Feuille d'exercices n$^{\text{o}}$10 Corrigé

Exercice 1 : petites questions

1. En appliquant la règle de la chaîne, il vient que

$$u'(x) = \partial_x f(x, -x) - \partial_y f(x, -x).$$

On peut aussi le voir en multipliant les jacobiennes :

$$\text{Jac} f = (\partial_x f \ \partial_y f) \text{ et Jac} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

De même,

$$\partial_x g(x, y) = \partial_y f(y, x),$$
$$\partial_y g(x, y) = \partial_x f(y, x).$$

2. La fonction est bien différentiable puisque c'est un polynôme. On peut écrire

$$\begin{aligned} q(x + h) &= q(x) + 2b(x, h) + q(h) \\ &= q(x) + 2b(x, h) + o(h). \end{aligned}$$

La différentielle est donc donnée par la forme bilinéaire, ce que l'on peut aussi voir en calculant les dérivées partielles.

3. a) Les deux applications

$$\varphi(x, y) := (\sqrt{x^2 + y^2}, 2\tan\left(\frac{y}{x + \sqrt{x^2 + y^2}}\right),$$

$$\psi(r, \theta) := (r\cos\theta, r\sin\theta)$$

sont bien $\mathcal{C}^1$ sur leurs domaines de définition respectifs $\mathbb{R}^2 - \mathbb{R}_-$ et $\mathbb{R}_+^* \times] - \pi; \pi[$, et sont réciproques l'une de l'autre. Ce sont donc bien des difféomorphismes.
On pourrait également utiliser le théorème d'inversion global en calculant la différentielle de l'une des applications pour s'apercevoir que c'est un difféomorphisme local. Comme le changement de coordonnées est injectif, c'est un difféomorphisme sur son image, qui est ouverte. Comme l'application est propre, elle est également fermée, donc l'application est surjective par connexité.

b) Il suffit de calculer

$$\begin{aligned}\partial_r g &= \cos\theta \partial_x f + \sin\theta \partial_y f \\ \partial_\theta g &= -r\sin\theta \partial_x f + r\cos\theta \partial_y f\end{aligned}$$

où l'argument des fonctions est implicitement (r, θ) à gauche et $(r\cos\theta, r\sin\theta)$ à droite. On peut inverser la formule pour obtenir

$$\begin{aligned}\partial_x f &= \cos\theta \partial_r g - \sin\theta \frac{\partial_\theta g}{r} \\ \partial_y f &= \sin\theta \partial_r g + \cos\theta \frac{\partial_\theta g}{r}.\end{aligned}$$

A l'aide de formes différentielles, on a

$$\mathrm{d}f = \partial_x f \mathrm{d}x + \partial_y f \mathrm{d}y,$$

et on remplace ensuite en utilisant

$$\mathrm{d}x = \mathrm{d}(r\cos\theta) = \cos\theta \mathrm{d}r - r\sin\theta \mathrm{d}\theta.$$

c) Le changement de variable n'est plus le même ici. Comme la dérivée partielle est obtenue en fixant le second argument, la réponse ne sera a priori pas la même, bien que l'on dérive la "même" fonction par rapport à la "même" variable.

$$h(x, \theta) = f(x, x\tan\theta),$$

d'où

$$\partial_x h(x\theta) = \partial_x f + \tan\theta \partial_y f.$$

Exercice 2 : différentielle du déterminant

1. Le déterminant est un polynôme en les coefficients de la matrice, il est donc bien différentiable (et même $\mathcal{C}^\infty$).
2. On pourrait calculer les dérivées partielles en chacun des coefficients. On peut à la place calculer pour $t \neq 0$:

$$\begin{aligned}\det(I_n + tH) &= t^n \det\left(H + \frac{1}{t} I_n\right) \\ &= t^n \chi_H\left(-\frac{1}{t}\right) \\ &= 1 + t\mathrm{tr}H + o(H).\end{aligned}$$

La différentielle est donc la trace.

