Devoir surveillé 7(4h)

I La fonction digamma pour exprimer une loi de probabilité

Rappels sur Γ et digamma

On note, pour tout $x \in]0,+\infty[$,

$$\Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1} \, \mathrm{d}t.$$

Q 1) Démontrer que la fonction Γ est de classe $\mathcal{C}^1$ sur $]0,+\infty[$ et exprimer Γ' sous forme intégrale.

Pour tout $x > 0$, on pose :

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

Malgré sa notation ψ cette fonction est appelée *digamma* (lettre grecque disparue...)

Q 2) Démontrer que $(\sum_{k=1}^{n} \frac{1}{k}) - \ln(n)$ converge quand $n \to +\infty$.

On note :

$$\gamma = \lim_{n\to+\infty} (\sum_{k=1}^{n} \frac{1}{k}) - \ln(n).$$

Au D.S. 4, on a montré que pour tout $x \in]0,+\infty[$,

$$\psi(x) = \frac{-1}{x} - \gamma + \sum_{k=1}^{+\infty} \left(\frac{1}{k} - \frac{1}{k+x}\right) \qquad (\dagger)$$

Q 3) En déduire une relation entre $\int_0^{+\infty} e^{-t} \ln(t) dt$ et γ.

Un peu de probabilités

Une urne contient n boules numérotées de 1 à n.

On effectue un premier tirage d'un boule dans l'urne et on adopte le protocole suivant :

si on a tiré la boule numéro k, on la remet alors dans l'urne avec k nouvelles boules toutes numérotées k.

Par exemple, si on a tiré la boule numéro 3, on remet quatre boules de numéro 3 dans l'urne (la boule tirée plus 3 nouvelles boules numéro 3).

On effectue ensuite un deuxième tirage d'une boule. On note X (respectivement Y) la variable aléatoire égale au numéro de la boule choisie au premier tirage (respectivement au deuxième tirage).

Q 4) Déterminer la loi de la variable aléatoire X ainsi que son espérance $E(X)$.

Q 5) Déterminer la loi de la variable aléatoire Y et vérifier que

$$\forall\, k \in [\![1,n]\!], P(Y = k) = \frac{1}{n}\left(\psi(2n+1) - \psi(n+1) + \frac{k}{n+k}\right).$$

Q 6) Exprimer l'espérance $E(Y)$ à l'aide de la fonction ψ.

On pourra utiliser, sans démonstration, que

$$\sum_{k=1}^{n} \frac{k^2}{n(n+k)} = \frac{1-n}{2} + n(\psi(2n+1) - \psi(n+1))$$

II Matrices stochastiques et processus de Markov

On dit qu'une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ est *stochastique* lorsque elle vérifie les conditions :

$$\forall i \in [\![1,n]\!], \sum_{j=1}^{n} a_{i,j} = 1 \tag{1}$$

$$\forall (i,j) \in [\![1,n]\!]^2,\ a_{i,j} \geq 0 \tag{2}$$

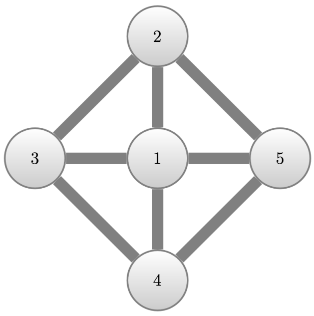
Ces matrices, et plus précisément les matrices stochastiques strictement positives, qui vérifient $a_{i,j} > 0$ pour tout $(i,j) \in [\![1,n]\!]^2$ ont été étudiées au D.S.2.

Il y a été montré que

> **Résultat du D.S. 2 :** Pour $A \in M_n(\mathbb{R})$ stochastique strictement positive :
> - 1 est valeur propre de A de multiplicité algébrique 1,
> - $\forall \lambda \in \mathrm{Sp}_{\mathbb{C}}(A) \setminus \{1\}$, $|\lambda| < 1$

Nous allons voir une conséquence de ces résultats en probabilité.

