Correction de l'épreuve Mines MP 1 (année 2016)
Autour de l'inégalité de Hoffman-Wielandt

Frédéric Morlot et Jean Nougayrède

Partie A - Un exemple

1) • Notons σ le cycle $(n, n-1, ..., 2, 1)$ qui est un élément de S_n.
Alors $J = M_\sigma$ donc $\boxed{J \text{ est une matrice de permutation}}$

• Le calcul du polynôme caractéristique donne par développement par rapport à la première colonne :
$\det(XI_n - J) = X^n + (-1)^{n+1)} \times (-1) \times (-1)^{n-1} = X^n - 1$
ce qui donne $\boxed{Sp(J) = \{e^{2ik\pi/n}, k \in [\![0, n-1]\!]\}}$
Le polynôme $X^n - 1$ est scindé à racines simples dans $\mathbb{C}[X]$ donc $\boxed{J \text{ est diagonalisable sur } \mathbb{C}}$.

2) Dans cette question, on pose $w = e^{2i\pi/n}$.

Soit $k \in [\![0, n-1]\!]$ et $V_k = \begin{pmatrix} 1 \\ w^k \\ w^{2k} \\ \vdots \\ w^{(n-1)k} \end{pmatrix}$

Alors $JV_k = \begin{pmatrix} w^k \\ w^{2k} \\ w^{3k} \\ \vdots \\ w^{(n-1)k} \\ 1 \end{pmatrix} = \begin{pmatrix} w^k \\ w^{2k} \\ w^{3k} \\ \vdots \\ w^{(n-1)k} \\ w^{nk} \end{pmatrix} = w^k V_k$

On note également que $V_k \neq 0$.
Ainsi, $(V_0, ..., V_{n-1})$ est une famille de vecteurs propres associés aux n valeurs propres distinctes deux à deux $1, w, ..., w^{n-1}$ de J.
C'est donc une famille libre de n vecteurs de $\mathbb{C}^n$ et $\dim(\mathbb{C}^n) = n$.
Donc $\boxed{(V_0, ..., V_{n-1}) \text{ constitue une base de } \mathbb{C}^n \text{ de vecteurs propres de } J}$.

3) • On a directement $U_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

• Soit $m \in \mathbb{N}$.
Soit $k \in [\![0, n-1]\!]$.
On travaille ici modulo n et on note que les événements $\{X_m = k-1\}$ et $\{X_m = k+1\}$ sont incompatibles car $k - 1 \neq k + 1$ modulo n ($n \geq 3$).
D'après la définition, $P(X_{m+1} = k) = \dfrac{1}{2}P(X_m = k-1) + \dfrac{1}{2}P(X_m = k+1)$.

Donc $U_{m+1} = \begin{pmatrix} 0 & 1/2 & & & 1/2 \\ 1/2 & 0 & \ddots & (0) & \\ & 1/2 & \ddots & 1/2 & \\ (0) & & \ddots & 0 & 1/2 \\ 1/2 & & & 1/2 & 0 \end{pmatrix} \times U_m = AU_m$ en posant $\boxed{A = \dfrac{1}{2}(J + J^{n-1})}$

4) • On reprend les notations de la question 2)
J est semblable à la matrice $Diag(1, w, ..., w^{n-1})$.

Donc $\frac{1}{2}(J+J^{n-1})$ est semblable à la matrice $Diag\left(1,\frac{1}{2}(w+w^{n-1}),...,\frac{1}{2}(w^{n-1}+w^{(n-1)(n-1)})\right)$.

Pour $k\in[\![0,n-1]\!], w^k+w^{(n-1)k}=w^k+w^{-k}=2\mathrm{Re}(w^k)=2\cos\left(\frac{2k\pi}{n}\right)$

Ainsi, A est diagonalisable et ses valeurs propres (comptées avec leur ordre de multiplicité) sont les réels de la liste $(\cos(2k\pi/n))_{k\in[\![0,n-1]\!]}$

- Toutes les valeurs propres de A sont inférieures ou égales à 1 en module et 1 est valeur propre de A.