3. On utilise la propriété de multiplicativité du déterminant :

$$\begin{aligned}\det(M_0+H) &= \det M_0 \det(I_n + M_0^{-1}H)\\ &= \det M_0 (1 + \operatorname{tr}(M_0^{-1}H) + o(H))\\ &= \det M_0 + \operatorname{tr}(\det M_0 \ M_0^{-1}H) + o(H).\end{aligned}$$

4. La différentielle du déterminant est égale à $\operatorname{tr}((\mathrm{com} M)^T H$ sur $\mathrm{GL}_n(\mathbb{R})$, qui est dense. Comme la comatrice est continue, c'est le cas partout puisque le déterminant est $\mathcal{C}^1$, donc sa différentielle est continue.

5. La différentielle s'annule là où la comatrice s'annule, c'est à dire en les matrices de rang inférieur à $n-2$.

Exercice 3 : différentielle chez les matrices

1. La fonction inverse est bien différentiable et même $\mathcal{C}^\infty$ puisque c'est une fraction rationnelle (donnée par la formule de la transcomatrice. Pour calculer sa différentielle on écrit

$$(A+H)(A+H)^{-1} = I_n$$

d'où

$$I_n + HA^{-1} + Ad_Af(H) + o(H) = I_n$$

où on a noté f la fonction inverse. Ainsi,

$$d_Af(H) = -A^{-1}HA^{-1}.$$

2. On calcule la dérivée de la fonction $t \mapsto e^{tB}e^{(1-t)A}$. En utilisant la règle de dérivation d'un produit et celle de l'exponentielle, il vient que la dérivée vaut

$$e^{tB}Be^{(1-t)A} - e^{tB}Ae^{(1-t)A} = e^{tB}(B-A)e^{(1-t)A}.$$

Ainsi, en intégrant on obtient

$$e^B - e^A = \int_0^1 e^{tB}(B-A)e^{(1-t)A}.$$

On remplace maintenant B par $A+H$, il vient

$$e^{A+H} - e^A = d_A \exp(H) + o(H) = \int_0^1 e^{t(A+H)}He^{(1-t)A}.$$

Par continuité de l'exponentielle en A on peut remplacer $e^{t(A+H)}$ par $e^{tA} + o(1)$ où le $o(1)$ est un o en H et pas en t, de sorte que lorsqu'on intègre sur t, il reste un $o(1)$, mais comme il est multiplié par H, il devient un $o(H)$, donc

$$\mathrm{d}_A \exp \cdot H + o(H) = \int_0^1 e^{tA}He^{(1-t)A}\mathrm{d}t + o(H).$$

En identifiant les parties linéaires, il vient l'expression demandée, que l'on aurait aussi pu trouver en développant en série les exponentielles.

$$\mathrm{d}_A \exp \cdot H = \int_0^1 e^{tA} H e^{(1-t)A} \mathrm{d}t.$$

3. On effectue un développement limité :

$$\begin{aligned}\left(e^{\frac{A}{n}} e^{\frac{B}{n}}\right)^n &= \left(\left(I + \frac{A}{n} + O(\frac{1}{n^2})\right)\left(I + \frac{B}{n} + O(\frac{1}{n^2})\right)\right)^n \\ &= \left(I + \frac{A+B}{n} + O(\frac{1}{n^2})\right)^n \\ &= e^{A+B} + o(1)\end{aligned}$$

En effet, soit A_n une suite convergente vers A :

$$\begin{aligned}\left\| e^A - \left(I + \frac{A_n}{n}\right)^n \right\| &\leq \|e^A - a^{A_n}\| + \left\| \sum_0^n \frac{1}{k!}(1 - \frac{(n-k)!}{n!} A_n^k + \sum_{n+1}^{\infty} \frac{A_n^k}{k!} \right\| \\ &\leq \|e^A - a^{A_n}\| + \sum_0^n \frac{1}{k!}(1 - \frac{(n-k)!}{n!} \|A_n\|^k + \sum_{n+1}^{\infty} \frac{\|A_n\|^k}{k!} \\ &= \|e^A - a^{A_n}\| + e^{\|A_n\|} - \left(1 + \frac{\|A_n\|}{n}\right)^n \to 0\end{aligned}$$

car le résultat est vrai dans $\mathbb{R}$. On montre de même la seconde formule.

