# HEC Paris – ESSEC BS

Mathématiques I – Mathématiques approfondies

## Corrigé du sujet adapté – 2026

---

La numérotation des questions et des sous-questions est celle du sujet fourni. Tous les vecteurs sont réels. On conserve les notations $P_\sigma$, $D(X)$, $\mathrm{diag}(A)$, $\mathrm{DG}(A)$ et $\widehat{\lambda}(A)$ de l'énoncé. Les produits scalaires et les normes sont euclidiens. Pour une matrice orthogonale $Q$, on utilise $Q^{-1} = {}^tQ$ et donc $Q\,{}^tQ = {}^tQQ = I_n$.

# Partie I – Préliminaires

**Partie I – Matrices de permutation et réordonnement**

On rappelle que $(P_\sigma)_{i,j} = \delta_{i,\sigma(j)}$. Le vecteur $\widehat{x}$ est obtenu en ordonnant les composantes de $x$ dans l'ordre décroissant. Questions 1 à 6.

**Corrigé de la partie I.**

**1. Image des vecteurs de la base canonique.**

La $j$-ième colonne de $P_\sigma$ a pour $i$-ième coefficient $\delta_{i,\sigma(j)}$ : c'est donc le vecteur $e_{\sigma(j)}$. Or la $j$-ième colonne d'une matrice est l'image de $e_j$ par l'endomorphisme canoniquement associé. Ainsi

$$\boxed{\forall j \in \{1,\dots,n\}, \quad \varphi_\sigma(e_j) = e_{\sigma(j)}.}$$

**2. Produit et inverse des matrices de permutation.**

Pour tout $j$,

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

Les deux endomorphismes coïncident sur une base, donc leurs matrices sont égales. Avec $\tau = \sigma^{-1}$, les deux produits valent $I_n$.

$$\boxed{P_\sigma P_\tau = P_{\sigma\circ\tau}, \qquad P_\sigma^{-1} = P_{\sigma^{-1}} \in \mathscr{P}_n.}$$

**3. Orthogonalité.**

Les colonnes de $P_\sigma$ sont les vecteurs $e_{\sigma(1)},\dots,e_{\sigma(n)}$. Comme $\sigma$ est bijective, ils constituent une base orthonormée. Plus explicitement,

$$({}^tP_\sigma P_\sigma)_{j,k} = \langle e_{\sigma(j)}, e_{\sigma(k)}\rangle = \delta_{j,k}.$$

Ainsi ${}^tP_\sigma = P_\sigma^{-1}$, d'où également $P_\sigma\,{}^tP_\sigma = I_n$.

$$\boxed{P_\sigma \in O_n.}$$

**4. Existence d'un réordonnement décroissant.**

Pour $n = 1$, la permutation identité convient. Supposons la propriété vraie au rang $n$, et considérons $(x_1, \dots, x_{n+1})$. Par l'hypothèse de récurrence, il existe $\alpha \in \mathcal{P}_n$ tel que $x_{\alpha(1)} \geq \cdots \geq x_{\alpha(n)}$.

Posons $r = \#\{i \in \{1, \dots, n\} : x_{\alpha(i)} \geq x_{n+1}\}$. Ces indices forment les $r$ premiers indices de la liste ordonnée. Insérons $n+1$ entre $\alpha(r)$ et $\alpha(r+1)$, en début de liste si $r = 0$, en fin de liste si $r = n$. La liste

$$(\alpha(1), \dots, \alpha(r), n+1, \alpha(r+1), \dots, \alpha(n))$$

contient chaque indice exactement une fois et ordonne les valeurs dans l'ordre décroissant. La propriété est donc héréditaire.

$$\boxed{\forall n \geq 1,\ \forall x \in \mathbb{R}^n,\ \exists \alpha \in \mathcal{P}_n, \quad x_{\alpha(1)} \geq \cdots \geq x_{\alpha(n)}.}$$

**5. Unicité de la liste des valeurs ordonnées.**

Fixons $i$. Si $x_{\alpha(i)} < x_{\beta(i)}$, les $i$ valeurs $x_{\beta(1)}, \dots, x_{\beta(i)}$ sont toutes strictement supérieures à $x_{\alpha(i)}$. Mais dans la liste ordonnée selon $\alpha$, seules les $i-1$ premières valeurs peuvent être strictement supérieures à $x_{\alpha(i)}$. C'est impossible, puisque les deux listes contiennent les mêmes valeurs avec les mêmes répétitions. Donc $x_{\alpha(i)} \geq x_{\beta(i)}$. En échangeant $\alpha$ et $\beta$, on obtient l'inégalité inverse.

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

Les permutations peuvent différer lorsque des valeurs sont égales, mais $\widehat{x}$ est bien défini.

**6. Tri par sélection en Python.**

À chaque étape $i$, on cherche l'indice du maximum parmi les valeurs situées aux indices $i, \dots, n-1$, puis on échange cette valeur avec $Y[i]$. On applique le même échange aux indices enregistrés dans `alpha`. L'initialisation `np.array(X)` crée un nouveau tableau contenant les valeurs de `X`. Les indices Python commencent à 0 : $\alpha_{\text{math}}(j) =$ `alpha`$[j-1] + 1$.

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

Au début de l'étape $i$, les $i$ premières valeurs sont déjà les plus grandes, dans l'ordre décroissant, et $Y[k] = X[\texttt{alpha}[k]]$ pour tout $k$. L'échange conserve cette dernière relation et place la plus grande valeur restante en position $i$.

$$\boxed{Y[0] \geq \cdots \geq Y[n-1], \qquad Y[i] = X[\texttt{alpha}[i]].}$$

# Partie II – Matrices bistochastiques

**Partie II – Théorème de Birkhoff–Von Neumann en dimensions 2 et 3**

Une matrice bistochastique a des coefficients positifs ou nuls et chaque somme de ligne et de colonne vaut 1. Une matrice orthostochastique est de la forme $S_{i,j} = Q_{i,j}^2$ avec $Q \in O_n$. Questions 7 à 10.