$\boxed{\text{Le vecteur } V=\frac{1}{\sqrt{n}}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} \text{ est vecteur propre unitaire associé à la valeur propre } 1}$.

5) n est impair donc pour tout $k\in[\![1,n-1]\!], |\cos(2k\pi/n)|<1$.
La matrice A est symétrique réelle donc diagonalisable dans une base orthonormale.
D'après la question précédente, il existe une matrice $P\in O_n(\mathbb{R})$, de première colonne V telle que $PD^tP=A$ avec $D=\mathrm{Diag}(1,\cos(2\pi/n),...,\cos(2(n-1)\pi/n))$.
Pour tout $m\in\mathbb{N}, A^m=PD^m\,{}^tP=P\mathrm{diag}(1,\cos^m(2\pi/n),...,\cos^m(2(n-1)\pi/n))^tP$(on note que ${}^tP=P^{-1}$)
La suite $(A^m)_{m\in\mathbb{N}}$ converge alors vers $P\times\mathrm{diag}(1,0,...,0)\times{}^tP$ (continuité du produit matriciel)
puis, comme $\forall m\in\mathbb{N}, U_m=A^mU_0$ et par continuité du produit matriciel,

$(U_m)_{m\in\mathbb{N}}$ converge vers $P\times\mathrm{diag}(1,0,...,0)\times{}^tP\times U_0=\begin{bmatrix}V & (0) & \dots & (0)\end{bmatrix}\times\begin{bmatrix}{}^tV\\ L_2\\ \vdots\\ L_n\end{bmatrix}\times U_0=\boxed{\frac{1}{n}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}}$

Partie B

6) Soit $M,N\in\mathcal{B}_n$, $\lambda\in[0,1]$, et posons $A=(1-\lambda)M+\lambda N$. On a bien :

$$\begin{cases}\forall(i,j)\in[\![1,n]\!]^2,\ A_{i,j}=(1-\lambda)M_{i,j}+\lambda N_{i,j}\geq 0\\ \forall i\in[\![1,n]\!],\ \sum_{j=1}^n A_{i,j}=(1-\lambda)\sum_{j=1}^n M_{i,j}+\lambda\sum_{j=1}^n N_{i,j}=1\\ \forall j\in[\![1,n]\!],\ \sum_{i=1}^n A_{i,j}=(1-\lambda)\sum_{i=1}^n M_{i,j}+\lambda\sum_{i=1}^n N_{i,j}=1\end{cases}$$

Donc $\boxed{A\in\mathcal{B}_n}$, ce qui montre la convexité. Passons à la compacité. Puisque $\mathcal{M}_n(\mathbb{R})$ est de dimension finie, toutes les normes sont équivalentes. Choisissons par exemple la norme infinie et soit $B\in\mathcal{B}_n$. On a :

$$\forall(i,j)\in[\![1,n]\!]^2,\ 0\leq B_{i,j}=1-\sum_{\substack{k=1\\k\neq j}}^n B_{i,k}\leq 1$$

Donc $\mathcal{B}_n$ est borné. De plus, si on se donne une suite $(C_k)_{k\in\mathbb{N}}$ à valeurs dans $\mathcal{B}_n$ qui converge vers $C\in\mathcal{M}_n(\mathbb{R})$, alors :

- par passage à la limite, C est à termes dans $\mathbb{R}^+$
- par somme sur les limites, on a :

$$\begin{cases}\forall i\in[\![1,n]\!],\ \sum_{j=1}^n C_{i,j}=1\\ \forall j\in[\![1,n]\!],\ \sum_{i=1}^n C_{i,j}=1\end{cases}$$

Donc $\mathcal{B}_n$ est fermé. Comme enfin $\mathcal{M}_n(\mathbb{R})$ est de dimension finie, $\mathcal{B}_n$ est compact. Notons que $\mathcal{B}_n$ ne peut pas être un sous espace vectoriel de $M_n(\mathbb{R})$, car il ne contient pas la matrice nulle.

