

Corrigé du Sujet HEC 2026

Partie I : Préliminaires

1. Soit $j \in [[1, n]]$. Par définition de P_σ , la j -ème colonne de P_σ est composée des coordonnées de $\phi_\sigma(e_j)$ dans la base canonique. On a donc pour tout $i \in [[1, n]]$,

$$(\phi_\sigma(e_j))_i = (P_\sigma)_{i,j} = (\delta_{i,\sigma(j)}) = \begin{cases} 1 & \text{si } i = \sigma(j) \\ 0 & \text{sinon} \end{cases}$$

donc :

$$\phi_\sigma(e_j) = (0, \dots, 0, 1, 0, \dots, 0) = e_{\sigma(j)}$$

2. • Soit $(\sigma, \tau) \in \mathcal{P}_n^2$. Pour tout $j \in [[1, n]]$, d'après la question précédente,

$$\phi_\sigma \circ \phi_\tau(e_j) = \phi_\sigma(e_{\tau(j)}) = e_{\sigma \circ \tau(j)} = \phi_{\sigma \circ \tau}(e_j)$$

Comme les applications linéaires $\phi_\sigma \circ \phi_\tau$ et $\phi_{\sigma \circ \tau}$ coïncident sur une base de E , elles sont égales. Par conséquent leurs matrices canoniquement associées sont égales et donc $P_\sigma P_\tau = P_{\sigma \circ \tau}$

- Soit P une matrice de permutation. Notons σ la permutation associée. On remarque alors que

$$P_\sigma P_{\sigma^{-1}} = P_{\sigma \circ \sigma^{-1}} = P_{Id} = I_n$$

donc la matrice P est inversible, d'inverse $(P_\sigma)^{-1} = P_{\sigma^{-1}}$.

Bilan : toute matrice de permutation P_σ est inversible et son inverse est $P_{\sigma^{-1}}$

3. Soit $\sigma \in \mathcal{P}_n$. On remarque que :

$$({}^t P_\sigma)_{i,j} = (P_\sigma)_{j,i} = \delta_{j,\sigma(i)} = \delta_{\sigma(i),j} = \delta_{i,\sigma^{-1}(j)} = (P_{\sigma^{-1}})_{i,j}$$

En effet, $\sigma(i) = j \Leftrightarrow i = \sigma^{-1}(j)$ donc $\delta_{\sigma(i),j} = \delta_{i,\sigma^{-1}(j)}$.

On a donc ${}^t P_\sigma = P_{\sigma^{-1}}$ et donc d'après ce qui précède :

$${}^t P_\sigma P_\sigma = P_{Id} = I_n$$

Bilan : toute matrice de permutation est orthogonale

4. On procède comme demandé et on montre par récurrence. Mais dans l'idée, il suffit de classer les valeurs $(x_1, \dots, x_n)$ dans l'ordre croissant !!

- Si $n = 1$, il suffit de prendre $\alpha = Id$ et on a bien $(x_1) = (x_{\alpha(1)})$ qui convient.
- Supposons la propriété vraie pour un entier n . On considère $n + 1$ nombres réels $(x_1, \dots, x_{n+1})$. Soit i_0 un entier tel que pour tout $i \in [[1, n]]$, $x_i \geq x_{i_0}$. On note alors $\alpha(n + 1) = i_0$. Il reste donc n nombres $(x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_{n+1})$ que l'on va noter $(y_1, \dots, y_n)$. Par HR, il existe une permutation β de l'ensemble $I = [[1, n]]$ telle que

$$y_{\beta(1)} \geq \dots \geq y_{\beta(n)}$$

Soit $i \in [[1, n]]$. On note alors :

$$\alpha(i) = \begin{cases} \beta(i) & \text{si } \beta(i) \leq i_0 - 1 \\ \beta(i) + 1 & \text{si } \beta(i) \geq i_0 \end{cases}$$

Alors α est une bijection et on a bien :

$$x_{\alpha(1)} \geq \dots \geq x_{\alpha(n)} \geq x_{\alpha(n+1)}$$

- Le résultat souhaité est donc démontré.

5. Les termes étant rangés dans l'ordre croissant et représentant la même suite d'entiers $(x_1, \dots, x_n)$, on a nécessairement les deux termes les plus grands qui sont égaux, donc $x_{\alpha(1)} = x_{\beta(1)}$. Puis en réitérant cet argument, $x_{\alpha(2)} = x_{\beta(2)}$ etc... Finalement par récurrence on obtient bien le résultat souhaité.

```
6. def permutevecteur(X):
    n=len(X)
    alpha=np.arange(0,n,1)
    Y=X.copy()
    for i in range(n):
        imax=i
        for k in range(i,n):
            if Y[k]>Y[imax]:
                imax=k
        if imax>i:
            Y[i],Y[imax]=Y[imax],Y[i]
            alpha[i],alpha[imax]=alpha[imax],alpha[i]
    return Y, alpha
```

Idée : dans la 1ère boucle, lorsque $i = 0$, on suppose par défaut que la valeur la plus grande est dans la case $Y[0]$. Puis on parcourt le tableau dans la seconde boucle et on remplace **imax** par **k** à chaque fois que l'on trouve une valeur plus grande. On se retrouve donc avec ma variable **imax** qui va contenir l'indice de la valeur la plus grande du tableau. Si les valeurs de **i** (ici 0) et **imax** sont différentes, on va permuter les valeurs contenues dans les deux cases $Y[i]$ et $Y[imax]$ de sorte que la plus grande valeur soit en 1ère position; et on échange les valeurs de $\alpha[0]$ et $\alpha[imax]$ de sorte que la première case de α vaut imax .