**Corrigé de la partie II.**

**7. Exemples et contre-exemple.**

**a)** Si $S_{i,j} = Q_{i,j}^2$, ses coefficients sont positifs ou nuls. L'orthogonalité de $Q$ donne

$$\sum_{j=1}^{n} S_{i,j} = (Q\,{}^tQ)_{i,i} = 1, \qquad \sum_{i=1}^{n} S_{i,j} = ({}^tQQ)_{j,j} = 1.$$

Une matrice de permutation possède exactement un coefficient égal à 1 dans chaque ligne et dans chaque colonne, tous les autres étant nuls. Elle vérifie donc aussi les trois conditions.

Toute matrice orthostochastique est bistochastique.
Toute matrice de permutation est bistochastique.

**b)** Considérons

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

Ses coefficients sont positifs ou nuls et les sommes de ses lignes et colonnes valent 1. Si $S_{i,j} = Q_{i,j}^2$ avec $Q$ orthogonale, alors $Q_{1,1}$ et $Q_{2,1}$ valent chacun $\pm 1/\sqrt{2}$, tandis que $Q_{2,2} = Q_{1,3} = 0$. Le produit scalaire des deux premières lignes serait alors

$$Q_{1,1}Q_{2,1} + Q_{1,2}Q_{2,2} + Q_{1,3}Q_{2,3} = Q_{1,1}Q_{2,1} = \pm\frac{1}{2} \neq 0.$$

C'est contraire à leur orthogonalité.

$S$ est bistochastique et n'est pas orthostochastique.

**8. Dimension 2.**

**a)** Les deux permutations sont $(1,2)$ et $(2,1)$, d'où

$$\mathscr{P}_2 = \left\{ I_2, J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

**b)** Écrivons $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Les sommes de lignes et de colonnes imposent $b = c = 1 - a$ et $d = a$. La positivité équivaut à $0 \leq a \leq 1$. Ainsi

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

Réciproquement, pour $a \in [0,1]$, cette matrice est bistochastique. Plus généralement, si $P \in \mathscr{P}_2$, les coefficients de $aI_2 + (1-a)P$ sont positifs ou nuls et ses sommes de lignes et de colonnes sont $a + (1-a) = 1$.

$$\boxed{S \text{ bistochastique } \iff \exists a \in [0,1],\ \exists P \in \mathscr{P}_2, \quad S = aI_2 + (1-a)P.}$$

**9. Dimension 3.**

**a)** On choisit l'image de 1 de trois façons, celle de 2 de deux façons, puis celle de 3 est imposée. Il y a donc

$$\boxed{3 \times 2 \times 1 = 6 \text{ permutations de } \{1,2,3\}.}$$

**b)** La $j$-ième colonne de $P_\sigma$ est $e_{\sigma(j)}$. En lisant successivement les colonnes des matrices proposées, on obtient

$$\boxed{\begin{array}{lll} P_1 = P_{(1,3,2)}, & P_2 = P_{(3,2,1)}, & P_3 = P_{(2,1,3)}, \\ P_4 = P_{(3,1,2)}, & P_5 = P_{(2,3,1)}. & \end{array}}$$

Chaque triplet contient une seule fois chacun des indices $1, 2, 3$ : il définit une bijection. Avec $I_3$, ces matrices représentent les six permutations.

**c)** Posons $a = S_{1,1}$, $b = S_{1,2}$, $c = S_{2,1}$ et $d = S_{2,2}$. Les sommes des deux premières lignes donnent $S_{1,3} = 1 - a - b$ et $S_{2,3} = 1 - c - d$. Les sommes des deux premières colonnes donnent $S_{3,1} = 1 - a - c$ et $S_{3,2} = 1 - b - d$. Enfin, la troisième ligne impose

$$S_{3,3} = 1 - (1-a-c) - (1-b-d) = a + b + c + d - 1.$$

Ainsi

$$\boxed{S = \begin{pmatrix} a & b & 1-a-b \\ c & d & 1-c-d \\ 1-a-c & 1-b-d & a+b+c+d-1 \end{pmatrix}.}$$

Les conditions nécessaires sont

$$\boxed{\begin{array}{lll} a,b,c,d \geq 0, & a+b \leq 1, & c+d \leq 1, \\ a+c \leq 1, & b+d \leq 1, & a+b+c+d \geq 1. \end{array}}$$

Elles sont également suffisantes : tous les coefficients de la matrice précédente sont alors positifs ou nuls, et le calcul donne 1 pour chaque somme de ligne et de colonne.

**d)** Posons $e = a + b + c + d - 1 = S_{3,3}$ et $t = \min(a, d, e) = \beta_0$. Choisissons

$$\boxed{\begin{array}{lll} \beta_0 = t, & \beta_1 = a - t, & \beta_2 = d - t, \\ \beta_3 = e - t, & \beta_4 = b - e + t, & \beta_5 = c - e + t. \end{array}}$$

Les quatre premiers coefficients sont positifs ou nuls. Pour les deux derniers, distinguons les trois possibilités :

| | $\beta_4$ | $\beta_5$ |
|---|---|---|
| $t = a$ | $1-c-d$ | $1-b-d$ |
| $t = d$ | $1-a-c$ | $1-a-b$ |
| $t = e$ | $b$ | $c$ |

Tous les termes du tableau sont positifs ou nuls par la question précédente. En cas d'égalité entre plusieurs minima, chacune des lignes correspondantes fournit la même valeur.

