Corrigé de l'exercice 14 du TD n°9

3 Un problème de synthèse

Exercice 14.

a) On assimile $M_n(\mathbb{R})$ à $\mathbb{R}^{n^2}$. Montrer que $(A|B) = \mathrm{Tr}({}^tAB)$ définit un produit scalaire sur $M_n(\mathbb{R})$. Tout du long nous fixons la norme qui en dérive.

b) Montrer que l'application $\det : M_n(\mathbb{R}) \to \mathbb{R}$ est C^∞, et calculer sa différentielle en Id, puis en tout point (en A, interviendra la comatrice A).

c) Montrer que $\mathrm{GL}_n(\mathbb{R})$ est ouvert dans $M_n(\mathbb{R})$. $\mathrm{GL}_n(\mathbb{R})$ est-il compact ? connexe ? connexe par arcs ?

d) Soit $\mathrm{SL}_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) : \det M = 1\}$. Montrer que $\mathrm{SL}_n(\mathbb{R})$ est fermé dans $\mathrm{GL}_n(\mathbb{R})$. $\mathrm{SL}_n(\mathbb{R})$ est-il compact ? connexe ? connexe par arcs ?

e) Montrer que $\mathrm{SL}_n(\mathbb{R})$ est une hypersurface de $\mathbb{R}^n$ et déterminer le plan tangent en Id.

f) Montrer que $\|\cdot\|$ atteint sur $\mathrm{SL}_n(\mathbb{R})$ un minimum global. Atteint-elle un maximum global ?

g) Minimiser $\|\cdot\|$ sur $\mathrm{SL}_n(\mathbb{R})$ (on pensera bien sûr aux multiplicateurs de Lagrange).

Solution. Le déterminant ayant soulevé un tollé inattendu, nous n'avons pas eu le temps de finir ce problème instructif. En voici donc la correction écrite.

a) L'application proposée :

$$\begin{array}{ccc} (M_n(\mathbb{R}))^2 & \to & \mathbb{R} \\ (A, B) & \mapsto & \mathrm{Tr}({}^tAB) \end{array}$$

est une forme bilinéaire. Comme $(B|A) = \mathrm{Tr}({}^tBA) = \mathrm{Tr}({}^t({}^tBA)) = \mathrm{Tr}({}^tAB) = (A|B)$, elle est symétrique. Vérifions qu'elle est définie positive. Or en effet, notant $A = (a_{i,j})_{i,j}$, on a :

$$({}^tAA)_{i,j} = \sum_{k=1}^n a_{k,i}a_{k,j}$$

d'où :

$$(A|A) = \mathrm{Tr}({}^tAA) = \sum_{i=1}^n ({}^tAA)_{i,i} = \sum_{i=1}^n \sum_{k=1}^n a_{k,i}a_{k,i} = \sum_{i,j=1\ldots n} a_{i,j}^2$$

Il est dès lors évident que la forme bilinéaire symétrique proposée est définie positive : c'est un produit scalaire sur $M_n(\mathbb{R})$.

Remarque. La norme ainsi définie coïncide avec la norme 2 sur $\mathbb{R}^{n^2}$. Mais ce n'est pas la norme d'opérateurs associée à la norme 2 sur $\mathbb{R}^n$: $\|\cdot\|_{\mathbb{R}^{n^2},2} \neq |\!|\!|\cdot|\!|\!|_{L(\mathbb{R}^n,\|\cdot\|_2)}$

b) La question qui fâche. Reprenons.

L'application différentielle est polynomiale en les coefficients : elle est donc C^∞ sur $\mathbb{R}^{n^2}$.

Il y a un autre argument, peut-être moins simple : le déterminant est n-linéaire sur l'espace des colonnes $\mathbb{R}^n$, donc est continue en fonction des colonnes, *et* les colonnes sont continues en fonction des coefficients, donc le déterminant est continu en fonction des coefficients. Ce second argument peut fournir un calcul de la différentielle, comme suit.

