Correction – Mathématiques approfondies ESSEC/HEC 2026

Filière ECG – voie générale

Cette correction est rédigée dans un style volontairement rigoureux et dense, au niveau attendu d'un sujet de concours. Pour alléger certaines écritures, on note souvent $\mathfrak{S}_n$ l'ensemble des permutations de $\{1, \ldots, n\}$ et P_σ la matrice de permutation associée à σ.

Partie I – Préliminaires

1. Par définition, la j-ième colonne de P_σ est $e_{\sigma(j)}$. Or l'endomorphisme φ_σ associé à P_σ envoie précisément chaque vecteur de base sur la colonne correspondante. Donc

$$\forall j \in \{1, \ldots, n\}, \qquad \varphi_\sigma(e_j) = e_{\sigma(j)}.$$

2. Pour tout $j \in \{1, \ldots, n\}$,

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

Comme deux endomorphismes coïncidant sur la base canonique sont égaux, on obtient

$$P_\sigma P_\tau = P_{\sigma\circ\tau}.$$

En prenant $\tau = \sigma^{-1}$, il vient

$$P_\sigma P_{\sigma^{-1}} = P_{\text{id}} = I_n,$$

donc $P_\sigma^{-1} = P_{\sigma^{-1}}$. L'inverse d'une matrice de permutation est donc encore une matrice de permutation.

3. D'après la question précédente, $P_\sigma^{-1} = P_{\sigma^{-1}}$. D'autre part,

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

Ainsi,

$$P_\sigma^T = P_{\sigma^{-1}} = P_\sigma^{-1}.$$

Donc $P_\sigma^T P_\sigma = I_n$: toute matrice de permutation est orthogonale.

4. On raisonne par récurrence sur $n \geq 1$.

Initialisation. Pour $n = 1$, c'est immédiat.

Hérédité. Supposons le résultat vrai à l'ordre n et considérons $x = (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1}$. Choisissons i_0 tel que

$$x_{i_0} = \max(x_1, \ldots, x_{n+1}).$$

On place cet indice en première position. En supprimant la composante x_{i_0}, on obtient un n-uplet auquel on applique l'hypothèse de récurrence : il existe une permutation des indices restants qui classe les n composantes restantes dans l'ordre décroissant. En recollant avec i_0 en tête, on construit une permutation $\alpha \in \mathfrak{S}_{n+1}$ telle que

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

Le résultat est donc vrai à l'ordre $n + 1$.

La propriété est établie pour tout $n \geq 1$.

5. On veut montrer que si deux permutations $\alpha, \beta \in \mathfrak{S}_n$ vérifient

$$x_{\alpha(1)} \geq \cdots \geq x_{\alpha(n)} \qquad \text{et} \qquad x_{\beta(1)} \geq \cdots \geq x_{\beta(n)},$$

alors

$$\forall i \in \{1, \dots, n\}, \qquad x_{\alpha(i)} = x_{\beta(i)}.$$

On raisonne par récurrence sur n.

Pour $n = 1$, c'est immédiat. Supposons le résultat vrai à l'ordre $n-1$. Les deux nombres $x_{\alpha(1)}$ et $x_{\beta(1)}$ sont le maximum de l'ensemble $\{x_1, \dots, x_n\}$, donc

$$x_{\alpha(1)} = x_{\beta(1)}.$$

En retirant une occurrence de ce maximum, les suites

$$(x_{\alpha(2)}, \dots, x_{\alpha(n)}) \quad \text{et} \quad (x_{\beta(2)}, \dots, x_{\beta(n)})$$

sont encore deux réordonnancements décroissants du même multiensemble à $n-1$ éléments. L'hypothèse de récurrence donne alors l'égalité terme à terme sur les positions 2 à n.

Ainsi, le réordonnancement décroissant est bien défini de manière unique au niveau des valeurs. Le vecteur $\bar{x}$ n'est donc pas ambigu.

6. On complète la fonction Python par un tri par sélection. Une réponse correcte est par exemple :

```
def permutevecteur(X):
    n = len(X)
    alpha = np.arange(0, n, 1)
    Y = np.copy(X)
    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
```

À la fin de l'algorithme, les composantes de Y sont rangées par ordre décroissant et $Y[i] = X[\alpha[i]]$ pour tout i.

