



PRÉPASIMPLE

prepasimple.fr

HEC

Mathématiques

2014

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Préliminaire

1. Dénombrement

(a) Nombre d'applications

Comme demandé, la réponse est donnée sans démonstration : $\#\mathcal{E}_N = N^N$.

(b) Applications injectives et strictement monotones

Le nombre d'applications injectives est $N!$.

Parmi elles, exactement deux sont strictement monotones : l'application strictement croissante $k \mapsto k$ et l'application strictement décroissante $k \mapsto N + 1 - k$.

Ces réponses sont données sans démonstration, conformément à la consigne.

2. Partie entière d'une variable exponentielle

(a) Loi de Y

La variable $Y = \lfloor pX \rfloor$ prend des valeurs entières et est positive ou nulle presque sûrement, puisque $X \geq 0$ presque sûrement. Elle est donc discrète, de loi portée par $\mathbb{N}$.

Pour $n \in \mathbb{N}$, comme $p > 0$,

$$\{Y = n\} = \{n \leq pX < n + 1\} = \{n/p \leq X < (n + 1)/p\}.$$

La loi exponentielle de taux 1 donne $P(X \geq t) = e^{-t}$ pour $t \geq 0$. Elle n'a pas d'atomes. Ainsi

$$P(Y = n) = e^{-n/p} - e^{-(n+1)/p} = (1 - e^{-1/p})e^{-n/p}.$$

Chaque probabilité est strictement positive, et leur somme vaut 1 par la série géométrique.

(b) Loi de $Y + 1$

Posons $q = 1 - e^{-1/p} \in]0, 1[$. Pour chaque entier $k \geq 1$,

$$P(Y + 1 = k) = P(Y = k - 1) = q(1 - q)^{k-1}.$$

Il s'agit de la loi géométrique de paramètre q sur les entiers strictement positifs :
 $Y + 1 \sim \mathcal{G}(1 - e^{-1/p})$.

(c) Encadrement strict de l'espérance

Méthode 1 — Espérance de la loi géométrique

La variable géométrique $Y + 1$ a une espérance finie égale à $1/q$. Ainsi Y a une espérance finie et

$$E(Y) = \frac{1}{q} - 1 = \frac{e^{-1/p}}{1 - e^{-1/p}} = \frac{1}{e^{1/p} - 1} > 0.$$

Pour $t > 0$, on a $e^t > 1 + t$: la fonction $e^t - 1 - t$ est nulle en zéro et a une dérivée strictement positive sur $]0, +\infty[$. Avec $t = 1/p$, cela donne $e^{1/p} - 1 > 1/p > 0$.

En passant aux inverses de quantités strictement positives, $E(Y) < p$. Ainsi $0 < E(Y) < p$.

Méthode 2 — Domination par la variable avant arrondi

La partie entière donne $0 \leq Y \leq pX$ presque sûrement. Comme $E(pX) = p < +\infty$, Y possède une espérance finie.

De plus, $P(Y \geq 1) = P(X \geq 1/p) = e^{-1/p} > 0$, donc $E(Y) \geq P(Y \geq 1) > 0$.

Pour prouver l'inégalité supérieure stricte, considérons $B = \{1/(4p) \leq X < 1/(2p)\}$. Cet événement a une probabilité strictement positive. Sur B , $Y = 0$ et $pX - Y \geq 1/4$. Partout ailleurs, $pX - Y \geq 0$. Donc

$$p - E(Y) = E(pX - Y) \geq \frac{1}{4}P(B) > 0.$$

On obtient à nouveau $0 < E(Y) < p$. La première méthode fournit une formule exacte ; la seconde montre directement la perte d'espérance produite par la partie entière.

3. Intégrales logarithmiques

(a) Convergence

Fixons $r, s \in \mathbb{N}$. L'intégrande est continue sur $]0, 1]$; seule la borne zéro peut être impropre.

Pour $0 < \varepsilon < 1$, le changement de variable $u = -\ln x$ est de classe C^1 , strictement décroissant et bijectif de $[\varepsilon, 1]$ sur $[0, -\ln \varepsilon]$, avec orientation renversée. Son inverse est $x = e^{-u}$ et $dx = -e^{-u} du$. Le changement de variable est donc valide sur ce segment.

Il donne pour l'intégrale absolue

$$\int_{\varepsilon}^1 x^r |\ln x|^s dx = \int_0^{-\ln \varepsilon} u^s e^{-(r+1)u} du.$$

L'intégrande de droite est continue sur $[0, +\infty[$. Pour u assez grand, la croissance comparée de l'exponentielle et des puissances donne $u^s \leq e^{(r+1)u/2}$. Ainsi $u^s e^{-(r+1)u} \leq e^{-(r+1)u/2}$, dont l'intégrale converge à l'infini.

Par comparaison, l'intégrale absolue converge. Lorsque ε tend vers zéro, $-\ln \varepsilon$ tend vers $+\infty$; l'intégrale demandée est donc absolument convergente.

(b) Valeur exacte

Notons $I_{r,s} = \int_0^1 x^r (\ln x)^s dx$.

