



PRÉPASIMPLE

prepasimple.fr

HEC / ESSEC

Mathématiques

2026

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Partie I : préliminaires

On conserve les notations du sujet. Une permutation σ envoie la colonne e_j sur $e_{\sigma(j)}$; en revanche, l'action de P_σ sur les coordonnées d'un vecteur fait intervenir σ^{-1} . Cette distinction sera essentielle aux questions 11, 15 et 23.

1. L'image de e_j par l'endomorphisme de matrice P_σ est la j -ème colonne de cette matrice. Son coefficient de ligne i vaut $\delta_{i,\sigma(j)}$. Cette colonne possède donc un seul coefficient égal à 1, situé à la ligne $\sigma(j)$, et tous ses autres coefficients sont nuls.

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

2. Méthode 1 — Action sur la base canonique

Pour chaque j , la question 1 donne

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

Deux applications linéaires qui coïncident sur une base sont égales. Ainsi, $P_\sigma P_\tau = P_{\sigma \circ \tau}$.

La permutation identité est associée à I_n . En prenant successivement $\tau = \sigma^{-1}$ et en échangeant les deux permutations, on obtient $P_\sigma P_{\sigma^{-1}} = P_{\sigma^{-1}} P_\sigma = I_n$. Donc

$$P_\sigma^{-1} = P_{\sigma^{-1}} \in P_n.$$

Méthode 2 — Calcul des coefficients

La formule du produit matriciel donne

$$(P_\sigma P_\tau)_{i,j} = \sum_{k=1}^n \delta_{i,\sigma(k)} \delta_{k,\tau(j)} = \delta_{i,\sigma(\tau(j))}.$$

Le seul terme éventuellement non nul est celui d'indice $k = \tau(j)$. Le coefficient obtenu est celui de $P_{\sigma \circ \tau}$. En appliquant cette identité à σ et σ^{-1} , on retrouve l'inverse indiqué ci-dessus.

La première méthode évite les indices ; la seconde permet de contrôler explicitement l'ordre de composition.

3. Méthode 1 — Colonnes orthonormées

Les colonnes de P_σ sont $e_{\sigma(1)}, \dots, e_{\sigma(n)}$. Elles constituent une base orthonormée, puisque σ est bijective. Le coefficient (i, j) de ${}^tP_\sigma P_\sigma$ est leur produit scalaire :

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

Donc ${}^tP_\sigma P_\sigma = I_n$. La question 2 assure que P_σ est inversible ; cette égalité donne ${}^tP_\sigma = P_\sigma^{-1}$, puis $P_\sigma {}^tP_\sigma = I_n$. Il est donc orthogonal au sens de la définition du sujet.

Méthode 2 — Transposée et permutation inverse

On a

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

La question 2 donne alors $P_\sigma {}^tP_\sigma = P_\sigma P_{\sigma^{-1}} = I_n$. Cette méthode établit directement la formule utile ${}^tP_\sigma = P_{\sigma^{-1}}$.

4. La méthode demandée est une récurrence sur la dimension.

Pour $n = 1$, la permutation identité convient.

Supposons la propriété vraie en dimension n , et considérons $x_1, \dots, x_{n+1}$. Une famille finie non vide de réels possède un maximum. Choisissons j tel que $x_j = \max_{1 \leq k \leq n+1} x_k$.

Énumérons les autres indices sous la forme $r_1, \dots, r_n$. L'hypothèse de récurrence appliquée à $(x_{r_1}, \dots, x_{r_n})$ donne une permutation γ de $\{1, \dots, n\}$ telle que $x_{r_{\gamma(1)}} \geq \dots \geq x_{r_{\gamma(n)}}$.

Définissons $\alpha(1) = j$ et $\alpha(k+1) = r_{\gamma(k)}$ pour $1 \leq k \leq n$. Chaque indice apparaît une fois, donc α est une permutation. Le choix du maximum donne

$$x_{\alpha(1)} \geq x_{\alpha(2)} \geq \dots \geq x_{\alpha(n+1)}.$$

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

5. Méthode 1 — Comparaison du nombre de valeurs dépassant un seuil

Posons $a_i = x_{\alpha(i)}$ et $b_i = x_{\beta(i)}$. Ces deux listes contiennent les mêmes valeurs, avec les mêmes nombres d'occurrences.

S'il existait i tel que $a_i > b_i$, les i premiers termes de la liste a seraient strictement supérieurs à b_i . Dans la liste b , seuls les $i - 1$ premiers termes pourraient être strictement supérieurs à b_i , puisque b est décroissante. Cela contredit l'égalité des nombres d'occurrences au-dessus de ce seuil.

On ne peut donc avoir $a_i > b_i$. En échangeant a et b , on exclut $b_i > a_i$. Ainsi $a_i = b_i$ pour tout i .

Méthode 2 — Retrait successif des maxima

Les premiers termes sont tous deux égaux à $\max_k x_k$. Retirons une occurrence de cette valeur dans chacune des listes. Les listes restantes sont décroissantes et contiennent encore les mêmes valeurs avec les mêmes nombres d'occurrences. Leurs premiers termes sont donc égaux, ce qui donne $a_2 = b_2$.

En répétant ce raisonnement jusqu'à épuisement des listes, on obtient $a_i = b_i$ pour tout i . Le retrait d'une seule occurrence traite aussi les valeurs répétées.

La première méthode compare directement un rang quelconque ; la seconde explique l'unicité du tri par extraction des maxima. Les permutations de tri peuvent être différentes, mais le vecteur trié est unique.

6. On effectue un tri par sélection : à l'étape i , on cherche le maximum parmi les positions de i à $n-1$, puis on le place à la position i .

Python

```
import numpy as np

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
```

À tout instant, les échanges simultanés préservent $Y[j] == X[\alpha[j]]$. Le tableau α , initialement constitué des indices de 0 à $n-1$, reste une permutation de ces indices.

Après l'étape i , les positions de 0 à i contiennent les $i+1$ plus grandes valeurs dans l'ordre décroissant. Une récurrence sur les étapes justifie cet invariant : le suffixe contient les valeurs restantes, et la boucle interne en sélectionne le maximum. À la fin, Y est donc décroissant et contient exactement les valeurs de X .

