Théorème de Weierstrass par les polynômes de Bernstein

↳ Ref : ZQ, p 114-115 (lemme), p 518-519 (thm), p 246-247 (inégalité)

Lemme : Soit $f : [0,1] \to \mathbb{R}$ une fonction continue. On définit son module de continuité par $\omega_f(\delta) = \sup\{|f(u) - f(v)|,\ |u-v| \leq \delta\}$ pour $\delta > 0$.
Alors : 1) ω_f est une fonction croissante et $\lim_{h \to 0} \omega_f(h) = 0$.

(dans $]0,1[$ en fait mais on ne va pas s'embêter avec dans l'idée, δ est petit)

2) Pour $h_1, h_2 \geq 0$, $\omega_f(h_1 + h_2) \leq \omega_f(h_1) + \omega_f(h_2)$

3) Pour $h \geq 0$, $\lambda \in \mathbb{R}_+$, $\omega_f(\lambda h) \leq (\lambda + 1)\omega_f(h)$.

Théorème (Bernstein) : Soit $f : [0,1] \to \mathbb{R}$ une fonction continue. Pour $n \geq 1$, on définit le n-ième polynôme de Bernstein de f de la façon suivante :
$B_n(f, x) = B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$ pour x dans $[0,1]$.

Alors 1) B_n converge uniformément vers f sur $[0,1]$, et plus précisément, on a l'inégalité : $\|f - B_n\|_\infty \leq \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right)$.

2) L'estimation précédente est optimale : il existe une fonction lipschitzienne f telle que $\|f - B_n\|_\infty \geq \text{cste} \times \omega_f\left(\frac{1}{\sqrt{n}}\right)$

Corollaire : (Weierstrass) Soit $a < b$ deux réels et $f : [a,b] \to \mathbb{R}$ continue, alors f est limite uniforme de polynômes sur $[a,b]$. Autrement dit, $\mathbb{R}[X]$ est dense dans $(\mathcal{C}^0([a,b], \mathbb{R}), \|\cdot\|_\infty)$.

Preuve du lemme : 1) Soit $0 \leq h_1 \leq h_2$. Alors on a l'inclusion suivante : $\{|f(u) - f(v)|,\ |u-v| \leq h_1\} \subset \{|f(u) - f(v)|,\ |u-v| \leq h_2\}$, d'où $\omega(h_1) \leq \omega(h_2)$ en passant au sup. La limite en 0 provient de l'uniforme continuité de f sur $[0,1]$ (par théorème de Heine) (UC ⇒ pour $\varepsilon > 0$, $\exists \delta > 0$ tq $\forall x, y$ ... ⇒ $\forall \eta < \delta$, $\omega(\eta) < \varepsilon$ par croissance
2) Soit $0 \leq h_1, h_2$ et $u, v \in [0,1]$ tels que $|u - v| \leq h_1 + h_2$. Il existe alors $w \in [\min(u,v), \max(u,v)]$ tel que $|u - w| \leq h_1$ et $|w - v| \leq h_2$, alors
$|f(u) - f(v)| \leq |f(u) - f(w)| + |f(w) - f(v)| \leq \omega_f(h_1) + \omega_f(h_2)$, d'où le résultat en passant au sup à gauche.
3) D'après le deuxième point, on a pour $n \in \mathbb{N}^*$, $h > 0$: $\omega_f(nh) \leq n\omega(h)$.
Ainsi, si $\lambda \in \mathbb{R}_+$ et $h > 0$, on a $\lfloor \lambda \rfloor \leq \lambda < \lfloor \lambda \rfloor + 1$, donc d'après le premier point,
$\omega_f(\lambda h) \leq \omega_f((\lfloor \lambda \rfloor + 1)h) \leq (\lfloor \lambda \rfloor + 1)\omega_f(h) \leq (\lambda + 1)\omega_f(h)$.

→ Tout l'intérêt du module de continuité est d'écrire, pour $x, y \in [0,1]$, $\boxed{|f(x) - f(y)| \leq \omega(|x-y|)}$ et c'est cela que l'on va utiliser (ainsi que les propriétés précédentes) pour la suite.

Rq : Si f est L-lipschitzienne, pour tout $\delta > 0$, $\omega_f(\delta) \leq L\delta$.

Preuve du théorème : 1) Soit $x \in [0,1]$. On considère $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires iid de loi de Bernoulli de paramètre x. On définit pour $n \geq 1$, $S_n = \sum_{i=1}^{n} X_i \sim \mathcal{B}(n,x)$ et par formule de transfert, $\mathbb{E}(f(\frac{S_n}{n})) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(\frac{k}{n}) = B_n(x)$. Ainsi, $B_n(x)$ (qui est bien un polynôme en x) tend (d'après la loi des grands nombres) vers $f(x)$, mais il faudrait montrer que la convergence est uniforme, c'est ce que nous allons justifier à présent. On a, par inégalité triangulaire

$$|f(x) - B_n(x)| = |\mathbb{E}(f(x)) - \mathbb{E}(f(\tfrac{S_n}{n}))| \leq \mathbb{E}(|f(x) - f(\tfrac{S_n}{n})|)$$
$$\leq \mathbb{E}(\omega_f(|x - \tfrac{S_n}{n}|))$$

on fait apparaître $\sqrt{n}$ parce qu'on sait que c'est ce qu'il va apparaître en calculant $\|x - \frac{S_n}{n}\|_2 = \sqrt{\mathrm{Var}(\frac{S_n}{n})}$

