Théorème de Weierstrass

Références : Zuily, Queffélec, *Analyse pour l'agrégation*, p 518-519 (et p 114-115)

Théorème.

Soit $f : [0,1] \to \mathbb{C}$ continue. On lui associe ses polynômes de Bernstein définis pour tout $n \geqslant 1$ par

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

Alors pour tout $n \geqslant 1$,

$$||f - B_n||_\infty \leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right),$$

où ω est le module de continuité uniforme de f [a]. De plus, cette estimation est optimale.

[a] Pour tout $h \in \mathbb{R}$, $\omega(h) = \sup_{|x-y|\leqslant h} (|f(x) - f(y)|)$. Celui-ci tend vers 0 en 0 par uniforme continuité de f (théorème de Heine). On le définit sur $\mathbb{R}$ pour ne pas avoir de problème ensuite.

Démonstration. Commençons par fixer un $x \in [0,1]$. On considère une suite $(X_n)_n$ de variables aléatoires indépendantes de Bernoulli de paramètre x. Pour $n \geqslant 1$ on notera

$$S_n = \sum_{k=1}^{n} X_k.$$

Remarquons que l'on a, pour $n \geqslant 1$,

$$\mathbb{E}(S_n) = nx,$$
$$\mathrm{Var}(S_n) = nx(1-x),$$

et

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

Soit $n \geqslant 1$. Par le constat précédent on peut écrire

$$\begin{aligned} |f(x) - B_n(x)| &= \left|\mathbb{E}\left(f(x) - f\left(\frac{S_n}{n}\right)\right)\right| \\ &\leqslant \mathbb{E}\left(\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right) \\ &\leqslant \mathbb{E}\left(\omega\left(\left|x - \frac{S_n}{n}\right|\right)\right). \end{aligned}$$

On aura besoin du lemme suivant :

Lemme.

Pour tous $\lambda, h \in [0,1]$, on a $\omega(\lambda h) \leqslant (\lambda + 1)\omega(h)$.

Démonstration. La fonction ω est croissante. En effet, si $h \leqslant h'$ alors

$$\{|f(x) - f(y)|, |x-y| \leqslant h\} \subset \{|f(x) - f(y)|, |x-y| \leqslant h'\}$$

d'où $\omega(h) \leqslant \omega(h')$.

Ensuite, celle-ci est sous-additive :

Soient t_1 et t_2 dans $[0,1]$ tels que $t_1+t_2 \in [0,1]$. Soient u et v dans $[0,1]$ tels que $|u-v| \leqslant t_1+t_2$ et $u \leqslant v$, et soit w dans $[0,1]$ tel que $|u-w| \leqslant t_1$ et $|v-w| \leqslant t_2$. (On peut prendre $w = u+t_1$ par exemple)

Alors on a

$$\begin{aligned} |f(u)-f(v)| &\leqslant |f(u)-f(w)| + |f(w)-f(v)| \\ &\leqslant \omega(t_1) + \omega(t_2) \end{aligned}$$

d'où $\omega(t_1+t_2) \leqslant \omega(t_1)+\omega(t_2)$ en passant à la borne supérieure.
On en déduit par une récurrence immédiate que si $n \in \mathbb{N}$ et $h \in [0,1]$ sont tels que $nh \in [0,1]$ alors $\omega(nh) \leqslant n\omega(h)$.

Soit donc λ et h dans $[0,1]$. Comme $\lfloor \lambda \rfloor \leqslant \lambda \leqslant \lfloor \lambda \rfloor + 1$ on obtient

$$\begin{aligned} \omega(\lambda h) &\leqslant \omega((\lfloor \lambda \rfloor + 1)h) \\ &\leqslant (\lfloor \lambda \rfloor + 1)\omega(h) \\ &\leqslant (\lambda+1)\omega(h). \end{aligned}$$

□

Par le lemme on a

$$\omega\left(\left|x-\frac{S_n}{n}\right|\right) \leqslant \left(\sqrt{n}\left|x-\frac{S_n}{n}\right|+1\right)\omega\left(\frac{1}{\sqrt{n}}\right).$$

Ainsi, en utilisant l'inégalité de Cauchy-Schwarz ($||.||_1 \leqslant ||.||_2$ comme on a une mesure de probabilité), on obtient

$$\begin{aligned} |f(x)-B_n(x)| &\leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\mathbb{E}\left(\sqrt{n}\left|x-\frac{S_n}{n}\right|+1\right) \\ &= \omega\left(\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\left|\left|x-\frac{S_n}{n}\right|\right|_1\right) \\ &\leqslant \omega\left(\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\left|\left|x-\frac{S_n}{n}\right|\right|_2\right). \end{aligned}$$

De plus, $\mathbb{E}\left(\frac{S_n}{n}\right) = x$ donc $\left|\left|x-\frac{S_n}{n}\right|\right|_2$ n'est autre que la racine carrée de la variance de $\frac{S_n}{n}$, qui vaut $\frac{1}{n^2}\mathrm{Var}(S_n) = \frac{1}{n^2}nx(1-x) = \frac{x(1-x)}{n}$.

