

1 Différentielle de l'inverse

Montrer que $\mathcal{GL}_n(\mathbb{R})$ est un ouvert, puis que l'application $M \in \mathcal{GL}_n(\mathbb{R}) \mapsto M^{-1}$ est différentiable et déterminer sa différentielle en tout point.

On pourra commencer par la différentielle en I_n .

Solution de 1 : Différentielle de l'inverse

Soit $M \in \mathcal{GL}_n(\mathbb{K})$. Munissons $\mathcal{M}_n(\mathbb{K})$ d'une norme d'algèbre $\|\cdot\|$.

Pour toute matrice H , on peut factoriser $M + H$ par M

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

Si H est telle que $\|H\|$ soit suffisamment petite, on a $\| -M^{-1}H \| \leq \|M^{-1}\| \|H\| < 1$.

D'après le cours, comme $\| -M^{-1}H \| < 1$, la matrice $I_n - (-M^{-1}H) = I_n + M^{-1}H$ est inversible et son inverse est donné par la somme de la série absolument convergente

$$(I_n + M^{-1}H)^{-1} = \sum_{k=0}^{+\infty} (-M^{-1}H)^k = I_n - M^{-1}H + \sum_{k=2}^{+\infty} (-M^{-1}H)^k.$$

Majorons le reste de cette série pour montrer qu'il s'agit d'un $o(\|H\|)$. Par inégalité triangulaire et sous-multiplicativité,

$$\left\| \sum_{k=2}^{+\infty} (-M^{-1}H)^k \right\| \leq \sum_{k=2}^{+\infty} \|M^{-1}H\|^k = \|M^{-1}H\|^2 \sum_{j=0}^{+\infty} \|M^{-1}H\|^j = \frac{\|M^{-1}H\|^2}{1 - \|M^{-1}H\|}$$

Puisque $\|M^{-1}H\| \leq \|M^{-1}\| \|H\|$, ce reste est majoré par une quantité de l'ordre de $\|H\|^2$. C'est donc bien un $o_{H \rightarrow 0_n}(\|H\|)$. On obtient ainsi le développement limité

$$(I_n + M^{-1}H)^{-1} = I_n - M^{-1}H + o_{H \rightarrow 0_n}(H)$$

En reportant ce résultat dans l'expression de l'inverse de $M + H$

$$(M + H)^{-1} = (I_n + M^{-1}H)^{-1} M^{-1} = \left(I_n - M^{-1}H + o_{H \rightarrow 0_n}(H) \right) M^{-1} = M^{-1} - M^{-1}HM^{-1} + o_{H \rightarrow 0_n}(H)$$

L'application $H \mapsto -M^{-1}HM^{-1}$ étant linéaire, on conclut que l'application f est différentiable en M et que sa différentielle s'écrit $df(M): H \mapsto -M^{-1}HM^{-1}$

2 Optimisation et convexité

Soit \mathcal{U} un ouvert convexe de \mathbb{R}^2 et $f: \mathcal{U} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^1 , convexe (la définition est la même que pour les fonctions d'une variable).

1. Montrer que si $a, b \in \mathcal{U}$, $df(a)(b-a) \leq f(b) - f(a)$.
2. Montrer que tout point critique est un minimum global.
3. Montrer que les points critiques forment un ensemble convexe fermé.

Solution de 2 : Optimisation et convexité

3 Différentielle du déterminant

- Justifier que $\det : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ est de classe \mathcal{C}^∞ .
- Une première méthode**
Calculer la différentielle et le gradient en toute matrice $A \in \mathcal{M}_n(\mathbb{R})$ de l'application \det en passant par un calcul de dérivées partielles.
- Deux autres méthodes**
 - En utilisant les valeurs propres complexes, calculer, pour $H \neq O_n$, $D_H \det(I_n)$ et en déduire $d \det(I_n)$.
 - Retrouver ce résultat en montrant que $\det \circ \exp = \exp \circ \text{tr}$ et en utilisant la différentielle en O_n de \exp .
 - En déduire $d \det(A)$ pour toute matrice $A \in \mathcal{GL}_n(\mathbb{R})$, puis pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$.