Calculons maintenant la combinaison :

$$\beta_0 I_3 + \sum_{i=1}^{5} \beta_i P_i = \begin{pmatrix} \beta_0+\beta_1 & \beta_3+\beta_4 & \beta_2+\beta_5 \\ \beta_3+\beta_5 & \beta_0+\beta_2 & \beta_1+\beta_4 \\ \beta_2+\beta_4 & \beta_1+\beta_5 & \beta_0+\beta_3 \end{pmatrix}.$$

En substituant les valeurs choisies, les trois coefficients diagonaux valent $a, d, e$ ; les autres valent respectivement $b, 1-a-b, c, 1-c-d, 1-a-c, 1-b-d$. On retrouve donc $S$.

$$\boxed{S = \beta_0 I_3 + \sum_{i=1}^{5} \beta_i P_i, \qquad \forall i \in \{0,\dots,5\},\ \beta_i \geq 0.}$$

**e)** Les coefficients vérifient

$$\begin{aligned} \sum_{i=0}^{5} \beta_i &= t + (a-t) + (d-t) + (e-t) + (b-e+t) + (c-e+t) \\ &= a+b+c+d-e = 1. \end{aligned}$$

Avec la question 8, cela établit

$$\boxed{\text{Le théorème de Birkhoff–Von Neumann est vrai pour } n=2 \text{ et } n=3.}$$

Si l'on souhaite des coefficients strictement positifs, on supprime simplement les termes dont le coefficient est nul.

**10. Normalisations successives en Python.**

On simule une matrice initiale strictement positive. Pour chaque ligne, on calcule sa somme avant d'en modifier les coefficients ; on procède ensuite de même pour chaque colonne.

```
def bistochastique(n, iter):
    A = rd.random([n, n])
    while np.min(A) == 0:
        A = rd.random([n, n])
    for k in range(iter):
        for i in range(n):
            s = np.sum(A[i, :])
            for j in range(n):
                A[i, j] = A[i, j] / s
        for j in range(n):
            s = np.sum(A[:, j])
            for i in range(n):
                A[i, j] = A[i, j] / s
    return A
```

La boucle initiale renouvelle le tirage si le générateur a renvoyé une valeur nulle. Toutes les sommes utilisées comme diviseurs sont alors strictement positives.

Après chaque normalisation des lignes, leurs sommes valent 1 ; après celle des colonnes, les sommes des colonnes valent 1. Cette dernière opération peut modifier les sommes des lignes. La fonction effectue exactement les `iter` étapes prescrites.

Elle renvoie $A^{(2\,\texttt{iter})}$, considérée comme bistochastique pour `iter` assez grand.

Pour un nombre fini d'itérations, l'énoncé demande donc une approximation : les sommes des lignes ne sont pas nécessairement exactement égales à 1.

## Partie III – Fonctions symétriques et S-convexes

**Partie III – Symétrie, convexité et dérivées partielles**

Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est S-convexe lorsque $f(BX) \leq f(X)$ pour toute matrice bistochastique $B$ et tout vecteur $X$. Le théorème de Birkhoff–Von Neumann est désormais admis en dimension $n$. Questions 11 à 21.

**Corrigé de la partie III.**

**11. Action d'une matrice de permutation sur un vecteur.**

Pour tout $i$,

$$(P_\sigma X)_i = \sum_{j=1}^{n} \delta_{i,\sigma(j)} x_j.$$

L'unique terme éventuellement non nul correspond à $j = \sigma^{-1}(i)$. Ainsi

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

En particulier, c'est $P_{\sigma^{-1}} X$ qui représente $x_\sigma$.

**12. Caractérisation matricielle de la symétrie.**

Si $f$ est symétrique, la question 11 donne $f(P_\sigma X) = f(x_{\sigma^{-1}}) = f(x)$. Réciproquement, si $f(PX) = f(X)$ pour toute matrice de permutation, le choix $P = P_{\sigma^{-1}}$ donne $f(x_\sigma) = f(x)$.

$$\boxed{f \text{ symétrique } \iff \forall X,\ \forall P \in \mathscr{P}_n, \quad f(PX) = f(X).}$$

**13. Factorisation par le réordonnement décroissant.**

Supposons $f$ symétrique et définissons $\widehat{f} = f_{|H_n}$, la restriction de $f$ à $H_n$. Pour $x \in \mathbb{R}^n$, il existe une permutation $\alpha$ telle que $\widehat{x} = x_\alpha \in H_n$, donc $f(x) = f(\widehat{x}) = \widehat{f}(\widehat{x})$.

Réciproquement, si $f(x) = \widehat{f}(\widehat{x})$, les vecteurs $x$ et $x_\sigma$ ont les mêmes composantes avec les mêmes répétitions. La question 5 assure $\widehat{x_\sigma} = \widehat{x}$, d'où $f(x_\sigma) = \widehat{f}(\widehat{x_\sigma}) = f(x)$.

$$\boxed{f \text{ symétrique } \iff \exists \widehat{f} : H_n \to \mathbb{R}, \quad \forall x \in \mathbb{R}^n,\ f(x) = \widehat{f}(\widehat{x}).}$$

De plus, pour $y \in H_n$, $\widehat{y} = y$ ; nécessairement $\widehat{f}(y) = f(y)$, ce qui donne l'unicité de $\widehat{f}$.

**14. Étude des trois exemples.**

Pour $f_1$, la bijectivité de $\sigma$ permet de réindexer la somme :

$$f_1(x_\sigma) = \sum_{k=1}^{n} x_{\sigma(k)} = \sum_{j=1}^{n} x_j = f_1(x).$$

Pour $f_2$, prenons $x = (2, 0, \dots, 0)$ et la permutation qui échange 1 et 2 : alors $f_2(x) = 4$, tandis que $f_2(x_\sigma) = 2$. Enfin, quand $(i, j)$ décrit $\{1, \dots, n\}^2$, $(\sigma(i), \sigma(j))$ décrit le même ensemble, donc

$$f_3(x_\sigma) = \max_{1\le i,j\le n} |x_{\sigma(i)} - x_{\sigma(j)}| = \max_{1\le i,j\le n} |x_i - x_j| = f_3(x).$$

$$\boxed{f_1 \text{ et } f_3 \text{ sont symétriques ; } f_2 \text{ ne l'est pas.}}$$

**15. Dérivées partielles d'une fonction symétrique.**

Dans la sous-question a), $g_i$ désigne $g_{i,x}$.

