LES EXERCICES §6.26 ET §7 SUPPLÉMENT 0.2

Tout est la formule pour une composée. Commençons avec “quelques” remarques générales concernant le calcul des dérivées et des dérivées partielles. La première remarque à faire est que **toutes** les formules pour le calcul de ces dérivées (partielles) sont des cas particuliers de la formule pour la dérivée d'une fonction composée ! Et cela déjà pour les fonctions d'une seule variable.

Regardons le cas de la formule $(f+g)' = f'+g'$. Évidemment il s'agit de fonctions f et g définies sur un même domaine, disons $I \subset \mathbf{R}$. Mais on peut définir la fonction $S : \mathbf{R}^2 \to \mathbf{R}$ par

$$S(u,v) = u+v$$

et la fonction $F : I \to \mathbf{R}^2$ par

$$F(x) = \big(f(x), g(x)\big) \ .$$

Et avec ces définition on a la formule

$$f+g = S \circ F \ .$$

Quand on calcule les différentielles de F et de S, on trouve les matrices

$$\begin{aligned}(DS)(u,v) &= \big((\partial_1 S)(u,v)\,;\,(\partial_2 S)(u,v)\big) = (1;1)\\ (DF)(x) &= \begin{pmatrix}(\partial_1 F_1)(x)\\(\partial_1 F_2)(x)\end{pmatrix} = \begin{pmatrix}f'(x)\\g'(x)\end{pmatrix} \ ,\end{aligned}$$

où on a été “cohérent” en parlant de la dérivée partielle $\partial_1 F_i$: la dérivée partielle de la i-ième composante de F par rapport à la première variable. Bien sûr c'est exagéré, car il n'y a qu'une seule variable, donc on aurait pu directement parler de la dérivée. Mais le noter comme $\partial_1 F_i$ au lieu de F_i' n'est pas faux (du tout).

Pour la différentielle de la composée $S \circ F$ on trouve donc

$$\big(D(S\circ F)\big)(x) = (DS)\big(F(x)\big)\cdot(DF)(x) = (1;1)\cdot\begin{pmatrix}f'(x)\\g'(x)\end{pmatrix} = 1\cdot f'(x) + 1\cdot g'(x) \ ,$$

ce qui est la formule bien connue.

On peut jouer le même jeu avec le produit en définissant l'application $P : \mathbf{R}^2 \to \mathbf{R}$ par

$$P(u,v) = uv \qquad \text{avec} \qquad (DP)(u,v) = \big((\partial_1 S)(u,v)\,;\,(\partial_2 S)(u,v)\big) = (v;u) \ .$$

La fonction produit $f\cdot g$ s'écrit alors comme la composée

$$f\cdot g = P\circ F \ ,$$

ce qui nous donne pour sa différentielle la formule

$$\begin{aligned}\big(D(P\circ F)\big)(x) &= (DP)\big(F(x)\big)\cdot(DF)(x) = (DP)\big(f(x),g(x)\big)\cdot(DF)(x)\\ &= \big(g(x);f(x)\big)\cdot\begin{pmatrix}f'(x)\\g'(x)\end{pmatrix} = g(x)\cdot f'(x) + f(x)\cdot g'(x) \ ,\end{aligned}$$

ce qui est de nouveau la formule bien connue.

Et de même pour le quotient : on définit l'application $Q : \mathbf{R} \times \mathbf{R}^* \to \mathbf{R}$ par

$$Q(u,v) = \frac{u}{v} \qquad \text{avec} \qquad (QP)(u,v) = \big((\partial_1 S)(u,v)\,;\,(\partial_2 S)(u,v)\big) = \left(\ \frac{1}{v}\ ;\ \frac{-u}{v^2}\ \right)\ .$$

La fonction produit f/g s'écrit alors comme la composée

$$\frac{f}{g} = Q \circ F\ ,$$

ce qui nous donne pour sa différentielle la formule

$$\begin{aligned}\big(D(Q\circ F)\big)(x) &= (DQ)\big(F(x)\big)\cdot(DF)(x) = (DQ)\big(f(x),g(x)\big)\cdot(DF)(x)\\ &= \left(\ \frac{1}{g(x)}\ ;\ -\frac{f(x)}{\big(g(x)\big)^2}\ \right)\cdot\begin{pmatrix} f'(x)\\ g'(x)\end{pmatrix} = \frac{f'(x)}{g(x)} + \frac{-f(x)}{\big(g(x)\big)^2}\cdot g'(x)\\ &= \frac{f'(x)\cdot g(x) - f(x)\cdot g'(x)}{\big(g(x)\big)^2}\ .\end{aligned}$$

Quand on passe à des fonctions de plusieurs variables à valeurs dans $\mathbf{R}^p$, il faut "changer" un peu les fonctions qu'on utilise. Pour la somme, ce n'est plus l'application $S : \mathbf{R}^2 \to \mathbf{R}$, mais l'application $S_v : \mathbf{R}^p \times \mathbf{R}^p \to \mathbf{R}^p$ définie par

$$S_v(u,v) = u + v\ ,$$

où maintenant u et v représentent des vecteurs dans $\mathbf{R}^p$ (d'où l'indice v dans S_v pour "vectorielle"). Et bien sûr, il faut "changer" la définition de la fonction F en $F : U \to \mathbf{R}^p \times \mathbf{R}^p$ définie par

$$F(x) = \big(f(x), g(x)\big)\ ,$$

où maintenant $U \subset \mathbf{R}^n$ désigne le domaine de définition de f et de g.

Dans ce cas la formule pour la différentielle de la composée $S_v \circ F$ nous donne directement le résultat 6.7.

Pour le produit il faut prendre ses précautions pour bien multiplier un vecteur par un nombre. On définit donc l'application $P_v : \mathbf{R}^p \times \mathbf{R} \to \mathbf{R}^p$ par

$$P_v(u,v) = u \cdot v\ ,$$

avec l'avertissement que u représente un vecteur et v un nombre. Autrement dit, on doit écrire en coordonnées

$$P_v(u_1,\ldots,u_p,v) = (u_1v, u_2v, \ldots, u_nv)\ .$$

Si on a des fonctions $f : U \to \mathbf{R}^p$ et $g : U \to \mathbf{R}$, on définira l'application $F : U \to \mathbf{R}^{p+1}$ par

$$F(x) = \big(f(x), g(x)\big) \equiv \big(f_1(x),\ldots,f_p(x),g(x)\big)\ .$$

On peut donc calculer les différentielles comme suit :

$$(DF)(x) = \underbrace{\begin{pmatrix} \partial_1 f_1(x) & \partial_2 f_1(x) & \ldots & \partial_n f_1(x)\\ \partial_1 f_2(x) & \partial_2 f_2(x) & \ldots & \partial_n f_2(x)\\ \vdots & \vdots & & \vdots\\ \partial_1 f_p(x) & \partial_2 f_p(x) & \ldots & \partial_n f_p(x)\\ \partial_1 g(x) & \partial_2 g(x) & \ldots & \partial_n g(x)\end{pmatrix} = \begin{pmatrix} Df(x)\\ Dg(x)\end{pmatrix}}_{\text{matrice de taille } (p+1)\times n}\ .$$

