TD II – Dérivées partielles et extrema

---

1 Dérivées partielles

Exercice 1.1 (Caractérisation des fonctions de classe $\mathcal{C}^1$)**.** La fonction $x \mapsto J_f(x)$ est continue si et seulement si toutes ses coordonnées le sont. Or, ses coordonnées sont les fonctions

$$x \mapsto \frac{\partial f_i}{\partial x_j}(x)$$

qui par définition sont continues si et seulement si f est de classe $\mathcal{C}^1$.

Exercice 1.2 (Applications linéaires)**.** Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire et $(e_i)_{1\leqslant i\leqslant n}$ la base canonique de $\mathbb{R}^n$.

1. On a $f(x+te_i) - f(x) = tf(e_i)$, donc

$$\frac{f(x+te_i)-f(x)}{t} = f(e_i) \underset{t\to 0}{\longrightarrow} f(e_i) = \frac{\partial f}{\partial x_i}(x).$$

2. La différentielle de f est par définition l'application linéaire

$$\begin{aligned} x \mapsto & \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(x) \\ &= \sum_{i=1}^{n} x_i f(e_i) \\ &= f\left(\sum_{i=1}^{n} x_i e_i\right) \\ &= f(x). \end{aligned}$$

3. Les dérivées partielles sont constantes, donc en particulier continues. Ainsi, f est de classe $\mathcal{C}^1$.

Exercice 1.3 (Applications bilinéaires)**.** Soit $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ une forme bilinéaire.

1. Si $(e_i)_{1\leqslant i\leqslant n}$ désigne la base canonique de $\mathbb{R}^n$, on a pour $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$ que

$$B(x,y) = \sum_{i,j=1}^{n} x_i y_j B(e_i, e_j).$$

Autrement dit, B est polynomiale en les coefficients de x et y, donc de classe $\mathcal{C}^\infty$ d'après le cours.

2. On a

$$B(x+h,y+k) = B(x,y) + B(h,y) + B(x,k) + B(h,k).$$

Comme $(h,k) \mapsto B(h,y)+B(x,k)$ est linéaire, il suffit de montrer que le dernier terme est négligeable. Pour cela, calculons[1] pour $h,k \in \mathbb{R}^n$

$$\begin{aligned}
|B(h,k)| &= \left| \sum_{i,j=1}^{n} h_i k_j B(e_i,e_j) \right| \\
&\leqslant \left(\max_{1\leqslant i,j\leqslant n} |B(e_i,e_j)| \right) \sum_{i,j=1}^{n} |h_i||k_j| \\
&\leqslant \left(\max_{1\leqslant i,j\leqslant n} |B(e_i,e_j)| \right) \left(\sum_{i=1}^{n} |h_i| \right) \left(\sum_{i=j}^{n} |k_j| \right) \\
&\leqslant \left(\max_{1\leqslant i,j\leqslant n} |B(e_i,e_j)| \right) n^2 \|h\|\|k\| \\
&= C\|h\|\|k\| \\
&\leqslant C \max(\|h\|,\|k\|)^2 \\
&= C\|(h,k)\|.
\end{aligned}$$

On a utilisé ici l'Inégalité de Cauchy-Schwarz à la troisième et à la quatrième ligne. Autrement dit, $B(h,k) = o(h,k)$ et la différentielle de B est donc $(h,k) \mapsto B(h,y) + B(x,k)$.

3. Les dérivées partielles de B sont données par

$$\begin{aligned}
\frac{\partial B}{\partial(x_i,y_j)}(x,y) &= D_{(x,y)}B(e_i,e_j) \\
&= B(e_i,y) + B(x,e_j).
\end{aligned}$$

Il s'agit d'une application linéaire, donc continue. Par conséquent, B est de classe $\mathcal{C}^1$.

Exercice 1.4 (Avec un produit scalaire)**.** On a

$$\begin{aligned}
\frac{F(x+te_i)-F(x)}{t} &= \frac{1}{t}\langle f(x+te_i),x\rangle + \frac{1}{t}\langle f(x+te_i),te_i\rangle - \frac{1}{t}\langle f(x),x\rangle \\
&= \left\langle \frac{f(x+te_i)-f(x)}{t}, x \right\rangle + \langle f(x+te_i),e_i\rangle \\
&\underset{t\to 0}{\longrightarrow} \left\langle \frac{\partial f}{\partial x_i}(x), x \right\rangle + \langle f(x),e_i\rangle.
\end{aligned}$$

Comme f est de classe $\mathcal{C}^1$, les dérivées partielles de F sont continues comme composées de fonctions continues (le produit scalaire est une forme bilinéaire). Ainsi, F est de classe $\mathcal{C}^1$.

Exercice 1.5 (Identité d'Euler)**.** **1**. On considère, pour $x \in U$ fixé, la fonction $g : t \mapsto f(tx)$. On a

$$g'(t) = \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(tx).$$

Par ailleurs, par homogénéité, $g(t) = t^r f(x)$, dont la dérivée vaut $rt^{r-1}f(x)$ si $r \neq 0$ et 0 sinon. En égalisant ces deux dérivées et en prenant $t = 1$, on obtient le résultat.