Partie II – Matrices bistochastiques

7 (a). Soit S orthostochastique. Il existe donc $Q = (q_{ij}) \in O_n$ tel que $s_{ij} = q_{ij}^2$. Alors :

— $s_{ij} \geq 0$ pour tous i, j ;

— pour tout i, comme la i-ième ligne de Q est de norme 1,

$$\sum_{j=1}^{n} s_{ij} = \sum_{j=1}^{n} q_{ij}^2 = 1;$$

— pour tout j, comme la j-ième colonne de Q est de norme 1,

$$\sum_{i=1}^{n} s_{ij} = \sum_{i=1}^{n} q_{ij}^2 = 1.$$

Donc S est bistochastique.

Si maintenant P est une matrice de permutation, alors chaque coefficient vaut 0 ou 1, chaque ligne contient exactement un 1 et chaque colonne aussi. Les trois conditions de bistochasticité sont donc satisfaites. Ainsi, toute matrice de permutation est bistochastique.

7 (b). Considérons

$$S = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Cette matrice est clairement bistochastique : ses coefficients sont positifs et la somme des coefficients de chaque ligne et de chaque colonne vaut 1.

Montrons qu'elle n'est pas orthostochastique. Supposons qu'il existe $Q = (q_{ij}) \in O_3$ tel que $s_{ij} = q_{ij}^2$. Alors la première ligne de Q est de la forme

$$\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, 0\right),$$

et la deuxième de la forme

$$\left(\pm\frac{1}{\sqrt{2}}, 0, \pm\frac{1}{\sqrt{2}}\right).$$

Leur produit scalaire vaut alors nécessairement $\pm\frac{1}{2}$, donc ne peut pas être nul. Les deux lignes ne sont donc pas orthogonales, contradiction avec $Q \in O_3$.

Ainsi, une matrice bistochastique n'est pas toujours orthostochastique.

8 (a). Dans $M_2(\mathbb{R})$, les seules permutations de $\{1, 2\}$ sont l'identité et la transposition $(1\ 2)$. Les matrices de permutation correspondantes sont

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{et} \qquad J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

8 (b). Soit

$$S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

bistochastique. Les sommes des lignes et des colonnes valent 1, donc

$$a + b = 1, \qquad c + d = 1, \qquad a + c = 1, \qquad b + d = 1.$$

On en déduit

$$b = 1 - a, \qquad c = 1 - a, \qquad d = a.$$

Comme tous les coefficients sont positifs, on a $a \in [0, 1]$. Ainsi,

$$S = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix} = aI_2 + (1-a)J.$$

La réciproque est immédiate : si $a \in [0, 1]$, alors $aI_2 + (1-a)J$ est bien bistochastique.

On a donc montré qu'une matrice de $M_2(\mathbb{R})$ est bistochastique si et seulement s'il existe $a \in [0, 1]$ et une matrice de permutation P tels que

$$S = aI_2 + (1-a)P.$$

9 (a). Le nombre de permutations de $\{1, 2, 3\}$ est

$$3! = 6.$$

9 (b). Les cinq matrices données sont précisément les cinq matrices de permutation non égales à l'identité. Les bijections associées sont :

$$\begin{aligned}
P_1 &: 1 \mapsto 1,\ 2 \mapsto 3,\ 3 \mapsto 2, \\
P_2 &: 1 \mapsto 3,\ 2 \mapsto 2,\ 3 \mapsto 1, \\
P_3 &: 1 \mapsto 2,\ 2 \mapsto 1,\ 3 \mapsto 3, \\
P_4 &: 1 \mapsto 3,\ 2 \mapsto 1,\ 3 \mapsto 2, \\
P_5 &: 1 \mapsto 2,\ 2 \mapsto 3,\ 3 \mapsto 1.
\end{aligned}$$

9 (c). Si $S = (s_{ij}) \in M_3(\mathbb{R})$ est bistochastique, les sommes des deux premières lignes donnent

$$s_{13} = 1 - s_{11} - s_{12}, \qquad s_{23} = 1 - s_{21} - s_{22}.$$

Les sommes des deux premières colonnes donnent

$$s_{31} = 1 - s_{11} - s_{21}, \qquad s_{32} = 1 - s_{12} - s_{22}.$$

Enfin, la somme de la troisième ligne (ou de la troisième colonne) donne

$$s_{33} = s_{11} + s_{12} + s_{21} + s_{22} - 1.$$

Donc S s'écrit bien

$$S = \begin{pmatrix} s_{11} & s_{12} & 1 - s_{11} - s_{12} \\ s_{21} & s_{22} & 1 - s_{21} - s_{22} \\ 1 - s_{11} - s_{21} & 1 - s_{12} - s_{22} & s_{33} \end{pmatrix}$$

