Agrégation Interne de Mathématiques, Session 2014

Deuxième Épreuve

Correction proposée par G. Dupont.

Introduction

Cette seconde épreuve de l'agrégation interne 2014 de mathématiques propose essentiellement de revisiter quelques grands résultats d'analyse à l'aide de méthodes probabilistes.

La première partie est dédiée à une preuve probabiliste du théorème d'approximation de Weierstrass [1] à l'aide des polynômes de Bernstein [2], à savoir que toute fonction d'une variable réelle continue sur un compact est limite uniforme d'une suite de fonctions polynomiales. On pourra se référer à [3, Exercice 5.6] pour retrouver une preuve similaire dans la littérature, laquelle peut faire l'objet un excellent développement lors des épreuves orales.

La deuxième partie du problème vise à étudier la qualité de l'approximation polynomiale obtenue à la partie 1 pour une classe particulière de fonctions continues, les fonctions α-Hölderiennes [3], qui sont une généralisation des fonctions lipschitziennes [4].

La troisième partie vise à montrer que pour les fonctions de classe $\mathcal{C}^1$, l'approximation uniforme obtenue à l'aide des polynômes de Bernstein commute avec la dérivation. Pour éviter tout mésusage des théorèmes de dérivabilité des limites de suites de fonctions, les premières questions de cette partie visent à montrer qu'il n'y a *a piori* pas de lien général entre convergence uniforme et régularité.

La quatrième partie est plus directement motivée par le calcul des probabilités. On y montre sous certaines hypothèses que, si l'on considère une subdivision en n segments du support d'une variable aléatoire réelle X, alors la courbe de Bézier associée aux points du graphe de la fonction de répartition F_X correspondant à cette subdivision est elle-même le graphe de la fonction de répartition d'une variable aléatoire X_n. On y montre ensuite que si la subdivision est régulière, alors la suite (X_n) converge en loi vers X puis on étudie la convergence en moyenne quadratique. Les premières questions de cette dernière partie sont dédiées à l'étude géométrique des courbes de Bézier [5], indépendamment de leur application aux probabilités.

Les parties sont relativement dépendantes les unes des autres mais le sujet est organisé de manière à ce que les résultats essentiels à la continuation de l'épreuve se retrouvent dans l'énoncé. Il est donc possible d'avancer dans le sujet tout en laissant des questions de côté.

1. Karl Weierstrass, mathématicien allemand né à Osterfelde (Westphalie) en 1815 et décédé à Berlin en 1897. Ses premiers travaux portent sur les intégrales elliptiques et les fonctions abéliennes. Il sera un professeur très influent à l'Université de Berlin. Schwarz fut l'un de ses élèves. L'un de ses résultats les plus célèbres est le théorème d'approximation polynomiale uniforme des fonctions continues sur des segments de $\mathbb{R}$, généralisé par Stone dans les années 1950.

2. Sergeï Bernstein, mathématicien ukrainien né à Odessa en 1880 et décédé à Moscou en 1968. Ses travaux portent principalement sur les équations différentielles, l'analyse fonctionnelle et la théorie des probabilités.

3. Otto Hölder, mathématicien allemand né à Stuttgart en 1859 et décédé à Leipzig en 1927. Élève de Weierstrass, Hölder travaille sur les fonctions d'une variable réelle ou complexe et en théorie de Galois. Il est à l'origine de la notion de groupe quotient.

4. Rudolphe Lipschitz, mathématicien allemand né près de Königsberg en 1832 et décédé à Bonn en 1903. Ses travaux ont principalement porté sur les équations différentielles et la géométrie riemannienne.

5. Pierre Bézier, né en 1910 et décédé en 1999 à Paris, est un ingénieur électricien et mécanicien français. Il est principalement connu pour la découverte des courbes et des surfaces qui portent son nom.

Notations et rappels

- $\mathbb{C}$ désigne l'ensemble des nombres complexes.
- $\mathbb{R}$ désigne l'ensemble des nombres réels, $\mathbb{R}_+$ l'ensemble des nombres réels positifs ou nuls, $\mathbb{R}_+^*$ l'ensemble des nombres réels positifs non nuls.
- $\mathbb{N}$ désigne l'ensemble des entiers naturels.
- $\mathbb{N}^*$ désigne l'ensemble des entiers naturels non nuls.
- $\mathbb{Z}$ désigne l'ensemble des entiers relatifs.
- $\mathbb{C}[X]$ désigne l'espace vectoriel des polynômes à coefficients complexes.
- Si $n \in \mathbb{N}$, $\mathbb{C}_n[X]$ désigne le sous-espace vectoriel de $\mathbb{C}[X]$, des polynômes de degrés inférieurs ou égaux à n.
- Pour deux réels a, b vérifiant $a \leq b$, on désigne par $[a, b]$ l'intervalle fermé d'extrémités a et b. $\mathcal{C}([a,b],\mathbb{C})$ désigne l'espace vectoriel des fonctions continues de $[a,b]$ dans $\mathbb{C}$. On notera en particulier $\mathcal{C}^0$ l'ensemble des fonctions à valeurs complexes continues sur $[0,1]$ et $\mathcal{C}^1$ l'ensemble des fonctions à valeurs complexes de classe $\mathcal{C}^1$ sur $[0,1]$. Pour $g \in \mathcal{C}^0$, on pose
$$\|g\|_\infty = \sup_{x\in[0,1]} |g(x)|.$$
- Pour tous entiers naturels p et q vérifiant $p \leq q$, l'ensemble $\{n \in \mathbb{N} \mid p \leq n \leq q\}$ est noté $[\![p, q]\!]$.
- Si $(k, n) \in \mathbb{Z} \times \mathbb{N}$, $\binom{n}{k}$ désigne le coefficient binomial dont la valeur est
$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!(n-k)!} & \text{si } 0 \leq k \leq n, \\ 0 & \text{sinon.} \end{cases}$$
- On rappelle que si $(u_k)_{k\in\mathbb{N}}$ est une suite complexe, $\displaystyle\sum_{k\in\emptyset} u_k = 0$.

Notations

- **Dans tout le problème**, f désigne une fonction continue sur $[0,1]$ à valeurs complexes.
- Pour $(k,n) \in \mathbb{N}^2$ on appelle k-ème polynôme de Bernstein d'ordre n le polynôme $B_{k,n}$ donné par :
$$B_{k,n}(X) = \binom{n}{k} X^k (1-X)^{n-k}.$$
- On considérera dans ce problème un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On rappelle qu'une variable aléatoire réelle X suit une loi de Bernoulli [6] si $X(\Omega) = \{0,1\}$. On dit que $p \in [0,1]$ est le paramètre de cette loi si $\mathbb{P}(X=1) = \mathbb{P}\left(X^{-1}\{1\}\right) = p$.
- On considérera $(X_k)_{k\in\mathbb{N}}$ une suite de variables aléatoires, réelles, indépendantes et identiquement distribuées suivant toute une loi de Bernoulli de paramètre x, où $x \in [0,1]$.

  Pour n entier naturel non nul on notera $S_n = \displaystyle\sum_{k=1}^n X_k$ et $T_n = \dfrac{S_n}{n}$.
- Pour X variable aléatoire réelle, on note, sous réserve d'existence, $\mathbb{E}(X)$ son espérance mathématique et $\mathbb{V}(X)$ sa variance.
- Soit $(Y_k)_{k\in\mathbb{N}}$ et Y des variables aléatoires réelles, de fonctions de répartition respectives F_k et F. On rappelle que $(Y_k)_{k\in\mathbb{N}}$ converge en loi vers Y si, en tout point x où F est continue, la suite $(F_k(x))_{k\in\mathbb{N}}$ converge vers $F(x)$.
- Soit $(Y_k)_{k\in\mathbb{N}}$ et Y des variables aléatoires réelles. On dit que $(Y_k)_{k\in\mathbb{N}}$ converge en moyenne quadratique vers Y si la suite $(\mathbb{E}((Y_k - Y)^2))_{k\in\mathbb{N}}$ converge vers 0.

6. Jakob (ou Jacques) Bernoulli, mathématicien suisse né à Bâle en 1654 où il décéda en 1705. Jakob Bernoulli est le premier d'une longue lignée de grands mathématiciens. Il fait partie des grands développeurs du calcul infinitésimal, en particulier dans le domaine des courbes. Il a aussi travaillé sur les séries et sur le calcul des probabilités. C'est à lui que se réfère la *loi de Bernoulli* en probabilités, la *lemniscate de Bernoulli* en géométrie, les *nombres de Bernoulli* en théorie des nombres ou encore la fameuse *inégalité de Bernoulli* qui fait régulièrement son apparition dans les sujets de concours.

Partie I : Une démonstration probabiliste du théorème de Weierstrass

1. Donner, en la justifiant, la loi de S_n.

S_n est la somme de n variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(x)$, autrement dit, elle compte le nombre de succès dans un schéma de Bernoulli de paramètre x à n épreuves. Il s'ensuit que S_n suit une loi binomiale $\mathcal{B}(n, x)$.

En d'autres termes, $S_n(\Omega) = [\![0, n]\!]$ et, pour tout $k \in [\![0, n]\!]$, on a

$$\mathbb{P}(S_n = k) = \binom{n}{k} x^k (1-x)^{n-k} = B_{k,n}(x).$$

$\square$

2. Déduire de ce premier résultat les propriétés suivantes (seules les démonstrations utilisant la question précédente seront acceptées ici) :

2.a. $\forall n \in \mathbb{N}$, $\forall k \in \mathbb{N}$, $0 \leq B_{k,n}(x) \leq 1$**,**

Si $k \notin [\![0, n]\!]$, alors $B_{k,n} = 0$. Sinon, on a vu à la question précédente que

$$B_{k,n}(x) = \mathbb{P}(S_n = k),$$

or la probabilité d'un événement est toujours un réel compris entre 0 et 1. $\square$

2.b. $\forall n \in \mathbb{N}$, $\forall k \in \mathbb{N}$, $B_{k,n}(x) = B_{n-k,n}(1-x)$**,**

Si $k \notin [\![0, n]\!]$, alors $(n-k) \notin [\![0, n]\!]$ et donc

$$B_{k,n} = B_{n-k,n} = 0.$$

Sinon, il suit de la question 2.a. que $B_{k,n}(x)$ est égal à la probabilité de faire k succès dans un schéma de Bernoulli de paramètre x à n épreuves. Or faire k succès dans un tel schéma revient à faire $n-k$ échecs. Chaque échec ayant une probabilité $(1-x)$ de se réaliser, $B_{k,n}(x)$ est aussi égal à la probabilité de faire $n-k$ succès dans un schéma de Bernoulli de paramètre $(1-x)$ à n épreuves. Il suit donc de la question 2.a. que

$$B_{k,n}(x) = B_{n-k,n}(1-x).$$

$\square$

2.c. et les valeurs de :

2.c.i. $\sum_{k=0}^{n} B_{k,n}(x)$**,**

On a

$$\begin{aligned}\sum_{k=0}^{n} B_{k,n}(x) &= \sum_{k=0}^{n} \mathbb{P}\left(S_n = k\right) \\ &= \mathbb{P}\left(\bigsqcup_{k=0}^{n} (S_n = k)\right) \\ &= \mathbb{P}\left(S_n \in [\![0,n]\!]\right) \\ &= \mathbb{P}\left(S_n^{-1}(S_n(\Omega))\right) \\ &= \mathbb{P}\left(\Omega\right) \\ &= 1.\end{aligned}$$

□

2.c.ii. $\sum_{k=0}^{n} kB_{k,n}(x)$**,**

On a

$$\sum_{k=0}^{n} kB_{k,n}(x) = \sum_{k=0}^{n} k\mathbb{P}\left(S_n = k\right) = \mathbb{E}(S_n) = nx,$$

la dernière égalité venant du fait que $S_n \rightsquigarrow \mathcal{B}(n,x)$. □

2.c.iii. $\sum_{k=0}^{n} (k-nx)^2 B_{k,n}(x)$**,**

On a

$$\begin{aligned}\sum_{k=0}^{n}(k-nx)^2 B_{k,n}(x) &= \sum_{k=0}^{n}(k-\mathbb{E}(S_n))^2\mathbb{P}\left(S_n = k\right), \text{ car } \mathbb{E}S_n = nx, \\ &= \mathbb{E}((S_n - \mathbb{E}(S_n))^2), \text{ d'après le théorème de transfert,} \\ &= \mathbb{V}(X), \text{ par définition,} \\ &= nx(1-x), \text{ car } S_n \rightsquigarrow \mathcal{B}(n,x).\end{aligned}$$

□

3. Pour $n \in \mathbb{N}^*$ et $k \in [\![0,n]\!]$ calculer $\mathbb{P}\left(S_n = k\right)$ en fonction de $\mathbb{P}\left(S_{n-1} = k\right)$ et $\mathbb{P}\left(S_{n-1} = k-1\right)$, et en déduire l'expression de $B_{k,n}$ en fonction de $B_{k,n-1}$ et $B_{k-1,n-1}$.

Traitons d'abord le cas $k = 0$. On a alors

$$(S_n = 0) = (S_{n-1} = 0,\, X_n = 0)$$

et donc

$$\mathbb{P}\left(S_n = 0\right) = \mathbb{P}\left(S_{n-1} = 0, X_n = 0\right).$$

Puisque les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes, les variables aléatoires S_{n-1} et X_n sont aussi indépendantes et donc

$$\mathbb{P}\left(S_n = 0\right) = \mathbb{P}\left(S_{n-1} = 0\right)\mathbb{P}\left(X_n = 0\right) = (1-x)B_{0,n-1}(x) = (1-x)B_{k,n-1}(x).$$

Supposons maintenant $k \in [\![1,n]\!]$. On a

$$(S_n = k) = (S_{n-1} = k,\, X_n = 0) \sqcup (S_{n-1} = k-1,\, X_n = 1)$$

donc

$$\mathbb{P}\left(S_n = k\right) = \mathbb{P}\left(S_{n-1} = k,\, X_n = 0\right) + \mathbb{P}\left(S_{n-1} = k-1,\, X_n = 1\right).$$

Puisque les variables aléatoires $X_1, \dots, X_n$ sont indépendantes, les variables aléatoires S_{n-1} et X_n sont aussi indépendantes. Il s'ensuit que

$$\mathbb{P}(S_n = k) = \mathbb{P}(S_{n-1} = k)\,\mathbb{P}(X_n = 0) + \mathbb{P}(S_{n-1} = k-1)\,\mathbb{P}(X_n = 1).$$

Or $X_n \rightsquigarrow \mathcal{B}(x)$, $S_{n-1} \rightsquigarrow \mathcal{B}(n-1, x)$ et $S_n \rightsquigarrow \mathcal{B}(n, x)$ donc il suit de la question 1 que

$$B_{k,n}(x) = (1-x)B_{k,n-1}(x) + xB_{k-1,n-1}(x).$$

Puisque $B_{-1,n} = 0$ pour tout $n \in \mathbb{N}^*$, on a bien

$$B_{k,n}(x) = (1-x)B_{k,n-1}(x) + xB_{k-1,n-1}(x), \quad \forall n \in \mathbb{N}^*, \forall k \in [\![0, n]\!].$$

$\square$

4. Pour $n > 0$, on définit une application B_n de $\mathcal{C}^0$ vers $\mathbb{C}[X]$ par

$$\forall g \in \mathcal{C}^0, \quad B_n(g) = \sum_{k=0}^{n} g\left(\frac{k}{n}\right) B_{k,n}.$$

4.a. Montrer que $(B_{k,n})_{0 \leq k \leq n}$ est une base de $\mathbb{C}_n[X]$.