7) Soit $\sigma\in\mathcal{S}_n$ et montrons que $M_\sigma\in\mathcal{P}_n$. Notons que : $\forall(i,j),\ (M_\sigma)_{i,j}=\delta_{i,\sigma(j)}$. Ainsi :

- pour $i\in[\![1,n]\!]$ fixé, on a $(M_\sigma)_{i,j}=1$ ssi $\sigma(j)=i$, ssi $j=\sigma^{-1}(i)$. On a donc $\sum_{j=1}^n(M_\sigma)_{i,j}=1$

- pour $j \in [\![1,n]\!]$ fixé, on a $(M_\sigma)_{i,j} = 1$ ssi $i = \sigma(j)$. On a donc $\sum_{i=1}^{n}(M_\sigma)_{i,j} = 1$

Montrons maintenant que $\mathcal{P}_n$ est un sous-groupe de $GL_n(\mathbb{R})$:

- d'une part on a $I_n = M_{Id} \in \mathcal{P}_n$
- d'autre part, si on se donne $\sigma, \tau \in \mathcal{S}_n$, évaluons $M_\sigma M_\tau$. Si on fixe $(i,j) \in [\![1,n]\!]^2$ on a :

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

  Si $i \neq \sigma(\tau(j))$, on ne pourra jamais avoir simultanément $k = \tau(j)$ et $i = \sigma(k)$ donc on aura nécessairement $(M_\sigma M_\tau)_{i,j} = 0$. Sinon on aura $(M_\sigma M_\tau)_{i,j} = 1$. Finalement on a $\boxed{M_\sigma M_\tau = M_{\sigma\tau}}$ et donc $M_\sigma M_\tau \in \mathcal{P}_n$
- en particulier, si on se donne $\sigma \in \mathcal{S}_n$ on a $M_\sigma M_{\sigma^{-1}} = M_{Id} = I_n$, ce qui montre que $\boxed{M_\sigma \in GL_n(\mathbb{R})}$ et que $(M_\sigma)^{-1} = M_{\sigma^{-1}} \in \mathcal{P}_n$.

Soit $\sigma \in \mathcal{S}_n$ et notons d son ordre. Alors on a $(M_\sigma)^d = M_{\sigma^d} = M_{Id} = I_n$, ce qui montre que M_σ annule le polynôme $X^d - 1$, qui est scindé simple sur $\mathbb{C}$. Donc M_σ est diagonalisable sur $\mathbb{C}$.

En revanche, $\mathcal{P}_n$ n'est pas convexe comme le montre le contre-exemple suivant. Si on pose $\sigma = (1,2)$ (transposition qui est légitime car $n \geq 2$) puis $A = \frac{1}{2}(M_\sigma + I_n)$, on a $A_{1,1} = \frac{1}{2} \notin \{0,1\}$.

8) Soit $\sigma \in \mathcal{S}_n$, $(A,B) \in \mathcal{B}_n^2$, $\lambda \in]0,1[$, et supposons que $M_\sigma = \lambda A + (1-\lambda)B$. Fixons $(i,j) \in [\![1,n]\!]^2$.

   - si $i = \sigma(j)$ on a $1 = \lambda A_{i,j} + (1-\lambda)B_{i,j}$. Or nous avons vu que A et B étaient à coefficients dans $[0,1]$. Si $A_{i,j} < 1$ ou $B_{i,j} < 1$ alors $\lambda A_{i,j} + (1-\lambda)B_{i,j} < \lambda + (1-\lambda)$. Donc nécessairement on a $A_{i,j} = B_{i,j} = 1$
   - si $i = \sigma(j)$ on a $1 = \lambda A_{i,j} + (1-\lambda)B_{i,j}$. Si $A_{i,j} > 0$ ou $B_{i,j} > 0$ alors $\lambda A_{i,j} + (1-\lambda)B_{i,j} > 0$. Donc nécessairement on a $A_{i,j} = B_{i,j} = 0$

   Finalement on a $\boxed{M_\sigma = A = B}$.