avec

$$s_{33} = s_{11} + s_{12} + s_{21} + s_{22} - 1.$$

Les conditions nécessaires et suffisantes sur $s_{11}, s_{12}, s_{21}, s_{22}$ sont alors :

$$s_{11}, s_{12}, s_{21}, s_{22} \geq 0,$$

$$s_{11} + s_{12} \leq 1, \qquad s_{21} + s_{22} \leq 1,$$

$$s_{11} + s_{21} \leq 1, \qquad s_{12} + s_{22} \leq 1,$$

$$s_{11} + s_{12} + s_{21} + s_{22} \geq 1.$$

9 (d). Posons

$$\beta_0 = \min(s_{11}, s_{22}, s_{33}).$$

On considère

$$T = S - \beta_0 I_3.$$

Alors T est à coefficients positifs, et la somme des coefficients de chaque ligne et de chaque colonne vaut $1 - \beta_0$. De plus, l'un des coefficients diagonaux de T est nul.

En conjuguant éventuellement T par une matrice de permutation (ce qui ne fait que permuter les six matrices de permutation d'ordre 3), on peut supposer sans perte de généralité que

$$t_{33} = 0.$$

Dans ce cas, en reprenant la forme obtenue à la question précédente, on a

$$t_{11} + t_{12} + t_{21} + t_{22} = 1 - \beta_0.$$

Posons alors

$$\beta_1 = t_{11}, \qquad \beta_2 = t_{22}, \qquad \beta_3 = 0, \qquad \beta_4 = t_{12}, \qquad \beta_5 = t_{21}.$$

Tous ces coefficients sont positifs. Un calcul direct montre alors que

$$T = \beta_1 P_1 + \beta_2 P_2 + \beta_3 P_3 + \beta_4 P_4 + \beta_5 P_5.$$

On en déduit

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

9 (e). La question précédente montre que toute matrice bistochastique $S \in M_3(\mathbb{R})$ s'écrit comme combinaison linéaire à coefficients positifs des six matrices de permutation d'ordre 3.

De plus, en sommant tous les coefficients de l'égalité

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

on obtient

$$3 = 3\beta_0 + 3\sum_{i=1}^{5} \beta_i,$$

donc

$$\beta_0 + \beta_1 + \cdots + \beta_5 = 1.$$

Ainsi, S est une combinaison convexe de matrices de permutation. Le théorème de Birkhoff–Von Neumann est donc démontré pour $n = 3$.

10. Une écriture possible est :

```
def bistochastique(n, iter):
    A = rd.random([n, n])
    for k in range(iter):
        A = A / np.sum(A, axis=1).reshape((n, 1))
        A = A / np.sum(A, axis=0).reshape((1, n))
    return A
```

À chaque itération, on normalise d'abord les lignes puis les colonnes. Pour un nombre d'itérations assez grand, la matrice renvoyée est numériquement proche d'une matrice bistochastique.

Partie III – Fonctions symétriques et fonctions S-convexes

On rappelle que

$$H_n = \{y = (y_1, \ldots, y_n) \in \mathbb{R}^n \mid y_1 \geq y_2 \geq \cdots \geq y_n\}.$$

11. Si $X = (x_1, \ldots, x_n)^T$, alors

$$P_\sigma X = \sum_{j=1}^{n} x_j P_\sigma e_j = \sum_{j=1}^{n} x_j e_{\sigma(j)}.$$

Ainsi, la i-ième composante de $P_\sigma X$ est $x_{\sigma^{-1}(i)}$. Donc

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

Autrement dit, P_σ permute les coordonnées de X selon σ.

12. Si f est symétrique, par définition elle est invariante par permutation des coordonnées. Or $P_\sigma X$ est obtenu en permutant les coordonnées de X. Donc

$$\forall X \in M_{n,1}(\mathbb{R}),\ \forall P \in P_n, \qquad f(PX) = f(X).$$

Réciproquement, si cette propriété matricielle est vraie, alors pour tout $x \in \mathbb{R}^n$ et toute permutation σ, le vecteur $P_\sigma X$ est obtenu en permutant les coordonnées de x, donc f est symétrique. Les deux définitions sont bien équivalentes.