Méthode 1 — Intégration par parties et récurrence

Pour $s = 0$, $I_{r,0} = 1/(r+1)$. Pour $s \geq 1$, effectuons une intégration par parties sur $[\varepsilon, 1]$ avec $u(x) = (\ln x)^s$ et $v'(x) = x^r$. On a $u'(x) = s(\ln x)^{s-1}/x$ et $v(x) = x^{r+1}/(r+1)$.

Ainsi

$$\int_{\varepsilon}^1 x^r (\ln x)^s dx = \left[\frac{x^{r+1} (\ln x)^s}{r+1} \right]_{\varepsilon}^1 - \frac{s}{r+1} \int_{\varepsilon}^1 x^r (\ln x)^{s-1} dx.$$

Le terme à la borne 1 est nul. Le terme à la borne ε tend vers zéro, car $\varepsilon^{r+1} |\ln \varepsilon|^s \rightarrow 0$ par croissance comparée. Les deux intégrales impropres convergent, par (a), ce qui justifie le passage à la limite :

$$I_{r,s} = -\frac{s}{r+1} I_{r,s-1}.$$

En itérant cette relation jusqu'à $s = 0$, on obtient

$$I_{r,s} = \frac{(-1)^s s!}{(r+1)^{s+1}}.$$

Méthode 2 — Fonction gamma

Le changement de variable déjà justifié en (a) donne

$$I_{r,s} = (-1)^s \int_0^{+\infty} u^s e^{-(r+1)u} du.$$

Le changement de variable affine $v = (r+1)u$, effectué d'abord sur des segments puis prolongé par convergence, donne

$$I_{r,s} = \frac{(-1)^s}{(r+1)^{s+1}} \int_0^{+\infty} v^s e^{-v} dv.$$

L'intégrale est $\Gamma(s+1) = s!$, résultat du programme. On retrouve la formule annoncée. La première méthode établit une récurrence exploitable dans d'autres intégrales ; la seconde reconnaît une intégrale gamma.

Partie I — Transport dans une situation aléatoire

Dans cette partie, $X \sim \mathcal{U}(]0, 1[)$.

4. Transport vers une loi de Bernoulli

(a) Loi de $T_a(X)$

La variable $T_a(X)$ vaut seulement 0 ou 1. L'intervalle $]a, a+p[$ est contenu dans $[0, 1]$ et a longueur p . Les bornes ont une probabilité nulle pour la loi uniforme. Donc

$$P(T_a(X) = 1) = P(a < X < a+p) = p.$$

Ainsi $T_a(X)$ suit la loi de Bernoulli de paramètre p , indépendamment de a . Chaque T_a transporte X vers cette même loi.

(b) Coût

Les variables X et $T_a(X)$ sont bornées, donc le coût existe. On utilise $T_a(X)^2 = T_a(X)$:

$$C(T_a) = E(X^2) + E(T_a(X)) - 2E(XT_a(X)).$$

Par transfert pour la loi uniforme,

$$E(X^2) = \int_0^1 x^2 dx = \frac{1}{3}, \quad E(T_a(X)) = p,$$

$$E(XT_a(X)) = \int_a^{a+p} x dx = ap + \frac{p^2}{2}.$$

Par substitution,

$$C(T_a) = \frac{1}{3} + p - 2ap - \frac{p^2}{2} = \frac{1}{3} + p(1 - p) - 2ap.$$

(c) Coût minimal dans la famille

Pour $p > 0$, le coefficient de a dans cette expression affine est $-2p < 0$. Le coût est donc strictement décroissant sur $[0, 1 - p]$.

Son minimum dans la famille est atteint uniquement pour $a = 1 - p$, et vaut

$$C(T_{1-p}) = \frac{1}{3} - p(1 - p) = \frac{1}{3} - p + p^2.$$

Il est positif : $p(1 - p) \leq 1/4$, donc $C(T_{1-p}) \geq 1/12$.

5. Transport vers les lois exponentielle et géométrique

(a) Lois de $T_1(X)$ et $T_2(X)$

Les deux transformations sont strictement positives sur $]0, 1[$. Pour $y < 0$, leurs fonctions de répartition sont donc nulles. Pour $y \geq 0$,

$$P(-\ln X \leq y) = P(X \geq e^{-y}) = 1 - e^{-y},$$

$$P(-\ln(1 - X) \leq y) = P(X \leq 1 - e^{-y}) = 1 - e^{-y}.$$

Les deux variables suivent la loi exponentielle de taux 1. Les fonctions T_1 et T_2 transportent donc X vers la même loi exponentielle.

(b) Comparaison des coûts

Les transformations suivent une loi exponentielle de taux 1, dont le second moment vaut 2. Les coûts existent, car $(X - T_i(X))^2 \leq 2X^2 + 2T_i(X)^2$ et les deux espérances du majorant sont finies.

Par la question 3,

$$\int_0^1 x \ln x \, dx = -\frac{1}{4}, \quad \int_0^1 (\ln x)^2 \, dx = 2.$$

Pour T_1 ,

$$C(T_1) = \int_0^1 (x + \ln x)^2 \, dx = \frac{1}{3} + 2 + 2 \left(-\frac{1}{4} \right) = \frac{11}{6}.$$

Pour T_2 , le changement de variable affine $u = 1 - x$, sur des segments puis par convergence absolue, donne

$$\int_0^1 x \ln(1 - x) \, dx = \int_0^1 (1 - u) \ln u \, du = -1 - \left(-\frac{1}{4} \right) = -\frac{3}{4},$$

$$\int_0^1 (\ln(1 - x))^2 \, dx = \int_0^1 (\ln u)^2 \, du = 2.$$

Ainsi

$$C(T_2) = \frac{1}{3} + 2 + 2 \left(-\frac{3}{4} \right) = \frac{5}{6}.$$

Donc $C(T_2) < C(T_1)$, et $C(T_1) - C(T_2) = 1$. La transformation croissante T_2 a ici le coût le plus faible des deux.