Exercice 4 ⫽⫽⫽ : isométries de $\mathbb{R}^n$

1. Si f est une isométrie, f est de la forme $f(x) = \alpha + Ax$, où $\alpha \in \mathbb{R}^n$ et $A \in \mathcal{O}_n(\mathbb{R})$. On a alors $df(x) = A$, donc $df(x)$ est une isométrie pour tout x.
2. Soient h, k fixés. Posons $F(x) = \langle df(x).h, df(x).k\rangle$ pour tout $x \in \mathbb{R}^n$. Puisque $df(x)$ est une isométrie pour tout x, F est constante :

$$\forall h, k, \quad F(x) = \langle h, k\rangle$$

On doit donc avoir, pour tous x, l, $dF(x).l = 0$.
Or $dF(x).l = \langle d^{(2)}f(x).(h, l), df(x).k\rangle + \langle df(x).h, d^{(2)}f(x).(k, l)\rangle$.
3. L'égalité de la question précédente est exactement équivalente à $g(h, k, l) + g(k, h, l) = 0$ donc $g(h, k, l) = -g(k, h, l)$ pour tous h, k, l.
Puisque la différentielle seconde est symétrique, on a $g(h, k, l) = g(l, k, h)$ pour tous h, k, l.
Donc $g(h, k, l) = -g(k, h, l) = -g(l, h, k) = g(h, l, k) = g(k, l, h) = -g(l, k, h) = -g(h, k, l)$. Cela implique $g(h, k, l) = 0$ pour tous h, k, l.

4. Pour tout x, $df(x)$ est une isométrie de $\mathbb{R}^n$. C'est donc un isomorphisme. En particulier, $df(x)$ est surjective. Pour tous h, l, il existe k tel que $df(x).k = d^{(2)}f(x).(h,l)$. Puisque $g(h,k,l) = 0$, on doit avoir :

$$0 = g(h,k,l) = \langle d^{(2)}f(x).(h,l), df(x).k\rangle = \langle d^{(2)}f(x).(h,l), d^{(2)}f(x).(h,l)\rangle = ||d^{(2)}f(x).(h,l)||^2$$

Donc $d^{(2)}f(x).(h,l) = 0$ pour tous h, l. Donc $d^{(2)}f(x) = 0$ (pour tout x).
Puisque $d^{(2)}f = 0$, la fonction $x \to df(x)$ est constante. Il existe donc $A \in \mathcal{O}_n(\mathbb{R})$ telle que $df(x) = A$ pour tout $x \in \mathbb{R}^n$.
Alors $x \to f(x) - Ax$ est une fonction de différentielle nulle, c'est-à-dire une fonction constante. On en déduit que, pour un certain $\alpha \in \mathbb{R}^n$, on a :

$$\forall x \in \mathbb{R}^n, \quad f(x) = \alpha + Ax$$

Donc f est une isométrie.

Exercice 5 : lois de groupe sur $\mathbb{R}$

1. L'associativité nous donne $f(f(x,y),z) = f(x, f(y,z))$, donc en dérivant par rapport à la troisième coordonnée, on obtient

$$\partial_2 f(x * y, z) = \partial_2 f(y,z) \times \partial_2 f(x, y * z),$$

d'où la formule souhaitée en prenant $z = e$.
L'application $x \mapsto y * x = f(y,x)$ est $\mathcal{C}^1$, et c'est un difféomorphisme car sa réciproque $x \mapsto y^{-1} * x$ l'est également. Sa dérivée ne s'annule donc pas et garde donc un signe constant. Comme c'est le cas pour tout y, $\partial_2 f(y,e)$ garde un signe constant par continuité, nécessairement positif par l'égalité précédente.
2. On dérive l'égalité par rapport à y :

$$\varphi'(x * y) \times \partial_2 f(x,y) = \varphi'(y).$$

On évalue ensuite en $y = e$ pour obtenir

$$\varphi'(x) = \frac{\varphi'(e)}{\partial_2 f(x,e)}.$$

Comme $\varphi(e) = 0$, la formule est donc nécessaire par intégration.
3. Réciproquement, comme $\partial_2 f(t,e)$ est strictement positive, φ est bien définie, et $\mathcal{C}^1$ de dérivée strictement positive. C'est donc bien un difféomorphisme. Il ne reste qu'à vérifier qu'il conserve bien les opérations. En utilisant la formule de la première question, on voit que les deux fonctions $\varphi(x * y)$ et $\varphi(x) + \varphi(y)$ ont la même dérivée en y, de plus elles coïncident lorsque $y = e$. Elles sont donc égales et le tour est joué.

