Banque BCPST Inter-ENS / ENPC - Session 2019

Rapport sur l'épreuve écrite de mathématiques
(épreuve comptant uniquement pour l'admission)

Écoles concernées : ENS (Paris) - ENS de Lyon - ENS de Paris-Saclay - ENPC

Cœfficients (en pourcentage du total d'admission) :

ENS de Paris-Saclay : 6.2 %

ENS de Lyon : 6.6 % pour les deux options

ENS (Paris) : 11.3 % pour les deux options

ENPC : 20 %

Membre du jury : François Bolley

—— o ——

L'épreuve est composée de trois parties. Elle propose l'étude d'un problème d'optimisation, dans un cadre discret dans la partie 1, et dans un cadre continu dans la partie 2. Une partie 3 propose de comparer les résultats et les méthodes des parties 1 et 2.

Le sujet a permis de tester les candidats sur leur aisance à manipuler les techniques et outils classiques d'analyse, probabilités et calcul matriciel au programme des classes préparatoires BCPST. Les notes obtenues par les candidats admissibles sont comprises entre 1.50 et 19 sur 20, avec une moyenne de 8.80 et un écart type de 3.80. La précision de la rédaction et l'honnêteté des candidats dans leurs démonstrations ont été récompensées. Inversement, les rares candidats ayant tenté d'imposer leurs résultats par des affirmations non justifiées ont été sanctionnés.

Dans la **partie 1** on considère un entier naturel n plus grand que 2 fixé, et $2n$ nombres réels $x_1 \leqslant \cdots \leqslant x_n$ et $y_1 \leqslant \cdots \leqslant y_n$. On se propose de montrer que, parmi les matrices bistochastiques M de taille (n, n) et de cœfficients M_{ij}, la quantité

$$C(M) = \sum_{i=1}^{n} \sum_{j=1}^{n} \big(y_j - x_i\big)^2 M_{ij}$$

est minimale en la matrice identité; si de plus $x_1 < \cdots < x_n$ et $y_1 < \cdots < y_n$, alors la matrice identité est l'unique matrice bistochastique rendant minimale la quantité $C(M)$.

Pour cela on procède progressivement, par des exemples en questions 1.3 et 1.4 et les cas particuliers des matrices de permutation en question 1.5 et bistochastiques symétriques en question 1.6, le cas général étant traité en question 1.7. Pour cela on montre notamment en question 1.1 qu'une matrice de permutation est une matrice bistochastique; on remarque de plus que, pour la matrice M de la permutation p de $[\![1, n]\!]$, c'est-à-dire pour la matrice de cœfficients M_{ij} tels que $M_{ij} = 1$ si $j = p(i)$ et $M_{ij} = 0$ sinon, on a

$$C(M) = \sum_{i=1}^{n} \big(y_{p(i)} - x_i\big)^2.$$

La partie 1 débute par deux questions préliminaires qui ont été résolues par un grand nombre de candidats. Dans la **question 1.1** on montre donc qu'une matrice de permutation est une matrice bistochastique. La plupart des candidats ont bien vérifié que la somme des cœfficients de la matrice, par ligne ou colonne, est égale à 1, mais on s'attendait à ce que les candidats mentionnent que les cœfficients sont positifs.

La **question 1.2**, préliminaire également, demande de montrer une égalité qui est simple à vérifier, en développant par exemple la double somme du membre de droite et en changeant d'indices, et sera utile dans les questions 1.5 et 1.7. Cette égalité peut être interprétée comme une formule d'intégration par parties discrète, ainsi qu'on le verra dans la partie 3.

Dans la **question 1.3** on montre, dans un exemple numérique pour $n = 3$ et des $x_1 < x_2 < x_3$ et $y_1 < y_2 < y_3$, que la permutation identité rend minimale la quantité $\sum_{i=1}^{3} \left(y_{p(i)} - x_i\right)^2$ parmi toutes les permutations de $[\![1, 3]\!]$, et que c'est la seule permutation rendant minimale cette quantité; autrement dit, la matrice identité, qui est bien une matrice de permutation (pour n'importe quel n), est l'unique matrice rendant minimale la quantité $C(M)$ parmi toutes les matrices de permutation M de taille $(3, 3)$. Pour cela on détermine explicitement les $3! = 6$ permutations de $[\![1, 3]\!]$ et on calcule la valeur de la somme pour chacune de ces permutations; on obtient les valeurs $4, 6, 10, 14, 18, 20$, la valeur 4 étant (uniquement) obtenue pour la matrice identité. Cette question a été bien traitée par un grand nombre de candidats.