(c) Accessibilité de toutes les lois géométriques

Soit $q \in]0, 1[$ un paramètre géométrique. Posons $c = -1/\ln(1 - q) > 0$ et

$$T_q(x) = 1 + \lfloor -c \ln x \rfloor, \quad x \in]0, 1[.$$

La variable $E = -\ln X$ suit la loi exponentielle de taux 1, par (a). Le calcul de la question 2(a)-(b) utilise seulement la positivité du facteur de multiplication, et non sa majoration par 1. Vérifions-le ici pour le facteur c , qui peut être supérieur à 1.

Pour chaque $k \geq 1$,

$$P(T_q(X) = k) = P((k-1)/c \leq E < k/c) = e^{-(k-1)/c}(1 - e^{-1/c}).$$

Par définition de c , $e^{-1/c} = 1 - q$. Ainsi $P(T_q(X) = k) = q(1 - q)^{k-1}$, et $T_q(X) \sim \mathcal{G}(q)$.

Toutes les lois géométriques de paramètre dans $]0, 1[$ sont donc accessibles. Si la convention inclut le paramètre $q = 1$, la loi correspondante est certaine égale à 1 et s'obtient par la fonction constante $T(x) = 1$.

La restriction $0 < p < 1$ de la question 2 ne limite donc pas les paramètres accessibles dans cette question; le calcul particulier a été explicitement étendu à tout facteur $c > 0$.

6. Transport par l'inverse de la fonction de répartition

(a) Bijection

La densité f_Y est continue, donc F_Y est de classe C^1 sur $\mathbb{R}$ et $F'_Y = f_Y > 0$. La fonction F_Y est donc continue et strictement croissante.

Comme toute fonction de répartition, elle tend vers 0 à $-\infty$ et vers 1 à $+\infty$. Pour chaque réel y , la stricte positivité de la densité sur des intervalles situés à gauche et à droite de y donne $0 < F_Y(y) < 1$.

Par le théorème de la bijection et ces limites, F_Y réalise une bijection de $\mathbb{R}$ sur $]0, 1[$.

(b) Loi de la transformation inverse

Notons $Q = F_Y^{-1}$. La fonction Q est continue et strictement croissante sur $]0, 1[$. Pour chaque réel y , la monotonie et la réciprocité donnent

$$\{Q(X) \leq y\} = \{X \leq F_Y(y)\}.$$

Puisque $F_Y(y) \in]0, 1[$ et X est uniforme,

$$P(Q(X) \leq y) = F_Y(y).$$

La variable $Q(X)$ a donc la même fonction de répartition que Y , et par conséquent la même loi. Ainsi F_Y^{-1} est une fonction de transport vers la loi de Y .

7. Cas de la loi normale centrée réduite

On écrit $F = F_Y$ et $\varphi(y) = \frac{1}{\sqrt{2\pi}}e^{-y^2/2}$. On a $F' = \varphi$ et $\varphi'(y) = -y\varphi(y)$.

(a) Convergence et intégration par parties

Comme $0 \leq F \leq 1$, $|yF(y)\varphi(y)| \leq |y|\varphi(y)$. L'intégrale de ce majorant est finie : par parité,

$$\int_{-\infty}^{+\infty} |y|\varphi(y) dy = 2 \int_0^{+\infty} y\varphi(y) dy = 2\varphi(0) < +\infty.$$

La dernière égalité suit de la primitive $-\varphi$ et de sa limite nulle à l'infini. L'intégrale demandée est donc absolument convergente.

Sur un segment $[-A, B]$, avec $A, B > 0$, l'intégration par parties donne

$$\int_{-A}^B yF(y)\varphi(y) dy = -[F(y)\varphi(y)]_{-A}^B + \int_{-A}^B \varphi(y)^2 dy.$$

Le terme aux bornes tend vers zéro, puisque F est bornée et φ tend vers zéro aux deux infinis. De plus, $\varphi^2 \leq \varphi(0)\varphi$, donc son intégrale converge. Le passage à la limite est justifié et donne

$$\int_{-\infty}^{+\infty} yF(y)\varphi(y) dy = \int_{-\infty}^{+\infty} \varphi(y)^2 dy.$$

Pour calculer la dernière intégrale, la substitution affine $z = \sqrt{2}y$ et la normalisation de la densité normale donnent $\int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi}$. Ainsi

$$\int_{-\infty}^{+\infty} \varphi(y)^2 dy = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-y^2} dy = \frac{1}{2\sqrt{\pi}}.$$

C'est la valeur demandée.

(b) Calcul de l'intégrale quadratique

L'inégalité $(y - F(y))^2 \leq 2y^2 + 2$, avec $0 \leq F(y) \leq 1$, montre que l'intégrande est dominée par $2y^2\varphi(y) + 2\varphi(y)$, dont l'intégrale est finie. L'intégrale converge.

