Exercices du chapitre XIII (Calcul différentiel) – corrigés

✓ : exercice d'application des méthodes, ★ : exercice classique, ☢ : exercice difficile.

★ **Exercice 27. (Factorisation par les zéros)** Soit B la boule unité ouverte de $\mathbb{R}^n$ pour la norme euclidienne usuelle. Soit $f : B \to \mathbb{R}^n$ une application de classe C^∞ telle que : $f(0) = 0$. Montrer qu'il existe des fonctions $g_i : B \to \mathbb{R}^n$ de classe C^∞ telles que : $\forall \vec{x} = (x_1, \dots, x_n) \in B,\ f(\vec{x}) = \displaystyle\sum_{i=1}^{n} x_i g_i(\vec{x})$.

Corrigé 27. Fixons provisoirement $\vec{x} = (x_1, \dots, x_n) \in B$, et soit γ l'application $t \mapsto t\vec{x}$. Il est clair que γ est de classe C^∞ sur $[0,1]$ et à valeurs dans B (on peut invoquer la convexité des boules pour cela). Par composition, $f \circ \gamma$ est de classe C^∞ sur $[0,1]$ et on a :

$$f(\vec{x}) = f(\vec{x}) - f(\vec{0}) = f \circ \gamma(1) - f \circ \gamma(0) = \int_0^1 (f \circ \gamma)'(t)\mathrm{d}t = \sum_{i=1}^{n} \int_0^1 x_i \partial_i f(t\vec{x})\mathrm{d}t.$$

Cela incite à poser : $\forall i \in [\![1, n]\!],\ \forall \vec{x} \in B,\ g_i(\vec{x}) = \displaystyle\int_0^1 \partial_i f(t\vec{x})\mathrm{d}t$. Comme voulu, on a : $\forall \vec{x} \in B,\ f(\vec{x}) = \displaystyle\sum_{i=1}^{n} x_i g_i(\vec{x})$. Or l'application g_i est de classe C^∞ pour tout $i \in [\![1, n]\!]$ grâce au théorème de dérivation des intégrales à paramètres : en effet, l'application $h_i : (t, \vec{x}) \mapsto \partial_i f(t\vec{x})$ est de classe C^∞ sur $[0,1] \times B$ par composition des applications $(t, \vec{x}) \mapsto t\vec{x}$ (de classe C^∞ car bilinéaire) et $\partial_i f$ (par hypothèse sur f), ce qui assure :

— la classe C^∞ par rapport à $\vec{x}$ et à $t \in [0,1]$ fixé de h_i, par composition avec $\vec{x} \mapsto (t, \vec{x})$;

— la continuité par morceaux par rapport à t et à $\vec{x} \in B$ fixé de h_i et de toutes ses dérivées partielles (à n'importe quel ordre), par composition avec $t \mapsto (t, \vec{x})$;

— que pour tout compact K inclus dans B, pour tout $k \in \mathbb{N}$ et pour tout $(j_1, \dots, j_k) \in [\![1, n]\!]^k$, l'application $\partial_{j_1,\dots,j_k} h_i$ est continue sur le compact $[0,1] \times K$ et par le théorème des bornes atteintes il existe donc $M \in \mathbb{R}_+$ tel que :

$$\forall (t, \vec{x}) \in [0,1] \times K, \quad |\partial_{j_1,\dots,j_k} h_i(t, \vec{x})| \leqslant M, \qquad \text{(HYPOTHÈSE DE DOMINATION)}$$

l'application $t \mapsto M$ étant évidemment intégrable car continue sur le segment $[0,1]$.

Ceci vaut pour tout $i \in [\![1, n]\!]$: d'où le résultat : l'application g_i est de classe C^∞ sur B pour tout $i \in [\![1, n]\!]$.

Exercice 59. (Caractérisation de $\mathrm{SO}_n(\mathbb{R})$) On note $\|\cdot\|$ la norme euclidienne usuelle de $\mathrm{M}_n(\mathbb{R})$.

1. Soit $A \in \mathrm{SL}_n(\mathbb{R})$. Déterminer l'espace tangent à $\mathrm{SL}_n(\mathbb{R}) = \{M \in \mathrm{M}_n(\mathbb{R}) \mid \det(M) = 1\}$ en A.
2. Montrer : $\mathrm{SO}_n(\mathbb{R}) = \left\{ M \in \mathrm{M}_n(\mathbb{R}) \mid \|M\| = \inf\limits_{A \in \mathrm{SL}_n(\mathbb{R})} \|A\| \right\}$.