Dans la **question 1.4** on montre, dans un exemple numérique pour $n = 2$ et des $x_1 < x_2$ et $y_1 < y_2$, que la matrice identité, qui est bien une matrice bistochastique (pour n'importe quel n), rend minimale la quantité $C(M)$ parmi toutes les matrices bistochastiques de taille $(2, 2)$, et que c'est la seule telle matrice rendant minimale cette quantité. Pour cela on vérifie tout d'abord que $C(M) = 10 + 6r$ pour la matrice

$$M = \begin{pmatrix} 1-r & r \\ r & 1-r \end{pmatrix}.$$

On doit ensuite montrer qu'une matrice de taille $(2, 2)$ est une matrice bistochastique si et seulement si elle est de cette forme pour un $r \in [0, 1]$, ce qui n'a été effectué que par peu de candidats. Le sujet indique, sans même le justifier, qu'une telle matrice est bistochastique dès que $r \in [0, 1]$, mais encore faut-il s'en assurer, même brièvement, et le mentionner dans le raisonnement, et surtout montrer la réciproque. La quantité $10 + 6r$ atteignant sur $[0, 1]$ son unique minimum en $r = 0$, on en déduit alors que $C(M)$ est minimal en une unique matrice, à savoir la matrice bistochastique paramétrée par $r = 0$, qui est la matrice identité. (En lien avec la question 1.6 on peut remarquer que pour $n = 2$ les matrices bistochastiques sont toutes symétriques.)

Dans les questions 1.5 à 1.7 on fixe un entier n plus grand que 2 et $2n$ nombres réels quelconques $x_1 \leqslant \cdots \leqslant x_n$ et $y_1 \leqslant \cdots \leqslant y_n$.

Dans la **question 1.5** on montre que la permutation identité rend minimale la quantité $\sum_{i=1}^{n} \left(y_{p(i)} - x_i\right)^2$ parmi toutes les permutations p de $[\![1, n]\!]$; autrement dit, la matrice identité I rend minimale la quantité $C(M)$ parmi toutes les matrices de permutation M de taille (n, n). Comme

$$C(I) = \sum_{i=1}^{n} \left(y_i - x_i\right)^2,$$

ceci équivaut à montrer que

$$\sum_{i=1}^{n} \left(y_{p(i)} - x_i\right)^2 \geqslant \sum_{i=1}^{n} \left(y_i - x_i\right)^2 \qquad (1)$$

pour toute permutation p de $[\![1, n]\!]$.

Pour cela on développe

$$\sum_{i=1}^{n} \left(y_{p(i)} - x_i\right)^2 = \sum_{i=1}^{n} y_{p(i)}^2 + \sum_{i=1}^{n} x_i^2 - 2 \sum_{i=1}^{n} x_i \, y_{p(i)}.$$

On remarque alors que les n entiers $p(i)$ décrivent $[\![1,n]\!]$ quand i décrit $[\![1,n]\!]$ puisque p est une permutation, de sorte que

$$\sum_{i=1}^{n} y_{p(i)}^2 = \sum_{i=1}^{n} y_i^2.$$

L'égalité de la question 1.2 avec $b_i = x_i$ et $c_i = y_{p(i)}$ assure alors que

$$\sum_{i=1}^{n} x_i\, y_{p(i)} = x_1 \sum_{j=1}^{n} y_{p(j)} + \sum_{i=2}^{n} (x_i - x_{i-1}) \sum_{j=i}^{n} y_{p(j)}.$$

