - continuité

- $n=1$

$$f(A+H)-f(A)=h\frac{\partial f}{\partial x}(a,b)+k\frac{\partial f}{\partial y}(a,b)+o(\sqrt{h^2+k^2})$$

  - différentiabilité.

- $n=2$

$$\begin{aligned} f(A+H) \quad - f(A) &= h\frac{\partial f}{\partial x}(a,b)+k\frac{\partial f}{\partial y}(a,b)\\ &+\frac{1}{2}\left(h^2\frac{\partial^2 f}{\partial x^2}(a,b)+2hk\frac{\partial^2 f}{\partial x\partial y}(a,b)+k^2\frac{\partial^2 f}{\partial y^2}(a,b)\right)+o(h^2+k^2) \end{aligned} \qquad (8)$$

2.4 Vecteur normal et plan tangent à un graphe d'une fonction de 2 variables

A. Surfaces et coordonnées curvilignes (Ici je suis le cours [**Zorich**]).

Soit $f:D\to$ une fonction continue définie sur une partie D de ^{2}. L'ensemble des points de ^{3} :

$$S=\{(x,y,z)\in R^3|(x,y)\in D,\ z=f(x,y)\}$$

est le graphe de la fonction f sur D (définition **??**). Il est évident que l'application :

$$F : D \to S,\ F(x, y) = (x, y, f(x, y))$$

est une bijection. Puisque les points de S sont donnés par des paires de nombres (x, y), l'ensemble S est une surface de dimension 2 dans ^{3}.

Si on a un chemin $\Gamma : I \to D$, alors automatiquement on a un chemin $F \circ \Gamma : I \to S$ sur la surface S. Si

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

est une représentation paramétrique de Γ alors le chemin $F \circ \Gamma$ sur S est donné par les trois fonctions :

$$\begin{cases} x & = x(t) \\ y & = y(t) \\ z & = f(x(t), y(t)) \end{cases}$$

Soit $(x_0, y_0) \in D$. On peut trouver un chemin :

$$x = x_0 + t,\ y = y_0,\ z = f(x_0 + t, y_0)$$

sur la surface S pour lequel la coordonnée $y = y_0$ ne change pas et un autre chemin :

$$x = x_0,\ y = y_0 + t,\ z = f(x_0, y_0 + t)$$

pour lequel la coordonnée $x = x_0$ ne change pas. Ces chemins partant de points différents de la surface S tracent des lignes de coordonnées sur S. Pour cette raison on appelle (x, y) les coordonnées curvilignes sur S.

B. Plan tangent

Si la fonction $z = f(x, y)$ est différentiable en $(x_0, y_0) \in D$, alors, quand $(x, y) \to (x_0, y_0)$ on a :

$$f(x, y) = f(x_0, y_0) + \alpha(x - x_0) + \beta(y - y_0) + o\left(\sqrt{(x - x_0)^2 + (y - y_0)^2}\right) \quad (9)$$

où α et β sont des constantes égales aux dérivées partielles au point (x_0, y_0).

Considérons un plan dans ^{3} donné par une équation

$$z = z_0 + \alpha(x - x_0) + \beta(y - y_0) \quad (10)$$

où $z_0 = f(x_0, y_0)$. On voit que le graphe (9) de la fonction f autour du point (x_0, y_0) est éloigné du plan (10) d'une valeur négligeable devant $\sqrt{(x - x_0)^2 + (y - y_0)^2}$.

Le plan

$$f(x, y) = f(x_0, y_0) + \alpha(x - x_0) + \beta(y - y_0) \quad (11)$$

avec $\alpha = \dfrac{\partial f}{\partial x}(x_0, y_0), \beta = \dfrac{\partial f}{\partial y}(x_0, y_0)$ est appellé le plan tangent au graphe de la fonction $z = f(x, y)$ au point $(x_0, y_0, f(x_0, y_0))$.

