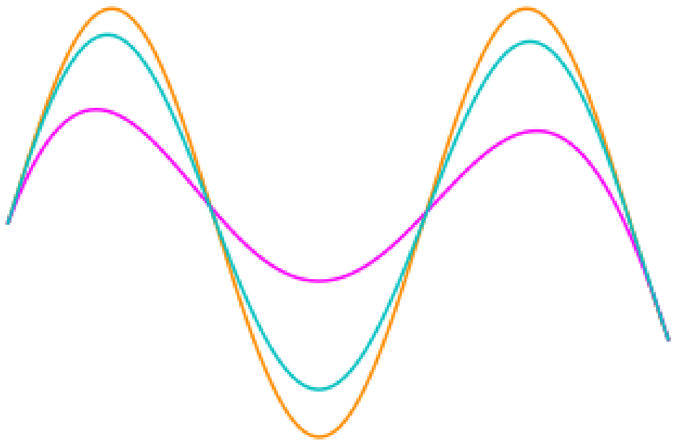
Autour des polynômes de Bernstein

Travail encadré de recherche

Sous la direction de Pr. Olivier Goubet

❀

Université de Lille



Céline Wang

6 avril 2023

Résumé

En 1885, **Weierstrass** (1815 – 1897) publie son célèbre Théorème selon lequel toute fonction continue sur un segment est limite uniforme d'une suite de polynômes. Il faudra attendre 1912 pour que **Sergueï Natanovitch Bernstein** (1880 – 1968) introduise des polynômes portant son nom pour apporter une preuve constructive au Théorème. Ces derniers paraissent pour la première fois dans une petite note de deux pages aux *Communications de la Société Mathématique de Kharkov*, volume 13, sous le titre de « Démonstration du théorème de Weierstrass fondé sur le calcul des probabilités ».

Ces polynômes ne sont pas seulement connus comme étant une preuve du Théorème de Weierstrass, ils sont également notables pour diverses propriétés.

Dans ce mémoire, on se propose de présenter deux démonstrations de ce Théorème et d'étudier quelques propriétés des opérateurs $f \longmapsto B_n(f)$, pour $n \in \mathbb{N}$, notamment sur les fonctions convexes en dimension 1 et 2.

Mots-clés : polynôme, Bernstein, opérateur, probabilités, approximation, Weierstrass, convexité, modèle de Wright-Fisher

Table des matières

Introduction

Dans ce chapitre, nous introduisons les polynômes de Bernstein, constituant une approche constructive de l'approximation des fonctions continues sur un segment. Nous présenterons ici les premières définitions, et quelques propriétés remarquables seront établies.

0.1 Définition

Définition 0.1 (Polynômes de Bernstein) : *Soit f une fonction définie et continue sur $[0,1]$. On appelle **polynômes de Bernstein** associés à f les polynômes de la forme*

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

où $n \in \mathbb{N}^*$.
La suite $(B_n)_{n\in\mathbb{N}^*}$ définie par

$$B_n : \begin{array}{ccc} (\mathcal{C}([0,1],\mathbb{R}), \|\cdot\|_\infty) & \longrightarrow & (\mathcal{C}([0,1],\mathbb{R}), \|\cdot\|_\infty) \\ f & \longmapsto & B_n(f) \end{array}$$

est une suite d'opérateurs linéaires.

✍ **Notation 1.** Afin d'alléger l'écriture, on introduit également, pour $n \in \mathbb{N}^*$, $k \in [\![0,n]\!]$ et $x \in [0,1]$, la quantité

$$b_k^n(x) = \binom{n}{k} x^k (1-x)^{n-k}.$$

Donc la Définition 0.1 se réécrit

$$B_n(f)(x) = \sum_{k=0}^{n} b_k^n(x) f\left(\frac{k}{n}\right).$$

0.2 Premières propriétés

Propriété 0.1 : *Soient $n \in \mathbb{N}$ et $x \in [0,1]$. On a les propriétés remarquables suivantes :*

1. *partition de l'unité :*
$$\sum_{k=0}^{n} b_k^n(x) = 1;$$
2. *positivité :* $\forall k \in [\![0,n]\!]$, $b_n^k(x) \geqslant 0$ *;*
3. *symétrie :* $\forall k \in [\![0,n]\!]$, $b_k^n(x) = b_{n-k}^n(1-x)$ *;*
4. *valeurs particulières :* $b_k^n(0) = \delta_{k,0}$ *et* $b_k^n(1) = \delta_{k,n}$*, où* δ *est le symbole de Kronecker.*

Démonstration. Soient $n \in \mathbb{N}$ et $x \in [0,1]$.

1. D'après la formule du binôme de Newton,
$$\sum_{k=0}^{n} b_k^n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = (x+1-x)^n = 1.$$
2. Immédiat.
3. Par la symétrie du binôme
$$b_{n-k}^n(1-x) = \binom{n}{n-k}(1-x)^{n-k}x^k = \binom{n}{k}x^k(1-x)^{n-k}.$$
4. Immédiat.

□

On peut également établir les relations de récurrence suivantes :

Proposition 0.1 (Deux relations de récurrence) : *Soient* $n \in \mathbb{N}$ *et* $x \in [0,1]$.
$$b_k^{n+1}(x) = \begin{cases} (1-x)b_k^n(x) + x b_{k-1}^n(x) & \text{si } 1 \leqslant k \leqslant n \\ (1-x)b_0^n(x) & \text{si } k = 0 \\ x b_n^n(x) & \text{si } k = n+1. \end{cases} \tag{1}$$
Pour tout $k \in [\![0,n]\!]$,
$$b_k^n(x) = \frac{n+1-k}{n+1} b_k^{n+1}(x) + \frac{k+1}{n+1} b_{k+1}^{n+1}(x). \tag{2}$$

Démonstration. Soient $n \in \mathbb{N}$ et $x \in [0,1]$. Soit $k \in [\![0,k]\!]$.
Si $k \in [\![1,n]\!]$, d'après la formule de Pascal,
$$b_k^{n+1}(x) = \binom{n+1}{k} x^k (1-x)^{n+1-k} = \left(\binom{n}{k} + \binom{n}{k-1}\right) x^k (1-x)^{n+1-k} = (1-x)b_k^n(x) + x b_{k-1}^n(x).$$
Si $k = 0$,
$$b_0^{n+1}(x) = (1-x)^{n+1} = (1-x)b_0^n(x).$$
Si $k = n+1$,
$$b_{n+1}^{n+1}(x) = x^{n+1} = x b_n^n(x).$$

Pour $k \in [\![0, n]\!]$,

$$\frac{n+1-k}{n+1} b_k^{n+1}(x) + \frac{k+1}{n+1} b_{k+1}^{n+1} = \frac{n!}{k!(n-k)!} x^k (1-x)^{n+1-k} + \frac{n!}{k!(n-k)!} x^{k+1} (1-x)^{n-k}$$
$$= \binom{n}{k} x^k (1-x)^{n-k} (1-x+x) = b_k^n(x).$$

□

Proposition 0.2 (Dérivation) : *Pour tout $n \in \mathbb{N}^*$, pour tout $x \in [0,1]$, pour tout $k \in [\![0, n]\!]$,*

$$(b_k^n)'(x) = \begin{cases} n(b_{k-1}^{n-1}(x) - b_k^{n-1}(x)) & \text{si } 1 \leqslant k \leqslant n-1 \\ nx^{n-1} & \text{si } k = n \\ -n(1-x)^{n-1} & \text{si } k = 0. \end{cases}$$

Démonstration. On a, pour tous $n \in \mathbb{N}^*$, $x \in [0,1]$ et $k \in [\![1, n-1]\!]$,

$$\begin{aligned}(b_k^n)'(x) &= \binom{n}{k} k x^{k-1} (1-x)^{n-k} - \binom{n}{k} x^k (n-k)(1-x)^{n-k-1} \\ &= n\binom{n-1}{k-1} x^{k-1}(1-x)^{n-k} - n\binom{n-1}{k} x^k (1-x)^{n-k-1} \\ &= n(b_{k-1}^{n-1}(x) - b_k^{n-1}(x)).\end{aligned}$$

□

Propriété 0.2 : *Soit $n \in \mathbb{N}^*$. On a les résultats suivants pour les premiers monômes : pour tout $x \in [0,1]$,*

$$B_n(1)(x) = 1, \tag{3}$$
$$B_n(X)(x) = x, \tag{4}$$
$$B_n(X^2)(x) = \frac{1}{n}\big((n-1)x^2 + x\big). \tag{5}$$

Démonstration. Soient $n \in \mathbb{N}^*$ et $x \in [0,1]$.

$$B_n(1)(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1 \text{ par la formule du binôme de Newton}$$

$$\begin{aligned}B_n(X)(x) &= \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \frac{k}{n} \\ &= \sum_{k=1}^{n} \binom{n-1}{k-1} x^k (1-x)^{n-k}\end{aligned}$$

$$= \sum_{k=0}^{n-1} \binom{n-1}{k} x^{k+1}(1-x)^{n-1-k}$$

$$= x \text{ par la formule du binôme de Newton}$$

$$\begin{aligned}
B_n(X^2)(x) &= \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \frac{k^2}{n^2} \\
&= \frac{1}{n} \sum_{k=1}^{n} k \binom{n-1}{k-1} x^k (1-x)^{n-k} \\
&= \frac{1}{n} \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} x^{k+1} (1-x)^{n-1-k} \\
&= \frac{1}{n} \left(\sum_{k=0}^{n-1} k \binom{n-1}{k} x^{k+1} (1-x)^{n-1-k} + \sum_{k=0}^{n-1} \binom{n-1}{k} x^{k+1} (1-x)^{n-1-k} \right) \\
&= \frac{1}{n} \left((n-1) \sum_{k=1}^{n-1} \binom{n-1}{k-1} x^{k+1} (1-x)^{n-1-k} + x \right) \\
&= \frac{1}{n} \big((n-1)x^2 + x \big).
\end{aligned}$$

□

Définition 0.2 (Moment) : *Soient $n \in \mathbb{N}^*$ et $k \in \mathbb{N}^*$. On définit le moment d'ordre k par*

$$\forall x \in [0,1],\ S_n^k(x) = B_n((X-x)^k)(x) = \sum_{\ell=0}^{n} b_\ell^n(x) \left(\frac{\ell}{n} - x \right)^k .$$

Propriété 0.3 : *On a, pour les quatre premiers ordres :*

$$\begin{cases}
B_n(X-x)(x) = 0 \\
B_n((X-x)^2)(x) = B_n(X^2)(x) - 2xB_n(X) + x^2 = \frac{x(1-x)}{N} \\
B_n((X-x)^3)(x) = \frac{x(1-x)(1-2x)}{n^2} \ \text{[Bus17]} \\
B_n((X-x)^4)(x) = \frac{x(1-x)}{n^2} \left(\left(3 - \frac{6}{n}\right) x(1-x) + \frac{1}{n} \right) \ \text{[Bus17].}
\end{cases}$$

Définition 0.3 (Opérateur de différence) : *On considère une subdivison régulière de $[0,1]$, $0 = x_0 < x_1 < \cdots < x_n = 1$ et l'on note ε le pas. On définit la différence première Δ_ε par*

$$\Delta_\varepsilon f(x_i) = f(x_{i+1}) - f(x_i),$$

pour tout $i \in [\![0, n-1]\!]$.

On définit, de même, la différence seconde Δ_ε^2 par

$$\Delta_\varepsilon^2 f(x_i) = \Delta_\varepsilon(\Delta_\varepsilon f(x_i)) = f(x_{i+2}) - 2f(x_{i+1}) + f(x_i),$$

pour tout $i \in [\![0, n-2]\!]$.
Ainsi, par récurrence, pour tout $k \in [\![0, n-1]\!]$,

$$\Delta_\varepsilon^{k+1} f(x_i) = \Delta_\varepsilon(\Delta_\varepsilon^k f(x_i)) = \Delta_\varepsilon^k f(x_{i+1}) - \Delta_\varepsilon^k f(x_i),$$

pour tout $i \in [\![0, n-k-1]\!]$.
On impose $\Delta_\varepsilon^0 f(x_i) = f(x_i)$.