[1] Il existe plusieurs normes équivalentes décrivant la structure d'espace vectoriel produit. On peut adapter le calcul suivant en fonction de celle choisie.

2. (a) On calcule en utilisant l'identité d'Euler :

$$\begin{aligned} F'(t) &= \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(tx) - rt^{r-1} f(x) \\ &= \frac{1}{t} \sum_{i=1}^{n} tx_i \frac{\partial f}{\partial x_i}(tx) - \frac{1}{t} rt^r f(x) \\ &= \frac{1}{t} \left(rf(tx) - rt^r f(x) \right) \\ &= \frac{r}{t} F(t). \end{aligned}$$

(b) On a $F(1) = 0$.

(c) On résout l'équation différentielle obtenue pour F, ce qui donne

$$F(t) = Ce^{r \ln(t)} = Ct^r.$$

Comme $F(1) = 0$, on a $C = 0$ et donc $F = 0$. Ainsi, f est homogène de degré r.

3. On dérive par rapport à la j-ème variable dans l'identité d'Euler, ce qui donne

$$\sum_{i=1}^{n} x_i \frac{\partial^2 f}{\partial x_j \partial x_i}(x) + \frac{\partial f}{\partial x_j}(x) = r \frac{\partial f}{\partial x_j}(x).$$

Il suffit alors de regrouper les dérivées partielles premières pour obtenir l'identité d'Euler pour la dérivée partielle avec un degré $r - 1$.

On peut également éviter l'emploi des dérivées partielles secondes de la façon suivante : si f est homogène d'ordre r, alors pour tout $x \in \mathbb{R}^n$, $1 \leqslant i \leqslant n$ et $t, u > 0$, on a

$$\begin{aligned} \frac{f(tx + tue_i)}{tu} &= \frac{f(t(x + ue_i)) - f(tx)}{tu} \\ &= t^{r-1} \frac{f(x + ue_i) - f(x)}{u} \end{aligned}$$

En faisant tendre $u \to 0$, on obtient le résultat.

2 Dérivées partielles d'ordre supérieur

Exercice 2.1 (Équation des ondes). **1**. Il suffit de calculer les dérivées partielles. On a

$$\frac{\partial F_1}{\partial x}(x,t) = f'(x - ct) \quad \& \quad \frac{\partial F_1}{\partial t}(x,t) = -cf'(x - ct)$$

et donc

$$\frac{\partial^2 F_1}{\partial x^2}(x,t) = f''(x - ct) \quad \& \quad \frac{\partial^2 F_1}{\partial t^2}(x,t) = c^2 f''(x - ct).$$

Ainsi, F_1 vérifie bien l'équation des ondes, et le calcul pour F_2 est similaire.

2. (a) Il faut calculer les dérivées partielles d'une composée, ce que l'on peut faire parce que F et Φ sont de classe $\mathcal{C}^1$. Allons-y donc :

$$\frac{\partial \widetilde{F}}{\partial X} = \frac{1}{2} \frac{\partial F}{\partial x} \circ \Phi(X,Y) - \frac{1}{2c} \frac{\partial F}{\partial t} \circ \Phi(X,Y)$$

et donc

$$\begin{aligned}
\frac{\partial \widetilde{F}}{\partial Y \partial X} &= \frac{1}{4}\frac{\partial^2 F}{\partial x^2} \circ \Phi(X,Y) + \frac{1}{4c}\frac{\partial^2 F}{\partial t \partial x} \circ \Phi(X,Y) \\
&\quad - \frac{1}{4c}\frac{\partial^2 F}{\partial x \partial t} \circ \Phi(X,Y) - \frac{1}{4c^2}\frac{\partial^2 F}{\partial t^2} \circ \Phi(X,Y) \\
&= \frac{1}{4}\frac{\partial^2 F}{\partial x^2} \circ \Phi(X,Y) - \frac{1}{4c^2}\frac{\partial^2 F}{\partial t^2} \circ \Phi(X,Y) \\
&= 0.
\end{aligned}$$

(b) De l'équation précédente, on tire l'existence d'une fonction $g : \mathbb{R} \to \mathbb{R}$ telle que

$$\frac{\partial \widetilde{F}}{\partial X}(X,Y) = g(Y)$$

pour tout $(X,Y) \in \mathbb{R}^2$. Par conséquent, il existe une fonction $f : \mathbb{R} \to \mathbb{R}$ telle que pour tout $(X,Y) \in \mathbb{R}^2$,

$$\widetilde{F}(X,Y) = f(X) + g(Y).$$

Il suffit maintenant de remarquer que $\Phi^{-1}(x,t) = (x - ct, x + ct)$ pour conclure que

$$\begin{aligned}
F(x,t) &= \widetilde{F} \circ \Phi^{-1}(x,t) \\
&= f(x - ct) + g(x + ct).
\end{aligned}$$