C.Vecteur normal

Soit $(x, y, z) \in D \subset^3$ $F(x, y, z) = 0$ l'équation implicite d'une surface S (précédemment on avait une surface : $z = f(x, y)$ pour laquelle $F(x, y, z) = f(x, y) - z$.) Soit

$$t \in I \subset, \ \gamma : t \mapsto \begin{cases} x = f(t) \\ y = g(t) \\ z = h(t) \end{cases}$$

l'équation paramétrique d'une courbe de la surface passant par le point $P_0(x_0, y_0, z_0)$, c'est-à-dire qu'il existe

$$t_0 \in I, (x_0, y_0, z_0) = (f(t_0), g(t_0), h(t_0))(x, y, z) = (f(t), g(t), h(t))$$

qui satisfont l'équation $F(x, y, z) = 0$ pour tout $t \in I$. Soit $u(t) = F(f(t), g(t), h(t))$ une fonction composée de $I \to$, qui est identiquement nulle sur I. Donc au point $t = t_0$ on a

$$0 = \ = \frac{\partial F}{\partial x} \cdot \frac{\mathrm{d}f}{} + \frac{\partial F}{\partial y} \cdot \frac{\mathrm{d}g}{} + \frac{\partial F}{\partial z} \cdot \frac{\mathrm{d}h}{} \tag{12}$$

De l'équation (12) suit que le vecteur ${}^t\left(\frac{\mathrm{d}f}{}, \frac{\mathrm{d}g}{}, \frac{\mathrm{d}h}{}\right)\Big|_{t=t_0}$ est orthogonal au vecteur ${}^t\left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right) \equiv F(P_0)$. Le vecteur ${}^t\left(\frac{\mathrm{d}f}{}, \frac{\mathrm{d}g}{}, \frac{\mathrm{d}h}{}\right)\Big|_{t=t_0}$ est un vecteur quelconque dans l'espace tangent à S au point P_0. Donc le vecteur

$$F(P_0) = \ {}^t\left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)(P_0)$$

est orthogonal à tout vecteur tangent à la surface S passant par P_0. Cela signifie exactement que le vecteur gradient est normal à la surface S.

L'équation du plan tangent à la surface donnée par l'équation $F(x, y, z) = 0$ est facile à établir : c'est le plan passant par P_0 tel que tout vecteur de ce plan est orthogonal à $F(P_0)$. Les coordonnées d'un point $M(x, y, z)$ du plan vérifient : $\overrightarrow{P_0M} \cdot F(P_0) = 0$. Ce produit scalaire donne l'équation du plan tangent :

$$(x - x_0, y - y_0, z - z_0) \cdot \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z}\right)(P_0) = 0.$$

De façon plus explicite :

$$(x - x_0)\frac{\partial F}{\partial x}(P_0) + (y - y_0)\frac{\partial F}{\partial y}(P_0) + (z - z_0)\frac{\partial F}{\partial z}(P_0) = 0.$$

On peut comparer cette formule à la formule (11).

3 Extrema

3.1 Extrema locaux et globaux. Définition

On étudie le comportement d'une fonction de plusieurs variables à valeurs réelles. Une telle fonction peut avoir des valeurs extrémales : des minima (des valeurs les plus petites) ou des maxima (des valeurs les plus grandes) sur tout le domaine de définition ou bien sur une certaine partie. On les appelle des extrema.

1. Soit $f : D \to$ une fonction définie sur une partie $D \subset^p$. On dit que f admet un maximum (resp. minimum) global au point $A \in D$ si pour tout $X \in D$ on a $f(X) \le f(A)$ (resp. $f(X) \ge f(A)$). Le maximum (resp. minimum) est appelé strict si $f(X) < f(A)$ (resp. $f(X) > f(A)$).

2. On dit que f admet un maximum (resp. minimum) local au point $A \in D$ si on peut trouver un nombre $r > 0$ tel que $X \in D\|X - A\| < r$ entraîne $f(X) \le f(A)$ (resp. $f(X) \ge f(A)$).

Les extrema globaux sont appelés aussi extrema absolus.

3.2 Théorème des extrema sur un compact

Soit $f : K \to$ une fonction continue sur un compact $K \subset^p$. Alors f admet un maximum et un minimum sur K.

En dimension $p = 1$ la fonction a des points extrémaux sur un intervalle. Soit ils sont à l'intérieur de l'intervalle, auquel cas ils vérifient $f'(x) = 0$, soit ils sont au bord de l'intervalle (sur le bord, la condition $f'(x) = 0$ n'est pas forcément satisfaite). Donc pour trouver les extrema on cherche d'abord des points critiques (où la derivée s'annule), puis on compare la valeur des points critiques avec les valeurs sur le bord de l'intervalle. Les valeurs max et min se trouvent parmi ces valeurs-là.