On peut développer et intégrer séparément, car les trois intégrales obtenues convergent :

$$\int_{-\infty}^{+\infty} (y - F(y))^2 \varphi(y) dy = \int y^2 \varphi(y) dy - 2 \int yF(y)\varphi(y) dy + \int F(y)^2 \varphi(y) dy.$$

Toutes les intégrales de cette ligne portent sur $\mathbb{R}$. Le premier terme vaut $E(Y^2) = 1$. Le deuxième est connu par (a). Pour le dernier, $F^2\varphi$ a pour primitive $F^3/3$, donc

$$\int_{-\infty}^{+\infty} F(y)^2 \varphi(y) dy = \left[\frac{F(y)^3}{3} \right]_{-\infty}^{+\infty} = \frac{1}{3}.$$

On obtient finalement

$$\int_{-\infty}^{+\infty} (y - F(y))^2 \varphi(y) dy = \frac{4}{3} - \frac{1}{\sqrt{\pi}}.$$

(c) Coût du transport inverse

Posons $V = F^{-1}(X)$. La question 6 assure que V suit la loi normale centrée réduite. De plus, $F(V) = X$ pour chaque issue, puisque $X(\Omega) =]0, 1[$.

Le coût existe, car $(X - V)^2 \leq 2X^2 + 2V^2$ et les deux variables ont un second moment fini. En utilisant $X = F(V)$, puis le transfert pour la loi normale,

$$C(F^{-1}) = E((F(V) - V)^2) = \int_{-\infty}^{+\infty} (F(y) - y)^2 \varphi(y) dy.$$

Par (b), $C(F_Y^{-1}) = \frac{4}{3} - \frac{1}{\sqrt{\pi}}.$

Cette preuve utilise une identité entre variables aléatoires et le transfert; elle ne requiert aucun changement de variable non affine supplémentaire.

Partie II — Transport optimal dans une situation déterministe

8. Inégalité pondérée

(a) Comparaison de deux appariements

La différence des deux membres est

$$d_k a_k + d_\ell a_\ell - d_k a_\ell - d_\ell a_k = (d_k - d_\ell)(a_k - a_\ell).$$

Les deux suites sont strictement croissantes. Les deux facteurs ont donc le même signe lorsque $k \neq \ell$, et sont tous deux nuls lorsque $k = \ell$. Leur produit est positif ou nul, ce qui établit l'inégalité.

Le même argument reste valable pour deux suites seulement croissantes au sens large : cette extension sera utilisée à la question 11.

(b) Double sommation

Multiplions l'inégalité de (a) par $p_k p_\ell \geq 0$ et sommons sur tous les couples (k, ℓ) . Toutes les sommes sont finies.

Comme

$$\sum_{\ell} p_{\ell} = 1,$$

on a

$$\sum_{k, \ell} p_k p_{\ell} d_k a_k = \sum_k p_k d_k a_k, \quad \sum_{k, \ell} p_k p_{\ell} d_{\ell} a_{\ell} = \sum_{\ell} p_{\ell} d_{\ell} a_{\ell}.$$

Les deux sommes croisées valent chacune

$$\left(\sum_k p_k d_k \right) \left(\sum_{\ell} p_{\ell} a_{\ell} \right).$$

La double sommation donne donc

$$2 \sum_k p_k d_k a_k \geq 2 \left(\sum_k p_k d_k \right) \left(\sum_k p_k a_k \right).$$

En divisant par 2, on obtient l'inégalité (1).

Cette preuve établit aussi l'identité particulière

$$\sum_k p_k d_k a_k - \left(\sum_k p_k d_k \right) \left(\sum_k p_k a_k \right) = \frac{1}{2} \sum_{k,\ell} p_k p_\ell (d_k - d_\ell)(a_k - a_\ell),$$

qui rend la positivité explicite.

9. Réordonnement croissant

Posons $b_k = a_{t(k)}$ et $c_k = a_{\hat{t}(k)}$. Les deux listes contiennent les mêmes valeurs, avec les mêmes répétitions, et $c_1 \leq \dots \leq c_N$.

(a) Comparaison des sommes de fin de liste

Pour $n = 1$, les sommes totales sont égales, puisque les deux listes diffèrent seulement par leur ordre.

Pour $n \geq 2$, les $n - 1$ premiers termes de la liste b sont un choix de $n - 1$ occurrences dans la liste complète. Après avoir trié ces occurrences, elles peuvent s'écrire $c_{i_1}, \dots, c_{i_{n-1}}$, avec $i_1 < \dots < i_{n-1}$. Les indices vérifient $i_j \geq j$, donc $c_{i_j} \geq c_j$.

Ainsi

$$\sum_{k=1}^{n-1} b_k \geq \sum_{k=1}^{n-1} c_k.$$

En soustrayant de la somme totale commune, on obtient

$$\sum_{k=n}^N b_k \leq \sum_{k=n}^N c_k.$$

Cette preuve distingue les occurrences et reste donc valide même si plusieurs valeurs de t sont identiques.