La permutation mathématique de la question 4 s'obtient par $\alpha_{\text{math}}(j) = \alpha[j-1] + 1$, pour $1 \leq j \leq n$. Le programme utilise la convention Python commençant à zéro.

Code exécuté et contrôlé sur des tableaux avec valeurs négatives, répétées et déjà triées.

Par exemple, pour $X = \text{np.array}([2., -1., 2., 0.])$, il renvoie $Y = [2., 2., 0., -1.]$ et $\alpha = [0, 2, 3, 1]$. Le tableau d'entrée est conservé grâce à `copy()`.

Partie II : matrices bistochastiques

Dans cette partie, « positif » signifie « positif ou nul » pour les coefficients des combinaisons convexes.

7. (a) Commençons par une matrice orthostochastique S , définie par $S_{i,j} = Q_{i,j}^2$ avec $Q^t Q = I_n$. Ses coefficients sont non négatifs.

L'égalité $Q^t Q = I_n$ implique que l'application linéaire associée à Q est surjective : $Z = Q(^t Q Z)$ pour tout Z . En dimension finie, elle est donc inversible, et $^t Q = Q^{-1}$; ainsi $^t Q Q = I_n$ également.

Les coefficients diagonaux de ces deux produits donnent

$$\sum_{j=1}^n Q_{i,j}^2 = 1, \quad \sum_{i=1}^n Q_{i,j}^2 = 1.$$

Les sommes des lignes et des colonnes de S valent donc 1 : S est bistochastique.

Une matrice de permutation possède exactement un coefficient égal à 1 par colonne, et aussi par ligne puisque la permutation est bijective. Tous les autres coefficients valent 0. Elle est donc bistochastique.

On peut également traiter les matrices de permutation par le résultat précédent : d'après la question 3, P est orthogonale et $P_{i,j}^2 = P_{i,j}$, donc P est orthostochastique, puis bistochastique.

7. (b) Prenons

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

Ses coefficients sont non négatifs et chaque ligne et colonne contient deux coefficients $1/2$, donc S est bistochastique.

Supposons $S_{i,j} = Q_{i,j}^2$ pour une matrice orthogonale Q . Les première et deuxième lignes auraient les formes

$$(\varepsilon_1/\sqrt{2}, \varepsilon_2/\sqrt{2}, 0), \quad (\varepsilon_3/\sqrt{2}, 0, \varepsilon_4/\sqrt{2}),$$

avec $\varepsilon_1, \dots, \varepsilon_4 \in \{-1, 1\}$. Leur produit scalaire serait $\varepsilon_1\varepsilon_3/2$, qui n'est jamais nul. Or $Q^tQ = I_3$ impose l'orthogonalité des lignes. Cette contradiction montre que S n'est pas orthostochastique.

8. (a) Il existe deux permutations de $\{1, 2\}$: l'identité et l'échange des deux éléments. Leurs matrices sont

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

8. (b) Écrivons $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Si S est bistochastique, les sommes de la première ligne et de la première colonne donnent $b = c = 1 - a$, puis la seconde ligne donne $d = a$.

La non-négativité impose $0 \leq a \leq 1$, et

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

On peut donc prendre $\alpha = a$ et $P = J$.

Réciproquement, si $\alpha \in [0, 1]$ et P est une matrice de permutation, les coefficients de $\alpha I_2 + (1-\alpha)P$ sont non négatifs et chaque somme de ligne ou de colonne vaut $\alpha + (1-\alpha) = 1$. Cette matrice est bistochastique. L'équivalence est établie et prouve le théorème de Birkhoff-Von Neumann en dimension 2.

9. (a) On choisit successivement l'image de 1 parmi trois valeurs, celle de 2 parmi les deux restantes, puis celle de 3. Le nombre de permutations vaut $3 \times 2 \times 1 = 3! = 6$.
9. (b) Pour retrouver $\sigma(j)$, on repère la ligne du coefficient 1 dans la j -ème colonne. Chaque matrice proposée comporte un coefficient 1 par colonne et par ligne, donc correspond à une permutation.

Matrice	σ représentée par son triplet
P_1	(1, 3, 2)
P_2	(3, 2, 1)
P_3	(2, 1, 3)
P_4	(3, 1, 2)
P_5	(2, 3, 1)

Avec I_3 , ces cinq matrices représentent les six permutations de la question 9(a).

9. (c) Posons $a = S_{1,1}$, $b = S_{1,2}$, $c = S_{2,1}$, $d = S_{2,2}$. Les sommes des deux premières lignes et colonnes imposent

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

En effet, la somme de la troisième ligne impose

$S_{3,3} = 1 - (1-a-c) - (1-b-d) = a+b+c+d-1$. La somme de la troisième colonne vaut alors également 1.

Les conditions de non-négativité des neuf coefficients sont

$$a, b, c, d \geq 0, \quad a+b \leq 1, \quad c+d \leq 1, \quad a+c \leq 1, \quad b+d \leq 1, \quad a+b+c+d \geq 1.$$

Elles sont nécessaires, et également suffisantes : elles assurent la non-négativité de la matrice affichée, dont toutes les sommes de lignes et de colonnes valent 1.

9. (d) Posons $e = a+b+c+d-1 = S_{3,3}$ et $t = \beta_0 = \min(a, d, e)$. En comparant d'abord les diagonales, puis les coefficients de positions $(1, 2)$ et $(2, 1)$, on est conduit à poser

$$\beta_1 = a - t, \quad \beta_2 = d - t, \quad \beta_3 = e - t, \quad \beta_4 = b - e + t, \quad \beta_5 = c - e + t.$$

Les trois premiers coefficients sont non négatifs puisque t est le minimum des trois termes diagonaux. Pour les deux autres, distinguons selon la valeur réalisant ce minimum.

Si $t = e$, alors $\beta_4 = b \geq 0$ et $\beta_5 = c \geq 0$.

Si $t = a$, alors $\beta_4 = 1 - c - d = S_{2,3} \geq 0$ et $\beta_5 = 1 - b - d = S_{3,2} \geq 0$.