9) On rappelle encore une fois que tous les coefficients de A appartiennent à $[0,1]$. Notons :

$$X = \{(i,j) \in [\![1,n]\!]^2 / A_{i,j} \in]0,1[\,\}$$

   Attention à une coquille dans l'énoncé : il fallait lire $r \geq 2$ et non pas $r \geq 1$ (sans quoi le résultat est immédiat et ne présente guère d'intérêt).

   - premier point : soit $(i,j) \in X$
     - supposons que $\forall k \in [\![1,n]\!] \backslash \{j\},\ A_{i,k} = 0$. Alors on aurait $\sum_{k=1}^{n} A_{i,k} = A_{i,j} \neq 1$, absurde. On peut donc choisir $j' \neq j$ tel que $A_{i,j'} > 0$.
     - supposons qu'on ait $A_{i,j'} = 1$. Alors on aurait $\sum_{k=1}^{n} A_{i,k} \geq A_{i,j} + A_{i,j'} > 1$, absurde. Donc on a $A_{i,j'} \in]0,1[$.
     - on construit ainsi une fonction $h : (i,j) \mapsto (i,j')$ (comme "horizontal") de X dans X
     - symétriquement, on construit une fonction $v : (i,j) \mapsto (i',j)$ (comme "vertical"), qui à tout couple $(i,j) \in X$ associe un couple $(i',j) \in X$ avec $i' \neq i$.
   - deuxième point : X est non vide
     En effet, supposons par l'absurde que X soit vide et soit $j \in [\![1,n]\!]$. Puisque tous les coefficients de A seraient égaux à 0 ou 1 et puisque $\sum_{i=1}^{n} A_{i,j} = 1$, la j-ème colonne de A contiendrait un et un seul terme égal à 1. Cela permettrait de définir une application $\sigma : [\![1,n]\!] \to [\![1,n]\!]$ telle que :

$$\forall (i,j) \in [\![1,n]\!]^2,\ A_{i,j} = \delta_{\sigma(j),j}$$

     De plus σ serait injective, sans quoi on pourrait trouver $j \neq j'$ tels que $\sigma(j) = \sigma(j')$, après quoi on aurait $\sum_{k=1}^{n} A_{\sigma(j),k} \geq 2$. Donc par comparaison de cardinaux, σ serait bijective si bien qu'on aurait $A = M_\sigma$, ce qui est exclu.

- troisième point : définissons par récurrence une suite $(i_k, j_k)_{k\in\mathbb{N}^\star}$ à valeurs dans X de la manière suivante :
  - initialisation : puisque X est non vide, on peut choisir $(i_1, j_1) \in X$
  - hérédité : soit $k \geq 1$ et supposons avoir construit (i_k, j_k). On pose alors

$$(i_{k+1}, j_{k+1}) = (v \circ h)[(i_k, j_k)]$$

  Par construction, pour tout $k \in \mathbb{N}^\star$ on a comme demandé : $(i_k, j_k) \in X$ et $(i_k, j_{k+1}) \in X$.
- quatrième point : $[\![1, n]\!]$ est fini donc par le principe des tiroirs, il existe deux entiers $k < l$ tels que $i_k = i_l$. *A fortiori*, il existe deux entiers $k < l$ tels que $(i_k = i_l$ ou $j_k = j_l)$. Choisissons-les tels que $l - k$ soit minimal. Puis quitte à tronquer la suite, supposons que $k = 1$. Enfin, posons $r = l - 1$ et montrons que $i_1, \dots, i_r$ et $j_1, \dots, j_r$ ainsi définis conviennent (ou peu s'en faut).
  - on ne peut pas avoir $r = 1$, car par construction on a $i_1 \neq i_2$ et $j_1 \neq j_2$
  - toujours par construction, $i_1, \dots, i_r$ sont bien deux à deux distincts ; de même avec $j_1, \dots, j_r$.
  - il reste à vérifier que $(i_r, j_1) \in X$. Par hypothèse on a $i_{r+1} = i_1$ ou $j_{r+1} = j_1$, et on a aussi $(i_r, j_{r+1}) \in X$. Puis :
    * si $j_{r+1} = j_1$ c'est terminé
    * si $i_{r+1} = i_1$, remplaçons j_1 par j_{r+1} ce qui permet de se ramener au premier cas. Pour conclure, il reste à remarquer que par minimalité de $l - k = r$, l'entier j_{r+1} est bien distinct de $j_2, \dots, j_r$.