Or d'une part $\displaystyle\sum_{i=1}^{n} y_{p(i)} = \sum_{i=1}^{n} y_i$ pour la même raison que ci-dessus pour la somme des carrés. D'autre part, pour $i \geqslant 2$ les indices $p(j)$ pour $j = i, \dots, n$ sont $n-i+1$ indices distincts parmi $1, \dots, n$ puisque p est bijective, donc les réels $y_{p(j)}$ pour $j = i, \dots, n$ sont $n-i+1$ réels parmi $y_1, \dots, y_n$ pour des indices tous distincts; par conséquent leur somme est plus petite que la somme des $n-i+1$ plus grands termes parmi $y_1, \dots, y_n$, qui sont $y_i, \dots, y_n$ puisque les y_j sont croissants. Autrement dit $\displaystyle\sum_{j=i}^{n} y_{p(j)} \leqslant \sum_{j=i}^{n} y_j$. Il faut alors mentionner le fait que $x_i - x_{i-1} \geqslant 0$ pour finalement en déduire l'inégalité (1). Cette question a été résolue et rédigée proprement par une moitié des candidats.

Dans la **question 1.6** on montre que la matrice identité I, qui est bien une matrice bistochastique symétrique, rend minimale la quantité $C(M)$ parmi toutes les matrices bistochastiques symétriques M de taille (n,n). Comme de nouveau $C(I) = \displaystyle\sum_{i=1}^{n} \big(y_i - x_i\big)^2$, ceci équivaut à montrer que

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \big(y_j - x_i\big)^2 M_{ij} \geqslant \sum_{i=1}^{n} \big(y_i - x_i\big)^2 \qquad (2)$$

pour toute matrice bistochastique symétrique M.

Pour cela on vérifie tout d'abord que pour tout i, j on a

$$x_i\, y_i + x_j\, y_j \geqslant x_i\, y_j + x_j\, y_i. \qquad (3)$$

Or, par exemple, cette inégalité est équivalente, par développement, à l'inégalité

$$(x_i - x_j)(y_i - y_j) \geqslant 0$$

qui est vraie par croissance des deux familles (x_k) et (y_k).

Pour en déduire (2), on multiplie alors l'inégalité (3) par M_{ij} et on somme sur i et j, ce qui donne

$$\sum_{i=1}^{n} x_i\, y_i \sum_{j=1}^{n} M_{ij} + \sum_{j=1}^{n} x_j\, y_j \sum_{i=1}^{n} M_{ij} \geqslant \sum_{i,j=1}^{n} x_i\, y_j\, M_{ij} + \sum_{i,j=1}^{n} x_j\, y_i\, M_{ij}.$$

Or la matrice M est bistochastique donc les sommes du membre de gauche sont toutes deux égales à $\displaystyle\sum_{i=1}^{n} x_i\, y_i$.

De plus $M_{ij} = M_{ji}$ par symétrie de la matrice M, donc en échangeant les indices muets i et j on remarque que les deux sommes du membre de droite sont égales. On obtient alors (2) en développant

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \big(y_j - x_i\big)^2 M_{ij} = \sum_{j=1}^{n} y_j^2 \sum_{i=1}^{n} M_{ij} + \sum_{i=1}^{n} x_i^2 \sum_{j=1}^{n} M_{ij} - 2 \sum_{i,j=1}^{n} x_i\, y_j\, M_{ij} = \sum_{j=1}^{n} y_j^2 + \sum_{i=1}^{n} x_i^2 - 2 \sum_{i,j=1}^{n} x_i\, y_j\, M_{ij}.$$

Ainsi, d'après ce qui précède,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \big(y_j - x_i\big)^2 M_{ij} \geqslant \sum_{j=1}^{n} y_j^2 + \sum_{i=1}^{n} x_i^2 - 2 \sum_{i,j=1}^{n} x_i\, y_i = \sum_{i=1}^{n} \big(y_i - x_i\big)^2,$$

c'est-à-dire (2). Si les questions 1.6.1 et 1.6.3 ont été résolues dans de nombreuses copies, la question 1.6.2 ne l'a été de manière satisfaisante que dans les meilleures copies.