On recommence ensuite dans le sous-tableau des cases entre 1 et $n-1$, etc...

Partie II : Matrices bistochastiques.

Théorème de Birkhoff- Von Neumann en basses dimensions

7. (a) • Soit S une matrice orthostochastique. Par définition, ses coefficients sont tous positifs d'où (a). Comme Q est orthogonale, on sait que $Q \cdot {}^t Q = I$. On a donc pour tout $(i, j) \in [[1, n]]$

$$(Q \cdot {}^t Q)_{i,j} = \sum_{k=1}^n Q_{i,k} \cdot Q_{j,k} = \delta_{i,j}$$

D'où si $i = j$:

$$\sum_{k=1}^n (Q_{i,k})^2 = 1 \Leftrightarrow \sum_{k=1}^n S_{i,k} = 1$$

d'où (c). De même en calculant

$$({}^t Q \cdot Q)_{i,i} = \sum_{k=1}^n (Q_{k,i})^2 = 1$$

on trouve (b).

Donc toute matrice orthostochastique est bistochastique

- Soit $P = P_\sigma$ une matrice de permutation. Pour tout $(i, j) \in [[1, n]]^2$, $(P_\sigma)_{i,j} = \delta_{i,\sigma(j)}$ donc les coefficients de P sont positifs et (a) est vraie. Par ailleurs, comme σ est une permutation, sur la ligne i il n'y aura d'une seule valeur de j pour laquelle $\delta_{i,\sigma(j)} = 1$, les autres valeurs étant nulles. Par conséquent on obtient (b).

Enfin, pour tout $j \in [[1, n]]$, sur la colonne numéro $\sigma(j)$ il y a aussi une unique valeur qui vaut 1, les autres étant égales à 0. Comme σ est une permutation de $[[1, n]]$, ceci sera vrai pour toutes les colonnes de la matrice, on obtient donc la relation (c).

Donc toute matrice de permutation est bistochastique

- (b) Soit $S = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$. Alors S est bistochastique. Supposons que S soit orthostochastique. Alors il existe une matrice orthogonale Q telle que pour tout (i, j) on ait $(Q_{i,j})^2 = (S)_{i,j}$, donc $Q_{i,j} = \pm \frac{1}{\sqrt{3}}$. Alors comme Q est orthogonale, on devrait avoir $Q \cdot Q = I_3$ et en particulier par exemple :

$$(Q \cdot Q)_{1,2} = 0 \Leftrightarrow \pm \frac{1}{3} \pm \frac{1}{3} \pm \frac{1}{3} = 0$$

ce qui est impossible, quels que soient les signes considérés ! On a donc une absurdité et la matrice S ci-dessus n'est pas orthostochastique.

Bilan : une matrice bistochastique n'est pas toujours orthostochastique.

8. (a) Si $n = 2$, il n'y a que deux permutations de $[1, 2]$ à savoir $(1, 2)$ (c'est l'identité) et $(2, 1)$. Les matrices associées sont

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ et } P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

- (b) En utilisant la condition sur les lignes, une matrice bistochastique de $\mathcal{M}_2(\mathbb{R})$ s'écrit nécessairement sous la forme suivante :

$$S = \begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix}$$

où $(a, b) \in [0, 1]$. De plus, avec la condition sur les colonnes : $a + b = 1$ et $1 - a + 1 - b = 1$ donc $b = 1 - a$ et

$$S = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix} = a \cdot I_2 + (1-a) \cdot P$$

Donc si S est bistochastique, alors S s'écrit sous la forme ci-dessus. Réciproquement, si S est de la forme ci-dessus, il est évident que S est bistochastique.

Le raisonnement par équivalences va être particulièrement attendu ici !!!

9. (a) Il y a $3! = 6$ permutations de l'ensemble $\{1, 2, 3\}$.
 (b) Chacune de ces matrices a un seul 1 sur chaque ligne et sur chaque colonne et des 0 ailleurs, il s'agit donc bien d'une matrice de permutation.
 Si l'on considère la matrice P_1 , on voit que $(P_1)_{1,1} = \delta_{1,\sigma(1)} = 1$ donc $\sigma(1) = 1$. On a aussi $(P_1)_{2,3} = \delta_{2,\sigma(3)} = 1$ donc $\sigma(3) = 2$. La matrice P_1 correspond donc à la permutation $(1, 3, 2)$. De même, P_2 correspond à $(3, 2, 1)$, P_3 à $(2, 1, 3)$, et P_4 et P_5 correspondent aux deux permutations circulaires : $(3, 1, 2)$ et $(2, 3, 1)$.
 Il manque la matrice I_3 qui correspond à la permutation $(1, 2, 3)$ c'est à dire à l'application identité de $\{1, 2, 3\}$.
 (c) Soit $S \in \mathcal{M}_3(\mathbb{R})$ une matrice bistochastique. Comme la somme des coefficients sur chaque ligne et sur chaque colonne vaut 1, il est clair que S s'écrit sous la forme

