Devoir à la maison n°22

- Le devoir devra être rédigé sur des copies *doubles*.
- Les copies ne devront comporter ni rature, ni renvoi, ni trace d'effaceur.
- Toute copie ne satisfaisant pas à ces exigences devra être intégralement récrite.

Problème 1

1 En développant χ_J par rapport à sa première colonne, on obtient $\chi_J = X^n - 1$. Ainsi $\mathrm{Sp}(J) = \mathbb{U}_n$ et comme χ_J est scindé sur $\mathbb{C}$ à racines simples, J est diagonalisable sur $\mathbb{C}$.

2 Posons $\omega = e^{\frac{2i\pi}{n}}$ de sorte que $\mathbb{U}_n = \{\omega^k,\ k \in [\![0, n-1]\!]\}$. On vérifie qu'en posant $X_k = (1, \omega^k, \omega^{2k}, \dots, \omega^{(n-1)k})$, $JX_k = \omega^k X_k$. Ainsi X_k est un vecteur propre associé à la valeur propre ω^k. Comme J est diagonalisable, $(X_0, \dots, X_{n-1})$ est une base de $\mathbb{C}^n$ formée de vecteurs propres de J.

3 On a évidemment $U_0 = (1, 0, \dots, 0)$. De plus,

$$\mathbb{P}(X_{m+1} = k) = \begin{cases} \frac{1}{2}\mathbb{P}(X_m = k-1) + \frac{1}{2}\mathbb{P}(X_m = k+1) & \text{si } 0 \leq k \neq n-2 \\ \frac{1}{2}\mathbb{P}(X_m = n-1) + \frac{1}{2}\mathbb{P}(X_m = 1) & \text{si } k = 0 \\ \frac{1}{2}\mathbb{P}(X_m = n-2) + \frac{1}{2}\mathbb{P}(X_m = 0) & \text{si } k = n-1 \end{cases}$$

On en déduit que $U_{m+1} = AU_m$ en posant $A = \frac{1}{2}(J^\top + J)$.

4 Les colonnes de J forment une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ (pour le produit scalaire usuel) puisqu'elles forment la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$. On en déduit que J est orthogonale. Notamment, $J^{-1} = J^\top$. En posant $D = \mathrm{diag}(1, \omega, \dots, \omega^{n-1})$, il existe $P \in GL_n(\mathbb{C})$ telle que $J = PDP^{-1}$. On en déduit que $A = \frac{1}{2}P(D+D^{-1})P^{-1}$ où $\frac{1}{2}(D+D^{-1}) = \mathrm{diag}(1, \cos(2\pi/n), \dots, \cos(2(n-1)\pi/n))$. Par conséquent,

$$\mathrm{Sp}(A) = \left\{\cos\left(\frac{2k\pi}{n},\ k \in [\![0, n-1]\!]\right)\right\}$$

Il est clair que la valeur propre de A de module maximal est 1. On vérifie aisément qu'un vecteur propre unitaire associé à cette valeur propre est $\frac{1}{\sqrt{n}}(1, \dots, 1)$.

5 Comme la matrice A est symétrique réelle, il existe une base orthonormée $(Y_0, \dots, Y_{n-1})$ de $\mathbb{R}^n$ où Y_k est la valeur propre associée à la valeur propre $\lambda_k = \cos\left(\frac{2k\pi}{n}\right)$. Il existe $(\alpha_0, \dots, \alpha_{n-1}) \in \mathbb{R}^n$ tel que $U_0 = \sum_{k=0}^{n-1} \alpha_k Y_k$. Alors

$$U_m = A^m U_0 = \sum_{k=0}^{n-1} \alpha_k \lambda_k^m Y_k \underset{n \to +\infty}{\longrightarrow} \alpha_0 Y_0$$

car $\lambda_0 = 1$ et $|\lambda_k| < 1$ pour $k \in [\![1, n-1]\!]$. Or $\alpha_0 Y_0$ est le projeté orthogonal de $U_0 = (1, 0, \dots, 0)$ sur $\mathrm{vect}(Y_0)$ où $Y_0 = \frac{1}{\sqrt{n}}(1, \dots, 1)$ est unitaire de sorte que

$$\alpha_0 Y_0 = \langle U_0, Y_0 \rangle Y_0 = \frac{1}{n}(1, \dots, 1)$$

En résumé, (U_m) converge vers $\frac{1}{n}(1, \dots, 1)$.