Corrigé 59.

1. On a : $\mathrm{SL}_n(\mathbb{R}) = \det^{-1}(\{1\})$, or on a vu en exemple du cours que le gradient du déterminant en A est $\mathrm{Com}(A)$ (il faut savoir le redémontrer), donc : $T_A\mathrm{SL}_n(\mathbb{R}) = \mathrm{Com}(A)^\perp$, pour la structure euclidienne usuelle sur $\mathrm{M}_n(\mathbb{R})$.
2. Montrons d'abord que $\|\cdot\|$ ou, ce qui revient au même, $\|\cdot\|^2$ (considérer le carré simplifiera plus tard le calcul de gradient), atteint effectivement sa borne inférieure sur $\mathrm{SL}_n(\mathbb{R})$. La démarche est classique. Posons $d = \inf\limits_{A \in \mathrm{SL}_n(\mathbb{R})} \|A\|^2$. Comme $\|\cdot\|^2$ est continue sur le compact $\mathrm{SL}_n(\mathbb{R}) \cap \mathrm{B}_f\left(0_{\mathrm{M}_n(\mathbb{R})}, d+1\right)$ (c'est un fermé en tant qu'intersection de fermés, la fermeture de $\mathrm{SL}_n(\mathbb{R})$ provenant de sa description comme image réciproque par le déterminant d'un singleton ; il est borné puisqu'il est inclus dans une boule fermée, et en dimension finie cela assure le caractère compact), il existe $M \in \mathrm{SL}_n(\mathbb{R}) \cap \mathrm{B}_f\left(0_{\mathrm{M}_n(\mathbb{R})}, d+1\right)$ tel que :

$$\|M\|^2 = \inf_{A \in \mathrm{SL}_n(\mathbb{R}) \cap \mathrm{B}_f\left(0_{\mathrm{M}_n(\mathbb{R})}, d+1\right)} \|A\|^2 = d.$$

Comme, de plus, on a $\|A\|^2 > d + 1 > \|M\|^2$ pour tout $A \notin \mathrm{B}_f\left(0_{\mathrm{M}_n(\mathbb{R})}, d+1\right)$, on a même :

$$\|M\|^2 = \inf_{A \in \mathrm{SL}_n(\mathbb{R}) \cap \mathrm{B}_f\left(0_{\mathrm{M}_n(\mathbb{R})}, d+1\right)} \|A\|^2 = \inf_{A \in \mathrm{SL}_n(\mathbb{R})} \|A\|^2.$$

Montrons à présent que M est une matrice spéciale orthogonale. Comme $\|\cdot\|^2$ atteint un extremum local en M sur une ligne de niveau du déterminant, et qu'il s'agit d'un point régulier puisque :

$$\nabla \det(M) = \mathrm{Com}(M) = \det(M)\left(M^\top\right)^{-1} = \left(M^\top\right)^{-1} \neq 0_{\mathrm{M}_n(\mathbb{R})},$$

par le théorème des extremums liés il existe $\lambda \in \mathbb{R}$ tel que : $\nabla\|\cdot\|^2(M) = \lambda \nabla \det(M)$. Or on connaît le gradient de $\|\cdot\|^2$ (revoir le cours si besoin : c'est très important de savoir le retrouver !), donc :

$$2M = \lambda \left(M^\top\right)^{-1}.$$

Déduisons-en que M est spéciale orthogonale ; comme on connaît déjà son déterminant, il suffit de montrer qu'on a : $M^\top M = \mathrm{I}_n$. L'égalité précédente implique déjà : $M^\top M = \frac{\lambda}{2}\mathrm{I}_n$. Il suffit donc de montrer qu'on a $\lambda = 2$. Pour cela, on note que prendre le déterminant dans cette égalité donne : $1 = \left(\frac{\lambda}{2}\right)^n$, donc : $\lambda \in \{2, -2\}$. Or le cas $\lambda = -2$ est exclu : dans le cas contraire, on aurait $M^\top M = -\mathrm{I}_n$ et donc $M^\top M$ serait une matrice symétrique définie négative : on sait classiquement que c'est faux, puisque $X^\top M^\top MX = \|MX\|^2 \geqslant 0$ pour tout $X \in \mathrm{M}_{n,1}(\mathbb{R})$. Ainsi $\lambda = 2$, donc : $M^\top M = \mathrm{I}_n$, et comme $\det(M) = 1$ cela montre le résultat voulu : $M \in \mathrm{SO}_n(\mathbb{R})$.
Ainsi le minimum de $\|\cdot\|^2$ sur $\mathrm{SL}_n(\mathbb{R})$ est atteint en une matrice M spéciale orthogonale, et il vaut :