10) Utilisons la matrice B de l'énoncé. Déjà, notons que B est bien définie et n'est pas la matrice nulle (c'est ici qu'intervient l'hypothèse $r \geq 2$). Ensuite, posons :

$$\varepsilon = \min(A_{i_1,j_1}, \dots, A_{i_r,j_r}, A_{i_1,j_2}, \dots, A_{i_r,j_{r+1}}) > 0$$

Quel que soit $\lambda \in [-\varepsilon, \varepsilon]$, la matrice $A + \lambda B$ est à coefficients dans $\mathbb{R}^+$. Si on peut montrer qu'elle est bistochastique, cela suffira à conclure puisque A sera le milieu du segment $[A - \varepsilon B, A + \varepsilon B]$ et puisque ce segment n'est pas réduit à un point.

Par linéarité de la somme, il suffit de montrer que les lignes et les colonnes de B ont des sommes nulles :

- soit $i \in [\![1, n]\!]$ et montrons que $\displaystyle\sum_{j=1}^{n} B_{i,j} = 0$
  - si i est égal à un certain i_k avec $k \in [\![1, r]\!]$, on a $\displaystyle\sum_{j=1}^{n} B_{i,j} = B_{i_k,j_k} + B_{i_k,j_{k+1}} = 0$
  - sinon la ligne est entièrement nulle
- soit $j \in [\![1, n]\!]$ et montrons que $\displaystyle\sum_{i=1}^{n} B_{i,j} = 0$
  - si j est égal à un certain j_k avec $k \in [\![2, r+1]\!]$, on a $\displaystyle\sum_{i=1}^{n} B_{i,j} = B_{i_k,j_k} + B_{i_{k-1},j_k} = 0$
  - sinon la colonne est entièrement nulle

On en déduit qu'une matrice de $\mathcal{B}_n$ est extrémale ssi c'est une matrice de permutation.

11) Utilisons le résultat admis ; soit donc $p, q \in \mathbb{N}^\star$ tels que $p + q = n + 1$, choisissons une sous-matrice A' quelconque de taille (p, q), et par l'absurde supposons que $A' = 0$.

Quitte à échanger les lignes de A, supposons que les p lignes de A' correspondent aux p premières lignes de A. De même, supposons que les q colonnes de A' correspondent aux q premières colonnes de A. On a :

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

En sommant ces égalités et en intervertissant les sommes on obtient :

$$\sum_{i=p+1}^{n} \left(\sum_{j=1}^{q} A_{i,j} \right) = q.$$

Or : $\forall i \in [\![p+1, n]\!]$, $\displaystyle\sum_{j=1}^{q} A_{i,j} \leq \sum_{j=1}^{n} A_{i,j} = 1$. Finalement on obtient : $q \leq n - p$, d'où contradiction.

12) - supposons que $\lambda_0 = 1$. Alors on aurait nécessairement : $\forall j \in [\![1,n]\!],\ A_{\sigma(j),j} = 1$. Ainsi, par le même raisonnement qu'en 9) (deuxième point), on en déduirait que $A = M_\sigma$, ce qui est exclu. Donc $\lambda_0 < 1$, et A_0 est bien définie.
    - puisque $\lambda_0 < 1$, pour montrer que A_0 est à coefficients dans $\mathbb{R}^+$ il suffit de montrer que $A - \lambda_0 M_\sigma$ est à coefficients dans $\mathbb{R}^+$. Soit donc $(i,j) \in [\![1,n]\!]^2$.
        - si $i = \sigma(j)$ alors $A_{i,j} - \lambda_0 (M_\sigma)_{i,j} = A_{\sigma(j),j} - \lambda_0 \geq 0$
        - sinon alors $A_{i,j} - \lambda_0 (M_\sigma)_{i,j} = A_{i,j} \geq 0$
    - par linéarité de la somme, on remarque que chaque ligne et chaque colonne de A_0 a pour somme :