Exercice 6 : fonction à dérivées successives prescrites

1. Notons χ la fonction telle que $\chi(x) = 0$ si $x \leq 0$ et $\chi(x) = \exp(-1/x)$ si $x > 0$. C'est une fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, dont l'ensemble des zéros est exactement $\mathbb{R}^-$.

Pour montrer que χ est de classe $\mathcal{C}^\infty$, on donne seulement le principe de la démonstration : il suffit de démontrer que, pour toute fonction P polynomiale, $x \to \frac{P(x)}{x^n}e^{-1/x}$ est une fonction continue et dérivable, dont la dérivée est de la forme $x \to \frac{Q(x)}{x^n}e^{-1/x}$ pour une autre fonction polynomiale Q. Ce résultat permet de démontrer par récurrence que χ est de classe $\mathcal{C}^n$ pour tout n.
Posons $\chi_2(x) = \chi(x+1/2)\chi(-1/3-x)$, pour tout $x \in \mathbb{R}$. C'est une fonction $\mathcal{C}^\infty$, nulle sur $]-\infty;-1/2] \cup [-1/3;+\infty[$ et strictement positive sur $]-1/2;-1/3[$. On note F sa primitive qui vaut 0 sur $]-\infty;-1/2]$. La fonction F est constante non-nulle sur $[-1/3;+\infty[$.
La fonction $\phi(x) = \frac{F(x)F(-x)}{F(0)^2}$ est donc de classe $\mathcal{C}^\infty$. Elle vaut 0 sur $\mathbb{R} - [-1/2;1/2]$ donc est à support dans $]-1;1[$. Sur $[-1/3;1/3]$, elle est constante, de valeur 1.

2. Sur $\mathbb{R}-]-\epsilon_n;\epsilon_n[$, g_n est identiquement nulle donc l'inégalité est vérifiée. Il suffit de montrer que, pour ϵ_n assez petit, elle est aussi vérifiée sur $]-\epsilon_n;\epsilon_n[$.
Par la formule de dérivée des produits :

$$g_n^{(\alpha)}(x) = c_n \sum_{s=0}^{\alpha} \binom{\alpha}{s} \frac{1}{\epsilon_n^s} \phi^{(s)}\left(\frac{x}{\epsilon_n}\right) \frac{x^{n-(\alpha-s)}}{(n-(\alpha-s))!}$$

Donc, lorsque $|x| < \epsilon_n$:

$$|g_n^{(\alpha)}(x)| \leq |c_n| \epsilon_n^{n-\alpha} \sum_{s=0}^{\alpha} \binom{\alpha}{s} \left| \phi^{(s)}\left(\frac{x}{\epsilon_n}\right) \right| \frac{1}{(n-\alpha+s)!}$$

Pour tout s, la fonction $\phi^{(s)}$ est uniformément bornée sur $\mathbb{R}$ (car elle est à support compact). Pour tout $\alpha < n$, le terme précédent tend donc vers 0 uniformément en x lorsque $\epsilon_n \to 0$. En particulier, si on choisit ϵ_n assez petit, on peut avoir $||g_n^{(\alpha)}||_\infty \leq 2^{-n}$ pour tout $\alpha < n$.

3. On choisit les ϵ_n comme dans la question précédente, de sorte que $||g_n^{(\alpha)}||_\infty \leq 2^{-n}$ pour tous n et α tels que $\alpha < n$.
Pour tout $\alpha \in \mathbb{N}$, la série $\sum_n g_n^{(\alpha)}$ converge alors normalement. La fonction f est donc de classe $\mathcal{C}^\infty$ et, pour tout α :

$$\begin{aligned} & f^{(\alpha)} = \sum_n g_n^{(\alpha)} \\ \Rightarrow \qquad & f^{(\alpha)}(0) = g_\alpha^{(\alpha)}(0) = c_\alpha \end{aligned}$$

En effet, pour tout n, $g_n(x) = c_n \frac{x^n}{n!}$ au voisinage de 0 donc $g_n^{(\alpha)}(0) = c_n$ si $\alpha = n$ et 0 sinon.