$$\|M\|^2 = \mathrm{tr}\left(M^\top M\right) = \mathrm{tr}(\mathrm{I}_n) = n.$$

Réciproquement, pour toute matrice spéciale orthogonale N, ce même calcul donne $\|N\|^2 = n$, donc le minimum de $\|\cdot\|^2$ est aussi atteint en N. Donc :

$$\left\{M \in \mathrm{M}_n(\mathbb{R}) \mid \|M\| = \inf_{A \in \mathrm{SL}_n(\mathbb{R})} \|A\|\right\} = \mathrm{SO}_n(\mathbb{R}).$$

Convexité

★ **Exercice 60. (Fonctions convexes)** Soit U un ouvert non vide convexe de E. On dit que $f : U \to \mathbb{R}$ est *convexe* si : $\forall(\vec{x}, \vec{y}) \in U^2$, $\forall t \in [0,1]$, $f(t\vec{x} + (1-t)\vec{y}) \leqslant tf(\vec{x}) + (1-t)f(\vec{y})$. On dit que f est *strictement convexe* si l'inégalité précédente est stricte lorsque $x \neq y$ et $0 < t < 1$.

1. Soit $\|\cdot\|$ une norme sur E. Montrer que $\|\cdot\|^2$ est convexe sur E.
2. On suppose que f est convexe. Soient $x \in U$, $\vec{h} \in E$ et $t \in [0,1]$ tels que $\vec{x} - \vec{h} \in U$ et $\vec{x} + \vec{h} \in U$. Montrer : $(1+t)f(\vec{x}) - tf(\vec{x} - \vec{h}) \leqslant f(\vec{x} + t\vec{h}) \leqslant (1-t)f(\vec{x}) + tf(\vec{x} + \vec{h})$.
3. Montrer que si f est convexe, alors f est continue.
4. Si f est de classe C^1, montrer que f est convexe si et seulement si : $\forall(\vec{x}, \vec{y}) \in U^2$, $f(\vec{y}) - f(\vec{x}) \geqslant \mathrm{d}f(\vec{x})(\vec{y} - \vec{x})$, si et seulement si : $\forall(\vec{x}, \vec{y}) \in U^2$, $\langle \nabla f(\vec{y}) - \nabla f(\vec{x}), \vec{y} - \vec{x}\rangle \geqslant 0$, où $\langle\cdot,\cdot\rangle$ est un produit scalaire sur E.

Corrigé 60.

1. Vu en travaux dirigés.
2. L'application $g : u \mapsto f(\vec{x} + u\vec{h})$ est définie sur $[-1,1]$ grâce aux hypothèses de l'énoncé (on a $\vec{x} - \vec{h}$ et $\vec{x} + \vec{h}$ dans U, donc les points de $[\vec{x} - \vec{h}, \vec{x}]$ et $[\vec{x}, \vec{x} + \vec{h}]$ le sont aussi car U est convexe), et g est même convexe sur cet intervalle en tant que composée d'une fonction convexe (à droite) par une fonction affine, donc elle est convexe. Si l'on désire plus de détails à ce sujet, voir la démonstration en remarque plus bas. Par le théorème des pentes croissantes, on a donc pour $t \in]0,1]$:

$$\frac{g(t) - g(0)}{t - 0} \leqslant \frac{g(1) - g(0)}{1 - 0}, \quad \text{et :} \quad \frac{g(0) - g(-1)}{0 - (-1)} \leqslant \frac{g(t) - g(0)}{t - 0},$$

c'est-à-dire :

$$f(\vec{x} - \vec{h}) - f(\vec{x}) \leqslant \frac{f(\vec{x} + t\vec{h}) - f(\vec{x})}{t} \leqslant f(\vec{x} + \vec{h}) - f(\vec{x}),$$