6 Tout d'abord, $\mathcal{B}_n$ n'est pas un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ puisqu'il ne contient pas la matrice nulle.

Notons $\varphi_i : M \in \mathcal{M}_n(\mathbb{R}) \mapsto \sum_{j=1}^{n} M_{i,j}$ et $\psi_j : M \in \mathcal{M}_n(\mathbb{R})$ et $\xi_{i,j} : M \in \mathcal{M}_n(\mathbb{R}) \mapsto M_{i,j}$. Ces applications sont clairement des formes linéaires. De plus,

$$M \in \mathcal{B}_n \iff \begin{cases} \forall i \in [\![1,n]\!],\ \varphi_i(M) = 1 \\ \forall j \in [\![1,n]\!],\ \psi_j(M) = 1 \\ \forall (i,j) \in [\![1,n]\!]^2,\ \xi_{i,j}(M) \geq 0 \end{cases}$$

Soient $(M,N) \in \mathcal{B}_n^2$ et $\lambda \in [0,1]$. Alors pour tout $(i,j) \in [\![1,n]\!]^2$,

$$\begin{aligned} \varphi_i(\lambda M + (1-\lambda)N) &= \lambda\varphi_i(M) + (1-\lambda)\varphi_i(N) = \lambda + 1 - \lambda = 1 \\ \psi_j(\lambda M + (1-\lambda)N) &= \lambda\psi_j(M) + (1-\lambda)\psi_j(N) = \lambda + 1 - \lambda = 1 \\ \xi_{i,j}(\lambda M + (1-\lambda)N) &= \lambda\xi_{i,j}(M) + (1-\lambda)\xi_{i,j}(N) \geq 0 \end{aligned}$$

Ainsi $\lambda M + (1-\lambda)N \in \mathcal{B}_n$. On en déduit que $\mathcal{B}_n$ est convexe.
$\mathcal{B}_n$ est bornée puisqu'une matrice bistochastique est clairement à valeurs dans $[0,1]$. Enfin,

$$\mathcal{B}_n = \left(\bigcap_{i=1}^{n} \varphi_i^{-1}(\{1\})\right) \cap \left(\bigcap_{j=1}^{n} \psi_j^{-1}(\{1\})\right) \cap \left(\bigcap_{(i,j)\in[\![1,n]\!]^2} \xi_{i,j}^{-1}(\mathbb{R}_+)\right)$$

Or les applications φ_i, ψ_j et $\xi_{i,j}$ sont continues car linéaires sur $\mathcal{M}_n(\mathbb{R})$ qui est de dimension finie. Donc les images réciproques des fermés $\{1\}$ et $\mathbb{R}_+$ par ces applications sont des fermés. Enfin, $\mathcal{B}_n$ est fermé comme intersection de fermés. Comme $\mathcal{M}_n(\mathbb{R})$ est de dimension finie, $\mathcal{B}_n$ est compact en tant que fermé borné.

7 On vérifie aisément que $\mathcal{P}_n \subset \mathcal{B}_n$. On laisse le lecteur vérifié que, pour $(\sigma,\tau) \in S_n^2$, $M_\sigma M_\tau = M_{\sigma\circ\tau}$. Notamment, $M_\sigma M_{\sigma^{-1}} = M_{\mathrm{Id}} = I_n$. Ainsi $\mathcal{P}_n \subset GL_n(\mathbb{R})$. Enfin, $\mathcal{P}_n$ est un sous-groupe de $GL_n(\mathbb{R})$ en tant qu'image du morphisme de groupes $\sigma \in S_n \mapsto M_\sigma$.
Soit $\sigma \in S_n$. Comme S_n est un groupe d'ordre fini, σ est également d'ordre fini. Notons p cet ordre. Alors $(M_\sigma)^p = M_{\sigma^p} = M_{\mathrm{Id}} = I_n$. Ainsi le polynôme $X^p - 1$ annule M_σ. Comme ce polynôme est simplement scindé sur $\mathbb{C}$, M_σ est diagonalisable sur $\mathbb{C}$.
Enfin, $\mathcal{P}_n$ n'est évidemment pas convexe. En effet, si on prend deux matrices de permutation M et N distinctes, $\frac{1}{2}(M+N)$ possèdera des coefficients égaux à $1/2$ (prendre I_n et J par exemple).