Exercice 2.2 (Équation de la chaleur)**.** Pour $x \in \mathbb{R}$ et $t > 0$, on pose

$$f(x,t) = \int_0^{x/2\sqrt{t}} e^{-s^2} \mathrm{d}s.$$

1. Posons

$$F(y) = \int_0^y e^{-s^2} \mathrm{d}s,$$

de sorte que $F'(y) = e^{-y^2}$ et $F''(y) = -2ye^{-y^2}$. On a alors $f(x,t) = F(x/2\sqrt{t})$, d'où

$$\begin{aligned}
\frac{\partial f}{\partial x}(x,t) &= \frac{1}{2\sqrt{t}} F'\left(\frac{x}{2\sqrt{t}}\right) \\
\frac{\partial^2 f}{\partial x^2}(x,t) &= \frac{1}{2\sqrt{t}} \times \frac{1}{2\sqrt{t}} F''\left(\frac{x}{2\sqrt{t}}\right) \\
&= \frac{1}{4t}\frac{-x}{\sqrt{t}} e^{-x^2/4t}
\end{aligned}$$

et

$$\begin{aligned}
\frac{\partial f}{\partial t}(x,t) &= \frac{-1}{2}\frac{x}{2t^{3/2}} e^{-x^2/4t} \\
&= \frac{\partial^2 f}{\partial x^2}(x,t).
\end{aligned}$$

2. Pour une solution à variables séparées, il faut avoir

$$g_1(x) g_2'(t) = g_1''(x) g_2(t).$$

Si on cherche une solution qui ne s'annule pas, on peut réécrire cette égalité sous la forme

$$\frac{g_1''(x)}{g_1(x)} = \frac{g_2'(t)}{g_2(t)}.$$

Le membre de gauche ne dépend que de x, mais est égal à une quantité qui ne dépend que de t. Par conséquent, les deux membres doivent être constants. En notant $-\omega^2$ cette constante – qu'on choisit donc négative – on déduit que

$$g_1(x) = \alpha_1 \cos(\omega x + \theta_1) \quad \& \quad g_2(t) = \alpha_2 \cosh(-\omega^2 t + \theta_2).$$

En prenant $\alpha_1 = \alpha_2 = 1$ et $\theta_1 = 0 = \theta_2$, on trouve la solution $(x,t) \mapsto \cos(\omega x)\cosh(-\omega^2 t)$.

Exercice 2.3 (Inversion de matrices – le retour)**.** On considère l'application $M \mapsto \det(M)$ sur $\mathrm{M}_n(\mathbb{R})$.

1. La formule explicite du déterminant montre qu'il s'agit d'une application polynomiale, donc de classe $\mathcal{C}^\infty$.

2. Soit E_{ij} la matrice dont le coefficient en position (i,j) vaut 1 et tous les autres 0. Ces matrices forment la base canonique de $\mathrm{M}_n(\mathbb{R})$ et nous allons calculer les dérivées partielles correspondantes. Pour cela, on observe que tout d'abord que si $i \neq j$, alors

$$\det(\mathrm{I}_n + tE_{ij}) = 1 = \det(\mathrm{I}_n),$$

donc la dérivée partielle vaut 0. Par ailleurs, si $i = j$, alors

$$\det(\mathrm{I}_n + tE_{ii}) = 1 + t,$$

donc la dérivée partielle vaut 1. En conclusion, la différentielle du déterminant en I_n est donnée par

$$\begin{aligned} D_{\mathrm{I}_n} \det(H) &= \sum_{i,j=1}^n H_{ij} \frac{\partial \det}{\partial E_{ij}}(\mathrm{I}_n) \\ &= \sum_{i=1}^n H_{ii} \\ &= \mathrm{Tr}(H). \end{aligned}$$

3. Si A est inversible, alors on a

$$\begin{aligned} \det(A + H) &= \det\left(A(\mathrm{I}_n + A^{-1}H)\right) \\ &= \det(A)\det(\mathrm{I}_n + A^{-1}H) \\ &= \det(A)\left(1 + \mathrm{Tr}(A^{-1}H) + o(\|A^{-1}H\|)\right) \\ &= \det(A) + \det(A)\mathrm{Tr}(A^{-1}H) + \det(A)o(\|A^{-1}H\|). \end{aligned}$$

L'application $H \mapsto \det(A)\mathrm{Tr}(A^{-1}H)$ est linéaire, donc il suffit de montrer que le dernier terme est un $o(\|H\|)$. Or, le dernier terme est de la forme $\|A^{-1}H\|\varepsilon(\|A^{-1}H\|)$ avec $\varepsilon \to 0$ quand son argument tend vers 0. Comme $\|A^{-1}H\| \leqslant \|A^{-1}\|\|H\|$, et que $\|A^{-1}\|\varepsilon(\|A^{-1}H\|) \to 0$ quand $\|H\| \to 0$, le dernier terme est bien négligeable devant H. Ainsi, nous avons trouvé un développement limité au premier ordre au point A, et la partie linéaire est donc sa différentielle. Autrement dit,

$$D_A \det : H \mapsto \det(A)\mathrm{Tr}(A^{-1}H).$$

4. On sait que l'application $A \mapsto D_A H$ est continue. De plus, elle coïncide sur $\mathrm{GL}_n(\mathbb{R})$ avec l'application $A \mapsto \mathrm{Tr}(\mathrm{Com}(A)^t H)$, qui est également continue. Comme $\mathrm{GL}_n(\mathbb{R})$ est dense, on conclut que ces deux applications coïncident partout.