Dans le cadre des polynômes de Bernstein, on s'intéresse aux subdivisions régulières de pas de type $\frac{1}{n}$ ($n \in \mathbb{N}^*$), pour les polynômes $B_n(f)$, ce qui nous permettra de réécrire les coefficients de $B_n(f)$ en fonction de cet opérateur.

✍ **Notation 2.** Dans la suite, on notera Δ pour $\Delta_{1/n}$.

Proposition 0.3 : *Soit $f \in \mathcal{C}([0,1], \mathbb{R})$. Soit $n \in \mathbb{N}^*$. Pour tout $k \in [\![0, n]\!]$,*

$$\Delta^k f(0) = \sum_{\ell=0}^{k} \binom{k}{\ell} (-1)^{k-\ell} f\left(\frac{\ell}{n}\right).$$

Démonstration. Soient f une fonction continue sur $[0,1]$ et $n \in \mathbb{N}^*$.
Montrons par récurrence sur $k \in [\![0, n]\!]$ que l'assertion ($\mathcal{H}_k$) suivante est vraie :

$$\Delta^k f(0) = \sum_{\ell=0}^{k} \binom{k}{\ell} (-1)^{k-\ell} f\left(\frac{\ell}{n}\right). \qquad (\mathcal{H}_k)$$

- $k = 0$: $\Delta^0 f(0) = f(0)$ par définition.
- Soit $k \in [\![0, n-1]\!]$ tel que ($\mathcal{H}_k$) soit vraie.
  On vérifie de même que
  $$\Delta^k f\left(\frac{1}{n}\right) = \sum_{\ell=0}^{k} \binom{k}{\ell} (-1)^{k-\ell} f\left(\frac{\ell+1}{n}\right)$$
  par translation.

$$\begin{aligned}
\Delta^{k+1} f(0) &= \Delta^k f\left(\frac{1}{n}\right) - \Delta^k f(0) \\
&= \sum_{\ell=0}^{k} \binom{k}{\ell} (-1)^{k-\ell} f\left(\frac{\ell+1}{n}\right) - \sum_{\ell=0}^{k} \binom{k}{\ell} (-1)^{k-\ell} f\left(\frac{\ell}{n}\right) \text{ par hypothèse } (\mathcal{H}_k) \\
&= \sum_{\ell=1}^{k+1} \binom{k}{\ell-1} (-1)^{k+1-\ell} f\left(\frac{\ell}{n}\right) - \sum_{\ell=0}^{k} \binom{k}{\ell} (-1)^{k-\ell} f\left(\frac{\ell}{n}\right)
\end{aligned}$$

$$= f\left(\frac{k+1}{n}\right) + \sum_{\ell=1}^{k}(-1)^{k+1-\ell}\left(\binom{k}{\ell}+\binom{k}{\ell-1}\right) f\left(\frac{\ell}{n}\right) + (-1)^{k+1}f(0)$$

$$= \sum_{\ell=0}^{k+1}(-1)^{k+1-\ell}\binom{k+1}{\ell} f\left(\frac{\ell}{n}\right) \text{ par la formule de Pascal.}$$

□

Corollaire 0.1 : *Soit $f \in \mathcal{C}([0,1],\mathbb{R})$. Alors, pour tout $n \in \mathbb{N}^*$ et pour tout $x \in [0,1]$,*

$$B_n(f)(x) = \sum_{k=0}^{n} \Delta^k f(0)\binom{n}{k} x^k.$$

Démonstration. On a, d'après la Proposition 0.3,

$$\begin{aligned}
\sum_{k=0}^{n}\binom{n}{k}\Delta^k f(0)x^k &= \sum_{k=0}^{n}\binom{n}{k}\sum_{\ell=0}^{k}\binom{k}{\ell}(-1)^{k-\ell} f\left(\frac{\ell}{n}\right)x^k \\
&= \sum_{\ell=0}^{n}\sum_{k=\ell}^{n}\frac{n!}{(n-k)!\ell!(k-\ell)!}(-1)^{k-\ell} f\left(\frac{\ell}{n}\right)x^k \\
&= \sum_{\ell=0}^{n}\sum_{k=\ell}^{n}\frac{n!}{\ell!(n-\ell)!}\times\frac{(n-\ell)!}{(n-k)!(k-\ell)!}(-1)^{k-l} f\left(\frac{\ell}{n}\right)x^k \\
&= \sum_{\ell=0}^{n}\binom{n}{\ell}\sum_{k=\ell}^{n}\binom{n-\ell}{k-\ell}(-1)^{k-\ell}x^k f\left(\frac{\ell}{n}\right) \\
&= \sum_{\ell=0}^{n}\binom{n}{\ell}\sum_{k=0}^{n-\ell}\binom{n-\ell}{k}(-1)^k x^\ell x^k f\left(\frac{\ell}{n}\right) \\
&= \sum_{l=0}^{n}\binom{n}{\ell}x^\ell(1-x)^{n-\ell} f\left(\frac{\ell}{n}\right).
\end{aligned}$$

□

1

Convergence des polynômes de Bernstein

Ce chapitre est consacré à la démonstration du Théorème d'approximation de Weierstrass et à quelques propriétés des opérateurs de Bernstein appliqués à certains types de fonctions.

1.1 Approximation des fonctions continues

Théorème 1.1 (Weierstrass) : *Toute fonction continue sur $[0,1]$ est limite uniforme d'une suite de polynômes.*

Remarque.

- Ce résultat reste valable pour toute fonction continue sur tout segment $[a,b]$, $a < b \in \mathbb{R}$: pour une fonction f définie et continue sur un segment $[a,b]$, on se ramène à une fonction continue sur $[0,1]$, $g(x) = f((b-a)x+a)$, pour tout $x \in [0,1]$.
- Ce théorème a fait l'objet de plusieurs démonstrations et de généralisations, le Théorème de Stone-Weierstrass ou le Théorème de Müntz.
- Il s'agit d'un théorème d'existence qui n'indique rien sur la façon de construire une telle suite de polynômes.

On se propose de donner deux démonstrations du Théorème Théorème 1.1, toutes deux faisant intervenir les polynômes de Bernstein. Pour la première, énonçons le Lemme suivant :

Lemme 1.1 : *Soit f continue sur $[0,1]$. Alors, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que*

$$|f(x) - f(y)| \leqslant \varepsilon + \frac{2\|f\|_\infty}{\delta^2}(x-y)^2.$$

Démonstration. Soit f une fonction continue sur $[0,1]$.
Alors, d'après le Théorème de Heine, elle y est uniformément continue.
Soit $\varepsilon > 0$.

Par l'uniforme continuité de f, il existe $\delta > 0$ tel que pour tout $x, y \in [0,1]$, si $|x-y| \leqslant \delta$, alors $|f(x) - f(y)| \leqslant \varepsilon$.
Soient $x, y \in [0,1]$.

- *1er cas* : si $|x-y| \leqslant \delta$, alors
$$|f(x) - f(y)| \leqslant \varepsilon \leqslant \varepsilon + \frac{2\|f\|_\infty}{\delta^2}(x-y)^2.$$
- *2e cas* : si $|x-y| > \delta$, alors $\left|\frac{x-y}{\delta}\right| \geqslant 1$.
Donc
$$|f(x) - f(y)| \leqslant 2\|f\|_\infty \leqslant 2\|f\|_\infty \left|\frac{x-y}{\delta}\right| \leqslant 2\|f\|_\infty \frac{(x-y)^2}{\delta^2} \leqslant \varepsilon + 2\|f\|_\infty \frac{(x-y)^2}{\delta^2}.$$

□

Procédons maintenant à la première démonstration du Théorème 1.1.

Démonstration. Montrons que pour tout $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$ et pour tout $x \in [0,1]$,
$$|B_n(f)(x) - f(x)| \leqslant \varepsilon.$$
Soit $\varepsilon > 0$.
Alors, d'après le Lemme 1.1, il existe $\delta > 0$ tel que pour tous $x, y \in [0,1]$,
$$|f(x) - f(y)| \leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{\delta^2}(x-y)^2.$$
Soit $x \in [0,1]$.
$$\begin{aligned}
|B_n(f)(x) - f(x)| &= \left|\sum_{k=0}^{n} b_k^n(x) f\left(\frac{k}{n}\right) - f(x)\right| \\
&= \left|\sum_{k=0}^{n} b_k^n(x)\left(f\left(\frac{k}{n}\right) - f(x)\right)\right| \text{ car } \sum_{k=0}^{n} b_k^n = 1 \\
&\leqslant \sum_{k=0}^{n} b_k^n(x) \left|f\left(\frac{k}{n}\right) - f(x)\right| \text{ par l'inégalité triangulaire} \\
&\leqslant \sum_{k=0}^{n} b_k^n(x) \left(\frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{\delta^2}(\frac{k}{n} - x)^2\right) \text{ d'après le Lemme 1.1} \\
&= \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{\delta^2}(B_n(X^2)(x) - 2xB_n(X)(f) + x^2 B_n(1)).
\end{aligned}$$
En utilisant les résultats (3), (4) et (5),
$$\begin{aligned}
|B_n(f)(x) - f(x)| &\leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{\delta^2}(\frac{1}{n}((n-1)x^2 + x) - 2x^2 + x^2) \\
&\leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{\delta^2}(\frac{1}{n}((n-1)x^2 + 1) - x^2) \\
&\leqslant \frac{\varepsilon}{2} + \frac{2\|f\|_\infty}{\delta^2 n}.
\end{aligned}$$

Donc il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, pour tout $x \in [0,1]$,

$$|B_n(f)(x) - f(x)| \leqslant \varepsilon.$$

On a ainsi montré que $(B_n(f))_{n\in\mathbb{N}^*}$ converge uniformément vers f sur $[0,1]$. $\square$

La seconde démonstration fait intervenir les probabilités. Voici une première approche :

Soit $x \in [0,1]$. En considérant une variable aléatoire réelle $S_n(x)$ suivant la loi binomiale de taille n et de paramètre x, modélisant par exemple n lancers indépendants d'une pièce de monnaie, avec une probabilité x de tomber sur « pile », on peut établir, par la loi forte des grands nombres, que $\frac{S_n(x)}{n}$ converge presque sûrement vers x. Donc en composant par f, puisque f est continue, puis en appliquant le théorème de convergence dominé,

$$f(x) = \lim_{n\to\infty} \mathbb{E}\left(\frac{S_n}{n}\right) = \sum_{k=1}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}.$$

Malheureusement, on n'a montré qu'une convergence simple...

Voici donc la démonstration de la convergence uniforme.