Si $t = d$, alors $\beta_4 = 1 - a - c = S_{3,1} \geq 0$ et $\beta_5 = 1 - a - b = S_{1,3} \geq 0$. En cas d'égalité entre minima, au moins l'un de ces cas s'applique.

Enfin, la combinaison demandée s'écrit

$$tI_3 + \sum_{i=1}^5 \beta_i P_i = \begin{pmatrix} t + \beta_1 & \beta_3 + \beta_4 & \beta_2 + \beta_5 \\ \beta_3 + \beta_5 & t + \beta_2 & \beta_1 + \beta_4 \\ \beta_2 + \beta_4 & \beta_1 + \beta_5 & t + \beta_3 \end{pmatrix}.$$

Les coefficients diagonaux sont a, d, e ; ceux de positions $(1, 2)$ et $(2, 1)$ sont b, c . Les quatre autres sont

$$\beta_2 + \beta_5 = d + c - e = 1 - a - b,$$

$$\beta_1 + \beta_4 = a + b - e = 1 - c - d,$$

$$\beta_2 + \beta_4 = d + b - e = 1 - a - c,$$

$$\beta_1 + \beta_5 = a + c - e = 1 - b - d.$$

On retrouve exactement la matrice de la question 9(c), donc

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

avec tous les coefficients non négatifs.

9. (e) En sommant les coefficients d'une ligne dans l'égalité précédente, chaque matrice de permutation contribuant une somme égale à 1, on obtient

$$\sum_{i=0}^5 \beta_i = 1.$$

Ainsi S est une combinaison convexe de matrices de permutation. Le théorème de Birkhoff-Von Neumann est établi en dimension 3. Au moins un coefficient est non nul, puisque leur somme vaut 1 ; on peut omettre les coefficients nuls si l'on souhaite une décomposition sans terme nul.

10. Question adaptée au programme actuel : la sortie est dite approximativement bistochastique, conformément à la réserve déjà présente à la fin de la question originale. À un nombre fini d'itérations, les deux normalisations n'imposent pas en général simultanément les deux familles de sommes.

Voici un code qui suit exactement l'ordre demandé. Le retraitage d'un coefficient nul assure que la matrice initiale est strictement positive malgré la possibilité informatique de tirer zéro avec `rd.random()`.

Python

```
import numpy as np
import numpy.random as rd

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