Or d'après le lemme, $\omega_f(|x - \frac{S_n}{n}|) = \omega_f(\frac{1}{\sqrt{n}} \times \sqrt{n}|x - \frac{S_n}{n}|)$
$$\leq \omega_f(\tfrac{1}{\sqrt{n}}) \times (1 + \sqrt{n}|x - \tfrac{S_n}{n}|)$$

D'où $|f(x) - B_n(x)| \leq \omega_f(\frac{1}{\sqrt{n}}) \times (1 + \sqrt{n}\,\mathbb{E}(|x - \frac{S_n}{n}|))$
$$\leq \omega_f(\tfrac{1}{\sqrt{n}}) \times (1 + \sqrt{n} \times \sqrt{\mathbb{E}((x - \tfrac{S_n}{n})^2)})$$ par Cauchy-Schwarz

Or puisque $x = \mathbb{E}(\frac{S_n}{n})$, $\mathbb{E}((x - \frac{S_n}{n})^2) = \mathrm{Var}(\frac{S_n}{n}) = \frac{1}{n^2}\mathrm{Var}(S_n) = \frac{x(1-x)}{n}$,

donc $|f(x) - B_n(x)| \leq \omega_f(\frac{1}{\sqrt{n}}) \times (1 + \sqrt{x(1-x)})$
$$\leq \frac{3}{2}\,\omega_f(\tfrac{1}{\sqrt{n}})$$ ← pour f lipschitzienne, on obtient une vitesse en $\frac{1}{\sqrt{n}}$

Puisque la borne est uniforme en x, on obtient $\|f - B_n\|_\infty \leq \frac{3}{2}\omega_f(\frac{1}{\sqrt{n}})$ ce qui implique en particulier $\lim_{n\to+\infty} \|f - B_n\|_\infty = 0$ puisque $\omega_f(\delta) \xrightarrow[\delta\to 0]{} 0$. Le théorème de Weierstrass s'en déduit directement, quitte à composer par la fonction $\begin{array}{l} [0,1] \to [a,b] \\ x \mapsto a + (b-a)x \end{array}$ (qui est polynomiale !)

2) Pour le deuxième point, on va considérer la fonction $f : \begin{array}{l} [0,1] \to [0,\frac{1}{2}] \\ x \mapsto |x - \frac{1}{2}| \end{array}$ (il faut une fonction pas trop régulière sinon la vitesse est mieux que $\frac{1}{\sqrt{n}}$ et on symétrise pour travailler sur $x = \frac{1}{2}$ parce qu'on sait que $\|f - B_n\|_\infty$ sera atteint loin du bord puisque $B_n(0) = f(0)$ et $B_n(1) = f(1)$). Tout d'abord, f est 1-lipschitzienne, donc pour tout $h > 0$, $\omega_f(h) \leq h$. D'autre part, on a

$$\|f - B_n\|_\infty \geq |f(\tfrac{1}{2}) - B_n(\tfrac{1}{2})| = |B_n(\tfrac{1}{2})| = \mathbb{E}(|\tfrac{S_n}{n} - \tfrac{1}{2}|) = \tfrac{1}{2n}\mathbb{E}(|2S_n - n|)$$
$$= \tfrac{1}{2n}\mathbb{E}(|\varepsilon_1 + \dots + \varepsilon_n|)$$

où, puisque les $(X_i)_{i\in\mathbb{N}}$ suivent une loi de Bernoulli de paramètre $\frac{1}{2}$, les $\varepsilon_i := 2X_i - 1$ sont iid de loi de Rademacher.

Or d'après l'inégalité de Khintchine (A), $\mathbb{E}(|\varepsilon_1+\dots+\varepsilon_n|) \geq \frac{1}{\sqrt{e}}\sqrt{\mathbb{E}((\varepsilon_1+\dots+\varepsilon_n)^2)} = \sqrt{\frac{n}{e}}$

Ainsi, $\|f-B_n\|_\infty \geq \frac{1}{2n}\times\sqrt{\frac{n}{e}} = \frac{1}{2\sqrt{ne}} \geq \frac{1}{2\sqrt{e}}\,\omega_f\left(\frac{1}{\sqrt{n}}\right)$.

Ainsi, la vitesse en $\omega\left(\frac{1}{\sqrt{n}}\right)$ est optimale.

(A) On veut montrer que $\mathbb{E}(|T_n|)\geq\sqrt{\frac{n}{e}}$, où $T_n=\sum_{i=1}^n \varepsilon_i$. Pour cela, on pose $Y=\prod_{j=1}^n\left(1+i\frac{\varepsilon_j}{\sqrt{n}}\right)$ (ce Y en particulier, c'est de la sorcellerie, mais l'idée derrière c'est de minorer $\|T_n\|_1$ par "dualité" en écrivant $|\mathbb{E}(T_nY)|\leq\|T_n\|_1\|Y\|_\infty$ et ensuite $\|T_n\|_1\geq\frac{|\mathbb{E}(T_nY)|}{\|Y\|_\infty}$, et Y marche bien pour avoir $\|Y\|_\infty$ petit et $|\mathbb{E}(T_nY)|$ grand (magie noire)).