Soit $f : D \to$ une fonction de classe C^1 sur une partie D de ^{p}. On dit que $A \in D$ est un point critique de f si toutes les dérivées partielles s'annulent en A (équivalent à dire que le gradient de f est nul en A, équivalent à dire aussi que la différentielle de f est nulle en A).

[Condition nécessaire d'extremum local] Soit $f : U \to$ une fonction de classe C^2 définie sur un ouvert $U \subset^p$ admettant un maximum ou un minimum local au point $A \in U$. Alors A est un point critique de f. Reprenons la formule de Taylor (8) à l'ordre 2 en dimension 2. La preuve se généralise sans problème aux dimensions supérieures.

$$
\begin{aligned}
&f(a+h,b+k) - f(a,b) = h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b) \\
&+\frac{1}{2}\left(h^2\frac{\partial^2 f}{\partial x^2}(a,b) + 2hk\frac{\partial^2 f}{\partial x\partial y}(a,b) + k^2\frac{\partial^2 f}{\partial y^2}(a,b)\right) + o(\|(h,k)\|^2)
\end{aligned}
$$

Si on a un maximum local en A, alors $f(a+h,b+k) - f(a,b) \le 0$ pour tout (h,k) suffisamment petit. La valeur de la fonction linéaire de deux variables $h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b)$, si elle n'est pas 0, est grande par rapport aux termes suivants. Donc cette valeur, si elle n'est pas égale à 0, doit être négative. Pourtant pour h,k positifs il faut que les constantes $\frac{\partial f}{\partial x_i}(a,b) \le 0$, $i = 1,2$ et pour h, k négatifs il faut que les mêmes valeurs $\dfrac{\partial f}{\partial x_i}(a,b) \geq 0,\ i = 1, 2$, d'où $\dfrac{\partial f}{\partial x_i} \equiv 0,\ i = 1, 2$. On peut refaire le même raisonnement pour un min local.

3.3 Extrema de fonctions de 2 variables - critère par le déterminant de matrice Hessienne

Soit $f: D \subset R^p \to R$ et $X_0 \in D$. Quand $p = 1$, pour savoir si un point critique X_0 est un maximum local ou un minimum local, on étudie la dérivée seconde (quand elle existe):

- si $f''(X_0) > 0$, alors $f(X_0)$ est un minimum local,
- si $f''(X_0) < 0$, alors $f(X_0)$ est un maximum local,
- si $f''(X_0) = 0$, il faut faire des calculs supplémentaires de dérivées supérieures - ce peut être un point d'inflexion, un maximum ou un minimum.

Dans le cas de plusieurs variables à la place de f'', on étudie la Hessienne. Soit $f: D \subset R^p \to R$ et $X_0 \in D$ un point critique de f. On suppose que la Hessienne $Hf(X_0)$ existe. Alors

- si toutes les valeurs propres de $Hf(X_0)$ sont strictement positives, $f(X_0)$ est un minimum local,
- si toutes les valeurs propres de $Hf(X_0)$ sont strictement négatives, $f(X_0)$ est un maximum local,
- sinon, et si toutes les valeurs propres ne sont pas 0, il n'y a pas d'extrema. Si toutes les valeurs propres sont 0, il faut étudier des termes d'ordre supérieur dans la décomposition de Taylor en X_0.

Pour $p = 2$ on fait le calcul de la formule de Taylor. Au point critique $X_0(a, b)$ on a

$$f(a+h, b+k) - f(a,b) = \frac{1}{2}\left(h^2 \frac{\partial^2 f}{\partial x^2}(a,b) + 2hk\frac{\partial^2 f}{\partial x \partial y}(a,b) + k^2 \frac{\partial^2 f}{\partial y^2}(a,b)\right) + o(\|(h,k)\|^2)$$

Donc le signe de la forme quadratique (la forme hessienne)

$$\frac{1}{2}\left(h^2 \frac{\partial^2 f}{\partial x^2}(a,b) + 2hk\frac{\partial^2 f}{\partial x \partial y}(a,b) + k^2 \frac{\partial^2 f}{\partial y^2}(a,b)\right)$$

va déterminer si on a un maximum, un minimum ou ni l'un ni l'autre. Pour avoir un maximum (resp. minimum) il faut que la forme soit négative (resp. positive) pour tout (h, k) au voisinage de $(0, 0)$. Si la forme hessienne n'est pas de signe défini on a des couples (h, k) pour lesquelles la valeur de $f(a+h, b+k) - f(a, b)$ est positive et d'autres pour lesquelles cette valeur est négative. Donc on a des directions (h,k) dans lesquelles la fonction a un maximum au point (a,b) et d'autres où la fonction a un minimum au même point. Ce type de point critique s'appelle un point selle (comme une selle de cheval) ou bien point col (comme dans les montagnes).