```
for j in range(n):
    s = np.sum(A[:, j])
    A[:, j] = A[:, j] / s
return A
```

On suppose, comme dans la question, que n et $iter$ sont des entiers strictement positifs. Chaque coefficient initial est strictement positif ; les normalisations conservent cette propriété. Toutes les sommes utilisées comme dénominateurs sont donc strictement positives.

Le calcul de s précède la modification de la ligne ou de la colonne. Les différentes lignes ne se modifient pas entre elles pendant la première boucle ; les différentes colonnes ne se modifient pas entre elles pendant la seconde. Le code réalise donc les deux formules du sujet.

Après chaque normalisation des lignes, les sommes des lignes valent 1. Après la normalisation des colonnes, ce sont les sommes des colonnes qui valent 1, mais celles des lignes ont pu être modifiées. Par exemple, en partant de

$$\frac{1}{5} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

un cycle donne

$$\begin{pmatrix} 7/16 & 7/13 \\ 9/16 & 6/13 \end{pmatrix}.$$

Les colonnes somment à 1, tandis que les sommes des lignes sont $203/208$ et $213/208$. Une itération finie ne garantit donc pas une bistochasticté exacte.

Résultat admis dans cette version adaptée, déjà admis dans le sujet original : « On admet que si $iter$ est assez grand, on peut considérer que cette matrice est bistochastique. » On

utilise cette approximation pour interpréter la sortie, sans démontrer ici la convergence de l'algorithme ni affirmer un seuil universel d'itérations.

Code exécuté et contrôlé sur plusieurs tailles, notamment $n = 1$, et plusieurs nombres d'itérations. Les coefficients restent non négatifs, les sommes des colonnes sont égales à 1 à l'arrondi près, et les écarts des sommes des lignes à 1 deviennent très petits pour les exemples exécutés avec de nombreuses itérations. Ces observations vérifient le programme ; elles ne constituent pas une preuve du résultat admis.

Partie III : fonctions symétriques et fonctions S-convexes

Dans cette partie, $n \geq 2$. Le théorème de Birkhoff-Von Neumann en dimension quelconque est admis dans l'énoncé : toute matrice bistochastique est une combinaison convexe finie de matrices de permutation. Il sera utilisé à la question 18 puis, par ce résultat, à la question 27.

11. Pour chaque i , on a

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

En effet, l'unique indice pour lequel $\sigma(j) = i$ est $j = \sigma^{-1}(i)$. Donc

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

Le tuple associé est $x_{\sigma^{-1}}$, et non en général x_{σ} .

12. Si f est symétrique, la question 11 donne $f(P_{\sigma}X) = f(x_{\sigma^{-1}}) = f(x) = f(X)$ pour toute permutation σ .

Réciproquement, supposons $f(PX) = f(X)$ pour toute matrice de permutation. Pour $\sigma \in \mathcal{P}_n$, la question 11 appliquée à σ^{-1} donne $P_{\sigma^{-1}}X$ de tuple x_σ . Ainsi $f(x_\sigma) = f(P_{\sigma^{-1}}X) = f(X) = f(x)$. C'est exactement la symétrie de f .

13. Supposons f symétrique. Définissons $\hat{f}$ comme la restriction de f à H_n . Pour chaque x , il existe α tel que $\hat{x} = x_\alpha$ par la question 4; ainsi $f(x) = f(\hat{x}) = \hat{f}(\hat{x})$.

Réciproquement, supposons $f(x) = \hat{f}(\hat{x})$. Les tuples x et x_σ contiennent les mêmes composantes. L'unicité du tri décroissant établie à la question 5 donne $\widehat{x_\sigma} = \hat{x}$. Donc

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

La fonction f est symétrique. La représentation obtenue a pour fonction $\hat{f}$ la restriction de f à H_n : pour $y \in H_n$, on a $\hat{y} = y$.

14. Pour f_1 , le changement d'indice par une bijection conserve la somme :

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

Donc f_1 est symétrique.

Pour f_2 , prenons $x = (2, 0, \dots, 0)$ et la permutation échangeant 1 et 2. Alors $f_2(x) = 4$, tandis que $f_2(x_\sigma) = 2$. Ainsi f_2 n'est pas symétrique, pour tout $n \geq 2$.

Pour f_3 , le changement d'indices $(i, j) \mapsto (\sigma(i), \sigma(j))$ est une bijection de $\{1, \dots, n\}^2$. Il laisse donc inchangé l'ensemble des valeurs $|x_i - x_j|$ dont on prend le maximum : $f_3(x_\sigma) = f_3(x)$. Ainsi f_3 est symétrique.

Une autre description éclaire cette dernière fonction : si $M = \max_i x_i$ et $m = \min_i x_i$, alors $|x_i - x_j| \leq M - m$ pour tous i, j , et l'égalité est atteinte entre un indice de maximum et un indice de minimum. Donc $f_3(x) = M - m$, quantité conservée par permutation des coordonnées.

15. (a) Question adaptée au programme actuel : on écrit $g_{i,x}$ au lieu du raccourci g_i employé ici dans le PDF, afin de conserver explicitement le paramètre x .

Comme f est de classe C^1 sur $\mathbb{R}^n$, sa restriction à la droite $t \mapsto x + te_i$ est de classe C^1 . La dérivation le long de cette droite donne

$$g'_{i,x}(t) = \sum_{k=1}^n \partial_k f(x + te_i)(e_i)_k = \partial_i f(x + te_i).$$

15. (b) Pour tout t , la coordonnée de rang k du tuple $(x + te_{\sigma(i)})_\sigma$ est $x_{\sigma(k)} + t\delta_{\sigma(k),\sigma(i)} = x_{\sigma(k)} + t\delta_{k,i}$. Ainsi

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

La symétrie de f donne donc

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

Les deux fonctions sont égales sur $\mathbb{R}$.

15. (c) En dérivant l'égalité précédente en $t = 0$ et en utilisant la question 15(a), on obtient

$$\partial_i f(x_\sigma) = g'_{i,x_\sigma}(0) = g'_{\sigma(i),x}(0) = \partial_{\sigma(i)} f(x).$$

Cette identité vaut pour tout x et tout indice i .

16. Soit f S-convexe. Pour une matrice de permutation P , les questions 2 et 7(a) montrent que P et P^{-1} sont bistochastiques. Donc $f(PX) \leq f(X)$.