et

$$\begin{aligned}
(DP_v)(u,v) &= (DP_v)(u_1,\ldots,u_p,v) \\
&= \big(\ \partial_1 P_v(u,v)\,;\ \partial_2 P_v(u,v)\,;\ \ldots\ ;\ \partial_p P_v(u,v)\,;\ \partial_{p+1} P_v(u,v)\ \big) \\
&= \begin{pmatrix} \partial_1(P_v)_1(u,v) & \partial_2(P_v)_1(u,v) & \ldots & \partial_n(P_v)_1(u,v) & \partial_{p+1}(P_v)_1(u,v) \\ \vdots & \vdots & & \vdots & \\ \partial_1(P_v)_p(u,v) & \partial_2(P_v)_p(u,v) & \ldots & \partial_n(P_v)_p(u,v) & \partial_{p+1}(P_v)_p(u,v) \end{pmatrix} \\
&= \begin{pmatrix} v & 0 & \ldots & 0 & u_1 \\ 0 & v & \ldots & 0 & u_2 \\ \vdots & & & & \vdots \\ 0 & 0 & \ldots & v & u_p \end{pmatrix} = \Big(\ v\cdot \mathrm{Id}\,;\ u\ \Big) \quad \text{matrice de taille } p\times(p+1)\ .
\end{aligned}$$

Pour la différentielle de la composée $P_v \circ F = f\cdot g$ on trouve donc :

$$\begin{aligned}
\big(D(P_v\circ F)\big)(x) &= (DP_v)\big(F(x)\big)\cdot(DF)(x) = (DP_v)\big(f(x),g(x)\big)\cdot(DF)(x) \\
&= \Big(\ g(x)\cdot\mathrm{Id}\,;\ f(x)\ \Big)\cdot\begin{pmatrix} Df(x) \\ Dg(x)\end{pmatrix} \\
&= g(x)\cdot\mathrm{Id}\cdot(Df)(x) + f(x)\cdot(Dg)(x) \\
&= g(x)\cdot(Df)(x) + f(x)\cdot(Dg)(x)\ ,
\end{aligned}$$

ce qui est bien le résultat 6.7 (avec les symboles f et g échangés).

Quelques questions de l'exercice 6.26. • (ii) Pour voir la fonction F_2 comme une composée, on introduit la fonction $G : \mathbf{R}\to\mathbf{R}^2$ par

$$G(x) = \big(g(x),h(x)\big)\ ,$$

ce qui transforme la définition de la fonction F_2 en

$$F_2(x) = (f\circ G)(x) = f\big(G(x)\big) = f\big(g(x),h(x)\big)\ .$$

Pour sa différentielle (dérivée !) on trouve donc

$$(DF_2)(x) = \big(\,(\partial_1F_2)(x)\,\big) = (Df)\big(G(x)\big)\cdot(DG)(x)\ ,$$

et pour les différentielles de f et de G on a

$$(Df)(x,y) = \big(\,(\partial_1 f)(x,y)\,;\,(\partial_2 f)(x,y)\,\big) \quad \text{et} \quad (DG)(x) = \begin{pmatrix}(\partial_1G_1)(x)\\(\partial_1G_2)(x)\end{pmatrix} = \begin{pmatrix}g'(x)\\h'(x)\end{pmatrix}\ .$$

Au final on trouve donc

$$\begin{aligned}
(DF_2)(x) &= (Df)\big(G(x)\big)\cdot(DG)(x) = (Df)\big(g(x),h(x)\big)\cdot(DG)(x) \\
&= \Big(\,(\partial_1 f)\big(g(x),h(x)\big)\,;\,(\partial_2 f)\big(g(x),h(x)\big)\,\Big)\cdot\begin{pmatrix}g'(x)\\h'(x)\end{pmatrix} \\
&= (\partial_1 f)\big(g(x),h(x)\big)\cdot g'(x) + (\partial_2 f)\big(g(x),h(x)\big)\cdot h'(x)\ .
\end{aligned}$$

• (iii) Pour voir la fonction F_3 comme une composée, on définit la fonction $G : \mathbf{R}^2 \to \mathbf{R}^2$ par

$$G(x,y) = \Big(\ g\big(\pi_1(x,y)\big)\,,\ h\big(\pi_2(x,y)\big)\ \Big) \qquad \text{avec} \qquad \pi_1(x,y) = x \text{ et } \pi_2(x,y) = y\ .$$

Avec cette définition on a immédiatement l'égalité $F_3 = f \circ G$. Pour les différentielles de f, π_1, π_2 et G on trouve

$$\begin{aligned}
(Df)(x,y) &= \big(\,(\partial_1 f)(x,y)\,;\,(\partial_2 f)(x,y)\,\big)\\
(D\pi_1)(x,y) &= \big(\partial_1\pi_1(x,y)\,;\,\partial_2\pi_1(x,y)\big) = (1;0)\\
(D\pi_2)(x,y) &= \big(\partial_1\pi_2(x,y)\,;\,\partial_2\pi_2(x,y)\big) = (0;1)\\
(DG)(x) &= \begin{pmatrix}(\partial_1 G_1)(x,y) & (\partial_2 G_1)(x,y)\\ (\partial_1 G_2)(x,y) & (\partial_2 G_2)(x,y)\end{pmatrix} = \begin{pmatrix} g'(x) & 0\\ 0 & h'(x)\end{pmatrix}\ .
\end{aligned}$$

La différentielle DG se trouve aussi (officiellement ?) par l'application de la formule pour une composée :

$$\begin{aligned}
(DG)(x,y) &= \begin{pmatrix}(DG_1)(x,y)\\ (DG_2)(x,y)\end{pmatrix} = \begin{pmatrix}\big(D(g\circ\pi_1)\big)(x,y)\\ \big(D(h\circ\pi_2)\big)(x,y)\end{pmatrix}\\
&= \begin{pmatrix}(Dg)\big(\pi_1(x,y)\big)\cdot(D\pi_1)(x,y)\\ (Dh)\big(\pi_2(x,y)\big)\cdot(D\pi_2)(x,y)\end{pmatrix} = \begin{pmatrix} g'(x)\cdot(1;0)\\ h'(y)\cdot(0;1)\end{pmatrix}\\
&= \begin{pmatrix} g'(x) & 0\\ 0 & h'(x)\end{pmatrix}\ .
\end{aligned}$$