    $$\frac{1}{1-\lambda_0} \times (1 - \lambda_0 \times 1) = 1.$$

    Donc A_0 est bistochastique.
    - soit $(i,j) \in [\![1,n]\!]^2$ tel que $A_{i,j} = 0$. Alors nécessairement $i \neq \sigma(j)$, donc $(M_\sigma)_{i,j} = 0$ puis $(A_0)_{i,j} = 0$. Donc A_0 contient au moins autant de coefficients nuls que A. De plus, λ_0 est atteint en un certain j_0 : $\lambda_0 = A_{\sigma(j_0),j_0}$. On a alors :

    $$A_{\sigma(j_0),j_0} - \lambda_0 (M_\sigma)_{\sigma(j_0),j_0} = 0$$

    donc $(A_0)_{\sigma(j_0),j_0} = 0$.
    Comme enfin $A_{\sigma(j_0),j_0} > 0$, A_0 contient au moins un coefficient nul de plus que A.

13) Pour cette question, il est plus commode de supprimer l'hypothèse que A n'est pas une matrice de permutation. On suppose donc seulement que $A \in \mathcal{B}_n$ et on raisonne par récurrence forte sur le nombre k de coefficients non nuls de A.
    - initialisation : si $k \leq n$, notons que A admet au moins un coefficient non nul par colonne (puisque la somme de chaque colonne vaut 1). Et sur une colonne donnée, il ne peut pas y avoir deux coefficients non nuls sans quoi on aurait $k \geq n+1$. Donc il y en a un seul et il est nécessairement égal à 1. Toujours par le même raisonnement, on en conclut que A est une matrice de permutation, ce qui achève l'initialisation.
    - hérédité : soit $k > n$, et supposons que l'hypothèse de récurrence est vérifiée à tous les rangs strictement inférieurs à k. Puisque $k \neq n$, A n'est pas une matrice de permutation. On peut donc construire :

    $$A_0 = \frac{1}{1-\lambda_0}(A - \lambda_0 M_\sigma)$$

    comme dans la question précédente. Par hypothèse de récurrence, elle peut s'écrire :

    $$A_0 = \tilde{\lambda}_0 \widetilde{M_0} + \ldots + \tilde{\lambda}_s \widetilde{M_s}$$

    avec $\tilde{\lambda}_0, \ldots, \tilde{\lambda}_s > 0$ et $\tilde{\lambda}_0 + \ldots + \tilde{\lambda}_s = 1$. On a ensuite :

    $$A = \lambda_0 M_\sigma + (1-\lambda_0)\tilde{\lambda}_0 \widetilde{M_0} + \ldots + (1-\lambda_0)\tilde{\lambda}_s \widetilde{M_s}$$

    Il suffit alors de poser :

    $$\begin{cases} M_0 = M_\sigma \\ \forall i \in [\![1, s+1]\!],\ M_i = \widetilde{M}_{i-1} \\ \forall i \in [\![1, s+1]\!],\ \lambda_i = (1-\lambda_0)\tilde{\lambda}_{i-1} \end{cases}$$

    Puisque $0 < \lambda_0 < 1$ les différents λ_i sont bien strictement positifs, et leur somme vaut bien 1.

14) - Puisque $\mathcal{P}_n$ est fini (et non vide), $\inf\limits_{M \in \mathcal{P}_n} \varphi(M)$ existe et c'est même un minimum. Notons-le m.
    - Soit $A \in \mathcal{B}_n$ et montrons que $\varphi(A) \geq m$.
        - si $A \in \mathcal{P}_n$ c'est par définition de m.
        - sinon, par la question précédente on peut écrire : $\varphi(M) = \displaystyle\sum_{i=0}^{s} \lambda_i \varphi(M_i) \geq \sum_{i=0}^{s} \lambda_i m = m$
    - On en déduit que $\inf\limits_{M \in \mathcal{B}_n} \varphi(M)$ existe, que c'est un minimum et qu'il vaut m. De plus, il est effectivement atteint en une matrice de permutation (par définition de m)