Appliquons aussi la S-convexité à P^{-1} et au vecteur PX :

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

Les deux inégalités donnent $f(PX) = f(X)$. La question 12 montre que f est symétrique. Aucune continuité n'est nécessaire.

17. Démontrons par récurrence sur m l'inégalité de convexité pour une combinaison finie.

Pour $m = 1$, on a $\alpha_1 = 1$ et l'inégalité est une égalité. Supposons le résultat acquis pour m termes. Considérons $m + 1$ termes et posons

$$s = \sum_{k=1}^m \alpha_k = 1 - \alpha_{m+1}.$$

Si $s = 0$, tous les α_k pour $k \leq m$ sont nuls et $\alpha_{m+1} = 1$: le résultat est encore une égalité.

Si $s > 0$, les coefficients α_k/s , pour $k \leq m$, sont non négatifs et somment à 1. En posant

$$z = \sum_{k=1}^m (\alpha_k/s) z_k,$$

l'hypothèse de récurrence donne

$$g(z) \leq \sum_{k=1}^m (\alpha_k/s) g(z_k).$$

Puis la définition de la convexité, applicable avec $s \in]0, 1]$, donne

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

La récurrence établit l'inégalité pour tout $m \geq 1$, y compris en présence de coefficients nuls.

18. Soit B bistochastique. Le théorème de Birkhoff-Von Neumann admis fournit une décomposition

$$B = \sum_{k=1}^r a_k P_k,$$

avec $a_k \geq 0$,

$$\sum_k a_k = 1$$

et P_k des matrices de permutation.

La question 17, appliquée à la fonction convexe f et aux vecteurs $P_k X$, donne

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

Par la symétrie de f et la question 12, chaque $f(P_k X) = f(X)$. Ainsi

$$f(BX) \leq f(X) \sum_k a_k = f(X).$$

C'est la S-convexité demandée.

19. (a) Posons $z(t) = ((1-t)x_1 + tx_2, (1-t)x_2 + tx_1)$. La fonction z est affine et f est de classe C^1 , donc g est de classe C^1 sur $\mathbb{R}$. La dérivation le long d'une droite donne

$$g'(t) = (x_2 - x_1)\partial_1 f(z(t)) + (x_1 - x_2)\partial_2 f(z(t))$$

$$= (x_2 - x_1)(\partial_1 f(z(t)) - \partial_2 f(z(t))).$$

19. (b) Pour $t \in [0, 1]$, la matrice $B_t = \begin{pmatrix} 1-t & t \\ t & 1-t \end{pmatrix}$ est bistochastique : ses coefficients

sont non négatifs et ses sommes de lignes et colonnes valent 1. Son action sur

$X = (x_1, x_2)^t$ donne le tuple $z(t)$.

Par S-convexité, $g(t) = f(B_t X) \leq f(X) = g(0)$. Pour $t \in]0, 1]$, on peut donc diviser par $t > 0$:

$$\frac{g(t) - g(0)}{t} \leq 0.$$

La dérivabilité en 0 permet de passer à la limite à droite, et donne $g'(0) \leq 0$. On utilise un maximum sur un intervalle situé à droite de 0, sans prétendre que $g'(0)$ serait nécessairement nul.

19. (c) Puisque $z(0) = x$, les deux questions précédentes donnent

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

En multipliant par -1 , on obtient $(x_1 - x_2)(\partial_1 f(x) - \partial_2 f(x)) \geq 0$. Cette conclusion couvre aussi le cas $x_1 = x_2$.

20. Si $i = j$, le produit demandé vaut 0. Supposons donc $i \neq j$.

Soit P la matrice de permutation échangeant les coordonnées i et j et laissant toutes les autres inchangées. Pour $t \in [0, 1]$, $B_t = (1-t)I_n + tP$ est bistochastique, car il est une combinaison convexe de deux matrices bistochastiques.

Posons $z(t) = x + t(x_j - x_i)e_i + t(x_i - x_j)e_j$ et $g(t) = f(z(t))$. Le tuple $z(t)$ correspond à $B_t X$. La S-convexité donne $g(t) \leq g(0)$ pour $t \in [0, 1]$, et le quotient de différence à droite donne $g'(0) \leq 0$ comme à la question 19(b).

La dérivation le long de la droite donne

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

On conclut $(x_i - x_j)(\partial_i f(x) - \partial_j f(x)) \geq 0$ pour tous i, j et tout x .

21. Montrons d'abord la nécessité. Comme h est de classe C^1 ,

$$f(x) = \sum_k h(x_k)$$

est de classe C^1 et $\partial_i f(x) = h'(x_i)$. Si f est S-convexe, la question 20 donne, en prenant les deux premières coordonnées égales à des réels quelconques a, b ,

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

Pour $a > b$, on en déduit $h'(a) \geq h'(b)$. La dérivée h' est donc croissante sur $\mathbb{R}$. Le critère de convexité d'une fonction dérivable sur un intervalle montre que h est convexe.

Pour la suffisance, deux approches naturelles sont possibles.

Méthode 1 — Symétrie et convexité de la somme

Si h est convexe, pour $x, y \in \mathbb{R}^n$ et $t \in [0, 1]$, on a, pour chaque k ,
 $h(tx_k + (1 - t)y_k) \leq th(x_k) + (1 - t)h(y_k)$. En sommant ces inégalités, on obtient
 $f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y)$: f est convexe.

La fonction f est symétrique, puisque toute permutation ne fait que réordonner les termes de sa somme. La question 18 donne alors sa S-convexité.

Méthode 2 — Convexité appliquée à chaque ligne

Soit B bistochastique. Pour chaque i , les coefficients $B_{i,j}$ sont non négatifs et somment à 1. La question 17 appliquée à la fonction convexe h donne

$$h\left(\sum_{j=1}^n B_{i,j}x_j\right) \leq \sum_{j=1}^n B_{i,j}h(x_j).$$

En sommant sur i et en réordonnant des sommes finies, on obtient

$$f(BX) \leq \sum_{j=1}^n \left(\sum_{i=1}^n B_{i,j}\right) h(x_j) = \sum_{j=1}^n h(x_j) = f(X).$$

La dernière égalité utilise les sommes des colonnes, égales à 1. Cette preuve directe évite Birkhoff-Von Neumann et montre le rôle distinct des sommes de lignes et de colonnes. Dans les deux approches, h convexe implique f S-convexe ; avec la nécessité déjà établie, on obtient l'équivalence.

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

Toutes les matrices étudiées ici sont réelles et symétriques. Le théorème spectral réel du programme assure l'existence d'une base orthonormée de vecteurs propres.