13. Supposons f symétrique. On définit $\tilde{f} : H_n \to \mathbb{R}$ par restriction :

$$\tilde{f}(y) = f(y), \qquad y \in H_n.$$

Soit maintenant $x \in \mathbb{R}^n$. Par définition de $\bar{x}$, il existe une permutation σ telle que $\bar{x}$ soit le réordonnancement décroissant des coordonnées de x, donc $x = P_\sigma \bar{x}$ pour une certaine permutation. Comme f est symétrique,

$$f(x) = f(\bar{x}) = \tilde{f}(\bar{x}).$$

Réciproquement, si l'on dispose d'une fonction $\tilde{f} : H_n \to \mathbb{R}$ telle que $f(x) = \tilde{f}(\bar{x})$ pour tout $x \in \mathbb{R}^n$, alors pour toute permutation σ,

$$\overline{P_\sigma x} = \bar{x},$$

puisque permuter les coordonnées ne change pas le multiensemble des composantes. Donc

$$f(P_\sigma x) = \tilde{f}(\overline{P_\sigma x}) = \tilde{f}(\bar{x}) = f(x).$$

Ainsi f est symétrique.

14.

— $f_1(x_1, \dots, x_n) = \sum_{k=1}^n x_k$ est symétrique, car une permutation des coordonnées ne change pas la somme.

— f_2 n'est pas symétrique : cette expression privilégie l'ordre des coordonnées. Par exemple, si l'on prend $x = (1, 0, 0, \dots, 0)$ et $y = (0, 1, 0, \dots, 0)$, qui se déduisent l'un de l'autre par permutation, on obtient deux valeurs différentes dès que l'expression contient de manière spéciale x_1 ou le produit $x_1 x_2$. Donc f_2 n'est pas symétrique.

— $f_3(x_1, \dots, x_n) = \max_{1 \leq i < j \leq n} |x_i - x_j|$ est symétrique, car une permutation des coordonnées ne change pas l'ensemble des différences $|x_i - x_j|$ lorsque (i, j) parcourt les couples distincts.

15. On peut écrire, pour tout $t \in \mathbb{R}$,

$$g_{i,\sigma,x}(t) = f\big(P_\sigma(x + te_i)\big) = f\big(P_\sigma x + te_{\sigma(i)}\big).$$

(a) Comme f est de classe $\mathcal{C}^1$, la fonction $g_{i,\sigma,x}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et, par dérivation d'une composée affine,

$$g'_{i,\sigma,x}(t) = \partial_{\sigma(i)} f\big(P_\sigma x + te_{\sigma(i)}\big).$$

(b) Puisque f est symétrique,

$$g_{i,\sigma,x}(t) = f\big(P_\sigma(x + te_i)\big) = f(x + te_i).$$

(c) En dérivant l'égalité précédente en $t = 0$, on obtient

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

16. Soit f une fonction S-convexe. D'après la question 7(a), toute matrice de permutation est bistochastique. Donc, pour tout $P \in P_n$ et tout X,

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

Appliquons maintenant cette inégalité à P^{-1} et au vecteur PX :

$$f(X) = f(P^{-1}PX) \leq f(PX).$$

Ainsi

$$f(PX) = f(X)$$

pour toute matrice de permutation P. D'après la question 12, f est symétrique.

17. On raisonne par récurrence sur m.

Pour $m = 2$, c'est exactement la définition de la convexité. Supposons la propriété vraie à l'ordre $m - 1$ et considérons des réels positifs $\alpha_1, \dots, \alpha_m$ de somme 1. Posons

$$\beta = \alpha_1 + \cdots + \alpha_{m-1} = 1 - \alpha_m.$$

Si $\beta = 0$, il n'y a rien à démontrer. Sinon,

$$\sum_{k=1}^{m} \alpha_k x_k = \beta \left(\sum_{k=1}^{m-1} \frac{\alpha_k}{\beta} x_k \right) + \alpha_m x_m.$$

Par convexité de g,

$$g\left(\sum_{k=1}^{m} \alpha_k x_k \right) \leq \beta\, g\left(\sum_{k=1}^{m-1} \frac{\alpha_k}{\beta} x_k \right) + \alpha_m g(x_m).$$

L'hypothèse de récurrence appliquée aux coefficients α_k/β donne alors

$$g\left(\sum_{k=1}^{m} \alpha_k x_k \right) \leq \sum_{k=1}^{m-1} \alpha_k g(x_k) + \alpha_m g(x_m) = \sum_{k=1}^{m} \alpha_k g(x_k).$$

La propriété est démontrée.