La famille $(B_{k,n})_{0 \leq k \leq n}$ est de cardinal $(n+1)$ et $\mathbb{C}_n[X]$ est de dimension $(n+1)$, il suffit donc de montrer qu'elle est libre ou génératrice pour montrer que c'est une base. Montrons qu'elle est libre.

Supposons qu'il existe des complexes $\lambda_0, \dots, \lambda_n$ tels que

$$\Lambda(X) = \sum_{k=0}^{n} \lambda_k B_{k,n}(X) = 0.$$

Puisque Λ est nul, il s'ensuit en particulier que $\Lambda(0) = 0$. Or $\Lambda(0) = \sum_{k=0}^{n} \lambda_k B_{k,n}(0) = \lambda_0$ car $B_{k,n}(0) = 0$ pour tout $k > 0$ et $B_{0,0}(0) = 1$.

Ainsi,

$$0 = \Lambda(X) = \sum_{k=1}^{n} \lambda_k B_{k,n}(X) = X \sum_{k=1}^{n} \lambda_k \binom{n}{k} X^{k-1}(1-X)^{n-k}.$$

On pose

$$\Lambda_1(X) = \sum_{k=1}^{n} \lambda_k \binom{n}{k} X^{k-1}(1-X)^{n-k}.$$

Puisque $0 = \Lambda(X) = X\Lambda_1(X)$ et que $\mathbb{C}[X]$ est intègre, il vient $\Lambda_1(X) = 0$. On évalue alors en 0 et on obtient $\lambda_1 = 0$. Une récurrence forte montre alors que $\lambda_0 = \lambda_1 = \dots = \lambda_n = 0$. $\square$

4.b. Montrer que la restriction de B_n à $\mathbb{C}_n[X]$ induit un automorphisme linéaire $\overline{B_n}$ de $\mathbb{C}_n[X]$.

Soient $f, g \in \mathcal{C}^0$ et $\lambda \in \mathbb{C}$. Il suit immédiatement de la définition que $B_n(f) \in \mathbb{C}_n[X]$, $B_n(g) \in \mathbb{C}_n[X]$ et

$$B_n(\lambda f + g) = \lambda B_n(f) + B_n(g)$$

de sorte que $B_n \in \mathrm{Hom}_{\mathbb{C}}(\mathcal{C}^0, \mathbb{C}_n[X])$ induit $\overline{B_n} \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}_n[X])$.

Puisque $\overline{B_n}$ est un endomorphisme d'un espace vectoriel de dimension finie, il s'agit d'un automorphisme si et seulement s'il est injectif. Déterminons $\mathrm{Ker}(\overline{B_n})$.

Soit $P \in \mathbb{C}_n[X]$. Alors

$$\begin{aligned}
\overline{B_n}(P) = 0 &\Leftrightarrow B_n(P) = 0 \\
&\Leftrightarrow \sum_{k=0}^{n} P\left(\frac{k}{n}\right) B_{k,n} = 0 \\
&\Leftrightarrow P\left(\frac{k}{n}\right) = 0,\ \forall k \in [\![0, n]\!],\ \text{ du fait de la liberté de } (B_{k,n})_{0 \leq k \leq n}, \\
&\Leftrightarrow P = 0,
\end{aligned}$$

car P, de degré inférieur ou égal à n, ne peut avoir $(n+1)$ racines s'il est non-nul. □

4.c. Montrer que

$$\forall P \in \mathbb{C}[X],\ \exists Q \in \mathbb{C}[X],\ \exists n \in \mathbb{N},\ P = B_n(Q).$$

Un tel Q est-il unique ?

Soit $P \in \mathbb{C}[X]$. Posons $n = \deg(P)$ de sorte que $P \in \mathbb{C}_n[X]$. Puisque $\overline{B_n}$ est une bijection de $\mathbb{C}_n[X]$ dans $\mathbb{C}_n[X]$, il existe alors (un unique) $Q_n \in \mathbb{C}_n[X]$ tel que $B_n(Q_n) = P$.

Considérons le polynôme

$$R = \prod_{k=0}^{n} \left(X - \frac{k}{n}\right) \in \mathbb{C}_{n+1}[X].$$

Alors

$$B_n(R) = 0$$

et, par linéarité de B_n,

$$B_n(R + Q_n) = B_n(Q_n) = P.$$

Il n'y a donc pas unicité de $Q \in \mathbb{C}[X]$ tel que $B_n(Q) = P$.

Remarque : On notera qu'il existe bien un unique $Q_n \in \mathbb{C}_n[X]$ tel que $B_n(Q_n) = P$, mais on perd cette unicité dès que l'on ôte la restriction sur le degré.

□

5. On rappelle ici que $S_n = \sum_{k=1}^{n} X_k$, que $T_n = \frac{S_n}{n}$ et que les variables aléatoires $(X_k)_{k \in \mathbb{N}}$ sont indépendantes et identiquement distribuées selon une loi de Bernoulli de paramètre x.
Soit f une fonction continue de $[0, 1]$ dans $\mathbb{C}$. Montrer que

$$\mathbb{E}(f(T_n)) = B_n(f)(x)$$

puis que

$$B_n(f)(x) - f(x) = \mathbb{E}(f(T_n) - f(\mathbb{E}(T_n))).$$

On a

$$\begin{aligned}
\mathbb{E}(f(T_n)) &= \mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right) \\
&= \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \mathbb{P}\left(S_n = k\right),\ \text{ d'après le théorème de transfert,} \\
&= \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{k,n}(x),\ \text{ d'après la question 1,} \\
&= B_n(f)(x).
\end{aligned}$$

D'autre part,

$$\mathbb{E}(T_n) = \mathbb{E}\left(\frac{S_n}{n}\right) = \frac{\mathbb{E}(S_n)}{n} = \frac{nx}{n} = x$$

donc

$$B_n(f)(x) - f(x) = \mathbb{E}(f(T_n)) - f(\mathbb{E}(T_n)) = \mathbb{E}\left[f(T_n) - f(\mathbb{E}(T_n))\right].$$

□

6. Monter que $\forall \delta > 0$, $\mathbb{P}\left(|T_n - x| \geq \delta\right) \leq \dfrac{1}{4n\delta^2}$.

Ce point est une conséquence directe de la célèbre *inégalité de Tchebychev* : Si une variable aléatoire réelle X admet un moment d'ordre 2, alors

$$\mathbb{P}\left(|X - \mathbb{E}X| \geq \delta\right) \leq \frac{\mathbb{V}(X)}{\delta^2}.$$

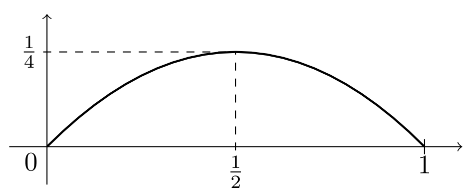
Nous avons vu à la question précédente que $\mathbb{E}(T_n) = x$. Par ailleurs, la variable aléatoire S_n suivant une loi binomiale, elle admet un moment d'ordre 2 et donc T_n aussi. En outre

$$\mathbb{V}(T_n) = \mathbb{V}\left(\frac{S_n}{n}\right) = \frac{\mathbb{V}(S_n)}{n^2} = \frac{nx(1-x)}{n^2} = \frac{x(1-x)}{n}.$$

En appliquant l'inégalité de Tchebychev à T_n, il vient

$$\begin{aligned}\mathbb{P}\left(|T_n - x| \geq \delta\right) &= \mathbb{P}\left(|T_n - \mathbb{E}T_n| \geq \delta\right) \\ &\leq \frac{\mathbb{V}(T_n)}{\delta^2} \\ &= \frac{x(1-x)}{n\delta^2}.\end{aligned}$$

Il ne reste plus qu'à observer que la fonction $x \mapsto x(1-x)$ atteint son maximum sur $[0,1]$ en $\dfrac{1}{2}$ où elle vaut $\dfrac{1}{4}$.



Ainsi,

$$\mathbb{P}\left(|T_n - x| \geq \delta\right) \leq \frac{1}{4n\delta^2}.$$

□

7. Soit X une variable aléatoire réelle à valeurs dans un ensemble fini et ϕ une fonction définie et convexe sur un intervalle contenant $X(\Omega)$. On suppose de plus que X et $\phi(X)$ possèdent une espérance. Montrer qu'alors

$$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X)).$$

On note que, pour tout $k \in X(\Omega)$, $\mathbb{P}(X = k) \geq 0$ et $\sum_{k \in X(\Omega)} \mathbb{P}(X = k) = 1$. Ainsi, $\sum_{k \in X(\Omega)} k\mathbb{P}(X = k)$ est une combinaison barycentrique des éléments de $X(\Omega)$. Il suit donc de l'inégalité de Jensen [7] (qui se prouve par

7. Johan Ludwig William Valdemar Jensen, ingénieur danois, né en 1859 à Nakskov et décédé en 1925 à Copenhague. Il n'était mathématicien qu'à temps perdu.

récurrence à partir de l'inégalité de convexité) que

$$\begin{aligned}\phi(\mathbb{E}(X)) &= \phi\left(\sum_{k\in X(\Omega)} k\mathbb{P}(X=k)\right)\\ &\leq \sum_{k\in X(\Omega)} \phi(k)\mathbb{P}(X=k)\\ &= \mathbb{E}(\phi(X)), \text{ d'après le théorème de transfert.}\end{aligned}$$

□

8. Déduire de ce qui précède que $(B_n(f))_{n\in\mathbb{N}^*}$ converge uniformément vers f.

Soit $n \in \mathbb{N}^*$ et $x \in [0,1]$. Il suit de la question 5 que

$$|B_n(f)(x) - f(x)| = |\mathbb{E}[f(T_n) - f(\mathbb{E}(T_n))]|.$$

Fixons $\epsilon > 0$.

f étant continue sur le compact $[0,1]$, elle y est uniformément continue. Il existe donc $\delta > 0$ tel que

$$|t-x| \leq \delta \Rightarrow |f(t) - f(x)| < \frac{\epsilon}{2}, \forall t \in [0,1],$$

uniformément en x.

Ainsi,

$$\begin{aligned}|\mathbb{E}[f(T_n) - f(\mathbb{E}(T_n))]| &\leq \mathbb{E}|f(T_n) - f(\mathbb{E}(T_n))|\\ &= \mathbb{E}\left[\mathbb{1}_{(|T_n - x|<\delta)}|f(T_n) - f(x)|\right]\\ &\quad + \mathbb{E}\left[\mathbb{1}_{(|T_n - x|\geq\delta)}|f(T_n) - f(x)|\right]\\ &< \frac{\epsilon}{2}\mathbb{P}(|T_n - x| < \delta) + M\mathbb{P}(|T_n - x| \geq \delta),\end{aligned}$$

où

$$M = \sup_{(x,y)\in[0,1]^2} |f(y) - f(x)|$$

est bien défini par f est continue sur le segment $[0,1]$.

Il suit alors de la question 6 que

$$|\mathbb{E}(f(T_n) - f(\mathbb{E}(T_n))| < \frac{\epsilon}{2} + \frac{M}{4n\delta^2}.$$

Puisque $\frac{M}{4n\delta^2} \xrightarrow[n\to+\infty]{} 0$, il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$,

$$n \geq n_0 \Rightarrow \frac{M}{4n\delta^2} < \frac{\epsilon}{2}.$$

Ainsi, pour tout $n \in \mathbb{N}$,

$$n \geq n_0 \Rightarrow |\mathbb{E}(f(T_n) - f(\mathbb{E}(T_n))| < \epsilon.$$

Ceci étant valable uniformément en $x \in [0,1]$, n_0 ne dépend pas de x et il s'ensuit que, pour tout $n \in \mathbb{N}$,

$$n \geq n_0 \Rightarrow \|B_n(f) - f\|_\infty < \epsilon,$$

autrement dit,

$$\|B_n(f) - f\|_\infty \xrightarrow[n\to+\infty]{} 0.$$

□

9. Démontrer le théorème de Weierstrass : Soient $(a,b) \in \mathbb{R}^2$, avec $a < b$. Si g est continue sur $[a,b]$, alors g est limite uniforme sur $[a,b]$ d'une suite de polynômes.

Soit $g \in \mathcal{C}([a,b],\mathbb{C})$. Considérons l'homéomorphisme affine

$$\phi : \left\{ \begin{array}{rcl} [0,1] & \longrightarrow & [a,b] \\ t & \mapsto & ta+(1-t)b \end{array} \right.$$

et posons

$$f = g \circ \phi$$

de sorte que $f \in \mathcal{C}^0$. Il suit de la question 8 que $B_n(f)$ converge uniformément vers f. En composant avec ϕ^{-1}, il s'ensuit que $B_n(f) \circ \phi^{-1}$ converge uniformément vers $f \circ \phi^{-1} = g$. Puisque ϕ est affine, ϕ^{-1} l'est aussi et donc chaque $B_n(f) \circ \phi^{-1}$ est polynomiale, ce qui montre le théorème de Weierstrass. $\square$

10. Montrer que ce résultat n'est pas valable si on remplace $[a,b]$ par $\mathbb{R}$.

Considérons la fonction sin et supposons qu'elle est limite uniforme de polynômes sur $\mathbb{R}$. Alors il existe un polynôme P tel que

$$\| \sin - P\|_\infty < 1,$$

la norme infinie étant prise sur $\mathbb{R}$.

Puisque, pour tout $k \in \mathbb{Z}$, on a

$$\sin\left(\frac{\pi}{2} + 2k\pi\right) = 1 \text{ et } \sin\left(-\frac{\pi}{2} + 2k\pi\right) = -1,$$

il s'ensuit que

$$P\left(\frac{\pi}{2} + 2k\pi\right) > 0 \text{ et } P\left(-\frac{\pi}{2} + 2k\pi\right) < 0.$$

Le polynôme P étant continu, il suit du théorème des valeurs intermédiaires qu'il admet un zéro sur chaque intervalle de la forme

$$\left]-\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + 2k\pi\right[, \quad k \in \mathbb{Z}.$$

Ainsi P admet une infinité de racines, c'est donc nécessairement le polynôme nul. On obtient alors

$$1 = \| \sin \|_\infty < 1,$$

une contradiction.

