# Matrices minimisant la norme sur $\mathrm{SL}_n(\mathbb{R})$ :

## I Le développement

Le but de ce développement est de déterminer les matrices minimisant une certaine norme sur $\mathrm{SL}_n(\mathbb{R})$.

On commence tout d'abord par deux résultats préliminaires :

**Lemme 1 : [Gourdon, p.330]**
Soient $E$,$F$ et $G$ trois espace vectoriels de dimension finie et $\varphi : E\times F \longrightarrow G$ une application bilinéaire.
L'application $\varphi$ est différentiable sur $E\times F$ et :

$$\forall (x,y)\in E\times F,\ \forall (h,k)\in E\times F,\ d\varphi_{(x,y)}(h,k) = \varphi(x,k)+\varphi(h,y)$$

**Preuve :**
Soient $E$,$F$ et $G$ trois espace vectoriels de dimension finie et $\varphi : E\times F \longrightarrow G$ une application bilinéaire.

Soit $(x,y)\in E\times F$.
Pour tout $(h,k)\in E\times F$, on a (par bilinéarité de $\varphi$) :

$$\varphi((x,y)+(h,k)) = \varphi(x+h,y+k) = \varphi(x,y) + (\varphi(x,k)+\varphi(h,y)) + \varphi(h,k)$$

Or, l'application $(h,k)\longmapsto \varphi(x,k)+\varphi(h,y)$ est linéaire et puisque $\varphi$ est bilinéaire, il existe $C>0$ tel que $\|\varphi(h,k)\|_G \leq C\,\|(h,k)\|_{E\times F}^2$. On a alors :

$$\varphi((x,y)+(h,k)) = \varphi(x,y) + (\varphi(x,k+\varphi(h,y)) + o\left(\|(h,k)\|_{E\times F}\right)$$

Finalement, l'application $\varphi$ est différentiable sur $E\times F$ et :

$$\forall (x,y)\in E\times F,\ \forall (h,k)\in E\times F,\ d\varphi_{(x,y)}(h,k) = \varphi(x,k)+\varphi(h,y)$$

■

**Proposition 2 : [Gourdon, p.332]**
Soit $n\in\mathbb{N}^*$.
L'application $\det : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}$ est de classe $\mathcal{C}^\infty$ sur $\mathcal{M}_n(\mathbb{R})$ et :

$$\forall M\in\mathcal{M}_n(\mathbb{R}),\ \forall H\in\mathcal{M}_n(\mathbb{R}),\ d(\det)_M(H) = \mathrm{Tr}(\mathrm{Com}(M)^\intercal H)$$

**Preuve :**
Soit $n\in\mathbb{N}^*$.
L'application $\det : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}$ est de classe $\mathcal{C}^\infty$ sur $\mathcal{M}_n(\mathbb{R})$ car polynomiale en ses coefficients.

Soit $M\in\mathcal{M}_n(\mathbb{R})$.
Pour tout $t\in\mathbb{R}$ et tout $(i,j)\in [\![1;n]\!]^2$, on a par $n$-linéarité du déterminant que :

$$\det(M+tE_{i,j}) = \det(M) + \left|\begin{pmatrix} m_{1,1} & \cdots & 0 & \cdots & m_{1,n} \\ \vdots & & t & & \vdots \\ m_{n,1} & \cdots & 0 & \cdots & m_{n,n}\end{pmatrix}\right| = \det(M) + t\,\mathrm{Com}(M)_{i,j}$$

On a alors : $\frac{\partial \det}{\partial E_{i,j}}(M) = \mathrm{Com}(M)_{i,j}$.

Ainsi, pour tout $H=(h_{i,j})_{i,j\in[\![1;n]\!]}\in\mathcal{M}_n(\mathbb{R})$, on a :

$$\begin{aligned} d(\det)_M(H) &= \sum_{i,j=1}^n h_{i,j}\frac{\partial\det}{\partial E_{i,j}}(M) = \sum_{i=1}^n\left(\sum_{j=1}^n h_{i,j}\left(\mathrm{Com}(M)^\intercal\right)_{j,i}\right) \\ &= \sum_{i=1}^n \left(H\,\mathrm{Com}(M)^\intercal\right)_{i,i} = \mathrm{Tr}\left(\mathrm{Com}(M)^\intercal H\right)\end{aligned}$$

■

**Remarque 3 :**
En particulier, pour $M\in\mathrm{GL}_n(\mathbb{R})$, on a $\mathrm{Com}(M)^\intercal = \det(M)M^{-1}$, donc pour tout $H\in\mathcal{M}_n(\mathbb{R})$ on a : $d(\det)_M(H) = \det(M)\,\mathrm{Tr}\left(M^{-1}H\right)$.