Un labyrinthe est constitué de cinq salles, numérotées de 1 à 5 , qui communiquent par des tubes selon le schéma ci-dessous :



Un rat se déplace dans ce labyrinthe, et on relève sa position en des instants numérotés $0, 1, 2, \cdots, k, \cdots$ pour $k \in \mathbf{N}$. On admet que, si le rat se trouve à l'instant k (pour $k \in \mathbf{N}$) dans la salle numéro i avec $1 \leq i \leq 5$, alors il empruntera aléatoirement l'un des tubes de la salle i et se trouvera donc, à l'instant $k+1$, avec équiprobabilité, dans l'une quelconque des salles communiquant avec la salle i (il ne reste pas dans la même salle). On admet que l'on peut introduire, pour tout k entier naturel, une variable aléatoire S_k donnant le numéro de la salle où se trouve le rat à l'instant k. À titre d'exemple, on aura donc

$$\forall k \in \mathbf{N},$$
$$P(S_{k+1} = 1 \mid S_k = 2) = P(S_{k+1} = 3 \mid S_k = 2) = P(S_{k+1} = 5 \mid S_k = 2) = \frac{1}{3}.$$

Pour tout $k \in \mathbf{N}$, on introduit la matrice-colonne

$$X_k = \begin{pmatrix} P(S_k = 1) \\ P(S_k = 2) \\ P(S_k = 3) \\ P(S_k = 4) \\ P(S_k = 5) \end{pmatrix} \in \mathcal{M}_{5,1}(\mathbf{R}).$$

Premiers pas

Q 7) montrer que $P(S_{k+1} = 1)$ s'écrit comme une combinaison linéaire des $(P(S_k = i), i = 1, \cdots, 5)$.

Q 8) Expliciter la matrice carrée $B \in \mathcal{M}_5(\mathbf{R})$ telle que $X_{k+1} = BX_k$ pour tout k entier naturel, et vérifier que $B^\top$ est un matrice stochastique.

Q 9) Montrer que $B^{\top}$ admet 1 comme valeur propre et expliciter un vecteur propre associé.

On suppose que la loi de la variable S_0 est donnée par : $X_0 = \begin{pmatrix} 1/4 \\ 3/16 \\ 3/16 \\ 3/16 \\ 3/16 \end{pmatrix}$.

Q 10) Montrer qu'alors les variables aléatoires S_k ont toutes la même loi.

Q 11) Est-ce que S_0 et S_1 sont indépendantes ?

Propriété générale de Convergence en moyenne de Cesaro dans $\mathcal{M}_n(\mathbf{R})$ pour les matrices de norme d'opérateur au plus un

Soit u un endomorphisme d'un $\mathbb{R}$-espace vectoriel E de dimension finie. On suppose qu'il existe une norme $\|\cdot\|$ sur E telle que l'inégalité suivante soit satisfaite pour tout $x \in E$,

$$\|u(x)\| \leq \|x\|.$$

Pour tout entier naturel k non nul, on considère l'endomorphisme de moyenisation des itérés de u :

$$r_k = \frac{1}{k}\sum_{l=0}^{k-1} u^l = \frac{1}{k}\left(\mathrm{I}_E + u + u^2 + \cdots + u^{k-1}\right),$$

où I_E représente l'endomorphisme identité de E.

Q 12) a) Soit $x \in \ker(u - \mathrm{I}_E)$. Déterminer $\lim\limits_{k\to\infty} r_k(x)$.

b) Soit $x \in \mathrm{Im}\,(u - \mathrm{I}_E)$. Montrer que $\lim\limits_{k\to\infty} r_k(x) = 0_E$.

c) En déduire que $E = \ker(u - \mathrm{I}_E) \oplus \mathrm{Im}\,(u - \mathrm{I}_E)$.

d) Soit $x \in E$, un vecteur quelconque. Montrer que la suite $(r_k(x))_{k\in\mathbf{N}^*}$ converge vers un vecteur de E, que l'on notera $p(x)$.

Interpréter géométriquement l'application $p : E \longrightarrow E$ ainsi définie.

Soit $A \in \mathcal{M}_n(\mathbf{R})$ une matrice carrée d'ordre n à coefficients réels. On suppose qu'il existe une norme, aussi notée $\|\cdot\|$, sur l'espace vectoriel $\mathcal{M}_{n,1}(\mathbf{R})$ identifié à $\mathbf{R}^n$, telle que, pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$, on ait $\|AX\| \leq \|X\|$. Pour tout k entier naturel non nul, on considère la matrice

$$R_k = \frac{1}{k}\sum_{l=0}^{k-1} A^l = \frac{1}{k}\left(I_n + A + A^2 + \cdots + A^{k-1}\right), \tag{3}$$

où I_n est la matrice identité dans $\mathcal{M}_n(\mathbf{R})$.

Q 13) Montrer que la suite de matrices $(R_k)_{k\in\mathbf{N}^*}$ converge dans $\mathcal{M}_n(\mathbf{R})$ vers une matrice P, telle que $P^2 = P$.

Retour aux matrices stochastiques

On fixe dans cette partie, un entier $n \geq 2$.