On étudie alors la forme hessienne. On choisit des notations standard :

$$R = \frac{\partial^2 f}{\partial x^2}(a,b),\ S = \frac{\partial^2 f}{\partial x \partial y}(a,b),\ T = \frac{\partial^2 f}{\partial y^2}(a,b).$$

On suppose que $R \neq 0$ et on réécrit la forme hessienne :

$$\begin{aligned} Rh^2 + 2Shk + Tk^2 &= R\left(h^2 + 2\frac{S}{R}hk + \frac{T}{R}k^2\right) \\ &= R\left(h^2 + 2\frac{S}{R}hk + \left(\frac{S}{R}\right)^2 k^2 - \left(\frac{S}{R}\right)^2 k^2 + \frac{T}{R}k^2\right) \\ &= R\left(\left(h + \frac{S}{R}k\right)^2 + \left(\frac{T}{R} - \frac{S^2}{R^2}\right)k^2\right) \end{aligned}$$

Puisque le premier terme $\left(h + \frac{S}{R}k\right)^2 \geq 0$, c'est le deuxième terme qui définit si la forme est de signe défini. Alors,

- Si $(\frac{T}{R} - \frac{S^2}{R^2}) > 0$ ($\Leftrightarrow RT - S^2 > 0$) on a un maximum si $R < 0$ et minimum si $R > 0$.
- Si $RT - S^2 < 0$ on a un point selle.

Si $RT - S^2 > 0$ la condition $R > 0$ ($R < 0$) est équivalente à la condition $R + T > 0$ ($R + T < 0$) i.e. la condition sur la trace de la matrice hessienne.

Recherche des extrema :

- Déterminer des points où f n'est pas de classe C^1 et regarder les valeurs de f en ces points. Par exemple, la fonction $f(x,y) = 1 - \sqrt{x^2+y^2}$ admet un maximum à l'origine mais on ne le trouve pas parmi les points critiques.
- Rechercher les points critiques.
- Etudier les points critiques.

Extrema locaux et globaux de $f(x,y) = 2x^2y + 2x^2 + y^2$ sur ^{2}. Points critiques :

$$\left\{ \begin{array}{ll} \dfrac{\partial f}{\partial x} & = 4xy + 4x = 0 \\ \dfrac{\partial f}{\partial y} & = 2x^2 + 2y = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{ll} x(y+1) & = 0 \\ x^2 + y & = 0 \end{array} \right.$$

On trouve alors trois points critiques $(0,0), (-1,-1)(1,-1)$.

pt critique	$(0,0)$	$(-1,-1)$	$(1,-1)$
$R = 4y+4$	4	0	0
$S = 4x$	0	-4	4
$T = 2$	2	2	2
$RT - S^2$	8	-16	-16
Signe de R	> 0		
Nature du pt critique:	min	pt selle	pt selle

Les extrema globaux : on voit que

$$\lim_{x\to\pm\infty} f(x,0) = \lim_{x\to\pm\infty} 2x^2 = +\infty$$

donc pas de maximum global. Pas de minimum global non plus car

$$\lim_{x\to\pm\infty} f(x,-2) = \lim_{x\to\pm\infty} -2x^2 + 4 = -\infty$$

Ici on a utilisé un critère par le signe du déterminant (et de la trace) de la matrice hessienne pour déterminer la nature du point critique. Si le déterminant est 0 on doit regarder la formule de Taylor à l'ordre supérieur (à l'ordre 3 et parfois plus).

On cherche des extrema locaux de $g(x,y) = x^4 + y^4 - 2x^2$ sur ^{2}.

On trouve 3 points critiques $(-1,0)$, $(0,0)$, $(1,0)$ pour lesquels on ne peut pas utiliser le critère car $RT - S^2 = 0$ mais $g(x,y) = (x^2-1)^2 + y^4 - 1$ donc en $(\pm 1, 0)$ il y a un minimum local. En $(0,0)$ on a $g(0,0) = 0$ et au voisinage de $(0,0)$ on a des valeurs positives et négatives $g(0,y) = y^4 > 0$ et $g(y,0) = x^4 - 2x^2 < 0$ pour x suffisamment petit. Donc $(0,0)$ n'est pas un max ni un min, c'est un point-selle.