d'où le résultat après multiplication par t et réarrangement des termes. Si $t = 0$ alors l'encadrement est trivial vrai.
Remarque. Nous démontrons ci-dessous que la composée à droite par une application affine d'une fonction convexe reste convexe. En fait, il suffit de montrer que c'est vrai quand on compose à droite par une application linéaire puis par une translation. Nous le faisons en deux temps par souci de clarté, mais une démonstration directement dans le cas affine se fait semblablement.
Soient $f : U \to \mathbb{R}$ une fonction convexe et $\ell : E \to U$ une application linéaire. Pour tout $(\vec{x}, \vec{y}) \in E^2$ et tout $t \in [0,1]$, on a :

$$f \circ \ell(t\vec{x} + (1-t)\vec{y}) = f(t\ell(\vec{x}) + (1-t)\ell(\vec{y})) \leqslant tf(\ell(\vec{x})) + (1-t)f(\ell(\vec{y})) = tf \circ \ell(\vec{x}) + (1-t)f \circ \ell(\vec{y}),$$

donc $f \circ \ell$ est convexe.
Montrons à présent que la composition à droite par une translation $T_{\vec{a}} : \vec{x} \mapsto \vec{x} + \vec{a}$ est convexe (chose naturelle géométriquement : on ne fait que décaler le graphe, donc l'épigraphe reste convexe). Pour tout $(\vec{x}, \vec{y}) \in E^2$ et tout $t \in [0,1]$ tel que $\vec{a} + \vec{x} \in U$ et $\vec{a} + \vec{y} \in U$, on a $\vec{a} + t\vec{x} + (1-t)\vec{y} \in U$ puisque :

$$\vec{a} + t\vec{x} + (1-t)\vec{y} = t(\vec{a} + \vec{x}) + (1-t)(\vec{a} + \vec{y}) \in [\vec{a} + \vec{x}, \vec{a} + \vec{y}] \subseteq U$$

et de plus, par convexité de f :

$$f \circ T(t\vec{x} + (1-t)\vec{y}) = f(\vec{a} + t\vec{x} + (1-t)\vec{y}) = f(t(\vec{a} + \vec{x}) + (1-t)(\vec{a} + \vec{y})) \leqslant tf(\vec{a} + \vec{x}) + (1-t)f(\vec{a} + y) = tf \circ T(\vec{x}) + (1-t)f \circ T(\vec{y}),$$

d'où le résultat.