Démonstration. Soit f une fonction continue sur $[0,1]$.
On considère une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ indépendantes et identiquement distribuées sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On suppose que X_1 suit la loi de Bernoulli de paramètre $p \in [0,1]$.
Posons $S_n = X_1 + X_2 + \cdots + X_n$.
S_n suit la loi binomiale de taille n et de paramètre p : $\mathbb{P}(S_n = k) = b_n^k(p)$.
Donc,

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{n}\mathbb{E}(S_n) = p,$$

et

$$\operatorname{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\operatorname{Var}(S_n) = \frac{p(1-p)}{n}.$$

f étant une fonction continue sur un compact, d'après le Théorème de Heine, elle y est uniformément continue.
Soit $\varepsilon > 0$.
Comme $p \in [0,1]$, par l'uniforme continuité de f, il existe $\delta > 0$ tel que

$$\forall x \in [0,1],\ |x-p| \leqslant \delta \Rightarrow |f(x) - f(p)| \leqslant \frac{\varepsilon}{2}.$$

On a

$$\begin{aligned}
|B_n(f)(p) - f(p)| &= \left| \sum_{k=0}^{n} b_n^k(p)\left(f\left(\frac{k}{n}\right) - f(p)\right)\right| \\
&\leqslant \sum_{k=0}^{n} b_n^k(p)\left|f\left(\frac{k}{n}\right) - f(p)\right| \text{ par l'inégalité triangulaire} \\
&= \mathbb{E}\left(\left|f\left(\frac{S_n}{n}\right) - f(p)\right|\right) \text{ par le Théorème de transfert.}
\end{aligned}$$

Or,

$$\begin{aligned}
\mathbb{E}\left(\left|f\left(\frac{S_n}{n}\right)-f(p)\right|\right) &= \sum_{k=1}^{n}\left|f\left(\frac{k}{n}\right)-f(p)\right|b_k^n(p) \\
&= \sum_{|\frac{k}{n}-p|\leqslant\delta} b_k^n(p)\left|f\left(\frac{k}{n}\right)-f(p)\right| + \sum_{|\frac{k}{n}-p|>\delta} b_k^n(p)\left|f\left(\frac{k}{n}\right)-f(p)\right| \\
&\leqslant \frac{\varepsilon}{2}\sum_{k=1}^{n} b_k^n(p) + 2\|f\|_\infty \sum_{|\frac{k}{n}-p|>\delta} b_k^n(p) \\
&= \frac{\varepsilon}{2} + 2\|f\|_\infty \mathbb{P}\left(\left|\frac{S_n}{n}-p\right|>\delta\right).
\end{aligned}$$

Et d'après l'inégalité de Bienaymé-Tchebytchev,

$$\mathbb{P}\left(\left|\frac{S_n}{n}-p\right|>\delta\right) \leqslant \frac{\mathrm{Var}(\frac{S_n}{n})}{\delta^2} = \frac{p(1-p)}{n\delta^2} \leqslant \frac{1}{4n\delta^2}.$$

On a ainsi majoré $\mathbb{P}\left(\left|\frac{S_n}{n}-p\right|>\delta\right)$ indépendamment de p.
Donc, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geqslant n_0$, pour tout $p \in [0,1]$,

$$\frac{1}{4n\delta^2} \leqslant \frac{\varepsilon}{4\|f\|_\infty}.$$

Par conséquent, pour tout $\varepsilon > 0$, pour tout $n \geqslant n_0$, avec n_0 ainsi fixé, pour tout $p \in [0,1]$,

$$|B_n(f)(p) - f(p)| = \mathbb{E}\left(\left|f\left(\frac{S_n}{n}\right)-f(p)\right|\right) \leqslant \varepsilon.$$

Le Théorème 1.1 est ainsi démontré. □

1.2 Propriétés de régularité

1.2.1 Fonctions lipschitziennes

Soit $k \geqslant 0$. Une fonction $f \in \mathcal{C}([0,1],\mathbb{R})$ est dite k-lipschitzienne si pour tous $x, y \in [0,1]$,

$$|f(x) - f(y)| \leqslant k|x-y|.$$

Propriété 1.1 : *Soit $M \geqslant 0$. Soit f une fonction M-lipschitzienne. Alors, pour tout $n \in \mathbb{N}^*$,*

$$\|f - B_n(f)\|_\infty \leqslant \frac{M}{2\sqrt{n}}.$$

Remarque. Cette propriété est donc vraie en particulier pour les fonctions continues sur $[0,1]$, dérivables sur $]0,1[$ et de dérivée bornée, puisque l'on peut appliquer l'inégalité des accroissements finis.

Démonstration. On a

$$\begin{aligned}
\left|f(x)-B_n(f)(x)\right| &= \Big|\sum_{k=0}^{n} b_k^n(x)(f\left(\frac{k}{n}\right)-f(x))\Big| \\
&\leqslant \sum_{k=0}^{n} b_k^n(x)\Big|f\left(\frac{k}{n}\right)-f(x)\Big| \\
&\leqslant M\sum_{k=0}^{n} b_k^n(x)\Big|\frac{k}{n}-x\Big| \text{ par l'inégalité des accroissements finis} \\
&= \|f'\|_\infty \mathbb{E}\left(\Big|\frac{S_n(x)}{n}-x\Big|\right) \\
&\leqslant M\mathbb{E}\left(\Big|\frac{S_n(x)}{n}-x\Big|^2\right)^{1/2} \text{ par l'inégalité de Cauchy-Schwarz} \\
&= M\sqrt{\operatorname{Var}\left(\frac{S_n}{n}\right)} \\
&\leqslant M\sqrt{\frac{x(1-x)}{n}} \\
&\leqslant \frac{M}{2\sqrt{n}}.
\end{aligned}$$

□

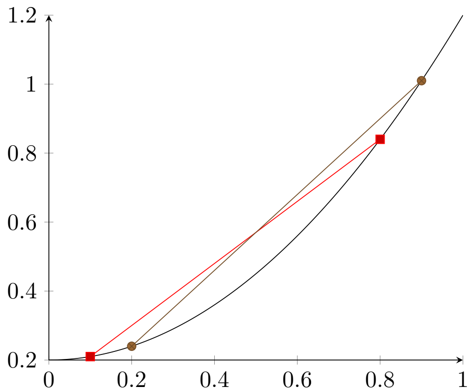
1.2.2 Fonctions convexes

Rappelons la définition d'une fonction convexe :

Définition 1.1 (Fonction convexe) : *Une fonction f définie sur un intervalle I est dite convexe si, pour tout $\lambda \in [0,1]$, pour tout $x, y \in I$,*

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y).$$

f est dite concave si $-f$ est convexe.



Graphiquement, cela se traduit par le fait que toutes les cordes sont au dessus de la courbe représentative de f.

Remarque. Dans la suite, nous allons énoncer des propriétés pour les fonctions convexes, mais il est facile de transposer ces résultats aux fonctions concaves puisque si f est concave, $-f$ est convexe.

Théorème 1.2 (Caractérisation de la convexité avec la dérivée seconde) : *oit f une fonction définie sur un intervalle I deux fois dérivable. Alors f est convexe si, et seulement si $f'' \geqslant 0$.*

Proposition 1.1 : *Si f est convexe, alors ses polynômes de Bernstein sont convexes.*

Soit $n \in \mathbb{N}^*$. $B_n(f)$ étant un polynôme, nous pouvons calculer sa dérivée seconde pour conclure. Afin de rendre le calcul de sa dérivée moins fastidieux, introduisons le lemme suivant :

Lemme 1.2 : *Soit $n \geqslant 2$. Si f est convexe sur $[0,1]$, alors $\Delta^2 f\left(\frac{k}{n}\right) \geqslant 0$ pour tout $k \in [\![0, n-2]\!]$.*

Démonstration. Soit $k \in [\![0, n-2]\!]$.

$$\Delta^2 f\left(\frac{k}{n}\right) = f\left(\frac{k+2}{n}\right) - f\left(\frac{k+1}{n}\right) - \left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right)$$

Notons, pour $a, b \in [0,1]$,

$$p(f; a, b) = \frac{f(b) - f(a)}{b - a}.$$

D'après l'inégalité des pentes,

$$p(f; \frac{k}{n}, \frac{k+1}{n}) \leqslant p(f; \frac{k+1}{n}, \frac{k+2}{n}),$$

i.e.

$$n\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right) \leqslant n\left(f\left(\frac{k+2}{n}\right) - f\left(\frac{k+1}{n}\right)\right),$$

d'où le résultat. □

Remarque. L'opérateur Δ^2 agit comme une dérivée seconde discrète aux points $\frac{k}{n}$, ce qui permet d'étudier les variations d'une fonction continue mais non dérivable, comme par exemple $|\cdot|$. Il n'est donc pas surprenant d'avoir l'implication présentée dans le Lemme.

Procédons maintenant à la preuve de la Proposition 1.1.

Démonstration. Supposons que f est convexe.
$B_0(f)$ et $B_1(f)$ sont immédiatement convexes puisque leur dérivée seconde est nulle.
Soit $n \geqslant 2$. Montrons que $B_n''(f)$ est positive.
Pour tout $x \in [0,1]$,

$$
\begin{aligned}
(B_n(f))'(x) &= \sum_{k=0}^{n}(b_k^n)'(x) f\left(\frac{k}{n}\right) \\
&= nx^{n-1}f(1) - n(1-x)^{n-1}f(0) + \sum_{k=1}^{n-1}(b_{k-1}^{n-1}(x) - b_k^{n-1}(x)) f\left(\frac{k}{n}\right) \\
&\quad \text{d'après la Proposition 0.2} \\
&= nb_{n-1}^{n-1}f(1) - nb_0^{n-1}f(0) + n\sum_{k=1}^{n-2} b_k^{n-1}(x)\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right) \\
&\quad + n(1-x)^{n-1}f\left(\frac{1}{n}\right) - nx^{n-1}f\left(\frac{n-1}{n}\right) \\
&= n\sum_{k=0}^{n-1} b_k^{n-1}\left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right) \\
&= n\sum_{k=0}^{n-1} b_k^{n-1}(x)\Delta f\left(\frac{k}{n}\right).
\end{aligned}
$$

De la même façon, on calcule $B_n''(f)$:

$$
\begin{aligned}
B_n(f)''(x) = n(n-1)\sum_{k=0}^{n-2} b_k^{n-2}(x)\left(f\left(\frac{k+2}{n}\right) - 2f\left(\frac{k+1}{n}\right) + f\left(\frac{k}{n}\right)\right) \\
= n(n-1)\sum_{k=0}^{n-2} b_k^{n-2}(x)\Delta^2 f\left(\frac{k}{n}\right) \geqslant 0
\end{aligned}
$$

d'après le Lemme 1.2.

Donc pour tout $x \in [0,1]$, $B_n(f)''(x) \geqslant 0$.

□

Proposition 1.2 : *Soit f une fonction convexe. Alors, pour tout $n \in \mathbb{N}^*$ et pour tout $x \in [0,1]$,*

$$f(x) \leqslant B_n(f)(x).$$

Démonstration. Pour tout $x \in [0,1]$,

$$\sum_{k=0}^{n} b_k^n(x) = 1.$$

Donc d'après l'inégalité de Jensen,

$$f\left(\sum_{k=0}^{n} b_k^n(x)\frac{k}{n}\right) \leqslant \sum_{k=0}^{n} b_k^n(x) f\left(\frac{k}{n}\right),$$

i.e.

$$f(B_n(X)(x)) \leqslant B_n(f)(x),$$

i.e.

$$f(x) \leqslant B_n(f)(x).$$

□

Étudions maintenant la monotonie des polynômes de Bernstein. Intuitivement, si f est une fonction convexe et si $(B_n(f))_{n\in\mathbb{N}}$ est monotone, elle ne peux qu'être décroissante, puisqu'elle converge vers f et que f est en dessous de ses polynômes de Bernstein.

Proposition 1.3 : *Soit $f \in \mathcal{C}([0,1],\mathbb{R})$. Si f est convexe, alors $(B_n(f))_{n\in\mathbb{N}^*}$ est une suite décroissante.*

Démonstration. [Bus17] Soient $f : [0,1] \longrightarrow \mathbb{R}$ une fonction continue et convexe, $n \in \mathbb{N}^*$ et $x \in [0,1]$.
Supposons dans un premier temps que $f(0) = f(1) = 0$. On pose $f(x) = 0$ pour tout $x < 0$. Montrons que

$$\sum_{k=0}^{n} f\left(\frac{k}{n}\right) b_k^n(x) = \sum_{k=0}^{n+1}\left(\frac{n+1-k}{n+1} f\left(\frac{k}{n}\right) + \frac{k}{n+1} f\left(\frac{k-1}{n}\right)\right) b_k^{n+1}(x).$$

On a

$$\begin{aligned}
\sum_{k=0}^{n+1}\left(\frac{n+1-k}{n+1} f\left(\frac{k}{n}\right) + \frac{k}{n+1} f\left(\frac{k-1}{n}\right)\right) b_k^{n+1}(x) &= \sum_{k=0}^{n} \frac{n+1-k}{n+1} f\left(\frac{k}{n}\right) b_k^{n+1}(x) \\
&\quad + \sum_{k=1}^{n+1} \frac{k}{n+1} f\left(\frac{k-1}{n}\right) \\
&= \sum_{k=0}^{n} \frac{n+1-k}{n+1} f\left(\frac{k}{n}\right) \\
&\quad + \sum_{k=0}^{n} \frac{k+1}{n+1} f\left(\frac{k}{n}\right) b_{k+1}(x) \\
&= \sum_{k=0}^{n} (1-x) f\left(\frac{k}{n}\right) b_k^n(x) \\
&\quad + \sum_{k=0}^{n} x f\left(\frac{k}{n}\right) b_{k+1}^{n+1}(x) \\
&= \sum_{k=0}^{n} f\left(\frac{k}{n}\right) b_k^n(x).
\end{aligned}$$

Alors,

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

Si maintenant $f(0) \neq 0$ ou $f(1) \neq 0$, en considérant la fonction affine telle que $P(0) = f(0)$ et $P(1) = f(1)$, alors $g := f - P$ vérifie $g(0) = g(1) = 0$.
On applique le résultat précédent à $B_n(f-P) - B_{n+1}(f-P) = B_n(f) - B_{n+1}(f)$.
Ainsi, dans tous les cas,

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

Or, par convexité de f pour tout $k \in [\![1, n]\!]$,

$$\begin{aligned} \frac{n+1-k}{n+1} f\left(\frac{k}{n}\right) + \frac{k}{n+1} f\left(\frac{k-1}{n}\right) &\geqslant f\left(\frac{k(n+1-k)}{n(n+1)} + \frac{k(k-1)}{n(n+1)}\right) \\ &= f\left(\frac{k}{n+1}\right). \end{aligned}$$

Donc $B_n(f)(x) - B_{n+1}(f)(x) \geqslant 0$. □

✍ **Exemple 1.** Étudions la fonction $f(x) = |x - \frac{1}{2}|$, pour tout $x \in [0, 1]$.