18. Soit $f : \mathbb{R}^n \to \mathbb{R}$ symétrique et convexe, et soit $B \in M_n(\mathbb{R})$ bistochastique. Le théorème de Birkhoff–Von Neumann admis dans cette partie donne

$$B = \sum_{r=1}^{N} \alpha_r P_r,$$

où $\alpha_r \geq 0$, $\sum_r \alpha_r = 1$, et les P_r sont des matrices de permutation.

Alors, pour tout X,

$$BX = \sum_{r=1}^{N} \alpha_r P_r X.$$

Par convexité de f et par la question 17,

$$f(BX) \leq \sum_{r=1}^{N} \alpha_r f(P_r X).$$

Or f est symétrique, donc $f(P_r X) = f(X)$ pour tout r. Ainsi

$$f(BX) \leq \sum_{r=1}^{N} \alpha_r f(X) = f(X).$$

Donc f est S-convexe.

19. On est dans le cas $n = 2$ et l'on pose

$$g(t) = f\big((1-t)x_1 + tx_2,\ (1-t)x_2 + tx_1\big).$$

(a) La fonction g est de classe $\mathcal{C}^1$ comme composée d'une application affine et d'une fonction $\mathcal{C}^1$. En posant

$$u(t) = (1-t)x_1 + tx_2, \qquad v(t) = (1-t)x_2 + tx_1,$$

on a $u'(t) = x_2 - x_1$ et $v'(t) = x_1 - x_2$, donc

$$g'(t) = (x_2 - x_1)\, \partial_1 f(u(t), v(t)) + (x_1 - x_2)\, \partial_2 f(u(t), v(t)).$$

(b) Pour $t \in [0,1]$, la matrice

$$B_t = \begin{pmatrix} 1-t & t \\ t & 1-t \end{pmatrix}$$

est bistochastique. Donc, par S-convexité,

$$g(t) = f(B_t x) \leq f(x) = g(0).$$

Ainsi $t = 0$ est un point de maximum de g sur $[0,1]$, donc

$$g'(0) \leq 0.$$

(c) En évaluant la formule de (a) en $t = 0$, on obtient

$$g'(0) = (x_2 - x_1)\partial_1 f(x_1, x_2) + (x_1 - x_2)\partial_2 f(x_1, x_2).$$

Comme $g'(0) \leq 0$, il vient

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

20. Fixons $i \neq j$ et $x \in \mathbb{R}^n$. Pour $t \in [0,1]$, considérons la matrice bistochastique

$$B_t = I_n + t(E_{ij} + E_{ji} - E_{ii} - E_{jj}).$$

Elle agit sur x en ne modifiant que les coordonnées i et j :

$$(B_t x)_i = (1-t)x_i + t x_j, \qquad (B_t x)_j = (1-t)x_j + t x_i,$$

et les autres coordonnées restent inchangées.

Posons

$$g(t) = f(B_t x).$$

Par S-convexité, on a $g(t) \leq g(0)$ pour tout $t \in [0,1]$, donc $g'(0) \leq 0$. Or la dérivation donne

$$g'(0) = (x_j - x_i)\partial_i f(x) + (x_i - x_j)\partial_j f(x).$$

Ainsi

$$(x_i - x_j)\big(\partial_i f(x) - \partial_j f(x)\big) \geq 0.$$

Le résultat est donc vrai pour tous i, j.

21. On pose

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} h(x_i).$$

Sens direct. Si h est convexe, alors f est convexe sur $\mathbb{R}^n$ comme somme de fonctions convexes de chacune des coordonnées. De plus, f est symétrique puisque la somme ne dépend pas de l'ordre des coordonnées. La question 18 donne alors que f est S-convexe.

Réciproque. Supposons f S-convexe. Par la question 20, pour tout $x \in \mathbb{R}^n$ et tout i, j,

$$(x_i - x_j)\big(h'(x_i) - h'(x_j)\big) \geq 0.$$

En particulier, en choisissant un vecteur dont seules les coordonnées i et j valent respectivement u et v, on obtient

$$(u - v)\big(h'(u) - h'(v)\big) \geq 0 \qquad (u, v \in \mathbb{R}).$$

Donc h' est croissante sur $\mathbb{R}$, ce qui équivaut à la convexité de h.

Ainsi, f est S-convexe si et seulement si h est convexe.