3.4 Extrema liés

Soit K un compact de ^{2}. Soit $f : K \to$ une fonction de classe C^2. Soit $g(x,y) = 0$ l'équation de la courbe $C \subset K$. Si C est le bord de K, on a une notation $C = \partial K$. On regarde la restriction de f sur la courbe C. Si la courbe C a pour équation $g(x,y) = 0$, tous les points de la courbe satisfont cette équation. Quand on cherche les extrema de la fonction f sur C on dit qu'on étudie les extrema de f assujettie à la contrainte $g(x,y) = 0$. Ce sont des extrema liés.

Exemple A. Voici un exemple de problème de recherche d'extrema liés : parmi des rectangles avec la somme de cotés $2p$ (où p est un nombre positif donné), trouver un rectangle à l'aire maximale. Soient x, y les cotés du rectangle. Alors on a $\sigma(x,y) = xy$ l'aire, qui doit être maximale tandis que (x,y) sont sousmis à la condition $x + y = p$. Ici, il est facile d'exprimer y par x et trouver un maximum d'une fonction d'une variable ainsi obtenue.

Il est rare que l'on puisse exprimer y directement comme une fonction de x en utilisant la contrainte.

Exemple B. Regardons un exemple de la page 362 [**Rouv**] : la fonction $f(x,y) = x^2 + y^2$ et la contrainte, la courbe C, définie par une équation $g(x,y) = 0$. Il s'agit de trouver un minimum de f, lié par cette relation $g(x,y) = 0$. C'est un minimum de f sur la courbe C. Géométriquement on peut résoudre le problème en traçant des lignes de niveau de f. Les lignes de niveau de f sont des cercles concentriques du centre $(0,0)$. Si on trace des cercles de rayons croissants, jusqu'à leur rencontre avec la courbe C, la valeur critique est sur le cercle qui touche la courbe. Faites un dessin - c'est instructif (dessinez une courbe quelconque et tracez les cercles).

La méthode générale utilise la considération suivante. Soit $P(a,b)$ un point extremum de f restreint à la courbe C. Le vecteur tangent à la courbe au point P doit être aussi tangent à la ligne de niveau $f(a,b)$ (on le voit clairement dans l'exemple B). Mais les lignes de niveau sont normales au gradient de f, de l'autre côté le vecteur tangent à C est normal au gradient de g. Donc ces deux gradients sont proportionnels. On appelle le coefficient de proportionnalité le multiplicateur de Lagrange.

Soit $f : U \to, \ g : U \to$ deux fonctions de classe C^1 sur un ouvert U de ^{2}. Soit (a,b) un point de U tel que :

1. f soumise à la contrainte $g(x,y) = 0$ admet un extremum au point(a,b).
2. $g(a,b) \neq 0$

Alors il existe un nombre réel $\lambda \neq 0$ tel que $f(a,b) = \lambda \ \ g(a,b)$. Les nombres a, b, λ sont des solutions du système d'équations suivant : les dérivées partielles de $f(x,y) - \lambda g(x,y)$ par rapport à x, y, λ doivent être égales à 0.

$$\begin{cases} \dfrac{\partial f}{\partial x}(a,b) - \lambda \dfrac{\partial g}{\partial x}(a,b) & = 0 \\ \dfrac{\partial f}{\partial y}(a,b) - \lambda \dfrac{\partial g}{\partial y}(a,b) & = 0 \\ g(a,b) & = 0 \end{cases}$$

Trouver le point de la courbe $y = x^2$ qui est le plus près du point $(0,h)$. Alors, ici $g(x,y) = y - x^2$, $f(x,y) = x^2 + (y-h)^2$ - le carré de la distance. Les gradients nous donnent

$$\begin{cases} 2x + 2\lambda x & = 0 \\ 2(y-h) - \lambda & = 0 \\ y - x^2 & = 0 \end{cases}$$

Les solutions : soit $x = 0$, et alors $y = 0$ aussi, ou bien $\lambda = -1$ et $y = h - 1/2$, $x = \pm\sqrt{h - 1/2}$. Alors pour $h \geq 1/2$, les points $(\pm\sqrt{h-1/2}, h - 1/2)$ sont à la distance minimale de $(0,h)$. Si $h < 1/2$ on a $(0,0)$ comme point le plus proche. Soit f une fonction C^2 sur un compact $K \subset^2$, alors f atteint un minimum et un maximum globaux sur K. Ces points d'extrema sont

- soit des points intérieurs de K, auquel cas ce sont des points critiques ($f = 0$ en ces points)

- soit ils sont sur le bord ∂K de K auquel cas ils sont donnés par le calcul des extrema liés en utilisant des multiplicateurs de Lagrange.