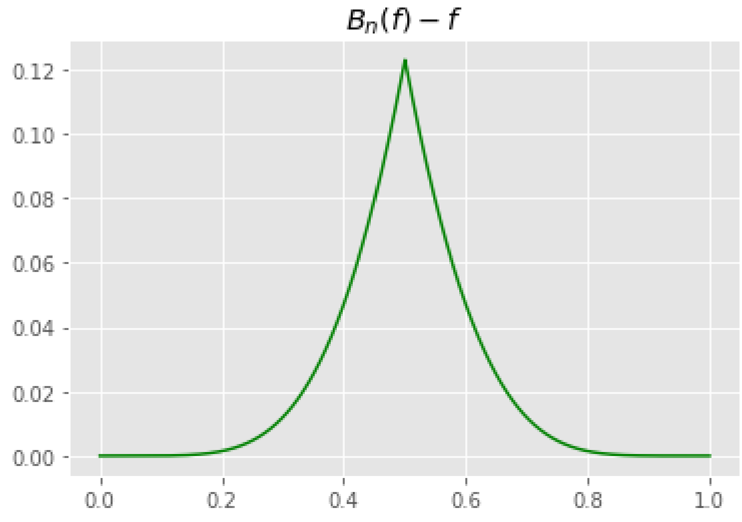


Figure 1.1. Graphe de $B_n(f) - f$ pour $n = 10$ sur Python

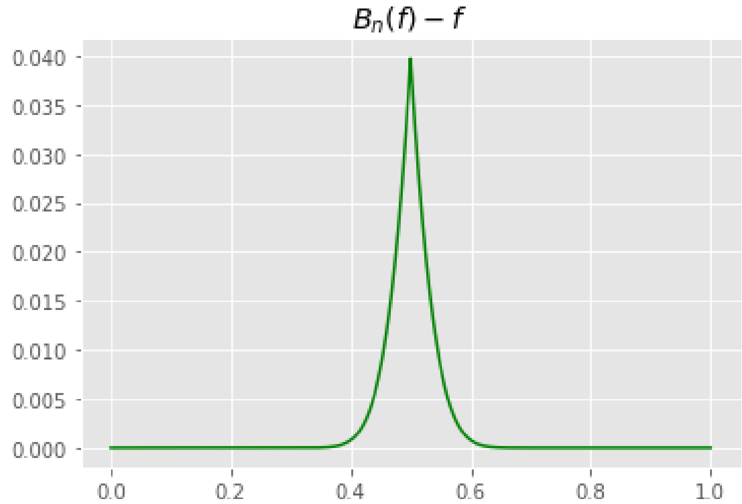


Figure 1.2. Graphe de $B_n(f) - f$ pour $n = 100$ sur Python

On peut constater graphiquement que $\|B_n(f) - f\|_\infty$ est atteint en $x = 1/2$.
Pour tout $p \in \mathbb{N}$,

$$\begin{aligned}
B_{2p+1}(f)\left(\frac{1}{2}\right) &= \frac{1}{2^{2p+1}} \sum_{k=0}^{2p+1} \binom{2p+1}{k} \left|\frac{k}{n} - \frac{1}{2}\right| \\
&= \frac{1}{2^{2p+1}} \left(\sum_{k=0}^{p} \binom{2p+1}{k} \left(\frac{1}{2} - \frac{k}{2p+1}\right) + \sum_{k=p+1}^{2p+1} \binom{2p+1}{k} \left(\frac{k}{2p+1} - \frac{1}{2}\right) \right) \\
&= \frac{1}{2^{2p+1}} \left(\sum_{k=p+1}^{2p+1} \binom{2p+1}{k} \frac{k}{2p+1} - \sum_{k=1}^{p} \binom{2p+1}{k} \frac{k}{2p+1} \right) \\
&= \frac{1}{2^{2p+1}} \left(\sum_{k=0}^{p} \binom{2p+1}{k+p+1} \frac{k+p+1}{2p+1} - \sum_{k=1}^{p} \binom{2p+1}{k} \frac{k}{2p+1} \right) \\
&= \frac{(2p)!}{2^{2p+1}} \left(\sum_{k=0}^{p} \left(\frac{1}{(p-k)!(p+k)!} - \frac{1}{k!(2p-k)!}\right) + \frac{1}{(p!)^2} \right) \\
&= \frac{(2p)!}{2^{2p+1}} \left(\sum_{k=0}^{p} \left(\frac{1}{k!(2p-k)!} - \frac{1}{k!(2p-k)!}\right) + \frac{1}{(p!)^2} \right) \\
&\underset{p \to \infty}{\sim} \frac{(\frac{2p}{e})^{2p}\sqrt{4\pi p}}{(\frac{p}{e})^{2p}(2\pi p)2^{2p+1}} \text{ par la formule de Stirling} \\
&\underset{p \to \infty}{\sim} \frac{1}{2\sqrt{\pi p}}.
\end{aligned}$$

Donc

$$\sqrt{2p+1} B_{2p+1}(f)\left(\frac{1}{2}\right) \underset{p \to \infty}{\sim} \frac{1}{\sqrt{2\pi}}.$$

En particulier, il existe $C \in \mathbb{R}_+^*$ tel que

$$\frac{C}{\sqrt{2p+1}} \leqslant B_{2p+1}(f)\left(\frac{1}{2}\right).$$

Or,

$$B_{2p+1}(f)\left(\frac{1}{2}\right) = \left|B_{2p+1}(f)\left(\frac{1}{2}\right) - f\left(\frac{1}{2}\right)\right| \leqslant \|B_{2p+1}(f) - f\|_\infty.$$

Donc

$$\frac{C}{\sqrt{2p+1}} \leqslant \|B_{2p+1}(f) - f\|_\infty.$$

Comme f est convexe, la monotonie de $(B_n(f))_{\in\mathbb{N}^*}$ suffit pour conclure qu'il existe $C \in \mathbb{R}_+^*$ tel que

$$\frac{C}{\sqrt{n}} \leqslant \|B_n(f) - f\|_\infty.$$

pour tout $n \in \mathbb{N}^*$

Remarque. Nous venons d'exhiber une fonction dont la convergence uniforme est en $\mathcal{O}(\frac{1}{\sqrt{n}})$. Cette fonction ayant des propriétés de régularité (lipschitzienne et convexe), on pourrait s'attendre à ce que l'erreur d'approximation soit moins importante. Cela est dû à la conservation de ces propriétés par l'opérateur de Bernstein : comme les polynômes de Bernstein associés à une fonction convexe f constituent une suite de polynômes croissante restant au dessus de f, celle-ci n'oscille pas autour de f.

2

Fonction d'entropie et modèle de Wright-Fisher

2.1 Définition et quelques propriétés intéressantes

Dans cette section, on considère l'ensemble $E := \{f \in \mathcal{C}([0,1],\mathbb{R}),\ f(0) = f(1) = 0\}$, muni de la norme $\|\cdot\|_\infty$.
On définit la fonction d'entropie comme suit [Bus17] [1] :

Définition 2.1 (Fonction d'entropie) : *La fonction d'entropie $\tau : [0,1] \longrightarrow \mathbb{R}$ est définie par*

$$\forall x \in]0,1[,\ \tau(x) = -2(x\ln x + (1-x)\ln(1-x)) \text{ et } \tau(0) = \tau(1) = 0.$$

On pose, pour $x \in]0,1[$, $\tau^1(x) = -x\ln x$, de sorte que $\tau(x) = \tau^1(x) + \tau^1(1-x)$.

Remarque. Calculons, pour tout $x \in]0,1[$, les trois premières dérivées, qui nous seront bien utiles par la suite :

- $\tau'(x) = -2(\ln x - \ln(1-x))$;
- $\tau''(x) = -2(\frac{1}{x} + \frac{1}{1-x}) = -\frac{2}{x(1-x)}$;
- $\tau^{(3)}(x) = -2(\frac{1}{(1-x)^2} - \frac{1}{x^2}) = 2\frac{1-2x}{x^2(1-x)^2}$.

Par le calcul de dérivée seconde, on établit que τ est une fonction concave.
Donc d'après la Proposition 1.2, τ est au-dessus de ses polynômes de Bernstein.
Soit $N \geqslant 2$.
On définit la fonction $\mathbb{1}_N$ affine par morceaux par

$$\forall k \in [\![1, N-1]\!],\ \mathbb{1}_N\left(\frac{k}{N}\right) = 1 \text{ et } \mathbb{1}_N(0) = \mathbb{1}_N(1) = 0.$$

1. à un facteur multiplicatif près

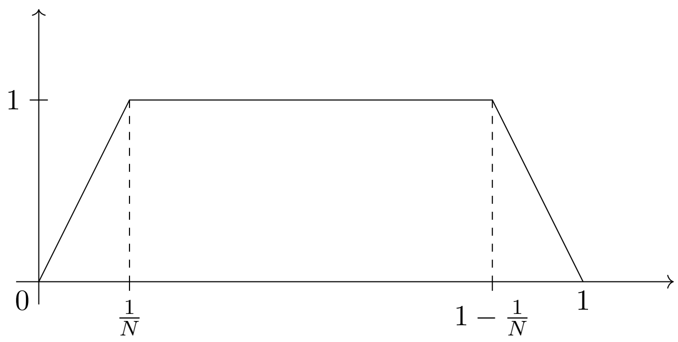


Figure 2.1. Courbe représentative de $\mathbb{1}_N$

On s'intéresse ici aux opérateurs $F_N := N(\mathrm{Id} - B_N)$ définis sur E.

Théorème 2.1 : *L'équation $F_N(f_N) = \mathbb{1}_N$ d'inconnue f_N admet une unique solution τ_N dans E. Cette fonction est concave.*

Démonstration. Pour toute fonction $f \in E$ vérifiant $\|f\|_\infty = 1$,

$$\forall x \in [0,1], |B_N(f)(x)| = \left| \sum_{k=1}^{N-1} b_k^N(x) f\left(\frac{k}{N}\right) \right| \leqslant \sum_{k=1}^{N-1} b_k^N(x) \|f\|_\infty < 1.$$

Donc $\|B_N(f)\|_{\mathcal{L}(E)} < 1$.
Ainsi, l'opérateur $\mathrm{Id} - B_N \in \mathcal{L}(E)$ est inversible, et

$$(\mathrm{Id} - B_N)^{-1} = \sum_{k=0}^{+\infty} B_N^k.$$

Donc

$$F_N^{-1} = \frac{1}{N} \sum_{k=0}^{+\infty} B_N^k.$$

Comme $\mathbb{1}_N \in E$, il existe un unique élément $\tau_N \in E$ tel que

$$F_N(\tau_N) = \mathbb{1}_N,$$

et

$$\tau_N = F_N^{-1}(\mathbb{1}_N) = \frac{1}{N} \sum_{k=0}^{+\infty} B_N^k(\mathbb{1}_N).$$

Comme $\mathbb{1}_N$ est concave, $B_N(\mathbb{1}_N)$ est concave. On établit par récurrence que, pour tout $k \in \mathbb{N}$, $B_N^k(\mathbb{1}_N)$ est concave (puisque B_N conserve la concavité).
Donc, par somme, τ_N est concave. □

On obtient, par construction, une suite de fonctions concaves $(\tau_N)_{N \in \mathbb{N}^*}$.