Soient $H = (h_{i,j})_{i,j}$ et $M = \mathrm{Id} + H$. On cherche un développement limité de $\det M$. Si A est une matrice, on note $C_j(A)$ la $j^{\text{ème}}$ colonne de A. Comme le déterminant est n-linéaire en les colonnes,

$$\det M = \det(C_1(M), \ldots, C_n(M)) = \det(C_1(\mathrm{Id}) + C_1(H), \ldots, C_n(\mathrm{Id}) + C_n(H))$$

Préparons-nous à développer ; dans chaque $C_i(\mathrm{Id})+C_i(H)$, on choisit $C_i(\mathrm{Id})$ ou $C_i(H)$. Littéralement cela ferait 2^n termes, mais beaucoup vont disparaître.

– Si dans chaque $C_i(\mathrm{Id}) + C_i(H)$ on garde $C_i(\mathrm{Id})$, on récupère $\det(C_1(\mathrm{Id}), \ldots, C_n(\mathrm{Id})) = 1$.

– Si l'on garde exactement un $C_i(H)$, on récupère

$$\det(C_1(\mathrm{Id}),\dots,C_i(H),\dots,C_n(\mathrm{Id})) = \begin{bmatrix} 1 & 0\dots & h_{1,i} & \dots & 0 \\ 0 & 1 & h_{2,i} & & \vdots \\ & \ddots & & & \\ \vdots & & \vdots & 1 & 0 \\ 0 & \dots & h_{n,i} & 0 & 1 \end{bmatrix} = h_{i,i}$$

– Si l'on garde au moins deux $C_i(H)$, on récupère un $o(H)$.
Au bilan,

$$\det M = 1 + \sum_{i=1}^{n} h_{i,i} + o(H)$$

En particulier, $\det M = \det \mathrm{Id} + \mathrm{Tr}\, H + o(H)$, ce qui signifie que $D_{\mathrm{Id}} \det = \mathrm{Tr}$.

Remarque.

i) Les calculs de o sont menés pour la norme $\|H\|_\infty = \max_{i,j} |h_{i,j}|$, qui les rend transparents. Comme elle est équivalente à la norme de l'exercice, il n'y a pas de piège.

ii) Je persiste à trouver la définition combinatoire du déterminant bien plus adaptée. Rappelons que si S_n désigne le groupe symétrique (les bijections de n) et ε la signature (cette chose responsable de l'alternance des signes dans votre développement ; qu'importe) :

$$\det M = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{i,\sigma(i)}$$

Or si $\sigma \neq \mathrm{Id}$, σ admet au moins *deux* i tels que $\sigma(i) \neq i$. Donc le produit $\prod_{i=1}^{n} m_{i,\sigma(i)}$ passe au moins deux fois hors de la diagonale ; il englobe au moins deux termes en $h_{k,\ell}$; c'est un $o(H)$. Le seul terme qui ne soit pas à vue un $o(H)$ est ainsi celui où $\sigma = \mathrm{Id}$. On a donc :

$$\det M = \prod_{i=1}^{n} m_{i,i} + o(H) = \prod_{i=1}^{n}(1 + h_{i,i}) + o(H) = 1 + \sum_{i=1}^{n} h_{i,i} + o(H)$$

Il s'agit maintenant de calculer la différentielle en un point autre que l'identité. On pourrait avec de la combinatoire procéder comme en Id, mais analytiquement parlant il est plus instructif (et peut-être plus facile) de raisonner comme suit.
Soient $M \in \mathrm{GL}_n(\mathbb{R})$ et $H \in M_n(\mathbb{R})$. Alors :

$$\begin{aligned} \det(M + H) &= \det(M(\mathrm{Id} + M^{-1}H)) = \det M \cdot \det(\mathrm{Id} + M^{-1}H) \\ &= \det M \cdot (1 + \mathrm{Tr}(M^{-1}H) + o(H)) = \det M + \det M\, \mathrm{Tr}(M^{-1}H) + o(H) \end{aligned}$$