8 Soit $A \in \mathcal{P}_n$. Supposons qu'il existe $\lambda \in]0,1[$ et $(M,N) \in \mathcal{B}_n^2$ tel que $A = \lambda M + (1-\lambda)N$. Soit $i \in [\![1,n]\!]^2$. On sait alors qu'il existe un unique $j \in [\![1,n]\!]^2$ tel que $A_{i,j} = 1$. De plus,

$$\lambda(A_{i,j} - M_{i,j}) + (1-\lambda)(A_{i,j} - N_{i,j}) = 0$$

Les deux termes de cette somme sont positis puisque $A_{i,j} = 1$ et M et N sont à coefficients dans $[0,1]$. Ces deux termes sont donc nuls. Comme λ et $1-\lambda$ ne sont pas nuls, $M_{i,j} = N_{i,j} = A_{i,j} = 1$. Comme la somme des coefficients de la ligne i de M et N vaut 1, les autres coefficients de cette ligne sont nuls dans M et N. Ainsi les lignes i de A, M et N sont égales. Ceci étant vrai pour tout $i \in [\![1,n]\!]$, $A = M = N$.
Les matrices de $\mathcal{P}_n$ sont donc bien extrémales dans $\mathcal{B}_n$.

9

10 Notons m le plus petit des coefficients A_{i_k,j_k} et $A_{i_k,j_{k+1}}$ pour $k \in [\![1,r]\!]$. Remarquons que $A - mB$ et $A + mB$ sont alors à coefficients positifs. De plus, les sommes des coefficients de chaque ligne et de chaque colonne de B sont nulles donc les sommes des coefficients de chaque ligne et de chaque colonne de $A - mB$ et $A + mB$ valent 1. Ainsi $A - mB$ et $A + mB$ sont bistochastiques. Comme $A = \frac{1}{2}(A - mB) + \frac{1}{2}(A + mB)$, A n'est pas extrémale dans $\mathcal{B}_n$.

11 Soit M une matrice extraite de A à p ligne et q colonnes avec $p + q = n + 1$. Comme le caractère bistochastique d'une matrice est invariant par permutation des lignes ou des colonnes, on peut supposer que M est constituée des p premières lignes et des q premières colonnes de A. Supposons que $M = 0$. Alors

$$\forall i \in [\![1,p]\!],\ \sum_{j=q+1}^{n} A_{i,j} = 1$$

donc, en sommant,

$$\sum_{i=1}^{p}\sum_{j=q+1}^{n} A_{i,j} = p$$

Or en intervertissant les sommes

$$\sum_{i=1}^{p}\sum_{j=q+1}^{n} \mathrm{A}_{i,j} = \sum_{j=q+1}^{n}\sum_{i=1}^{p} \mathrm{A}_{i,j} \leq \sum_{j=q+1}^{n}\sum_{i=1}^{n} \mathrm{A}_{i,j} = n-q$$

Ainsi $p \leq n-q$, ce qui contredit $p+q = n+1$.
D'après le résultat admis, M admet un chemin strictement positif.

$\boxed{\mathbf{12}}$ Si $\lambda_0 = 1$, alors $\mathrm{A}_{\sigma(j),j} = 1$ pour tout $j \in [\![1,n]\!]$. Par bistochasticité de A, on a alors $\mathrm{A}_{i,j} = \delta_{i,\sigma(j)}$ pour tout $(i,j) \in [\![1,n]\!]^2$ ou encore $\mathrm{A} = \mathrm{M}_\sigma$. Mais A n'est pas une matrice de permutation donc $\lambda_0 \neq 1$ et A_0 est bien définie.