3. Soit $(\vec{e}_1,\dots,\vec{e}_n)$ une base de E. Munissons E de la norme $\|\cdot\| = \|\cdot\|_1$: le choix est indifférent puisque toutes les normes sont équivalentes sur E qui est de dimension finie. Le choix de cette norme sera clair lorsque que nous introduirons les coordonnées de $\vec{h}$ plus bas et appliquerons avec elles une inégalité de convexité.
Soit $\vec{x} \in U$. Montrons la continuité de f en $\vec{x}$. Comme U est un ouvert, il existe $r > 0$ tel que : $\mathrm{B}(\vec{x}, r) \subseteq U$. On peut supposer $r < 1$ quitte à remplacer r par $\min\left(r, \frac{1}{2}\right)$. Soit $\vec{y} \in \mathrm{B}(\vec{x}, r)$ différent de $\vec{x}$. On a $\vec{y} \in U$ (donc $f(\vec{y})$ existe) et de plus $\vec{y} - \vec{x}$ est de norme strictement inférieure à 1, ce qui permet d'écrire : $\vec{y} - \vec{x} = t\vec{h}$ avec $t \in]0,1[$ et $\vec{h}$ unitaire. Plus précisément :
$$t = \|\vec{y} - \vec{x}\|, \quad \vec{h} = \frac{1}{\|\vec{y} - \vec{x}\|}(\vec{y} - \vec{x}).$$
On veut montrer que $f(\vec{y})$ est proche de $f(\vec{x})$ lorsque $\vec{y}$ est proche de $\vec{x}$. Pour cela, la question précédente permet d'écrire :
$$f(\vec{y}) - f(\vec{x}) = f(\vec{x} + (\vec{y} - \vec{x})) - f(\vec{x}) = f(\vec{x} + t\vec{h}) - f(\vec{x}) \leqslant t\left(f(\vec{x} + \vec{h}) - f(\vec{x})\right).$$
Or, si l'on écrit $\vec{h} = \sum\limits_{i=1}^{n} h_i\vec{e}_i = \sum\limits_{i=1}^{n} |h_i|\varepsilon_i\vec{e}_i$ avec $\varepsilon_i \in \{1, -1\}$ pour tout $i \in [\![1, n]\!]$ (on réécrit ainsi $\vec{h}$ de sorte à avoir des coefficients dans $[0,1]$; comme $\vec{h}$ est unitaire pour la norme 1, on a : $\sum\limits_{i=1}^{n} |h_i| = 1$), alors on a par l'inégalité de Jensen (qui se démontre par récurrence sur n à l'aide de l'inégalité de convexité, comme dans le cas réel ; c'est l'associativité du barycentre qui permet de comprendre l'étape de l'hérédité), appliquée à $\vec{u} \mapsto f(\vec{x} + \vec{u})$ qui est convexe en tant que composée de f à droite par une translation :
$$f(\vec{x} + \vec{h}) = f\left(\vec{x} + \sum_{i=1}^{n} |h_i|\varepsilon_i\vec{e}_i\right) \leqslant \sum_{i=1}^{n} |h_i| f(\vec{x} + \varepsilon_i\vec{e}_i) \leqslant \max_{\substack{i \in [\![1,n]\!] \\ \varepsilon_i \in \{-1,1\}}} f(\vec{x} + \varepsilon_i\vec{e}_i) \sum_{i=1}^{n} |h_i| \leqslant \max_{\substack{i \in [\![1,n]\!] \\ \varepsilon_i \in \{-1,1\}}} f(\vec{x} + \varepsilon_i\vec{e}_i).$$
Posons : $M = \max\limits_{\substack{i \in [\![1,n]\!] \\ \varepsilon_i \in \{-1,1\}}} f(\vec{x} + \varepsilon_i\vec{e}_i) - f(\vec{x})$. D'après ce qui précède, on a :
$$f(\vec{y}) - f(\vec{x}) \leqslant tM = M\|\vec{y} - \vec{x}\|.$$
En utilisant la première inégalité de l'encadrement de la question précédente, on a plus généralement :
$$0 \leqslant |f(\vec{y}) - f(\vec{x})| \leqslant M\|\vec{y} - \vec{x}\| \underset{\vec{y} \to \vec{x}}{\longrightarrow} 0,$$
donc par le théorème des gendarmes : $\lim\limits_{\vec{y} \to \vec{x}} f(\vec{y}) = f(\vec{x})$: cela montre la continuité en $\vec{x}$. Ceci vaut pour tout $\vec{x} \in U$, donc f est continue sur U.
Remarque. En affinant le raisonnement (notamment pour éliminer la dépendance de M en $\vec{x}$), on peut démontrer un peu mieux : une fonction convexe est localement lipschitzienne.
Remarque. La démonstration de la continuité est bien plus technique que dans le cas d'une fonction d'une seule variable. Pourquoi donc ?
4. Vu en travaux dirigés.

★ **Exercice 61. (Convexité et classe C^2)** Soient U un ouvert non vide convexe de E et $f : U \to \mathbb{R}$ une application de classe C^2.

1. Montrer que f est convexe si et seulement si sa matrice hessienne est symétrique positive en tout point (voir l'exercice 60 pour une définition de la convexité).
2. Soit $A \in \mathrm{S}_n(\mathbb{R})$. Montrer que $X \mapsto X^\top AX$ est convexe sur $\mathrm{M}_{n,1}(\mathbb{R})$ si et seulement si A est symétrique positive.
3. Soient $A \in \mathrm{S}_n^{++}(\mathbb{R})$ et $B \in \mathrm{M}_{n,1}(\mathbb{R})$. Montrer que l'application $X \mapsto \frac{1}{2}X^\top AX + B^\top X$ atteint un minimum sur $\mathrm{M}_{n,1}(\mathbb{R})$ et qu'il est unique.

Corrigé 61.