Théorème 2.2 (Voronovskaya) : *Pour toute fonction $f \in E \cap \mathcal{C}^2([0,1],\mathbb{R})$, pour tout $x \in [0,1]$,*

$$\lim_{N\to\infty} N(f(x) - B_N f(x)) = -\frac{f''(x)}{2}x(1-x).$$

Démonstration. Soit $f \in E \cap \mathcal{C}^2(]0,1[,\mathbb{R})$. Si f est un polynôme de degré inférieur ou égal à 1, le résultat est immédiat. Supposons que f n'est pas polynomial de degré inférieur à 1.
Montrons que pour tout $x \in [0,1]$,

$$\lim_{N\to\infty} N(B_N(f)(x) - f(x)) = \frac{f''(x)x(1-x)}{2}.$$

Soit $x \in]0,1[$.
Soit $k \in [\![0,N]\!]$.
D'après la formule de Taylor-Lagrange, si $x \neq \frac{k}{N}$, il existe $\zeta_k \in [0,1]$ tel que $|\zeta_k - x| \leqslant \frac{k}{N}$, et

$$\begin{aligned} f\left(\frac{k}{N}\right) &= f(x) + f'(x)\left(\frac{k}{N} - x\right) + \frac{f''(\zeta_k)}{2}\left(\frac{k}{N} - x\right)^2 \\ &= f(x) + f'(x)\left(\frac{k}{N} - x\right) + \frac{f''(x)}{2}\left(\frac{k}{N} - x\right)^2 + \frac{f''(\zeta_k) - f''(x)}{2}\left(\frac{k}{N} - x\right)^2. \end{aligned}$$

Donc par linéarité de l'opérateur de Bernstein B_N,

$$\begin{aligned} B_N(f)(x) &= f(x) + f'(x)(B_N(X)(x) - x) + \frac{f''(x)}{2}(B_N(X^2)(x) - 2xB_N(X)(x) + x^2) \\ &\quad + \sum_{k=0}^{N} b_k^N(x)\frac{f''(\zeta_k)}{2}\left(\frac{k}{N} - x\right)^2 \\ &= f(x) + \frac{f''(x)}{2N}x(1-x) + \sum_{k=0}^{N} b_k^N(x)\frac{f''(\zeta_k) - f''(x)}{2}\left(\frac{k}{N} - x\right)^2. \end{aligned}$$

Donc

$$\left| N(B_N(f)(x) - f(x)) - \frac{f''(x)}{2}x(1-x) \right| \leqslant \frac{N}{2}\sum_{k=0}^{N} b_k^N(x)\left|f''(\zeta_k) - f''(x)\right|\left(\frac{k}{N} - x\right)^2.$$

Soit $\varepsilon > 0$.
D'après le Lemme 1.1, il existe $\delta > 0$ tel que pour tout $(x_1, x_2) \in [0,1]^2$,

$$|f''(x_1) - f''(x_2)| \leqslant \frac{2\varepsilon}{\|f''\|_\infty x(1-x)} + \frac{2\|f''\|_\infty}{\delta^2}(x_1 - x_2)^2.$$

Donc

$$\sum_{k=0}^{N} b_k^N(x)\left|f''(\zeta_k) - f''(x)\right|\left(\frac{k}{N} - x\right)^2 \leqslant \sum_{k=0}^{N} b_k^N(x)(1 + \frac{2\|f''\|_\infty}{\delta^2}(\zeta_k - x)^2)\left(\frac{k}{N} - x\right)^2$$

$$\leqslant \frac{\varepsilon}{2} + \frac{2\|f''\|_\infty}{\delta^2} \sum_{k=0}^{N} b_k^N(x) \left(\frac{k}{N} - x\right)^4.$$

D'après la Propriété 0.3,

$$\sum_{k=0}^{N} b_k^N(x) \left(\frac{k}{N} - x\right)^4 = \frac{x(1-x)}{N^2} \left(\left(3 - \frac{6}{N}\right) x(1-x) + \frac{1}{N}\right) = \mathcal{O}\left(\frac{1}{N^2}\right).$$

Donc il existe $N_0 \in \mathbb{N}$ tel que pour tout $N \geqslant N_0$,

$$N \frac{2\|f''\|_\infty}{\delta^2} \sum_{k=0}^{N} b_k^N(x) \left(\frac{k}{N} - x\right)^4 \leqslant \frac{\varepsilon}{2}.$$

Donc pour tout $N \geqslant N_0$,

$$\left| N(B_N(f)(x) - f(x)) - \frac{f''(x)}{2} x(1-x) \right| \leqslant \varepsilon.$$

□

Théorème 2.3 : *La suite $(\tau_N)_{N \in \mathbb{N}^*}$ converge simplement vers la fonction τ.*

Montrons le résultat suivant :

Lemme 2.1 : *Soient f et $(f_n)_{n\in\mathbb{N}}$ des éléments de E, respectivement une fonction positive et une suite de fonctions concaves. On suppose de plus que f est dérivable sur $]0,1[$, de dérivée bornée sur $]0,1[$, et qu'il existe $C > 0$ tel que pour tout $x \in [0,1]$, pour tout $n \in \mathbb{N}$,*

$$f_n(x) \leqslant C f(x).$$

Alors $(f_n)_{n\in\mathbb{N}}$ est équicontinue.

Remarque. $(f_n)_{n\in\mathbb{N}}$ étant une suite de fonctions concaves de E, elles sont nécessairement positives.

Démonstration. Montrons que pour tout $x \in [0,1]$, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $y \in]x-\delta, x+\delta[\cap [0,1]$, pour tout $n \in \mathbb{N}$,

$$|f_n(x) - f_n(y)| \leqslant \varepsilon.$$

Comme f est continue sur $[0,1]$ et dérivable sur $]0,1[$ de dérivée bornée sur $]0,1[$ d'après l'inégalité des accroissements finis, il existe $M > 0$ tel que pour tout $x, y \in [0,1]$,

$$|f(x) - f(y)| \leqslant M|x-y|.$$

Soit $x \in [0,1]$.

Soit $\varepsilon > 0$.
Comme f est continue en 0 et 1, et $f(0) = f(1) = 0$, il existe $\delta_0 > 0$ et $\delta_1 > 0$ tels que pour tout $y \in [0, \delta_0[$,

$$|f(y) - f(0)| = f(y) \leqslant \frac{\varepsilon}{C}$$

et pour tout $y \in]1 - \delta_1, 1]$,

$$|f(y) - f(1)| = f(y) \leqslant \frac{\varepsilon}{C}.$$

Posons $\delta_\varepsilon = \min(\frac{\delta_0}{C}, \frac{\delta_1}{C}, \frac{\varepsilon}{MC}) > 0$.
Supposons dans un premier temps que $x \in]0, 1[$.
Soit $y \in]x - \delta_\varepsilon, x + \delta_\varepsilon[\cap [0, 1]$. Si $y = x$, il n'y a rien à montrer. Supposons que $y \neq x$.

- *1er cas* : $y > x$. Pour tout $n \in \mathbb{N}$, d'après l'inégalité des pentes,

$$\frac{f_n(x) - f_n(y)}{x - y} \geqslant \frac{f_n(x) - f_n(1)}{x - 1} = \frac{f_n(x)}{x - 1},$$

  donc

$$f_n(x) - f_n(y) \leqslant \frac{x - y}{x - 1} f_n(x) = \frac{y - x}{1 - x} f_n(x) \leqslant \frac{y - x}{1 - x} Cf(x) \leqslant \delta_\varepsilon CM = \varepsilon.$$

  De même, d'après l'inégalité des pentes,

$$f_n(y) - f_n(x) \leqslant \frac{y - x}{x} f_n(x) \leqslant \frac{y - x}{x} Cf(x) \leqslant \varepsilon.$$

  Ainsi, pour tout $y \in]x, x + \delta_\varepsilon[$

$$|f_n(x) - f_n(y)| \leqslant \varepsilon.$$

- *2e cas* : $y < x$. Pour tout $n \in \mathbb{N}$,

$$f_n(x) - f_n(y) \leqslant (x - y) \frac{f_n(x)}{x} \leqslant \delta_\varepsilon MC \leqslant \varepsilon,$$

  et

$$f_n(y) - f_n(x) \leqslant \frac{x - y}{1 - x} f_n(x) \leqslant \varepsilon.$$

  Ainsi, pour tout $y \in]x - \delta_\varepsilon, x[$,

$$|f_n(x) - f_n(y)| \leqslant \varepsilon.$$

Traitons maintenant les cas particuliers $x = 0$ et $x = 1$. Pour tout $n \in \mathbb{N}$, pour tout $y \in [0, \delta_\varepsilon[$,

$$|f_n(0) - f_n(y)| = f_n(y) \leqslant Cf(y) \leqslant \delta_0 C \leqslant \varepsilon.$$

De même, pour tout $y \in]1 - \delta_\varepsilon, 1]$,

$$|f_n(1) - f_n(y)| \leqslant \varepsilon.$$

On a ainsi montré que pour tout $x \in [0, 1]$, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $y \in]x - \delta, x + \delta[\cap [0, 1]$, pour tout $n \in \mathbb{N}$,

$$|f_n(x) - f_n(y)| \leqslant \varepsilon.$$

□

Démonstration. On vérifie que τ est dans $\mathcal{C}^4(|]0,1[,\mathbb{R})$, et que $\tau^{(4)}_{]0,1[} \leqslant 0$, puisque τ'' est concave, ce qui permet d'appliquer le Théorème de Taylor avec reste intégral, et ainsi d'écrire, pour tout $x \in]0,1[$, et tout $k \in [\![1, N-1]\!]$,

$$\tau\left(\frac{k}{N}\right) \leqslant \tau(x) + \tau'(x)\left(\frac{k}{N} - x\right) + \frac{\tau''(x)}{2}\left(\frac{k}{N} - x\right)^2 + \frac{\tau^{(3)}(x)}{6}\left(\frac{k}{N} - x\right)^3.$$

Donc, par somme sur k,

$$B_N(\tau)(x) \leqslant \tau(x) + \tau'(x)\sum_{k=0}^{N} b_k^N(x)\left(\frac{k}{N} - x\right) + \frac{\tau''(x)}{2}\sum_{k=0}^{N} b_k^N(x)\left(\frac{k}{N} - x\right)^2 + \sum_{k=0}^{N} b_k^N(x)\frac{\tau^{(3)}(x)}{6}\left(\frac{k}{N} - x\right)^3.$$

D'après la Propriété 0.3, pour tout $x \in]0,1[$,

$$B_N(\tau)(x) \leqslant \tau(x) + \frac{\tau''(x)}{2}\frac{x(1-x)}{N} + \frac{\tau^{(3)}(x)}{6}\frac{x(1-x)(1-2x)}{N^2},$$

i.e.

$$\tau(x) - B_N(\tau)(x) \geqslant \frac{1}{N} - \frac{1}{3N^2 x(1-x)}.$$

Ainsi,

$$F_N(\tau)(x) \geqslant 1 - \frac{1}{3x(1-x)N}.$$

Alors, pour tout $x \in [\frac{1}{N}, 1 - \frac{1}{N}]$,

$$N(\tau(x) - B_N(\tau)(x)) \geqslant \left(1 - \frac{1}{3(1-\frac{1}{N})}\right)\mathbb{1}_N(x) = \left(1 - \frac{1}{3(1-\frac{1}{N})}\right)N(\tau_N(x) - B_N(\tau_N)(x)),$$

i.e.

$$\tau(x) - B_N(\tau)(x) \geqslant \left(1 - \frac{1}{3(1-\frac{1}{N})}\right)(\tau_N(x) - B_N(\tau_N)(x)).$$