On considère maintenant l'application $M \mapsto M^{-1}$ sur $\mathrm{GL}_n(\mathbb{R})$.

5. Rappelons la Formule de Cramer :

$$M^{-1} = \frac{1}{\det(M)}\mathrm{Com}(M)^t.$$

Les coefficients de la comatrice sont polynomiaux donc de classe $\mathcal{C}^\infty$ et le déterminant est une fonction de classe $\mathcal{C}^\infty$, donc il en est de même pour l'inversion.

6. On commence par calculer les dérivées partielles. Si $i \neq j$, alors

$$(\mathrm{I}_n + tE_{ij})^{-1} = \mathrm{I}_n - tE_{ij},$$

donc la dérivée partielle correspondante est $-E_{ij}$. Si par contre $i = j$, alors

$$(\mathrm{I}_n + tE_{ii})^{-1} = \mathrm{I}_n - \frac{t}{1+t}E_{ii},$$

donc la dérivée partielle correspondante est $-E_{ii}$. Il suit que la différentielle de l'inversion en l'identité est l'application $H \mapsto -H$. Si maintenant $A \in \mathrm{GL}_n(\mathbb{R})$, on a

$$\begin{aligned}(A+H)^{-1} &= \left(A\left(\mathrm{I}_n + A^{-1}H\right)\right)^{-1} \\ &= \left(\mathrm{I}_n + A^{-1}H\right)^{-1}A^{-1} \\ &= \left(\mathrm{I}_n - A^{-1}H + o(H)\right)A^{-1} \\ &= A^{-1} - A^{-1}HA^{-1} + o(H)\end{aligned}$$

donc la différentielle en A est $H \mapsto -A^{-1}HA^{-1}$.

Exercice 2.4 (Puissance)**.** Pour un entier k, on considère l'application $M \mapsto M^k$.

1. Chaque coefficient est une fonction polynomiale, donc de classe $\mathcal{C}^\infty$. Il en est donc de même pour l'application puissance k-ième.

2. Soient $A, H \in \mathrm{M}_n(\mathbb{R})$. En gardant à l'esprit que ces deux matrices ne commutent pas nécessairement, on voit que

$$(A+H)^k = A^k + A^{k-1}H + A^{k-2}HA + \cdots + HA^{k-1} + g(A,H),$$

où $g(A,H)$ est un polynôme dont tous les termes sont au moins de degré deux en H. Il reste à montrer que $g(A,H) = o(H)$. Pour cela, observons qu'il s'agit d'une somme d'un nombre fini $(2^k - k)$ de termes dont la norme est majorée par $\|H\|^p\|A\|^{n-p}$ pour $2 \leqslant p \leqslant k$. Chacune de ces normes est négligeable devant $\|H\|$, donc $g(A,H) = o(H)$. Ainsi, nous avons trouvé le développement limité au premier ordre de notre fonction.

3. On peut écrire le développement complet dans ce cas :

$$(A+H)^2 = A^2 + AH + HA + H^2.$$

Ainsi, dans ce cas, $D_Af(H) = AH + HA$. Il s'agit d'une application linéaire dont la différentielle est donc aisée à calculer. Faisons-le néanmoins explicitement pour plus de clarté :

$$\begin{aligned}D_{A+K}f(H+K) &= (A+K)H + H(A+K) \\ &= AH + KH + HA + HK \\ &= D_Af(H) + KH + HK.\end{aligned}$$

Ainsi, la différentielle seconde est la forme bilinéaire $(H,K) \mapsto HK + KH$ et la différentielle troisième est nulle.

Exercice 2.5 (Fonction de production de Cobb-Douglas). On considère la fonction de production de Cobb-Douglas générale

$$f(K, L) = \lambda K^\alpha L^{1-\alpha}$$

pour un $0 < \alpha < 1$ fixé.

1. On a

$$\frac{\partial f}{\partial K}(K, L) = \alpha\lambda K^{\alpha-1} L^{1-\alpha} \quad \& \quad \frac{\partial f}{\partial L}(K, L) = (1-\alpha)\lambda K^\alpha L^{-\alpha}.$$

et

$$\begin{aligned}
\frac{\partial^2 f}{\partial K^2} &= \lambda\alpha(\alpha-1)K^{\alpha-2}L^{1-\alpha} \\
\frac{\partial^2 f}{\partial L \partial K} &= \lambda\alpha(1-\alpha)K^{\alpha-1}L^{-\alpha} = \frac{\partial^2 f}{\partial K \partial L} \\
\frac{\partial^2 f}{\partial L^2} &= -\lambda\alpha(1-\alpha)K^\alpha L^{-\alpha-1}.
\end{aligned}$$

2. Elle est positive, ce qui signifie que si on augmente le Capital, on augmente la production, ce qui est raisonnable.

3. Elle est négative à cause du terme $(\alpha - 1)$. Cela signifie que certes, augmenter le Capital améliore la production, mais que cet effet décroît quand le Capital augmente. Autrement dit, plus on augmente le Capital plus l'augmentation correspondante de production est faible.