Pour la différentielle DF_3 on trouve donc

$$\begin{aligned}
(DF_3)(x,y) &= (Df)\big(G(x,y)\big)\cdot(DG)(x,y)\\
&= \Big(\ (\partial_1 f)\big(G(x,y)\big)\,;\,(\partial_2 f)\big(G(x,y)\big)\ \Big)\cdot\begin{pmatrix} g'(x) & 0\\ 0 & h'(x)\end{pmatrix}\\
&= \Big(\ (\partial_1 f)\big(g(x),h(y)\big)\cdot g'(x)\,;\,(\partial_2 f)\big(g(x),h(y)\big)\cdot h'(y)\ \Big)
\end{aligned}$$

• (vi) Pour voir F_6 comme une composée, on introduit la fonction $G : \mathbf{R}^2 \to \mathbf{R}^2$ définie par

$$G(x,y) = \Big(\ k(x,y), g\big(\pi_1(x,y)\big)\ \Big)\ .$$

Pour la différentielle de ce G on trouve

$$(DG)(x,y) = \begin{pmatrix}(\partial_1 G_1)(x,y) & (\partial_2 G_1)(x,y)\\ (\partial_1 G_2)(x,y) & (\partial_2 G_2)(x,y)\end{pmatrix} = \begin{pmatrix}(\partial_1 k)(x,y) & (\partial_2 k)(x,y)\\ g'(x) & 0\end{pmatrix}\ .$$

Comme dans le cas précédent, on peut calculer la différentielle $(DG_2)(x,y)$ aussi via la dérivée d'une composée :

$$\begin{aligned}
(DG_2)(x,y) &= \big(D(g\circ\pi_1)\big)(x,y) = (Dg)\big(\pi_1(x,y)\big)\cdot(D\pi_1)(x,y)\\
&= g'(x)\cdot(1;0) = \big(g'(x);0\big)\ ,
\end{aligned}$$

conformément à ce qu'on a écrit précédemment. Pour la différentielle DF_6 on trouve donc

$$\begin{aligned}
(DF_6)(x,y) &= (Df)\big(G(x,y)\big)\cdot(DG)(x,y)\\
&= \Big(\ (\partial_1 f)\big(G(x,y)\big)\,;\,(\partial_2 f)\big(G(x,y)\big)\ \Big)\cdot\begin{pmatrix}(\partial_1 k)(x,y) & (\partial_2 k)(x,y)\\ g'(x) & 0\end{pmatrix}\\
&= \Big(\ (\partial_1 f)\big(k(x,y),g(x)\big)\cdot(\partial_1 k)(x,y) + (\partial_2 f)\big(k(x,y),g(x)\big)\cdot g'(x)\ ;\\
&\qquad\qquad (\partial_1 f)\big(k(x,y),g(x)\big)\cdot(\partial_2 k)(x,y)\ \Big)
\end{aligned}$$

Un regard différent sur les résultats de 6.26. Il est "facile" de généraliser le résultat de 6.26.ii à la situation d'une fonction $f : \mathbf{R}^n \to \mathbf{R}$ et une fonction $g : \mathbf{R} \to \mathbf{R}^n$ avec $F(x) = f\big(g(x)\big)$, c'est-à-dire

$$F : \mathbf{R} \to \mathbf{R} \qquad \text{avec} \qquad F(x) = \big(g_1(x), g_2(x), \dots, g_n(x)\big)\ .$$

Avec exactement les mêmes arguments on trouve

$$\begin{aligned} F'(x) &= (Df)\big(g(x)\big) \cdot (Dg)(x) \\ &= \Big(\ (\partial_1 f)\big(g(x)\big)\ ;\ (\partial_2 f)\big(g(x)\big)\ ;\ \dots\ ;\ (\partial_n f)\big(g(x)\big)\ \Big) \cdot \begin{pmatrix} g_1'(x) \\ \vdots \\ g_n'(x) \end{pmatrix} \\ &= (\partial_1 f)\big(g(x)\big) \cdot g_1'(x) + \cdots + (\partial_n f)\big(g(x)\big) \cdot g_n'(x) \\ &= \sum_{i=1}^{n} (\partial_i f)\big(g(x)\big) \cdot g_i'(x)\ . \end{aligned} \tag{0.1}$$

Regardons de plus près la formule $F(x)$: dans l'expression

$$F(x) = f\big(g_1(x), g_2(x), \dots, g_n(x)\big)$$

on voit apparaître la variable x un certain nombre de fois (n pour être précis). Mettons quelques uns des ces occurrences en rouge :

$$f\big(g_1(\color{red}{x}), g_2(x), \dots, g_n(x)\big) \text{ ou } f\big(g_1(x), g_2(\color{red}{x}), \dots, g_n(x)\big) \text{ ou } f\big(g_1(x), g_2(x), \dots, g_n(\color{red}{x})\big)\ .$$

Il y a n endroits où apparaît x et il y a n termes dans la formule pour $F'(x)$. Mettons ces termes l'un à côté de l'autre pour comparer :

$$\begin{array}{lcl} f\big(g_1(\color{red}{x}), g_2(x), \dots, g_n(x)\big) & \leftrightarrow & (\partial_1 f)\big(g_1(\color{red}{x}), g_2(x), \dots, g_n(x)\big) \cdot g_1'(\color{red}{x}) \\ f\big(g_1(x), g_2(\color{red}{x}), \dots, g_n(x)\big) & \leftrightarrow & (\partial_2 f)\big(g_1(x), g_2(\color{red}{x}), \dots, g_n(x)\big) \cdot g_2'(\color{red}{x}) \\ f\big(g_1(x), g_2(x), \dots, g_n(\color{red}{x})\big) & \leftrightarrow & (\partial_n f)\big(g_1(x), g_2(x), \dots, g_n(\color{red}{x})\big) \cdot g_n'(\color{red}{x})\ . \end{array}$$

Quand on ne garde que les variables en rouges et qu'on n'écrit pas les autres, on aura les comparaisons

$$\begin{array}{lcl} f\big(g_1(\color{red}{x})\big) & \overset{\text{dérivée}}{\leftrightarrow} & (\partial_1 f)\big(g_1(\color{red}{x})\big) \cdot g_1'(\color{red}{x}) \\ f\big(g_2(\color{red}{x})\big) & \overset{\text{dérivée}}{\leftrightarrow} & (\partial_2 f)\big(g_2(\color{red}{x})\big) \cdot g_2'(\color{red}{x}) \\ f\big(g_n(\color{red}{x})\big) & \overset{\text{dérivée}}{\leftrightarrow} & (\partial_n f)\big(g_n(\color{red}{x})\big) \cdot g_n'(\color{red}{x})\ . \end{array}$$

À part le fait qu'on note des dérivées partielles, c'est exactement la formule our la dérivée d'une fonction composée ! Et même la dérivée partielle se comprend facilement : quand la variable en rouge se trouve à la i-ième place, on prend la dérivée partielle par rapport à cette même i-ième variable.