En appliquant successivement l'opérateur de Bernstein à l'inégalité qui reste vraie en 0 et en 1, on a, pour tout $x \in [0,1]$ et tout $k \geqslant 0$,

$$B_N^k(\tau)(x) - B_N^{k+1}(\tau)(x) \geqslant \left(1 - \frac{1}{3(1-\frac{1}{N})}\right)(B_N^k(\tau_N)(x) - B_N^{k+1}(\tau_N)(x)),$$

et par somme sur k et passage à la limite,

$$\tau(x) \geqslant \left(1 - \frac{1}{3(1-\frac{1}{N})}\right)\tau_N(x).$$

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

$$\tau_N(x) \leqslant \frac{3(N-1)}{2N-3}\tau(x) \leqslant 3\tau(x).$$

$(\tau_N)_{N\in\mathbb{N}^*}$ est uniformément bornée par $C\|\tau\|_\infty$ sur $[0,1]$, par hypothèse. De plus, en appliquant le lemme précédent à la suite (τ_N) et τ, (τ_N) est équicontinue. En vertu du Théorème d'Ascoli, $(\tau_N)_{N\in\mathbb{N}}$ admet une sous-suite $(\tau_{\sigma(N)})$ uniformément convergente vers une certaine fonction τ^0.
Donc en particulier,

$$\tau^0(0) = \tau^0(1) = 0.$$

On considère $\langle\cdot,\cdot\rangle$ le produit scalaire usuel sur $L^2([0,1])$. On admettra le résultat suivant :

Lemme 2.2 : *On pose $\mathcal{K} := E \cap \{f \in \mathcal{C}^\infty(]0,1[,\mathbb{R}),\ f \text{ à support compact}\}$.*
Soit $(L_N)_{N\in\mathbb{N}}$ une suite d'opérateurs linéaires sur $\mathcal{K}$ telle que

$$\forall\varphi\in\mathcal{K},\ \lim_{N\to\infty} L_N(\varphi) = L(\varphi),$$

où L est un opérateur linéaire sur $\mathcal{K}$.
Soit $(f_N)_{N\in\mathbb{N}}$ une suite de E qui converge uniformément vers une fonction $f \in E$.
Alors, pour tout $\varphi\in\mathcal{K}$,

$$\lim_{N\to\infty} \langle f_N, L_N(\varphi)\rangle = \langle f, L(\varphi)\rangle.$$

On admet également que $(F_N^*)_{N\in\mathbb{N}^*}$ converge simplement vers un F^*.
On a donc, d'une part,

$$\langle F_N(\tau_N),\varphi\rangle = \langle \mathbb{1}_N,\varphi\rangle \underset{N\to\infty}{\longrightarrow} \int_0^1 \varphi(x)\,\mathrm{d}x,$$

et d'autre part,

$$\langle \tau_N, F_N^*(\varphi)\rangle \underset{N\to\infty}{\longrightarrow} \langle \tau, F^*(\varphi)\rangle = \langle F(\tau),\varphi\rangle.$$

Donc, pour tout $\varphi\in\mathcal{K}$,

$$\int_0^1 \varphi(x)\,\mathrm{d}x = \int_0^1 F(\tau^0)(x)\varphi(x)\,\mathrm{d}x.$$

donc (résultat admis de régularité elliptique de Schauder) τ^0 est $\mathcal{C}^2(]0,1[,\mathbb{R})$, et est solution de l'équation différentielle

$$\begin{cases} \forall x\in\,]0,1[,\ x(1-x)y''(x) = -2 \\ y(0) = y(1) = 0 \end{cases}$$

Or, si g est solution de l'équation différentielle, en intégrant une fois, il existe $a\in\mathbb{R}$ tel que

$$g'(x) = -2(\ln x - \ln(1-x)) + a.$$

En intégrant à nouveau l'équation, il existe $b\in\mathbb{R}$ tel que

$$g(x) = -2(x\ln x + (1-x)\ln(1-x) + ax + b.$$

En considérant les conditions aux bords,

$$a = b = 0.$$

L'équation admettant une unique solution, $\tau^0 = \tau$.
Ainsi, pour tout $x \in [0,1]$, la suite $(\tau_N(x))_{N\in\mathbb{N}^*}$ admet une unique valeur d'adhérence $\tau(x)$ dans un compact, donc $(\tau_N)_{N\in\mathbb{N}^*}$ converge simplement vers τ sur $[0,1]$. $\square$

La fonction d'entropie définie précédemment intervient dans le modèle de Wright-Fisher. Il s'agit d'un modèle qui peut, par exemple, représenter, en génétique, l'évolution d'un allèle [2] dans une population de reproduction cellulaire haploïde [3].

Avant d'introduire le modèle, énonçons quelques définitions et résultats sur les chaînes de Markov [Gal06]. Tous les résultats de la section qui suit seront admis.

2.2 Quelques éléments de probabilités

Définition 2.2 (Matrice stochastique) : *Soit E un ensemble fini ou dénombrable, muni de la tribu $\mathcal{P}(E)$. $(Q(x,y))_{x,y\in E}$ est une matrice stochastique (ou noyau de transition) si*

1. *pour tout $x, y \in E$, $0 \leqslant Q(x,y) \leqslant 1$,*
2. *pour tout $x \in E$, $\sum_{y\in E} Q(x,y) = 1$.*

De la même façon que pour les matrices classiques, on définit la puissance sur les matrices stochastiques de la façon suivante :

$$\begin{cases} Q^1 = Q \\ \forall n \in \mathbb{N},\ Q^{n+1} = \sum_{z\in E} Q^n(x,z)Q(z,y). \end{cases}$$

Définition 2.3 (Chaîne de Markov) : *Soit Q une matrice stochastique sur E. Soit $(X_n)_{n\in\mathbb{N}}$ un processus aléatoire dans E. On dit que $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition Q si, pour tout $n \in \mathbb{N}$, la loi de X_{n+1} connaissant $(X_0, X_1, \ldots, X_n)$ est $Q(X_n, y)$.*

Les chaînes de Markov ont la propriété intéressante suivante : pour prédire le futur X_{n+1} en connaissant le passé $(X_0, X_1, \ldots, X_n)$, seule l'information contenue dans X_n suffit. Il s'agit de la propriété de Markov.

✍ **Notation 3.** Soient $x, y \in E$. On note $\mathbb{P}_x$ la probabilité telle que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}_x(X_n = y) = Q^n(x,y).$$

Cette mesure de probabilité est obtenue par construction sur les chaînes de Markov canoniques.

Définition 2.4 (Classe de communication) : *Soient $x, y \in E$. On dit que l'on peut atteindre y en partant de x s'il existe $n \in \mathbb{N}$ tel que $Q^n(x,y) > 0$. Concrètement, cela signifie qu'il existe un chemin allant de x vers y.*

2. version d'un gène
3. chaque cellule ne comporte qu'un exemplaire de chaque chromosome

Si de plus, on peut atteindre x en partant de y, on dit que x et y communiquent, et l'on note $x \sim y$.
$\sim$ est une relation d'équivalence : $x \sim y$ si, et seulement si x et y sont dans la même classe dite de communication. On peut donc partitionner E en classes de communication.

Soient $T_x := \inf\{n \geqslant 1,\ X_n = x\}$ le temps mis par la chaîne pour aller à l'état x, et $N_x = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{X_n=x\}}$ le nombre de visites de l'état x.
Deux cas de figure se présentent :

- ou bien $\mathbb{P}_x(T_x < \infty) = 1$ et dans ce cas $\mathbb{P}_x(N_x = \infty) = 1$. On dit que x est un état récurrent ;
- ou bien $\mathbb{P}_x(T_x < \infty) < 1$ et dans ce cas $\mathbb{P}_x(N_x < \infty) = 1$. On dit que x est un état transitoire.

Il est possible de partitionner l'ensemble des états récurrents :

Théorème 2.4 (Classification des états) : *Soit R l'ensemble des états récurrents. Alors, il existe une partition de R,*

$$R = \bigcup_{i\in I} R_i$$

telle que

- *si x appartient à l'un des R_i, alors, $\mathbb{P}_x$-presque sûrement :*
  - $\forall y \in R_i,\ N_y = +\infty$;
  - $\forall y \in E \setminus R_i,\ N_y = 0$;
- *si $x \in E \setminus R$, alors, en notant $T = \inf\{n \geqslant 0,\ X_n \in R\}$ $\mathbb{P}_x$-presque sûrement,*
  - *ou bien $T = \infty$ et $\forall y \in E \setminus R$, $N_y < \infty$;*
  - *ou bien $T < \infty$ et il existe $j \in I$ tel que pour tout $n \geqslant T$, $X_n \in R_j$.*

Ce théorème formalise le fait suivant : en partant d'un état récurrent, on restera indéfiniment dans la classe de communication ; en partant d'un état transitoire, soit on ne visite que des états transitoires, soit on visite au moins un état récurrent et dans ce cas, on restera dans cette classe de récurrence.

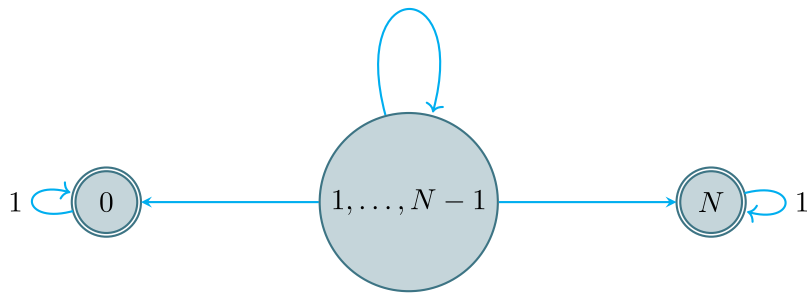
2.3 Modèle simplifié de Wright-Fisher

[Cha06][Daw06] Soit $N \in \mathbb{N}^*$. On considère une population de taille fixée N, comportant des individus qui sont soit de type A, soit de type B. Soit X_n une variable aléatoire réelle modélisant le nombre d'individus de type A à la génération n, $n \in \mathbb{N}$.

On suppose que le passage de la génération n à la génération $n+1$ se fait de la façon suivante : on effectue N tirages indépendants avec remise d'un individu dans la population de la génération n pour constituer une nouvelle génération $n+1$ comportant N individus. Cela implique que les N individus de la génération précédente meurent tous en même temps. Ainsi, $X_{n+1}|X_n$ suit une loi binomiale de paramètre $\frac{X_n}{N}$. On vérifie que $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice de transition Q donnée par

$$\forall k, \ell \in [\![0, N]\!],\ Q(k,l) = \binom{N}{\ell} \left(\frac{k}{N}\right)^{\ell} \left(1 - \frac{k}{N}\right)^{N-\ell}.$$

Remarquons que si à la génération n, aucun individu n'est de type A, aucun individu ne sera de type A dans les générations qui suivent. De même, si tous les individus de la génération n sont de type A, alors tous les individus seront de type A dans les générations qui suivent. Donc 0 et N sont des états absorbants. Les états $0, 1, \ldots, N-1$ communiquent et sont transitoires. On a donc deux états récurrents et $N-1$ états transitoires :



D'après le Théorème de classification des états, en partant de l'un des états transitoires k dans $1, 2, \ldots, N-1$, $T := \inf\{n \geqslant 0, X_n \in \{0, N\}\}$ est un temps d'arrêt ou bien fini $\mathbb{P}_k$-presque sûrement, ou bien infini $\mathbb{P}_k$-presque sûrement.

Or,

$$\mathbb{P}_k(T < \infty) \geqslant \mathbb{P}_k(T = 1) = \left(1 - \frac{k}{N}\right)^N + \left(\frac{k}{N}\right)^N > 0.$$

Donc T est finie $\mathbb{P}_k$-presque sûrement. Cela signifie qu'à un moment donné, ou bien tous les individus seront de type A, ou bien plus aucun individu ne sera de type A.

On définit le temps moyen d'extinction $t_N(k) = \mathbb{E}_k(T)$.

On établit le résultat important suivant :

Proposition 2.1 : *Pour tout $k \in [\![0, N]\!]$,*

$$t_N(k) = 1 + \sum_{\ell=0}^{N} Q(k,l) t_N(\ell),$$

autrement dit, si $T_N = \begin{pmatrix} t_N(0) \\ t_N(1) \\ \vdots \\ t_N(N) \end{pmatrix}$, *on a*

$$T_N - QT_N = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Ainsi, aux points $k \in [\![0, N]\!]$,

$$t_N(k) = N\tau_N\left(\frac{k}{N}\right).$$

Pour $x \in]0,1[$, en faisant tendre $\frac{k}{N}$ vers x lorsque $N \to \infty$, $\frac{t_N}{N}$ converge vers τ.
Donc

$$t_N(x) \underset{N\to\infty}{\sim} N\tau\left(\frac{x}{N}\right).$$

Cela signifie, en supposant que N est grand, que si à l'état initial, le nombre d'individus de type A est proche de la moitié du nombre total N, on peut approcher le temps moyen d'extinction par $2\ln(2)N \simeq 1,39N$. De même, si l'état initial est proche de 1, le temps moyen d'extinction est approché par $2\ln N$.