Remarque : On peut aller un peu plus loin et montrer que les limites uniformes de polynômes sur $\mathbb{R}$ sont les polynômes. En effet, si $f \in \mathcal{C}(\mathbb{R},\mathbb{C})$ est limite uniforme d'une suite $(P_n)_{n\in\mathbb{N}}$ de fonctions polynomiales, alors la suite $(P_n)_{n\in\mathbb{N}}$ étant uniformément convergente (vers f), elle satisfait le critère de Cauchy uniforme. En particulier, il existe $n_0 \in \mathbb{N}$ tel que, pour tous $n, p \in \mathbb{N}$, on a

$$n \geq n_0 \Rightarrow \|P_{n+p} - P_n\|_\infty \leq 1,$$

la norme infinie étant prise sur $\mathbb{R}$.

Ainsi, pour tout $n \geq n_0$ et tout $p \in \mathbb{N}$, $(P_{n+p} - P_n)$ est un polynôme borné, il est donc constant. Il existe donc $c_{n,p} \in \mathbb{C}$ tel que

$$P_{n+p} - P_n = c_{n,p}.$$

Puisque la suite $(P_{n+p})_{p\in\mathbb{N}}$ converge vers f, la suite $(c_{n,p})_{p\in\mathbb{N}}$ converge aussi vers un nombre complexe c_n. En passant à la limite en p dans l'égalité ci-dessus, on obtient

$$f - P_n = c_n,$$

autrement dit $f = P_n + c_n$ est polynomiale. $\square$

Partie II : Approximation des fonctions Hölderiennes

Soit $\alpha \in]0,1]$. Une fonction g, définie sur $[0,1]$ et à valeurs complexes, est dite α-Hölderienne sur $[0,1]$ s'il existe $L > 0$ tel que

$$\forall (x,y) \in [0,1]^2,\ |g(x) - g(y)| \leq L|x-y|^\alpha.$$

Notons Lip_α l'ensemble des fonctions α-Hölderiennes sur $[0,1]$.

A. Généralités sur les fonctions Hölderiennes

11. Soit $\alpha \in]0,1]$. Montrer que la fonction g définie sur $[0,1]$ à valeurs dans $\mathbb{C}$ par $g(x) = x^\alpha$ est α-Hölderienne puis que Lip_α est un $\mathbb{C}$-espace vectoriel.

Soit $x \in [0,1]$ et considérons la fonction

$$\Delta_x : \left\{ \begin{array}{rcl} [x,1] & \longrightarrow & \mathbb{R} \\ y & \mapsto & y^\alpha - x^\alpha - (y-x)^\alpha. \end{array} \right.$$

Alors Δ_x est dérivable et, pour tout $y \in [x,1]$, on a

$$\Delta_x'(y) = \alpha \left(y^{\alpha-1} - (y-x)^{\alpha-1} \right)$$

or $\alpha - 1 \leq 0$ donc $t \mapsto t^{\alpha-1}$ est décroissante sur $[0,1]$ et $(y-x) \leq y$ de sorte que

$$\Delta_x'(y) \leq 0, \quad \forall y \in [x,1].$$

Puisque $\Delta_x(x) = 0$, il s'ensuit que

$$\Delta_x(y) \leq 0, \quad \forall y \in [x,1],$$

c'est-à-dire

$$y^\alpha - x^\alpha \leq (y-x)^\alpha$$

et donc $g \in \mathrm{Lip}_\alpha$.

Montrons maintenant que Lip_α est un $\mathbb{C}$-espace vectoriel. Pour cela, nous allons montrer qu'il s'agit d'un sous-espace vectoriel de l'ensemble des applications de $[0,1]$ à valeurs complexes. La fonction nulle est clairement α-Hölderienne donc $\mathrm{Lip}_\alpha \neq \emptyset$. Fixons $f, g \in \mathrm{Lip}_\alpha$ et $\lambda \in \mathbb{C}$. Alors il existe $L_f, L_g > 0$ tels que, pour tous $x, y \in [0,1]$, on a

$$|f(y) - f(x)| \leq L_f |y-x|^\alpha \text{ et } |g(y) - g(x)| \leq L_g |y-x|^\alpha.$$

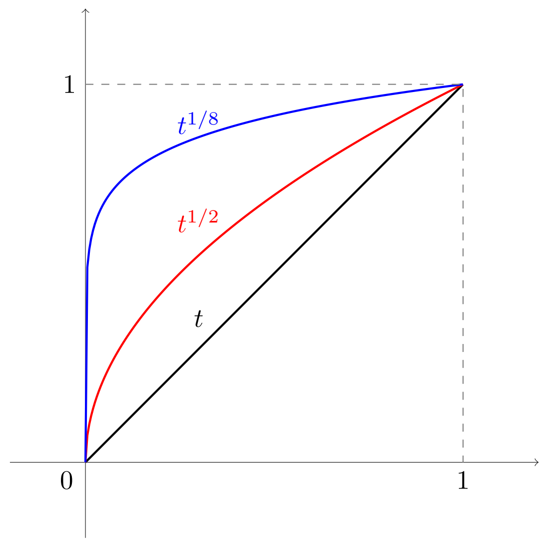
Ainsi,

$$\begin{aligned} |(\lambda f + g)(y) - (\lambda f + g)(x)| &= |\lambda(f(y) - f(x)) + (g(y) - g(x))| \\ &\leq |\lambda(f(y) - f(x))| + |g(y) - g(x)| \\ &= |\lambda||f(y) - f(x)| + |g(y) - g(x)| \\ &\leq |\lambda| L_f |y-x|^\alpha + L_g |y-x|^\alpha \\ &= (|\lambda| L_f + L_g)|y-x|^\alpha. \end{aligned}$$

Et puisque $(|\lambda| L_f + L_g) > 0$, ceci montre que $\lambda f + g \in \mathrm{Lip}_\alpha$. □

12. Soient α et β deux réels tels que $0 < \alpha \leq \beta \leq 1$. Montrer que $\mathrm{Lip}_\beta \subset \mathrm{Lip}_\alpha$. En déduire que l'ensemble des $\alpha \in]0,1]$ tels qu'une fonction est α-Hölderienne sur $[0,1]$ est soit vide soit un sous-ensemble convexe de $]0,1]$.

On commence par noter que, pour tout $t \in [0,1]$, $\alpha \leq \beta \Rightarrow t^\beta \leq t^\alpha$. En particulier, pour tous $x, y \in [0,1]$, puisque $|y-x| \leq 1$, on a $|y-x|^\beta \leq |y-x|^\alpha$.



Soit $f \in \mathrm{Lip}_\beta$. Il existe alors $L > 0$ tel que, pour tous $x, y \in [0,1]$, on a

$$|f(y) - f(x)| \leq L|y-x|^\beta \leq L|y-x|^\alpha$$

et donc $f \in \mathrm{Lip}_\alpha$.

Soit f une fonction de $[0,1]$ dans $\mathbb{C}$. Considérons

$$A = \{\alpha \in]0,1] \mid f \in \mathrm{Lip}_\alpha\}.$$

Supposons A non-vide. Alors A étant majoré par 1, il possède une borne supérieure que l'on note α_0. Il suit de la discussion précédente qu'alors $A =]0, \alpha_0[$ ou $A =]0, \alpha_0]$. Dans les deux cas A est convexe.

Remarque : Une conséquence de ce résultat est que $f \in \mathrm{Lip}_1 \Rightarrow f \in \mathrm{Lip}_\alpha$, pour tout $\alpha \in]0,1]$. Mais comme le sous-entend la notation, Lip_1 n'est autre que l'ensemble des fonctions lipschitziennes sur $[0,1]$. Ainsi, une fonction lipschitzienne sur $[0,1]$ est α-Hölderienne pour n'importe quel $\alpha \in]0,1]$.

$\square$

13. Montrer que pour tout $\alpha \in]0,1]$,

$$\mathcal{C}^1 \subset \mathrm{Lip}_\alpha \subset \mathcal{C}^0.$$

Les inclusions précédentes sont-elles strictes ? Justifier votre réponse.

Supposons $f \in \mathcal{C}^1$ de sorte que f est dérivable et $f' \in \mathcal{C}^0$. En particulier, f' est bornée sur $[0,1]$. L'inégalité des accroissements finis implique donc que, pour tous $x, y \in [0,1]$, on a

$$|f(y) - f(x)| \leq \|f'\|_\infty |y-x|.$$

Ainsi, f est lipschitzienne et donc α-Hölderienne, pour tout $\alpha \in]0,1]$ (voir remarque à la question 12).

Montrons la seconde inclusion. Soit $\alpha \in]0,1]$ et $f \in \mathrm{Lip}_\alpha$. Fixons $\epsilon > 0$ et posons $\delta = \epsilon^{1/\alpha}$. Alors puisque $f \in \mathrm{Lip}_\alpha$, il existe $L > 0$ tel que, pour tous $x, y \in [0,1]$, on a

$$|y - x| \leq \delta \Rightarrow |f(y) - f(x)| \leq L|y-x|^\alpha \leq \delta^\alpha = \epsilon.$$

En particulier, f est uniformément continue sur $[0,1]$ et donc continue.

Montrons que les inclusions sont strictes. Pour tout $\alpha \in]0,1]$, la fonction $x \mapsto x^\alpha$ est dans Lip_α d'après la question 11 mais elle n'est pas dérivable en 0. Ainsi, la première inclusion est stricte.

Pour la seconde, considérons $\alpha \in]0,1]$ et posons $\beta = \frac{\alpha}{2}$ de sorte que $\beta < \alpha$. Alors la fonction $g : x \mapsto x^\beta$ est une fonction continue mais elle n'est pas α-Hölderienne car, pour tout $x \in]0,1[$,

$$\frac{|g(x) - g(0)|}{|x - 0|^\alpha} = x^{\beta - \alpha} \xrightarrow[x:0^+]{} +\infty.$$

Il ne peut donc exister $L > 0$ tel que, pour tous $x, y \in [0,1]$,

$$|g(x) - g(y)| \leq L|x-y|^\alpha.$$

Ainsi, $g \in \mathcal{C}^0 \setminus \mathrm{Lip}_\alpha$.

□

B. Une majoration de l'erreur

Pour $L > 0$, considérons l'ensemble

$$\mathrm{Lip}_\alpha(L) = \left\{ g : [0,1] \longrightarrow \mathbb{C} \mid \forall (x,y) \in [0,1]^2,\ |g(x) - g(y)| \leq L|x-y|^\alpha \right\}.$$

14. Montrer que pour tout $g \in \mathrm{Lip}_\alpha(L)$, la suite de terme général $B_n(g)$ converge uniformément vers g.

Si $g \in \mathrm{Lip}_\alpha(L)$, alors $g \in \mathrm{Lip}_\alpha$ et donc $g \in \mathcal{C}^0$ d'après la question 13. Il suit alors de la question 8 que la suite $(B_n(g))_n$ converge uniformément vers g. □

On s'intéresse alors à une majoration de l'erreur commise lorsqu'on remplace $g \in \mathrm{Lip}_\alpha(L)$ par son approximation $B_n(g)$.

15. Démontrer que si $g \in \mathrm{Lip}_\alpha(L)$ alors, pour tout $x \in [0,1]$, on a $|g(x) - B_n(g)(x)| \leq L\mathbb{E}(|x - T_n|^\alpha)$.

Si $g \in \mathrm{Lip}_\alpha(L)$, alors elle est continue et il suit de la question 5 que, pour tout $x \in [0,1]$,

$$\begin{aligned}
|B_n(g)(x) - g(x)| &= |\mathbb{E}\left[g(T_n) - g(\mathbb{E}(T_n))\right]| \\
&\leq \mathbb{E}\left|g(T_n) - g(\mathbb{E}(T_n))\right| \\
&\leq \mathbb{E}\left[L\left|T_n - \mathbb{E}(T_n)\right|^\alpha\right], \text{ car } g \in \mathrm{Lip}_\alpha(L), \\
&= L\mathbb{E}\left[|T_n - \mathbb{E}(T_n)|^\alpha\right], \text{ par linéarité de l'espérance}, \\
&= L\mathbb{E}\left[|T_n - x|^\alpha\right], \text{ car } \mathbb{E}(T_n) = x.
\end{aligned}$$

□

16. Inégalité de Hölder

16.a. Soit g une fonction définie sur $\mathbb{R}_+$ à valeurs dans $\mathbb{R}$ et strictement concave (i.e. vérifiant $\forall x \neq y$, $\forall \alpha \in]0,1[$, $g(\alpha x + (1-\alpha)y) > \alpha g(x) + (1-\alpha)g(y)$) et soit h la fonction définie sur $\mathbb{R}_+ \times \mathbb{R}_+^*$ par

$$h(x,y) = yg\left(\frac{x}{y}\right).$$

Montrer par récurrence sur $n \in \mathbb{N}^*$ que, pour tout $(x_1, \ldots, x_n) \in (\mathbb{R}_+)^n$ et tout $(y_1, \ldots, y_n) \in (\mathbb{R}_+^*)^n$, on a

$$\sum_{i=1}^n h(x_i, y_i) \leq h\left(\sum_{i=1}^n x_i, \sum_{i=1}^n y_i\right),$$

avec égalité si, et seulement si, les n-uplets

$$(x_1, \ldots, x_n) \in (\mathbb{R}_+)^n$$

et

$$(y_1, \ldots, y_n) \in (\mathbb{R}_+^*)^n$$

sont proportionnels.

Si $n = 1$, il n'y a rien à montrer. Montrons la propriété pour $n = 2$. On a

$$h(x_1, y_1) + h(x_2, y_2) = y_1 g\left(\frac{x_1}{y_1}\right) + y_2 g\left(\frac{x_2}{y_2}\right)$$

mais

$$\frac{1}{y_1 + y_2}\left(y_1 g\left(\frac{x_1}{y_1}\right) + y_2 g\left(\frac{x_2}{y_2}\right)\right) = \frac{y_1}{y_1 + y_2} g\left(\frac{x_1}{y_1}\right) + \frac{y_2}{y_1 + y_2} g\left(\frac{x_2}{y_2}\right).$$

Or si, $\frac{x_1}{y_1} \neq \frac{x_2}{y_2}$, c'est-à-dire si (x_1, y_1) et (x_2, y_2) ne sont pas proportionnels, la stricte concavité de g implique que

$$\frac{y_1}{y_1 + y_2} g\left(\frac{x_1}{y_1}\right) + \frac{y_2}{y_1 + y_2} g\left(\frac{x_2}{y_2}\right) < g\left(\frac{x_1}{y_1 + y_2} + \frac{x_2}{y_1 + y_2}\right) = g\left(\frac{x_1 + x_2}{y_1 + y_2}\right).$$

Puisque $y_1 + y_2 > 0$, on obtient

$$y_1 g\left(\frac{x_1}{y_1}\right) + y_2 g\left(\frac{x_2}{y_2}\right) < (y_1 + y_2) g\left(\frac{x_1 + x_2}{y_1 + y_2}\right),$$

c'est-à-dire

$$h(x_1, y_1) + h(x_2, y_2) < h(x_1 + x_2, y_1 + y_2).$$

Il reste à observer que l'on a bien égalité quand les deux couples sont proportionnels, c'est-à-dire quand $\frac{x_1}{y_1} = \frac{x_2}{y_2}$.

