TD: normes vectorielles et matricielles, matrices orthogonales

(*) très facile (1-4 lignes), (**) moyen, (***) des subtilités

Exercice 1 : Normes de vecteurs

On rappelle les 3 normes usuelles pour les vecteurs $\mathbf{x}$ de $\mathbb{R}^n$:

- la 2-norme ou norme Euclidienne, définie ainsi $\|\mathbf{x}\|_2 := \left(x_1^2 + \cdots + x_1^n\right)^{\frac{1}{2}}$
- la ∞-norme, définie ainsi $\|\mathbf{x}\|_\infty := \max_{i=1,\ldots,n} |x_i|$
- la 1-norme, définie ainsi $\|\mathbf{x}\|_1 := |x_1| + |x_2| + \cdots + |x_n|$

1. (*) Soient les vecteurs $\mathbf{x} = \begin{bmatrix}-2 & 1\end{bmatrix}^\mathsf{T}, \mathbf{y} = \begin{bmatrix}3 & -4 & 1\end{bmatrix}^\mathsf{T}, \mathbf{z} = \begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}^\mathsf{T}$. Calculez leurs $1, 2, \infty$-normes. Vérifiez que $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$, $\|\mathbf{y}\|_\infty \leq \|\mathbf{y}\|_2 \leq \|\mathbf{y}\|_1$, $\|\mathbf{z}\|_\infty \leq \|\mathbf{z}\|_2 \leq \|\mathbf{z}\|_1$.
2. (*) Montrez que, pour tout vecteur $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2$
3. (*) Montrez que, pour tout vecteur $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$
4. (*) Montrez que $\forall p = 1, 2$ (et en fait n'importe quel entier), $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_p \leq n^{\frac{1}{p}} \|\mathbf{x}\|_\infty$, pour $\mathbf{x} \in \mathbb{R}^n$. Les $1, 2, \infty$-normes sont donc équivalentes à une constante n ou $\sqrt{n}$ près.
5. (*) Montrez que si P est une matrice orthogonale (i.e. $PP^\mathsf{T} = P^\mathsf{T}P = \mathbf{I}_n$), alors $\|P\mathbf{x}\|_2 = \|\mathbf{x}\|_2$. Utilisez le fait que $\|\mathbf{u}\|_2^2 = \mathbf{u}^\mathsf{T}\mathbf{u}$ si $\mathbf{u}$ est un vecteur.

Take-away message

- Les normes servent à mesurer la longueur des vecteurs
- Les trois normes usuelles $(1, 2, \infty)$ sont équivalentes à constante $\sqrt{n}$ près
- La 2-norme (ou norme Euclidienne) d'un vecteur n'est pas changée par toute application d'une matrice orthogonale. Autrement dit les matrices orthogonales représentent les rotations et les réflexions.

Exercice 2 : Normes matricielles

On rappelle qu'on peut définir la norme matricielle subordonnée à une norme quelconque $\|\cdot\|_*$ ainsi, si A est une matrice:

$$\|A\|_* := \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|_*}{\|\mathbf{x}\|_*}, \text{ ou (équivalent) } \|A\|_* := \sup_{\mathbf{x}\neq\mathbf{0}, \|\mathbf{x}\|_*=1} \|A\mathbf{x}\|_*$$

On vérifie que:

- la 1-norme est la norme "somme des colonnes": $\|A\|_1 = \max_{j=1\ldots n} \sum_{i=1}^m |a_{ij}|$.
- la ∞-norme est la norme "somme des lignes": $\|A\|_\infty = \max_{i=1\ldots m} \sum_{j=1}^n |a_{ij}|$.
- la 2-norme ou *norme spectrale* est associé à la plus grande valeur singulière de $\mathbf{A}$.
- Une autre norme matricielle est la norme de Fröbenius: $\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}$.

1. (*) Vérifiez que si $\mathbf{I}_n$ est la matrice identité d'ordre n, $\|\mathbf{I}_n\|_* = 1$ et $\|\mathbf{I}\|_F = \sqrt{n}$.
2. (*) Vérifiez que ces normes satisfont $\|\lambda A\|_* = |\lambda| \|A\|_*$.

