Centrale Maths 1 MP 2021 — Corrigé

Ce corrigé est proposé par Angèle Niclas (ENS Lyon) ; il a été relu par Tristan Poullaouec (professeur en CPGE) et Benjamin Monmege (enseignant-chercheur à l'université).

L'objet de ce problème est l'étude de la répartition des valeurs propres des matrices symétriques réelles aléatoires de grande taille et la preuve de la loi du demi-cercle. Cette loi affirme que pour toute matrice symétrique réelle constituée de variables aléatoires indépendantes identiquement distribuées, d'espérance nulle et de variance égale à 1, les valeurs propres normalisées par un facteur $1/\sqrt{n}$ se répartissent suivant un profil en demi-cercle de rayon 2.

Le sujet est composé de quatre parties. Les deux premières sont indépendantes et s'attachent à montrer des résultats intermédiaires utiles dans les deux dernières parties.

- Dans la partie I, on montre l'inégalité de Hoffman-Wielandt liant l'écart entre les valeurs propres de deux matrices symétriques et la norme de la différence des deux matrices. Pour cela, on est amené à minimiser une fonction sur l'ensemble des matrices bistochastiques. Cette partie utilise les chapitres sur la réduction des matrices et sur les espaces vectoriels normés.
- La partie II porte sur le dénombrement des mots bien parenthésés pour retrouver les nombres de Catalan. Si le début fait appel à quelques techniques classiques de dénombrement, il faut surtout maîtriser le cours sur les séries entières.
- Dans la partie suivante, on montre la loi du demi-cercle dans le cas simplifié où les variables sont uniformément bornées. On commence par quelques calculs d'intégrales puis on fait des dénombrements plus complexes que dans la partie II. Après avoir prouvé la loi du demi-cercle pour les polynômes, on l'étend aux fonctions bornées. Cette partie est assez longue et demande une bonne compréhension de la démarche car les questions s'appuient les unes sur les autres.
- La partie IV étend les résultats de la partie III au cas général. Pour cela, on introduit de nouvelles variables tronquées qui vérifient les hypothèses de la partie III, et avec les résultats de la partie I on prouve finalement la loi du demi-cercle dans le cas général.

Ce problème est long, assez difficile pour un sujet de la banque Centrale-Supélec et a dû déconcerter beaucoup de candidats par son originalité. La partie II est une bonne révision sur les séries entières, tandis que les parties II et III permettent de travailler le dénombrement. Le sujet dans sa globalité fait appel à de nombreux chapitres du programme de deuxième année. Pour le réussir, il fallait prendre le temps de comprendre le chemin suivi par l'énoncé pour démontrer la loi du demi-cercle.

Indications

Partie I

2 Appliquer le théorème spectral à A et B, puis utiliser le résultat de la question 1.

4 Montrer que $\mathscr{B}_n(\mathbb{R})$ est un compact de $\mathscr{M}_n(\mathbb{R})$. On pourra utiliser la caractérisation séquentielle des fermés puis majorer $\|\mathrm{M}\|_{\mathrm{F}}$ pour tout $\mathrm{M} \in \mathscr{B}_n(\mathbb{R})$.

6 Construire la matrice recherchée par récurrence en annulant un par un les coefficients $m_{i,j}$ et $m_{j,i}$ pour $j \in [\![\, i+1\,;n\,]\!]$. Pour cela, on pourra se servir de la matrice définie en question 5.

7 Pour tout $\mathrm{M} \in \mathscr{B}_n(\mathbb{R})$, utiliser la question 6 pour construire par récurrence une suite de matrices $\mathrm{M}^{(k)}$ se rapprochant de I_n et vérifiant $f(\mathrm{M}^{(k)}) \leqslant f(\mathrm{M})$.

8 Utiliser le résultat des questions 2 et 3 puis montrer que $\mathrm{M} = ({p_{i,j}}^2)_{1\leqslant i,j\leqslant n}$ est une matrice de $\mathscr{B}_n(\mathbb{R})$ pour appliquer le résultat de la question 7.

Partie II

10 Utiliser le lemme d'Abel pour minorer le rayon de convergence de la série entière $\sum \mathrm{C}_k x^k$.

12 Utiliser le résultat de la question 11 et la formule du produit de Cauchy pour développer $(\mathrm{F}(x))^2$ en série entière.

13 Dériver la relation de la question 12 puis raisonner par l'absurde.

14 Utiliser la relation de la question 12 pour reconnaître une équation du second degré en $\mathrm{F}(x)$, puis le résultat de la question 13 pour choisir la racine réelle à conserver.

15 Utiliser le développement en série entière de $(1+x)^\alpha$ puis lier le produit des nombres impairs au produit des nombres pairs pour faire apparaître les factorielles demandées.

16 Développer en série entière chaque membre du résultat de la question 14 en utilisant le développement en série entière de la question 15.