Dans la **question 1.7** on considère le cas général, et on montre que la matrice identité I, qui est bien une matrice bistochastique, rend minimale la quantité $C(M)$ parmi toutes les matrices bistochastiques M de taille (n, n), étendant ainsi les résultats des exemples et cas particuliers précédents. Pour cela on suit la démarche de la question 1.5, mais dans le cas plus général de matrices qui ne sont pas a priori de permutation. Ainsi, l'inégalité clé

$$\sum_{j=i}^{n} y_{p(j)} \leqslant \sum_{j=i}^{n} y_j \qquad (4)$$

de la question 1.5.1 est-elle généralisée dans la question 1.7.1 en

$$\sum_{j=i}^{n} \sum_{k=1}^{n} M_{jk} y_k \leqslant \sum_{j=i}^{n} y_j$$

qui est bien l'inégalité (4) si M est la matrice de la permutation p.

Dans la question 1.7.2 on déduit facilement de 1.7.1 que la matrice identité I, qui est bien une matrice bistochastique, rend minimale la quantité $C(M)$ parmi toutes les matrices bistochastiques M de taille (n, n), de même que le résultat de la question 1.5.2 (resp. 1.6.3) est la simple conséquence de la question 1.5.1 (resp. 1.6.2).

Dans la question 1.7.3 on montre que la matrice identité est la seule matrice rendant minimale la quantité $C(M)$ dès que les familles (x_k) et (y_k) sont strictement croissantes. Pour résoudre cette question on pouvait, en guise d'entraînement, se poser la question et la résoudre dans le cas plus simple de la question 1.5. Une solution est la suivante : on se donne une matrice bistochastique M telle que $C(M) = C(I)$, et on montre que nécessairement $M = I$. L'égalité $C(M) = C(I)$ impose d'avoir égalité dans toutes les inégalités qui ont été sommées en question 1.7.1 en vue d'obtenir l'inégalité $C(M) \geqslant C(I)$. On utilise alors le fait que les familles (x_k) et les (y_k) sont strictement croissantes pour, par exemple, montrer que nécessairement $M_{nn} = 1$, puis $M_{n-1,n} = 0$, puis $M_{n-1,j} = 0$ sauf pour $j = n-1$, et $M_{n-1,n-1} = 1$. De proche en proche on en déduit que $M_{ij} = 0$ si $i \neq j$ et $M_{ii} = 1$, c'est-à-dire que M est la matrice identité. Seules quelques rares excellentes copies ont su mener à bien ces délicates questions 1.7.1 et 1.7.3, qui s'appuyaient notamment sur une bonne compréhension du cas plus simple de la question 1.5.

La **partie 1** propose l'étude d'un problème d'optimisation, dans un cadre discret, portant sur des permutations et des matrices. La **partie 2** propose l'étude d'une version continue de ce problème, ainsi qu'on le verra dans la partie 3, portant sur des fonctions de $[0, 1]$ dans $[1, 2]$.

La **question 2.1** propose l'étude d'un exemple, dans lequel on peut résoudre complètement le problème d'optimisation : on montre l'existence et l'uncité de l'optimiseur, et on le donne explicitement. La question 2.1.1 demande de montrer l'inégalité de Cauchy-Schwarz $\big(\mathbb{E}[g(U)]\big)^2 \leqslant \mathbb{E}[g(U)^2]$. Ainsi que de nombreux candidats l'ont remarqué, c'est une conséquence simple du fait que la variable aléatoire $\big(g(U) - \mathbb{E}[g(U)]\big)^2$ est positive, donc que son espérance, qui est la variance de la variable $g(U)$, est positive. Le sujet demande alors de montrer qu'il y a égalité si et seulement si la fonction g est de plus constante. Le sens réciproque a été obtenu par de nombreux candidats, alors que le sens direct, qui nécessite un petit mais délicat raisonnement, n'a été montré proprement que dans de rares copies. Le programme de mathématiques contient néanmoins tous les outils nécessaires, notamment le fait qu'une fonction continue positive sur $[0, 1]$, d'intégrale nulle, est nulle. La question 2.1.2 s'en déduit en posant notamment $g(x) = S(x) - x$. Là encore, cette question délicate n'a été traitée de manière satisfaisante que dans quelques excellentes copies.