**a)** Par définition de la dérivée partielle, pour tout $t$,

$$\lim_{s\to 0} \frac{f(x + (t+s)e_i) - f(x + te_i)}{s} = \partial_i f(x + te_i).$$

Ainsi $g_{i,x}$ est dérivable, et la continuité de $\partial_i f$ assure la continuité de sa dérivée.

$$\boxed{g_{i,x} \in \mathcal{C}^1(\mathbb{R}), \qquad g'_{i,x}(t) = \partial_i f(x + te_i).}$$

**b)** On vérifie composante par composante que

$$x_\sigma + te_i = (x + te_{\sigma(i)})_\sigma.$$

En effet, la composante d'indice $k$ à droite est $x_{\sigma(k)} + t\delta_{\sigma(k),\sigma(i)} = x_{\sigma(k)} + t\delta_{k,i}$. La symétrie de $f$ donne alors

$$g_{i,x_\sigma}(t) = f((x + te_{\sigma(i)})_\sigma) = f(x + te_{\sigma(i)}) = g_{\sigma(i),x}(t).$$

$$\boxed{g_{i,x_\sigma} = g_{\sigma(i),x}.}$$

**c)** Les dérivées en 0 de ces deux fonctions égales sont égales. La sous-question a) fournit

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

**16. Toute fonction S-convexe est symétrique.**

Soit $P \in \mathscr{P}_n$. D'après la question 7, $P$ est bistochastique, donc $f(PX) \le f(X)$. La question 2 montre que $P^{-1}$ est aussi une matrice de permutation, donc est bistochastique. En appliquant la S-convexité à $PX$ et $P^{-1}$,

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

Les deux inégalités donnent $f(PX) = f(X)$ ; la question 12 conclut.

$$\boxed{f \text{ S-convexe} \Longrightarrow f \text{ symétrique.}}$$

**17. Inégalité de convexité pour une combinaison finie.**

Procédons par récurrence sur $m \geq 1$. Pour $m = 1$, $\alpha_1 = 1$ et l'inégalité est une égalité. Supposons la propriété vraie pour $m$ termes et considérons $m+1$ termes. Posons $s = \sum_{k=1}^{m} \alpha_k = 1 - \alpha_{m+1}$. Si $s = 0$, les $m$ premiers coefficients sont nuls et le résultat est immédiat. Si $s > 0$, posons

$$z = \sum_{k=1}^{m} \frac{\alpha_k}{s} z_k.$$

Les coefficients de cette somme sont positifs ou nuls et de somme 1. Par convexité, puis par l'hypothèse de récurrence,

$$\begin{aligned} g\left(\sum_{k=1}^{m+1} \alpha_k z_k\right) &= g(sz + \alpha_{m+1} z_{m+1}) \\ &\leq s g(z) + \alpha_{m+1} g(z_{m+1}) \\ &\leq \sum_{k=1}^{m+1} \alpha_k g(z_k). \end{aligned}$$

La propriété est héréditaire, donc

$$\boxed{g\left(\sum_{k=1}^{m} \alpha_k z_k\right) \leq \sum_{k=1}^{m} \alpha_k g(z_k).}$$

**18. Une fonction symétrique et convexe est S-convexe.**

Soit $B$ bistochastique. Le théorème de Birkhoff–Von Neumann donne $B = \sum_{k=1}^{m} \alpha_k P_k$, où $\alpha_k \geq 0$, $\sum_k \alpha_k = 1$ et $P_k \in \mathscr{P}_n$. Par la question 17 et la symétrie de $f$,

$$f(BX) = f\left(\sum_{k=1}^{m} \alpha_k P_k X\right) \leq \sum_{k=1}^{m} \alpha_k f(P_k X) = \sum_{k=1}^{m} \alpha_k f(X) = f(X).$$

$$\boxed{f \text{ symétrique et convexe} \Longrightarrow f \text{ S-convexe.}}$$

**19. Condition différentielle en dimension 2.**

**a)** Posons $a = x_2 - x_1$, $b = x_1 - x_2$, $u(t) = x_1 + at$ et $v(t) = x_2 + bt$. Nous calculons la dérivée directement, par accroissements finis. Pour $s \neq 0$,

$$\begin{aligned} g(t+s) - g(t) = f(u(t) + as, v(t) + bs) - f(u(t), v(t) + bs) \\ + f(u(t), v(t) + bs) - f(u(t), v(t)). \end{aligned}$$

Pour $a \neq 0$, la fonction d'une variable $r \mapsto f(r, v(t) + bs)$ est de classe $\mathcal{C}^1$ sur le segment d'extrémités $u(t)$ et $u(t) + as$, et sa dérivée vaut $\partial_1 f(r, v(t) + bs)$. Le théorème des accroissements finis fournit un réel $\xi_s$ entre ces extrémités tel que le premier accroissement soit $as\partial_1 f(\xi_s, v(t) + bs)$. Pour $a = 0$, cet accroissement est nul. De même, le second vaut $bs\partial_2 f(u(t), \eta_s)$ pour un $\eta_s$ entre $v(t)$ et $v(t) + bs$, ou vaut 0 si $b = 0$.

Comme $\xi_s \to u(t)$ et $\eta_s \to v(t)$ lorsque $s \to 0$, la continuité des dérivées partielles donne

$$\boxed{\begin{aligned} g'(t) = (x_2 - x_1)\partial_1 f(u(t), v(t)) \\ + (x_1 - x_2)\partial_2 f(u(t), v(t)). \end{aligned}}$$

Cette expression est continue en $t$, donc $g \in \mathcal{C}^1(\mathbb{R})$.