Supposons maintenant la propriété montrée pour un certain $n \in \mathbb{N}^*$ et montrons la pour $n + 1$. On a

$$\begin{aligned}
\sum_{i=1}^{n+1} h(x_i, y_i) &= h(x_{n+1}, y_{n+1}) + \sum_{i=1}^{n} h(x_i, y_i) \\
&\leq h(x_{n+1}, y_{n+1}) + h\left(\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} y_i\right), \text{ d'après l'hypothèse au rang } n, \\
&\leq h\left(\sum_{i=1}^{n+1} x_i, \sum_{i=1}^{n+1} y_i\right), \text{ car la propriété a été montrée au rang 2,}
\end{aligned}$$

ce qui prouve l'hérédité. Il ne reste plus qu'à observer que l'on a égalité si et seulement si les deux inégalités dans la chaîne sont des égalités, auquel cas on a proportionnalité à chacune des deux étapes. Il s'ensuit que l'on a égalité si et seulement si les n-uplets sont proportionnels. $\square$

16.b. Soient p et q deux réels strictement positifs tels que $\frac{1}{p} + \frac{1}{q} = 1$. Déduire de ce qui précède que pour tout $(u_1, \ldots, u_n) \in \mathbb{R}^n$ et pour tout $(v_1, \ldots, v_n) \in \mathbb{R}^n$,

$$\sum_{k=1}^{n} |u_k v_k| \leq \left(\sum_{k=1}^{n} |u_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |v_k|^q\right)^{\frac{1}{q}}.$$

À quelle condition y a-t-il égalité dans l'inégalité précédente ?

On pose $g : x \mapsto x^{1/p}$ qui est strictement concave sur $\mathbb{R}_+$ car, pour tout $x \in \mathbb{R}_+^*$, on a

$$g''(x) = \frac{1}{p}\left(\frac{1}{p} - 1\right) x^{1/p-2} < 0,$$

puisque $\frac{1}{p} < 1$.

Considérons la fonction h définie à la question 16.a de sorte que

$$h(x, y) = y \frac{x^{1/p}}{y^{1/p}} = x^{1/p} y^{1-1/p} = x^{1/p} y^{1/q}.$$

Pour tous $(u_1, \ldots, u_n), (v_1, \ldots, v_n) \in \mathbb{R}^n$, on a

$$(|u_1|, \ldots, |u_n|), (|v_1|, \ldots, |v_n|) \in \mathbb{R}_+^n$$

et quitte à réduire n, on peut supposer tous les v_i non-nuls de sorte que l'on a

$$(|v_1|, \dots, |v_n|) \in (\mathbb{R}_+^*)^n.$$

Il suit alors de la question 16.a que

$$\sum_{i=1}^n h(|u_i|^p, |v_i|^q) \leq h\left(\sum_{i=1}^n |u_i|^p, \sum_{i=1}^n |v_i|^q\right) = \left(\sum_{i=1}^n |u_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |v_i|^q\right)^{\frac{1}{q}}.$$

En outre, on a montré à la question 16.a que l'on a égalité si et seulement si les n-uplets $(|u_1|^p, \dots, |u_n|^p)$ et $(|v_1|^q, \dots, |v_n|^q)$ sont colinéaires. $\square$

17. Montrer que

$$\mathbb{E}\left(|x - T_n|^\alpha\right) \leq \left(\sum_{k=0}^n \left|x - \frac{k}{n}\right|^2 \mathbb{P}\left(T_n = \frac{k}{n}\right)\right)^{\frac{\alpha}{2}} \left(\sum_{k=0}^n \mathbb{P}\left(T_n = \frac{k}{n}\right)\right)^{\frac{2-\alpha}{2}}.$$

En déduire que

$$\mathbb{E}\left(|x - T_n|^\alpha\right) \leq \left(\mathbb{E}|x - T_n|^2\right)^{\frac{\alpha}{2}}.$$

D'après le théorème de transfert, on a

$$\mathbb{E}\left(|x - T_n|^\alpha\right) = \sum_{k=0}^n \left|x - \frac{k}{n}\right|^\alpha \mathbb{P}\left(T_n = \frac{k}{n}\right).$$

Posons $p = \dfrac{2}{\alpha}$ et $q = \dfrac{2}{2-\alpha}$ de sorte que

$$\frac{1}{p} + \frac{1}{q} = \frac{\alpha}{2} + \frac{2-\alpha}{2} = 1.$$

Pour tout $k \in [\![0, n]\!]$, on pose

$$u_k = \left|x - \frac{k}{n}\right|^\alpha \mathbb{P}\left(T_n = \frac{k}{n}\right)^{\frac{\alpha}{2}} \geq 0$$

et

$$v_k = \mathbb{P}\left(T_n = \frac{k}{n}\right)^{\frac{2-\alpha}{2}} > 0.$$

Alors,

$$\begin{aligned}
\mathbb{E}\left(|x-T_n|^\alpha\right) &= \sum_{k=0}^{n} u_k v_k \\
&= \sum_{k=0}^{n} |u_k v_k| \\
&\leq \left(\sum_{k=0}^{n} |u_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=0}^{n} |v_k|^q\right)^{\frac{1}{q}}, \text{ d'après l'inégalité de Hölder,} \\
&= \left(\sum_{k=0}^{n} \left|x-\frac{k}{n}\right|^2 \mathbb{P}\left(T_n=\frac{k}{n}\right)\right)^{\frac{\alpha}{2}} \left(\sum_{k=0}^{n} \mathbb{P}\left(T_n=\frac{k}{n}\right)\right)^{\frac{2-\alpha}{2}} \\
&= \left(\sum_{k=0}^{n} \left|x-\frac{k}{n}\right|^2 \mathbb{P}\left(T_n=\frac{k}{n}\right)\right)^{\frac{\alpha}{2}}, \text{ car } \sum_{k=0}^{n} \mathbb{P}\left(T_n=\frac{k}{n}\right)=1, \\
&= \left(\mathbb{E}(|x-T_n|^2)\right)^{\frac{\alpha}{2}}, \text{ d'après le théorème de transfert.}
\end{aligned}$$

□

18. En déduire que si $g \in \mathrm{Lip}_\alpha(L)$ alors $\|g-B_n(g)\|_\infty \leq L\left(\frac{1}{4n}\right)^{\frac{\alpha}{2}}$.

Soit $g \in \mathrm{Lip}_\alpha(L)$. Alors, pour tout $x \in [0,1]$, il suit de la question 15 et de la question 17 que

$$|g(x)-B_n(g)(x)| \leq L\mathbb{E}(|x-T_n|^\alpha) \leq L\mathbb{E}(|x-T_n|^2)^{\frac{\alpha}{2}} = L\mathbb{V}(T_n)^{\frac{\alpha}{2}}.$$

Or,

$$\mathbb{V}(T_n) = \mathbb{V}\left(\frac{S_n}{n}\right) = \frac{\mathbb{V}(S_n)}{n^2} = \frac{x(1-x)}{n} \leq \frac{1}{4n}$$

car, comme cela a déjà été observé, la fonction $x \mapsto x(1-x)$ est majorée par $\frac{1}{4}$.

Ainsi, pour tout $x \in [0,1]$,

$$|g(x)-B_n(g)(x)| \leq L\left(\frac{1}{4n}\right)^{\frac{\alpha}{2}}.$$

□

19. Notons

$$e_n(L,\alpha) = \sup_{f \in \mathrm{Lip}_\alpha(L)} \|f-B_n(f)\|_\infty.$$

Montrer que $e_n(L,\alpha) = O(n^{-\frac{\alpha}{2}})$ lorsque n tend vers $+\infty$.

Pour toute fonction $f \in \mathrm{Lip}_\alpha(L)$, on a montré à la question 18 que

$$\|f-B_n(f)\|_\infty \leq L\left(\frac{1}{4n}\right)^{\frac{\alpha}{2}}.$$

Ainsi,

$$e_n(L,\alpha) \leq L\left(\frac{1}{4n}\right)^{\frac{\alpha}{2}}$$

et donc

$$e_n(L,\alpha) = O(n^{-\frac{\alpha}{2}}).$$

□

Partie III : Approximation de la dérivée par les polynômes de Bernstein

Dans cette partie, après avoir souligné l'absence de lien entre convergence uniforme et régularité de la fonction limite, nous montrons que si f est une fonction de classe $\mathcal{C}^1$ sur $[0,1]$, alors la suite $(B_n(f)')_{n\in\mathbb{N}}$ converge uniformément vers f'.

20. ***Un premier contre-exemple, où l'on considère une suite de fonctions dérivables $(k_n)_{n\in\mathbb{N}^*}$ convergeant uniformément vers une fonction k.***
On considère donc la suite $(k_n)_{n\in\mathbb{N}^*}$ définie par : $\forall n \in \mathbb{N}^*$, $\forall x \in \mathbb{R}$, $k_n(x) = x\arctan(nx)$.

20.a. Montrer que $(k_n)_{n\in\mathbb{N}^*}$ converge simplement vers une fonction k à déterminer.

Si $x = 0$, on a $k_n(x) = 0$, pour tout $n \in \mathbb{N}^*$. Ainsi, $k_n(0) \xrightarrow[n\to+\infty]{} 0$.
Si $x > 0$, $nx \xrightarrow[n\to+\infty]{} +\infty$ donc $\arctan(nx) \xrightarrow[n\to+\infty]{} \frac{\pi}{2}$ et $k_n(x) \xrightarrow[n\to+\infty]{} \frac{\pi}{2}x$.
Si $x < 0$, $nx \xrightarrow[n\to+\infty]{} -\infty$ donc $\arctan(nx) \xrightarrow[n\to+\infty]{} -\frac{\pi}{2}$ et $k_n(x) \xrightarrow[n\to+\infty]{} -\frac{\pi}{2}x$.
Ainsi, la suite $(k_n)_{n\in\mathbb{N}^*}$ converge simplement vers la fonction

$$k : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & \frac{\pi}{2}|x|. \end{array} \right.$$

□

20.b. Montrer que $\forall x \in \mathbb{R}^*$,

$$|k_n(x) - k(x)| \leq \left| x\arctan\left(\frac{1}{nx}\right) \right|$$

et montrer que $(k_n)_{n\in\mathbb{N}^*}$ converge uniformément vers k sur $\mathbb{R}$.

Pour tout $t \in \mathbb{R}^*$, on a les identités bien connues

$$\arctan(t) + \arctan\left(\frac{1}{t}\right) = \frac{\pi}{2}, \text{ si } t > 0,$$

$$\arctan(t) + \arctan\left(\frac{1}{t}\right) = -\frac{\pi}{2}, \text{ si } t < 0$$

(au besoin, on remontrera cette formule en étudiant la fonction

$$t \mapsto \arctan(t) + \arctan\left(\frac{1}{t}\right)$$

sur $\mathbb{R}^*$).
Ainsi, pour tout $x \in \mathbb{R}^*$, et tout $n \in \mathbb{N}^*$,

$$x\arctan(nx) + x\arctan\left(\frac{1}{nx}\right) = \frac{\pi}{2}|x|$$

et donc

$$|k_n(x) - k(x)| = \left| \frac{\pi}{2}|x| - x\arctan\left(\frac{1}{nx}\right) - \frac{\pi}{2}|x| \right| = \left| x\arctan\left(\frac{1}{nx}\right) \right|.$$

Pour tout $x \in \mathbb{R}^*$, on a $\arctan'(x) = \dfrac{1}{1+x^2} \leq 1$ et $\arctan(0) = 0$. Il suit donc de l'inégalité des accroissements finis que

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

Ainsi,

$$\left| x \arctan\left(\frac{1}{nx}\right) \right| \leq \frac{1}{n}.$$

Ainsi, puisque $k_n(0) = k(0) = 0$, pour tout $n \in \mathbb{N}^*$, on obtient

$$\|k_n - k\|_\infty \leq \frac{1}{n} \xrightarrow[n\to+\infty]{} 0,$$

la norme infinie étant ici prise sur $\mathbb{R}$. □

20.c. Conclure cette question.

La suite $(k_n)_{n\in\mathbb{N}^*}$ est une suite de fonctions dérivables sur $\mathbb{R}$, convergeant uniformément sur $\mathbb{R}$ vers la fonction k qui n'est pas dérivable en 0. Ainsi, la dérivabilité n'est pas stable par convergence uniforme.

Remarque : Il est possible d'aller plus loin et de construire une suite de fonctions dérivables sur $\mathbb{R}$, convergeant uniformément vers une fonction f dérivable nulle part. Nous renvoyons à [2, Chapitre 9, §11] pour un exemple explicite. □

21. ***Un second contre-exemple, où l'on considère une suite de fonctions dérivables*** $(q_n)_{n\in\mathbb{N}}$ ***convergeant uniformément vers une fonction dérivable*** q**.**
On considère ici la suite $(q_n)_{n\in\mathbb{N}}$ **définie par :** $\forall n \in \mathbb{N}, \forall x \in [0,1], q_n(x) = \dfrac{x}{1+n^2x^2}$**. Montrer que l'on définit ainsi une suite de fonctions pour laquelle la limite de la suite des fonctions dérivées n'est pas égale à la dérivée de la limite.**

Soit $n \in \mathbb{N}^*$. Étudions la fonction q_n. Elle est dérivable sur $[0,1]$ et, pour tout $x \in [0,1]$, on a

$$q_n'(x) = \frac{1-n^2x^2}{(1+n^2x^2)^2}.$$

On en déduit aisément que q_n atteint son maximum sur $[0,1]$ en $\dfrac{1}{n}$ où elle vaut $\frac{1}{2n}$. Ainsi,

$$\|q_n\|_\infty = \frac{1}{2n} \xrightarrow[n\to+\infty]{} 0.$$

Autrement dit, $(q_n)_{n\in\mathbb{N}}$ converge uniformément sur $[0,1]$ vers la fonction nulle dont la dérivée est bien évidemment nulle. Or on a $q_n'(0) = 1$, ce qui montre que même si la fonction limite est dérivable, la dérivée de la fonction limite n'est pas nécessairement la limite de la suite des fonctions dérivées.