Le schéma qu'o observe est donc le suivant : pour trouver la dérivée par rapport à la variable x, on note chaque endroit où figure cette variable. Ensuite on ignore tous les autres variables et autres occurrences de la variable x et on calcule sa dérivée comme d'habitude. Et ensuite on prend la somme sur toutes les occurrences de x.

Repassons les questions (ii), (iii) et (vi) avec cette idée. • Pour (ii) on a donc la fonction

$$F_2(x) = f\big(g(x), h(x)\big)$$

avec deux occurrences de x :

$$f\big(g(\color{red}{x}), h(x)\big) \qquad \text{et} \qquad f\big(g(x), h(\color{red}{x})\big)\ .$$

On regarde "donc" deux composées pour dériver :

$$f\big(g(x)\big) \qquad \text{et} \qquad f\big(h(x)\big) \ ,$$

mais on se rappelle que pour $g(x)$ c'est la première variable de f qui est concernée et pour $h(x)$ c'est la deuxième variable qui est concernée. On écrit "donc"

$$\begin{aligned} F'(x) &\cong \underbrace{f'\big(g(x)\big)}_{\text{première var.}} \cdot g'(x) + \underbrace{f'\big(h(x)\big)}_{\text{deuxième var.}} \cdot h(x) \\ \text{on remet les variables} \quad &= (\partial_1 f)\big(g(x), h(x)\big) \cdot g'(x) + (\partial_2 f)\big(g(x), h(x)\big) \cdot h(x) \ . \end{aligned}$$

• Mais cette méthode s'applique aussi bien aux cas où il y a plusieurs variables ! Car pour déterminer une différentielle, on calcule les dérivées partielles, ce qui revient à calculer une dérivée "ordinaire" en supposant que les autres variables sont de paramètres constants. Regardons ce que cela donne pour (iii). On a la fonction F_3 donnée par l'expression

$$F_3(x, y) = f\big(g(x), h(y)\big) \ .$$

Sa différentielle consiste de la matrice des deux dérivées partielles

$$(DF_3)(x, y) = \Big(\ (\partial_1 F_3)(x, y) \ ; \ (\partial_2 F_3) \ \Big) \ .$$

Et pour calculer ses dérivées partielles, on calcule une dérivée ordinaire en mettant les autres variables comme paramètres. Ainsi on a "donc"

$$(\partial_1 F_3(x, y) = H'(x) \qquad \text{avec} \qquad H(x) = F_3(x, y) = f\big(g(x), h(y)\big) \ .$$

On a donc une seule occurrence de x :

$$f\big(g(x), h(y)\big) \ ,$$

qu'on traite donc comme une composée en ignorant tous le reste :

$$f\big(g(x)\big) \quad \overset{\text{dérivée}}{\leftrightarrow} \quad \underbrace{f'\big(g(x)\big)}_{\text{première var.}} \cdot g'(x) \overset{\text{on remet}}{=} (\partial_1 f)\big(g(x), h(y)\big) \cdot g'(x) \ .$$

Et ceci est bien le résultat qu'on a trouvé ci-dessus comme première composante de $(DF_3)(x, y)$.

Pour la dérivée partielle par rapport à la deuxième variable on procède de la même manière (en raccourci) : on a une seule occurrence de la (deuxième) variable y :

$$f\big(g(x), h(y)\big) \overset{\text{on oublie}}{\to} f\big(h(y)\big) \overset{\text{on dérive}}{\to} \underbrace{f'\big(h(y)\big)}_{\text{deuxième var.}} \cdot h'(y) \overset{\text{on remet}}{\to} (\partial_2 f)\big(g(x), h(y)\big) \cdot h'(y) \ .$$

Et on retrouve bien la deuxième composante de $(DF_3)(x, y)$ déjà trouvé ci-dessus par une autre méthode.

• Regardons finalement la fonction $F_6(x,y) = f\big(k(x,y), g(x)\big)$. Pour la dérivée partielle $\partial_1 F_6$ il faut dériver par rapport à la variable x qui apparaît deux fois :

$$\begin{array}{rcl}
f\big(k({\color{red}x},y),g(x)\big) & \text{et} & f\big(k(x,y),g({\color{red}x})\big) \\
 & \downarrow \text{ on oublie} & \\
f\big(k(x)\big) & \text{et} & f\big(g(x)\big) \\
 & \downarrow \text{ on dérive et on somme} & \\
\underbrace{f'\big(k(x)\big)}_{\text{première var.}} \cdot \underbrace{k'(x)}_{\text{première var.}} & + & \underbrace{f'\big(g(x)\big)}_{\text{deuxième var.}} \cdot g'(x) \\
 & \downarrow \text{ on remet} & \\
(\partial_1 f)\big(k({\color{red}x},y),g(x)\big)\cdot(\partial_1 k)({\color{red}x},y) & + & (\partial_2 f)\big(k(x,y),g({\color{red}x})\big)\cdot g'({\color{red}x}) \ .
\end{array}$$

Ce résultat est bien la première composante dans l'expression pour $(DF_6)(x,y)$ qu'on a trouvé auparavant.

Pour la dérivée partielle $\partial_2 F_6$ on procède de la même manière, sauf qu'ici la deuxième variable y n'apparaît qu'une seule fois :

$$\begin{array}{rl}
 & f\big(k(x,{\color{red}y}),g(x)\big) \\
\text{on oublie } \downarrow & \\
 & f\big(k(y)\big) \\
\text{on dérive et on somme } \downarrow & \\
 & \underbrace{f'\big(k(y)\big)}_{\text{première var.}} \cdot \underbrace{k'(y)}_{\text{deuxième var.}} \\
\text{on remet } \downarrow & \\
 & (\partial_1 f)\big(k(x,{\color{red}y}),g(x)\big)\cdot(\partial_2 k)(x,{\color{red}y}) \ .
\end{array}$$

Et de nouveau on retrouve bien le résultat qu'on connaît déjà.