Solution de 3 : Différentielle du déterminant

- En tant qu'application polynomiale en les coefficients de la matrice, l'application $\det : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathbb{R}$ est de classe \mathcal{C}^∞ .
- Une première méthode**
L'application \det est affine quand on la considère comme fonction d'un seul coefficient de la matrice (une application partielle, donc). Autrement dit, il va s'agir de dériver une fonction $t \mapsto at + b$.

Si (i, j) est un couple fixé,

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \det A = a_{i,j} [\text{Com} A]_{i,j} + \phi_{i,j}(A)$$

où ni $[\text{Com} A]_{i,j}$, ni $\phi_{i,j}(A)$ ne dépend de $a_{i,j}$. Ceci est conséquence par exemple de la formule sur le développement

du déterminant par rapport à la ligne L_i , ou par rapport à la colonne C_j . Et on a donc $\frac{\partial \det}{\partial a_{i,j}}(A) = [\text{Com} A]_{i,j}$.

Comme \det est de classe \mathcal{C}^1 donc différentiable, on peut écrire

$$\forall H \in \mathcal{M}_n(\mathbb{R}), d \det(A)(H) = \sum_{1 \leq i, j \leq n} \frac{\partial \det}{\partial a_{i,j}}(A) H_{i,j} = \sum_{1 \leq i, j \leq n} [\text{Com} A]_{i,j} H_{i,j} = \sum_{j=1}^n \left(\sum_{i=1}^n [\text{Com} A]_{i,j} H_{i,j} \right)$$

Et on arrive ainsi à une formule célèbre

$$\forall H \in \mathcal{M}_n(\mathbb{R}), d \det(A)(H) = \text{tr}((\text{Com} A)^T H) = (\text{Com} A | H)$$

pour le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$.

Donc $\nabla \det(A) = \text{Com} A$.

- Deux autres méthodes**
 - Soit $t \in \mathbb{R}$. Si on note $\lambda_1, \dots, \lambda_n$ les valeurs propres complexes de $H \in \mathcal{M}_n(\mathbb{R})$ comptées avec multiplicité, celles de $I_n + tH$ sont $1 + t\lambda_1, \dots, 1 + t\lambda_n$.
Comme les matrices sont trigonalisables dans \mathbb{C} ,

$$\det(I_n + tH) = \prod_{k=1}^n (1 + t\lambda_k) = \underbrace{1}_{=\det I_n} + t \sum_{k=1}^n \lambda_k + o_{t \rightarrow 0}(t).$$

On a donc $D_H \det(I_n) = d \det(I_n)(H) = \text{tr}(H)$.

Remarque : et donc, en particulier, l'ensemble $T_n(\mathcal{SL}_n(\mathbb{R}))$ des vecteurs tangents à $\mathcal{SL}_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}), \det M = 1\}$ est $\text{Ker tr} = \{H \in \mathcal{M}_n(\mathbb{R}), \text{tr} H = 0\}$.

- La formule $\det \circ \exp = \exp \circ \text{tr}$ s'obtient en trigonalisant dans $\mathcal{M}_n(\mathbb{C})$.

Comme toutes les fonctions sont différentiables et en sachant que $d \exp(O_n) = \text{id}_{\mathcal{M}_n(\mathbb{R})}$ (c'est classique et pas difficile en isolant les deux premiers termes de la série exponentielle et en utilisant une norme d'algèbre), on peut écrire

$$d(\det \circ \exp)(O_n) = d \det(\exp(O_n)) \circ d \exp(O_n) = d \det(I_n)$$

et

$$d(\exp \circ \text{tr})(O_n) = \exp'(\text{tr}(O_n)) \times d \text{tr}(O_n) = 1 \times \text{tr}$$

car la trace étant linéaire, sa différentielle en tout point est elle-même.