Remarque : Néanmoins, sous des hypothèses plus fines, il existe des résultats sur la dérivabilité d'une limite de suite de fonctions. On pourra par exemple se référer à [1, §IV.3.2, p.124] ou plus généralement à [4, §2.2.3, p.66] pour retrouver les principaux résultats à ce sujet. □

22. Soit ϕ **la fonction définie sur** $\mathbb{R}$ **par** $\phi(x) = x^4(1-x) + x(1-x)^4$**. Démontrer que :**

$$\forall x \in \mathbb{R},\ x^4(1-x) + x(1-x)^4 \leq \frac{1}{12}.$$

(On pourra remarquer que, pour tout $h \in \mathbb{R}$**,** $\phi\left(\frac{1}{2}+h\right) = \phi\left(\frac{1}{2}-h\right)$**.)**

Soit $x \in \mathbb{R}$ et posons $t = x - \frac{1}{2}$ de sorte que

$$\phi(x) = \phi\left(t + \frac{1}{2}\right).$$

Soit ψ la fonction définie définie sur $\mathbb{R}$ par $\psi(t) = \phi\left(t + \frac{1}{2}\right)$.

On a donc

$$\psi(t) = \frac{1}{16}(-48t^4 + 8t^2 + 1), \quad \forall t \in \mathbb{R}.$$

Soit P la fonction polynomiale de degré 2 définie sur $\mathbb{R}$ par $P(u) = -48u^2 + 8u + 1$. Alors P atteint son maximum en $u = \dfrac{1}{2}$ où elle vaut $\frac{4}{3}$.

Ainsi,

$$\phi(x) = \psi(t) = \frac{1}{16}P(t^2) \leq \frac{1}{16} \cdot \frac{4}{3} = \frac{1}{12}.$$

□

23. On souhaite, dans cette question, montrer que

$$\mathbb{E}((T_n - x)^4) \leq \frac{1}{12n^3} + \frac{3(n-1)}{16n^3} \leq \frac{1}{n^2}.$$

23.a. Si $(\alpha_i)_{i\in\mathbb{N}^*}$ est une suite de réels, montrer que :

$$\left(\sum_{i=1}^{n} \alpha_i\right)^4 = \sum_{k=1}^{n} \alpha_k^4 + 6 \sum_{1 \leq i < j \leq n} \alpha_i^2 \alpha_j^2 + \sum_{(i_1,i_2,i_3,i_4)\in I} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4}$$

où I est un ensemble à déterminer.

Une récurrence immédiate sur n montre que

$$\left(\sum_{i=1}^{n} \alpha_i\right)^4 = \sum_{(i_1,i_2,i_3,i_4)\in[\![1,n]\!]^4} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4}.$$

Partitionnons $[\![1,n]\!]^4$ en trois sous-ensembles. On note K le sous-ensemble formé des quadruplets d'entiers tous égaux, J le sous-ensemble des quadruplets qui sont dans la $\mathfrak{S}_4$-orbite d'un élément de la forme (i,i,j,j), avec $i \neq j \in [\![1,n]\!]$ et I le complémentaire de $J \sqcup K$. On a donc

$$\begin{aligned}\left(\sum_{i=1}^{n} \alpha_i\right)^4 &= \sum_{(i_1,i_2,i_3,i_4)\in K} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4} \\ &\quad + \sum_{(i_1,i_2,i_3,i_4)\in J} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4} \\ &\quad + \sum_{(i_1,i_2,i_3,i_4)\in I} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4}.\end{aligned}$$

L'application $[\![1,n]\!] \longrightarrow K$ qui à k associe (k,k,k,k) est une bijection de sorte que

$$\sum_{(i_1,i_2,i_3,i_4)\in K} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4} = \sum_{k=1}^{n} \alpha_k^4.$$

D'autre part, $\mathfrak{S}_4$ agit naturellement sur J. Chaque orbite pour cette action est de cardinal 6 et les orbites sont naturellement étiquetées par les couples (i,j) tels que $1 \leq i < j \leq n$. En outre, tous les quadruplets d'une orbite étiquetée par (i,j) contribuent pour $\alpha_i^2\alpha_j^2$. On obtient donc

$$\sum_{(i_1,i_2,i_3,i_4)\in J} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\alpha_{i_4} = 6\sum_{1\le i<j\le n}\alpha_i^2\alpha_j^2.$$

Décrivons maintenant I plus explicitement. Un quadruplet de I est un quadruplet qui n'est ni formé de 4 entiers identiques, ni formé de deux paires d'entiers identiques. Il est ainsi formé ou bien de trois entiers identiques et d'un entier unique distinct, ou bien de deux entiers identiques et de deux autres entiers distincts, ou bien de quatre entiers distincts deux à deux. Nous voyons donc que I est précisément formé des quadruplets tels qu'il existe un entier n'apparaissant qu'une seule fois dans le quadruplet. En termes mathématiques, on a :

$$I = \left\{(i_1,i_2,i_3,i_4)\in [\![1,n]\!]^4 \mid \exists l\in[\![1,n]\!], \exists! j\in[\![1,4]\!], i_j = l\right\}.$$

et on peut écrire

$$\left(\sum_{i=1}^n \alpha_i\right)^4 = \sum_{k=1}^n \alpha_k^4 + 6\sum_{1\le i<j\le n}\alpha_i^2\alpha_j^2 + \sum_{l=1}^n \alpha_l \left(\sum_{\substack{(i_1,i_2,i_3)\\ \in([\![1,n]\!]\setminus\{l\})^3}} \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}\right).$$

$\square$

23.b. Montrer que :

$$n^4\mathbb{E}((T_n-x)^4) = \sum_{k=1}^n \mathbb{E}((X_k-x)^4) + 6\sum_{1\le i<j\le n}\mathbb{E}[(X_i-x)^2(X_j-x)^2].$$

On a

$$\begin{aligned}
n^4\mathbb{E}((T_n-x)^4) &= n^4\mathbb{E}\left[\frac{1}{n^4}(S_n-nx)^4\right]\\
&= \mathbb{E}\left[(S_n-nx)^4\right]\\
&= \mathbb{E}\left[\left(\sum_{k=1}^n (X_k-x)\right)^4\right]\\
&= \mathbb{E}\Bigg[\sum_{k=1}^n (X_k-x)^4 + 6\sum_{1\le i<j\le n}(X_i-x)^2(X_j-x)^2\\
&\qquad + \sum_{l=1}^n (X_l-x)\left(\sum_{\substack{(i_1,i_2,i_3)\\ \in([\![1,n]\!]\setminus\{l\})^3}} (X_{i_1}-x)(X_{i_2}-x)(X_{i_3}-x)\right)\Bigg]\\
&= \sum_{k=1}^n \mathbb{E}[(X_k-x)^4] + 6\sum_{1\le i<j\le n}\mathbb{E}[(X_i-x)^2(X_j-x)^2]\\
&\qquad + \sum_{l=1}^n \mathbb{E}\left[(X_l-x)\left(\sum_{\substack{(i_1,i_2,i_3)\\ \in([\![1,n]\!]\setminus\{l\})^3}} (X_{i_1}-x)(X_{i_2}-x)(X_{i_3}-x)\right)\right],
\end{aligned}$$

la quatrième égalité étant vraie d'après la question 23.a et la dernière égalité étant vraie par linéarité de l'espérance.

On sait que les variables aléatoires $(X_k)_{k\in\mathbb{N}}$ sont indépendantes et on sait par ailleurs que l'espérance est multiplicative pour des variables aléatoires indépendantes. Ainsi, si on pose, pour tout $l \in [\![1,n]\!]$,

$$Y_l = \sum_{(i_1,i_2,i_3)\in([\![1,n]\!]\setminus\{l\})^3} (X_{i_1}-x)(X_{i_2}-x)(X_{i_3}-x),$$

la variable Y_l est indépendante de X_l et donc de $(X_l - x)$. Ainsi,

$$\begin{aligned} n^4\mathbb{E}((T_n-x)^4) &= \sum_{k=1}^n \mathbb{E}[(X_k-x)^4] + 6\sum_{1\le i<j\le n} \mathbb{E}[(X_i-x)^2(X_j-x)^2] \\ &\quad + \sum_{l=1}^n \mathbb{E}\left[(X_l-x)Y_l\right] \\ &= \sum_{k=1}^n \mathbb{E}[(X_k-x)^4] + 6\sum_{1\le i<j\le n} \mathbb{E}[(X_i-x)^2(X_j-x)^2] \\ &\quad + \sum_{l=1}^n \mathbb{E}(X_l-x)\mathbb{E}(Y_l). \end{aligned}$$

Or, pour tout $l \in [\![1,n]\!]$, on a $\mathbb{E}(X_l) = x$ de sorte que $\mathbb{E}(X_l - x) = 0$ et donc

$$n^4\mathbb{E}((T_n-x)^4) = \sum_{k=1}^n \mathbb{E}[(X_k-x)^4] + 6\sum_{1\le i<j\le n} \mathbb{E}[(X_i-x)^2(X_j-x)^2].$$

□

23.c. Montrer que $\mathbb{E}[(X_k-x)^4] \le \dfrac{1}{12}$ et $\mathbb{E}[(X_i-x)^2(X_j-x)^2] \le \dfrac{1}{16}$, pour tout $(i,j,k) \in [\![0,n]\!]^3$, $i \neq j$.

D'après le théorème de transfert, on a

$$\begin{aligned} \mathbb{E}[(X_k-x)^4] &= (1-x)^4\mathbb{P}\left(X_k=1\right) + (-x)^4\mathbb{P}\left(X_k=0\right) \\ &= x(1-x)^4 + x^4(1-x) \\ &\le \frac{1}{12},\ \text{d'après la question 22.} \end{aligned}$$

D'autre part, puisque X_i et X_j sont indépendantes si $i \neq j$, on a

$$\begin{aligned} \mathbb{E}[(X_i-x)^2(X_j-x)^2] &= \mathbb{E}[(X_i-x)^2]\mathbb{E}[(X_j-x)^2] \\ &= \mathbb{V}(X_i)\mathbb{V}(X_j) \\ &= (x(1-x))^2 \\ &\le \left(\frac{1}{4}\right)^2 = \frac{1}{16}, \end{aligned}$$

la dernière inégalité venant du fait, déjà mentionné plusieurs fois, que la fonction $x \mapsto x(1-x)$ est majorée par $\frac{1}{4}$.
□

23.d. Conclure.

On a donc, d'après la question 23.b,

$$\begin{aligned}
\mathbb{E}((T_n - x)^4) &= \frac{1}{n^4}\left(\sum_{k=1}^{n} \mathbb{E}[(X_k - x)^4] + 6\sum_{1\le i<j\le n} \mathbb{E}[(X_i - x)^2(X_j - x)^2]\right) \\
&\le \frac{1}{n^4}\left(\frac{n}{12} + 6\frac{n(n-1)}{2}\cdot\frac{1}{16}\right), \text{ d'après la question 23.c,} \\
&= \frac{1}{12n^3} + \frac{3(n-1)}{16n^3} \\
&= \frac{9n-5}{48n^3} \\
&< \frac{9}{48}\cdot\frac{1}{n^2} \\
&< \frac{1}{n^2}.
\end{aligned}$$

En particulier,

$$\mathbb{E}((T_n - x)^4) \le \frac{1}{12n^3} + \frac{3(n-1)}{16n^3} \le \frac{1}{n^2},$$

comme annoncé au début de la question 23. □

24.

24.a. Si $n \in \mathbb{N}^*$, exprimer $B'_{k,n}$ en fonction de $B_{k-1,n-1}$ et $B_{k,n-1}$ pour tout $k \in [\![0, n]\!]$.

Supposons d'abord $k \in [\![1, n-1]\!]$. Alors

$$\begin{aligned}
B'_{k,n} & \left[\binom{n}{k} X^k (1-X)^{n-k}\right]' \\
&= \binom{n}{k} k X^{k-1}(1-X)^{n-k} - \binom{n}{k}(n-k)X^k(1-X)^{n-1-k} \\
&= \frac{n!}{(k-1)!(n-k)!}X^{k-1}(1-X)^{n-k} - \frac{n!}{k!(n-1-k)!}X^k(1-X)^{n-1-k} \\
&= n\left[\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!}X^{k-1}(1-X)^{(n-1)-(k-1)}\right. \\
&\quad \left. -\frac{(n-1)!}{k!((n-1)-k)!}X^k(1-X)^{(n-1)-k}\right] \\
&= n\left(B_{k-1,n-1} - B_{k,n-1}\right).
\end{aligned}$$

D'autre part, on a $B_{0,n}(X) = (1-X)^n$ et donc $B'_{0,n}(X) = -n(1-X)^{n-1} = -nB_{0,n-1}$. Puisque les conventions de notations adoptées font que $B_{-1,n-1} = 0$, on peut écrire

$$B'_{0,n} = n\left(B_{-1,n-1} - B_{0,n-1}\right).$$

De même, $B_{n,n} = X^n$ de sorte que $B'_{n,n}(X) = nX^{n-1} = nB_{n-1,n-1}$. Mais puisque $B_{n,n-1} = 0$, on a

$$B'_{n,n} = n\left(B_{n-1,n-1} - B_{n,n-1}\right).$$

Ainsi, pour tout $k \in [\![0, n]\!]$, on a

$$B'_{k,n}(X) = n\left(B_{k-1,n-1} - B_{k,n-1}\right).$$

□

24.b. Montrer que $B_n(f)$ est dérivable sur $[0,1]$ et que

$$B_n(f)' = n\sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right) B_{k,n-1}.$$

La fonction $B_n(f)$ est polynomiale donc dérivable. En outre,

$$\begin{aligned}
B_n(f)' &= \left[\sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{k,n}\right]' \\
&= \sum_{k=0}^{n} f\left(\frac{k}{n}\right) n\left(B_{k-1,n-1} - B_{k,n-1}\right), \text{ d'après la question 24.a,} \\
&= n\left[\sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{k-1,n-1} - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) B_{k,n-1}\right] \\
&= n\left[\sum_{k=1}^{n} f\left(\frac{k}{n}\right) B_{k-1,n-1} - \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) B_{k,n-1}\right], \text{ car } B_{-1,n-1} = B_{n,n-1} = 0, \\
&= n\left[\sum_{k=0}^{n-1} f\left(\frac{k+1}{n}\right) B_{k,n-1} - \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) B_{k,n-1}\right] \\
&= n\sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right) B_{k,n-1}.
\end{aligned}$$

□

25. Dans la suite de cette partie nous supposerons que f est dérivable en un point $x_0 \in [0,1]$.

25.a. Dans le cas où $x_0 \in \{0,1\}$, montrer que $(B_n(f)'(x_0))_{n\in\mathbb{N}^*}$ converge vers $f'(x_0)$.

Pour tout $n \in \mathbb{N}^*$, on a montré en 24.b que

$$B_n(f)' = n\sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right) B_{k,n-1}.$$

Or pour tout $k \in [\![0, n-1]\!]$, on a

$$B_{k,n-1} = \binom{n-1}{k} X^k (1-X)^{n-1-k}$$

de sorte que

$$B_{k,n-1}(0) = \delta_{k,0}$$

et

$$B_{k,n-1}(1) = \delta_{k,n-1}$$

où δ est le symbole de Kronecker [8].