Concrètement, en prenant par exemple $N = k.10^9$ ($k \in [0, 10[$), le temps d'extinction du type A est de l'ordre de N lorsque la moitié de la population est de ce type, tandis qu'il sera de l'ordre de k s'il y a à l'origine un individu de type A.

3

Polynômes de Bernstein en dimension 2

Nous avons vu dans les chapitres précédents des propriétés remarquables sur les polynômes de Bernstein associés à des fonctions réelles à une variable. Dans ce chapitre, nous allons voir si ces mêmes propriétés peuvent être étendues en dimension 2 et nous intéresser à d'autres propriétés.

3.1 Définitions

Définition 3.1 : *Soit $f : [0,1]^2 \longrightarrow \mathbb{R}$ une fonction continue. On appelle polynômes de Bernstein associés à f les polynômes de la forme*

$$B_{n_1,n_2}(f)(x_1,x_2) = \sum_{k=0}^{n_1}\sum_{\ell=0}^{n_2}\binom{n_1}{k}\binom{n_2}{\ell}x_1^k(1-x_1)^{n_1-k}x_2^\ell(1-x_2)^{n_2-\ell}f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right),$$

où $n_1, n_2 \in \mathbb{N}^*$

✍ **Notation 4.** Avec les notations des chapitres précédents, la définition se réécrit

$$B_{n_1,n_2}(f)(x_1,x_2) = \sum_{k=0}^{n_1}\sum_{\ell=0}^{n_2} b_k^{n_1}(x_1)b_\ell^{n_2}(x_2)f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right).$$

Nous noterons $B_n(f)$ le polynôme $B_{n,n}(f)$.

Comme en dimension 1, on introduit les opérateurs de différence par rapport à chaque variable :

Définition 3.2 (Opérateur de différence) : *Soit $f : [0,1]^2 \longrightarrow \mathbb{R}$ une fonction continue. Soient $n_1, n_2 \in \mathbb{N}^*$ et $B_{n_1,n_2}(f)$ un polynôme de Bernstein associé à f.*
De la même manière qu'en dimension 1, on définit :

$$\Delta_1 f\left(\frac{k}{n_1},x_2\right) = \Delta f(\cdot,x_2)\left(\frac{k}{n_1}\right) = f\left(\frac{k+1}{n_1},x_2\right) - f\left(\frac{k}{n_1},x_2\right)$$

et
$$\Delta_2 f\left(x_1, \frac{\ell}{n_2}\right) = f\left(x_1, \frac{\ell+1}{n_2}\right) - f\left(x_1, \frac{\ell}{n_2}\right).$$

Propriété 3.1 : *Soient $f : [0,1] \longrightarrow \mathbb{R}$ une fonction, $n_1, n_2 \in \mathbb{N}^*$ et $B_{n_1,n_2}(f)$ un polynôme de Bernstein associé à f. Alors, pour tous $k \in [\![0, n_1-1]\!]$ et $\ell \in [\![0, n_2-1]\!]$,*
$$\Delta_1\Delta_2 f\left(\frac{k}{n_2}, \frac{\ell}{n_2}\right) = \Delta_2\Delta_1 f\left(\frac{k}{n_2}, \frac{\ell}{n_2}\right).$$

Remarque. Cette propriété ressemble au Théorème de Schwarz, ce qui n'est pas surprenant puisqu'il s'agit d'opérateurs qui se comportent comme les opérateurs standards de différenciation.

3.2 Approximation

Théorème 3.1 : *Soit f une fonction continue sur $[0,1]^2$. La suite de polynômes $(B_{n_1,n_2}(f))_{(n_1,n_2)\in\mathbb{N}^2}$ converge uniformément vers f.*

Démonstration. Soit $f : [0,1]^2 \to \mathbb{R}$ une fonction continue et non nulle.
Comme $[0,1]^2$ est compact, f y est uniformément continue, d'après le Théorème de Heine.
Soit $\varepsilon > 0$.
Alors il existe $\delta > 0$ tel que pour tout $(x_1, x_2), (y_1, y_2) \in [0,1]^2$, si $\max(|x_1 - y_1|, |x_2 - y_2|) \leqslant \delta$, alors $|f(x_1, x_2) - f(y_1, y_2)| \leqslant \frac{\varepsilon}{4}$.
Soient $(X_n)_{n\in\mathbb{N}^*}$ et $(Y_n)_{n\in\mathbb{N}^*}$ deux suites de variables indépendantes, identiquement distribuées et indépendantes entre elles, suivant respectivement les lois de Bernoulli de paramètres x_1 et x_2. On pose $S_n = X_1 + \cdots + X_n$ et $S'_n = Y_1 + \cdots + Y_n$.
De la même façon qu'en dimension 1, on a

$$\begin{aligned}
\left|B_{n_1,n_2}(f)(x_1,x_2) - f(x_1,x_2)\right| \leqslant & \mathbb{E}\left(\left|f\left(\frac{S_{n_1}}{n_1}, \frac{S'_{n_2}}{n_2}\right) - f(x_1,x_2)\right|\right) \\
= & \mathbb{E}\left(\left|f\left(\frac{S_{n_1}}{n_1}, \frac{S'_{n_2}}{n_2}\right) - f(x_1,x_2)\right| \mathbb{1}_{\{|\frac{S_{n_1}}{n_1} - x_1| \leqslant \delta\}} \mathbb{1}_{\{|\frac{S'_{n_2}}{n_2} - x_2| \leqslant \delta\}}\right) \\
& + \mathbb{E}\left(\left|f\left(\frac{S_{n_1}}{n_1}, \frac{S'_{n_2}}{n}\right) - f(x_1,x_2)\right| \mathbb{1}_{\{|\frac{S_{n_1}}{n_1} - x_1| \leqslant \delta\}} \mathbb{1}_{\{|\frac{S'_{n_2}}{n_2} - x_2| > \delta\}}\right) \\
& + \mathbb{E}\left(\left|f\left(\frac{S_{n_1}}{n_1}, \frac{S'_{n_2}}{n_2}\right) - f(x_1,x_2)\right| \mathbb{1}_{\{|\frac{S_{n_1}}{n_1} - x_1| > \delta\}} \mathbb{1}_{\{|\frac{S'_{n_2}}{n_2} - x_2| \leqslant \delta\}}\right) \\
& + \mathbb{E}\left(\left|f\left(\frac{S_{n_1}}{n_1}, \frac{S'_{n_2}}{n_2}\right) - f(x_1,x_2)\right| \mathbb{1}_{\{|\frac{S_{n_1}}{n_1} - x_1| > \delta\}} \mathbb{1}_{\{|\frac{S'_{n_2}}{n_2} - x_2| > \delta\}}\right) \\
\leqslant & \frac{\varepsilon}{4} + 2\|f\|_\infty \Bigg(\mathbb{P}(\left|\frac{S_{n_1}}{n_1} - x_1\right| \leqslant \delta)\mathbb{P}(\left|\frac{S'_{n_2}}{n_2} - x_2\right| > \delta)
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{P}(\big|\frac{S_{n_1}}{n_1}-x_1\big|>\delta)\mathbb{P}(\big|\frac{S'_{n_2}}{n_2}-x_2\big|\leqslant\delta) \\
&+\mathbb{P}(\big|\frac{S_{n_1}}{n_1}-x_1\big|>\delta)\mathbb{P}(\big|\frac{S'_{n_2}}{n_2}-x_2\big|>\delta)\Big) \text{ par indépendance} \\
\leqslant\ &\frac{\varepsilon}{4}+2\|f\|_\infty\Big(\frac{\mathbb{P}(\big|\frac{S_{n_1}}{n_1}-x_1\big|\leqslant\delta)}{4n_2\delta^2}+\frac{\mathbb{P}(\big|\frac{S'_{n_2}}{n_2}-x_1\big|\leqslant\delta)}{4n_1\delta^2} \\
&+\frac{1}{16n_1n_2\delta^4}\Big).
\end{aligned}
$$

Comme $\mathbb{P}(\big|\frac{S_n}{n}-x_1\big|\leqslant\delta)$ et $\mathbb{P}(\big|\frac{S'_n}{n}-x_1\big|\leqslant\delta)$ sont bornées, il existe $n_0\in\mathbb{N}^*$ tel que pour tout $n_1,n_2\geqslant n_0$,

$$
\frac{\mathbb{P}(\big|\frac{S_{n_1}}{n_1}-x_1\big|\leqslant\delta)}{4n_1\delta^2}\leqslant\frac{\varepsilon}{8\|f\|_\infty};\ \frac{\mathbb{P}(\big|\frac{S'_{n_2}}{n_2}-x_2\big|\leqslant\delta)}{4n_1\delta^2}\leqslant\frac{\varepsilon}{8\|f\|_\infty}\text{ et }\frac{1}{16n_1n_2\delta^4}\leqslant\frac{\varepsilon}{8\|f\|_\infty}.
$$

Donc pour tout $n_1,n_2\geqslant n_0$ avec n_0 ainsi fixé,

$$
\big|B_{n_1,n_2}(f)(x_1,x_2)-f(x_1,x_2)\big|\leqslant\varepsilon.
$$

□

3.3 Convexité

Nous avons vu en dimension 1 que les polynômes de Bernstein préservent la convexité. On peut se demander si cette propriété est toujours vraie en dimension 2.

Intéressons-nous donc à la différentielle seconde des polynômes de Bernstein $B_{n_1,n_2}(f)$ associés à une fonction continue et convexe f. $B_{n_1,n_2}(f)$ est convexe si, et seulement si sa Hessienne est symétrique positive.

Proposition 3.1 : *Soient* $n_1,n_2\in\mathbb{N}^*$. *Pour tous* $x_1,x_2\in[0,1]$,

$$
\frac{\partial^2 B_{n_1,n_2}(f)}{\partial x_1\partial x_2}(x_1,x_2)=n_1n_2\sum_{k=0}^{n_1-1}\sum_{\ell=0}^{n_2-1}b_k^{n_1-1}(x_1)b_\ell^{n_2-1}(x_2)\Delta_2\Delta_1 f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right).
$$

Démonstration. On a

$$
\begin{aligned}
\frac{\partial B_{n_1,n_2}(f)}{\partial x_1}(x_1,x_2)&=\sum_{\ell=0}^{n_2}b_\ell^{n_2}(x_2)\left(\sum_{k=0}^{n_1}(b_k^{n_1})'(x_1)f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right)\right) \\
&=n_1\sum_{\ell=0}^{n_2}b_\ell^{n_2}(x_2)\sum_{k=0}^{n_1-1}b_k^{n_1-1}(x_1)\Delta_1 f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right).
\end{aligned}
$$

De même,

$$
\frac{\partial^2 B_{n_1,n_2}(f)}{\partial x_2^2}(x_1,x_2)=n_2(n_2-1)\sum_{\ell=0}^{n_2-2}b_\ell^{n_2-2}(x_2)\sum_{k=0}^{n_1}b_k^{n_1}(x_1)\Delta_2^2 f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right).
$$

Et, en procédant de la même manière,

$$\begin{aligned}\frac{\partial^2 B_{n_1,n_2}(f)}{\partial x_2 \partial x_1}(x_1,x_2) &= n_1 \sum_{k=0}^{n_1-1} b_k^{n_1-1}\left(\sum_{\ell=0}^{n_2}(b_\ell^{n_2})'(x_2)\Delta_1 f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right)\right)\\ &= n_1 n_2 \sum_{k=0}^{n_1-1}\sum_{\ell=0}^{n_2-1} b_k^{n_1-1}(x_1) b_\ell^{n_2-1}(x_2)\Delta_2\Delta_1 f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right).\end{aligned}$$