3. (*) Vérifiez que $\|A\|_F = \sqrt{\operatorname{tr}(A^\intercal A)}$ pour une matrice carrée d'ordre 2, où $\operatorname{tr}(\cdot)$ désigne la *trace* de la matrice, donc la somme de ses coefficients diagonaux.

4. (*) Vérifiez que pour tout vecteur $\mathbf{y}$, $\|A\mathbf{y}\|_* \leq \|A\|\|\mathbf{y}\|_*$

5. (*) Soit $A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Pour chaque vecteur $\mathbf{x}(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$, quelle est la valeur de $\|Ax\|_2^2$? Quand est-ce que cette longueur est minimale ou maximale (dériver la fonction de t précédente), et quelles sont alors ces longueurs ? Déduisez $\|A\|_2$.

6. (*) Avec la même matrice A que précédemment, tracez la courbe $A\mathbf{x}(t)$. Soit R_α la matrice de rotation $\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$. Son inverse est sa transposée. Soit $B = R_\alpha A R_\alpha^\intercal$ pour $\alpha = 30$. Tracez la courbe $B\mathbf{x}(t)$. Que constatez-vous ? En déduire que $\|A\|_2 = \|B\|_2$ pour tout α. On pourra utiliser le code python suivant pour tracer les courbes.

```python
get_ipython().run_line_magic('matplotlib', 'inline') # jupyter-notebook
import numpy as np                    # bibliotheques utiles
import matplotlib as mp
import matplotlib.pyplot as plt
from math import *
# On construit les vecteurs x(t), y(t)
T = np.arange(0.0, 2*pi, 0.05)
X = np.cos(T)
Y = np.sin(T)
# On applique la matrice diagonale D à chaque vecteur
A = np.matrix([[1,0],[0,4]])
E= A*np.matrix([X,Y])
EX=np.array(E[0]).flatten() # coordonnées x
EY=np.array(E[1]).flatten() # coordonnées y
# On construit une matrice de rotation d'angle t=30°
t=30.*pi/180.
R=np.matrix([[cos(t),-sin(t)],[sin(t),cos(t)]])
R1=np.matrix([[cos(t),sin(t)],[-sin(t),cos(t)]])
# On applique la matrice diagonalisable A=RDR^t à chaque vecteur
B=R*A*R1
ER=B*np.matrix([X,Y])
ERX=np.array(ER[0]).flatten() # coordonnées x
ERY=np.array(ER[1]).flatten() # coordonnées y
# On affiche le tout
plt.figure(figsize=(10,10))
plt.plot(EX,EY,color='blue')
plt.plot(ERX,ERY,color='black')
plt.xlim([-4.5,4.5])
plt.ylim([-4.5,4.5])
```

7. (***) Soit A une matrice qui s'écrit $A = PDP^\intercal$, avec D matrice diagonale à coefficients réels et P orthogonale (i.e. $PP^\intercal = P^\intercal P = \mathbf{I}_n$). Montrez que $\|A\|_2 = \max_{i=1,\ldots,n} |d_i|$, c'est-à-dire la valeur absolue du plus grand coefficient de D.

Take-away message

- Les normes matricielles sont équivalentes à constante près. Seule la norme de Frobenius n'assigne pas 1 à la matrice identité.
- Les normes matricielles sont sous-additives $\|A+B\| \leq \|A\| + \|B\|$ et sous-multiplicatives $\|AB\| \leq \|A\|\|B\|$.
- La 2-norme matricielle n'est pas affectée par une matrice orthogonale.
- La 2-norme d'une matrice A est la valeur absolue de la plus grande valeur singulière de A, c'est-à-dire le plus grand coefficient en valeur absolue de sa diagonalisation.
- La 2-norme d'une matrice A exprime la plus grande dilatation possible qu'elle peut faire lorsqu'appliquée à un vecteur.