Nous avons donc :

$$D_M \det \cdot H = \det M\, \mathrm{Tr}(M^{-1}H)$$

Cela est bien, mais on voit mal comment étendre à $M \in M_n(\mathbb{R})$, qui ne serait pas forcément inversible. C'est ici qu'intervient la comatrice. Rappelons que la matrice des cofacteurs $\mathrm{com}\, M$ a pour coefficients :

$$(\mathrm{com}\, M)_{i,j} = (-1)^{i+j} \mu_{i,j}$$

où $\mu_{i,j}$ est le déterminant de la matrice obtenue de M en effaçant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne. Et l'on a la formule, qui est en fait le développement du déterminant :

$$M \cdot {}^t\mathrm{com}\, M = {}^t\mathrm{com}\, M \cdot M = \det M \cdot \mathrm{Id}$$

Notamment, si $M \in \mathrm{GL}_n(\mathbb{R})$, $M^{-1} = \frac{1}{\det M} {}^t\mathrm{com}\, M$, si bien que : $D_M \det \cdot H = \det M\, \mathrm{Tr}(M^{-1}H) = \mathrm{Tr}({}^t\mathrm{com}\, M H)$ et $D_M \det = (\mathrm{com}\, M|\cdot)$. Cette dernière expression se généralise à toute matrice M. En effet, comme det est C^∞, la fonction $D \det$ est continue. Comme $\mathrm{GL}_n(\mathbb{R})$ est dense dans $M_n(\mathbb{R})$, on a ainsi : pour toute $M \in M_n(\mathbb{R})$, $D_M \det = ({}^t\mathrm{com}\, M|\cdot)$.

Remarque. Le produit scalaire $(A|B)$ introduit une correspondance entre applications linéaires sur l'espace des matrices, et matrices. Le gradient est la matrice (ici, nos vecteurs sont des éléments de $M_n(\mathbb{R})$, i.e. des matrices) associée à la différentielle. Et comme $(\overrightarrow{\mathrm{grad}}_M \det |H) = D_M \det \cdot H = \mathrm{Tr}({}^t\mathrm{com}\, M H) = (\mathrm{com}\, M|H)$, on voit que $\overrightarrow{\mathrm{grad}}_M \det = \mathrm{com}\, M$.

c) Comme det : $M_n(\mathbb{R}) \to \mathbb{R}$ est continue, $\mathrm{GL}_n(\mathbb{R}) = \det^{-1}(\mathbb{R}^\times)$ est un ouvert de $M_n(\mathbb{R})$. $\mathrm{GL}_n(\mathbb{R})$ n'est pas compact : on peut voir que $(k\,\mathrm{Id})_{k\in\mathbb{N}}$ n'a pas de valeur d'adhérence, ou l'on peut raisonner de manière sophistiquée comme suit. Comme $M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$, $M_n(\mathbb{R})$ est convexe, connexe par arcs, et connexe. Comme $\varnothing \subsetneq \mathrm{GL}_n(\mathbb{R}) \subsetneq M_n(\mathbb{R})$ et que $\mathrm{GL}_n(\mathbb{R})$ est ouvert dans $M_n(\mathbb{R})$, il n'y est pas fermé. Et comme il n'y est pas fermé, il n'est pas compact.

   $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe : s'il l'était, $\det(\mathrm{GL}_n(\mathbb{R})) = \mathbb{R}^\times$ le serait, ce qui n'est bien sûr pas le cas. $\mathrm{GL}_n(\mathbb{R})$ n'est donc pas non plus connexe par arcs.

d) Comme det est continue, $\mathrm{SL}_n(\mathbb{R}) = \det^{-1}(\{1\})$ est fermé dans $M_n(\mathbb{R})$, et aussi dans $\mathrm{GL}_n(\mathbb{R})$. $\mathrm{SL}_n(\mathbb{R})$ n'est pas compact, car il n'est pas borné dans $M_n(\mathbb{R})$ (penser à la diagonale $(\lambda, \lambda^{-1}, 1, \ldots, 1)$).