On conclut avec la formule $d \det(I_n) = \text{tr}$

(c) Ensuite, pour A est inversible, on peut écrire

$$\det(A + H) = \det(A) \det(I_n + A^{-1}H) = \det(A) \left(1 + \operatorname{tr}(A^{-1}H) + \underset{H \rightarrow 0_n}{o}(\|H\|) \right)$$

En effet, un $\underset{H \rightarrow 0_n}{o}(\|A^{-1}H\|)$ est un $\underset{H \rightarrow 0_n}{o}(\|H\|)$, comme on le voit avec une norme d'algèbre pour laquelle

$$\|A^{-1}H\| \leq \|A^{-1}\| \times \|H\|$$

On en déduit par la formule de la comatrice que la différentielle en A est

$$d \det(A) : H \mapsto \det(A) \operatorname{tr}(A^{-1}H) = \operatorname{tr}((\operatorname{Com} A)^T H) = (\operatorname{Com} A | H)$$

Les applications $d \det$ et $A \mapsto (\operatorname{Com} A | \cdot)$ sont continues car \det est de classe \mathcal{C}^1 pour la première, et pour la deuxième, cela vient de la continuité de l'application comatrice (polynomiale) et de la bilinéarité du produit scalaire sur un espace de dimension finie. Elles coïncident sur $\mathcal{GL}_n(\mathbb{R})$ classiquement dense dans $\mathcal{M}_n(\mathbb{R})$, donc sont égales.

Finalement, pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, $d \det(A) = (\operatorname{Com} A | \cdot)$, c'est-à-dire $\nabla \det(A) = \operatorname{Com} A$.

4 Mines - X

Soit $f : M \in \mathcal{M}_n(\mathbb{R}) \mapsto (\operatorname{tr} M, \operatorname{tr}(M^2), \dots, \operatorname{tr}(M^n)) \in \mathbb{R}^n$.

1. Montrer que f est différentiable en tout $M \in \mathcal{M}_n(\mathbb{R})$ et déterminer $df(M)$.
2. Relier le rang de $df(M)$ et le degré du polynôme minimal de M .
3. Montrer que l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ dont le polynôme minimal est égal au polynôme caractéristique est un ouvert de $\mathcal{M}_n(\mathbb{R})$.

Solution de 4 : Mines - X

1. **Méthode directe** Remarquons d'abord que, par distributivité, si $k \geq 2$,

$$(M + H)^k = \underbrace{(M \text{ ou } H) \times (M \text{ ou } H) \times \dots \times (M \text{ ou } H)}_{k \text{ fois}} = \sum_{(\varepsilon_1, \dots, \varepsilon_k) \in \{0,1\}^k} M^{\varepsilon_1} H^{1-\varepsilon_1} M^{\varepsilon_2} H^{1-\varepsilon_2} \dots M^{\varepsilon_k} H^{1-\varepsilon_k}$$

Or, s'il y a au moins deux ε_i égaux à 0,

$$M^{\varepsilon_1} H^{1-\varepsilon_1} M^{\varepsilon_2} H^{1-\varepsilon_2} \dots M^{\varepsilon_k} H^{1-\varepsilon_k} = \underset{H \rightarrow 0_n}{o}(H)$$

(se précise facilement en utilisant une norme sous-multiplicative).

Et comme la trace est continue (linéaire sur $\mathcal{M}_n(\mathbb{R})$ qui est de dimension finie), il existe k tel que

$$\forall A \in \mathcal{M}_n(\mathbb{R}) \quad |\operatorname{tr}(A)| \leq k \|A\|$$

ce qui montre que, si au moins deux ε_i sont égaux à 0,

$$\operatorname{tr}(M^{\varepsilon_1} H^{1-\varepsilon_1} M^{\varepsilon_2} H^{1-\varepsilon_2} \dots M^{\varepsilon_k} H^{1-\varepsilon_k}) = \underset{H \rightarrow 0_n}{o}(H)$$

Donc

$$\begin{aligned} \operatorname{tr}((M + H)^k) &= \operatorname{tr}(M^k) + \sum_{i=1}^k \operatorname{tr}(M^{i-1} H M^{k-i}) + \underset{H \rightarrow 0_n}{o}(H) \\ &= \operatorname{tr}(M^k) + k \operatorname{tr}(M^{k-1} H) + \underset{H \rightarrow 0_n}{o}(H) \end{aligned}$$

ce qui permet de conclure :

$$df(M) : H \mapsto (\operatorname{tr}(H), 2 \operatorname{tr}(MH), \dots, n \operatorname{tr}(M^{n-1}H))$$

Par opérations Soit $g_k : M \mapsto M^k$. On a $f_k = \operatorname{tr} \circ g_k : M \mapsto \operatorname{tr}(M^k)$.