(b) Sommation par différences

En intervertissant deux sommes finies,

$$\sum_{n=1}^N (d_n - d_{n-1}) \sum_{k=n}^N b_k = \sum_{k=1}^N b_k \sum_{n=1}^k (d_n - d_{n-1}).$$

La somme intérieure est télescopique et vaut $d_k - d_0 = d_k$, car $d_0 = 0$. On retrouve donc

$$\sum_{k=1}^N d_k b_k,$$

ce qui prouve l'égalité demandée.

(c) Inégalité de réordonnement

Méthode 1 — Sommes de fin de liste et différences

En utilisant (b) pour les deux listes,

$$\sum_{k=1}^N d_k c_k - \sum_{k=1}^N d_k b_k = \sum_{n=1}^N (d_n - d_{n-1}) \left(\sum_{k=n}^N c_k - \sum_{k=n}^N b_k \right).$$

Pour $n = 1$, le facteur entre parenthèses est nul. Ce terme est donc nul même si d_1 est négatif. Pour $n \geq 2$, on a $d_n - d_{n-1} > 0$, et le facteur entre parenthèses est positif ou nul par (a).

La somme entière est donc positive ou nulle, ce qui donne

$$\sum_{n=1}^N d_n a_{t(n)} \leq \sum_{n=1}^N d_n a_{\hat{t}(n)}.$$

La prise en compte du terme $n = 1$ est nécessaire, car l'énoncé n'impose pas que les points de départ soient positifs.

Méthode 2 — Suppression des inversions

Si une liste $(b_1, \dots, b_N)$ n'est pas croissante, il existe deux termes voisins $b_i > b_{i+1}$. Les échanger modifie la somme pondérée de

$$d_i b_{i+1} + d_{i+1} b_i - d_i b_i - d_{i+1} b_{i+1} = (d_{i+1} - d_i)(b_i - b_{i+1}) > 0.$$

La somme pondérée augmente donc à chaque échange de voisins mal ordonnés. Le nombre de paires d'indices $j < k$ telles que $b_j > b_k$ est un entier positif ou nul. Un échange de voisins mal ordonnés le diminue exactement de 1 : leur inversion disparaît, et les comparaisons avec les autres positions ne changent pas dans leur nombre total.

Après un nombre fini d'échanges, la liste est croissante et devient la liste c . La somme finale est au moins la somme initiale, ce qui établit à nouveau (2).

La première méthode est directement liée aux questions précédentes ; la seconde décrit une procédure de tri. Les deux preuves restent valables pour des poids d_k seulement croissants au sens large et pour une liste initiale de valeurs réelles quelconques : il suffit alors de remplacer les accroissements strictement positifs par des accroissements positifs ou nuls. Pour $N = 1$, l'inégalité est une égalité.

10. Optimalité de l'appariement croissant

Méthode 1 — Développement du coût et utilisation de (2)

Tout programme de transport $T : D \rightarrow A$ est une bijection. Il existe donc une permutation t de $\{1, \dots, N\}$ telle que $T(d_k) = a_{t(k)}$. Sa liste réordonnée est $(1, \dots, N)$.

Le développement du coût donne

$$c(T) = \sum_{k=1}^N d_k^2 + \sum_{k=1}^N a_{t(k)}^2 - 2 \sum_{k=1}^N d_k a_{t(k)}.$$

La bijectivité donne

$$\sum_k a_{t(k)}^2 = \sum_k a_k^2.$$

Par (2),

$$\sum_k d_k a_{t(k)} \leq \sum_k d_k a_k.$$

Ainsi

$$c(T) \geq \sum_k d_k^2 + \sum_k a_k^2 - 2 \sum_k d_k a_k = \sum_k (d_k - a_k)^2 = c(\widehat{T}).$$

Le transport $\widehat{T}(d_k) = a_k$ est donc optimal.

Méthode 2 — Décroissement de deux affectations

Supposons qu'un programme affecte d_i à a_r et d_j à a_s , avec $i < j$ mais $r > s$. Échangeons ces deux arrivées, en conservant toutes les autres affectations.

Le coût avant échange moins le coût après échange vaut

$$(d_i - a_r)^2 + (d_j - a_s)^2 - (d_i - a_s)^2 - (d_j - a_r)^2 = 2(d_j - d_i)(a_r - a_s) > 0.$$

L'échange réduit strictement le coût. Un programme possédant une inversion ne peut donc être optimal. En supprimant successivement des inversions voisines, comme dans la seconde preuve de 9(c), on atteint la permutation croissante, qui est l'identité.

On obtient donc l'optimalité de $\widehat{T}$. Cette méthode montre aussi son unicité : toute autre permutation possède une inversion, qui permet de diminuer strictement son coût.

La première preuve utilise l'inégalité générale ; la seconde explique concrètement pourquoi croiser deux affectations augmente le coût quadratique.

11. Interprétation probabiliste

(a) Covariance avec une fonction croissante

Si X prend une seule valeur, les deux membres de l'inégalité sont égaux. Sinon, rangeons les valeurs de probabilité positive de X en $x_1 < \dots < x_m$, et posons $p_k = P(X = x_k)$.

Les valeurs $h(x_k)$ sont croissantes au sens large. La preuve de l'inégalité (1), explicitement étendue à ce cas en 8(a), s'applique avec $d_k = x_k$ et $a_k = h(x_k)$.