   $\mathrm{SL}_n(\mathbb{R})$ est connexe, et même connexe par arcs ; cette question est plus technique. Des méthodes auxquelles je pense, aucune n'a vraiment sa place ici, et je préfère ne pas vous malmener.

e) $\mathrm{SL}_n(\mathbb{R})$ est une hypersurface de $M_n(\mathbb{R})$. Soit en effet $g(M) = \det M - 1$; g est C^∞, et $\mathrm{SL}_n(\mathbb{R}) = g^{-1}(\{0\})$. En outre, si $M \in \mathrm{SL}_n(\mathbb{R})$, $D_M \det = (\mathrm{com}\, M|\cdot)$, mais comme M est inversible, $\mathrm{com}\, M \neq 0$, et donc g n'a pas de point critique dans $\mathrm{SL}_n(\mathbb{R})$. On a bien une hypersurface.

   Remarque. Les points critiques du déterminant sont en fait les matrices de rang $\leqslant n-2$.

   Le plan tangent à $\mathrm{SL}_n(\mathbb{R})$ en Id est $T_{\mathrm{Id}}\,\mathrm{SL}_n(\mathbb{R}) = \ker D_{\mathrm{Id}} g = \ker D_{\mathrm{Id}} \det = \ker Tr = \{H \in M_n(\mathbb{R}) : \mathrm{Tr}\, H = 0\}$.

f) Montrons que $\|\cdot\|$ admet sur $\mathrm{SL}_n(\mathbb{R})$ un minimum global. On peut en effet se restreindre à $\{M \in \mathrm{SL}_n(\mathbb{R}) : \|M\| \leqslant \|\mathrm{Id}\|\}$. Mais cet ensemble est fermé comme $\mathrm{SL}_n(\mathbb{R})$, et borné : il est donc compact. La fonction $\|\cdot\|$ y admet donc un minimum, qui est un minimum global sur tout $\mathrm{SL}_n(\mathbb{R})$.

   C'est bien sûr faux avec « maximum » : $\mathrm{SL}_n(\mathbb{R})$ n'est pas borné !

g) Au lieu de minimiser la fonction $\|\cdot\|$, nous minimisons bien sûr $\|\cdot\|^2$, qui est C^∞. Rappelons que la différentielle de $\|\cdot\|^2$ en M est :

$$D_M\|\cdot\|^2 \cdot H = (2M|H)$$

   Soit $A \in \mathrm{SL}_n(\mathbb{R})$ un minimum global relatif. C'est aussi un minimum local relatif, donc d'après la théorie de Lagrange, il existe $\lambda \in \mathbb{R}$ tel que :

$$\forall H \in M_n(\mathbb{R}), (2A|H) = \lambda(\mathrm{com}\, A|H)$$

   Il suit que $2A = \lambda\, \mathrm{com}\, A$. Multiplions par ${}^tA$: il vient $2{}^tAA = \lambda \det A\, \mathrm{Id} = \lambda\, \mathrm{Id}$ car $A \in \mathrm{SL}_n(\mathbb{R})$. Passant au déterminant, on trouve $2^n = \lambda^n$, donc $\lambda = \pm 2$. Ainsi $\|A\|^2 = \mathrm{Tr}({}^tAA) = \pm 1\, \mathrm{Tr}\, \mathrm{Id} = \pm n$, mais ce nombre est positif (c'est une norme au carré).
   Conclusion : le minimum global de $\|\cdot\|$ sur $\mathrm{SL}_n(\mathbb{R})$ est $\sqrt{n}$.

   Remarque. Les points réalisant ce minimum sont les points du groupe spécial orthogonal $\mathrm{SO}_n(\mathbb{R})$.