Comme tr est linéaire et $(M_1, \dots, M_k) \mapsto M_1 \times \dots \times M_k$ est k -linéaire, f_k est différentiable sur $\mathcal{M}_n(\mathbb{K})$ et $\forall M \in \mathcal{M}_n(\mathbb{K})$,

$$df_k(M) : H \mapsto \operatorname{tr}(dg_k(M)(H)) = \operatorname{tr}\left(\sum_{i=1}^k M^{i-1} d(\operatorname{id}_{\mathcal{M}_n(\mathbb{R})})(M)(H) M^{k-i}\right) = \operatorname{tr}\left(\sum_{i=1}^k M^{i-1} H M^{k-i}\right) = k \operatorname{tr}(M^{k-1} H)$$

ce qui redonne l'expression $df(M) : H \mapsto (\operatorname{tr}(H), 2 \operatorname{tr}(MH), \dots, n \operatorname{tr}(M^{n-1}H))$.

2. Pour calculer le rang de $df(M)$, on cherche par exemple son noyau. On remarque successivement :

$$H \in \text{Ker}(df(M)) \iff \forall k \in \llbracket 0, n-1 \rrbracket, \quad \text{tr}(M^k H) = 0$$

puis, par combinaison linéaire :

$$H \in \text{Ker}(df(M)) \iff \forall P \in \mathbb{R}_{n-1}[X], \quad \text{tr}(P(M)H) = 0$$

Or on sait d'après le cours que si $d = \deg \pi_M \leq n$, (I_n, M, \dots, M^{d-1}) est une base de $\mathbb{R}[M]$, c'est aussi bien sûr une base de $\mathbb{R}_{d-1}[M]$ et comme $d \leq n$, $\mathbb{R}[M] = \mathbb{R}_{d-1}[M] = \mathbb{R}_{n-1}[M]$ ce qui permet de conclure

$$H \in \text{Ker}(df(M)) \iff \forall P \in \mathbb{R}[X], \quad \text{tr}(P(M)H) = 0$$

Ceci signifie que

$$H \in \text{Ker}(df(M)) \iff H^\top \in \mathbb{R}[M]^\perp$$

en considérant le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ et donc

$$\text{rg}(df(M)) = n^2 - \dim \text{Ker}(df(M)) = n^2 - (n^2 - \dim \mathbb{R}[X]) = d = \deg \pi_M$$

On a montré que $\text{rg}(df(M))$ est bien le degré du polynôme minimal de M .

3. Il reste à montrer que

$$\{M \in \mathcal{M}_n(\mathbb{R}), \text{rg}(df(M)) = n\} \text{ est ouvert.}$$

Or c'est l'image réciproque par df , continue (car f est de classe \mathcal{C}^1 , polynomiale en les coefficients de M), de

$$\{u \in \mathcal{L}(\mathcal{M}_n(\mathbb{R}), \mathbb{R}^n), \text{rg } u = n\}.$$

Le fait que ce dernier ensemble soit un ouvert est plus facile à voir matriciellement :

$$\{A \in \mathcal{M}_{n,n^2}(\mathbb{R}), \text{rg } A = n\} \text{ est ouvert.}$$

En effet, si A est dans cet ensemble, on peut sélectionner n colonnes de A définissant une matrice extraite inversible. Au voisinage de A , par continuité du déterminant, cette même matrice extraite restera inversible. Elle est donc de rang au moins n , et vu son nombre de lignes, au plus n , donc exactement n .

Remarque : c'est la semi-continuité inférieure du rang. On montre plus généralement qu'en toutes circonstances,

$$\{A \in \mathcal{M}_{n,p}(\mathbb{R}), \text{rg}(A) \geq r\} \text{ est un ouvert}$$

en s'aidant que d'une matrice de cet ensemble on peut toujours extraire une matrice carrée $r \times r$ inversible, et en utilisant la continuité du déterminant.