Mais la méthode est aussi robuste et ne dépend pas (trop) comment on écrit la fonction. Regardons un exemple "facile" qu'on va écrire de deux façons différentes. Considérons la fonction $f(x) = x^5$. Ça paraît évident que sa dérivée est $5x^4$. Mais regardons cette fonction d'une autre façons :

$$f(x) = x^2 \cdot x^3 \ .$$

Écrit comme ça, la variable x apparaît deux fois. Appliquons notre recette :

$$\begin{array}{rcl}
{\color{red}x^2}\cdot x^3 & \text{et} & x^2\cdot {\color{red}x^3} \\
 & \downarrow \text{ on met le reste en constante} & \\
{\color{red}x^2}\cdot C_1 & \text{et} & C_2\cdot {\color{red}x^3} \\
 & \downarrow \text{ on dérive et on somme} & \\
{\color{red}2x}\cdot C_1 & + & C_2\cdot {\color{red}3x^2} \\
 & \downarrow \text{ on remet les constantes} & \\
{\color{red}2x}\cdot x^3 & + & x^2\cdot {\color{red}3x^2} \\
 & \downarrow \text{ on ramasse} & \\
2x^4 & + & 3x^4 \quad = \quad 5x^4 \ .
\end{array}$$

L'exercice §7 série supplémentaires 0.2.i. On a deux fonctions d'une seule variable réelle f et g et on définit la fonction $h : \mathbf{R}^2 \to \mathbf{R}$ par

$$h(x,y) = f(x) \cdot \cos\bigl(g(y)\bigr) \ .$$

On nous demande de calculer la matrice Hessienne

$$H(x,y) = \begin{pmatrix} (\partial_1\partial_1 h)(x,y) & (\partial_1\partial_2 h)(x,y) \\ (\partial_2\partial_1 h)(x,y) & (\partial_2\partial_2 h)(x,y) \end{pmatrix} \ .$$

Pour calculer ces dérivées partielles secondes, il faut "donc" commencer avec les dérivées partielles premières. Un calcul direct nous donne

$$(\partial_1 h)(x,y) = f'(x) \cdot \cos\bigl(g(y)\bigr) \qquad \text{et} \qquad (\partial_2 h)(x,y) = -f(x) \cdot \sin\bigl(g(y)\bigr) \cdot g'(y) \ .$$

De là le calcul des dérivées partielles secondes est aussi direct :

$$\begin{aligned} (\partial_1\partial_1 h)(x,y) &= f''(x) \cdot \cos\bigl(g(y)\bigr) \\ (\partial_2\partial_1 h)(x,y) &= -f'(x) \cdot \sin\bigl(g(y)\bigr) \cdot g'(y) = (\partial_1\partial_2 h)(x,y) \\ (\partial_2\partial_2 h)(x,y) &= -f(x) \cdot \cos\bigl(g(y)\bigr) \cdot g'(y) \cdot g'(y) - f(x) \cdot \sin\bigl(g(y)\bigr) \cdot g''(y) \ . \end{aligned}$$

On trouve donc pour la matrice Hessienne :

$$H(x,y) = \begin{pmatrix} f''(x)\cdot\cos(g(y)) & -f'(x)\cdot\sin(g(y))\cdot g'(y) \\ \\ -f'(x)\cdot\sin(g(y))\cdot g'(y) & -f(x)\cdot\cos(g(y))\cdot g'(y)^2 - f(x)\cdot\sin(g(y))\cdot g''(y) \end{pmatrix} \ .$$

L'exercice §7 série supplémentaires 0.2.ii. Dans cette question la fonction f est définie sur $\mathbf{R}^2$ à valeurs réelles et on définit $g : \mathbf{R}^2 \to \mathbf{R}$ par

$$g(x,y) = f\bigl(x^2+y, \sin(xy)\bigr) \ .$$

Comme dans le cas précédent, pour calculer les dérivées partielles secondes il faut d'abord calculer les dérivées partielles premières. Faisons le de deux façons différentes selon les deux procédures décrites ci-dessus.

- *Première méthode.* On définit l'application $F : \mathbf{R}^2 \to \mathbf{R}^2$ par

$$F(x,y) = \bigl(x^2+y, \sin(xy)\bigr) \ ,$$

et on constate qu'on a $g = f \circ F$. On calcule la différentielle de F comme

$$(DF)(x,y) = \begin{pmatrix} 2x & 1 \\ y\cos(xy) & x\cos(xy) \end{pmatrix}$$

ce qui nous permet de calculer Dg comme

$$\begin{aligned} (Dg(x,y) &= (Df)\bigl(F(x,y)\bigr) \cdot (DF)(x,y) \\ &= \Bigl(\ (\partial_1 f)\bigl(F(x,y)\bigr) \ ; \ (\partial_2 f)\bigl(F(x,y)\bigr) \ \Bigr) \cdot \begin{pmatrix} 2x & 1 \\ y\cos(xy) & x\cos(xy) \end{pmatrix} \\ &= \Bigl(\ (\partial_1 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot 2x + (\partial_2 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot y\cos(xy) \ ; \\ &\qquad (\partial_1 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot 1 + (\partial_2 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot x\cos(xy) \ \Bigr) \ . \end{aligned}$$

Autrement dit, on a les deux dérivées partielles

$$\begin{aligned} (\partial_1 g)(x,y) &= (\partial_1 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot 2x + (\partial_2 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot y\cos(xy) \\ (\partial_2 g)(x,y) &= (\partial_1 f)\bigl(x^2+y, \sin(xy)\bigr) + (\partial_2 f)\bigl(x^2+y, \sin(xy)\bigr) \cdot x\cos(xy) \ . \end{aligned}$$

• *Deuxième méthode.* Quand on regarde les occurrences de x et de y dans l'expression pour g, on voit que x et y apparaissent deux fois. Commençons avec x :

$$\begin{array}{rcl} f\big(x^2+y,\sin(xy)\big) & \text{et} & f\big(x^2+y,\sin(xy)\big) \\ & \downarrow \text{ on oublie} & \\ f(x^2+y) & \text{et} & f\big(\sin(xy)\big) \\ & \downarrow \text{ on dérive et on somme} & \\ \underbrace{f'(x^2+y)}_{\text{première var.}}\cdot 2x & + & \underbrace{f'\big(\sin(xy)\big)}_{\text{deuxième var.}}\cdot y\cos(xy) \\ & \downarrow \text{ on remet} & \\ (\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x & + & (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)\ . \end{array}$$

Et ensuite pour y :

$$\begin{array}{rcl} f\big(x^2+y,\sin(xy)\big) & \text{et} & f\big(x^2+y,\sin(xy)\big) \\ & \downarrow \text{ on oublie} & \\ f(x^2+y) & \text{et} & f\big(\sin(xy)\big) \\ & \downarrow \text{ on dérive et on somme} & \\ \underbrace{f'(x^2+y)}_{\text{première var.}}\cdot 1 & + & \underbrace{f'\big(\sin(xy)\big)}_{\text{deuxième var.}}\cdot x\cos(xy) \\ & \downarrow \text{ on remet} & \\ (\partial_1 f)\big(x^2+y,\sin(xy)\big) & + & (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)\ . \end{array}$$

Et on retrouve bien les mêmes réponses que précédemment.

Pour calculer les dérivées partielles secondes de g, on a de nouveau deux approches, voire trois si on fait des distinctions plus fines. La première est de considérer que la différentielle première Dg est une application sur $\mathbf{R}^2$ et à valeurs dans $\mathbf{R}^2$:

$$Dg : \mathbf{R}^2 \to \mathbf{R}^2\ , \qquad (Dg)(x,y) = \big(\ (\partial_1 g)(x,y)\ ,\ (\partial_2 g)(x,y)\ \big)$$

et de calculer la différentielle de cette application. En tant qu'application de $\mathbf{R}^2 \to \mathbf{R}^2$ sa différentielle $\big(D(Dg)\big)(x,y)$ en un point (x,y) sera donc une matrice de taille 2×2 qui consiste (bien évidemment) des quatre dérivées partielles secondes $(\partial_i\partial_j g)(x,y)$.