Partie IV – Fonctions spectrales. Théorèmes de Davis et de Fan

22. Pour $k \in \mathbb{N}^*$, on considère

$$F_k(A) = \mathrm{Tr}(A^k), \qquad A \in \mathcal{S}_n(\mathbb{R}).$$

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

$$F_k(QAQ^T) = \mathrm{Tr}\big((QAQ^T)^k\big) = \mathrm{Tr}(QA^kQ^T) = \mathrm{Tr}(A^kQQ^T) = \mathrm{Tr}(A^k) = F_k(A),$$

où l'on a utilisé l'invariance cyclique de la trace. Donc F_k est spectrale.

23. Soit $X = (x_1, \dots, x_n)^T$. Comme $D(X)$ est la matrice diagonale de diagonale X, on a pour tout j,

$$P_\sigma D(X) P_\sigma^T e_j = P_\sigma D(X) e_{\sigma^{-1}(j)} = P_\sigma\big(x_{\sigma^{-1}(j)} e_{\sigma^{-1}(j)}\big) = x_{\sigma^{-1}(j)} e_j.$$

La matrice $P_\sigma D(X) P_\sigma^T$ est donc diagonale, de diagonale

$$\big(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)}\big)^T = P_\sigma X.$$

Ainsi,

$$P_\sigma D(X) P_\sigma^T = D(P_\sigma X).$$

24. Comme $A \in \mathcal{S}_n(\mathbb{R})$, le théorème spectral assure l'existence d'une matrice orthogonale Q telle que

$$A = Q\, D\big(\lambda(A)\big)\, Q^T.$$

C'est exactement le résultat demandé.

25. On suppose ici que F est spectrale et l'on définit, pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$,

$$f(x) = F\big(D(x)\big).$$

(a) Soit $A \in \mathcal{S}_n(\mathbb{R})$. D'après la question 24, il existe $Q \in O_n$ tel que

$$A = QD(\lambda(A))Q^T.$$

Par spectralité de F,

$$F(A) = F\big(QD(\lambda(A))Q^T\big) = F\big(D(\lambda(A))\big).$$

(b) Soit $\sigma \in \mathfrak{S}_n$. Par la question 23,

$$D(P_\sigma x) = P_\sigma D(x) P_\sigma^T.$$

Donc, par spectralité de F,

$$f(P_\sigma x) = F\big(D(P_\sigma x)\big) = F\big(P_\sigma D(x) P_\sigma^T\big) = F\big(D(x)\big) = f(x).$$

Ainsi f est symétrique.

(c) Si F est convexe, alors pour tous $x, y \in \mathbb{R}^n$ et tout $t \in [0,1]$,

$$f((1-t)x + ty) = F\big(D((1-t)x + ty)\big) = F\big((1-t)D(x) + tD(y)\big) \le (1-t)f(x) + tf(y).$$

Donc f est convexe.

26. *Sens direct.* Supposons F spectrale. En définissant $f(x) = F(D(x))$, la question 25(b) montre que f est symétrique. De plus, par la question 25(a), pour tout $A \in \mathcal{S}_n(\mathbb{R})$,

$$F(A) = F\big(D(\lambda(A))\big) = f(\lambda(A)).$$

Sens réciproque. Supposons qu'il existe une fonction symétrique $f : \mathbb{R}^n \to \mathbb{R}$ telle que

$$F(A) = f(\lambda(A)) \qquad \text{pour tout } A \in \mathcal{S}_n(\mathbb{R}).$$

Si $Q \in O_n$, alors A et QAQ^T ont les mêmes valeurs propres, donc

$$\lambda(QAQ^T) = \lambda(A).$$

Par suite,

$$F(QAQ^T) = f(\lambda(QAQ^T)) = f(\lambda(A)) = F(A),$$

et F est spectrale.

Unicité. Si f et g sont deux fonctions symétriques telles que $F(A) = f(\lambda(A)) = g(\lambda(A))$, alors pour tout $x \in \mathbb{R}^n$,

$$f(x) = f(\bar{x}) = F(D(\bar{x})) = F(D(x)) = g(\bar{x}) = g(x),$$

car $D(x)$ et $D(\bar{x})$ sont orthogonalement semblables par une matrice de permutation. Donc $f = g$.