- Si $i \neq \sigma(j)$, alors $(\mathrm{A}_0)_{i,j} = \dfrac{1}{1-\lambda_0}\mathrm{A}_{i,j} \geq 0$.
- Si $i = \sigma(j)$, alors
$$(\mathrm{A}_0)_{i,j} = \frac{1}{1-\lambda_0}(\mathrm{A}_{\sigma(j),j} - \lambda_0) \geq 0$$

Ainsi A_0 est à coefficients positifs.
Par linéarité de φ_i et ψ_j, pour tout $(i,j) \in [\![1,n]\!]^2$,

$$\varphi_i(\mathrm{A}_0) = \frac{1}{1-\lambda_0}(\varphi_i(\mathrm{A}) - \lambda_0\varphi_i(\mathrm{M}_\sigma)) = \frac{1}{1-\lambda_0}(1-\lambda_0) = 1$$
$$\psi_j(\mathrm{A}_0) = \frac{1}{1-\lambda_0}(\psi_j(\mathrm{A}) - \lambda_0\psi_j(\mathrm{M}_\sigma)) = \frac{1}{1-\lambda_0}(1-\lambda_0) = 1$$

On en déduit que A_0 est bien bistochastique.
On note j_0 un indice tel que $\lambda_0 = \mathrm{A}_{\sigma(j_0),j_0}$. Soit $(i,j) \in [\![1,n]\!]^2$ tel que $\mathrm{A}_{i,j} = 0$. Nécessairement, $i \neq \sigma(j)$ donc $(\mathrm{A}_0)_{i,j} = \dfrac{1}{1-\lambda_0}\mathrm{A}_{i,j} = 0$. Enfin, $\mathrm{A}_{\sigma(j_0),j_0} > 0$ et $(\mathrm{A}_0)_{\sigma(j_0),j_0} = \dfrac{1}{1-\lambda_0}(\mathrm{A}_{\sigma(j_0),j_0} - \lambda_0) = 0$. Ainsi A_0 possède un coefficient nul de plus que A.

$\boxed{\mathbf{13}}$ On raisonne par récurrence sur le nombre de coefficients non nuls d'une matrice bistochastique. L'hypothèse de récurrence est donc la suivante :

> HR(k) : si $\mathrm{A} \in \mathcal{B}_n$ possède au plus k coefficients non nuls, alors A s'écrit comme une combinaison convexe de matrices de permutation à coefficients strictement positifs.

Soit $\mathrm{A} \in \mathcal{B}_n$ comportant au plus n coefficients non nuls. Comme chaque ligne de A comporte au moins un coefficient non nul (puisque la somme des coefficients de chaque ligne vaut 1), A possède en fait exactement un coefficient non nul par ligne. Ce coefficient vaut alors 1 et il est alors clair que A est une matrice de permutation. Ainsi HR(n) est vraie.
Supposons que HR(k) soit vraie pour un certain $k \geq n$. Soit alors $\mathrm{A} \in \mathcal{B}_n$ possédant au plus $k+1$ coefficients non nuls. Si A est une matrice de permutation, il n'y a rien à prouver. Sinon, on peut définir A_0 comme précédemment. A_0 possède alors au moins un coefficient non nul de moins que A. D'après HR(k), il existe des réels $\alpha_1, \dots, \alpha_s$ strictement positifs et de somme 1 ainsi que des matrices de permutations $\mathrm{M}_1, \dots, \mathrm{M}_s$ tels que $\mathrm{A}_0 = \sum_{i=1}^{s} \alpha_i \mathrm{M}_i$. De plus,

$$\mathrm{A} = (1-\lambda_0)\mathrm{A}_0 + \lambda_0\mathrm{M}_\sigma$$

En posant $\mathrm{M}_0 = \mathrm{M}_\sigma$ et $\lambda_k = (1-\lambda_0)\alpha_k > 0$ (car $\lambda_0 < 1$),

$$\mathrm{A} = \sum_{i=0}^{s} \lambda_i \mathrm{M}_i$$

De plus,

$$\sum_{i=0}^{s} \lambda_i = \lambda_0 + (1-\lambda_0)\sum_{i=1}^{s} \alpha_i = 1$$

Ainsi HR($k+1$) est vraie.
En conclusion, HR(k) est vraie pour tout $k \in \mathbb{N}$ (en fait $n \leq k \leq n^2$).