Les variables X , $h(X)$ et $Xh(X)$ ont des images finies, donc leurs espérances existent. La formule du transfert donne

$$E(Xh(X)) = \sum_k p_k x_k h(x_k) \geq \left(\sum_k p_k x_k \right) \left(\sum_k p_k h(x_k) \right) = E(X)E(h(X)).$$

Ainsi $\text{Cov}(X, h(X)) \geq 0$.

(b) Coefficient de corrélation

Si les deux variances sont strictement positives, le coefficient est défini par

$$\rho(X, h(X)) = \frac{\text{Cov}(X, h(X))}{\sqrt{\mathbb{V}(X)\mathbb{V}(h(X))}}.$$

Le dénominateur est strictement positif, et le numérateur est positif ou nul par (a). Donc $\rho(X, h(X)) \geq 0$. Avec la borne générale du programme, $0 \leq \rho(X, h(X)) \leq 1$.

Sous ces hypothèses, on peut même préciser que $\rho(X, h(X)) > 0$. La variance strictement positive de $h(X)$ impose deux valeurs $x_i < x_j$ de probabilité positive avec $h(x_i) < h(x_j)$. Dans l'identité de 8(b), le terme correspondant est alors strictement positif, et tous les autres sont positifs ou nuls. La covariance est donc strictement positive.

On a ainsi $0 < \rho(X, h(X)) \leq 1$. L'égalité avec 1 correspond au cas où $h(X) = aX + b$ presque sûrement pour certains réels $a > 0$ et b , par le cas d'égalité de la corrélation.

(c) Réordonnement d'une application d'une variable uniforme

Comme X est uniforme sur $\{1, \dots, N\}$,

$$E(h(X)t(X)) = \frac{1}{N} \sum_{k=1}^N h(k)t(k), \quad E(h(X)\hat{t}(X)) = \frac{1}{N} \sum_{k=1}^N h(k)\hat{t}(k).$$

Les poids $h(k)$ sont croissants au sens large. La liste $\hat{t}(k)$ est le réordonnement croissant de la liste $t(k)$. Appliquons la version de (2) établie en 9(c), avec ces poids et la liste réelle $t(k)$.

On obtient

$$\sum_k h(k)t(k) \leq \sum_k h(k)\hat{t}(k).$$

Après division par $N > 0$,

$$E(h(X)t(X)) \leq E(h(X)\hat{t}(X)).$$

Les espérances existent, puisque les variables concernées ont des images finies.

Partie III — Transport optimal dans une situation aléatoire

Dans cette partie, $U \sim \mathcal{U}([0, 1])$, g est de classe C^1 sur $[0, 1]$ et $Z = g(U)$ a la même loi que Y .

La densité de Y est nulle hors de $[\alpha, \beta]$, et continue et strictement positive sur ce segment. Par conséquent,

$$F_Y(y) = 0 \text{ pour } y \leq \alpha, \quad F_Y(y) = 1 \text{ pour } y \geq \beta.$$

Sur $[\alpha, \beta]$, la fonction est continue et strictement croissante : pour $\alpha \leq a < b \leq \beta$,

$$F_Y(b) - F_Y(a) = \int_a^b f_Y(t) dt > 0. \text{ Sa restriction est donc une bijection de } [\alpha, \beta] \text{ sur } [0, 1].$$

Notons $Q = F_Y^{-1}$ cette réciproque, qui est continue et croissante sur $[0, 1]$.

Toutes les variables à valeurs dans ces segments sont bornées, donc possèdent tous les moments utilisés ci-dessous.

12. Discrétisation de la variable uniforme

(a) Loi de X_N

Pour $1 \leq k < N$, l'événement $\{X_N = k\}$ est $\{(k-1)/N \leq U < k/N\}$. Pour $k = N$, il est $\{(N-1)/N \leq U \leq 1\}$, grâce à la valeur particulière prévue lorsque $U = 1$.

Ces intervalles ont tous longueur $1/N$, et une variable uniforme continue n'a pas d'atomes. Ainsi

$$P(X_N = k) = \frac{1}{N} \quad \text{pour } k \in \{1, \dots, N\}.$$

Donc X_N suit la loi uniforme discrète sur $\{1, \dots, N\}$. Pour $N = 1$, elle est certaine égale à 1, ce qui est cohérent avec la même formule.

(b) Contrôle uniforme de l'erreur

Pour chaque issue, la définition donne $0 \leq X_N/N - U \leq 1/N$. Lorsque $U = 1$, la différence est nulle ; lorsque $U < 1$, cette borne suit directement de la partie entière.

La dérivée g' est continue sur le segment $[0, 1]$, donc elle y est bornée et atteint un maximum en valeur absolue. Posons $L = \max_{u \in [0, 1]} |g'(u)|$ et choisissons $\lambda = L + 1 > 0$.

L'inégalité des accroissements finis donne, pour tous $u, v \in [0, 1]$,
 $|g(u) - g(v)| \leq L|u - v|$. Ainsi, pour chaque issue,

$$|Z - Y_N| = |g(U) - g(X_N/N)| \leq L|U - X_N/N| \leq \frac{L}{N} \leq \frac{\lambda}{N}.$$

La constante λ est strictement positive et indépendante de N .