Partie C

15) Soit $A \in M_n(\mathbb{R})$ et $P, Q \in O_n(\mathbb{R})$.
$\|PAQ\| = \sqrt{\mathrm{tr}({}^tQ{}^tA{}^tPPAQ)} = \sqrt{\mathrm{tr}(Q^{-1}{}^tAAQ)} = \sqrt{\mathrm{tr}({}^tAA)} = \|A\|$
On exploite ici que la trace est un invariant de similitude et que ${}^tP = P^{-1}, {}^tQ = Q^{-1}$

16) D'après le théorème spectral, il existe P_1, P_2 deux matrices orthogonales et D_A, D_B deux matrices diagonales telles que $A = P_1D_A{}^tP_1$ et $B = P_2D_B{}^tP_2$.
On pose $P = {}^tP_1P_2$ qui est une matrice orthogonale car $O_n(\mathbb{R})$ est stable par produit et par transposition.
On a alors d'après la question précédente :
$\|A - B\| = \|P_1D_A{}^tP_1 - P_2D_B{}^tP_2\| = \|{}^tP_1 \times (P_1D_A{}^tP_1 - P_2D_B{}^tP_2) \times P_2\| \boxed{= \|D_AP - PD_B\|}$

17)
- Tous les coefficients de R sont positifs (carrés de réels) et pour tout $k \in [\![1, n]\!], \sum_{i=1}^{n} R_{k,i} = \sum_{i=1}^{n} R_{i,k} = 1$ car les lignes et les colonnes de P sont des vecteurs unitaires de $\mathbb{R}^n$.
Donc $\boxed{R \text{ est une matrice bistochastique}}$
- On peut noter $\lambda_1(A), ..., \lambda_n(B)$ et $\lambda_1(B), ..., \lambda_n(B)$ les coefficients diagonaux des matrices D_A et D_B car ce sont les valeurs propres de A et de B (comptées avec leur ordre de multiplicité).
Le coefficient (i, j) de la matrice $D_AP - PD_B$ est $(\lambda_i(A) - \lambda_j(B))P_{i,j}$.
Donc $\|A - B\|^2 = \sum_{1 \leq i,j \leq n} (\lambda_i(A) - \lambda_j(B))^2(P_{i,j})^2 \boxed{= \sum_{1 \leq i,j \leq n} |\lambda_i(A) - \lambda_j(B)|^2 R_{i,j}}$
Notons que pour un réel x, $x^2 = |x|^2$

18) On va ici exploiter la question 14)
On pose φ définie sur $M_n(\mathbb{R})$ par $\varphi(M) = \sum_{1 \leq i,j \leq n} |\lambda_i(A) - \lambda_j(B)|^2 M_{i,j}$
φ est une forme linéaire de $M_n(\mathbb{R})$ et $\|A - B\|^2 = \varphi(R)$ avec $R \in B_n$.
D'après la question 14), il existe $\sigma \in S_n$ tel que $\|A - B\|^2 \geq \varphi(M_\sigma) = \sum_{j=1}^{n} |\lambda_{\sigma(j)}(A) - \lambda_j(B)|^2$.
De ceci, il vient que $\boxed{\min_{\sigma \in S_n} \sum_{j=1}^{n} |\lambda_{\sigma(j)}(A) - \lambda_j(B)|^2 \leq \|A - B\|^2}$