On a donc

$$|f(x)-B_n(x)| \leqslant \omega\left(\frac{1}{\sqrt{n}}\right)(1+\sqrt{x(1-x)}).$$

Or le polynôme $X(1-X)$ admet son maximum sur $[0,1]$ en $\frac{1}{2}$ (simple étude de fonction), ce maximum valant $\frac{1}{4}$. Finalement,

$$|f(x)-B_n(x)| \leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right),$$

et la majoration étant uniforme, on a bien

$$||f-B_n||_\infty \leqslant \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right).$$

Pour montrer que cette estimation est optimale, on va exhiber une fonction f telle que pour tout $n \geqslant 1$,

$$||f-B_n||_\infty \geqslant C\omega\left(\frac{1}{\sqrt{n}}\right),$$

où C est une constante strictement positive.

On considère la fonction $f : x \mapsto \left|x - \frac{1}{2}\right|$, continue sur $[0,1]$ et dont le module de continuité uniforme vérifie $\forall h \in [0,1]$, $\omega(h) \leqslant h$ (seconde inégalité triangulaire). En gardant les mêmes notations que précédemment, pour $x = \frac{1}{2}$ et $n \geqslant 1$, on a

$$\begin{aligned}
||f - B_n||_\infty &\geqslant \left|f\left(\frac{1}{2}\right) - B_n\left(\frac{1}{2}\right)\right| \\
&= \mathbb{E}\left(\left|\frac{1}{2} - \frac{S_n}{n}\right|\right) \\
&= \frac{1}{2n}\mathbb{E}\left(|2S_n - n|\right) \\
&= \frac{1}{2n}||\varepsilon_1 + \cdots + \varepsilon_n||_1,
\end{aligned}$$

où pour tout $i \in \{1, ..., n\}$, $\varepsilon_i = 2X_i - 1$ est une variable aléatoire de Rademacher.
On va montrer que $||\varepsilon_1 + \cdots + \varepsilon_n||_1 \geqslant \sqrt{\frac{n}{\mathrm{e}}}$ [1], ce qui prouvera que

$$||f - B_n||_\infty \geqslant \frac{1}{2\sqrt{\mathrm{e}}}\,\omega\left(\frac{1}{\sqrt{n}}\right)$$

et terminera la preuve.

Posons $f = \varepsilon_1 + \cdots + \varepsilon_n$ et $g = \prod_{j=1}^{n}\left(1 + i\frac{\varepsilon_j}{\sqrt{n}}\right)$. On a presque surement

$$|g| = \left|\prod_{j=1}^{n}\left(1 + i\frac{\varepsilon_j}{\sqrt{n}}\right)\right| = \prod_{j=1}^{n}\left(1 + \frac{\varepsilon_j^2}{n}\right)^{1/2} \leqslant \prod_{j=1}^{n}\exp\left(\frac{1}{n}\right)^{1/2} = \sqrt{\mathrm{e}}$$

car pour tout $j \in \{1, ..., n\}$, $\varepsilon_j^2 = 1$ ps et $\forall x \in \mathbb{R}, 1 + x \leqslant \exp(x)$.

De plus,

$$\begin{aligned}
|\mathbb{E}(fg)| &= \left|\sum_{j=1}^{n}\mathbb{E}\left(\varepsilon_j\prod_{k=1}^{n}\left(1 + i\frac{\varepsilon_k}{\sqrt{n}}\right)\right)\right| \\
&= \left|\sum_{j=1}^{n}\left(\prod_{\substack{k=1\\k\neq j}}^{n}\mathbb{E}\left(1 + i\frac{\varepsilon_k}{\sqrt{n}}\right) \times \mathbb{E}\left(\varepsilon_j + \frac{i}{\sqrt{n}}\right)\right)\right| \\
&= \sqrt{n}
\end{aligned}$$

car les ε_k sont centrées et indépendantes.
Finalement, on a $\sqrt{n} = |\mathbb{E}(fg)| \leqslant \mathbb{E}(|fg|) \leqslant ||f||_1||g||_\infty \leqslant \sqrt{\mathrm{e}}||f||_1$, d'où le résultat. □

Remarques : • En pratique, on ne prouve ni la croissance, ni la sous-additivité de ω pour avoir le temps de tout faire.

Adapté du travail d'Alexandre Bailleul.

1. Cette inégalité est appelée inégalité de Khintchine.