On énonce désormais le résultat qui nous intéresse :

**Proposition 4 : [Gourdon, p.341]**
Si l'on munit $\mathcal{M}_n(\mathbb{R})$ de la norme $\|\cdot\|_2 : M \longrightarrow \left(\sum_{i,j=1}^n m_{i,j}^2\right)^{\frac{1}{2}}$, alors le groupe des matrices orthogonales directes de $\mathrm{SO}_n(\mathbb{R})$ est l'ensemble des éléments de $\mathrm{SL}_n(\mathbb{R})$ de norme minimale.

**Preuve :**
On munit $\mathcal{M}_n(\mathbb{R})$ de la norme $\|\cdot\|_2 : M \longrightarrow \left(\sum_{i,j=1}^n m_{i,j}^2\right)^{\frac{1}{2}}$.
On remarque que pour tout $M\in\mathcal{M}_n(\mathbb{R})$, on a $\|M\|_2^2 = \mathrm{Tr}\left(M^\intercal M\right)$. Il s'agit donc de minimiser la fonction $f$ définie sur $\mathcal{M}_n(\mathbb{R})$ par $f(M) = \mathrm{Tr}\left(M^\intercal M\right)$ sous la contrainte $g(M)=0$, avec $g(M)=\det(M)-1$.

Or, l'application $f$ est une forme quadratique et $g$ est une forme multilinéaire sur un espace vectoriel de dimension finie, donc $f$ et $g$ sont continues. L'ensemble

$\mathrm{SL}_n(\mathbb{R}) = g^{-1}(\{0\})$ est donc un fermé de $\mathcal{M}_n(\mathbb{R})$ et ainsi le minimum de $f$ sur $\mathrm{SL}_n(\mathbb{R})$ est atteint en un point $A \in \mathrm{SL}_n(\mathbb{R})$.

Or par la proposition précédente, on a $dg_A(H) = \mathrm{Tr}\left(A^{-1}H\right)$. Donc $dg_A$ est non nulle et par le théorème des extrema liés, il existe un réel $\lambda$ tel que $df_A = \lambda dg_A$.

De plus, on a par le lemme que $df_A(H) = 2\,\mathrm{Tr}\left(A^\intercal H\right)$ et donc pour toute matrice $H \in \mathcal{M}_n(\mathbb{R})$ on a :

$$2\,\mathrm{Tr}\left(A^\intercal H\right) = \lambda\,\mathrm{Tr}\left(A^{-1}H\right)$$

soit :

$$\mathrm{Tr}\left(\left(2A^\intercal - \lambda A^{-1}\right)H\right) = 0$$

On a alors $2A^\intercal = \lambda A^{-1}$ (car $\mathrm{Tr}(A^\intercal B)$ est un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$). Ainsi, $2A^\intercal A = \lambda I_n$ et par $n$-linéarité du déterminant on a : $2^n = \lambda^n$.

Or, $A^\intercal A$ est une matrice positive, donc $n\lambda = 2\,\mathrm{Tr}\left(A^\intercal A\right) \geq 0$ et ainsi $\lambda = 2$. On a alors $A^\intercal A = I_n$, c'est-à-dire $A \in \mathrm{SO}_n(\mathbb{R})$ et le minimum vaut $f(A) = \mathrm{Tr}\left(A^\intercal A\right) = \mathrm{Tr}(I_n) = n$.

Réciproquement, toute matrice de $\mathrm{SO}_n(\mathbb{R})$ est de norme au carré égale à $n$ et on a donc le résultat voulu.

■

# II Remarques sur le développement

## II.1 Résultat(s) utilisé(s)

Dans ce développement, on a utilisé quelques résultats sur la comatrice. On rappelle ainsi quelques généralités sur celle-ci pour une matrice $A \in \mathcal{M}_n(\mathbb{R})$ :

**Définition 5 : Mineur et cofacteur [Deschamps, p.1248]**
On considère $(i,j) \in [\![1;n]\!]^2$.
On appelle :
$*$ **mineur de** $a_{i,j}$ le déterminant $\Delta_{i,j}$ de la matrice extraite de $A$ obtenue en supprimant la $i$-ème ligne et la $j$-ème colonne de $A$.
$*$ **cofacteur de** $a_{i,j}$ le scalaire $(-1)^{i+j}\Delta_{i,j}$.