4. Elle est positive. Ainsi, si on augmente le Travail, on améliore l'effet de l'augmentation du Capital. Ceci est assez raisonnable : si la force de travail ne suit pas, il est peu utile d'injecter de l'argent dans la production. Il faut donc augmenter le Travail pour que l'effet positif du Capital soit plus important.

Exercice 2.6 (Différentielle seconde). Soit $f : U \to \mathbb{R}^m$ une fonction de classe $\mathcal{C}^2$. Pour $h \in \mathbb{R}^n$, on considère l'application $g : U \to \mathbb{R}^m$ définie par

$$g(x) = D_x f(h).$$

1. On a

$$\begin{aligned}
\frac{g(x + te_i) - g(x)}{t} &= \frac{1}{t}\left(D_{x+te_i} f(h) - D_x f(h)\right) \\
&= \frac{1}{t}\sum_{j=1}^n h_j \left(\frac{\partial f}{\partial x_j}(x + te_i) - \frac{\partial f}{\partial x_j}(x)\right) \\
&= \sum_{j=1}^n h_j \frac{1}{t}\left(\frac{\partial f}{\partial x_j}(x + te_i) - \frac{\partial f}{\partial x_j}(x)\right) \\
&\xrightarrow[t\to 0]{} \sum_{j=1}^n h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x).
\end{aligned}$$

2. Pour $k \in \mathbb{R}^n$, on a donc

$$\begin{aligned}
D_x g(k) &= \sum_{i=1}^n k_i \frac{\partial g}{\partial x_i} \\
&= \sum_{i=1}^n \sum_{j=1}^n k_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j} \\
&= D_x^2 f(k, h).
\end{aligned}$$

3. Pour $h \in \mathbb{R}^n$, considérons l'application $T_h : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ définie par $T_h(L) = L(h)$. Alors, comme T_h est linéaire elle est sa propre différentielle, donc

$$\begin{aligned} D_x(T_h \circ F) &= D_{F(x)}T_h \circ D_xF \\ &= T_h \circ D_xF. \end{aligned}$$

Or, $T_h \circ F = g$, donc d'après ce qui précède,

$$T_h \circ D_xF(k) = D_x^2 f(k, h).$$

Autrement dit, D_xF est l'application linéaire envoyant k sur l'application linéaire envoyant h sur $D_x^2 f(k, h)$.

Exercice 2.7 (Différentielle seconde d'une composée). On commence par écrire

$$(g \circ f)(x + h) = g\left(f(x) + D_xf(h) + \frac{1}{2}D_x^2 f(h, h) + o(\|h\|^2)\right).$$

Puis, en notant $f(x) + h'$ le terme entre parenthèses,

$$\begin{aligned} (g \circ f)(x + h) &= g(f(x)) + D_{f(x)}g(h') + \frac{1}{2}D_{f(x)}^2 g(h', h') + o(\|h'\|^2) \\ &= (g \circ f)(x) + (D_{f(x)}g) \circ D_xf(h) + \frac{1}{2}(D_{f(x)}g) \circ D_x^2 f(h, h) + (D_{f(x)}g)(o(\|h\|^2)) \\ &+ \frac{1}{2}D_{f(x)}^2 g(D_xf(h), D_xf(h)) + \frac{1}{8}D_{f(x)}^2 g(D_x^2 f(h, h), D_x^2 f(h, h)) + \frac{1}{2}D_{f(x)}^2 g(o(\|h\|^2), o(\|h\|^2)) \\ &+ o(\|h'\|^2). \end{aligned}$$

On peut alors observer, en utilisant la continuité des applications linéaires et bilinéaires, que tous les termes ci-dessous contenant des $o(\cdot)$ sont des $o(\|h\|^2)$. De plus, comme $D_x^2 f(h, h) = O(\|h\|^2)$, on a

$$D_{f(x)}^2 g(D_x^2 f(h, h), D_x^2 f(h, h)) = O(\|h\|^4) = o(\|h\|^2).$$

Ainsi,

$$\begin{aligned} (g \circ f)(x + h) &= (g \circ f)(x) + (D_{f(x)}g) \circ D_xf(h) \\ &\quad + \frac{1}{2}(D_{f(x)}g) \circ D_x^2 f(h, h) + \frac{1}{2}D_{f(x)}^2 g(D_xf(h), D_xf(h)) + o(\|h\|^2) \end{aligned}$$

et on conclut que

$$D_x^2(g \circ f)(h, h) = (D_{f(x)}g) \circ D_x^2 f(h, h) + D_{f(x)}^2 g(D_xf(h), D_xf(h)).$$

Remarque : On aurait aussi pu raisonner sur les dérivées partielles, mais les calculs n'en sont pas plus simples.

3 Extrema

Exercice 3.1 (Échauffement). 1. On a $\nabla f(x) = (2x + y, 2y + x)$. Les points critiques sont donc donnés par les équations

$$\begin{cases} 2x + y &= 0 \\ 2y + x &= 0 \end{cases}$$

dont l'unique solution est $x = 0 = y$.