Trouver les extrema globaux de $f(x,y) = y + y^2 - x^2 + 3$ sur $B(0,1)$ disque de centre $(0,0)$ de rayon 1. On cherche les points critiques :

$$\begin{cases} \dfrac{\partial f}{\partial x} & = -2x & = 0 \\ \dfrac{\partial f}{\partial y} & = 1 + 2y & = 0 \end{cases}$$

On trouve un seul point critique $(0, -1/2)$. Ce point se trouve dans le disque et sa valeur est $f(0,-1/2) = 11/4$ La matrice hessienne donne :

$$\text{dét} \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix} = -4 < 0 \Rightarrow (0, -1/2) \text{ point selle.}$$

Il faut alors chercher les extrema globaux sur le bord $x^2 + y^2 - 1 = 0$. On a :

$$\begin{cases} -2x - 2\lambda x & = 0 \\ 1 + 2y - 2\lambda y & = 0 \\ x^2 + y^2 - 1 & = 0 \end{cases}$$

On trouve les points $(0, \pm 1)(\pm\dfrac{\sqrt{15}}{4}, -\dfrac{1}{4})$. Les valeurs : $f(0,1) = 5$, $f(0,-1) = 3$, $f(\pm\dfrac{\sqrt{15}}{4}, -\dfrac{1}{4}) = \dfrac{15}{8}$. On compare ces valeurs et conclut que le max se trouve au point $(0,1)$ et le min aux points $(\pm\dfrac{\sqrt{15}}{4}, -\dfrac{1}{4})$.

3.5 Extrema d'une fonction de $n > 2$ variables

En dimension n on procède de la même façon qu'en dimension 2. En utilisant la formule de Taylor en dimension n au voisinage d'un extremum on voit que la condition nécessaire est que le gradient s'annule aux points d'extrema locaux. La condition suffisante pour avoir un miminum (resp. maximum) est que la forme hessienne soit positivement (resp. négativement) définie.

Pour les extrema liés on a le théorème suivant ([**Rouv**]) : Soient $f, g_1, \cdots g_n$ des fonctions réelles de classe C^1 sur un ouvert U de ^{p}, et E un ensemble défini par les équations :

$$g_1(X) = 0, \cdots, g_n(X) = 0, \text{ avec } X \in U.$$

Si la restriction de f à E admet un extremum local en $A \in E$, et si les différentielles $Dg_1(A), \cdots, Dg_n(A)$ sont linéarement indépendantes sur ^{p}, alors nécessairement les formes linéaires $Df(A), Dg_1(A), \cdots, Dg_n(A)$ sont liées. En d'autres termes, il existe des coefficients réels $\lambda_1, \cdots \lambda_n$, appelés multiplicateurs de Lagrange, tels que

$$Df(A) = \lambda_1 Dg_1(A) + \cdots \lambda_n Dg_n(A)$$

4 Exercices corrigés

Exercice 1:

Justifier l'existence des dérivées partielles des fonctions suivantes, et les calculer.

1. $f(x, y) = e^x cosy$.
2. $f(x, y) = (x^2 + y^2)cos(xy)$.
3. $f(x, y) = \sqrt{1 + x^2y^2}$.

Corrigé

Pour tout réel y fixé, la fonction $x \mapsto e^x cosy$ est dérivable sur $\Re$, ce qui justifie l'existence de la dérivée partielle par rapport à la première variable dans le premier exemple. La justification est identique pour les autres fonctions et on trouve respectivement :

1. $$\frac{\partial f}{\partial x}(x, y) = e^x cosy \ et \ \frac{\partial f}{\partial y}(x, y) = -e^x siny.$$
2. $$\frac{\partial f}{\partial x}(x, y) = 2xcos(xy) - y(x^2 + y^2)sin(xy),$$
   et
   $$\frac{\partial f}{\partial y}(x, y) = 2ycos(xy) - x(x^2 + y^2)sin(xy).$$
3. $$\frac{\partial f}{\partial x}(x, y) = xy^2\sqrt{1 + x^2y^2}, et \frac{\partial f}{\partial y}(x, y) = yx^2\sqrt{1 + x^2y^2}$$
   .