**Théorème 6 : [Deschamps, p.1248]**
On a le développement suivant par rapport à une colonne/ligne :

$$\forall j \in [\![1;n]\!],\ \det(A) = \sum_{i=1}^{n} a_{i,j}(-1)^{i+j}\Delta_{i,j} \text{ et } \forall i \in [\![1;n]\!],\ \det(A) = \sum_{j=1}^{n} a_{i,j}(-1)^{i+j}\Delta_{i,j}$$

**Définition 7 : Comatrice [Deschamps, p.1251] :**
On appelle **comatrice de** $A$ (notée $\mathrm{Com}(A)$) la matrice des cofacteurs de $A$.

**Proposition 8 : [Deschamps, p.1251]**
On a la relation $A\,\mathrm{Com}(A)^\intercal = \mathrm{Com}(A)^\intercal A = \det(A)I_n$.
En particulier, si $A \in \mathrm{GL}_n(\mathbb{K})$, alors $A^{-1} = \frac{1}{\det(A)}\mathrm{Com}(A)^\intercal$.

**Exemple 9 : [Deschamps, p.1251]**
Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible, alors $A^{-1} = \frac{1}{ad-bd}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

On a également utilisé le théorème des extrema liés dont on rappelle l'énoncé ainsi que la preuve :

**Théorème 10 : Théorème des extrema liés [Gourdon, p.337] :**
Soient $f, g_1, ..., g_r$ des fonctions réelles de classe $\mathcal{C}^1$ définies sur un ouvert non vide $\mathcal{U}$ de $\mathbb{R}^n$ et à valeurs réelles et $\Gamma = \{x \in \mathcal{U} \text{ tq } g_1(x) = ... = g_r(x) = 0\}$.
Si $f|_\Gamma$ admet un extremum relatif en $a \in \Gamma$ et si $dg_{1,a}, ..., dg_{r,a}$ sont linéairement indépendants, alors il existe $\lambda_1, ..., \lambda_r \in \mathbb{R}$ (uniques et appelés **multiplicateurs de Lagrange**) tels que $df_a = \sum_{i=1}^{r}\lambda_i dg_{i,a}$.

**Preuve :**
Soient $f, g_1, ..., g_r$ des fonctions réelles de classe $\mathcal{C}^1$ définies sur un ouvert non vide $\mathcal{U}$ de $\mathbb{R}^n$ et à valeurs réelles et $\Gamma = \{x \in \mathcal{U} \text{ tq } g_1(x) = ... = g_r(x) = 0\}$.

Posons $s = n - r$ et identifions $\mathbb{R}^n$ à $\mathbb{R}^s \times \mathbb{R}^r$. On écrit alors les éléments de $\mathbb{R}^n$ sous la forme $(x, y) = (x_1, ..., x_s; y_1, ..., y_r)$ et on pose $a = (\alpha, \beta) \in \mathbb{R}^n$, où $\alpha \in \mathbb{R}^s$ et $\beta \in \mathbb{R}^r$.

Commençons par faire la remarque suivante :
On a nécessairement $r \geq n$ car les formes linéaire $dg_{i,a}$ forment une famille libre et la dimension du dual de $\mathbb{R}^n$ est $n$. Par ailleurs, si $r = n$, le théorème est évident car les $dg_{i,a}$ forment une base de $(\mathbb{R}^n)^*$. On peut donc supposer que $r \leq n - 1$, c'est-à-dire $s \geq 1$.

Les formes linéaires $dg_{i,a}$ forment une famille libre, la matrice :

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \cdots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \cdots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \cdots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}$$

est donc de rang $r$. On peut donc en extraire une sous-matrice de taille $r \times r$ inversible. Quitte à changer le nom des variables, on peut supposer que :

$$\det\left(\left(\frac{\partial g_i}{\partial y_j}(a)\right)_{1 \leq i,j \leq r}\right) = \frac{D(g_1, ..., g_r)}{D(y_1, ..., y_r)} \neq 0$$