$$S = \begin{pmatrix} S_{1,1} & S_{1,2} & 1 - S_{1,1} - S_{1,2} \\ S_{2,1} & S_{2,2} & 1 - S_{2,1} - S_{2,2} \\ 1 - S_{1,1} - S_{2,1} & 1 - S_{1,2} - S_{2,2} & S_{3,3} \end{pmatrix}$$

avec

$$S_{3,3} = 1 - (1 - S_{1,1} - S_{2,1}) - (1 - S_{1,2} - S_{2,2}) = S_{1,1} + S_{2,1} + S_{1,2} + S_{2,2} - 1$$

Pour que tous ces coefficients soient positifs, on doit donc avoir : $S_{1,1}, S_{1,2}, S_{2,1}, S_{2,2}$ qui appartiennent à $[0, 1]$,

$$0 \leq S_{1,1} + S_{1,2} \leq 1$$

$$0 \leq S_{1,1} + S_{2,1} \leq 1$$

$$0 \leq S_{2,1} + S_{2,2} \leq 1$$

$$0 \leq S_{1,2} + S_{2,2} \leq 1$$

$$1 \leq S_{1,1} + S_{2,1} + S_{1,2} + S_{2,2} \leq 2$$

- (d) On note β_0 le plus petit des coefficients diagonaux de S . Supposons que l'on ait par exemple $\beta_0 = S_{1,1}$. Alors :

$$S - \beta_0 \cdot I_3 = \begin{pmatrix} 0 & S_{1,2} & 1 - S_{1,1} - S_{1,2} \\ S_{2,1} & S_{2,2} - S_{1,1} & 1 - S_{2,1} - S_{2,2} \\ 1 - S_{1,1} - S_{2,1} & 1 - S_{1,2} - S_{2,2} & S_{3,3} - S_{1,1} \end{pmatrix}$$

d'où ensuite en travaillant sur les coefficients diagonaux :

$$\begin{aligned} S - S_{1,1} \cdot I_0 &= (S_{2,2} - S_{1,1}) \cdot P_2 + (S_{3,3} - S_{1,1}) \cdot P_3 \\ &+ \begin{pmatrix} 0 & S_{1,2} - S_{3,3} + S_{1,1} & 1 - S_{1,2} - S_{2,2} \\ S_{2,1} - S_{3,3} + S_{1,1} & 0 & 1 - S_{2,1} - S_{2,2} \\ 1 - S_{2,1} - S_{2,2} & 1 - S_{1,2} - S_{2,2} & 0 \end{pmatrix} \\ &= (S_{2,2} - S_{1,1}) \cdot P_2 + (S_{3,3} - S_{1,1}) \cdot P_3 + (1 - S_{2,1} - S_{2,2}) \cdot P_4 + (1 - S_{1,2} - S_{2,2}) \cdot P_5 \end{aligned}$$

où nous avons par exemple utilisé la relation :

$$S_{2,1} + S_{1,1} - S_{3,3} = S_{2,1} + S_{1,1} - (S_{1,1} + S_{2,1} + S_{1,2} + S_{2,2} - 1) = 1 - S_{1,2} - S_{2,2}$$

Les coefficients obtenus étant tous positifs, on obtient bien le résultat souhaité.

On obtient des résultats similaires lorsque $\beta_0 = S_{2,2}$ ou lorsque $\beta_0 = S_{3,3}$.

Bilan : il existe bien des réels positifs β_i tels que

$$S = \beta_0 \cdot I_3 + \sum_{i=1}^5 \beta_i P_i$$

- (e) Nous obtenons donc que toute matrice bistochastique de $\mathcal{M}_3(\mathbb{R})$ est une combinaison linéaire à coefficients positifs des matrices de permutation. Comme on a alors

$$S = \begin{pmatrix} \beta_0 + \beta_1 & \beta_3 + \beta_4 & \beta_2 + \beta_5 \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}$$

en sommant les termes de la première ligne de S , on obtient que $\beta_0 + \dots + \beta_5 = 1$.

Nous avons donc démontré le théorème de Birkoff-Von Neumann en dimension 3

A la question 8, nous avons déjà démontré ce théorème en dimension 2.

```
10. import numpy as np
import numpy.random as rd
def bistochastique(n,iter):
    A=rd.random([n,n])
    for k in range(iter):
        for i in range(n):
            S=np.sum(A,axis=1)[i]
            for j in range(n):
                A[i,j]=A[i,j]/S
        for j in range(n):
            T=np.sum(A,axis=0)[j]
            for i in range(n):
                A[i,j]=A[i,j]/T
    return A
```

Partie III : fonctions symétriques et fonctions S-convexes.