Dans la question 2.4 on montrera que le problème d'optimisation de cette partie 2 admet la fonction F_V^{-1} comme solution. À cette fin, dans la **question 2.2** on montre notamment que cette fonction F_V^{-1} est bien définie et est admissible, c'est-à-dire qu'elle appartient bien à l'ensemble $\mathscr{S}$ sur lequel on optimise. L'unicité de F_V^{-1} parmi les fonctions strictement croissantes de $\mathscr{S}$ ne sert pas directement dans la suite du sujet; elle se démontre simplement par une variante de la démonstration de l'appartenance de F_V^{-1} à l'ensemble $\mathscr{S}$, par exemple en remarquant que $F_V(S(x)) = x$ pour toute fonction S strictement croissante de $\mathscr{S}$.

Dans la **question 2.3** on considère une fonction S de $\mathscr{S}$, et on montre plusieurs égalités et inégalités qui, combinées, assureront en question 2.4 que

$$\mathbb{E}\big[(S(U) - U)^2\big] \geqslant \mathbb{E}\big[(F_V^{-1}(U) - U)^2\big] \qquad (5)$$

Ceci étant vrai pour pour toute fonction S de $\mathscr{S}$, et comme de nouveau F_V^{-1} est une fonction de $\mathscr{S}$, ceci signifie bien que la fonction F_V^{-1} minimise sur $\mathscr{S}$ la quantité $\mathbb{E}\big[(S(U) - U)^2\big]$.

Le résultat de la question 2.3.1 s'obtient par exemple par une simple intégration par parties, comme l'ont bien observé de nombreux candidats; plusieurs copies l'ont également astucieusement obtenu comme une conséquence du théorème de Fubini énoncé dans la question 2.3.2. L'égalité de la question 2.3.2 s'obtient facilement à partir de l'indication, ainsi que l'ont bien compris un quart des candidats : encore fallait-il avoir l'esprit frais en cette fin d'épreuve pour analyser cette intégrale double. La question 2.3.3, relativement délicate, peut être résolue en écrivant

$$\mathbb{P}[U \geqslant u, S(U) \geqslant y] = \mathbb{P}[E_1 \cap E_2] \leqslant \min\{\mathbb{P}[E_1], \mathbb{P}[E_2]\} = \min\{1 - u, 1 - F_V(y)\} = 1 - \max\{u, F_V(y)\}$$

pour les événements $E_1 = [U \geqslant u]$ et $E_2 = [S(U) \geqslant y]$, et en remarquant que l'on a égalité pour $S = F_V^{-1}$.

La **question 2.4** permet de conclure l'argument, en combinant les résultats de la question 2.3. Les deux questions 2.3.3 et 2.4 étaient délicates et demandaient un peu de recul sur cette partie 2. Il n'était pas nécessaire de les traiter pour obtenir une excellente note.

Dans la question 2.4 on demandait notamment si la méthode générale de la question 2.3 permet de retrouver le résultat obtenu pour l'exemple de la question 2.1. Elle le retrouve en effet puisque, dans cet exemple, la fonction de répartition est donnée par $F_V(x) = \displaystyle\int_1^x 1\,dt = x - 1$ pour $1 \leqslant x \leqslant 2$, donc son inverse par $F_V^{-1}(x) = x + 1 = S_0(x)$ pour $0 \leqslant x \leqslant 1$. On retrouve alors le fait vu en 2.1 que la fonction S_0 est un minimiseur. La méthode générale des questions 2.3 et 2.4 permet même de montrer que c'est le seul. En effet, et par analogie avec la question 1.7.3, soit S un minimiseur, c'est-à-dire une fonction de $\mathscr{S}$ telle que

$$\mathbb{E}\big[(S(U) - U)^2\big] = \mathbb{E}\big[(F_V^{-1}(U) - U)^2\big].$$

Comme en 1.7.3, on doit avoir égalité dans toutes les inégalités sommées (intégrées) menant à l'inégalité (5). En particulier pour tout u de $[0, 1]$ on doit avoir

$$\int_u^1 S(x)\,dx = \int_u^1 F_V^{-1}(x)\,dx$$

et donc nécessairement $S = F_V^{-1}$ sur $[0, 1]$.