1. Vu en travaux dirigés.
2. Vu en travaux dirigés.
3. Les deux méthodes proposées pour obtenir l'existence d'un minimum global sont très instructives. Celle n'utilisant pas la convexité s'adapte à de nombreuses configurations et est à avoir en tête impérativement pour des études qui ne sont pas sur un ouvert borné.
Première méthode (sans convexité). Montrons d'abord l'existence d'un minimum global de l'application $g : X \mapsto \frac{1}{2}X^\top AX + B^\top X$ sur $\mathrm{M}_{n,1}(\mathbb{R})$ (l'étude des points critiques et de la hessienne ne donnerait *a priori* que les minimums locaux : c'est pour cette raison que cette étape est cruciale) en nous ramenant à une étude sur un compact. L'idée est que g devient très grand au voisinage de l'infini, de sorte que la recherche d'un minimum

global peut se limiter à une étude sur un voisinage de 0, dont la compacité permettra d'utiliser le théorème des bornes atteintes.
Montrons donc que $g(X)$ tend vers l'infini quand la norme de X tend vers l'infini. Pour faciliter cette vérification, notons que A étant supposée symétrique définie positive, l'application $\mathrm{N} : X \mapsto \sqrt{X^\top AX}$ est une norme sur $\mathrm{M}_{n,1}(\mathbb{R})$ (euclidienne associée à un produit scalaire). Comme toutes les normes sont équivalentes sur $\mathrm{M}_{n,1}(\mathbb{R})$, nous pouvons volontiers privilégier cette norme dans notre étude.
Comme $X \mapsto B^\top X$ est une application linéaire sur $\mathrm{M}_{n,1}(\mathbb{R})$ qui est de dimension finie, elle est lipschitzienne : soit $K \geqslant 0$ tel que : $\forall X \in \mathrm{M}_{n,1}(\mathbb{R})$, $|B^\top X| \leqslant K\,\mathrm{N}(X)$. On a alors, pour tout $X \in \mathrm{M}_{n,1}(\mathbb{R})$:

$$g(X) \geqslant \frac{\mathrm{N}(X)^2}{2} - K\,\mathrm{N}(X).$$

Or on a aisément : $\lim\limits_{u\to+\infty}\left(\frac{u^2}{2} - Ku\right) = +\infty$. Par composition de limites et l'inégalité ci-dessus, on a donc : $\lim\limits_{\mathrm{N}(X)\to+\infty} g(X) = +\infty$. On en déduit qu'il existe $r > 0$ tel que pour tout $X \in \mathrm{M}_{n,1}(\mathbb{R})$ vérifiant $\mathrm{N}(X) \geqslant r$, on ait :

$$g(X) \geqslant g(0_{\mathrm{M}_{n,1}(\mathbb{R})}) + 1 = 1.$$

La fonction g ne peut donc pas admettre de minimum global sur $\mathrm{M}_{n,1}(\mathbb{R}) \setminus \mathrm{B}_f(0_{\mathrm{M}_{n,1}(\mathbb{R})}, r)$: elle est strictement supérieure en tout point à $g(0_{\mathrm{M}_{n,1}(\mathbb{R})})$. En revanche, la continuité de g (c'est la somme d'une norme au carré et d'une application linéaire sur un espace de dimension finie) sur le compact $\mathrm{B}_f(0_{\mathrm{M}_{n,1}(\mathbb{R})}, r)$ implique l'existence de $X_0 \in \mathrm{B}_f(0_{\mathrm{M}_{n,1}(\mathbb{R})}, r)$ tel que :

$$g(X_0) = \inf_{\mathrm{B}_f(0,r)} g = \inf_{\mathrm{M}_{n,1}(\mathbb{R})} g,$$

la dernière égalité découlant de l'observation ci-dessus. Comme il s'agit d'un minimum de g (global, donc local) sur $\mathrm{M}_{n,1}(\mathbb{R})$ qui est un ouvert, il s'agit d'un point critique (sous réserve que g soit différentiable, ce que nous justifions immédiatement), donc : $\Delta g(X_0) = 0_{\mathrm{M}_{n,1}(\mathbb{R})}$. Calculons le gradient de g pour en déduire ses points critiques et déterminer les valeurs possibles pour X_0. Soit $X \in \mathrm{M}_{n,1}(\mathbb{R})$. L'application N^2 est différentiable puisque N est une norme euclidienne (voir le cours ; on peut aussi reconnaître la composition de l'application linéaire $X \mapsto (X, AX)$ avec l'application bilinéaire $(X,Y) \mapsto X^\top Y$) et l'application $L : X \mapsto B^\top X$ est différentiable car linéaire. Par conséquent g est différentiable et on a :

$$\forall (X,H) \in (\mathrm{M}_{n,1}(\mathbb{R}))^2, \quad \mathrm{d}g(X)H = \frac{1}{2}\mathrm{d}\,\mathrm{N}^2(X)H + \mathrm{d}L(X)H = X^\top AH + L(H) = \left(X^\top A + B^\top\right) H.$$