11. Soient $\sigma \in \mathcal{P}_n$ et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. En utilisant l'application φ_σ définie au début du I., on a :

$$P_\sigma X = \text{Mat}_B(\phi_\sigma(x_1 e_1 + \cdots + x_n e_n))$$

Or

$$\begin{aligned} \phi_\sigma\left(\sum_{i=1}^n x_i e_i\right) &= \sum_{i=1}^n x_i \phi_\sigma(e_i) \text{ par linéarité de } \phi_\sigma \\ &= \sum_{i=1}^n x_i e_{\sigma(i)} \\ &= \sum_{j=1}^n x_{\sigma^{-1}(j)} e_{\sigma(\sigma^{-1}(j))} \\ &= \sum_{j=1}^n x_{\sigma^{-1}(j)} e_j \end{aligned}$$

le changement d'indices $i = \sigma^{-1}(j)$ étant bien autorisé puisque σ (et donc σ^{-1}) est une bijection de $[[1, n]]$. D'où

$$P_\sigma X = \begin{pmatrix} x_{\sigma^{-1}(1)} \\ \vdots \\ x_{\sigma^{-1}(n)} \end{pmatrix} = X_{\sigma^{-1}} = x_{\sigma^{-1}}$$

si on identifie matrices colonnes et vecteurs de $\mathbb{R}^n$.

12. Raisonnons par double implication.
 $\Rightarrow$ Supposons que f est symétrique. Soit $P \in \mathbb{P}_n$. Alors il existe $\sigma \in \mathcal{P}_n$ tel que $P = P_\sigma$ et d'après ce qui précède,

$$f(PX) = f(x_{\sigma^{-1}}) = f(X_{\sigma^{-1}}) = f(X)$$

$\Leftarrow$ Supposons que pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, pour tout $P \in \mathcal{P}_n$, $f(PX) = f(X)$.
 Soit $\sigma \in \mathcal{P}_n$. Alors

$$f(x_\sigma) = f(P_{\sigma^{-1}}X) = f(X) = f(x)$$

et donc f est symétrique.

Bilan : f est symétrique si et seulement si pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, pour tout $P \in \mathcal{P}_n$, $f(PX) = f(X)$.

13. Supposons f symétrique. Soit $x \in \mathbb{R}^n$. D'après la Partie I, il existe une permutation σ telle que

$$x_{\sigma(1)} \geq \cdots \geq x_{\sigma(n)}$$

Comme f est symétrique, on a donc

$$f(x) = f(x_\sigma) = \hat{f}(\hat{x})$$

où $\hat{f}$ est la restriction de f à H_n .

Réciproquement, supposons qu'il existe $\hat{f}$ de H_n dans $\mathbb{R}$ telle que pour tout $x \in \mathbb{R}^n$, $f(x) = \hat{f}(\hat{x})$.
 Alors pour toute permutation σ de $[[1, n]]$,

$$f(x_\sigma) = \hat{f}(\hat{x}_\sigma) = \hat{f}(\hat{x}) = f(x)$$

où $\hat{x}_\sigma = \hat{x}$ puisque ceci ne dépend pas de l'ordre des termes $x_1, \dots, x_n$ (on les range dans l'ordre).

14. f_1 est symétrique car pour tout $\sigma \in \mathcal{P}(n)$, pour tout $x \in \mathbb{R}^n$,

$$f(x_\sigma) = \sum_{i=1}^n x_{\sigma(i)} = \sum_{j=1}^n x_j$$

en posant $j = \sigma(i)$, autorisé puisque σ est une bijection de $[[1, n]]$ sur $[[1, n]]$.
 f_2 n'est pas symétrique car $f_2(2, 0, \dots, 0) = 4 \neq f(0, 2, 0, \dots, 0) = 2$.
 f_3 est symétrique car les ensembles

$$\{|x_i - x_j|; (i, j) \in [[1, n]]^2\} \text{ et } \{|x_{\sigma(i)} - x_{\sigma(j)}|; (i, j) \in [[1, n]]^2\}$$

sont égaux.

15. (a) D'après le cours, comme f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$, l'application $g_{i,x}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Pour tout $t \in \mathbb{R}$,

$$g'_{i,x}(t) = \langle \nabla f(x + te_i), e_i \rangle = \partial_i(f)(x + te_i)$$

La dérivée en 0 de cette fonction est égale à $\partial_i(f)(x)$ (par définition de la dérivée partielle!).

- (b) **Attention**, il faut remarquer que pour tout $i \in [[1, n]]$ et pour tout $\sigma \in \mathcal{P}_n$,

$$(e_i)_\sigma = ((e_i)_{\sigma(1)}, \dots, (e_i)_{\sigma(n)})$$

et un coefficient $(e_i)_{\sigma(j)}$ vaut 1 lorsque $\sigma(j) = i$, c'est-à-dire lorsque $j = \sigma^{-1}(i)$. Donc

$$(e_i)_\sigma = e_{\sigma^{-1}(i)}$$

Revenons à ce qui est demandé.