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

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

est bistochastique. Avec $X = {}^t(x_1, x_2)$, la S-convexité donne $g(t) = f(B_t X) \leq f(X) = g(0)$. Pour $t \in ]0,1]$, on a donc $(g(t) - g(0))/t \leq 0$. En passant à la limite à droite en 0, et puisque $g$ est dérivable,

$$\boxed{\forall t \in [0,1], \quad g(t) \leq g(0), \qquad g'_d(0) = g'(0) \leq 0.}$$

**c)** D'après a),

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

La sous-question b) entraîne donc

$$\boxed{\forall x \in \mathbb{R}^2, \quad (x_1 - x_2)(\partial_1 f(x) - \partial_2 f(x)) \geq 0.}$$

**20. Condition différentielle en dimension quelconque.**

Pour $i = j$, l'expression demandée est nulle. Supposons $i \neq j$ et notons $T$ la matrice de permutation qui échange les coordonnées $i$ et $j$ et conserve les autres. Pour $t \in [0,1]$, la matrice $B_t = (1-t)I_n + tT$ est bistochastique. Posons $\varphi(t) = f(B_t X)$ pour tout $t \in \mathbb{R}$. Seules les coordonnées $i$ et $j$ varient : elles valent $x_i + t(x_j - x_i)$ et $x_j + t(x_i - x_j)$.

En décomposant l'accroissement en une modification de la coordonnée $i$, puis une modification de la coordonnée $j$, le calcul par accroissements finis de la question 19 a) donne

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

En détail, le premier quotient d'accroissement tend vers $(x_j - x_i)\partial_i f(x)$, le second vers $(x_i - x_j)\partial_j f(x)$ ; la continuité des dérivées partielles justifie ces deux limites, les autres coordonnées étant fixes. La S-convexité assure $\varphi(t) \leq \varphi(0)$ pour $t \in [0,1]$, donc $\varphi'_d(0) = \varphi'(0) \leq 0$. Ainsi

$$\boxed{\forall x \in \mathbb{R}^n, \ \forall i,j \in \{1,\dots,n\}, \quad (x_i - x_j)(\partial_i f(x) - \partial_j f(x)) \geq 0.}$$

**21. Fonctions de la forme $f(x) = \sum_{k=1}^{n} h(x_k)$.**

Supposons d'abord $h$ convexe. Pour $x, y \in \mathbb{R}^n$ et $t \in [0,1]$,

$$f(tx + (1-t)y) = \sum_{k=1}^{n} h(tx_k + (1-t)y_k) \leq t\sum_{k=1}^{n} h(x_k) + (1-t)\sum_{k=1}^{n} h(y_k).$$

Ainsi $f$ est convexe. Elle est aussi symétrique par réindexation de la somme, donc S-convexe par la question 18.

Réciproquement, supposons $f$ S-convexe. Comme $h \in \mathcal{C}^1(\mathbb{R})$, $f \in \mathcal{C}^1(\mathbb{R}^n)$ et $\partial_i f(x) = h'(x_i)$. En choisissant $x_1 = a$, $x_2 = b$ et les autres coordonnées nulles, la question 20 donne

$$(a-b)(h'(a) - h'(b)) \geq 0.$$

Pour $a < b$, on en déduit $h'(a) \leq h'(b)$ : $h'$ est croissante.

Montrons directement que cela implique la convexité de $h$. Soient $a < z < b$. La fonction $h$ est de classe $\mathcal{C}^1$ sur $[a, z]$ et $[z, b]$. Par le théorème des accroissements finis, il existe $c \in ]a, z[$ et $d \in ]z, b[$ tels que

$$\frac{h(z) - h(a)}{z - a} = h'(c) \leq h'(d) = \frac{h(b) - h(z)}{b - z}.$$

En multipliant par les dénominateurs strictement positifs puis en regroupant,

$$(b - a)h(z) \leq (b - z)h(a) + (z - a)h(b).$$

Pour $z = ta + (1 - t)b$ avec $t \in ]0, 1[$, cela donne $h(ta + (1 - t)b) \leq th(a) + (1 - t)h(b)$. Les cas $t = 0$, $t = 1$ ou $a = b$ donnent une égalité ; le cas $a > b$ s'obtient en échangeant $a$ et $b$.

$$\boxed{f \text{ est S-convexe } \iff h \text{ est convexe.}}$$

## Partie IV – Fonctions spectrales : Davis et Fan

**Partie IV – Convexité des fonctions des valeurs propres**

Pour $A \in \mathcal{S}_n(\mathbb{R})$, $\widehat{\lambda}(A)$ désigne la liste décroissante des valeurs diagonales d'une matrice diagonale semblable à $A$, et $\Lambda(A) = D(\widehat{\lambda}(A))$. Une fonction $F$ est spectrale si $F(QA{}^tQ) = F(A)$ pour tout $Q \in O_n$. Questions 22 à 30.