Exercice 2:

Soit $f : \Re^2 \to \Re$ une fonction de classe C^1.

1. On définit $g : \Re \to \Re$ par $g(t) = f(2 + 2t, t2)$. Démontrer que g est C^1 et calculer $g^{'}(t)$ en fonction des dérivées partielles de f.
2. On définit $h : \Re^2 \to \Re$ par $h(u, v) = f(uv, u2 + v2)$. Démontrer que h est C^1 et exprimer les dérivées partielles $\frac{\partial h}{\partial u}$ et $\frac{\partial h}{\partial v}$ en fonction des dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$.

Corrigé

1. La fonction $t \mapsto (2+2t, t^2)$ est de classe C^1, car polynômiale, donc g est de classe C^1 par composition. On applique ensuite la formule de la dérivée d'une fonction composée. Si on note $u(t) = 2 + 2t$ et $v(t) = t^2$, alors

$$g^{'}(t) = u^{'}(t)\frac{\partial f}{\partial x}(u(t), v(t)) + v^{'}(t)\frac{\partial f}{\partial y}(u(t), v(t)),$$

soit

$$g^{'}(t) = 2\frac{\partial f}{\partial x}(2 + 2t, t^2) + 2t\frac{\partial f}{\partial y}(2 + 2t, t^2).$$

2. La fonction $(u, v) \mapsto (uv, u^2 + v^2)$ est de classe C^1 car polynômiale, donc h est de classe C^1. Notons $r(u, v) = uv$ et $q(u, v) = u2 + v2$. Le théorème de dérivation d'une composée dit que

$$\frac{\partial h}{\partial u}(u, v) = \frac{\partial r}{\partial u}(u, v)\times\frac{\partial f}{\partial x}(r(u, v), q(u, v)) + \frac{\partial q}{\partial u}(u, v)\times\frac{\partial f}{\partial y}(r(u, v), q(u, v)).$$

Ceci donne

$$\frac{\partial h}{\partial u}(u, v) = v\frac{\partial f}{\partial x}(uv, u^2 + v^2) + 2u\frac{\partial f}{\partial y}(uv, u^2 + v^2).$$

De même on trouve

$$\frac{\partial h}{\partial v}(u, v) = u\frac{\partial f}{\partial x}(uv, u^2 + v^2) + 2v\frac{\partial f}{\partial y}(uv, u^2 + v^2).$$

Exercice 3:

Pour les fonctions suivantes, démontrer qu'elles admettent une dérivée suivant tout vecteur en $(0, 0)$ sans pour autant y être continue.

$$f(x, y) = \begin{cases} y^2 \ln|x| & si\ x \neq 0 \\ 0 & sinon. \end{cases}$$

$$g(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & si\ (x, y) \neq (0, 0) \\ 0 & sinon. \end{cases}$$

Exercice 4:

Justifier que les fonctions suivantes sont différentiables, et calculer leur différentielle

1. $f(x, y) = e^{xy}(x + y)$
2. $f(x, y, z) = xy + yz + zx$
3. $f(x, y) = (y \sin jx, \cos x)$

Corrigé

1. f admet des dérivées partielles

$$\frac{\partial f}{\partial x}(x,y) = (y(x+y)+1)e^{xy}$$

   et

$$\frac{\partial f}{\partial y}(x,y) = (x(x+y)+1)e^{xy}.$$

   Ces fonctions sont continues sur $\Re^2$, la fonction f est différentiable, et on a :

$$df_{(x,y)}(h,k) = (h(y(x+y)+1) + k(x(x+y)+1))e^{xy}.$$

   Avec la notation différentielle, on a (ce qu'on peut obtenir aussi en différentiant directement f), on a aussi

$$df = (y(x+y)+1)e^{xy}dx + (x(x+y)+1)e^{xy}dy.$$

2. f est clairement C^∞, et est donc différentiable. On a donc

$$df = (y+z)dx + (x+z)dy + (x+y)dz.$$

3. f est clairement C^∞, et est donc différentiable. La différentielle de f est une application linéaire de $\Re^2$ dans $\Re^2$. On a

$$\frac{\partial f}{\partial x}(x,y) = (y\cos x, -\sin x)$$

   et

$$\frac{\partial f}{\partial y}(x,y) = (\sin x, 0)$$

   .