Notation : On notera $U \in \mathcal{M}_{n,1}(\mathbf{R})$, la matrice-colonne dont tous les coefficients sont égaux à 1.

Rappel : Une matrice carrée $A = (a_{i,j}) \in \mathcal{M}_n(\mathbf{R})$ est dite stochastique si elle vérifie les conditions (1) et (2) données au début de ce II.

Nous dirons aussi qu'une matrice-ligne $L = (\lambda_1, \cdots, \lambda_n) \in \mathcal{M}_{1,n}(\mathbf{R})$ est stochastique lorsque ses coefficients λ_i sont tous positifs ou nuls, et de somme égale à 1.

Q 14) a) Vérifier que la condition (1) équivaut à la condition $AU = U$.

b) En déduire que l'ensemble $\mathcal{E}$ des matrices stochastiques (carrées d'ordre n) est stable par le produit matriciel.

c) Montrer que cet ensemble $\mathcal{E}$ est une partie fermée et convexe de $\mathcal{M}_n(\mathbf{R})$.

On munit l'espace $\mathcal{M}_{n,1}(\mathbf{R})$ de la norme $\|\cdot\|_\infty$ définie par $\|X\|_\infty = \max\limits_{1\leq i\leq n} |x_i|$ si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Q 15) Montrer que, si $A \in \mathcal{M}_n(\mathbf{R})$ est stochastique, alors on a $\|AX\|_\infty \leq \|X\|_\infty$ pour tout $X \in \mathcal{M}_{n,1}(\mathbf{R})$.

Dans les questions suivantes, on note $A \in \mathcal{M}_n(\mathbf{R})$ une matrice stochastique, et on suppose qu'il existe un entier naturel non nul p tel que la matrice A^p ait tous ses coefficients strictement positifs. Pour tout k entier naturel non nul, on posera

$$R_k = \frac{1}{k}\sum_{l=0}^{k-1} A^l$$

Q 16) Avec le résultat du D.S. 2 encadré plus haut, on sait que $\ker(A^p - I_n)$ est de dimension 1. En déduire que $\ker(A - I_n) = \mathrm{Vect}(U)$.

Q 17) Montrer que, pour tout $k \in \mathbf{N}^*$, la matrice R_k est stochastique.

Q 18) Montrer que la suite $(R_k)_{k\in\mathbf{N}^*}$ converge dans $\mathcal{M}_n(\mathbf{R})$ vers une matrice P, stochastique, de rang 1.

Q 19) En déduire que l'on peut écrire $P = UL$, où $L = (\lambda_1, \cdots, \lambda_n) \in \mathcal{M}_{1,n}(\mathbf{R})$ est une matrice-ligne stochastique.

Q 20) Montrer que $PA = P$. En déduire que L est la seule matrice-ligne stochastique vérifiant $LA = L$.

Application au labyrinthe

On approfondit l'étude commencée au début de ce II grâce aux résultats qu'on vient d'obtenir. On pose $A = B^\top$ où B est la matrice construite à la question 8.

Un calcul qui n'est pas demandé, montre que les coefficients de la matrice A^2 sont tous strictement positifs, ce qui traduit le fait que le rat peut, en deux étapes, passer de n'importe quelle salle à n'importe quelle autre.

Q 21) Expliciter la limite P de la suite de matrices $(R_k)_{k\in\mathbf{N}^*}$ définie en (3) pour cette matrice A.

Q 22) Montrer qu'il existe une unique loi de probabilité sur l'ensemble $[\![1, 5]\!]$ telle que, si la variable aléatoire S_0 suit cette loi, alors les variables S_k suivent toutes la même loi (autrement dit, telle que la probabilité de présence du rat dans une salle soit la même à tous les instants $k, k \in \mathbf{N}$). (On parle de l'unique loi de probabilité invariante pour ce processus de Markov).

III Équation de diffusion et marche aléatoire

Cette partie donne une modélisation probabilistique donnant la même relation de récurrence que le schéma numérique d'approximation des solutions de l'équation de la chaleur vue au D.S. 5.

Le déplacement d'une particule dans une direction donnée sous l'action des chocs avec les particules voisines peut se modéliser par des déplacements successifs à droite ou à gauche équiprobables, d'une quantité strictement positive δ, qui interviennent à intervalles de temps réguliers, le temps entre deux chocs étant égal à $\tau > 0$.

Une variable aléatoire est dite de Rademacher si elle est à valeurs dans $\{1, -1\}$ et si elle prend les valeurs 1 et -1 avec la même probabilité 1/2.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables de Rademacher mutuellement indépendantes, définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On note, pour tout entier $n \geqslant 1, S_n = \sum_{j=1}^{n} X_j$. Ainsi, la variable aléatoire δS_n modélise la position de la particule au temps $n\tau$.