Alors Y est bornée presque sûrement :

$$|Y|=\prod_{j=1}^n\sqrt{1+\frac{\varepsilon_j^2}{n}}=\prod_{j=1}^n\sqrt{1+\frac{1}{n}}\leq\prod_{j=1}^n\sqrt{e^{1/n}}=\sqrt{e} \quad (\text{car } 1+x\leq e^x)$$

Donc $|\mathbb{E}(T_nY)|\leq\mathbb{E}(|T_nY|)\leq\sqrt{e}\,\mathbb{E}(|T_n|)$. Et d'autre part,

$$\mathbb{E}(T_nY)=\mathbb{E}\left(\sum_{i=1}^n\varepsilon_i\times\prod_{j=1}^n\left(1+i\frac{\varepsilon_j}{\sqrt{n}}\right)\right)=\sum_{i=1}^n\mathbb{E}\left(\varepsilon_i\times\left(1+\frac{i\varepsilon_i}{\sqrt{n}}\right)\times\prod_{j\neq i}\left(1+i\frac{\varepsilon_j}{\sqrt{n}}\right)\right)$$

$$\underset{\perp\!\!\!\perp \text{ des } \varepsilon_i}{=}\sum_{i=1}^n\left(\underbrace{\mathbb{E}(\varepsilon_i)}_{=0}+\frac{i}{\sqrt{n}}\underbrace{\mathbb{E}(\varepsilon_i^2)}_{=1}\right)\times\prod_{j\neq i}\left(1+i\frac{\overbrace{\mathbb{E}(\varepsilon_j)}^{=0}}{\sqrt{n}}\right)=\sum_{i=1}^n\frac{i}{\sqrt{n}}=i\sqrt{n}.$$

Donc $\sqrt{n}=|\mathbb{E}(T_nY)|\leq\sqrt{e}\,\mathbb{E}(|T_n|)$, d'où $\mathbb{E}(|T_n|)\geq\sqrt{\frac{n}{e}}$.

(on appelle Y produit de Riesz)

Remarques : * Le théorème n'est plus vrai sur $\mathbb{R}$ tout entier : Soit f limite uniforme d'une suite de polynômes $(P_n)_{n\in\mathbb{N}}$ sur $\mathbb{R}$. En particulier, puisque $(P_n)_{n\in\mathbb{N}}$ converge pour $\|\cdot\|_\infty$, elle est de Cauchy, donc $\exists N_0\in\mathbb{N}$, $\forall n,p\geq N_0$, $\|P_n-P_p\|_\infty\leq 1$ donc $P_n-P_p=\text{cste}$. En particulier, $P_n-P_{N_0}=C$, et donc en faisant tendre $[n\to+\infty]$, on obtient $f=C+P_{N_0}\in\mathbb{R}[X]$.

* Autre preuve du théorème par la convolution : pour $n\geq1$, on définit $P_n: t\mapsto\begin{cases}\frac{(1-t^2)^n}{a_n} & \text{si } |t|\leq1\\ 0 & \text{sinon}\end{cases}$ où $a_n=\int_{-1}^1(1-t^2)^n\,dt$, alors P_n est une approximation de l'unité, et si $f:[-\frac12,\frac12]\to\mathbb{R}$ continue, $f*P_n$ est un polynôme qui CVU vers f.

* On peut obtenir de meilleures vitesses de convergence dans certains cas particuliers : par exemple si f est dans $C^2([0,1],\mathbb{R})$ avec $\|f''\|_\infty\leq M$, on peut obtenir $\|B_n-f\|_\infty\leq\frac{M}{8n}$, on peut même obtenir une sorte de TCL si f'' est L-lipschitzienne : $\left\|n(B_n-f)(x)-\frac{f''(x)x(1-x)}{2}\right\|_\infty\leq\frac{3^{3/4}}{6}\frac{L}{\sqrt{n}}$ (ça devient dur !)

En revanche, on a vu que $\omega\left(\frac{1}{\sqrt{n}}\right)$ est optimal dans le cas général.

* Ce théorème est un cas particulier du théorème de Stone-Weierstrass : Soit X un espace compact et $C(X)$ l'algèbre de Banach des fonctions continues de X dans $\mathbb{R}$. Une sous algèbre est dense dans $C(X)$ ssi elle sépare les points et contient pour tout x dans X une fonction non nulle en x.

ie $\forall x \neq y \in X$, $\exists f \in A$, $f(x) \neq f(y)$.

Pour les leçons + analyse, faire le lemme + le théorème et laisser tomber l'inégalité si manque de temps, pour les leçons de proba, mettre sous la forme lemme sur ω puis thm 1) puis lemme inégalité Khintchine puis 2), et zapper le lemme sur ω (et le faire éventuellement à la fin s'il reste du temps).