□

Remarque. On peut également établir, par récurrence, que pour tous $p \in [\![1, n_1]\!]$ et $q \in [\![1, n_2]\!]$,

$$\frac{\partial^{p+q} B_{n_1,n_2}(f)}{\partial x_1^p \partial x_2^q}(x_1,x_2) = \frac{n_1! n_2!}{(n_1-p)!(n_2-q)!}\sum_{k=0}^{n_1-p}\sum_{\ell=0}^{n_2-q} b_k^{n_1-p}(x_1) b_\ell^{n_2-q}(x_2)\Delta_2^q\Delta_1^p f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right).$$

Ainsi, d'après la proposition, en réécrivant explicitement $\Delta_2\Delta_1 f\left(\frac{k}{n_1},\frac{l}{n_2})\right)$,

$$\begin{aligned}\frac{\partial^2 B_{n_1,n_2}(f)}{\partial x_1 \partial x_2}(x_1,x_2) = \frac{n_1! n_2!}{(n_1-1)!(n_2-1)!}\sum_{k=0}^{n_1-1}\sum_{\ell=0}^{n_2-1} b_k^{n_1-1}(x_1) b_\ell^{n_2-1}(x_2)\Bigg(f\left(\frac{k+1}{n_1},\frac{\ell+1}{n_2}\right)\\ - f\left(\frac{k}{n_1},\frac{\ell+1}{n_2}\right) - f\left(\frac{k+1}{n_1},\frac{\ell}{n_2}\right) + f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right)\Bigg).\end{aligned}$$

Réécrivons également la définition de $\partial_2\partial_1 f(x_1,x_2)$, pour $x_1, x_2 \in [0,1]$:

$$\partial_1 f(x_1,x_2) = \lim_{h\to 0}\frac{f(x_1+h,x_2)-f(x_1,x_2)}{h},$$

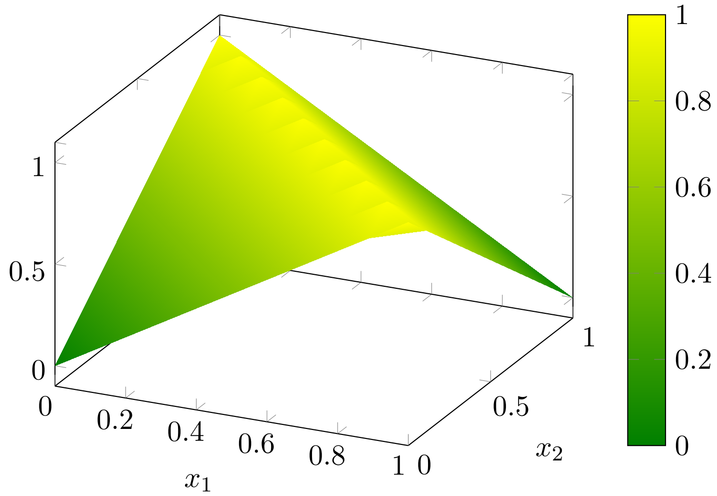
donc

$$\begin{aligned}\partial_2\partial_1(x_1,x_2) &= \lim_{k\to 0}\frac{\partial_1(x_1,x_2+k)-\partial_1(x_1,x_2)}{k}\\ &= \lim_{k\to 0}\lim_{h\to 0}\frac{f(x_1+h,x_2+k)-f(x_1,x_2+k)-f(x_1,+h,x_2)+f(x_1,x_2)}{hk}.\end{aligned}$$

On remarque alors que l'opérateur $\Delta_2\Delta_1$ agit comme une « dérivée discrète » suivant les deux directions, aux points $(\frac{k}{n_1},\frac{\ell}{n_2})$ pour $k,\ell \in [\![0, N-1]\!]$. Il est légitime de conjecturer que $\partial_2\partial_1 f \geqslant 0$ pourrait impliquer $\Delta_2\Delta_1 f\left(\frac{k}{n_1},\frac{\ell}{n_2}\right) \geqslant 0$, mais $\partial_2\partial_1 f \geqslant 0$ n'est pas une condition suffisante pour que f soit convexe. En voici un contre-exemple :

$$f : \begin{array}{ccc}[0,1]^2 & \longrightarrow & \mathbb{R}\\ (x_1,x_2) & \longmapsto & 1-|1-|x_1+x_2||\end{array}.$$

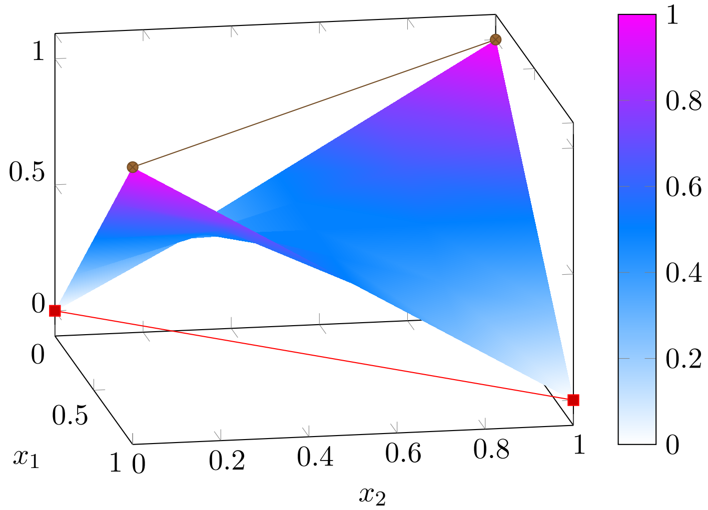
dont la représentation graphique est la figure ci-dessous :



f est une fonction concave.
Cependant,

$$B_1(f)(x_1, x_2) = x_1(1 - x_2) + x_2(1 - x_1),$$

et la Hessienne de $B_1(f)$ (qui existe puisque $B_{1,1}(f)$ est un polynôme) en tout point $x_1, x_2 \in [0, 1]$ est $\begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}$ qui a une trace nulle, et un déterminant non nul, donc n'est pas positive. Donc $B_1(f)$ n'est ni convexe ni concave : graphiquement on peut trouver par exemple une corde au-dessus de la surface représentant $B_1(f)$ et une corde en-dessous, comme on peut le voir sur la figure ci-dessous.



Ainsi, en dimension 2, les polynômes de Bernstein ne préservent pas la propriété de convexité.
On peut s'intéresser à une propriété moins forte que la convexité.
Considérons les fonctions $\mathcal{C}^2([0, 1]^2, \mathbb{R})$ $f : [0, 1]^2 \longrightarrow \mathbb{R}$ vérifiant $\partial_1\partial_2 f \leqslant 0$.
D'après le Théorème de Schwarz,

$$\partial_1\partial_2 f = \partial_2\partial_1 f.$$

Ainsi, à x_1 fixé, $x_2 \longmapsto \partial_1 f(x_1, x_2)$ est décroissante. De même, à x_2 fixé, $x_1 \longmapsto \partial_2 f(x_1, x_2)$ est décroissante.

Théorème 3.2 : *Soit f une fonction continue sur $[0,1]^2$ vérifiant $\partial_1\partial_2 f \leqslant 0$. Alors, pour tous $n_1, n_2 \in \mathbb{N}^*$, $B_{n_1,n_2}(f) \leqslant 0$.*

Démonstration. Soient $x_1, x_2 \in [0,1]$. Soient h et k deux réels positifs tels que $x_1 + h \leqslant 1$ et $x_2 + k \leqslant 1$.
D'après le Théorème fondamental de l'analyse,

$$\begin{aligned} f(x_1+h, x_2+k) - f(x_1, x_2+k) - f(x_1+h, x_2) + f(x_1, x_2) &= \int_{x_1}^{x_1+h} \partial_1 f(t, x_2+k) - \partial_1 f(t, x_2)\,\mathrm{d}t \\ &= \int_{x_1}^{x_1+h} \int_{x_2}^{x_2+k} \partial_2\partial_1 f(t,u)\,\mathrm{d}u\,\mathrm{d}t \\ &\leqslant 0 \qquad (\star) \end{aligned}$$

par positivité de l'intégrale (car $\partial_1\partial_2 f \leqslant 0$).
Or, pour tous $n_1, n_2 \in \mathbb{N}$.

$$\begin{aligned} \frac{\partial^2 B_{n_1,n_2}(f)}{\partial x_1 \partial x_2}(x_1,x_2) = \frac{n_1! n_2!}{(n_1-1)!(n_2-1)!} & \sum_{k=0}^{n_1-1}\sum_{\ell=0}^{n_2-1} b_k^{n_1-1}(x_1) b_\ell^{n_2-1}(x_2) \Bigg(f\left(\frac{k+1}{n_1}, \frac{\ell+1}{n_2}\right) \\ & - f\left(\frac{k}{n_1}, \frac{\ell+1}{n_2}\right) - f\left(\frac{k+1}{n_1}, \frac{\ell}{n_2}\right) + f\left(\frac{k}{n_1}, \frac{\ell}{n_2}\right) \Bigg). \end{aligned}$$

Et d'après ($\star$), pour tous $k \in [\![0, n_1-1]\!]$, $\ell \in [\![0, n_2-1]\!]$,

$$f\left(\frac{k+1}{n_1}, \frac{\ell+1}{n_2}\right) - f\left(\frac{k}{n_1}, \frac{\ell+1}{n_2}\right) - f\left(\frac{k+1}{n_1}, \frac{\ell}{n_2}\right) + f\left(\frac{k}{n_1}, \frac{\ell}{n_2}\right) \leqslant 0.$$

Donc $\partial_1\partial_2 B_{n_1,n_2}(f) \leqslant 0$ pour tous $n_1, n_2 \in \mathbb{N}$. □

Conclusion et perspectives

Dans ce mémoire, nous avons présenté la démonstration du Théorème de Weierstrass avec les polynômes de Bernstein. Nous avons vu que l'erreur d'approximation de degré n, pour une fonction lipschitzienne (Propriété 1.1) est en $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. Celle-ci est bien plus importante que celle des polynômes de Taylor qui sont en $\mathcal{O}\left(\frac{1}{(n+1)!}\right)$, par exemple. Comme nous l'avons vu plus loin, les polynômes de Bernstein conservent la convexité, et n'oscillent pas autour de f lorsque celle-ci est convexe (ou concave) ; cela explique en partie l'erreur d'approximation. Néanmoins, il existe des fonctions remarquables dont l'erreur d'approximation est en $\mathcal{O}\left(\frac{1}{n}\right)$: c'est le cas pour la fonction d'entropie τ (Définition 2.1), qui correspond à la limite du temps moyen d'extinction dans le modèle de Wright-Fisher, lorsqu'on fait tendre la taille de la population vers l'infini.

L'un des intérêts de ces polynômes est que nous en avons une expression explicite qui se programme facilement puisqu'elle ne nécessite que des opérations de faible complexité : addition, multiplication et évaluation en un nombre limité de points.

De plus, les polynômes de Bernstein permettent d'approcher des fonctions continues qui n'ont pas de propriétés de régularité plus poussées que la continuité.

Nous avons également vu que la convergence uniforme des polynômes de Bernstein vers une fonction f est encore vraie en dimension 2. En revanche, la convexité n'est plus conservée. Un axe de développement possible serait d'étudier l'erreur d'approximation dans ce cas, et voir si elle est moins importante qu'en dimension 1. Une autre piste envisageable serait d'établir un lien entre le processus de Wright-Fisher et les polynômes de Bernstein en dimension 2.

Bibliographie

[Ber12] Sergueï Natanovitch Bernstein. « Démonstration du théorème de Weierstrass fondé sur le calcul des probabilités ». In : Communications de la Société Mathématique de Kharkov 13 (1912). url : https://www.math.technion.ac.il/hat/fpapers/bern1.pdf.

[Cha06] Djalil Chafaï. Quelques aspects des chaînes de Markov. 2006. url : https://djalil.chafai.net/docs/biskra-markov.pdf.

[Daw06] Don Dawson. Stochastic Population Dynamics. 2006. url : http://www.math.ubc.ca/~db5d/SummerSchool09/lectures-dd/lectures6-7.pdf.

[Gal06] Jean-François Le Gall. Intégration, Probabilités et Processus Aléatoires. 2006. url : https://www.imo.universite-paris-saclay.fr/~jflegall/IPPA2.pdf.

[Bus17] Jorge Bustamante. Bernstein Operators And Their Properties. Birkhäuser, 2017, p. 91-92, 216. isbn : 978-3-319-55401-3. doi : 10.1007/978-3-319-55402-0.

Index