8. Leopold Kronecker, mathématicien allemand, né à Liegnitz en 1823 et décédé à Berlin en 1891. Ses travaux portèrent principalement sur la théorie algébrique des nombres. Il fut l'un des principaux opposants à la théorie des ensembles de Cantor.

Ainsi,

$$\begin{aligned}B_n(f)'(0) &= n\sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right)-f\left(\frac{k}{n}\right)\right)B_{k,n-1}(0)\\ &= n\left(f\left(\frac{1}{n}\right)-f(0)\right)\\ &= \frac{f\left(\frac{1}{n}\right)-f(0)}{\frac{1}{n}} \xrightarrow[n\to+\infty]{} f'(0)\end{aligned}$$

et

$$\begin{aligned}B_n(f)'(1) &= n\sum_{k=0}^{n-1}\left(f\left(\frac{k+1}{n}\right)-f\left(\frac{k}{n}\right)\right)B_{k,n-1}(1)\\ &= n\left(f(1)-f\left(\frac{n-1}{n}\right)\right)\\ &= \frac{f(1)-f\left(1-\frac{1}{n}\right)}{\frac{1}{n}} \xrightarrow[n\to+\infty]{} f'(1).\end{aligned}$$

□

25.b. Dans le reste de cette question, nous considérons que $x_0 \in]0,1[$.
25.b.i. Soit g la fonction définie sur $[0,1]$ par

$$\forall x \in [0,1],\ g(x) = f(x) - f(x_0) - (x-x_0)f'(x_0).$$

On définit alors pour tout $y \in [0,1]$, $g_1(y) = yg(y)$ et $g_2(y) = (1-y)g(y)$. Montrer que

$$\forall x \in]0,1[,\ B_n(g)'(x) = \frac{n}{x(1-x)}\left((1-x)B_n(g_1)(x) - xB_n(g_2)(x)\right).$$

On a

$$\begin{aligned}B_n(g_1)(x) &= \sum_{k=0}^{n}\frac{k}{n}g\left(\frac{k}{n}\right)B_{k,n}\\ &= \sum_{k=1}^{n}\frac{k}{n}g\left(\frac{k}{n}\right)B_{k,n}\\ &= \sum_{k=0}^{n-1}\frac{k+1}{n}g\left(\frac{k+1}{n}\right)B_{k+1,n}\\ &= \sum_{k=0}^{n-1}\frac{k+1}{n}g\left(\frac{k+1}{n}\right)\binom{n}{k+1}x^{k+1}(1-x)^{n-1-k}.\end{aligned}$$

Ainsi,

$$\frac{1}{x}B_n(g_1)(x) = \sum_{k=0}^{n-1}\frac{k+1}{n}g\left(\frac{k+1}{n}\right)\binom{n}{k+1}x^{k}(1-x)^{n-1-k}.$$

D'autre part,

$$\begin{aligned}
B_n(g_2)(x) &= \sum_{k=0}^{n}\left(1-\frac{k}{n}\right) g\left(\frac{k}{n}\right) B_{k,n} \\
&= \sum_{k=0}^{n-1}\left(\frac{n-k}{n}\right) g\left(\frac{k}{n}\right) B_{k,n} \\
&= \sum_{k=0}^{n-1}\left(\frac{n-k}{n}\right) g\left(\frac{k}{n}\right) \binom{n}{k} x^k(1-x)^{n-k}.
\end{aligned}$$

Ainsi,

$$\frac{1}{(1-x)} B_n(g_2)(x) = \sum_{k=0}^{n-1}\left(\frac{n-k}{n}\right) g\left(\frac{k}{n}\right) \binom{n}{k} x^k(1-x)^{n-1-k}.$$

Si on pose

$$(\star) = \frac{n}{x(1-x)}\left((1-x)B_n(g_1)(x) - xB_n(g_2)(x)\right),$$

on obtient

$$\begin{aligned}
(\star) &= \frac{n}{x} B_n(g_1)(x) + \frac{n}{(1-x)} B_n(g_2)(x) \\
&= n\sum_{k=0}^{n-1}\left[\frac{k+1}{n} g\left(\frac{k+1}{n}\right)\binom{n}{k+1} - \left(\frac{n-k}{n}\right) g\left(\frac{k}{n}\right)\binom{n}{k}\right] x^k(1-x)^{n-1-k}.
\end{aligned}$$

Or, pour tout $k \in [\![0, n-1]\!]$, si on pose

$$(\dagger_k) = \frac{k+1}{n} g\left(\frac{k+1}{n}\right)\binom{n}{k+1} - \left(\frac{n-k}{n}\right) g\left(\frac{k}{n}\right)\binom{n}{k},$$

on a

$$\begin{aligned}
(\dagger_k) &= \frac{k+1}{n}\cdot\frac{n!}{(k+1)!(n-(k+1))!} g\left(\frac{k+1}{n}\right) + \frac{n-k}{n}\cdot\frac{n!}{k!(n-k)!} g\left(\frac{k}{n}\right) \\
&= \frac{(n-1)!}{k!(n-1-k)!} g\left(\frac{k+1}{n}\right) - \frac{(n-1)!}{k!(n-1-k)!} g\left(\frac{k}{n}\right) \\
&= \binom{n-1}{k}\left(g\left(\frac{k+1}{n}\right) - g\left(\frac{k}{n}\right)\right).
\end{aligned}$$

Ainsi,

$$\begin{aligned}
(\star) &= n\sum_{k=0}^{n}\left(g\left(\frac{k+1}{n}\right) - g\left(\frac{k}{n}\right)\right)\binom{n-1}{k} x^k(1-x)^{n-1-k} \\
&= n\sum_{k=0}^{n}\left(g\left(\frac{k+1}{n}\right) - g\left(\frac{k}{n}\right)\right) B_{k,n-1}(x) \\
&= B_n(g)', \text{ d'après la question 24.b.}
\end{aligned}$$

□

25.b.ii. En déduire que

$$B_n(g)' = \frac{\mathbb{E}((T_n - x))g(T_n))}{\mathbb{V}(T_n)}.$$

On rappelle que

$$\mathbb{V}(T_n) = \mathbb{V}\left(\frac{S_n}{n}\right) = \frac{\mathbb{V}(S_n)}{n^2} = \frac{x(1-x)}{n}$$

de sorte que l'on peut récrire la formule trouvée à la question 25.b.i

$$B_n(g)' = \frac{1}{\mathbb{V}(T_n)}\left((1-x)B_n(g_1)(x) - xB_n(g_2)(x)\right).$$

D'autre part,

$$\begin{aligned}(1-x)B_n(g_1)(x) - xB_n(g_2)(x) &= B_n(g_1)(x) - xB_n(g_1)(x) - xB_n(g_2)(x)\\ &= \sum_{k=0}^{n}\frac{k}{n}g\left(\frac{k}{n}\right)B_{k,n}(x) - \sum_{k=0^n} x\frac{k}{n}g\left(\frac{k}{n}\right)B_{k,n}(x)\\ &\quad - \sum_{k=0}^{n} x\left(1-\frac{k}{n}\right)g\left(\frac{k}{n}\right)B_{k,n}(x)\\ &= \sum_{k=0}^{n}\left(\frac{k}{n}-x\right)g\left(\frac{k}{n}\right)B_{k,n}(x)\\ &= \sum_{k=0}^{n}\left(\frac{k}{n}-x\right)g\left(\frac{k}{n}\right)\mathbb{P}\left(T_n=\frac{k}{n}\right)\\ &= \mathbb{E}((T_n-x)g(T_n)),\end{aligned}$$

la dernière égalité découlant du théorème de transfert. □

25.b.iii. Montrer que $(B_n(g)'(x_0))_{n\in\mathbb{N}^*}$ converge vers 0.

Nous n'avons pas encore utilisé l'hypothèse de dérivabilité de f, peut-être est-il temps de le faire. Posons

$$\tau(x) = \frac{g(x)}{x-x_0}, \quad \forall x \in [0,1]\setminus\{x_0\}.$$

L'hypothèse de dérivabilité de f en x_0 implique que

$$\tau(x) \xrightarrow[x:x_0]{} 0.$$

Fixons $\epsilon > 0$. Il existe donc $\delta > 0$ tel que, pour tout $x \in]0,1[$,

$$|x-x_0| < \delta \Rightarrow |\tau(x)| \leq \frac{\epsilon}{2}.$$

Il suit de la question 25.b.ii que

$$|B_n(g)'(x_0)| = |\mathbb{E}((T_n-x_0)g(T_n))|.$$

Or

$$\begin{aligned}|\mathbb{E}((T_n-x_0)g(T_n))| &\leq \mathbb{E}|(T_n-x_0)g(T_n)|\\ &= \mathbb{E}|(T_n-x_0)^2\tau(T_n)|\\ &\leq \mathbb{E}|(T_n-x_0)^2\tau(T_n)\mathbb{1}_{(|T_n-x_0|<\delta)}|\\ &\quad + \mathbb{E}|(T_n-x_0)^2\tau(T_n)\mathbb{1}_{(|T_n-x_0|\geq\delta)}|\\ &\leq \frac{\epsilon}{2}\mathbb{E}((T_n-x_0)^2) + \mathbb{E}|(T_n-x_0)^2\tau(T_n)\mathbb{1}_{(|T_n-x_0|\geq\delta)}|\\ &= \frac{\epsilon}{2}\mathbb{V}(T_n) + \mathbb{E}|(T_n-x_0)^2\tau(T_n)\mathbb{1}_{(|T_n-x_0|\geq\delta)}|.\end{aligned}$$

Posons

$$(\spadesuit_n) = \mathbb{E}|(T_n-x_0)^2\tau(T_n)\mathbb{1}_{(|T_n-x_0|\geq\delta)}|.$$

La fonction τ prolongée par 0 en x_0 est continue sur $[0,1]$ donc il existe $M > 0$ tel que $|\tau(y)| \leq M$, pour tout $y \in [0,1]$. Puisque $T_n(\Omega) \subset [0,1]$, pour tout $n \in \mathbb{N}^*$, il vient

$$|\tau(T_n)| \leq M, \quad \forall n \in \mathbb{N}^*.$$

Ainsi,

$$\begin{aligned}
(\spadesuit_n) &\leq M\mathbb{E}|(T_n - x_0)^2 \mathbb{1}_{(|T_n - x_0| \geq \delta)}| \\
&\leq M\sqrt{\mathbb{E}[(T_n - x_0)^4]}\sqrt{\mathbb{E}(\mathbb{1}^2_{(|T_n - x_0| \geq \delta)})}, \text{ d'après la question 16 avec } p = q = 2, \\
&= M\sqrt{\mathbb{E}[(T_n - x_0)^4]}\sqrt{\mathbb{P}\left(|T_n - x_0| \geq \delta\right)} \\
&\leq M \cdot \frac{1}{n} \cdot \sqrt{\frac{1}{4n\delta^2}}, \text{ en vertu des questions 23.d et 6.} \\
&= \frac{M}{2\delta n^{3/2}} \xrightarrow[n \to +\infty]{} 0.
\end{aligned}$$

Ainsi, il existe $n_0 \in \mathbb{N}$ tel que

$$n \geq n_0 \Rightarrow (\spadesuit_n) \leq \frac{\epsilon}{2}$$

et donc

$$|B_n(g)'(x_0)| = |\mathbb{E}((T_n - x_0)g(T_n))| \leq \epsilon.$$

Autrement dit,

$$B_n(g)'(x_0) \xrightarrow[n \to +\infty]{} 0.$$

□

25.b.iv. Montrer que $(B_n(f)'(x_0))_{n \in \mathbb{N}^*}$ converge vers $f'(x_0)$.

Pour tout $n \in \mathbb{N}^*$, on a

$$g = f - f(x_0) - f'(x_0)(\mathrm{id} - x_0).$$

L'opérateur B_n étant linéaire, on obtient

$$B_n(g) = B_n(f) - f(x_0)B_n(1) - f'(x_0)\left(B_n(\mathrm{id}) - x_0 B_n(1)\right).$$

Or il suit de la question 2.c.i que, pour tout $x \in [0, 1]$,

$$B_n(1) = \sum_{k=0}^{n} B_{k,n}(x) = 1$$

et de la question 2.c.ii que

$$B_n(\mathrm{id})(x) = \sum_{k=0}^{n} \frac{k}{n} B_{k,n}(x) = \frac{1}{n}\sum_{k=0}^{n} k B_{k,n}(x) = \frac{nx}{n} = x.$$

Ainsi, pour tout $x \in [0, 1]$, on a

$$B_n(g)(x) = B_n(f)(x) - f(x_0) - f'(x_0)\left(x - x_0\right).$$

En particulier,

$$B_n(f)'(x_0) = B_n(g)'(x_0) + f'(x_0).$$

Il suit alors de la question 25.b.iii. que

$$B_n(f)'(x_0) \xrightarrow[n \to +\infty]{} f'(x_0).$$

□

Partie IV : Courbes et estimateurs de Bézier

A. Géométrie des courbes de Bézier

Définition : Pour $n \in \mathbb{N}$***,*** $t \in [0,1]$ ***et*** $(P_k)_{0\leq k\leq n}$***,*** $(n+1)$ ***nombres complexes distincts, on définit le nombre*** $G(t)$ ***par*** $G(t) = \sum_{k=0}^{n} B_{k,n}(t)P_k$***. En assimilant tout point du plan à son affixe complexe, on appelle courbe de Bézier associée aux points de contrôles*** $(P_k)_{0\leq k\leq n}$ ***le lieu des points*** $G(t)$ ***quand*** t ***varie dans*** $[0,1]$***.***

26. Quelle est la courbe de Bézier associée à un point de contrôle ? À deux points de contrôle ?

La courbe de Bézier associée à un point de contrôle P_0 est le lieu des points

$$G(t) = B_{0,0}(t)P_0 = P_0, \quad \forall t \in [0,1],$$

il s'agit donc du seul point P_0.