2. L'idée est d'essayer de factoriser l'expression en y trouvant le développement d'un carré, de la façon suivante :

$$\begin{aligned} f(x,y) &= \frac{x^2+y^2}{2} + 1 + \frac{x^2+2xy+y^2}{2} \\ &= \frac{x^2+y^2}{2} + 1 + \frac{(x+y)^2}{2}. \end{aligned}$$

Sous cette forme, on voit que $f(x,y) \geqslant 1$ et donc qu'il y a un minimum global en $(0,0)$.

On considère maintenant la fonction $g : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$x^2 + y^2 + 4xy - 2.$$

4. On a $\nabla g(x,y) = (2x+4y, 2y+4x)$. Encore une fois, le seul point critique est $(0,0)$.

5. Le long de la droite d'équation $x = y$, on a $g(x,y) = 6x^2 - 2 > -2 = g(0,0)$ tandis que le long de la droite d'équation $x = -y$, on a $g(x,y) = -2x^2 - 2 < -2 = g(0,0)$. Ainsi, il n'y a pas d'extremum local en -2 : on a un point selle.

Exercice 3.2 (Condition d'ordre deux). **1**. On a $\nabla f(x,y) = (6x^2+6y, 6x-6y)$. Si le gradient s'annule, on a $x = y$ et $6x^2 + 6y = 0$ donc $x = y = 0$ ou $x = y = -1$. La Hessienne de f est

$$H_f(x,y) = \begin{pmatrix} 12x & 6 \\ 6 & -6 \end{pmatrix}.$$

En $(0,0)$, son déterminant vaut -36 donc on a un point selle. En $(-1,-1)$, son déterminant vaut 36 donc les deux valeurs propres sont de même signe strict. De plus, le premier coefficient est négatif, donc les valeurs propres aussi, ce qui montre qu'on a un maximum local. On a $f(-1,-1) = 3$. Comme $f(x,0) \to +\infty$ quand $x \to +\infty$, ce maximum n'est pas global.

2. On a $\nabla g(x,y) = (2xy, x^2 + \ln(y)^2 + 2\ln(y))$. Si le gradient s'annule, x doit s'annuler (la fonction n'est pas définie pour $y = 0$) et on doit avoir $\ln(y)^2 + 2\ln(y) = 0$, donc soit $\ln(y) = 0$ soit $\ln(y) = -2$. La Hessienne de g est

$$H_g(x,y) = \begin{pmatrix} 2y & 2x \\ 2x & 2\ln(y)/y + 2/y \end{pmatrix}.$$

Pour $x = 0$, cette matrice est diagonale. Si $\ln(y) = 0$, c'est-à-dire $y = 1$, les coefficients diagonaux sont 2 et 2 donc on a un minimum local. Si $\ln(y) = -2$, c'est-à-dire $y = e^{-2}$, les coefficients diagonaux sont $2e^{-2}$ et $-2e^2$ donc on a un point selle. Dans le premier cas, la valeur du minimum local est $g(0,1) = 0$. Comme $g(x,y) \geqslant 0$ pour tous $(x,y) \in \mathbb{R} \times \mathbb{R}_+$, c'est un minimum global.

3. On a $\nabla h(x,y) = (4x^3 - 4y, 4y^3 - 4x)$. Si le gradient s'annule, on a $x = y^3$ et $x^3 = y$, donc $x = x^9$, ce qui donne $x = \pm 1$ dont on déduit $y = x$. La Hessienne de h est

$$H_h(x,y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}.$$

Pour $x = \pm 1$, le déterminant vaut $128 > 0$, donc on a deux extrema. De plus, le premier coefficient est positif, donc on deux minima locaux. Enfin, $h(1,1) = -2 = h(-1,-1)$. On peut démontrer que $h(x,y) \to +\infty$ quand $\|(x,y)\| \to +\infty$, et en déduire que h admet un minimum global. Par conséquent, ce minimum est atteint deux fois, en $(1,1)$ et en $(-1,-1)$.

Exercice 3.3 (The World Company)**.** Une entreprise cherche à maximiser son profit en produisant un bien. Si on note p le prix unitaire du bien, alors le profit est donné en fonction du travail L et du capital K par la formule
$$\pi(L, K) = pf(L, K) - \alpha L - \beta K,$$
où f est la fonction de production de l'entreprise et α et β sont des paramètres définissant la stratégie de l'entreprise.