D'après le théorème des fonctions implicites, on peut alors trouver un voisinage ouvert $\mathcal{U}'$ de $\alpha$ dans $\mathbb{R}^s$, un voisinage ouvert $\Omega$ de $a = (\alpha, \beta)$ dans $\mathbb{R}^n$ et une fonction $\varphi = (\varphi_1, ..., \varphi_r) : \mathcal{U}' \longrightarrow \mathbb{R}^r$ de classe $\mathcal{C}^1$ tels que (en notant $g = (g_1, ..., g_r)$) :

$$\big(g(x, y) = 0 \text{ avec } x \in \mathcal{U}' \text{ et } (x, y) \in \Omega\big) \iff (y = \varphi(x))$$

En d'autres termes, sur un voisinage de $a$, les éléments de $\Gamma = \{x \in \mathcal{U} \text{ tq } g(x) = 0\}$ s'écrivent $(x, \varphi(x))$. Posons alors $h(x) = f(x, \varphi(x))$. La fonction $h$ admet donc un extremum local en $x = \alpha$ (car $(\alpha, \varphi(\alpha)) = a$ et $(x, \varphi(x)) \in \Omega$), ce qui entraîne :

$$\forall i \in [\![1; s]\!],\ \frac{\partial h}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(\alpha) \frac{\partial f}{\partial y_j}(a) = 0 \quad (*)$$

Par ailleurs, en écrivant les dérivées partielles par rapport aux $x_i$ de $g(x, \varphi(x)) = 0$, on tire :

$$\forall k \in [\![1, r]\!],\ \forall i \in [\![1; s]\!],\ \frac{\partial g_k}{\partial x_i}(a) + \sum_{j=1}^{r} \frac{\partial \varphi_j}{\partial x_i}(\alpha) \frac{\partial g_k}{\partial y_j}(a) = 0 \quad (**)$$

Autrement dit, les s premiers vecteurs colonnes de la matrice :

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \cdots & \frac{\partial f}{\partial x_s}(a) & \frac{\partial f}{\partial y_1}(a) & \cdots & \frac{\partial f}{\partial y_r}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_s}(a) & \frac{\partial g_1}{\partial y_1}(a) & \cdots & \frac{\partial g_1}{\partial y_r}(a) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_r}{\partial x_1}(a) & \cdots & \frac{\partial g_r}{\partial x_s}(a) & \frac{\partial g_r}{\partial y_1}(a) & \cdots & \frac{\partial g_r}{\partial y_r}(a) \end{pmatrix}$$

s'expriment, d'après $(*)$ et $(**)$, linéairement en fonction de ses $r$ derniers vecteurs colonnes et donc $\mathrm{rg}(M) \leq r$.

Or le rang des vecteurs lignes est égal au rang des vecteurs colonnes de $M$ (car $\mathrm{rg}\,(M^\intercal) = \mathrm{rg}(M)$), donc les $r + 1$ vecteurs lignes de $M$ forment une famille liée, ce qui entraîne l'existence de $r$ réels $\mu_0, ..., \mu_r$ non tous nuls tels que $\sum_{i=1}^{r} \mu_i dg_{i,a} = 0$. Or, comme la famille $(dg_{i,a})_{i \in [\![1;r]\!]}$ est libre, on a $\mu_0 = 0$, et en posant pour tout $i \in [\![1; r]\!], \lambda_i = \frac{-\mu_i}{\mu_0}$ on en déduit que $df_a = \sum_{i=1}^{r} \lambda_i dg_{i,a}$.

■

## II.2 Pour aller plus loin...

### II.2.1 Différentielle du déterminant

Il est possible de calculer la différentielle du déterminant d'une autre manière :
$*$ Calcul de la différentielle en l'identité :
Soient $M \in \mathcal{M}_n(\mathbb{R})$ et $\lambda_1, ..., \lambda_n$ ses valeurs propres complexes.
Pour tout $t \in \mathbb{R}$, on a :

$$\det(I_n + tM) = \prod_{i=1}^{n} 1 + t\lambda_i = 1 + t\,\mathrm{Tr}(M) + O\left(t^2\right) = 1 + t\,\mathrm{Tr}(M) + o(t)$$

On a donc $d(\det)_{I_n}(M) = \mathrm{Tr}(M)$.