Mais on peut (et on le fait le plus souvent) effectuer ce calcul composante par composante selon 6.5 qui nous donne

$$(Dg)(x,y) = \begin{pmatrix} (\partial_1 g)(x,y) \\ (\partial_2 g)(x,y) \end{pmatrix} \qquad \Longrightarrow \qquad \big(D(Dg)\big)(x,y) = \begin{pmatrix} \big(D(\partial_1 g)\big)(x,y) \\ \big(D(\partial_2 g)\big)(x,y) \end{pmatrix}\ .$$

Chaque fonction $\partial_i g$ est à valeurs réelles et sa différentielle sera donnée par

$$\big(D(\partial_i g)\big)(x,y) = \Big(\ (\partial_1\partial_i g)(x,y)\ ;\ (\partial_2\partial_i g)(x,y)\ \Big)\ .$$

Et comme troisième méthode on peut regarder les occurrences de chaque variable et dériver-sommer pour trouver les dérivées partielles de chaque $\partial_i g$.

Ces trois façons de procéder mènent au même résultat. Mais avant d'entamer une de ces méthodes, il y a une **remarque importante** à faire : dans les expressions pour $\partial_i g$ il faut traiter les fonctions $\partial_i f$ comme des fonctions à part. En particulier il faut traiter une expression de la forme $(\partial_i f)\big(x^2+y,\sin(xy)\big)$ de la même façon qu'on a traité l'expression $f\big(x^2+y,\sin(xy)\big)$!

Vu la longueur des expressions de $\partial_i g$ on opte ici pour la dernière méthode qui prend le moins de place (mais qui est déjà un peu long). Commençons avec $\partial_1\partial_1 g$: dans $\partial_1 g$ la variable x apparaît à 6 endroits :

$$(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)\ .$$

L'expression pour $\partial_1\partial_1 g$ consistera donc d'une somme de 6 termes ; traitons les un par un.

occurrence 1 : $\underbrace{(\partial_1 f)\big(x^2+y,\sin(xy)\big)}_{\text{première var.}}\cdot 2x+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_1(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot 2x\cdot 2x$

occurrence 2 : $\underbrace{(\partial_1 f)\big(x^2+y,\sin(xy)\big)}_{\text{deuxième var.}}\cdot 2x+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_2(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)\cdot 2x$

occurrence 3 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2$

occurrence 4 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x+\underbrace{(\partial_2 f)\big(x^2+y,\sin(xy)\big)}_{\text{première var.}}\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_1(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot 2x\cdot y\cos(xy)$

occurrence 5 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x+\underbrace{(\partial_2 f)\big(x^2+y,\sin(xy)\big)}_{\text{deuxième var.}}\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_2(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)\cdot y\cos(xy)$

occurrence 6 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cdot\big(-y\sin(xy)\big)\ .$

Quand on prend la somme, on obtient l'expression (assez longue, c'est vrai)

$$\begin{aligned}(\partial_1\partial_1 g)(x,y) &= (\partial_1\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot(2x)2\\ &\quad+(\partial_2\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2xy\cos(xy)\\ &\quad+(\partial_1\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot 2xy\cos(xy)\\ &\quad+(\partial_2\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot\big(y\cos(xy)\big)^2\\ &\quad+(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot\big(-y^2\sin(xy)\big)\ .\end{aligned}$$

Bien sûr, on aurait pu combiner le deuxième et troisième terme dans cette somme ci-dessus en un seul en utilisant le théorème de Schwartz 7.4. La raison pour ne pas le faire ici est qu'en préservant les deux termes, il est plus facile de voir qu'on peut écrire ce résultat sous forme matricielle comme suit :

$$(\partial_1\partial_1 g)(x,y) = \big(\ 2x\ ;\ y\cos(xy)\ \big)\cdot\begin{pmatrix}(\partial_1\partial_1 f)(Z) & (\partial_1\partial_2 f)(Z)\\ (\partial_2\partial_1 f)(Z) & (\partial_2\partial_2 f)(Z)\end{pmatrix}\cdot\begin{pmatrix}2x\\ y\cos(xy)\end{pmatrix}$$

$$+\big(\ (\partial_1 f)(Z)\ ;\ (\partial_2 f)(Z)\ \big)\cdot\begin{pmatrix}2\\ -y^2\sin(xy)\end{pmatrix}\ ,$$

où on a abrégé l'argument $\big(x^2+y,\sin(xy)\big)$ comme Z pour améliorer la lisibilité.

Faisons maintenant le calcul analogue pour $\partial_2(\partial_1 g)$ en remarquant que la variable y apparaît aussi 6 fois dans l'expression pour $\partial_1 g$, ce qui donne les 6 calculs

occurrence 1 : $\underbrace{(\partial_1 f)\big(x^2+y,\sin(xy)\big)}_{\text{première var.}}\cdot 2x + (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_1(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot 1\cdot 2x$

occurrence 2 : $\underbrace{(\partial_1 f)\big(x^2+y,\sin(xy)\big)}_{\text{deuxième var.}}\cdot 2x + (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_2(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)\cdot 2x$

occurrence 3 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x + \underbrace{(\partial_2 f)\big(x^2+y,\sin(xy)\big)}_{\text{première var.}}\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_1(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot 1\cdot y\cos(xy)$

occurrence 4 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x + \underbrace{(\partial_2 f)\big(x^2+y,\sin(xy)\big)}_{\text{deuxième var.}}\cdot y\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_2(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)\cdot y\cos(xy)$

occurrence 5 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x + (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cdot\cos(xy)$

$\downarrow$ on dérive

$(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot 1\cdot\cos(xy)$

occurrence 6 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)\cdot 2x + (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cos(xy)$

$\downarrow$ on dérive

$(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot y\cdot\big(-x\sin(xy)\big)\ .$

Quand on prend la somme, on obtient l'expression (en abrégeant $\big(x^2+y,\sin(xy)\big)$ comme Z)

$$\begin{aligned}(\partial_2\partial_1 g)(x,y) &= \big(\partial_1(\partial_1 f)\big)(Z)\cdot 2x\\ &\quad+\big(\partial_2(\partial_1 f)\big)(Z)\cdot x\cos(xy)\cdot 2x\\ &\quad+\big(\partial_1(\partial_2 f)\big)(Z)\cdot y\cos(xy)\\ &\quad+\big(\partial_2(\partial_2 f)\big)(Z)\cdot x\cos(xy)\cdot y\cos(xy)\\ &\quad+(\partial_2 f)(Z)\cdot\big(\cos(xy)-xy\sin(xy)\big)\ .\end{aligned}$$

