2) Soit $f : \mathbb{R}^n \to \mathbb{R}$
$x \mapsto x^T A x + b^T x + c$ avec $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$

$$\begin{aligned}
f(x+h) &= (x+h)^T A (x+h) + b^T(x+h) + c \\
&= x^T A x + h^T A x + x^T A h + h^T A h + b^T x + b^T h + c \\
&= x^T A x + b^T x + c + x^T A h + h^T A x + b^T h + h^T A h \qquad \text{et } h^T A x = (h^T A x)^T = x^T A^T h \\
&= f(x) + x^T (A + A^T) h + b^T h + \underbrace{h^T A h}_{\text{c'est en } o_0(h)} \\
&= f(x) + (x^T (A + A^T) + b^T) h + o_0(h)
\end{aligned}$$

→ la différentielle est donc $\underline{df_x : h \mapsto (x^T(A+A^T) + b^T) h}$

→ le gradient $\nabla f(x)$ de f en x est tel que $df_x(h) = \nabla f(x)^T h$

Donc $\nabla f(x)^T h = (x^T(A+A^T) + b^T) h$

→ $\nabla f(x)^T = x^T(A+A^T) + b^T$

→ $\underline{\nabla f(x) = (A+A^T) x + b}$

Si de plus A est symétrique ($A = A^T$), alors $df_x(h) = (2x^T A + b^T) h$ et $\nabla f(x) = 2Ax + b$

3) Soit $f : \mathbb{R}^n \to \mathbb{R}$
$x \mapsto \|Ax - b\|_2^2$ avec $A \in \mathbb{R}^{m\times n}$ et $b \in \mathbb{R}^m$

On peut réécrire $f(x) = (Ax-b)^T (Ax-b)$

Donc
$$\begin{aligned}
f(x+h) &= (A(x+h) - b)^T (A(x+h) - b) \\
&= (Ax + Ah - b)^T (Ax + Ah - b) \\
&= ((Ax-b) + Ah)^T ((Ax-b) + Ah) \\
&= (Ax-b)^T(Ax-b) + (Ah)^T(Ax-b) + (Ax-b)^T(Ah) + (Ah)^T Ah \\
&= f(x) + 2(Ax-b)^T Ah + \underbrace{h^T A^T A h}_{o_0(h)} \qquad \text{puisque } (Ah)^T(Ax-b) = \left((Ah)^T(Ax-b)\right)^T = (Ax-b)^T Ah
\end{aligned}$$

→ la différentielle est donc $\underline{df_x : h \mapsto 2(Ax-b)^T Ah}$

→ le gradient $\nabla f(x)$ de f en x est tel que $df_x(h) = \nabla f(x)^T h$

Donc $\nabla f(x)^T = 2(Ax-b)^T A$

$\underline{\nabla f(x) = 2A^T(Ax-b)}$

4) Soit $f : \mathbb{R}^{n\times n} \to \mathbb{R}$
$X \mapsto (tr(X))^2$

$$\begin{aligned}
f(X+H) &= (tr(X+H))^2 \\
&= (tr(X) + tr(H))^2 \qquad \text{puisque } tr(A+B) = tr(A) + tr(B) \\
&= tr^2(X) + 2\,tr(X)\,tr(H) + tr^2(H)
\end{aligned}$$

$= f(X) + 2\,tr(X)tr(H) + o_0(H)$

On peut facilement vérifier que $H \mapsto 2tr(X)tr(H)$ est linéaire en H avec les propriétés de la trace

$\to$ la différentielle est donc $\underline{df_X : H \mapsto 2tr(X)tr(H)}$

$\to$ en revanche c'est nettement plus délicat d'exprimer le gradient de f à partir de la différentielle (on peut quand même s'en sortir en se référant à la section 2.5 du "matrix cookbook" pour la dérivée de la trace d'une matrice)

Exercice : différentielle d'un produit scalaire

Si $f, g : \mathbb{R}^n \to \mathbb{R}^m$ sont deux fonctions différentiables, alors leur produit scalaire $<f,g>$ est différentiable et
$d<f,g>_x = <df_x, g(x)> + <f(x), dg_x>$

$\to d<f,g>_x(h) = <df_x(h), g(x)> + <f(x), dg_x(h)>$

On reprend $f : \mathbb{R}^n \to \mathbb{R}$
$x \mapsto \|Ax-b\|_2^2 = (Ax-b)^T(Ax-b)$

$\to$ f peut s'écrire $f(x) = <Ax-b, Ax-b>$

On peut donc appliquer la formule de la différentielle d'un produit scalaire en prenant $f(x) = <g(x), g(x)>$ avec $g(x) = Ax-b$ et $dg_x : h \mapsto Ah$ (cf exercice "calcul d'une différentielle par le DL_1)

$$\to df_x = d<g,g>_x = <dg_x, g(x)> + <g(x), dg_x>$$
$$= 2<g(x), dg_x>$$

Autrement dit $df_x : h \mapsto 2<g(x), dg_x(h)> = 2<Ax-b, Ah> = \underline{2(Ax-b)^T Ah}$

On retrouve bien le résultat de l'exercice précédent

Exercice : différentielle d'une fonction composée

Si f différentiable en x_0 et g différentiable en $f(x_0)$ alors $g \circ f$ est différentiable en x_0 et $dg\circ f_x = dg_{f(x_0)} \circ df_{x_0}$

$\to dg\circ f_{x_0} : h \mapsto dg_{f(x_0)}(df_{x_0}(h))$

1) Soit $f : \mathbb{R}^n \to \mathbb{R}$
$x \mapsto \dfrac{1}{x^T x + 1}$

f peut s'écrire comme la composition de $a : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x^T x + 1$ et $b : \mathbb{R} \to \mathbb{R}$, $x \mapsto \dfrac{1}{x}$

$\to da_x : h \mapsto 2x^T h$ $\qquad \to db_x : h \mapsto h\,b'(x) = -\dfrac{h}{x^2}$

$\to f(x) = b\circ a(x)$

$\to df_x = db\circ a_x : h \mapsto db_{a(x)}(da_x(h)) = \underline{-\dfrac{2}{(x^T x + 1)^2}\, x^T h}$

2) Soit $f : \mathbb{R}^n \to \mathbb{R}$
$x \mapsto \cos^2(x^T A x)$ avec $A \in \mathbb{R}^{n\times n}$ symétrique

f peut s'écrire comme la composition de $a : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x^T A x$ et $b : \mathbb{R} \to \mathbb{R}$, $x \mapsto \cos^2 x$

$\to da_x : h \mapsto 2x^T A h$ $\quad \to db_x : h \mapsto h\,b'(x) = -2 \sin x \cos x\, h$
$= -\sin(2x) h$

$\to f(x) = b \circ a(x)$

$\to df_x = d\,b\circ a_x : h \mapsto db_{a(x)}(da_x(h)) = -\sin(2x^T A x)(2x^T A h) = \underline{-2\sin(2x^T A x) x^T A h}$