Si Q est orthogonale, alors ${}^tQ = Q^{-1}$, comme établi à la question 7(a). La matrice $QA{}^tQ$ est symétrique : sa transposée vaut $Q{}^tA{}^tQ = QA{}^tQ$. Si $AV = \lambda V$, alors $(QA{}^tQ)(QV) = \lambda QV$. L'application $V \mapsto QV$ réalise une bijection entre les espaces propres correspondants. Elle préserve leurs dimensions et transporte une base orthonormée de vecteurs propres vers une telle base. Ainsi, les listes de valeurs propres avec répétitions et leur ordre décroissant sont conservés :

$$\hat{\lambda}(QA{}^tQ) = \hat{\lambda}(A).$$

Ces faits n'utilisent ni déterminant ni polynôme caractéristique.

22. Pour A symétrique et Q orthogonale, une récurrence sur $k \geq 1$ donne

$$(QA^tQ)^k = QA^{k+1}Q.$$

L'initialisation est immédiate ; l'hérédité repose sur ${}^tQQ = I_n$, car

$$(QA^{k+1}Q)(QA^tQ) = QA^{k+1}Q.$$

La propriété du cours $\text{Tr}(MN) = \text{Tr}(NM)$, pour deux matrices carrées de même taille, donne

$$F_k(QA^tQ) = \text{Tr}(QA^{k+1}Q) = \text{Tr}({}^tQQA^k) = \text{Tr}(A^k) = F_k(A).$$

Donc F_k est spectrale.

23. Méthode 1 — Action sur la base canonique

Fixons j . La question 3 donne ${}^tP_\sigma e_j = e_{\sigma^{-1}(j)}$. Donc

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

La question 11 donne aussi $D(P_\sigma X) e_j = x_{\sigma^{-1}(j)} e_j$. Les deux matrices ont la même action sur chaque vecteur de la base canonique ; ainsi $P_\sigma D(X) {}^tP_\sigma = D(P_\sigma X)$.

Méthode 2 — Calcul des coefficients

La diagonalité de $D(X)$ donne

$$(P_{\sigma}D(X)^tP_{\sigma})_{i,j} = \sum_{k=1}^n x_k \delta_{i,\sigma(k)} \delta_{j,\sigma(k)}.$$

Si $i \neq j$, chaque produit de symboles est nul. Si $i = j$, l'unique terme non nul est $x_{\sigma^{-1}(i)}$. La matrice obtenue est donc diagonale, de diagonale $P_{\sigma}X$. Cette approche montre explicitement la disparition des coefficients hors diagonale.

- 24.** Puisque A est réelle symétrique, le théorème spectral fournit une base orthonormée de vecteurs propres. Réordonnons ses vecteurs en $V_1, \dots, V_n$ de façon que $AV_i = \hat{\lambda}_i(A)V_i$.

Soit Q la matrice dont les colonnes sont les V_i . Leurs produits scalaires donnent ${}^tQQ = I_n$, donc $Q^{-1} = {}^tQ$ et $Q \in \mathcal{O}_n$. Les égalités sur les colonnes s'écrivent $AQ = Q\Lambda(A)$. En multipliant à droite par tQ , on obtient $A = Q\Lambda(A){}^tQ$.

Les valeurs propres répétées ne posent aucune difficulté : on choisit une base orthonormée de chaque espace propre, puis on ordonne les vecteurs selon leur valeur propre.

- 25.** (a) La question 24 écrit $A = Q\Lambda(A){}^tQ$ avec Q orthogonale. La fonction F étant spectrale, $F(A) = F(Q\Lambda(A){}^tQ) = F(\Lambda(A))$.

- 25.** (b) Pour tout $P \in P_n$ et tout vecteur colonne X , la question 23 et l'orthogonalité de P donnent

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

La question 12 assure que f est symétrique.

- 25.** (c) L'application D est linéaire : $D(tX + (1-t)Y) = tD(X) + (1-t)D(Y)$. Les matrices $D(X)$ et $D(Y)$ sont symétriques. Si F est convexe, pour $t \in [0, 1]$,

$$f(tX + (1-t)Y) = F(tD(X) + (1-t)D(Y)) \leq tF(D(X)) + (1-t)F(D(Y)) = tf(X) + (1-t)f(Y).$$

Donc f est convexe.

- 26.** Supposons d'abord F spectrale. Définissons $f(x) = F(D(X))$ comme à la question 25. Elle est symétrique d'après 25(b), et 25(a) donne

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

Réciproquement, supposons qu'une fonction symétrique f satisfasse cette formule pour toute matrice symétrique. L'invariance de la liste ordonnée des valeurs propres, expliquée au début de la partie IV, donne

$$F(QA^tQ) = f(\hat{\lambda}(QA^tQ)) = f(\hat{\lambda}(A)) = F(A).$$

Donc F est spectrale.

Pour l'unicité, soit $x \in \mathbb{R}^n$ quelconque. Les valeurs propres de la matrice diagonale $D(X)$, avec répétitions, sont ses coefficients diagonaux, donc $\hat{\lambda}(D(X)) = \widehat{X}$. Toute fonction symétrique réalisant la représentation doit vérifier

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

Sa valeur en tout x est ainsi imposée. La fonction symétrique associée est unique et vaut $x \mapsto F(D(X))$.

- 27.** (a) Écrivons $A = Q\Lambda(A)^tQ$ comme à la question 24. Le coefficient diagonal de rang i vaut

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

Posons $S_{i,j} = Q_{i,j}^2$. La matrice S est orthostochastique et donc bistochastique par 7(a). Les égalités précédentes donnent $\text{diag}(A) = S\hat{\lambda}(A)$.

27. (b) Diagonalisons $C = A + B$ dans une base orthonormée ordonnée : $C = R\Lambda(C)^tR$. Posons $A' = {}^tRAR$ et $B' = {}^tRBR$. Ces matrices sont symétriques et vérifient $A' + B' = \Lambda(C)$.

Elles ont respectivement les mêmes listes de valeurs propres que A et B , puisque ce sont des conjugaisons orthogonales. Par 27(a), il existe des matrices bistochastiques S_1, S_2 telles que

$$\text{diag}(A') = S_1\hat{\lambda}(A), \quad \text{diag}(B') = S_2\hat{\lambda}(B).$$

La linéarité de l'application diag donne alors

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

27. (c) Fixons A, B symétriques et $t \in [0, 1]$. Appliquons 27(b) aux matrices tA et $(1-t)B$. Pour tout $s \geq 0$, on a $\hat{\lambda}(sA) = s\hat{\lambda}(A)$: si $s > 0$, la multiplication conserve l'ordre des valeurs propres ; si $s = 0$, elles sont toutes nulles.

Il existe donc deux matrices bistochastiques S_1, S_2 , dépendant éventuellement de A, B, t , telles que

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

La fonction f est symétrique par 26 et convexe par hypothèse ; elle est donc S-convexe par 18. Sa convexité, puis sa S-convexité, donnent

$$F(tA + (1-t)B) = f(tS_1\hat{\lambda}(A) + (1-t)S_2\hat{\lambda}(B))$$

$$\leq tf(S_1\hat{\lambda}(A)) + (1-t)f(S_2\hat{\lambda}(B))$$

$$\leq tf(\hat{\lambda}(A)) + (1-t)f(\hat{\lambda}(B)) = tF(A) + (1-t)F(B).$$