5 Vous reprendrez encore du principe variationnel ?

Soit E euclidien et $u \in \mathcal{S}(E)$.

1. Montrer que $f : x \in E \mapsto (u(x)|x)$ est différentiable sur E et calculer sa différentielle en tout point.
2. Montrer que $F : x \in E \setminus \{0_E\} \mapsto \frac{(u(x)|x)}{\|x\|^2}$ est différentiable sur $E \setminus \{0_E\}$ et que ses points critiques sont les vecteurs propres de u .

Solution de 5 : Vous reprendrez encore du principe variationnel ?

6 X Soient $n \in \mathbb{N} \setminus \{0, 1\}$ et

$$\Gamma = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 = 1 \text{ et } \sum_{i=1}^n x_i = 0 \right\}.$$

Maximiser $\Phi : (x_1, \dots, x_n) \mapsto x_1 x_2 + x_2 x_3 + \dots + x_{n-1} x_n + x_n x_1$ sur Γ .

Solution de 6 : X

Source : ddmaths

Considérons

$$J = \begin{pmatrix} 0 & 1 & & & (0) \\ \vdots & \ddots & \ddots & & \\ \vdots & & \ddots & & \\ 0 & & & 1 & \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

Pour $x = (x_1 \dots x_n)^\top \in \mathbb{R}^n$, on remarque

$$\Phi(x) = \frac{1}{2} x^\top A x \text{ avec } A = J + J^\top \in \mathcal{S}_n(\mathbb{R}).$$

Le polynôme caractéristique de J est $X^n - 1$. La matrice J est donc diagonalisable dans $\mathcal{M}_n(\mathbb{C})$ et l'on peut écrire $J = P D P^{-1}$ avec $D = \text{diag}(\omega^k)_{0 \leq k \leq n-1}$ où $\omega = e^{2i\pi/n}$.

On remarque $J^\top = J^{n-1}$ et donc $A = P \Delta P^{-1}$ avec $\Delta = \text{diag}(\omega^k + \omega^{(n-1)k})_{0 \leq k \leq n-1}$. Or, pour $k = 0, \dots, n-1$,

$$\omega^k + \omega^{(n-1)k} = \omega^k + \omega^{-k} = 2\text{Re}(\omega^k) = 2 \cos\left(\frac{2k\pi}{n}\right).$$

Cela détermine les valeurs propres de A et, puisque la matrice A est symétrique réelle, elle est orthogonalement semblable à la matrice Δ . On peut donc écrire $A = \Omega \Delta \Omega^{-1}$ avec $\Omega \in O_n(\mathbb{R})$. Notons que la première colonne de Ω est vecteur propre associé à la valeur propre 2. On remarque que le vecteur $U = (1 \dots 1)^\top$ est vecteur propre de A associé à la valeur propre 2. La première colonne de Ω est donc colinéaire à U . Quitte à réaliser un passage à l'opposé, on peut supposer que la première colonne (unitaire) de Ω vaut $\frac{1}{\sqrt{n}} U$.

Pour $x \in \mathbb{R}^n$,

$$\Phi(x) = \frac{1}{2} x^\top A x = \frac{1}{2} y^\top \Delta y = \sum_{k=0}^{n-1} \cos\left(\frac{2k\pi}{n}\right) y_{k+1}^2 \text{ avec } y = \Omega^\top x = (y_1, \dots, y_n).$$

On remarque $x^\top x = y^\top y$. On remarque aussi que $y_1 = \frac{1}{\sqrt{n}} U^\top x = \frac{1}{\sqrt{n}} (x_1 + \dots + x_n)$. On a donc

$$x \in \Gamma \Leftrightarrow y_1^2 + y_2^2 + \dots + y_n^2 = 1 \text{ et } y_1 = 0.$$

On peut alors simplifier

$$\Phi(x) = \sum_{k=1}^{n-1} \cos\left(\frac{2k\pi}{n}\right) y_{k+1}^2$$

et cette quantité est maximale pour $y_2 = 1$ et $y_3 = \dots = y_n = 0$.

La valeur maximale de Φ vaut donc $\cos\left(\frac{2\pi}{n}\right)$.