**Corrigé de la partie IV.**

**22. La fonction $F_k(A) = \mathrm{Tr}(A^k)$ est spectrale.**

Pour $Q \in O_n$, la matrice $QA{}^tQ$ est symétrique car ${}^t(QA{}^tQ) = Q{}^tA{}^tQ = QA{}^tQ$. Comme ${}^tQQ = I_n$, une récurrence sur $k \geq 1$ donne

$$(QA{}^tQ)^k = QA^k{}^tQ.$$

Pour deux matrices carrées $U, V$, le calcul des coefficients donne

$$\mathrm{Tr}(UV) = \sum_{i=1}^{n}\sum_{j=1}^{n} U_{i,j}V_{j,i} = \sum_{j=1}^{n}\sum_{i=1}^{n} V_{j,i}U_{i,j} = \mathrm{Tr}(VU).$$

Il s'agit ici de sommes finies. En appliquant cette identité,

$$F_k(QA{}^tQ) = \mathrm{Tr}(QA^k{}^tQ) = \mathrm{Tr}({}^tQQA^k) = \mathrm{Tr}(A^k) = F_k(A).$$

$$\boxed{\forall k \geq 1, \quad F_k \text{ est spectrale.}}$$

**23. Conjugaison d'une matrice diagonale par une permutation.**

Posons $X = {}^t(x_1, \ldots, x_n)$. Par les questions 1 et 2, ${}^tP_\sigma e_i = e_{\sigma^{-1}(i)}$. Ainsi

$$P_\sigma D(X){}^tP_\sigma e_i = P_\sigma\big(x_{\sigma^{-1}(i)}e_{\sigma^{-1}(i)}\big) = x_{\sigma^{-1}(i)}e_i.$$

La matrice à gauche est donc diagonale, avec pour $i$-ième coefficient diagonal $x_{\sigma^{-1}(i)} = (P_\sigma X)_i$.

$$\boxed{P_\sigma D(X) {}^tP_\sigma = D(P_\sigma X).}$$

**24. Diagonalisation orthogonale et ordre des valeurs propres.**

Le théorème spectral fournit une base orthonormée de vecteurs propres de $A$. On peut les réordonner de manière que les valeurs propres correspondantes soient décroissantes, par la question 4. Notons $V_1, \dots, V_n$ cette base et $Q$ la matrice dont les colonnes sont $V_1, \dots, V_n$. Alors $Q$ est orthogonale et $AQ = Q\Lambda(A)$. En multipliant à droite par ${}^tQ$,

$$\boxed{\exists Q \in O_n, \quad A = Q\Lambda(A) {}^tQ.}$$

La liste utilisée est indépendante de la base propre choisie. En effet, pour une matrice diagonalisable et une valeur propre $\lambda$, le nombre d'apparitions de $\lambda$ dans une diagonale semblable vaut $\dim \operatorname{Ker}(A - \lambda I_n)$ : pour une matrice diagonale, le noyau est engendré par les vecteurs canoniques correspondant à cette valeur, et la similitude conserve la dimension du noyau. Les valeurs et leurs répétitions sont donc déterminées par $A$ ; la question 5 assure ensuite l'unicité de leur liste décroissante.

**25. Fonction associée à une fonction spectrale.**

On suppose $F$ spectrale et on pose $f(x) = F(D(X))$.

**a)** La question 24 fournit $A = Q\Lambda(A) {}^tQ$ avec $Q$ orthogonale. L'invariance de $F$ donne immédiatement

$$\boxed{F(A) = F(\Lambda(A)).}$$

**b)** Pour $P \in \mathscr{P}_n$, les questions 23 et 3 donnent

$$f(PX) = F(D(PX)) = F(PD(X) {}^tP) = F(D(X)) = f(X).$$

La question 12 assure alors

$$\boxed{f \text{ est symétrique.}}$$

**c)** L'application $D$ est linéaire : $D(tX + (1-t)Y) = tD(X) + (1-t)D(Y)$. Pour $t \in [0,1]$, la convexité de $F$ donne donc

$$\begin{aligned} f(tX + (1-t)Y) &= F(tD(X) + (1-t)D(Y)) \\ &\leq tF(D(X)) + (1-t)F(D(Y)) \\ &= tf(X) + (1-t)f(Y). \end{aligned}$$

$$\boxed{F \text{ convexe} \Longrightarrow f \text{ convexe.}}$$

**26. Caractérisation et unicité des fonctions spectrales.**

Si $F$ est spectrale, définissons $f(x) = F(D(X))$. La question 25 montre que $f$ est symétrique et

$$F(A) = F(\Lambda(A)) = f(\widehat{\lambda}(A)).$$

Réciproquement, supposons qu'il existe une fonction symétrique $f$ telle que $F(A) = f(\widehat{\lambda}(A))$. Si $Q \in O_n$, les matrices $A$ et $QA{}^tQ$ sont semblables. Plus précisément, si $AV_i = \lambda_i V_i$, alors

$$(QA{}^tQ)(QV_i) = QAV_i = \lambda_i QV_i,$$

et les $QV_i$ constituent une base orthonormée. On retrouve donc la même liste de valeurs propres, avec les mêmes répétitions, puis le même réordonnement décroissant. Par conséquent $F(QA{}^tQ) = F(A)$.

$$\boxed{F \text{ spectrale } \iff \exists f \text{ symétrique}, \quad \forall A \in \mathcal{S}_n(\mathbb{R}),\ F(A) = f(\widehat{\lambda}(A)).}$$

Pour l'unicité, si $f$ représente $F$ ainsi, les valeurs diagonales de $D(X)$ sont les composantes de $X$, donc $\widehat{\lambda}(D(X)) = \widehat{X}$. Par symétrie,

$$F(D(X)) = f(\widehat{X}) = f(X).$$

Ainsi la seule fonction symétrique possible est

$$\boxed{\forall X \in \mathcal{M}_{n,1}(\mathbb{R}), \quad f(X) = F(D(X)).}$$

**27. Théorème de Davis.**

On suppose $F$ spectrale et sa fonction associée $f$ convexe. Par la question 25, $f$ est symétrique ; la question 18 assure donc qu'elle est S-convexe.