Et comme pour $\partial_1\partial_1 g$, on peut mettre ceci sous forme matricielle comme

$$\begin{aligned}(\partial_2\partial_1 g)(x,y) = \big(\ 1\ ;\ x\cos(xy)\ \big)\cdot\begin{pmatrix}(\partial_1\partial_1 f)(Z) & (\partial_1\partial_2 f)(Z)\\ (\partial_2\partial_1 f)(Z) & (\partial_2\partial_2 f)(Z)\end{pmatrix}\cdot\begin{pmatrix}2x\\ y\cos(xy)\end{pmatrix}\\ +\big(\ (\partial_1 f)(Z)\ ;\ (\partial_2 f)(Z)\ \big)\cdot\begin{pmatrix}0\\ \cos(xy)-xy\sin(xy)\end{pmatrix}\ .\end{aligned}$$

Selon le théorème de Schwartz, $\partial_1\partial_2 g = \partial_2\partial_1 g$, donc il nous reste le calcul de $\partial_2\partial_2 g$, qu'on fait de la même manière. La variable y n'apparaissant que 5 fois dans l'expression de $\partial_2 g$, on aura donc 5 termes :

occurrence 1 : $\underbrace{(\partial_1 f)\big(x^2+y,\sin(xy)\big)}_{\text{première var.}}+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_1(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot 1$

occurrence 2 : $\underbrace{(\partial_1 f)\big(x^2+y,\sin(xy)\big)}_{\text{deuxième var.}}+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_2(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)$

occurrence 3 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)+\underbrace{(\partial_2 f)\big(x^2+y,\sin(xy)\big)}_{\text{première var.}}\cdot x\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_1(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot 1\cdot x\cos(xy)$

occurrence 4 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)+\underbrace{(\partial_2 f)\big(x^2+y,\sin(xy)\big)}_{\text{deuxième var.}}\cdot x\cos(xy)$

$\downarrow$ on dérive

$\big(\partial_2(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)\cdot x\cos(xy)$

occurrence 5 : $(\partial_1 f)\big(x^2+y,\sin(xy)\big)+(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)$

$\downarrow$ on dérive

$(\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot x\cdot\big(-x\sin(xy)\big)\ .$

En prenant la somme, on obtient l'expression (toujours en abrégeant $\big(x^2+y, \sin(xy)\big)$ comme Z)

$$\begin{aligned}(\partial_2\partial_2 g)(x,y) &= \big(\partial_1(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\\ &\quad+ \big(\partial_2(\partial_1 f)\big)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)\\ &\quad+ \big(\partial_1(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot x\cos(xy)\\ &\quad+ \big(\partial_2(\partial_2 f)\big)\big(x^2+y,\sin(xy)\big)\cdot \big(x\cos(xy)\big)^2\\ &\quad+ (\partial_2 f)\big(x^2+y,\sin(xy)\big)\cdot x\cdot\big(-x\sin(xy)\big)\ .\end{aligned}$$

Et comme pour $\partial_1\partial_1 g$ et $\partial_2\partial_2 g$, on peut mettre ceci sous forme matricielle comme

$$\begin{aligned}(\partial_2\partial_2 g)(x,y) &= \big(\ 1\ ;\ x\cos(xy)\ \big)\cdot\begin{pmatrix}(\partial_1\partial_1 f)(Z) & (\partial_1\partial_2 f)(Z)\\ (\partial_2\partial_1 f)(Z) & (\partial_2\partial_2 f)(Z)\end{pmatrix}\cdot\begin{pmatrix}1\\ x\cos(xy)\end{pmatrix}\\ &\quad+\big(\ (\partial_1 f)(Z)\ ;\ (\partial_2 f)(Z)\ \big)\cdot\begin{pmatrix}0\\ -x^2\sin(xy)\end{pmatrix}\ .\end{aligned}$$

L'intérêt d'écrire les résultats des dérivées partielles secondes sous forme matricielle n'a pas d'utilité directe autre que de reconnaître une certaine structure dans ses résultats. Et on peut aller encore plus loin quand on remarque qu'on a interprété g comme la composée de f avec $F(x,y) = \big(x^2+y,\sin(xy)\big)$, ce qui permet d'écrire

$$(\partial_1 F)(x,y) = \begin{pmatrix}2x\\ y\cos(xy)\end{pmatrix}\ ,\quad (\partial_2 F)(x,y) = \begin{pmatrix}1\\ x\cos(xy)\end{pmatrix}$$

$$(\partial_1\partial_1 F)(x,y) = \begin{pmatrix}2\\ -y^2\sin(xy)\end{pmatrix}\ ,\qquad (\partial_2\partial_1 F)(x,y) = \begin{pmatrix}0\\ \cos(xy)-xy\sin(xy)\end{pmatrix}\ ,$$

$$(\partial_2\partial_2 F)(x,y) = \begin{pmatrix}0\\ -x^2\sin(xy)\end{pmatrix}\ .$$

On reconnaît dans ces "vecteurs" les vecteurs qui apparaissent dans les expressions matricielles pour $\partial_i\partial_j g$. On peut aller encore plus loin dans la description matricielle en écrivant la matrice Hessienne de g et termes des matrices Hessiennes de f et de F :

$$\begin{aligned}&\begin{pmatrix}(\partial_1\partial_1 g)(x,y) & (\partial_1\partial_2 g)(x,y)\\ (\partial_2\partial_1 g)(x,y) & (\partial_2\partial_2 g)(x,y)\end{pmatrix}\\ &= \begin{pmatrix}2x & y\cos(xy)\\ 1 & x\cos(xy)\end{pmatrix}\cdot\begin{pmatrix}(\partial_1\partial_1 f)(Z) & (\partial_1\partial_2 f)(Z)\\ (\partial_2\partial_1 f)(Z) & (\partial_2\partial_2 f)(Z)\end{pmatrix}\cdot\begin{pmatrix}2x & 1\\ y\cos(xy) & x\cos(xy)\end{pmatrix}\\ &\quad+(\partial_1 f)(Z)\cdot\begin{pmatrix}2x & 1\\ 1 & 0\end{pmatrix}\\ &\quad+(\partial_2 f)(Z)\cdot\begin{pmatrix}-y^2\sin(xy) & \cos(xy)-xy\sin(xy)\\ \cos(xy)-xy\sin(xy) & -x^2\sin(xy)\end{pmatrix}\ .\end{aligned}$$

Quand on note la matrice Hessienne d'une fonction h à valeurs réelles comme D^2h, alors ce résultat prend la forme

$$\begin{aligned}(D^2g)(x,y) = {}^{\mathrm{tr}}(DF)(x,y)\cdot(D^2f)\big(F(x,y)\big)\cdot(DF)(x,y)\\ +(\partial_1 f)\big(F(x,y)\big)\cdot(D^2F_1)(x,y)+(\partial_2 f)\big(F(x,y)\big)\cdot(D^2F_2)(x,y)\ .\end{aligned}$$