La courbe de Bézier associée à deux points de contrôle P_0 et P_1 est le lieu des points

$$G(t) = B_{0,1}(t)P_0 + B_{1,1}(t)P_1 = (1-t)P_0 + tP_1, \quad \forall t \in [0,1],$$

il s'agit donc du segment $[P_0P_1]$. □

27. Vérifier que la courbe de Bézier associée aux points de contrôles $(P_k)_{0\leq k\leq n}$ **contient** P_0 **et** P_n**.**

On a

$$G(0) = \sum_{k=0}^{n} B_{k,n}(0)P_k$$

or on a déjà vu que $B_{k,n}(0) = \delta_{k,0}$, pour tout $k \in [\![0,n]\!]$. Ainsi,

$$G(0) = P_0.$$

De même,

$$G(1) = \sum_{k=0}^{n} B_{k,n}(1)P_k$$

or on a déjà vu que $B_{k,n}(1) = \delta_{k,n}$, pour tout $k \in [\![0,n]\!]$. Ainsi,

$$G(1) = P_n.$$

Il s'ensuit que P_0 et P_n sont bien sur la courbe de Bézier. □

28. Montrer que la courbe de Bézier associée aux points de contrôles $(P_k)_{0\leq k\leq n}$ **(**$n \leq 2$**), est une portion de droite si et seulement si les points** $(P_k)_{0\leq k\leq n}$ **sont alignés.**

Nous commençons par observer que la condition ($n \leq 2$) proposée est très certainement une faute de frappe et nous répondrons plutôt à la question avec la condition ($n \geq 2$).

D'après la question 2, pour tout $t \in [0,1]$, $B_{k,n}(t)$ est un réel positif et $\sum_{k=0}^{n} B_{k,n}(t) = 1$. Il s'ensuit que $G(t)$ est une combinaison barycentrique de $P_0, \ldots, P_n$. Si ces derniers sont alignés, leur barycentre l'est aussi (le lecteur sceptique pourra redémontrer cette propriété à l'aide d'une récurrence sur n en utilisant l'associativité du barycentre).

Il nous reste donc à montrer la réciproque, à savoir que si la courbe de Bézier est une portion d'une droite Δ, alors les points $P_0, \ldots, P_k$ sont alignés. Plaçons nous dans le repère affine $(P_0, \overrightarrow{u}, i\overrightarrow{u})$ où $\overrightarrow{u} = (P_n - P_0)$. On commence par remarquer que puisque $G(0) = P_0$ et $G(1) = P_n$, alors

$$\Delta = (P_0P_n) = P_0 + \mathbb{R}\overrightarrow{u}.$$

Pour tout $k \in [\![0,n]\!]$, il existe un unique couple de réels (x_k, y_k) tel que

$$P_k = P_0 + (x_k + iy_k)\overrightarrow{u}.$$

Alors, pour tout $t \in [0,1]$, on a

$$\begin{aligned} G(t) &= \sum_{k=0}^{n} B_{k,n}(t) P_k \\ &= \sum_{k=0}^{n} B_{k,n}(t) P_0 + \sum_{k=0}^{n} B_{k,n}(t) x_k \overrightarrow{u} + i \sum_{k=0}^{n} B_{k,n}(t) y_k \overrightarrow{u} \\ &= P_0 + \left(\sum_{k=0}^{n} x_k B_{k,n}(t) \right) \overrightarrow{u} + i \left(\sum_{k=0}^{n} y_k B_{k,n}(t) \right) \overrightarrow{u} \end{aligned}$$

Si, pour tout $t \in [0,1]$, on a $G(t) \in \Delta = P_0 + \mathbb{R}\overrightarrow{u}$, on a

$$\sum_{k=0}^{n} y_k B_{k,n}(t) = 0, \quad \forall t \in [0,1].$$

Ainsi, le polynôme

$$\sum_{k=0}^{n} y_k B_{k,n}$$

admet une infinité de racines, il est donc nul. Mais nous avons montré à la question 4.a que la famille $(B_k)_{0 \leq k \leq n}$ est libre. Ainsi, pour tout $k \in [\![0,n]\!]$, on a $y_k = 0$ et donc

$$P_k = P_0 + x_k \overrightarrow{u} \in P_0 + \mathbb{R}\overrightarrow{u} = \Delta.$$

Tous les points de contrôle sont donc alignés. □

29. Montrer que la courbe de Bézier associée aux points de contrôle $(P_k)_{0 \leq k \leq n}$ admet une tangente en chacun de ses points, que sa tangente en P_0 est (P_0P_1) et que sa tangente en P_n est $(P_{n-1}P_n)$.

G est dérivable en tout point de $[0,1]$ donc la courbe de Bézier admet des tangentes en tout point. En outre,

$$\begin{aligned} G'(t) &= \sum_{k=0}^{n} B'_{k,n}(t) P_k \\ &= n \sum_{k=0}^{n} P_k \left(B_{k-1,n-1}(t) - B_{k,n-1}(t) \right). \end{aligned}$$

En particulier, $G'(0) = n(P_1 - P_0) \in \mathbb{R}(P_1 - P_0)$. Or on a vu à la question précédente que $G(0) = P_0$ de sorte que la tangente à la courbe de Bézier est $P_0 + \mathbb{R}(P_1 - P_0)$, c'est-à-dire la droite (P_0P_1).

Par symétrie, la tangente à la courbe de Bézier en P_n est $P_n + \mathbb{R}(P_n - P_{n-1})$, c'est-à-dire la droite $(P_{n-1}P_n)$. □

30. Montrer que la courbe de Bézier associée aux points de contrôle $(P_k)_{0 \leq k \leq n}$ est dans l'enveloppe convexe de $\{P_k \mid 0 \leq k \leq n\}$.

Comme nous l'avons remarqué à la question précédente, pour tout $t \in [0,1]$, $G(t)$ est une combinaison barycentrique de $P_0, \ldots, P_n$. L'enveloppe convexe Γ de la famille $\{P_k \mid 0 \leq k \leq n\}$ étant l'ensemble de toutes les combinaisons barycentriques de $P_0, \ldots, P_n$, il est clair que

$$G(t) \in \Gamma, \quad \forall t \in [0,1].$$

□

31. Soient α, β deux réels tels que $\alpha < \beta$ et soit $(P_k)_{0\leq k\leq n}$ $n+1$ nombres complexes tels que

$$P_0 = \alpha, \quad P_n = \beta + i, \quad \textbf{et} \quad \forall k \in [\![0, n-1]\!], \quad \mathrm{Re}(P_k) < \mathrm{Re}(P_{k+1}).$$

On considère C ; la courbe de Bézier associée à ces nombres complexes et, pour $t \in [0,1]$, on pose

$$G_n(t) = \sum_{k=0}^{n} B_{k,n}(t) P_k.$$

Avant d'aller plus loin dans la question 31, nous dessinons un exemple qui pourra aider le lecteur à se représenter la situation.

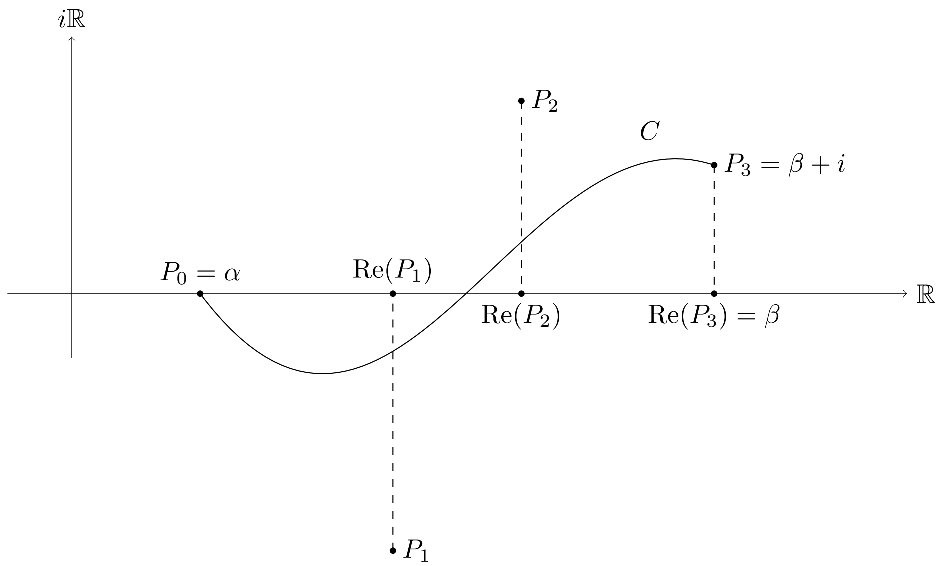


Figure 1. Un exemple avec $n = 3$

31.a. Montrer que la courbe C passe par les points d'affixe α et $\beta + i$.

C'est une conséquence directe de la question 27 puisque $\alpha = P_0 = G_n(0)$ et $\beta + i = P_n = G_n(1)$. □

31.b. Montrer que la fonction $t \mapsto \mathrm{Re}(G_n(t))$ est dérivable et strictement croissante sur $[0,1]$.

Pour tout $k \in [\![0, n]\!]$, il existe un unique couple $(x_k, y_k) \in \mathbb{R}^2$ tel que

$$P_k = x_k + i y_k.$$

Posons

$$R : \left\{ \begin{array}{ccc} [0,1] & \longrightarrow & \mathbb{R} \\ t & \mapsto & \mathrm{Re}(G_n(t)) \end{array} \right.$$

On a donc, pour tout $t \in [0,1]$,

$$R(t) = \sum_{k=0}^{n} x_k B_{k,n}(t)$$

de sorte que R est polynomiale en t, donc dérivable et,

$$\begin{aligned}
R'(t) &= \sum_{k=0}^{n} x_k B'_{k,n}(t) \\
&= n \sum_{k=0}^{n} x_k \left(B_{k-1,n-1}(t) - B_{k,n-1}(t)\right) \\
&= n \sum_{k=0}^{n} x_k B_{k-1,n-1}(t) - n \sum_{k=0}^{n} x_k B_{k,n-1}(t) \\
&= n \sum_{k=1}^{n} x_k B_{k-1,n-1}(t) - n \sum_{k=0}^{n-1} x_k B_{k,n-1}(t) \\
&= n \sum_{k=0}^{n-1} x_{k+1} B_{k,n-1}(t) - n \sum_{k=0}^{n-1} x_k B_{k,n-1}(t) \\
&= n \sum_{k=0}^{n-1} \left(x_{k+1} - x_k\right) B_{k,n-1}(t).
\end{aligned}$$

Or, pour tout $k \in [\![0, n-1]\!]$, on a

$$B_{k,n-1}(t) = \binom{n-1}{k} t^k (1-t)^{n-1-k} > 0, \quad \forall t \in]0,1[$$

et, par hypothèse,

$$x_{k+1} > x_k.$$

Ainsi, $R' > 0$ sur $]0,1[$. R étant continue sur $[0,1]$, elle y est strictement croissante. $\square$

31.c. En déduire que l'application $t \mapsto \mathrm{Re}(G_n(t))$ est un $\mathcal{C}^1$-difféomorphisme de $[0,1]$ sur $[\alpha, \beta]$.

R est une application polynomiale donc $\mathcal{C}^1$ sur $[0,1]$. Elle est strictement croissante sur $[0,1]$ et $R(0) = \alpha$, $R(1) = \beta$. Il s'ensuit que R induit une bijection de $[0,1]$ sur $[\alpha, \beta]$. En outre, $R'(t) \neq 0$ pour tout $t \in]0,1[$. Il s'ensuit que R est un $\mathcal{C}^1$-difféomorphisme de $[0,1]$ sur $[\alpha, \beta]$. $\square$

B. Estimateur de Bézier

Soient α, β deux réels tels que $\alpha < \beta$ et soit X une variable aléatoire réelle à densité continue et telle que $X(\Omega) = [\alpha, \beta]$. On appelle F_X sa fonction de répartition et on choisit $n+1$ réels $x_0, x_1, \ldots, x_n$ tels que $\alpha = x_0 < x_1 < \cdots < x_n = \beta$.
On définit alors la famille de nombres complexes $(P_k)_{0 \le k \le n}$ par :

$$\forall 0 \le k \le n,\ P_k = x_k + iF_X(x_k).$$

Enfin on considère la courbe de Bézier associée à ces points $(P_k)_{0 \le k \le n}$ en posant, pour tout $t \in [0,1]$,

$$G_n(t) = \sum_{k=0}^{n} B_{k,n}(t) P_k.$$

Avant d'attaquer cette partie, nous traçons une figure qui aidera le lecteur à se représenter la situation :

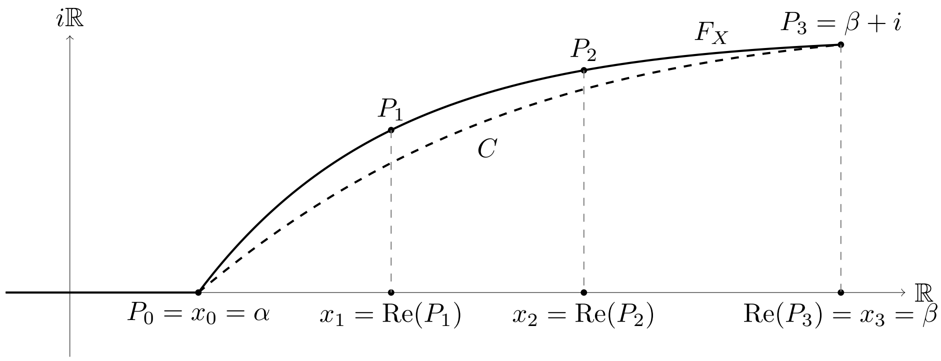


FIGURE 2. En pointillés, la courbe de Bézier C associée aux quatre points de contrôle $P_0, \ldots, P_3$ pris sur le graphe de la fonction de répartition (en trait plein).

32. Montrer que la fonction $\mathrm{Im}(G_n) \circ (\mathrm{Re}(G_n))^{-1}(t)$ est bien définie sur $[\alpha, \beta]$ et est la fonction de répartition d'une variable aléatoire à densité continue, notée X_n.

Pour tout $u \in [\alpha, \beta]$, il suit de la question 31.c que $t = \mathrm{Re}(G_n)^{-1}(u)$ est bien défini et appartient à $[0,1]$, il s'ensuit que l'on peut définir $\mathrm{Im}(G_n)(t) = \mathrm{Im}(G_n(t))$. Puisque $G_n(t)$ est dans l'enveloppe convexe de $\{x + iF_X(x_k) \mid k \in [\![0, n]\!]\}$ et que $0 \leq F_X(x_k) \leq 1$, pour tout $k \in [\![0, n]\!]$, il s'ensuit que $0 \leq \mathrm{Im}(G_n(t)) \leq 1$ et donc $\mathrm{Im}(G_n) \circ \mathrm{Re}(G_n)^{-1}(u) \in [0,1]$. L'application

$$F_n = \mathrm{Im}(G_n) \circ (\mathrm{Re}(G_n))^{-1} : [\alpha, \beta] \longrightarrow [0,1]$$

est donc bien définie.

Le diagramme (non-commutatif) ci-dessous aidera le lecteur à visualiser la situation :

$$\begin{tikzcd}
t \in [0,1] \arrow[r, shift left, "G_n"] & G_n([0,1]) \subset \mathbb{C} \arrow[l, shift left, "\mathrm{Im}"] \arrow[d, "\mathrm{Re}"] \\
& {[\alpha, \beta] \ni u} \arrow[ul, shift right, "F_X"'] \arrow[ul, shift left, "F_n"]
\end{tikzcd}$$

Il reste à montrer que F_n est une fonction de répartition.