On reconnaît le produit scalaire usuel de $AX + B$ (rappelons que A est symétrique) par H. Par définition du gradient :

$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}), \quad \nabla g(X) = AX + B.$$

On en déduit, comme A est inversible en tant que matrice définie positive :

$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}), \quad \left(\nabla g(X) = 0_{\mathrm{M}_{n,1}(\mathbb{R})} \iff X = -A^{-1}B\right).$$

Ainsi g n'admet qu'un seul point critique : c'est $-A^{-1}B$, et par ce qui précède X_0 doit lui être égal. Cela assure qu'il n'existe qu'un seul point où g atteint un minimum global.
Deuxième méthode (avec la convexité). On pourra juger ici de la simplicité des fonctions convexes dans la recherche des extremums. On commence par chercher les points critiques de g comme ci-dessus et on trouve qu'il n'existe qu'un seul point critique, à savoir $X_0 = -A^{-1}B$. Il reste à justifier que c'est un minimum global. Ce sera conséquence de la stricte convexité de g. En effet, comme : $\mathrm{d}g(X_0) = 0$, la dernière question de l'exercice 60 implique :

$$\forall X \in \mathrm{M}_{n,1}(\mathbb{R}), \quad g(X) - g(X_0) \geqslant \mathrm{d}g(X_0)(X - X_0) = 0,$$

d'où : $g(X) \geqslant g(X_0)$. En reprenant la résolution de cet exercice (et en utilisant la stricte convexité d'une fonction adaptée de la variable réelle), on montre que l'inégalité est stricte dès que $X \neq X_0$. Cela assure que X_0 est un minimum global, sous réserve que g est strictement convexe. Justifions-le.
Comme A est définie positive, par la question précédente l'application $f : X \mapsto X^\top AX$ est convexe. On a en fait un peu mieux : comme $\mathrm{N} : X \mapsto \sqrt{X^\top AX}$ est une norme euclidienne associée au produit scalaire $(X,Y) \mapsto X^\top AY$, on montre aisément que pour tout $(X,Y) \in (\mathrm{M}_{n,1}(\mathbb{R}))^2$ et pour tout $t \in [0,1]$, on a en posant $Z = tX + (1-t)Y$:

$$\begin{aligned}
f(Z) - tf(X) - (1-t)f(Y) &= \mathrm{N}(tX + (1-t)Y)^2 - t\,\mathrm{N}(X)^2 - (1-t)\,\mathrm{N}(Y)^2 \\
&= 2t(1-t)X^\top AY + t(t-1)\,\mathrm{N}(X)^2 + t(t-1)\,\mathrm{N}(Y)^2 \\
&= t(t-1)\left(\mathrm{N}(X)^2 + \mathrm{N}(Y)^2 - 2X^\top AY\right) \\
&= t(t-1)\,\mathrm{N}(X-Y)^2,
\end{aligned}$$

donc $f(Z) = tf(X) + (1-t)f(Y)$ si et seulement si $t \in \{0,1\}$ ou $X = Y$: cela signifie exactement que f est strictement convexe. C'est donc aussi le cas de g, puisqu'en conservant les mêmes notations que ci-dessus on a par linéarité de $X \mapsto B^\top X$:

$$g(Z) - tg(X) - (1-t)g(Y) = f(Z) - tf(X) - (1-t)f(Y),$$

donc cette quantité est négative parce que c'est le cas pour f, et nulle sous les mêmes conditions. Le raisonnement qui précède démontre donc bien que g admet un unique minimum en X_0 : d'où le résultat.
Conclusion. L'application g admet un unique minimum global. Il est en $-A^{-1}B$ et vaut :

$$g(-A^{-1}B) = \frac{1}{2}\left(B^\top A^{-1}\right)A(A^{-1}B) - B^\top A^{-1}B = -\frac{1}{2}B^\top A^{-1}B.$$