Pour tout $t \in \mathbb{R}$,

$$\begin{aligned} g_{\sigma(i),x}(t) &= f(x + te_{\sigma(i)}) \\ &= f(x + t(e_i)_{\sigma^{-1}}) \\ &= f((x_\sigma)_{\sigma^{-1}} + t(e_i)_{\sigma^{-1}}) \\ &= f((x_\sigma + te_i)_{\sigma^{-1}}) \\ &= f(x_\sigma + te_i) \text{ car } f \text{ est symétrique} \\ &= g_{i,x_\sigma}(t) \end{aligned}$$

- (c) En appliquant la formule du 15.(a) et en l'évaluant en $t = 0$, on a donc d'après l'égalité précédente :

$$\partial_{\sigma(i)}f(x) = \partial_i f(x_\sigma)$$

16. On a vu que toute matrice de permutation est bistochastique. Si P est une matrice de permutation, P^{-1} l'est aussi, donc c'est aussi une matrice bistochastique. Comme f est S-convexe, on a pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$f(PX) \leq f(X)$$

et pour tout $Y \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$f(P^{-1}Y) \leq f(Y)$$

Comme $Y \mapsto P^{-1}Y$ est un automorphisme de $\mathcal{M}_{n,1}(\mathbb{R})$, on a donc aussi (en posant $X = P^{-1}Y$) pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$f(X) \leq f(PX)$$

donc en fin $f(PX) \leq f(X) \leq f(PX)$ d'où $f(PX) = f(X)$.

On en déduit que toute fonction S-convexe de $\mathbb{R}^n$ dans $\mathbb{R}$ est symétrique

17. Il s'agit de l'inégalité de convexité généralisée, aussi appelée inégalité de Jensen. La preuve se fait classiquement par récurrence, avec un peu de subtilité dans la partie Hérité.
 (Cf cours de 1ère année)

18. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction symétrique et convexe. Soit B une matrice bistochastique. On veut montrer que pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, on a $f(BX) \leq f(X)$.
D'après le théorème de Birkhoff - Von Neumann, il existe des matrices de permutation $P_1, \dots, P_m$ et des réels positifs $a_1, \dots, a_m$ tels que $a_1 + \dots + a_m = 1$ et $B = a_1 P_1 + \dots + a_m P_m$.
On a alors

$$\begin{aligned} f(BX) &= f\left(\sum_{i=1}^m a_i P_i X\right) \\ &\leq \sum_{i=1}^m a_i f(P_i X) \quad \text{par convexité généralisée} \\ &\leq \sum_{i=1}^m a_i f(X) \quad \text{car } f \text{ est symétrique} \\ &= f(X) \quad \text{car } a_1 + \dots + a_m = 1 \end{aligned}$$

Donc toute fonction f symétrique et convexe est S -convexe.

19. (a) Pour tout $t \in \mathbb{R}$,

$$g(t) = f((x_1, x_2) + t.(x_2 - x_1, x_1 - x_2))$$

Il s'agit de la composée d'une fonction affine avec une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$, donc d'une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$. Pour tout $t \in \mathbb{R}$,

$$\begin{aligned} g'(t) &= \langle \nabla f((x_1, x_2) + t.(x_2 - x_1, x_1 - x_2)), (x_2 - x_1, x_1 - x_2) \rangle \\ &= \partial_1 f((x_1, x_2) + t.(x_2 - x_1, x_1 - x_2)).(x_2 - x_1) + \partial_2 f((x_1, x_2) + t.(x_2 - x_1, x_1 - x_2)).(x_1 - x_2) \\ &= (x_2 - x_1).(\partial_1 f((x_1, x_2) + t.(x_2 - x_1, x_1 - x_2)) - \partial_2 f((x_1, x_2) + t.(x_2 - x_1, x_1 - x_2))) \end{aligned}$$

- (b) On remarque que si $t \in [0, 1]$,

$$\begin{pmatrix} (1-t)x_1 + tx_2 \\ tx_1 + (1-t)x_2 \end{pmatrix} = \begin{pmatrix} 1-t & t \\ t & 1-t \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = PX$$

où P est une matrice bistochastique. Par S -convexité de f , on a donc :

$$g(t) = f(PX) \leq f(X) = f(x_1, x_2) = g(0)$$

Donc $g(t) \leq g(0)$ pour tout $t \in [0, 1]$.

En particulier on a pour tout $t \in]0, 1]$, $\frac{g(t)-g(0)}{t-0} \leq 0$, et par passage à la limite quand t tend vers 0^+ , la fonction g étant dérivable en 0, on obtient : $g'(0) \leq 0$.

- (c) Avec la relation du 19.(a), on trouve que : pour tout $x = (x_1, x_2) \in \mathbb{R}^2$,

$$(x_2 - x_1).(\partial_1 f(x_1, x_2) - \partial_2 f(x_1, x_2)) \leq 0$$

et donc

$$(x_1 - x_2).(\partial_1 f(x) - \partial_2 f(x)) \geq 0$$

20. Dans le cas général, avec $i < j$, on va considérer

$$g(t) = f(x_1, \dots, (1-t)x_i + tx_j, \dots, tx_i + (1-t)x_j, \dots, x_n)$$

Il suffit ensuite de suivre les mêmes étapes que dans le cas où $n = 2$ pour obtenir le résultat.

21. Supposons que f est S -convexe. Alors d'après la question précédente, on a pour tout $(i, j) \in [[1, n]]^2$,

$$(x_i - x_j)(\partial_i f(x) - \partial_j f(x)) \geq 0 \Leftrightarrow (x_i - x_j).(h'(x_i) - h'(x_j)) \geq 0$$

Donc si $x_i > x_j$, nécessairement $h'(x_i) \geq h'(x_j)$. Ceci montre que la fonction h' est croissante sur $\mathbb{R}$, ce qui est une caractérisation de la convexité : la fonction h est convexe sur $\mathbb{R}$. Réciproquement, supposons que h est convexe et montrons que f est S -convexe.
Tout d'abord, la fonction f est symétrique puisque changer l'ordre des variables ne modifie pas la valeur de f .