27. On suppose désormais que F est spectrale et que la fonction associée f est convexe.

(a) Soit $A \in \mathcal{S}_n(\mathbb{R})$. Écrivons, grâce au théorème spectral,

$$A = QD(\lambda(A))Q^T, \qquad Q = (q_{ij}) \in O_n.$$

Alors, pour tout i,

$$a_{ii} = \sum_{j=1}^{n} q_{ij}^2 \lambda_j(A).$$

Si l'on pose $S = (q_{ij}^2)$, la matrice S est orthostochastique donc bistochastique, et l'on a

$$\operatorname{diag}(A) = S\lambda(A).$$

(b) Soient $A, B \in \mathcal{S}_n(\mathbb{R})$ et $C = A + B$. Choisissons $Q \in O_n$ tel que

$$C = QD(\lambda(C))Q^T.$$

Alors

$$\lambda(C) = \operatorname{diag}(D(\lambda(C))) = \operatorname{diag}(Q^T CQ) = \operatorname{diag}(Q^T AQ) + \operatorname{diag}(Q^T BQ).$$

En appliquant (a) aux matrices symétriques $Q^T AQ$ et $Q^T BQ$, il existe deux matrices bistochastiques S_1, S_2 telles que

$$\operatorname{diag}(Q^T AQ) = S_1\lambda(A), \qquad \operatorname{diag}(Q^T BQ) = S_2\lambda(B).$$

Donc

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

(c) Soit $t \in [0, 1]$. En appliquant (b) à tA et $(1-t)B$, il existe S_1, S_2 bistochastiques telles que

$$\lambda(tA + (1-t)B) = tS_1\lambda(A) + (1-t)S_2\lambda(B).$$

Par convexité de f,

$$f\big(\lambda(tA + (1-t)B)\big) \le tf(S_1\lambda(A)) + (1-t)f(S_2\lambda(B)).$$

Or f est symétrique et convexe, donc S-convexe d'après la question 18. Ainsi

$$f(S_1\lambda(A)) \le f(\lambda(A)), \qquad f(S_2\lambda(B)) \le f(\lambda(B)).$$

Finalement,

$$F(tA + (1-t)B) = f\big(\lambda(tA + (1-t)B)\big) \le tF(A) + (1-t)F(B).$$

Donc F est convexe.

(d) Soit $A \in \mathcal{S}_n(\mathbb{R})$. D'après (a), il existe une matrice bistochastique S telle que

$$\operatorname{diag}(A) = S\lambda(A).$$

Comme $DG(A) = D(\operatorname{diag}(A))$, on a

$$F(DG(A)) = f(\lambda(DG(A))) = f\big(\overline{\operatorname{diag}(A)}\big) = f\big(\operatorname{diag}(A)\big),$$

puisque f est symétrique. Donc

$$F(DG(A)) = f(S\lambda(A)) \leq f(\lambda(A)) = F(A).$$

28. Pour $m \in \{1, \dots, n\}$, on pose

$$\Sigma_m(A) = \sum_{i=1}^{m} \lambda_i(A).$$

(a) Comme $(V_1, \dots, V_n)$ est une base orthonormée de vecteurs propres de A,

$$A = \sum_{i=1}^{n} \lambda_i(A) V_i V_i^T.$$

Pour tout k,

$$(X_k, AX_k) = \sum_{i=1}^{n} \lambda_i(A)(X_k, V_i)^2.$$

Comme $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$, on en déduit

$$(X_k, AX_k) \leq \lambda_k(A) + \sum_{i=1}^{k-1} (\lambda_i(A) - \lambda_k(A))(X_k, V_i)^2,$$

ce qui est bien l'inégalité (3).
En sommant pour $k = 1, \dots, m$, on obtient

$$\sum_{k=1}^{m} (X_k, AX_k) \leq \sum_{k=1}^{m} \lambda_k(A) + \sum_{k=1}^{m} \sum_{i=1}^{k-1} (\lambda_i(A) - \lambda_k(A))(X_k, V_i)^2.$$

Posons $a_{ki} = (X_k, V_i)^2$. Pour chaque i, on a $\sum_{k=1}^{n} a_{ki} = 1$, donc $\sum_{k=1}^{m} a_{ki} \leq 1$. En réorganisant la double somme, on obtient

$$\sum_{k=1}^{m} (X_k, AX_k) \leq \sum_{i=1}^{m} \lambda_i(A) = \Sigma_m(A).$$

C'est l'inégalité (4).