**a)** Écrivons $A = Q\Lambda(A){}^tQ$ par la question 24. Le calcul du $i$-ième coefficient diagonal donne

$$A_{i,i} = \sum_{j=1}^{n} Q_{i,j}\widehat{\lambda}_j(A)Q_{i,j} = \sum_{j=1}^{n} Q_{i,j}^2\widehat{\lambda}_j(A).$$

Posons $S_{i,j} = Q_{i,j}^2$. Cette matrice est orthostochastique, donc bistochastique par la question 7 a).

$$\boxed{\mathrm{diag}(A) = S\widehat{\lambda}(A), \qquad S \text{ bistochastique.}}$$

**b)** Posons $C = A + B$ et choisissons $R \in O_n$ tel que $C = R\Lambda(C){}^tR$. Alors

$$\Lambda(C) = {}^tRAR + {}^tRBR.$$

Les matrices ${}^tRAR$ et ${}^tRBR$ sont symétriques et ont respectivement les mêmes listes ordonnées de valeurs propres que $A$ et $B$. La sous-question a), appliquée séparément à ces deux matrices, fournit deux matrices bistochastiques $S_1, S_2$ telles que

$$\mathrm{diag}({}^tRAR) = S_1\widehat{\lambda}(A), \qquad \mathrm{diag}({}^tRBR) = S_2\widehat{\lambda}(B).$$

L'application diag étant linéaire,

$$\boxed{\widehat{\lambda}(A+B) = S_1\widehat{\lambda}(A) + S_2\widehat{\lambda}(B).}$$

**c)** Soient $A, B \in \mathcal{S}_n(\mathbb{R})$ et $t \in ]0,1[$. Appliquons b) aux matrices $tA$ et $(1-t)B$. Comme les deux facteurs sont positifs, ils conservent l'ordre des valeurs propres ; on a

$$\widehat{\lambda}(tA) = t\widehat{\lambda}(A), \qquad \widehat{\lambda}((1-t)B) = (1-t)\widehat{\lambda}(B).$$

Il existe donc des matrices bistochastiques $S_1, S_2$ telles que

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

Par convexité de $f$, puis par S-convexité,

$$\begin{aligned} F(tA + (1-t)B) &= f(tS_1\widehat{\lambda}(A) + (1-t)S_2\widehat{\lambda}(B)) \\ &\leq tf(S_1\widehat{\lambda}(A)) + (1-t)f(S_2\widehat{\lambda}(B)) \\ &\leq tf(\widehat{\lambda}(A)) + (1-t)f(\widehat{\lambda}(B)) \\ &= tF(A) + (1-t)F(B). \end{aligned}$$

Les cas $t = 0$ et $t = 1$ sont des égalités.

$$\boxed{F \text{ est convexe.}}$$

Avec 25 c), on obtient l'équivalence entre la convexité de $F$ et celle de sa fonction associée $f$, lorsque $F$ est spectrale.

**d)** Avec la matrice $S$ obtenue en a), et puisque $\mathrm{DG}(A) = D(\mathrm{diag}(A))$, on a

$$F(\mathrm{DG}(A)) = f(\mathrm{diag}(A)) = f(S\widehat{\lambda}(A)) \leq f(\widehat{\lambda}(A)) = F(A).$$

$$\boxed{F(\mathrm{DG}(A)) \leq F(A).}$$

**28. Sommes des plus grandes valeurs propres.**

**a)** Pour alléger les calculs, posons $\lambda_i = \widehat{\lambda}_i(A)$ et $c_{k,i} = \langle X_k, V_i\rangle$. Puisque $(V_i)$ est une base orthonormée,

$$X_k = \sum_{i=1}^{n} c_{k,i}V_i, \qquad \sum_{i=1}^{n} c_{k,i}^2 = \|X_k\|^2 = 1.$$

Comme $AV_i = \lambda_i V_i$, la bilinéarité et l'orthonormalité donnent

$$\langle X_k, AX_k\rangle = \sum_{i=1}^{n} \lambda_i c_{k,i}^2 = \lambda_m + \sum_{i=1}^{n} (\lambda_i - \lambda_m)c_{k,i}^2.$$

Pour $i \geq m$, $\lambda_i - \lambda_m \leq 0$. En supprimant ces termes, on majore la somme, d'où

$$\boxed{\langle X_k, AX_k\rangle \leq \lambda_m + \sum_{i=1}^{m-1} (\lambda_i - \lambda_m)\langle X_k, V_i\rangle^2.}$$

C'est l'inégalité (3), valable pour chaque $k \in \{1, \ldots, n\}$.
En sommant pour $k = 1, \ldots, m$, on obtient

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

Puisque $(X_1, \ldots, X_n)$ est une base orthonormée,

$$\sum_{k=1}^{m} \langle X_k, V_i\rangle^2 \leq \sum_{k=1}^{n} \langle X_k, V_i\rangle^2 = \|V_i\|^2 = 1.$$

Pour $i < m$, le coefficient $\lambda_i - \lambda_m$ est positif ou nul. Par conséquent,

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

Ainsi

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

Pour $m = 1$, les sommes de 1 à $m-1$ sont vides et valent 0 ; les deux démonstrations restent valables. Elles restent également valables pour $m = n$.

**b)** Comme les $X_k$ sont des vecteurs propres unitaires de $C = A + B$,

$$\Sigma_m(A + B) = \sum_{k=1}^{m} \langle X_k, (A + B)X_k \rangle.$$

La première somme portant sur $A$ est majorée par $\Sigma_m(A)$, d'après a). La même démonstration, dans une base orthonormée propre de $B$, majore la seconde par $\Sigma_m(B)$. Ainsi

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

Pour $s \geq 0$, les valeurs propres de $sA$ sont les $s\lambda_i$ avec le même ordre si $s > 0$, et toutes nulles si $s = 0$. Donc $\Sigma_m(sA) = s\Sigma_m(A)$. Pour $t \in [0, 1]$, la sous-additivité précédente appliquée à $tA$ et $(1 - t)B$ donne

$$\boxed{\Sigma_m(tA + (1 - t)B) \leq t\Sigma_m(A) + (1 - t)\Sigma_m(B).}$$

$$\boxed{\forall m \in \{1, \ldots, n\}, \quad \Sigma_m \text{ est convexe.}}$$

**29. Combinaison pondérée des valeurs propres ordonnées.**

La liste $\widehat{\lambda}(A)$ est inchangée sous conjugaison orthogonale, comme établi à la question 26. Donc

$$H(QA^tQ) = \sum_{k=1}^{n} \alpha_k \widehat{\lambda}_k(A) = H(A),$$

ce qui prouve que $H$ est spectrale.