1. La valeur $(\widetilde{L}, \widetilde{K})$ étant critique, on a
$$\begin{aligned} 0 &= \frac{\partial \pi}{\partial L}(\widetilde{L}, \widetilde{K}) \\ &= p\frac{\partial f}{\partial L}(\widetilde{L}, \widetilde{K}) - \alpha, \end{aligned}$$
d'où
$$\frac{\partial f}{\partial L}(\widetilde{L}, \widetilde{K}) = \frac{\alpha}{p}.$$

2. On dérive l'expression par rapport à α (en supposant que $\widetilde{K}$ ne dépend pas de α). On trouve alors
$$\frac{\partial^2 f}{\partial L^2}(\widetilde{L}, \widetilde{K})\frac{\partial \widetilde{L}}{\partial \alpha}(\alpha) = \frac{1}{p}.$$
et on conclut en remarquant que
$$\frac{\partial^2 \pi}{\partial L^2}(\widetilde{L}, \widetilde{K}) = p\frac{\partial^2 f}{\partial L^2}(\widetilde{L}, \widetilde{K}).$$

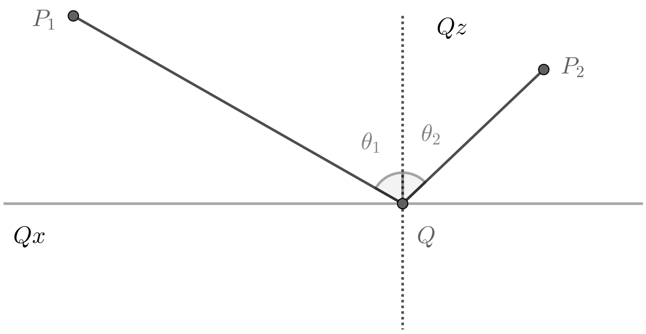
3. Dans le cas de la fonction de Cobb-Douglas, les dérivées partielles secondes non-croisées sont négatives. Si c'est le cas pour f, alors la dérivée de $\widetilde{L}$ par rapport à α doit également être négative. Autrement dit, $\widetilde{L}$ décroît quand α augmente. C'est logique : si le travail coûte plus cher, l'optimum sera obtenu avec moins de travail.

Exercice 3.4 (Un peu de réflexion)**.** On considère un rayon lumineux partant d'un point P_1 de coordonnées $(0, 0, z_1)$ et se réfléchissant sur un miroir formant le plan $\mathcal{P}$ d'équation $z = 0$ en un point Q de coordonnées $(x, y, 0)$. Il atteint ensuite un point P_2 de coordonnées $(x_2, 0, z_2)$ avec $z_2 > 0$.

1. Il suffit d'additionner les temps de trajet de P_1 à Q et de Q à P_2 : si v désigne la vitesse de la lumière dans le milieu (supposée constante), alors
$$\begin{aligned} T(x, y) &= \frac{P_1Q}{v} + \frac{P_2Q}{v} \\ &= \frac{1}{v}\left(\sqrt{x^2 + y^2 + z_1^2} + \sqrt{(x - x_2)^2 + y^2 + z_2^2}\right). \end{aligned}$$

2. Si le temps est minimal, alors les dérivées partielles de T s'annulent. Or,
$$v\frac{\partial T}{\partial y}(x, y) = \frac{y}{\sqrt{x^2 + y^2 + z_1^2}} + \frac{y}{\sqrt{(x - x_2)^2 + y^2 + z_2^2}},$$
donc si cette quantité s'annule on doit avoir $y = 0$. Alors, P_1, P_2 et Q sont bien tous les trois dans le plan d'équation $y = 0$.

3. Pour un point critique $(x, 0)$, on doit de plus avoir
$$0 = v\frac{\partial T}{\partial x}(x, y) = \frac{x}{\sqrt{x^2 + y^2 + z_1^2}} + \frac{x - x_2}{\sqrt{(x - x_2)^2 + y^2 + z_2^2}}$$
Pour mieux visualiser le sens de cette égalité, faisons une figure :



Les coordonnées de P_1 dans le repère (Q, Qx, Qz) sont $(-\sin(\theta_1)QP_1, \cos(\theta_1)QP_1)$ tandis que celles de P_2 sont $(\sin(\theta_2)QP_2, \cos(\theta_2)QP_2)$. De plus, l'abscisse de P_1 dans le repère (Q, Qx, Qz) est également $-x$, tandis que celle de P_2 est $x_2 - x$. Ainsi, on a

$$\begin{aligned} v\frac{\partial T}{\partial x}(x,y) &= \frac{\sin(\theta_1)QP_1}{QP_1} + \frac{-\sin(\theta_2)QP_2}{QP_2} \\ &= \sin(\theta_2) - \sin(\theta_1) \end{aligned}$$

et le résultat suit.

Exercice 3.5 (Une inégalité isopérimétrique)**.** Le but de cet exercice est de trouver, parmi tous les triangles de périmètre 2, ceux qui ont une aire maximale. On utilisera pour cela la Formule de Héron : pour un triangle dont les longueurs des côtés sont a, b et c, l'aire est égale à

$$A = \sqrt{p(p-a)(p-b)(p-c)},$$

où p est le demi-périmètre.