Exercice 7 : fonctions homogènes

1. Lorsque $h \in \mathbb{R}$ tend vers 0 : $f((1+h)x) = f(x) + df(x).(hx) + o(h) = f(x) + hdf(x).x + o(h)$.
Or $f((1+h)x) = (1+h)^k f(x) = (1+kh+o(h))f(x) = f(x) + h(kf(x)) + o(h)$.
Par unicité du développement limité, $df(x).x = kf(x)$.

$f(tx+u) = f(tx) + df(tx).u + o(u)$
De plus, $f(tx+u) = f(t(x+u/t)) = t^k f(x+u/t) = t^k(f(x) + df(x).(u/t) + o(u)) = f(tx) + t^{k-1}df(x).u + o(u)$.

Par unicité de la différentielle, on a donc $df(tx) = t^{k-1}df(x)$.

2. On procède par récurrence sur k.

Pour $k = 0$, c'est vrai : pour tout $x \in E$ et tout $t \in \mathbb{R}$, $f(tx) = f(x)$. Pour $t = 0$, cela implique que, pour tout x, $f(x) = f(0)$.

Si c'est vrai pour $k-1$, démontrons-le pour k. Si f est homogène de degré k et de classe $\mathcal{C}^k$, df est homogène de degré $k-1$ et de classe $\mathcal{C}^{k-1}$, d'après la deuxième inégalité de la première question (cette égalité est aussi valable en $t = 0$, par continuité).

On a donc, pour tout x, $df(x) = \frac{1}{(k-1)!}df^{(k)}(0)(x, ..., x)$, par hypothèse de récurrence.

D'après la première question, cela implique, pour tout $x \neq 0$, $f(x) = \frac{1}{k}df(x).x = \frac{1}{k!}df^{(k)}(0)(x, ..., x)$. C'est aussi vrai si $x = 0$, par continuité.

3. Soit $\phi : \mathbb{R} \to \mathbb{R}$ une application continue et 2π-périodique telle que $\phi(x + \pi) = -\phi(x)$ pour tout $x \in \mathbb{R}$.

L'application $f : \mathbb{R}^2 \approx \mathbb{C} \to \mathbb{R}$ telle que $f(\rho e^{i\theta}) = \rho\phi(\theta)$ est une application homogène de degré 1. Pourtant, elle n'est pas nécessairement linéaire.

Exercice 8 ⫽⫽⫽ : descente de gradient

1. Montrons d'abord que le minimum est atteint.

Soit $x_0 \in \mathbb{R}^n$ quelconque. Soit $R > 0$ tel que, pour tout x tel que $||x|| > R$, $f(x) > f(x_0)$. Soit m le minimum de f sur $\overline{B}(0, R)$. Il est atteint car la convexité de f implique sa continuité. C'est aussi le minimum sur tout $\mathbb{R}^n$.

Montrons maintenant que le point auquel le minimum est atteint est unique.

Supposons par l'absurde que le minimum est atteint en deux points distincts x_0 et x_1. Alors, comme f est strictement convexe, $f\left(\frac{x_0+x_1}{2}\right) < \frac{f(x_0)+f(x_1)}{2} = \min f$. C'est absurde.

2. a)

$$\begin{aligned}
f(x_{n+1}) - f(x_n) &= \int_0^1 \langle \nabla f(x_n + t(x_{n+1} - x_n)), x_{n+1} - x_n \rangle dt \\
&= \int_0^1 \langle \nabla f(x_n), x_{n+1} - x_n \rangle dt \\
&\quad + \int_0^1 \langle \nabla f(x_n + t(x_{n+1} - x_n)) - \nabla f(x_n), x_{n+1} - x_n \rangle dt \\
&\leq \int_0^1 \langle \nabla f(x_n), x_{n+1} - x_n \rangle dt + \int_0^1 tL||x_{n+1} - x_n||^2 dt \\
&= \langle \nabla f(x_n), x_{n+1} - x_n \rangle + \frac{L}{2}||x_{n+1} - x_n||^2 \\
&= -\alpha(1 - \frac{L\alpha}{2})||\nabla f(x_n)||^2
\end{aligned}$$

b) La fonction f étant convexe, on a, pour tous y_1, y_2, $f(y_1) - f(y_2) \geq \langle \nabla f(y_2), y_1 - y_2 \rangle$. Pour $y_1 = x^*$ et $y_2 = x_n$, on obtient $f(x_n) \leq f(x^*) + \langle \nabla f(x_n), x_n - x^* \rangle$.