(c) Comparaison des fonctions de répartition

Posons $\delta = \lambda/N$. La borne précédente implique $Y_N \leq Z + \delta$ pour chaque issue. Donc

$$\{Z < y - \delta\} \subset \{Y_N < y\}.$$

La variable Z a la même loi à densité que Y ; elle n'a donc aucun atome. Il s'ensuit $P(Z < y - \delta) = F_Y(y - \delta)$, même avec l'inégalité stricte à gauche.

En prenant les probabilités dans l'inclusion,

$$F_Y(y - \lambda/N) \leq P(Y_N < y).$$

L'absence d'atomes de Z est nécessaire pour conserver la forme stricte de l'événement demandé ; aucune absence d'atomes n'est supposée pour Y_N , qui est discrète.

13. Comparaison avec le transport croissant

(a) Encadrement d'une valeur réordonnée

La première inégalité est la question 12(c) avec $y = \hat{t}_N(k)$.

Comme $Y_N = t_N(X_N)$ et X_N est uniforme sur $\{1, \dots, N\}$,

$$P(Y_N < \hat{t}_N(k)) = \frac{1}{N} \#\{j \in \{1, \dots, N\} \mid t_N(j) < \hat{t}_N(k)\}.$$

Dans la liste triée, il y a au plus $k - 1$ occurrences strictement inférieures à sa k -ième valeur. Ce comptage reste valable en présence de répétitions. Ainsi

$$P(Y_N < \hat{t}_N(k)) \leq \frac{k-1}{N} < \frac{k}{N}.$$

On obtient les deux inégalités demandées, pour tout $1 \leq k \leq N$, y compris $k = N$ et $N = 1$.

(b) Inégalité entre les sommes pondérées

La liste t_N prend ici des valeurs réelles, et non nécessairement des valeurs dans $\{1, \dots, N\}$. La version du réordonnement pour une liste réelle a été établie dans la preuve de 9(c). Appliquons-la avec les poids croissants k/N .

Pour expliciter cette application, notons $b_k = t_N(k)$ et $c_k = \hat{t}_N(k)$. La comparaison des sommes de fin de liste de 9(a) utilise seulement le tri des occurrences, donc

$$\sum_{k=n}^N c_k \geq \sum_{k=n}^N b_k,$$

avec égalité pour $n = 1$. La sommation par différences de 9(b) donne alors

$$\sum_{k=1}^N \frac{k}{N} c_k - \sum_{k=1}^N \frac{k}{N} b_k = \frac{1}{N} \sum_{n=1}^N \left(\sum_{k=n}^N c_k - \sum_{k=n}^N b_k \right) \geq 0.$$

Ainsi

$$\sum_k (k/N)g(k/N) \leq \sum_k (k/N)\hat{t}_N(k).$$

Par (a), $F_Y(\hat{t}_N(k) - \lambda/N) < k/N$. Or $F_Y(Q(k/N)) = k/N$. Si $\hat{t}_N(k) - \lambda/N \geq Q(k/N)$, la croissance de F_Y sur $\mathbb{R}$ donnerait l'inégalité inverse entre ces valeurs, contradiction.

Donc

$$\hat{t}_N(k) < Q(k/N) + \lambda/N.$$

Ce raisonnement reste valable si $\hat{t}_N(k) - \lambda/N$ est inférieur à α : on n'applique pas la réciproque à un argument hors de son domaine. Il vaut également pour $k = N$, où $Q(1) = \beta$.

En multipliant par $k/N > 0$, en sommant puis en divisant par $N > 0$, on obtient

$$\frac{1}{N} \sum_{k=1}^N \frac{k}{N} g(k/N) \leq \frac{1}{N} \sum_{k=1}^N \frac{k}{N} \left(Q(k/N) + \frac{\lambda}{N} \right).$$

L'inégalité demandée est établie. Pour $N = 1$, le réordonnement ne change pas la liste et les mêmes arguments s'appliquent.

(c) Passage à la limite par sommes de Riemann

Les fonctions $u \mapsto ug(u)$ et $u \mapsto uQ(u)$ sont continues sur $[0, 1]$, puisque g et Q le sont. Leurs sommes de Riemann aux extrémités droites convergent donc respectivement vers

$$\int_0^1 ug(u) du, \quad \int_0^1 uQ(u) du.$$

Le terme supplémentaire dans le membre de droite de (b) vaut

$$\frac{\lambda}{N} \cdot \frac{1}{N} \sum_{k=1}^N \frac{k}{N} = \frac{\lambda(N+1)}{2N^2} \rightarrow 0,$$

car

$$\sum_{k=1}^N k = N(N+1)/2.$$

Le passage à la limite dans l'inégalité des sommes réelles donne

$$\int_0^1 ug(u) du \leq \int_0^1 uQ(u) du.$$

Les produits $Ug(U)$ et $UQ(U)$ sont bornés et ont donc des espérances. Le transfert pour la loi uniforme donne finalement

$$E(Ug(U)) \leq E(UF_Y^{-1}(U)).$$

Le passage à la limite utilise uniquement les sommes de Riemann de fonctions continues; aucun échange général entre une limite et une espérance n'est invoqué.

14. Transport optimal et exemple

(a) Fonction de transport de coût minimal

Considérons $T^* = Q = F_Y^{-1}$ sur $[0, 1]$.