La convexité de F est établie, y compris aux extrémités $t = 0$ et $t = 1$. C'est le théorème de Davis dans le cadre étudié.

27. (d) Par la définition de la fonction associée dans 26,

$$F(\text{DG}(A)) = F(D(\text{diag}(A))) = f(\text{diag}(A)).$$

La question 27(a) fournit une matrice bistochastique S telle que $\text{diag}(A) = S\hat{\lambda}(A)$. La S -convexité de f donne

$$F(\text{DG}(A)) = f(S\hat{\lambda}(A)) \leq f(\hat{\lambda}(A)) = F(A).$$

28. (a) Fixons k et décomposons le vecteur unitaire X_k dans la base orthonormée (V_i) :

$$X_k = \sum_{i=1}^n c_{k,i} V_i, \quad c_{k,i} = \langle X_k, V_i \rangle, \quad \sum_{i=1}^n c_{k,i}^2 = 1.$$

L'orthonormalité et les égalités $AV_i = \hat{\lambda}_i(A)V_i$ donnent

$$\begin{aligned}\langle X_k, AX_k \rangle &= \sum_{i=1}^n \hat{\lambda}_i(A) c_{k,i}^2 \\ &= \hat{\lambda}_m(A) + \sum_{i=1}^n (\hat{\lambda}_i(A) - \hat{\lambda}_m(A)) c_{k,i}^2.\end{aligned}$$

Pour $i \geq m$, les différences sont non positives. En supprimant ces termes, on majore la somme, ce qui donne exactement (3) :

$$\langle X_k, AX_k \rangle \leq \hat{\lambda}_m(A) + \sum_{i=1}^{m-1} (\hat{\lambda}_i(A) - \hat{\lambda}_m(A)) \langle X_k, V_i \rangle^2.$$

Pour chaque i , la décomposition de V_i dans la base orthonormée (X_k) donne

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

Ainsi

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

En sommant (3) sur $k = 1, \dots, m$, et puisque les différences pour $i < m$ sont non négatives, on obtient

$$\begin{aligned}\sum_{k=1}^m \langle X_k, AX_k \rangle &\leq m \hat{\lambda}_m(A) + \sum_{i=1}^{m-1} (\hat{\lambda}_i(A) - \hat{\lambda}_m(A)) \sum_{k=1}^m \langle X_k, V_i \rangle^2 \\ &\leq m \hat{\lambda}_m(A) + \sum_{i=1}^{m-1} (\hat{\lambda}_i(A) - \hat{\lambda}_m(A)) = \Sigma_m(A).\end{aligned}$$

C'est (4). Pour $m = 1$, la somme d'indice i est vide et vaut 0. Pour $m = n$, la somme sur une base orthonormée donne en fait l'égalité avec la trace.

La preuve de (4) utilise seulement l'orthonormalité de $X_1, \dots, X_m$, et non leur qualité de vecteurs propres de C . En effet, on peut compléter cette famille en base orthonormée pour obtenir la majoration des sommes de carrés. On a donc aussi établi, pour toute famille orthonormée $Z_1, \dots, Z_m$,

$$\sum_{k=1}^m \langle Z_k, AZ_k \rangle \leq \Sigma_m(A).$$

Cette observation sera utile à la question 30.

28. (b) Méthode 1 — Sous-additivité et homogénéité positive

Choisissons les m premiers vecteurs propres orthonormés de $C = A + B$ comme dans 28(a). Puisque $CX_k = \hat{\lambda}_k(C)X_k$ et $\|X_k\| = 1$,

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

L'inégalité (4) appliquée à A , puis à B , donne $\Sigma_m(A + B) \leq \Sigma_m(A) + \Sigma_m(B)$.

Pour tout $s \geq 0$, l'ordre des valeurs propres donne $\Sigma_m(sA) = s\Sigma_m(A)$, avec le cas $s = 0$ traité directement. Donc, pour $t \in [0, 1]$,

$$\Sigma_m(tA + (1 - t)B) \leq \Sigma_m(tA) + \Sigma_m((1 - t)B) = t\Sigma_m(A) + (1 - t)\Sigma_m(B).$$

Ainsi chaque Σ_m est convexe.

Méthode 2 — Maximum d'expressions linéaires

L'observation finale de 28(a) donne une majoration pour toute famille orthonormée $Z_1, \dots, Z_m$. L'égalité est atteinte en choisissant les m premiers vecteurs propres ordonnés de A . On a donc établi la formule particulière

$$\Sigma_m(A) = \max_{(Z_1, \dots, Z_m) \text{ orthonormée}} \sum_{k=1}^m \langle Z_k, AZ_k \rangle.$$

Ce maximum existe parce qu'une famille de vecteurs propres le réalise; aucun théorème topologique supplémentaire n'est nécessaire.

Pour une famille orthonormée fixée, la quantité

$$L_Z(A) = \sum_{k=1}^m \langle Z_k, AZ_k \rangle$$

est linéaire en A . Donc

$$L_Z(tA + (1-t)B) = tL_Z(A) + (1-t)L_Z(B) \leq t\Sigma_m(A) + (1-t)\Sigma_m(B).$$

En prenant le maximum sur les familles orthonormées, on retrouve la convexité. Cette méthode décrit Σ_m comme un maximum de fonctions linéaires; la première utilise directement les deux propriétés numériques déduites de (4).

29. La liste ordonnée des valeurs propres est conservée par conjugaison orthogonale, comme expliqué au début de la partie IV. On a donc $H(QA^tQ) = H(A)$: H est spectrale.

Pour établir la convexité, utilisons l'indication et posons $\alpha_{n+1} = 0$. Alors

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

En effet, le coefficient de $\hat{\lambda}_k(A)$ dans le membre de droite est

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

par télescopage.

Chaque différence $\alpha_m - \alpha_{m+1}$ est non négative, y compris la dernière qui vaut α_n . Chaque Σ_m est convexe d'après 28(b). En multipliant leurs inégalités de convexité par ces coefficients non négatifs et en les additionnant, on obtient celle de H . Ainsi H est spectrale et convexe.

30. On donne deux preuves : la première utilise les sommes partielles issues de la question 28 ; la seconde utilise les matrices bistochastiques de la question 27 et le théorème admis de Birkhoff-Von Neumann. Dans les deux cas, aucune positivité des valeurs propres ni des diagonales n'est supposée.

Méthode 1 — Sommes partielles et identité télescopique

Posons $a = \hat{\lambda}(A)$, $b = \hat{\lambda}(B)$, $u = \widehat{U}$ et $v = \widehat{V}$. Ces quatre vecteurs sont décroissants.

Pour chaque m , choisissons $i_1, \dots, i_m$ les indices des m plus grands coefficients diagonaux de A . Les vecteurs canoniques $e_{i_1}, \dots, e_{i_m}$ sont orthonormés. L'observation finale de 28(a) donne

$$\sum_{k=1}^m u_k = \sum_{k=1}^m \langle e_{i_k}, A e_{i_k} \rangle \leq \Sigma_m(A) = \sum_{k=1}^m a_k.$$

Pour $m = n$, l'égalité des traces d'une matrice et de sa diagonalisation donne

$$\sum_k u_k = \text{Tr}(A) = \sum_k a_k.$$

De même,