(b) En prenant $C = A + B$ et en utilisant les vecteurs propres $X_1, \dots, X_m$ associés aux m plus grandes valeurs propres de C, on a

$$\Sigma_m(C) = \sum_{k=1}^{m} (X_k, CX_k) = \sum_{k=1}^{m} (X_k, AX_k) + \sum_{k=1}^{m} (X_k, BX_k).$$

L'inégalité (4) appliquée à A puis à B donne

$$\Sigma_m(A + B) \leq \Sigma_m(A) + \Sigma_m(B).$$

Comme de plus, pour tout $t \geq 0$,

$$\Sigma_m(tA) = t\Sigma_m(A),$$

on en déduit que Σ_m est convexe sur $\mathcal{S}_n(\mathbb{R})$.

29. La fonction
$$H(A) = \sum_{i=1}^{n} \alpha_i \lambda_i(A) \qquad (\alpha_1 \geq \cdots \geq \alpha_n \geq 0)$$
est clairement spectrale puisqu'elle ne dépend que des valeurs propres ordonnées de A.

Posons $\alpha_{n+1} = 0$ et
$$\beta_m = \alpha_m - \alpha_{m+1} \geq 0.$$
Alors, par télescopage,
$$\sum_{m=1}^{n} \beta_m \Sigma_m(A) = \sum_{m=1}^{n} (\alpha_m - \alpha_{m+1}) \sum_{i=1}^{m} \lambda_i(A) = \sum_{i=1}^{n} \alpha_i \lambda_i(A) = H(A).$$
Comme chaque Σ_m est convexe et que les β_m sont positifs, H est une combinaison linéaire à coefficients positifs de fonctions convexes. Donc H est convexe.

30. Posons
$$U = \operatorname{diag}(A), \qquad V = \operatorname{diag}(B),$$
et notons $\bar{U}$ et $\bar{V}$ les vecteurs obtenus en réordonnant les composantes de U et V dans l'ordre décroissant.

Étape 1 : inégalité de réarrangement. On a
$$(U, V) \leq (\bar{U}, \bar{V}).$$
En effet, lorsque l'on échange dans un vecteur deux composantes mal ordonnées face à un vecteur décroissant, le produit scalaire ne diminue pas ; par une suite d'échanges adjacents, on obtient le vecteur décroissant $\bar{U}$, puis de même $\bar{V}$.

Étape 2 : comparaison avec les valeurs propres de B. Choisissons $c \geq 0$ tel que $\bar{u}_n + c \geq 0$, et considérons
$$H_c(M) = \sum_{i=1}^{n} (\bar{u}_i + c) \lambda_i(M).$$
Les coefficients sont décroissants et positifs, donc H_c est spectrale et convexe d'après la question 29. La question 27(d) appliquée à B donne
$$H_c(DG(B)) \leq H_c(B).$$
Or $\lambda(DG(B)) = \bar{V}$, donc
$$\sum_{i=1}^{n} (\bar{u}_i + c) \bar{v}_i \leq \sum_{i=1}^{n} (\bar{u}_i + c) \lambda_i(B).$$
Comme $\sum_i \bar{v}_i = \operatorname{Tr}(B) = \sum_i \lambda_i(B)$, le terme en c se simplifie et l'on obtient
$$(\bar{U}, \bar{V}) \leq (\bar{U}, \lambda(B)).$$

Étape 3 : comparaison avec les valeurs propres de A. Choisissons $d \geq 0$ tel que $\lambda_n(B) + d \geq 0$, et considérons
$$K_d(M) = \sum_{i=1}^{n} (\lambda_i(B) + d) \lambda_i(M).$$
Par la question 29, K_d est spectrale et convexe. En appliquant 27(d) à A, on obtient
$$K_d(DG(A)) \leq K_d(A).$$
Comme $\lambda(DG(A)) = \bar{U}$, il vient
$$\sum_{i=1}^{n} (\lambda_i(B) + d) \bar{u}_i \leq \sum_{i=1}^{n} (\lambda_i(B) + d) \lambda_i(A).$$

Les termes en d se simplifient grâce à l'égalité des traces, d'où

$$(\bar{U}, \lambda(B)) \leq (\lambda(A), \lambda(B)).$$

En combinant les trois étapes,

$$(U, V) \leq (\bar{U}, \bar{V}) \leq (\bar{U}, \lambda(B)) \leq (\lambda(A), \lambda(B)).$$

Ainsi,

$$(\mathrm{diag}(A), \mathrm{diag}(B)) \leq (\lambda(A), \lambda(B)).$$

C'est bien l'inégalité de Fan.

Fin de la correction.