On commence par remarquer qu'elle est continue comme composée de fonctions continues.

Ensuite, on a $G_n(0) = \alpha \in \mathbb{R}$ donc $\mathrm{Re}(G_n)^{-1}(\alpha) = 0$ et $\mathrm{Im}(G_n(0)) = 0$ de sorte que $F_n(\alpha) = 0$. D'autre part, on a $G_n(1) = \beta + i$ de sorte que $\mathrm{Re}(G_n)^{-1}(\beta) = 1$ et $\mathrm{Im}(G_n(1)) = \mathrm{Im}(\beta + i) = 1$ de sorte que $F_n(\beta) = 1$.

Enfin, elle est croissante comme composée de fonctions croissantes. En effet, on a déjà montré que $\mathrm{Re}(G_n)$ est une fonction croissante, donc $\mathrm{Re}(G_n)^{-1}$ l'est aussi. D'autre part, puisque F_X est une fonction de répartition, elle est croissante. Ainsi, si l'on pose $y_k = F_X(x_k)$, pour tout $k \in [\![0, n]\!]$, on obtient $y_{k+1} \geq y_k$, pour tout $k \in [\![0, n-1]\!]$. Un argument analogue à celui de la question 31 montre alors que $\mathrm{Im}(G_n)$ est croissante.

Ainsi, F_n est bien la fonction de répartition d'une variable aléatoire à densité continue X_n. $\square$

33. On choisit dans cette question, pour chaque entier $n \geq 1$, les $n+1$ réels $x_k = \alpha + \dfrac{k}{n}(\beta - \alpha)$. Montrer que $(X_n)_{n \geq 1}$ converge en loi vers X.

Il s'agit de montrer que, pour tout $u \in [\alpha, \beta]$, on a

$$F_{X_n}(u) \xrightarrow[n \to +\infty]{} F_X(u).$$

Pour tout $t \in [0,1]$, on a

$$\begin{aligned}\operatorname{Re}(G_n)(t) &= \sum_{k=0}^{n} x_k B_{k,n}(t)\\ &= \alpha + (\beta-\alpha)\frac{1}{n}\sum_{k=0}^{n} kB_{k,n}(t)\\ &= \alpha + t(\beta-\alpha), \text{ d'après la question 2.c.ii,}\\ &= (1-t)\alpha + t\beta.\end{aligned}$$

et

$$\begin{aligned}\operatorname{Im}(G_n)(t) &= \sum_{k=0}^{n} F_X(x_k)B_{k,n}(t)\\ &= \sum_{k=0}^{n} F_X\left(\alpha + \frac{k}{n}(\beta-\alpha)\right)B_{k,n}(t).\end{aligned}$$

Posons $G_X = F_X \circ \operatorname{Re}(G_n)$, qui est continue sur $[0,1]$. Alors il suit de la question 8 que, pour tout $t \in [0,1]$,

$$\operatorname{Im}(G_n)(t) = \sum_{k=0}^{n} G_X\left(\frac{k}{n}\right)B_{k,n}(t) \xrightarrow[n\to+\infty]{} G_X(t).$$

Soit $u \in [\alpha,\beta]$. Alors il existe un unique t tel que $u = \operatorname{Re}(G_n)(t)$ et

$$F_{X_n}(u) = F_{X_n}(\operatorname{Re}(G_n)(t)) = \operatorname{Im}(G_n)(t) \xrightarrow[n\to+\infty]{} G_X(t) = F_X(\operatorname{Re}(G_n)(t)) = F_X(u).$$

Ainsi F_{X_n} converge simplement vers F_X sur $[\alpha,\beta]$, c'est-à-dire, $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X. □

34. On revient au cas général. Montrer que

$$\mathbb{E}(X_n) = \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}x_k\sum_{i=0}^{n-1}\frac{\binom{n-1}{i}}{\binom{2n-1}{i+k}}\left(F_X(x_{i+1}) - F_X(x_i)\right).$$

La première chose à faire est de garder son calme face à cette formule assez peu intuitive et de s'y prendre méthodiquement.

Par définition, on a

$$\mathbb{E}(X_n) = \int_{-\infty}^{+\infty} xf_{X_n}(x)\,dx = \int_{\alpha}^{\beta} xf_{X_n}(x)\,dx,$$

car $X_n(\Omega) \subset [\alpha,\beta]$.

Pour alléger un peu les notations, posons $R = \operatorname{Re}(G_n)$ dont on sait qu'il s'agit d'un $\mathcal{C}^1$-difféomorphisme de $[0,1]$ vers $[\alpha,\beta]$ (voir question 31.c). On peut donc effectuer le changement de variable $x = R(u)$ dans le calcul de l'intégrale. Avec les notations différentielles usuelles, on a donc $dx = R'(u)du$.

Ainsi,

$$\mathbb{E}(X_n) = \int_0^1 R(u)f_{X_n}(R(u))R'(u)\,du.$$

L'astuce consiste maintenant à se souvenir que la densité de X_n est continue, ce qui signifie que l'on a la relation fondamentale $F'_{X_n} = f_{X_n}$. Ainsi,

$$(f_{X_n}\circ R)\cdot R' = (F'_{X_n}\circ R)\cdot R' = (F_{X_n}\circ R)'.$$

On obtient donc

$$\mathbb{E}(X_n) = \int_0^1 R(u)(F_{X_n}\circ R)'(u)\,du.$$

Or
$$F_{X_n} = \mathrm{Im}(G_n) \circ \mathrm{Re}(G_n)^{-1} = \mathrm{Im}(G_n) \circ R^{-1}$$
donc
$$F_{X_n} \circ R = \mathrm{Im}(G_n),$$
d'où l'expression grandement simplifiée
$$\mathbb{E}(X_n) = \int_0^1 \mathrm{Re}(G_n)(u) \cdot \mathrm{Im}(G_n)'(u)\ du.$$

Entrons maintenant dans le calcul à proprement parler. On a
$$\mathrm{Re}(G_n)(u) = \sum_{k=0}^{n} x_k B_{k,n}(u)$$
et
$$\begin{aligned}\mathrm{Im}(G_n)(u) &= \sum_{k=0}^{n} F_X(x_k) B_{k,n}(u)\\ &= \sum_{k=1}^{n} F_X(x_k) B_{k,n}(u),\ \text{ car } F_X(x_0) = F_X(\alpha) = 0.\end{aligned}$$
Ainsi,
$$\begin{aligned}\mathrm{Im}(G_n)'(u) &= \sum_{k=1}^{n} F_X(x_k) B'_{k,n}(u)\\ &= n\sum_{k=1}^{n} F_X(x_k)(B_{k-1,n-1}(u) - B_{k,n-1}(u))\\ &= n\sum_{i=0}^{n-1} \left(F_X(x_{i+1}) - F_X(x_i)\right) B_{i,n-1}.\end{aligned}$$

Il s'ensuit que
$$\begin{aligned}\mathbb{E}(X_n) &= \int_0^1 \sum_{k=0}^{n} x_k B_{k,n}(u)\ n\sum_{i=0}^{n-1} \left(F_X(x_{i+1}) - F_X(x_i)\right) B_{i,n-1}(u)\ du\\ &= \int_0^1 \sum_{k=0}^{n} x_k \binom{n}{k} u^k (1-u)^{n-k}\ n\\ &\quad\times \sum_{i=0}^{n-1} \left(F_X(x_{i+1}) - F_X(x_i)\right) \binom{n-1}{i} u^i (1-u)^{n-1-i}\ du\\ &= n\sum_{k=0}^{n} \binom{n}{k} x_k \sum_{i=0}^{n-1} \binom{n-1}{i} \left(F_X(x_{i+1}) - F_X(x_i)\right)\\ &\quad\times \int_0^1 u^{k+i}(1-u)^{2n-1-(k+i)}\ du.\end{aligned}$$

Or c'est un résultat classique que, pour tous $k, m \in \mathbb{N}$, on a
$$I_{k,m} = \int_0^1 u^k (1-u)^m\ du = \frac{1}{m+k+1} \frac{1}{\binom{m+k}{k}}.$$
(ceci se montre par récurrence en observant que
$$\forall k \in \mathbb{N}^*,\ \forall m \in \mathbb{N},\ I_{k,m} = \frac{k}{m+1} I_{k-1,m+1}$$
puis en calculant $I_{0,k+m}$.)

Il s'ensuit que

$$\int_0^1 u^{k+i}(1-u)^{2n-1-(k+i)}\ du = \frac{1}{2n} \cdot \frac{1}{\binom{2n-1}{k+i}}$$

et donc

$$\begin{aligned}\mathbb{E}(X_n) &= n \sum_{k=0}^{n} \binom{n}{k} x_k \sum_{i=0}^{n-1} \binom{n-1}{i} \left(F_X(x_{i+1}) - F_X(x_i)\right) \cdot \frac{1}{2n} \cdot \frac{1}{\binom{2n-1}{k+i}} \\ &= \frac{1}{2} \sum_{k=0}^{n} \binom{n}{k} x_k \sum_{i=0}^{n-1} \frac{\binom{n-1}{i}}{\binom{2n-1}{k+i}} \left(F_X(x_{i+1}) - F_X(x_i)\right).\end{aligned}$$

□

35. Y a-t-il convergence en moyenne quadratique ?

On rappelle qu'il y a convergence en moyenne quadratique si $\mathbb{E}((X_n - X)^2) \xrightarrow[n\to+\infty]{} 0$. Puisque nous venons de trouver une formule explicite pour $\mathbb{E}(X_n)$, le chemin semble tout tracé pour nous orienter vers un calcul direct. Cependant, étant donné la non-compacité de ladite formule, nous allons essayer de ruser et de conclure d'une autre manière.

On commence par rappeler que la convergence en moyenne quadratique implique la convergence en loi mais que la réciproque est fausse. On ne peut donc pas conclure à l'aide de la question 33, laquelle n'était de toute façon valable que dans le cas d'une subdivision régulière. C'est d'ailleurs cette latitude pour le positionnement des points $\alpha = x_0 < x_1 < \cdots < x_{n-1} < x_n = \beta$ que nous allons utiliser pour construire un exemple où l'on n'a pas convergence en moyenne quadratique.

Mais commençons par montrer un cas où il y a effectivement convergence en moyenne quadratique. Considérons une variable aléatoire X de loi uniforme sur $[0,1]$. La restriction de la fonction de répartition au support de X est donc la fonction linéaire $x \mapsto x$. Si nous considérons la subdivision régulière $x_0, \ldots, x_n$ telle que $x_k = \frac{k}{n}$, pour tout $k \in [\![0,n]\!]$, on obtient $G_n(t) = t$, pour tout $t \in [0,1]$ et donc $F_{X_n}(x) = x$, pour tout $x \in [0,1]$. Il s'ensuit que chaque X_n suit aussi une loi uniforme sur $[0,1]$ et que l'on a donc trivialement convergence en moyenne quadratique.

Construisons maintenant un exemple où l'on n'a pas convergence en moyenne quadratique. Intuitivement, la stratégie est la suivante : Nous avons toute latitude sur le positionnement des points $x_0, \ldots, x_n$ tant que $x_0 = \alpha$ et $x_n = \beta$. Nous allons donc considérer une variable aléatoire qui prend ses valeurs avec une bien plus forte probabilité dans au voisinage de β qu'au voisinage de α puis nous allons concentrer $x_0, \ldots, x_{n-1}$ dans le voisinage de α, ne laissant que x_n dans le voisinage de β. Il est alors intuitif que la courbe de Bézier correspondante ne donnera jamais une bonne approximation de la fonction de répartition.

Considérons par exemple la variable aléatoire $X : \Omega \longrightarrow [0,3]$ de fonction de répartition F_X définie par

$$F_X(x) = \mathbb{1}_{[2,3]}(x) \cdot (x-2).$$

Pour tout $n \in \mathbb{N}^*$, on considère la suite de points définis par

$$x_k = \frac{k}{n},\ \text{si } 0 \leq k \leq n-1$$

et $x_n = 3$.

Alors

$$F_X(x_k) = \begin{cases} 0 & \text{si } 0 \leq k \leq n-1, \\ 1 & \text{si } k = n. \end{cases}$$

En particulier, les points $P_0, \ldots, P_n$ sont tous dans le triangle T fermé formé des points d'affixes 0, 1 et $3+i$. Il suit donc de la question 30 que $G_n(t) \in T$, pour tout $t \in [0,1]$. Or, pour tout $x \in]2,3]$, $F_X(x) \notin T$ de sorte que $F_{X_n}(x)$ ne peut tendre vers $F_X(x)$ quand n tend vers $+\infty$. On n'a donc pas convergence en loi de X_n vers X et donc *a fortiori* pas non plus convergence en moyenne quadratique. La figure 3 ci-dessous aidera le lecteur à se représenter la situation.

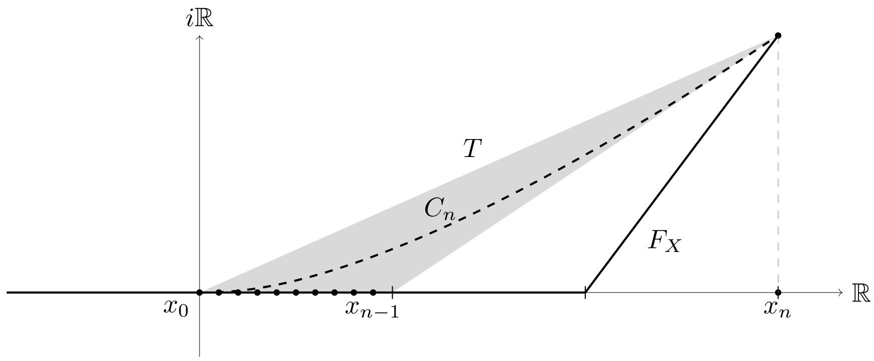


FIGURE 3. En pointillés, la courbe de Bézier C_n associée aux points de contrôle $P_0, \ldots, P_n$. En grisé, le triangle fermé où se situent tous les $G_n(t)$, avec $t \in [0,1]$, $n \in \mathbb{N}^*$. En gras, le graphe de la fonction de répartition de X.

En conclusion, on ne peut pas conclure... □

FIN DE L'ÉPREUVE

Références

[1] Xavier Gourdon. *Analyse*. Les maths en tête. Ellipses, 1994.

[2] Bertrand Hauchecorne. *Les contre-exemples en mathématiques*. Ellipses, 1988.

[3] Jean-Yves Ouvrard. *Probabilités 1*, volume 1. Licence - CAPES. Cassini, 2007.

[4] Edmond Ramis, Claude Deschamps, and Jacques Odoux. *Cours de mathématiques spéciales*, volume 4. Séries et équations différentielles. Masson, 3ème edition, 1993.