Vérifions d'abord qu'elle appartient à la classe demandée. Sur $]0, 1[$, le théorème de dérivation de la réciproque donne

$$Q'(u) = \frac{1}{f_Y(Q(u))}.$$

La restriction de f_Y à $[\alpha, \beta]$ est continue et strictement positive. Elle y atteint donc un minimum strictement positif. Comme Q est continue jusqu'aux extrémités, le membre de droite se prolonge continûment sur $[0, 1]$, avec valeurs $1/f_Y(\alpha)$ et $1/f_Y(\beta)$ aux bornes.

Le théorème du prolongement de la dérivée donne les dérivées unilatérales correspondantes en 0 et 1. Ainsi Q est de classe C^1 sur le segment, avec la convention usuelle de dérivées unilatérales aux extrémités.

Pour $y \in [\alpha, \beta]$, la croissance de Q donne $P(Q(U) \leq y) = P(U \leq F_Y(y)) = F_Y(y)$. Pour $y < \alpha$, les deux fonctions de répartition valent zéro ; pour $y \geq \beta$, elles valent 1. Ainsi $Q(U)$ suit la loi de Y : Q est bien une fonction de transport.

Pour comparer avec toutes les fonctions de transport de classe C^1 , remarquons qu'une telle fonction g prend nécessairement ses valeurs dans $[\alpha, \beta]$. En effet, si $g(u_0)$ était hors de ce segment, sa continuité fournirait un intervalle de longueur positive dans $[0, 1]$ où ses valeurs restent hors du segment. La loi uniforme donnerait alors une probabilité strictement positive d'obtenir $g(U)$ hors de $[\alpha, \beta]$, contrairement à la loi cible. Ce raisonnement vaut aussi lorsque u_0 est une extrémité.

La question 13(c) s'applique donc à toute fonction de transport g de classe C^1 . Les coûts existent, car U , $g(U)$ et $Q(U)$ sont bornés. En développant,

$$C(g) = E(U^2) + E(g(U)^2) - 2E(Ug(U)).$$

Comme $g(U)$ et $Q(U)$ suivent la même loi que Y , leurs seconds moments sont égaux à $E(Y^2)$. Par 13(c),

$$C(g) \geq E(U^2) + E(Y^2) - 2E(UQ(U)) = C(Q).$$

Une fonction de transport de coût minimal est donc

$$T^*(u) = F_Y^{-1}(u), \quad u \in [0, 1].$$

Son coût minimal peut s'écrire $C(T^*) = \int_0^1 (u - F_Y^{-1}(u))^2 du$.

(b) Exemple $Y = |4U - 2|$

Méthode 1 — Fonction de répartition et transport inverse

La variable Y prend ses valeurs dans $[0, 2]$. Pour $0 \leq y \leq 2$,

$$\{|4U - 2| \leq y\} = \left\{ \frac{2-y}{4} \leq U \leq \frac{2+y}{4} \right\}.$$

L'intervalle a longueur $y/2$, donc $F_Y(y) = y/2$ sur $[0, 2]$. Hors de ce segment, la fonction vaut zéro à gauche et 1 à droite. Ainsi Y suit la loi uniforme sur $[0, 2]$, avec une densité égale à $1/2$ sur ce segment.

La réciproque de la restriction de F_Y est $Q(u) = 2u$. Par (a), on peut donc choisir $T^*(u) = 2u$.

Son coût vaut

$$C(T^*) = E((U - 2U)^2) = E(U^2) = \int_0^1 u^2 du = \frac{1}{3}.$$

La fonction initiale $u \mapsto |4u - 2|$ n'est pas de classe C^1 au point $1/2$, mais cela ne fait pas obstacle à la question : sa loi admet la fonction de transport C^1 $u \mapsto 2u$.

Méthode 2 — Borne par la covariance

La loi cible, calculée ci-dessus, est uniforme sur $[0, 2]$. Pour toute fonction de transport g , posons $V = g(U)$. Les moments existent puisque V est presque sûrement dans $[0, 2]$. On a

$$E(U) = \frac{1}{2}, \quad \mathbb{V}(U) = \frac{1}{12}, \quad E(V) = 1, \quad \mathbb{V}(V) = \frac{1}{3}, \quad E(V^2) = \frac{4}{3}.$$

Cauchy-Schwarz pour les variables centrées donne

$$\text{Cov}(U, V) \leq \sqrt{\mathbb{V}(U)\mathbb{V}(V)} = \frac{1}{6}.$$

Donc $E(UV) = E(U)E(V) + \text{Cov}(U, V) \leq 1/2 + 1/6 = 2/3$. Par conséquent,

$$C(g) = E(U^2) + E(V^2) - 2E(UV) \geq \frac{1}{3} + \frac{4}{3} - \frac{4}{3} = \frac{1}{3}.$$

La fonction $g(u) = 2u$ transporte la loi uniforme sur $[0, 1]$ vers celle sur $[0, 2]$ et réalise exactement ce coût. Elle est donc optimale, en particulier dans la classe C^1 demandée.

La première méthode utilise le résultat général du problème ; la seconde vérifie l'optimalité de cet exemple directement à l'aide des moments et de Cauchy-Schwarz.

Relevé des adaptations

Aucune modification des questions n'a été nécessaire. Les extensions particulières utiles — facteur positif quelconque en 5(c), listes réelles et suites croissantes au sens large en 11 et 13 — sont établies dans les démonstrations avec les outils du référentiel.