Montrons que f est convexe. Comme h est convexe, pour tout $k \in [[1, n]]$, on a : $\forall (x_k, y_k) \in \mathbb{R}^2$, $\forall t \in [0, 1]$,

$$h(tx_k + (1-t)y_k) \leq th(x_k) + (1-t)h(y_k)$$

D'où en sommant : $\forall (x, y) \in \mathbb{R}^n$, $\forall t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

donc f est convexe.

D'après la question 18., f est symétrique et convexe donc f est S -convexe.

Partie IV : fonctions spectrales. Théorèmes de Davis et de Fan

22. Soit $k \in \mathbb{N}^*$. Tout d'abord, l'application F_k va bien de $\mathcal{S}_n(\mathbb{R})$ dans $\mathbb{R}$.

Soit $A \in \mathcal{S}_n(\mathbb{R})$ et $Q \in \mathcal{O}_n(\mathbb{R})$. Alors

$$\begin{aligned} F_k(QA^tQ) &= \text{Tr}((QA^tQ)^k) \\ &= \text{Tr}(QA^k{}^tQ) \quad \text{par récurrence classique} \\ &= \text{Tr}({}^tQQA^k) \\ &= \text{Tr}(A^k) = F_k(A) \end{aligned}$$

donc F_k est spectrale.

23. Soit $\sigma \in \mathcal{P}_n$. Pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$\begin{aligned} (P_\sigma D(X).{}^tP_\sigma)_{i,j} &= \sum_{k=1}^n \sum_{l=1}^n (P_\sigma)_{i,k}.(\text{Diag}(x_1, \dots, x_n))_{k,l}.({}^tP_\sigma)_{l,j} \\ &= \sum_{k=1}^n (P_\sigma)_{i,k}.x_k.(P_\sigma)_{j,k} \\ &= \sum_{k=1}^n \delta_{i,\sigma(k)}.x_k.\delta_{j,\sigma(k)} \\ &= \begin{cases} 0 & \text{si } \sigma(j) \neq \sigma(i) \Leftrightarrow i \neq j \\ x_{\sigma^{-1}(i)} & \text{si } i = j \end{cases} \end{aligned}$$

D'autre part,

$$(P_\sigma X)_i = \sum_{k=1}^n (P_\sigma)_{i,k}.x_k = x_{\sigma^{-1}(i)}$$

Donc

$$(D(P_\sigma X))_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ (P_\sigma X)_i = x_{\sigma^{-1}(i)} & \text{si } i = j \end{cases}$$

On a donc bien

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}), \quad P_\sigma D(X).{}^tP_\sigma = D(P_\sigma X)$$

24. C'est le théorème spectral pour les matrices symétriques !!

25. (a) Soit A une matrice symétrique. D'après la question précédente, on peut noter $A = Q\Lambda(A){}^tQ$.
Alors

$$F(A) = F(Q\Lambda(A){}^tQ) = F(\Lambda(A))$$

puisque F est spectrale.

- (b) Quelle que soit la permutation σ de $[[1, n]]$, quelle que soit $X \in \mathcal{M}_{n,1}$, on aura

$$\Lambda(D(X)) = \Lambda(D(X_\sigma))$$

En effet, les matrices diagonales $D(X)$ et $D(X_\sigma)$ auront le même spectre $\{x_1, \dots, x_n\}$.

On en déduit que

$$f(x) = F(D(X)) = F(D(X_\sigma)) = f(x_\sigma)$$

donc f est symétrique.

(c) Supposons F convexe. Soit $t \in [0, 1]$, et $(x, y) \in (\mathbb{R}^n)^2$. Alors

$$\begin{aligned} f(tx + (1-t)y) &= F\left(D \begin{pmatrix} tx_1 + (1-t)y_1 \\ \vdots \\ tx_n + (1-t)y_n \end{pmatrix}\right) \\ &= F(tD(X) + (1-t)D(Y)) \\ &\leq tF(D(X)) + (1-t)F(D(Y)) \\ &\leq tf(x) + (1-t)f(y) \end{aligned}$$

Donc si F est convexe, alors f est convexe aussi.

26. • Soit F une application de $\mathcal{S}_n(\mathbb{R})$ dans $\mathbb{R}$. Supposons qu'il existe une fonction symétrique $f : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que pour tout $A \in \mathcal{S}_n(\mathbb{R})$, $F(A) = f(\hat{\lambda}(A))$. Alors $F(A)$ ne dépend que du spectre de A . Comme A et QA^tQ ont le même spectre, on en déduit que $F(QA^tQ) = F(A)$ ainsi F est bien spectrale.
- D'après la question précédente, si on suppose que F est spectrale, la fonction f définie dans la question précédente est bien symétrique et vérifie : pour toute matrice symétrique A , $F(A) = f(\hat{\lambda}(A))$.
- Supposons qu'une autre fonction g vérifie cette condition. En prenant $A = D(x_1, \dots, x_n)$, on trouve que

$$f(x_1, \dots, x_n) = F(A) = g(\hat{\lambda}(A)) = g(x_1, \dots, x_n)$$

d'où l'unicité.