$$\begin{aligned}
f(x_{n+1}) &\leq f(x_n) - ||\nabla f(x_n)||^2 \alpha(1 - \frac{L\alpha}{2}) \\
&\leq f(x_n) - \frac{\alpha}{2}||\nabla f(x_n)||^2 \\
&\leq f(x^*) + \langle \nabla f(x_n), x_n - x^* \rangle - \frac{\alpha}{2}||\nabla f(x_n)||^2 \\
&= f(x^*) + \frac{1}{2\alpha}(2\alpha\langle \nabla f(x_n), x_n - x^* \rangle - \alpha^2||\nabla f(x_n)||^2) \\
&= f(x^*) + \frac{1}{2\alpha}(2\langle x_n - x_{n+1}, x_n - x^* \rangle - ||x_n - x_{n+1}||^2) \\
&= f(x^*) + \frac{1}{2\alpha}(||x_n - x^*||^2 - ||x_{n+1} - x^*||^2)
\end{aligned}$$

c) Puisque, pour tout n, $f(x_{n+1}) - f(x^*) \leq \frac{1}{2\alpha}(||x_n - x^*||^2 - ||x_{n+1} - x^*||^2)$:

$$\sum_{n=1}^{N}(f(x_n) - f(x^*)) \leq \frac{1}{2\alpha}(||x_0 - x^*||^2 - ||x_N - x^*||^2) \leq \frac{||x_0 - x^*||^2}{2\alpha}$$

Puisque $(f(x_n))_n$ est une suite décroissante (d'après la question a)), cela donne, pour tout $N \geq 1$:

$$f(x_N) - f(x^*) \leq \frac{1}{N}\sum_{n=1}^{N}(f(x_n) - f(x^*)) \leq \frac{1}{2N\alpha}||x_0 - x^*||^2$$

3. a) Notons $g(x) = f(x) - \frac{m}{2}||x||^2$. Cette fonction est de classe $\mathcal{C}^1$ (puisque f l'est) et vérifie $\nabla g(x) = \nabla f(x) - mx$.

Puisque g est convexe, on a, pour tous x, y :

$$g(y) - g(x) \geq \langle \nabla f(x) - mx, y - x \rangle = \langle \nabla f(x), y - x \rangle - m\langle x, y - x \rangle$$

On obtient donc :

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle + \frac{m}{2}||y||^2 - \frac{m}{2}||x||^2 - m\langle x, y - x \rangle = \langle \nabla f(x), y - x \rangle + \frac{m}{2}||x - y||^2$$

b) D'après la question a), $\langle \nabla f(x_n), x_n - x^* \rangle \geq f(x_n) - f(x^*) + \frac{m}{2}||x_n - x^*||^2 \geq \frac{m}{2}||x_n - x^*||^2$. Donc :

$$\begin{aligned}
||x_{n+1} - x^*||^2 &= ||x_n - x^*||^2 - 2\alpha\langle \nabla f(x_n), x_n - x^* \rangle + \alpha^2||\nabla f(x_n)||^2 \\
&\leq ||x_n - x^*||^2 - m\alpha||x_n - x^*||^2 + \alpha^2||\nabla f(x_n)||^2
\end{aligned}$$

Puisque ∇f est L-lipschitzienne et $\nabla f(x^*) = 0$, on a, pour tout n, $||\nabla f(x_n)|| \leq L||x_n - x^*||$. On obtient donc l'inégalité :

$$||x_{n+1} - x^*||^2 \leq (1 - m\alpha + L^2\alpha^2)||x_n - x^*||^2$$

Si α est suffisamment petit pour que $1 - m\alpha + L^2\alpha^2 < 1$, on a, pour tout n :

$$||x_{n+1} - x^*|| \leq c||x_n - x^*||$$

avec $c = \sqrt{1 - m\alpha + L^2\alpha^2}$.

Par récurrence, cela implique bien l'inégalité demandée pour tout n.