Pour tout $n \in \mathbb{N}^*$, on pose $Y_n = \frac{1}{2}(X_n + 1)$ et $Z_n = \sum_{j=1}^{n} Y_j$. Soit $n \in \mathbb{N}^*$.

Q 23) Déterminer la loi de la variable aléatoire Y_n et celle de la variable aléatoire Z_n. Soit k un entier tel que $-n \leqslant k \leqslant n$.

Q 24) Montrer que, si n et k ne sont pas de même parité, alors $\mathbb{P}(S_n = k) = 0$.

Q 25) Montrer que, si n et k sont de même parité, $\mathbb{P}(S_n = k) = \binom{n}{(k+n)/2}\frac{1}{2^n}$.

Pour x réel, on note $\lfloor x \rfloor$ la partie entière de x.

Q 26) Pour tous réels $\delta > 0$ et $\tau > 0$, calculer $\mathbb{V}\left(\delta S_{\lfloor 1/\tau \rfloor}\right)$, variance de la variable aléatoire $\delta S_{\lfloor 1/\tau \rfloor}$.

Q 27) Montrer que, pour tout réel $\delta, \mathbb{V}\left(\delta S_{\lfloor 1/\tau \rfloor}\right)$ est équivalent à $\frac{\delta^2}{\tau}$, lorsque τ tend vers 0 par valeurs supérieures.

Q 28) Pour tout $n \in \mathbb{N}^*$ et tout $k \in \mathbb{Z}$, en posant $p_n(k) = \mathbb{P}(S_n = k)$, montrer que

$$\frac{p_{n+1}(k) - p_n(k)}{\tau} = \frac{\delta^2}{2\tau}\,\frac{p_n(k+1) - 2p_n(k) + p_n(k-1)}{\delta^2}$$

Interprétation de ce résultat en lien avec l'équation de diffusion du D.S. 5 :

Remarquons d'abord que, pour tout $n \in \mathbb{N}^*$ et $k \in \mathbb{Z}, p_n(k)$ est la probabilité de trouver la particule à l'emplacement $k\delta$ à l'instant $n\tau$.

En passant à des variables continues, pour tout $(t, x) \in \mathbb{R}_+^* \times \mathbb{R}$, notons alors $\mathrm{P}(t, x)$ la probabilité de trouver la particule à l'emplacement x à la date t. Celle-ci vérifie

$$\forall n \in \mathbb{N}^*, \forall k \in \mathbb{Z} \quad \mathrm{P}(n\tau, k\delta) = p_n(k)$$

et on a donc, d'après le résultat de la question précédente, pour tous $n \in \mathbb{N}^*$ et $k \in \mathbb{Z}$,

$$\frac{\mathrm{P}((n+1)\tau, k\delta) - \mathrm{P}(n\tau, k\delta)}{\tau} = \frac{\delta^2}{2\tau}\,\frac{\mathrm{P}(n\tau, (k+1)\delta) - 2\mathrm{P}(n\tau, k\delta) + \mathrm{P}(n\tau, (k-1)\delta)}{\delta^2}$$

Supposons maintenant que δ et τ tendent tous deux vers 0 de façon à ce que le coefficient $D = \delta^2/2\tau$ reste constant.

Q 29) En supposant que les dérivées partielles secondes de P existent, montrer que P vérifie l'équation

$$\frac{\partial \mathrm{P}}{\partial t}(t, x) = \frac{\delta^2}{2\tau}\,\frac{\partial^2 \mathrm{P}}{\partial x^2}(t, x)$$

qui est une équation de diffusion, dont le coefficient de diffusion est $D = \delta^2/2\tau$

Par ailleurs, $\lfloor 1/\tau \rfloor$ représente le nombre de chocs subis par la particule entre les dates $t = 0$ et $t = 1$; la variable aléatoire $\delta \mathrm{S}_{\lfloor 1/\tau \rfloor}$ représente par conséquent la position de la particule à la date $t = 1$, et $\mathbb{V}\left(\delta S_{\lfloor 1/\tau \rfloor}\right)$ est un indicateur de la dispersion de ses valeurs. Or d'après le résultat de la question 27, si τ est proche de $0, \mathbb{V}\left(\delta\ \mathrm{S}_{[1/\tau]}\right)$ sera proche de $\delta^2/\tau = 2D$.

Le coefficient de diffusion D peut donc s'interpréter comme une quantité proportionnelle à la variance de la variable aléatoire donnant la position de la particule à une date donnée.