$\boxed{\mathbf{14}}$ $\mathcal{P}_n$ est fini (puisque isomorphe à S_n). Ainsi φ possède un minimum (et a fortiori une borne inférieure sur $\mathcal{P}_n$). Posons $m_1 = \min\limits_{\mathrm{M} \in \mathcal{P}_n} \varphi(\mathrm{M})$.
On sait également que $\mathcal{B}_n$ est compact et que φ est continue (en tant que forme linéaire sur un espace vectoriel de dimension finie). Ainsi φ possède également un minimum (et donc une borne inférieure sur $\mathcal{B}_n$). Posons $m_2 = \min\limits_{\mathrm{M} \in \mathcal{B}_n} \varphi(\mathrm{M})$.

Puisque $\mathcal{P}_n \subset \mathcal{B}_n$, $m_2 \leq m_1$. Soit $A \in \mathcal{B}_n$ telle que $\varphi(A) = m_2$. D'après la question précédente, il existe des réels $\lambda_0, \dots, \lambda_s$ strictement positifs et de somme 1 ainsi que des matrices de permutation $M_0, \dots, M_s$ tels que $A = \sum_{i=0}^{s} \lambda_i M_i$. Alors

$$m_2 = \varphi(A) = \sum_{i=0}^{s} \lambda_i \varphi(M_i) \geq \sum_{i=0}^{s} \lambda_s m_1 = m_1$$

Par conséquent $m_2 = m_1$ et le minimum de φ sur $\mathcal{B}_n$ est atteint sur $\mathcal{P}_n$.

$\boxed{\mathbf{15}}$ Soit $(A, P, Q) \in \mathcal{M}_n(\mathbb{R}) \times O_n(\mathbb{R})^2$. Alors

$$\|PAQ\|^2 = \operatorname{tr}((PAQ)^\top PAQ) = \operatorname{tr}(Q^\top AP^\top PAQ) = \operatorname{tr}(Q^\top A^\top AQ) = \operatorname{tr}(A^\top AQQ^\top) = \operatorname{tr}(A^\top A) = \|A\|^2$$

$\boxed{\mathbf{16}}$ D'après le théorème spectral, il existe $(Q_A, Q_B) \in O_n(\mathbb{R})^2$ et deux matrices diagonales réelles D_A et D_B telles que $A = Q_A D_A Q_A^\top$ et $B = Q_B D_B Q_B^\top$. D'après la question précédente,

$$\|A - B\|^2 = \|Q_A^\top (A - B) Q_B\|^2 = \|D_A Q_A^\top Q_B - Q_A^\top D_B\|^2 = \|D_A P - P D_B\|^2$$

en posant $P = Q_A^\top Q_B$. Comme $O_n(\mathbb{R})$ est un groupe, $P \in O_n(\mathbb{R})$.

$\boxed{\mathbf{17}}$ Les colonnes et les lignes de P sont unitaires pour le produit scalaire usuel de $\mathbb{R}^n$ ce qui prouve que R est bistochastique.
Remarquons que $D_A = \operatorname{diag}(\lambda_1(A), \dots, \lambda_n(A))$ et $D_B = \operatorname{diag}(\lambda_1(B), \dots, \lambda_n(B))$. De plus, pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})^2$, $\|M\|^2 = \sum_{1 \leq i,j \leq n} M_{i,j}^2$. Donc

$$\|A - B\|^2 = \sum_{1 \leq i,j \leq n} (D_A P - P D_B)_{i,j}^2 = \sum_{1 \leq i,j \leq n} (\lambda_i(A) P_{i,j} - \lambda_j(B) P_{i,j})^2 = \sum_{1 \leq i,j \leq n} R_{i,j} (\lambda_i(A) - \lambda_j(B))^2$$

$\boxed{\mathbf{18}}$ L'application $\varphi : M \in \mathcal{M}_n(\mathbb{R}) \mapsto \sum_{1 \leq i,j \leq n} M_{i,j} (\lambda_i(A) - \lambda_j(B))^2$ est une forme linéaire. D'après la question **14**, comme B est bistochastique,

$$\|A - B\|^2 = \varphi(R) \geq \min_{\mathcal{B}_n} \varphi = \min_{\mathcal{P}_n} \varphi = \min_{\sigma \in S_n} \varphi(M_\sigma)$$

Mais par ailleurs, pour $\sigma \in S_n$,

$$\varphi(M_\sigma) = \sum_{1 \leq i,j \leq n} (M_\sigma)_{i,j} (\lambda_i(A) - \lambda_j(B))^2 = \sum_{1 \leq i,j \leq n} \delta_{i,\sigma(j)} (\lambda_i(A) - \lambda_j(B))^2 = \sum_{j=1}^{n} (\lambda_{\sigma(j)}(A) - \lambda_j(B))^2$$

On en déduit le résultat demandé.