Partie III

17 Remarquer que la fonction intégrée est impaire.

19 Intégrer la fonction $x \mapsto x\sqrt{4-x^2}$ et dériver la fonction $x \mapsto x^{2k+1}$ pour obtenir une relation liant m_{2k+2} et m_{2k}.

20 Utiliser le résultat des questions 18 et 19 pour exprimer m_{2k} en fonction de k puis utiliser le résultat des questions 16 et 17.

21 Calculer par récurrence le coefficient $({\mathrm{M}_n}^k)_{i,j}$ pour tout $(i,j) \in [\![\, 1\,;n\,]\!]^2$.

23 Appliquer le résultat de la question 22 à $\ell = (k+1)/2$.

24 Utiliser l'indépendance mutuelle des $(\mathrm{X}_{ij})_{1\leqslant i\leqslant j\leqslant n}$ et le fait que $\mathbb{E}(\mathrm{X}_{ij}) = 0$ pour tout $(i,j) \in [\![\, 1\,;n\,]\!]^2$.

25 Remarquer que tout cycle comportant ℓ sommets comporte au moins $\ell - 1$ arêtes distinctes.

26 Utiliser le résultat de la question 21 puis séparer la somme entre $\mathscr{A}_k$ et $\mathscr{C}_k$. Utiliser le résultat de la question 24 pour traiter la somme sur $\mathscr{A}_k$ et le résultat des questions 23 et 25 pour la somme sur $\mathscr{C}_k$.

28 Étant donné un mot bien parenthésé, remarquer que tout cycle correspondant est associé à une unique liste ordonnée des arêtes distinctes, puis que toute liste ordonnée des arêtes distinctes est associée à une unique liste ordonnée des sommets distincts.

29 Réutiliser le raisonnement de la question 26 et séparer les cycles de $\mathscr{B}_k$ selon leur nombre de sommets distincts ℓ. Utiliser le résultat de la question 23 si $\ell < k/2+1$ et le résultat de la question 28 si $\ell = k/2+1$.

30 Utiliser le résultat des questions 20, 26 et 29.

32 Appliquer le résultat des questions 29 et 31 puis le résultat de la question 10 pour estimer $\mathrm{C}_{p+q}/\mathrm{A}^{p+2q}$ quand $q \to +\infty$.

34 Utiliser le résultat des questions 32 et 33.

35 Approcher uniformément sur un intervalle $[-\mathrm{A}\,;\mathrm{A}]$ la fonction f par un polynôme, puis utiliser le résultat des questions 30 et 34.

Partie IV

36 Utiliser la caractérisation séquentielle de la limite et le caractère sommable de la famille $(x\mathbb{P}(\mathrm{X}=x))_{x\in\mathrm{X}(\Omega)}$.

37 Appliquer le résultat de la question 36 à X_{ij} et $(\mathrm{X}_{ij})^2$.

38 Utiliser le résultat de la question 37 pour montrer que les variables $\widehat{\mathrm{X}}_{ij}$ sont bien définies.

40 Utiliser le résultat de la question 39 et l'inégalité $(a+b)^2 \leqslant 2a^2+2b^2$ pour tout $(a,b)\in\mathbb{R}^2$. Conclure en utilisant le résultat de la question 36.

41 S'aider de la concavité de la fonction $x \mapsto \sqrt{x}$ pour faire apparaître la quantité $\sum\limits_{i=1}^{n}(\Lambda_{i,n}-\widehat{\Lambda}_{i,n})^2$ puis utiliser le résultat de la question 8.

42 Calculer la limite du résultat de la question 41 à l'aide de la question 40. Montrer ensuite que les $\widehat{\mathrm{X}}_{ij}$ vérifient toutes les hypothèses de la partie III à l'aide du résultat de la question 38 pour ensuite appliquer le résultat de la question 35.

43 Prouver en introduisant une fonction lipschitzienne bien choisie que

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{|\Lambda_{i,n}|\geqslant 3}\right) \xrightarrow[n\to\infty]{} 0$$

Effectuer ensuite un raisonnement similaire à celui de la question 35 en remplaçant le polynôme P par une fonction lipschitzienne qui vaut P sur $[-3\,;3]$.