Cette formule est à comparer avec la dérivée seconde d'une fonction composée d'une seule variable réelle $g(x) = f\big(F(x)\big)$:

$$\begin{aligned}g'(x) &= f'\big(F(x)\big)\cdot F'(x)\\ g''(x) &= f''\big(F(x)\big)\cdot F'(x)\cdot F'(x) + f'\big(F(x)\big)\cdot F''(x)\\ &= F'(x)\cdot f''\big(F(x)\big)\cdot F'(x) + f'\big(F(x)\big)\cdot F''(x)\ .\end{aligned}$$

On voit clairement la transformation du terme $F'(x)\cdot f''\big(F(x)\big)\cdot F'(x)$ sous sa forme matricielle avec la matrice de la différentielle DF qui remplace F' et la matrice Hessienne D^2f qui remplace f''. La transformation du terme $f'\big(F(x)\big)\cdot F''(x)$ est peut-être un peu plus difficile à comprendre, car on ne voit pas clairement apparaître la différentielle Df qui "devrait" remplacer f'. Ceci est dû au fait qu'on note cette formule sur une feuille de papier en 2 dimensions ! Si on avait la possibilité d'écrire en 3 dimensions, on aurait pu faire apparaître d'une façon naturelle aussi la différentielle Df qui remplace f'.

L'exercice §7 série supplémentaires 0.2.iii. Contrairement à ce qu'on pourrait croire, cette question est plus facile que la précédente. Commençons donc avec les dérivées partielles d'ordre 1 de $h = g\circ f$. On écrit d'abord avec toutes les composantes

$$h(x) = g\big(f_1(x_1,\dots,x_n), f_2(x_1,\dots,x_n),\dots,f_p(x_1,\dots,x_n)\big)\ ,$$

et ensuite on écrit sa différentielle (avec x au lieu de $(x_1,\dots,x_n)$)

$$\begin{aligned}(Dh)(x) &= \big(\ (\partial_1 h)(x)\ ;\ \dots\ ;\ (\partial_n h)(x)\ \big)\\ &= \Big(\ (\partial_1 g)\big(f(x)\big)\ ;\ \dots\ ;\ (\partial_p g)\big(f(x)\big)\ \Big)\cdot\begin{pmatrix}(\partial_1 f_1)(x) & \dots & (\partial_n f)(x)\\ \vdots & & \vdots\\ (\partial_1 f_p)(x) & \dots & (\partial_n f_p)(x)\end{pmatrix}\ .\end{aligned}$$

On en déduit directement la formule

$$(\partial_i h)(x) = \sum_{j=1}^{p}(\partial_j g)\big(f(x)\big)\cdot(\partial_i f_j)(x)$$

ou encore

$$\big(\partial_i(g\circ f)\big)(x) = \sum_{j=1}^{p}\big((\partial_j g)\circ f\big)(x)\cdot(\partial_i f_j)(x)$$

ou encore

$$\partial_i(g\circ f) = \sum_{j=1}^{p}\big((\partial_j g)\circ f\big)\cdot(\partial_i f_j)\ . \tag{0.2}$$

On devrait reconnaître l'approche où on regarde chaque occurrence de la variable x_i, qu'on dérive "ordinairement" et on remet le reste. C'est l'équivalent direct de la formule (0.1).

Avec cette formule, le calcul des dérivées partielles secondes devient “facile,” car on ré-applique la même formule ! Cela nous donne donc (avec aussi la formule de Leibniz) le résultat demandé (en rouge)

$$\begin{aligned}
\partial_k\big(\partial_i(g\circ f)\big) &= \sum_{j=1}^{p} \partial_k\Big(\big((\partial_j g)\circ f\big)\cdot(\partial_i f_j)\Big) \\
\text{Leibniz} \quad &= \sum_{j=1}^{p} \partial_k\Big(\big((\partial_j g)\circ f\big)\Big)\cdot(\partial_i f_j) + \sum_{j=1}^{p}\big((\partial_j g)\circ f\big)\cdot\big(\partial_k(\partial_i f_j)\big) \\
(0.2) \quad &= \sum_{j=1}^{p}\Bigg(\sum_{\ell=1}^{p}\Big(\big(\partial_\ell(\partial_j g)\big)\circ f\Big)\cdot(\partial_k f_\ell)\Bigg)\cdot(\partial_i f_j) + \sum_{j=1}^{p}\big((\partial_j g)\circ f\big)\cdot\big(\partial_k(\partial_i f_j)\big) \;.
\end{aligned}$$

Quand on réécrit cela sous forme matricielle, on trouve :

$$\begin{pmatrix} (\partial_1\partial_1 h)(x) & \dots & (\partial_1\partial_n h)(x) \\ \vdots & & \vdots \\ (\partial_n\partial_1 h)(x) & \dots & (\partial_n\partial_n h)(x) \end{pmatrix} =$$

$$= \begin{pmatrix} \partial_1 f_1 & \dots & \partial_1 f_p \\ \vdots & & \vdots \\ \partial_n f_1 & \dots & \partial_n f_p \end{pmatrix} \cdot \begin{pmatrix} (\partial_1\partial_1 g)\circ f & \dots & (\partial_1\partial_p g)\circ f \\ \vdots & & \vdots \\ (\partial_p\partial_1 g)\circ f & \dots & (\partial_p\partial_p g)\circ f \end{pmatrix} \cdot \begin{pmatrix} \partial_1 f_1 & \dots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_p & \dots & \partial_n f_p \end{pmatrix}$$

$$+ \sum_{j=1}^{p}\big((\partial_j g)\circ f\big)\cdot \begin{pmatrix} (\partial_1\partial_1 f_j)(x) & \dots & (\partial_1\partial_n f_j)(x) \\ \vdots & & \vdots \\ (\partial_n\partial_1 f_j)(x) & \dots & (\partial_n\partial_n f_j)(x) \end{pmatrix} \;.$$

Ou encore, en notant la matrice Hessienne d’une fonction réelle F par D^2F, et en mettant l’argument x :

$$(D^2h)(x) = {}^{\mathrm{tr}}\big((Df)(x)\big)\cdot(D^2g)\big(f(x)\big)\cdot\big((Df)(x)\big) + \sum_{j=1}^{p}(\partial_j g)\big(f(x)\big)\cdot(D^2 f_j)(x) \;.$$

Cela est exactement la forme qu’on avait trouvé dans le cas particulier de la fonction de la question 0.2.ii. À noter que les matrices Hessiennes sont des matrices carrées (celles de h et de chaque f_j de taille $n\times n$ et celle de g de taille $p\times p$), mais que la matrice de la différentielle Df est de taille $p\times n$ qui sera (en général) rectangulaire non-carrée.