$$\sum_{k=1}^m v_k \leq \sum_{k=1}^m b_k, \quad \sum_{k=1}^n v_k = \sum_{k=1}^n b_k.$$

Établissons l'identité finie nécessaire. Pour des réels $d_1, \dots, d_n$, posons

$$D_m = \sum_{k=1}^m d_k$$

et $D_0 = 0$. Comme $d_k = D_k - D_{k-1}$, on a, pour tout vecteur w ,

$$\sum_{k=1}^n d_k w_k = \sum_{m=1}^{n-1} D_m (w_m - w_{m+1}) + D_n w_n.$$

Cette égalité résulte en développant les deux sommes et en regroupant le coefficient de chaque D_m . Si $D_m \geq 0$ pour $m < n$, $D_n = 0$ et w est décroissant, son membre de droite est non négatif, même lorsque certaines composantes de w sont négatives.

Appliquons-la d'abord à $d_k = a_k - u_k$ et $w = v$. Les inégalités de sommes partielles donnent $D_m \geq 0$ et l'égalité des totaux donne $D_n = 0$. Ainsi $\langle a, v \rangle - \langle u, v \rangle \geq 0$.

Appliquons-la ensuite à $d_k = b_k - v_k$ et $w = a$, également décroissant. On obtient $\langle a, b \rangle - \langle a, v \rangle \geq 0$.

En combinant,

$$\langle \widehat{U}, \widehat{V} \rangle = \langle u, v \rangle \leq \langle a, v \rangle \leq \langle a, b \rangle = \langle \widehat{\lambda}(A), \widehat{\lambda}(B) \rangle.$$

Méthode 2 — Birkhoff-Von Neumann et réarrangement prouvé par échanges

Par 27(a), il existe des matrices bistochastiques S, T telles que $U = Sa$ et $V = Tb$. Choisissons des matrices de permutation R, L réalisant le tri : $\widehat{U} = RU$ et $\widehat{V} = LV$. Leur existence résulte des questions 4 et 11 ; on prend la matrice associée à l'inverse de la permutation de tri.

Posons $S' = RS$ et $T' = LT$. Une permutation des lignes conserve la non-négativité, les sommes de lignes et les sommes de colonnes ; S' et T' sont donc bistochastiques.

La transposée d'une matrice bistochastique est bistochastique. Le produit de deux telles matrices l'est aussi : ses coefficients sont non négatifs, et pour D, E bistochastiques,

$$\sum_j (DE)_{i,j} = \sum_k D_{i,k} \sum_j E_{k,j} = \sum_k D_{i,k} = 1,$$

$$\sum_i (DE)_{i,j} = \sum_k \left(\sum_i D_{i,k} \right) E_{k,j} = \sum_k E_{k,j} = 1.$$

Donc $M = {}^t S' T'$ est bistochastique et

$$\langle \widehat{U}, \widehat{V} \rangle = \langle S'a, T'b \rangle = {}^t a M b.$$

Birkhoff-Von Neumann donne

$$M = \sum_{r=1}^q c_r P_r,$$

où $c_r \geq 0$ et

$$\sum_r c_r = 1.$$

Il suffit donc de montrer $\langle a, Pb \rangle \leq \langle a, b \rangle$ pour chaque permutation P .

Prouvons cette inégalité particulière. Le vecteur $z = Pb$ contient les mêmes composantes que b . Plaçons d'abord une composante égale à b_1 , le maximum, à la première position en l'échangeant avec sa position actuelle j . La variation du produit scalaire avec a vaut

$$(a_1 - a_j)(b_1 - z_1) \geq 0,$$

car $a_1 \geq a_j$ et $b_1 \geq z_1$. Une fois les premières positions fixées, à l'étape k le suffixe contient exactement les composantes restantes de b , et son maximum est b_k . Si celui-ci se trouve en position $j \geq k$, l'échange avec la position k augmente le produit scalaire de

$$(a_k - a_j)(b_k - z_k) \geq 0.$$

Après au plus n étapes, le vecteur est b . Le produit scalaire n'a jamais diminué, donc $\langle a, Pb \rangle \leq \langle a, b \rangle$. Les composantes répétées et les composantes négatives ne modifient pas l'argument, qui utilise seulement l'ordre.

Finalement,

$$\langle \widehat{U}, \widehat{V} \rangle = \sum_{r=1}^q c_r \langle a, P_r b \rangle \leq \sum_{r=1}^q c_r \langle a, b \rangle = \langle a, b \rangle.$$

L'inégalité de Fan est donc établie par les deux méthodes. La première permet de reconnaître une comparaison par sommes partielles ; la seconde explique comment le mélange par des matrices bistochastiques réduit le produit scalaire maximal entre deux listes ordonnées.

Formulaire Python

Question adaptée au programme actuel : les descriptions informatiques de l'annexe ont été précisées. `np.linspace`, `np.zeros(n)` et `np.arange` renvoient des tableaux à une dimension, et non des matrices lignes ou des listes Python. `np.sum(M, axis=...)` renvoie aussi un tableau à une dimension lorsque `M` est une matrice à deux dimensions.

La plage informatique de `rd.random` est $[0, 1[$. Pour la question 10, le retraitage éventuel des zéros permet de travailler avec des coefficients strictement positifs. La distinction entre intervalles ouverts ou fermés ne change pas la loi uniforme continue idéale, mais compte pour la division dans un programme.

Le paramètre `a` de `rd.gamma(m, a, ...)` est une échelle : le taux correspondant est $1/a$. La mention non paramétrée de la convention $\Gamma(m, a)$ dans le formulaire original a été remplacée par une description explicite et sa densité, afin d'éviter une confusion entre taux et échelle. Aucun calcul du problème ne dépend de cette commande.

Les alias `al` et `plt`, utilisés dans le formulaire sans import explicite dans le préambule original, sont définis par les imports correspondants dans l'annexe transcrite.

Relevé des adaptations

- 10 : la sortie est qualifiée d'approximativement bistochastique. Cette précision reprend la réserve finale de l'original et évite d'affirmer une propriété exacte après un nombre fini d'itérations ; l'algorithme et le résultat admis sont conservés.
- 15(a) : la notation abrégée g_i est remplacée par $g_{i,x}$, notation déjà définie dans le préambule de la question. Rectification locale de notation, sans modification de la demande mathématique.
- Formulaire Python, sections I et III : descriptions des tableaux à une dimension, de la plage de `rd.random` et de l'échelle de `rd.gamma` précisées pour correspondre aux commandes réelles de NumPy.

- Formulaire Python, sections II et IV : imports des alias `al` et `plt` explicités pour rendre utilisables les commandes de l'annexe.