   On en déduit que

$$df_{(x,y)}(h,k) = h\begin{pmatrix} y\cos x \\ -\sin x \end{pmatrix} + k\begin{pmatrix} \sin x \\ (0 \end{pmatrix}$$

Exercice 5:

Justifier que les fonctions suivantes sont différentiables, et calculer leur matrice jacobienne.

1. $f(x,y,z) = (\frac{1}{2}(x^2 - z^2), \sin x \sin y)$.
2. $f(x,y) = (xy, \frac{1}{2}x^2 + y, \ln(1+x^2))$.

Corrigé

Il suffit de vérifier que les fonctions coordonnées sont différentiables, et elles sont clairement C^∞. On a respectivement

1. $$J_{(x,y,z)}f = \begin{pmatrix} x & 0 & -z \\ \cos x \sin y & \sin x \cos y & 0 \end{pmatrix}$$

2. $$J_{(x,y,z)}f = \begin{pmatrix} y & x \\ x & 1 \\ \frac{2x}{1+x^2} & 0 \end{pmatrix}$$

Exercice 6:

On définit sur $\Re^2$ l'application suivante :

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & si\ (x,y) \neq (0,0) \\ 0 & sinon. \end{cases}$$

1. f est-elle continue en $(0,0)$?
2. f admet-elle des dérivées partielles en $(0,0)$?
3. f est-elle différentiable en $(0,0)$?

Corrigé

1. Remarquons que $f(x,x) = 1/2$, qui ne tend pas vers 0 si x tend vers 0 : f n'est pas continue en 0.
2. Puisque $f(x,0) = 0$, $\frac{\partial f}{\partial x}(0,0)$ existe et vaut 0. De même, $\frac{\partial f}{\partial y}(0,0)$ existe, et vaut 0.
3. f ne peut pas être différentiable puisqu'elle n'est pas continue!

Exercice 7:

1. Démontrer que, pour tous (x,y) réels, alors $|xy| \leq x^2 - xy + y^2$
2. Soit f la fonction de $\Re^2$ dans $\Re$ définie par $f(0,0) = 0$ et $f(x,y) = (x^p y^q)/(x^2 - xy + y^2)$ si $(x,y) \neq (0,0)$, où p et q sont des entiers naturels non nuls. Pour quelles valeurs de p et q cette fonction est-elle continue?
3. Montrer que si $p + q = 2$, alors f n'est pas différentiable.

4. On suppose que $p+q=3$, et que f est différentiable en $(0,0)$. Justifier qu'alors il existe deux constantes a et b telles que $f(x,y)=ax+by+o(\| (x,y) \|)$. En étudiant les applications partielles $x \mapsto f(x,0)$ et $y \mapsto f(0,y)$, justifier que $a=b=0$. Conclure, à l'aide de $x \mapsto f(x,x)$, que f n'est pas différentiable en $(0,0)$.

Corrigé

1. On a $(|x|-|y|)^2=|x|^2-2|xy|+y^2 \geq 0$, ce qui donne l'inégalité demandé!

2. Seule la continuité en $(0,0)$ pose problème. On a donc $|f(x,y)| \leq \frac{|xy||x^{p-1}y^{q-1}|}{(x^2-xy+y^2)} \leq |x|^{p-1}|y|^{q-1}$. Cette dernière quantité tend vers 0, sauf si $p-1=q-1=0$, c'est-à-dire sauf si $p+q=2$. Dans ce cas, on a $f(x,x)=1$, qui ne tend pas vers 0 si x tend vers 0. f n'est alors pas continue en $(0,0)$

3. Si $p+q=2$, la fonction n'est pas continue : a fortiori, elle ne peut pas être différentiable.

4. Si $p+q=3$, et que f est différentiable en $(0,0)$, alors $f(x,y)=f(0,0)+df_{(0,0)}(x,y)+o(\| (x,y) \|)$. Mais la différentielle est ici une application linéaire de $\Re^2$ dans $\Re$. On note $(a \quad b)$ sa matrice. On obtient alors le résultat demandé. Ensuite, puisque $f(x,0)=0=ax+o(|x|)$, on obtient, par unicité du développement limite, que $a=0$. De même en étudiant l'application $y \mapsto f(0,y)$, on trouve $b=0$. Ainsi, $f(x,x)=o(|x|)$. Mais $f(x,x)=x$, qui n'est pas un $o(|x|)$.