$*$ Calcul de la différentielle pour $X \in \mathrm{GL}_n(\mathbb{R})$ :
Pour tout $X \in \mathrm{GL}_n(\mathbb{R})$ et tout $H \in \mathcal{M}_n(\mathbb{R})$ on a :

$$\begin{aligned} \det(X + H) &= \det\left(X\left(I_n + X^{-1}H\right)\right) = \det(X)\left(1 + \mathrm{Tr}\left(X^{-1}H\right) + o\left(\|H\|\right)\right) \\ &= \det(X) + \mathrm{Tr}\left(\det(X)X^{-1}H\right) + o\left(\|H\|\right) \\ &= \det(X) + \mathrm{Tr}\left(\mathrm{Com}(X)^\intercal H\right) + o\left(\|H\|\right) \end{aligned}$$

On a donc $d(\det)_X(H) = \mathrm{Tr}\left(\mathrm{Com}(X)^\intercal H\right)$.

$*$ Calcul de la différentielle pour $X \in \mathcal{M}_n(\mathbb{R})$ :
L'application $X \longmapsto \mathrm{Com}(X)^\intercal$ est continue, donc $f : X \longmapsto \mathrm{Tr}\left(\mathrm{Com}(X)^\intercal\right)$ est continue également. De plus, $g : X \longmapsto d(\det)_X$ est également continue et puisque $f = g$ sur $\mathrm{GL}_n(\mathbb{R})$ qui est un ouvert dense de $\mathcal{M}_n(\mathbb{R})$, on en déduit que $f = g$ sur $\mathcal{M}_n(\mathbb{R})$ tout entier.
Ainsi, on a $d(\det)_X(H) = \mathrm{Tr}\left(\mathrm{Com}(X)^\intercal H\right)$.

### II.2.2 Une autre méthode...

On peut également chercher les matrices minimisant la norme $\|\cdot\|_2$ sur $\mathrm{SL}_n(\mathbb{R})$ d'une autre manière (en gardant les mêmes notations) :
Toute matrice $M \in \mathcal{M}_n(\mathbb{R})$ peut s'écrire sous la forme $M = \Omega S$ avec $\Omega$ une matrice orthogonale et $S$ une matrice symétrique positive. On a alors $\|M\|_2^2 = \|\Omega S\|_2^2 = \|S\|_2^2$.

Comme $S$ est symétrique réelle, elle se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe $P \in \mathrm{O}_n(\mathbb{R})$ telle que $P^\intercal SP = D$ est diagonale. En notant $\lambda_1, ..., \lambda_n$ les valeurs propres (réelles positives) de $S$, on a alors :

$$\|M\|_2^2 = \mathrm{Tr}\left(S^2\right) = \mathrm{Tr}\left(D^2\right) = \sum_{k=1}^{n} \lambda_k^2$$

En utilisant l'inégalité entre moyennes géométrique et arithmétique :

$$\left(\prod_{k=1}^{n} \lambda_k^2\right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{k=1}^{n} \lambda_k^2$$

On en déduit que :

$$\forall M \in \mathcal{M}_n(\mathbb{R}),\ \|M\|_2^2 \geq n \left(\prod_{k=1}^{n} \lambda_k^2\right)^{\frac{1}{n}} = n\left(\det\left(S^2\right)\right)^{\frac{1}{n}} = n\left(\det\left(M^2\right)\right)^{\frac{1}{n}}$$

Pour toute matrice $M \in \mathrm{SL}_n(\mathbb{R})$, on a alors $\|M\|_2^2 \geq n$ et ainsi la distance au carré de $\mathrm{SL}_n(\mathbb{R})$ à $0_{\mathcal{M}_n(\mathbb{R})}$ est minorée par $n$. Or, pour $M \in \mathrm{SO}_n(\mathbb{R})$, on a $\|M\|_2^2 = \mathrm{Tr}\left(M^\intercal M\right) = \mathrm{Tr}(I_n) = n$, d'où l'inégalité inverse.

Finalement, on retrouve bien le résultat que l'on avait démontré.

> **Remarque 11 : [Caldero, p.190 + 226]**
> Il est possible d'effectuer d'autres calculs de distance dans les espaces de matrices en utilisant les mêmes arguments que précédemment (décomposition ou théorème des extrema liés).

## II.3 Recasages

Recasages : 148 - 149 - 161 - 206 - 214 - 215 - 219.

# III Bibliographie


— Xavier Gourdon, *Les maths en tête, Analyse*.

— Claude Deschamps, *Maths MPSI Tout-en-un*.

— Philippe Caldero, *Carnet de voyage en Analystan*.