1. Si le périmètre vaut 2, $p = 1$ et $c = 2 - a - b$, d'où

$$A = \sqrt{(1-a)(1-b)(a+b-1)}.$$

2. Pour que cette fonction soit bien définie, il faut que $a, b \in [0,1]$ et que $a + b \geqslant 1$.

3. (a) On a

$$\begin{aligned} \nabla f(x,y) &= \begin{pmatrix} -(1-y)(x+y-1) + (1-x)(1-y) \\ -(1-x)(x+y-1) + (1-x)(1-y) \end{pmatrix} \\ &= \begin{pmatrix} (1-y)(1-x-x-y+1) \\ (1-x)(1-y-x-y+1) \end{pmatrix} \\ &= \begin{pmatrix} (1-y)(1-x-x-y+1) \\ (1-x)(1-y-x-y+1) \end{pmatrix} \\ &= \begin{pmatrix} (1-y)(2-2x-y) \\ (1-x)(2-2y-x) \end{pmatrix} \end{aligned}$$

Si (x,y) est un point critique, comme on veut $x, y \neq 1$, la seule possibilité est

$$\begin{aligned} 2 - 2x - y &= 0 \\ 2 - 2y - x &= 0 \end{aligned}$$

ce qui donne $x = 2/3 = y$ et $f(2/3, 2/3) = 1/27$.

(b) Nous allons étudier la fonction $f_y : x \mapsto (1-x)(1-y)(x+y-1)$. Calculons d'abord sa dérivée :

$$f_y'(x) = (1-x)(2-2x-y).$$

Cette fonction s'annule uniquement pour $x = 1 - y/2$ et il s'agit d'un maximum global d'après le tableau de variations ci-dessous :

x	0		$1-y/2$		1
f_y	$-(y-1)^2$	$\nearrow$	$M(y)$	$\searrow$	0

On a ainsi

$$M(y) = f_y(1-y/2) = \frac{y^2}{4}(1-y).$$

(c) À nouveau on étudie la fonction $M(y)$ en calculant sa dérivée :

$$\begin{aligned} M'(y) &= \frac{y}{2}(1-y) - \frac{y^2}{4} \\ &= \frac{y}{2}\left(1 - y - \frac{y}{2}\right) \\ &= \frac{y}{2}\left(1 - \frac{3y}{2}\right) \end{aligned}$$

qui ne s'annule qu'en $y = 2/3$. On vérifie aisément que c'est un maximum et que la valeur correspondante est $1/27$. Ainsi, pour tous $x, y \in]0;1[$,

$$\frac{1}{27} \geqslant M(y) \geqslant f_y(x) = f(x,y).$$

Nous avons donc bien trouvé le maximum.

4. La fonction que nous avons maximisée est le carré de l'aire. Comme l'aire est positive, le maximum est atteint au même point. Ainsi, l'aire sera maximale quand tous les côtés sont égaux, c'est-à-dire pour un triangle équilatéral.

Exercice 3.6 (Méthode de Héron)**.** Héron d'Alexandrie avait développé une méthode pour calculer les décimales du nombre irrationnel $\sqrt{2}$. Cette méthode consiste à calculer par récurrence les termes de la suite définie par $x_0 = 2$ et

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

1. Considérons la fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2 - 2$. Cette fonction ne s'annule qu'une seule fois, en $x = \sqrt{2}$. La Méthode de Newton donne alors une suite qui, si l'on part assez près de $\sqrt{2}$, convergera vers cette valeur. De plus, la suite en question est définie par la relation de récurrence

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\ &= x_n - \frac{x_n^2 - 2}{2x_n} \\ &= x_n - \frac{x_n}{2} + \frac{2}{2x_n} \\ &= \frac{x_n}{2} + \frac{1}{x_n} \\ &= \frac{1}{2}\left(x_n + \frac{2}{x_n}\right). \end{aligned}$$

Ainsi, la suite de Héron est la suite de la Méthode de Newton associée à f.

2. On procède par récurrence, le résultat étant vrai pour $x_0 = 2$. Supposons donc que $1 \leqslant x_n \leqslant 2$. Alors,

$$\begin{aligned} x_{n+1} &= \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \\ &\geqslant \frac{1}{2}\left(1 + \frac{2}{2}\right) \\ &= 1 \end{aligned}$$

et

$$\begin{aligned} x_{n+1} &= \frac{1}{2}\left(x_n + \frac{2}{x_n}\right) \\ &\leqslant \frac{1}{2}\left(2 + \frac{2}{1}\right) \\ &= 2. \end{aligned}$$

3. On a

$$\begin{aligned} x_{n+1} - \sqrt{2} &= \frac{x_n^2 + 2 - 2\sqrt{2}x_n}{2x_n} \\ &= \frac{(x_n - \sqrt{2})^2}{2x_n} \\ &\leqslant \frac{(x_n - \sqrt{2})^2}{2}. \end{aligned}$$

4. On a d'après ce qui précède,

$$\begin{aligned} |x_{n+1} - \sqrt{2}| &\leqslant \frac{1}{2}|x_n - \sqrt{2}|^2 \\ &\leqslant \left(\frac{1}{2}\right)^{1+2+4+\cdots+2^n} |x_0 - \sqrt{2}|^{2^n} \\ &= \left(\frac{1}{2}\right)^{2^{n+1}-1} |2 - \sqrt{2}|^{2^{n+1}}. \end{aligned}$$

Ainsi,

$$x_n \underset{n\to+\infty}{\longrightarrow} \sqrt{2}.$$