I. Inégalité de Hoffman-Wielandt

1 Soient $\mathrm{M} \in \mathscr{M}_n(\mathbb{R})$ et $(\mathrm{P}, \mathrm{Q}) \in \mathscr{O}_n(\mathbb{R})^2$, alors

$$\begin{aligned}
\|\mathrm{PMQ}\|_{\mathrm{F}} &= \sqrt{\mathrm{Tr}\,(\mathrm{PMQ}(\mathrm{PMQ})^{\mathrm{T}})} \\
&= \sqrt{\mathrm{Tr}\,(\mathrm{PMQQ^TM^TP^T})} && ((\mathrm{M_1M_2})^{\mathrm{T}} = \mathrm{M_2}^{\mathrm{T}}\mathrm{M_1}^{\mathrm{T}}) \\
&= \sqrt{\mathrm{Tr}\,(\mathrm{PMM^TP^T})} && (\mathrm{QQ^T} = \mathrm{I}_n) \\
&= \sqrt{\mathrm{Tr}\,(\mathrm{P^TPMM^T})} && (\mathrm{Tr}\,(\mathrm{M_1M_2}) = \mathrm{Tr}\,(\mathrm{M_2M_1})) \\
&= \sqrt{\mathrm{Tr}\,(\mathrm{MM^T})} && (\mathrm{P^TP} = \mathrm{I}_n)
\end{aligned}$$

$$\boxed{\|\mathrm{PMQ}\|_{\mathrm{F}} = \|\mathrm{M}\|_{\mathrm{F}}}$$

2 La matrice A est réelle et symétrique et le théorème spectral permet d'affirmer qu'elle est diagonalisable dans une base orthonormale. De même, la matrice B est réelle et symétrique donc diagonalisable dans une autre base orthonormale. Ainsi, il existe $(\mathrm{Q}, \mathrm{R}) \in \mathscr{O}_n(\mathbb{R})^2$ tel que

$$\mathrm{A} = \mathrm{QD_AQ^T} \qquad \text{et} \qquad \mathrm{B} = \mathrm{RD_BR^T}$$

Remarquons alors que

$$\begin{aligned}
\|\mathrm{A} - \mathrm{B}\|_{\mathrm{F}} &= \|\mathrm{QD_AQ^T} - \mathrm{RD_BR^T}\|_{\mathrm{F}} \\
&= \|\mathrm{Q}(\mathrm{D_AQ^T} - \mathrm{Q^TRD_BR^T})\|_{\mathrm{F}} && (\mathrm{Q}^{-1} = \mathrm{Q^T}) \\
&= \|\mathrm{Q}(\mathrm{D_AQ^TR} - \mathrm{Q^TRD_B})\mathrm{R^T}\|_{\mathrm{F}} && ((\mathrm{R^T})^{-1} = \mathrm{R}) \\
\|\mathrm{A} - \mathrm{B}\|_{\mathrm{F}} &= \|\mathrm{D_AQ^TR} - \mathrm{Q^TRD_B}\|_{\mathrm{F}} && (\text{question } 1)
\end{aligned}$$

Définissons $\mathrm{P} = \mathrm{Q^TR}$. Comme $\mathrm{Q^T}$ et R appartiennent à $\mathscr{O}_n(\mathbb{R})$, alors $\mathrm{P} \in \mathscr{O}_n(\mathbb{R})$ et

$$\boxed{\text{Il existe une matrice orthogonale P telle que } \|\mathrm{A} - \mathrm{B}\|_{\mathrm{F}}{}^2 = \|\mathrm{D_AP} - \mathrm{PD_B}\|_{\mathrm{F}}{}^2.}$$

3 Posons $\mathrm{M} = \mathrm{D_AP} - \mathrm{PD_B}$. En utilisant le résultat de la question 2, on a

$$\|\mathrm{A} - \mathrm{B}\|_{\mathrm{F}}{}^2 = \|\mathrm{M}\|_{\mathrm{F}}{}^2 = \mathrm{Tr}\,(\mathrm{MM^T}) = \sum_{i=1}^{n} (\mathrm{MM^T})_{i,i} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{M}_{i,j}(\mathrm{M^T})_{j,i} = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{M}_{i,j}{}^2$$

Soit $(i, j) \in [\![1\,;n]\!]^2$, remarquons que

$$\mathrm{M}_{i,j} = (\mathrm{D_AP} - \mathrm{PD_B})_{i,j} = \sum_{k=1}^{n} [(\mathrm{D_A})_{i,k}p_{k,j} - p_{i,k}(\mathrm{D_B})_{k,j}]$$

Les matrices $\mathrm{D_A}$ et $\mathrm{D_B}$ sont diagonales : pour tout $(r, s) \in [\![1\,;n]\!]^2$ avec $r \neq s$ alors $(\mathrm{D_A})_{r,s} = (\mathrm{D_B})_{r,s} = 0$. En particulier,

$$\mathrm{M}_{i,j} = (\mathrm{D_A})_{i,i}p_{i,j} - p_{i,j}(\mathrm{D_B})_{j,j} = (\lambda_i(\mathrm{A}) - \lambda_j(\mathrm{B}))p_{i,j}$$

On conclut que $\boxed{\|\mathrm{A} - \mathrm{B}\|_{\mathrm{F}}{}^2 = \sum\limits_{1 \leqslant i,j \leqslant n} p_{i,j}{}^2(\lambda_i(\mathrm{A}) - \lambda_j(\mathrm{B}))^2}$