La **partie 3** demande aux candidats de prendre du recul sur les deux premières parties. De nouveau il n'était pas nécessaire d'aborder cette partie ouverte pour avoir une excellente note, et cette partie n'a été traitée de manière conséquente que par une quinzaine de candidats. Voici quelques remarques que l'on pouvait faire.

Dans la **partie 1** on a cherché à transporter et répartir une masse 1, en chacun des n points $x_1 \leqslant \cdots \leqslant x_n$ au départ, sur les n points $y_1 \leqslant \cdots \leqslant y_n$ à l'arrivée, de sorte à minimiser un certain coût (problème dit de transport optimal).

Tout d'abord, dans la question 1.5, la masse en un point x_i est envoyée sur un unique point $y_{p(i)}$ pour une permutation p et on suppose que le coût de transport d'une masse 1 d'un point x_i à un point y_j est égal à $(y_j - x_i)^2$. On cherche alors la permutation p minimisant le coût total

$$\sum_{i=1}^{n}(y_{p(i)} - x_i)^2.$$

On montre dans la question 1.5.2 que la permutation optimale est l'identité, c'est-à-dire qu'on envoie la masse du point le plus à gauche x_1 sur le point le plus à gauche y_1, puis le suivant x_2 sur le suivant y_2, etc jusqu'à x_n et y_n. C'est ce qui a été remarqué dans l'exemple de la question 1.3.

Dans la question 1.7 on considère le cas plus général où la masse 1 en un point x_i peut être répartie sur les n points y_j à l'arrivée, la partie de la masse en x_i envoyée sur y_j étant notée $M_{ij} \geqslant 0$ et telle que

1. pour chaque i la masse 1 partant de x_i est bien intégralement répartie sur les y_j, c'est-à-dire $1 = \sum_{j=1}^{n} M_{ij}$;
2. pour chaque j le point y_j reçoit bien une masse totale 1 en provenance des x_i, c'est-à-dire $1 = \sum_{i=1}^{n} M_{ij}$.

Aussi la matrice M des M_{ij} est-elle une matrice bistochastique, et le cas particulier 1.5 ci-dessus d'une permutation p correspond au cas particulier où la masse totale 1 en x_i n'est pas séparée sur plusieurs y_j, mais envoyée sur le seul $y_{p(i)}$.

Le coût de transport d'une masse $m \in [0,1]$ d'un point x_i à un point y_j étant égal à $m\ (y_j - x_i)^2$, le coût total pour cette stratégie M est

$$\sum_{i,j=1}^{n} M_{ij}(y_j - x_i)^2$$

qu'on cherche à minimiser parmi les matrices bistochastiques M.

Après un exemple dans la question 1.4 et un cas particulier dans la question 1.6, on montre dans la question 1.7.2 que la matrice optimale est la matrice identité, c'est-à-dire qu'il n'y a en fait pas lieu de séparer la masse provenant de chaque x_i, mais qu'on doit l'envoyer sur y_i, comme ci-dessus.

Dans la **partie 2** on considère une version continue de la situation de la question 1.5 : au lieu de transporter une répartition discrète de masses sur une répartition discrète de points, on transporte une répartition continue sur une répartition continue. De nouveau la masse en $x \in [0,1]$ est transportée en un unique point $S(x) \in [1,2]$: aussi est-ce bien une version continue de la situation de la question 1.5, et non de la situation plus générale de 1.7 où l'on autorise la masse à être séparée.

Dans ce cadre continu on montre de nouveau que la stratégie optimale est la stratégie croissante, donnée par la fonction $S = F_V^{-1}$.

Dans certains cas particuliers ou situations (questions 1.3.2, 1.4.2, 1.7.3, 2.1.2, et même 2.4) on montre que cette stratégie croissante est l'unique stratégie optimale.

La démonstration est fondée sur une intégration par parties dans le cas continu (question 2.3.1), dont l'analogue discret est la question 1.2, et d'une comparaison des intégrales / sommes partielles portant sur la partie la plus à droite de la distribution : question 2.3.3 et fonction s dans le cas continu, première inégalité des question 1.5.1. et 1.7.1 dans le cas discret.