$\boxed{\mathbf{19}}$ On peut sans perte de généralité supposer que les a_i et les b_i sont déjà rangés par ordre croissant. Soit $(X, Y) \in V^2$ tel que $X \sim P_1$ et $Y \sim P_2$. D'après la formule de transfert appliquée au couple (X, Y),

$$\mathbb{E}((X - Y)^2) = \sum_{1 \leq i,j \leq n} (a_i - b_j)^2 \mathbb{P}((X, Y) = (a_i, b_j))$$

Pour tout $(i, j) \in [\![1, n]\!]^2$,

$$\sum_{j=1}^{n} \mathbb{P}((X, Y) = (a_i, b_j)) = \mathbb{P}(X = a_i) = \frac{1}{n}$$
$$\sum_{i=1}^{n} \mathbb{P}((X, Y) = (a_i, b_j)) = \mathbb{P}(Y = b_j) = \frac{1}{n}$$

On en déduit que la matrice $R = \left(n \mathbb{P}((X, Y) = (a_i, b_j))\right)_{1 \leq i,j \leq n}$ est bistochastique. Comme à la question précédente, on montre que

$$\mathbb{E}((X - Y)^2) \geq \frac{1}{n} \min_{\sigma \in S_n} \sum_{j=1}^{n} (a_{\sigma(j)} - b_j)^2$$

Notons m ce minimum. L'inégalité précédente montre que $d^2(P_1, P_2)$ est bien définie et que $d^2(P_1, P_2) \geq \frac{m}{n}$.

Motrons que m est atteint en $\sigma = \mathrm{Id}$. Pour cela, posons $f(\sigma) = \sum_{j=1}^{n} (a_{\sigma(j)} - b_j)^2$. Remarquons que la seule permutation croissante de $\llbracket 1, n \rrbracket$ est Id. Soit $\sigma \in S_n$ telle que $\sigma \neq \mathrm{Id}$. Il existe alors $(i, j) \in \llbracket 1, n \rrbracket^2$ tel que $i < j$ et $\sigma(i) > \sigma(j)$. Notons alors τ la transposition $(\sigma(i), \sigma(j))$. Alors

$$f(\tau \circ \sigma) - f(\sigma) = (a_{\sigma(i)} - b_j)^2 + (a_{\sigma(j)} - b_i)^2 - (a_{\sigma(j)} - b_j)^2 + (a_{\sigma(i)} - b_i)^2 = 2(a_{\sigma(j)} - a_{\sigma(i)})(b_j - b_i) < 0$$

On en déduit que m est bien atteint en $\sigma = \mathrm{Id}$ et ainsi $m = \sum_{j=1}^{n} (a_j - b_j)^2$.

On suppose alors qu'il existe un couple de variables aléatoires (X, Y) sur $(\Omega, \mathcal{A}, \mathbb{P})$ tel que

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \ \mathbb{P}((X, Y) = (a_i, b_j)) = \frac{\delta_{i,j}}{n}$$

> **Remarque.** Malheureusement, rien ne garantit l'existence d'un tel couple. Il y a probablement un problème dans l'énoncé.

On vérifie alors que les lois marginales de X et Y suivent bien les lois P_1 et P_2 respectivement. De plus, $\mathbb{E}((X - Y)^2) = \frac{m}{n}$ par formule de transfert. On a donc bien

$$d^2(P_1, P_2) = \frac{m}{n} = \frac{1}{n} \sum_{j=1}^{n} (a_j - b_j)^2$$

Avec la question précédente, on a également,

$$n d^2(P_1, P_2) \leq \sum_{j=1}^{n} (a_j - b_j)^2$$