27. (a) Soit $A \in \mathcal{S}_n(\mathbb{R})$. On repart de la décomposition classique : $A = Q.\Lambda(A).^tQ$ où Q est orthogonale. Pour tout $i \in \llbracket 1, n \rrbracket$,

$$\begin{aligned} A_{i,i} &= \sum_{k=1}^n \sum_{l=1}^n Q_{i,k} \Lambda(A)_{k,l} (^tQ)_{l,i} \\ &= \sum_{k=1}^n Q_{i,k} \lambda_k Q_{i,k} \text{ en notant } \Lambda(A) = \text{Diag}(\lambda_1, \dots, \lambda_n) \text{ pour plus de simplicité} \\ &= \sum_{k=1}^n (Q_{i,k})^2 \lambda_k \\ &= (S\hat{\lambda}(A))_i \end{aligned}$$

où S vérifie $S_{i,j} = (Q_{i,j})^2$. D'après le II, S est orthostochastique donc bistochastique (d'après 7.(a)). On a donc bien $A = S\hat{\lambda}(A)$ où S est bistochastique.

- (b) On part de $C = A + B$. D'après la décomposition du 24., il existe Q orthogonale telle que

$$C = Q.\Lambda(C).^tQ \Leftrightarrow \Lambda(C) = ^tQ.C.Q$$

D'où

$$\Lambda(C) = ^tQ.A.Q + ^tQ.B.Q$$

donc

$$\hat{\lambda}(C) = \text{Diag}(\Lambda(C)) = \text{Diag}(^tQ.A.Q) + \text{Diag}(^tQ.B.Q)$$

et comme $^tQ.A.Q$ et $^tQ.B.Q$ sont symétriques et de même spectre que A et B : il existe deux matrices bistochastiques S_1 et S_2 telles que

$$\text{Diag}(^tQAQ) = S_1.\hat{\lambda}(A) \text{ et } \text{Diag}(^tQBQ) = S_2.\hat{\lambda}(B)$$

donc enfin

$$\hat{\lambda}(C) = S_1.\hat{\lambda}(A) + S_2.\hat{\lambda}(B)$$

(c) Soit A et B deux matrices symétriques et $t \in [0, 1]$. Alors

$$\begin{aligned} F(tA + (1-t)B) &= f(\hat{\lambda}(tA + (1-t)B)) \\ &= f(S_1\hat{\lambda}(tA) + S_2\hat{\lambda}((1-t)B))) \\ &= f(tS_1\hat{\lambda}(A) + (1-t)S_2\hat{\lambda}(B)) \\ &\leq tf(S_1\hat{\lambda}(A)) + (1-t)f(S_2\hat{\lambda}(B)) \text{ car } f \text{ est convexe} \\ &\leq tf(\hat{\lambda}(A)) + (1-t)f(\hat{\lambda}(B)) \text{ car } f \text{ est symétrique et convexe donc } S\text{-convexe} \\ &\leq tF(A) + (1-t)F(B) \end{aligned}$$

donc F est convexe.

- (d) Soit $A \in \mathcal{S}_n(\mathbb{R})$. Alors

$$\begin{aligned} F(DG(A)) &= f(\hat{\lambda}(DG(A))) \\ &= f(P_\sigma.\text{diag}(A)) \text{ les valeurs propres d'une matrice diagonale} \\ &\quad \text{sont ses coefficients, puis il faut les ranger dans l'ordre} \\ &= f(\text{diag}(A)) \text{ car } f \text{ est symétrique} \\ &= f(S.\hat{\lambda}(A)) \\ &\leq f(\hat{\lambda}(A)) \text{ car } f \text{ est } S\text{-convexe} \\ &\leq F(A) \end{aligned}$$

28. On note $\Sigma_m(A) = \sum_{k=1}^m \hat{\lambda}_k(A)$: il s'agit donc de la somme des m plus grandes valeurs propres.

- (a) C est somme de matrices symétriques, donc est encore symétrique.

Suite de la correction prise sur celle de Pierre-Paul TACHER

<https://bitsflip.io/doc/mathsl1.hec.essec.2026.pdf>

(un peu adapté à ma sauce !)

Notons

$$F = \Sigma_m : A \mapsto \sum_{i=1}^m \hat{\lambda}_i(A)$$

F est une fonction spectrale puisque pour toute matrice orthogonale, tQAQ et A ont le même spectre, donc la même somme partielle des valeurs propres (dans l'ordre).

On note U la matrice composée des colonnes $X_1, \dots, X_n : U = (X_1 | \dots | X_n)$. Comme ces colonnes $X_1, \dots, X_n$ forment une BON de $\mathcal{M}_{n,1}(\mathbb{R})$, la matrice U est orthogonale. On considère la matrice tUAU :

$$^tUAU = \begin{pmatrix} ^tX_1 \\ ^tX_2 \\ \vdots \\ ^tX_n \end{pmatrix} \cdot \begin{pmatrix} AX_1 & AX_2 & \cdots & AX_n \end{pmatrix}$$

et on voit donc que les coefficients diagonaux de cette matrice sont les produits scalaires $\langle X_k, AX_k \rangle$.