Pour la convexité, l'identité utile est

$$\boxed{H(A) = \alpha_n \Sigma_n(A) + \sum_{m=1}^{n-1} (\alpha_m - \alpha_{m+1})\Sigma_m(A).}$$

En effet, le coefficient de $\widehat{\lambda}_k(A)$ à droite vaut

$$\alpha_n + \sum_{m=k}^{n-1} (\alpha_m - \alpha_{m+1}) = \alpha_k,$$

avec une somme vide pour $k = n$. Par hypothèse, $\alpha_n \geq 0$ et $\alpha_m - \alpha_{m+1} \geq 0$. Chaque $\Sigma_m$ est convexe par la question 28. En multipliant ses inégalités de convexité par ces coefficients puis en les additionnant, on obtient celle de $H$.

$$\boxed{H \text{ est spectrale et convexe.}}$$

**30. Inégalité de Fan.**

Posons $U = \mathrm{diag}(A)$, $V = \mathrm{diag}(B)$, $L = \widehat{\lambda}(A)$ et $M = \widehat{\lambda}(B)$. Nous établissons deux inégalités successives.

**Première étape : sommes partielles des diagonales ordonnées.** Par la question 27 a), il existe $S$ bistochastique telle que $U = SL$. Choisissons $P \in \mathscr{P}_n$ tel que $\widehat{U} = PU$. La matrice $T = PS$ est obtenue en permutant les lignes de $S$ ; elle est donc bistochastique, et $\widehat{U} = TL$. Pour $m \in \{1, \dots, n\}$, posons

$$c_j = \sum_{i=1}^{m} T_{i,j}.$$

La positivité et les sommes de colonnes de $T$ donnent $0 \leq c_j \leq 1$, tandis que les sommes des $m$ premières lignes donnent $\sum_{j=1}^{n} c_j = m$. Ainsi

$$\begin{aligned}
\sum_{i=1}^{m} \widehat{U}_i &= \sum_{j=1}^{n} c_j L_j \\
&= mL_m + \sum_{j=1}^{n} c_j (L_j - L_m) \\
&\leq mL_m + \sum_{j=1}^{m-1} c_j (L_j - L_m) \\
&\leq mL_m + \sum_{j=1}^{m-1} (L_j - L_m) = \sum_{j=1}^{m} L_j.
\end{aligned}$$

La première majoration supprime des termes négatifs ou nuls pour $j \geq m$ ; la seconde utilise $c_j \leq 1$ pour $j < m$, où $L_j - L_m \geq 0$. Pour $m = n$, les sommes de colonnes impliquent $c_j = 1$ pour tout $j$, donc il y a égalité des sommes totales. On obtient de même les relations pour $V$ et $M$ :

$$\boxed{\begin{gathered}
\forall m \in \{1, \dots, n\}, \quad \sum_{i=1}^{m} \widehat{U}_i \leq \sum_{i=1}^{m} L_i, \qquad \sum_{i=1}^{m} \widehat{V}_i \leq \sum_{i=1}^{m} M_i, \\
\sum_{i=1}^{n} \widehat{U}_i = \sum_{i=1}^{n} L_i, \qquad \sum_{i=1}^{n} \widehat{V}_i = \sum_{i=1}^{n} M_i.
\end{gathered}}$$

**Deuxième étape : une identité pour des poids décroissants.** Soient $a, b, w \in \mathbb{R}^n$ tels que

$$\sum_{i=1}^{m} a_i \leq \sum_{i=1}^{m} b_i \quad (1 \leq m < n), \qquad \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i,$$

et $w_1 \geq \cdots \geq w_n$. En posant $s_m = \sum_{i=1}^{m} (a_i - b_i)$, on a $s_m \leq 0$ et $s_n = 0$. Une réorganisation de sommes finies donne

$$\sum_{i=1}^{n} w_i (a_i - b_i) = \sum_{m=1}^{n-1} (w_m - w_{m+1}) s_m + w_n s_n \leq 0.$$

Pour vérifier l'identité, le coefficient de $a_i - b_i$ à droite est $\sum_{m=i}^{n-1} (w_m - w_{m+1}) + w_n = w_i$. On a donc établi

$$\boxed{\sum_{i=1}^{n} w_i a_i \leq \sum_{i=1}^{n} w_i b_i.}$$

Cette implication n'exige pas que les poids soient positifs : l'égalité des sommes totales annule le terme $w_n s_n$.

**Troisième étape : application deux fois de cette implication.** D'abord, prenons $a = \widehat{U}$, $b = L$ et $w = \widehat{V}$. Les hypothèses sur les sommes partielles sont vérifiées par la première étape, et les poids sont décroissants par définition. On obtient

$$\langle \widehat{U}, \widehat{V} \rangle \leq \langle L, \widehat{V} \rangle.$$

Ensuite, prenons $a = \widehat{V}$, $b = M$ et $w = L$, qui est également décroissant. Il vient

$$\langle L, \widehat{V} \rangle \leq \langle L, M \rangle.$$

En réunissant les deux inégalités,

$$\boxed{\langle \widehat{U}, \widehat{V} \rangle \leq \langle \widehat{\lambda}(A), \widehat{\lambda}(B) \rangle.}$$

C'est l'inégalité de Fan demandée, sans hypothèse de positivité sur les diagonales ou les valeurs propres.