19)
- On considère ici que $a_1, ..., a_n$ et $b_1, ..., b_n$ sont distincts deux à deux.
- Soit X et Y deux variables aléatoires de V vérifiant $X \sim P_1$ et $Y \sim P_2$.
Alors $nE(|X - Y|^2) = \sum_{1 \leq i,j \leq n} |a_i - b_{(j)}|^2 \times nP(X = a_i, Y = b_{(j)})$.
On note $R \in M_n(\mathbb{R})$ la matrice de coefficient $R_{i,j} = nP(X = a_i, Y = b_{(j)})$.
Alors pour tout $k \in [\![1, n]\!], \sum_{j=1}^{n} nP(X = a_k, Y = b_{(j)}) = nP(X = a_k) = 1$ et $\sum_{i=1}^{n} nP(X = a_i, Y = b_{(k)}) = nP(Y = b_{(k)}) = 1$.
On a aussi $R_{i,j} \geq 0$. Donc $R \in B_n$.
D'après la question 14), il existe $\sigma \in S_n$ tel que
$nE(|X - Y|^2) \geq \sum_{j=1}^{n} |a_{\sigma(j)} - b_{(j)}|^2 = \sum_{k=1}^{n} (a_k^2 + b_{(k)}^2) - 2\sum_{j=1}^{n} a_{\sigma(j)}b_{(j)}$.
Il nous reste à montrer que $\sum_{j=1}^{n} a_{\sigma(j)}b_{(j)} \leq \sum_{j=1}^{n} a_{(j)}b_{(j)}$ qui est une inégalité du réordonnement.
σ étant une permutation "quelconque" et $a_1, ..., a_n$ n'étant pas ordonné a priori, on se ramène à démontrer que $\sum_{j=1}^{n} a_j b_{(j)} \leq \sum_{j=1}^{n} a_{(j)}b_{(j)}$.
On peut démontrer cette propriété par récurrence sur n.
Pour $n = 1$, l'initialisation est évidente.
Supposons $n \geq 2$ et la propriété vérifiée au rang $n - 1$.
Si $a_{(1)} = a_1$, alors il reste à démontrer que $\sum_{j=2}^{n} a_j b_{(j)} \leq \sum_{j=2}^{n} a_{(j)}b_{(j)}$ qui est vérifié par hypothèse de récurrence.

Sinon, il existe $k \geq 2$ tel que $a_{(1)} = a_k$.
L'inégalité $(a_1 - a_{(1)})(b_{(k)} - b_{(1)}) \geq 0$ est vérifiée et donne en développant
$a_1 b_{(k)} + a_{(1)} b_{(k)} \leq a_{(1)} b_{(1)} + a_1 b_{(k)}$ ce qui permet de se ramener au cas précédent et de conclure : le résultat est vrai au rang n.

Ainsi, on a démontré que $nE(|X-Y|^2) \geq \sum_{j=1}^{n} |a_{(j)} - b_{(j)}|^2$

ce qui assure que $\boxed{d^2(P_1, P_2) \geq \frac{1}{n}\sum_{j=1}^{n} |a_{(j)} - b_{(j)}|^2}$

- On considère un couple de variable aléatoire (X, Y) dont la loi est donnée par $P(X = a_{(i)}, Y = b_{(j)}) = \frac{\delta_{i,j}}{n}$.
  On a alors directement $X \sim P_1$ et $Y \sim P_2$.
  De plus, $E(|X-Y|^2) = \sum_{j=1}^{n} \frac{1}{n} |a_{(j)} - b_{(j)}|^2$ ce qui démontre que $\boxed{d^2(P_1, P_2) \leq \frac{1}{n}\sum_{j=1}^{n} |a_{(j)} - b_{(j)}|^2}$

- Conclusion : $\boxed{d^2(P_1, P_2) = \frac{1}{n}\sum_{j=1}^{n} |a_{(j)} - b_{(j)}|^2}$

- On reprend la question 17) et ses notations.
  On peut alors définir un couple de variables aléatoires (X, Y) tel que $P(X = a_i, Y = b_j) = \frac{1}{n} R_{i,j}$
  Ainsi définies, X et Y suivent respectivement les lois P_1 et P_2 donc
  $d^2(P_1, P_2) \leq E(|X-Y|^2) = \sum_{1 \leq i,j \leq n} \frac{1}{n} R_{i,j} |a_i - b_j|^2 = \frac{1}{n} \|A - B\|^2$
  Finalement, $\boxed{n \times d^2(P_1, P_2) \leq \|A - B\|^2}$
  Remarque : Ce n'est pas tout à fait une déduction de l'égalité précédente, mais tant pis.