Enfin, la fonction symétrique f associée à la fonction F ci-dessus est simplement la fonction de $\mathbb{R}_n$ dans $\mathbb{R}$ qui somme les m termes les plus grands parmi $x_1, \dots, x_n$. Cette fonction est bien symétrique puisqu'elle ne dépend pas de l'ordre des termes x_i . Montrons qu'elle est bien convexe. Soit x et y deux éléments de $\mathbb{R}^n$ et σ_x, σ_y les deux permutations de $\llbracket 1, n \rrbracket$ qui permettent de remettre dans l'ordre les termes de x et de y . Soit $t \in [0, 1]$, et $z = tx + (1-t)y$. Notons encore σ_z la permutation qui remet dans l'ordre les termes de z . Alors

$$\begin{aligned} f(z) &= \sum_{k=1}^m z_{\sigma_z(k)} \\ &= \sum_{k=1}^m (tx + (1-t)y)_{\sigma_z(k)} \\ &= t \sum_{k=1}^m x_{\sigma_x(k)} + (1-t) \sum_{k=1}^m y_{\sigma_y(k)} \\ &\leq t.f(x) + (1-t).f(y) \end{aligned}$$

puisque $f(x)$ est la somme des m plus grands termes de x (idem pour $f(y)$).
 Cette fonction f est donc bien convexe. On peut alors appliquer la question 27.(d) et le Théorème de Davis :

$$F(DG({}^tUAU)) \leq F({}^tUAU)$$

D'une part,

$$F(DG({}^tUAU)) = \sum_{k=1}^m \langle X_k, AX_k \rangle$$

et d'autre part,

$$F({}^tUAU) = F(A) = \sum_{k=1}^m \hat{\lambda}_k(A)$$

car F est spectrale. On trouve bien que

$$\sum_{k=1}^m \langle X_k, AX_k \rangle \leq \Sigma_m(A)$$

Notons que nous n'avons pas démontré l'inégalité intermédiaire qui était demandée.

- (b) D'après ce que nous avons démontré précédemment et le théorème de Davis, la fonction Σ_m est convexe.

29. • Soit H comme dans l'énoncé. Alors

$$H = \alpha_n \cdot \Sigma_n(A) + (\alpha_{n-1} - \alpha_n) \Sigma_{n-1}(A) + (\alpha_{n-2} - \alpha_{n-1}) \cdot \Sigma_{n-2}(A) + \dots + (\alpha_1 - \alpha_2) \cdot \Sigma_1(A)$$

et tous les coefficients sont positifs d'après l'hypothèse $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$. Comme toutes les applications Σ_m sont spectrales, on en déduit immédiatement que H est spectrale. Enfin, une combinaison linéaire à coefficients positifs de fonctions convexes reste convexe (facile à montrer) donc H est convexe.

- Supposons tout d'abord B à valeurs propres positives. Alors on considère

$$H(M) = \sum_{k=1}^n \hat{\lambda}_k(B) \cdot \hat{\lambda}_k(M)$$

D'après le 29., cette application est spectrale et convexe.

Il existe une permutation P_σ telle que les coefficients diagonaux de la matrice $A' = P_\sigma \cdot A \cdot {}^tP_\sigma$ soient dans l'ordre décroissant. Comme H est spectrale, on a $H(A) = H(A')$ et on peut donc supposer que les coefficients de A sont dans l'ordre décroissant. En appliquant le 27.(d) à H on trouve alors que

$$H(DG(A)) \leq H(A)$$

donc

$$\sum_{k=1}^n \hat{\lambda}_k(B) \cdot A_{k,k} \leq \left\langle \hat{\lambda}(A), \hat{\lambda}(B) \right\rangle$$

- Notons que si B n'est pas à spectre positif, la matrice $B' = B - \hat{\lambda}_n \cdot I_n$ l'est. Notons $V' = \text{diag}(B')$. D'une part,

$$\left\langle \hat{U}, \hat{V}' \right\rangle = \left\langle \hat{U}, \hat{V} \right\rangle - \hat{\lambda}_n \left\langle \hat{U}, (1, \dots, 1) \right\rangle$$

et d'autre part,

$$\left\langle \hat{\lambda}(A), \hat{\lambda}(B - \hat{\lambda}_n \cdot I_n) \right\rangle = \left\langle \hat{\lambda}(A), \hat{\lambda}(B) \right\rangle - \hat{\lambda}_n \cdot \Sigma_n(A)$$

Comme $\left\langle \hat{U}, (1, \dots, 1) \right\rangle = \text{Tr}(A) = \Sigma_n(A)$, démontrer l'inégalité sur B ou sur B' revient bien au même.

- On a pas encore fini !

On considère cette fois l'application I telle que

$$I(M) = \sum_{k=1}^n \hat{\lambda}_k(M) \cdot A_{k,k}$$

On va supposer que A est à coefficients diagonaux positifs, on peut se ramener à ce cas via des manipulations semblables à celle faite précédemment. Cette application I est alors spectrale et convexe d'après le 29.