Théorème de Weierstrass

Référence : [QZ06] p.518-519.

Théorème 0.1 *Soit $f : [0,1] \longrightarrow \mathbb{C}$ une fonction continue. Soit ω son module de continuité uniforme, ie :*

$$\omega(h) = \sup\{|f(u) - f(v)|; |u - v| \leq h\}$$

Pour $n \geq 1$, on considère le polynôme :

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

le n-ième polynôme de Bernstein de f. Alors :

1. *B_n converge vers f uniformément sur $[0,1]$;*
2. *Plus précisément, on a :*
$$\|f - B_n\|_\infty \leq C\omega\left(\frac{1}{\sqrt{n}}\right)$$
où C est une constante numérique ;
3. *L'estimation (2) est optimale, ie il existe une fonction lipschitzienne f pour laquelle :*
$$\|f - B_n\| \geq \frac{\delta}{\sqrt{n}}$$
où δ est une constante numérique.

Démonstration

Première assertion

$\rightsquigarrow$ Soit $x \in [0,1]$. Soit X une loi de Bernoulli de paramètre x, ie $X \sim \mathcal{B}(1,x)$. Soient $X_1, \ldots, X_n$ un échantillon de X. Soit $S_n = X_1 + \ldots + X_n$.
On remarque tout d'abord que :

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

car S_n suit une loi binomiale de paramètres n et x, ie $S_n \sim \mathcal{B}(n,x)$, car elle est la somme de n variables de Bernoulli de paramètre x.
Or d'après la loi faible des grands nombres, on sait que $\frac{S_n}{n}$ converge en probabilités vers $\mathbb{E}(X) = x$, donc :

$$\mathbb{E}(f(x)) = f(x) = \lim_{n \longrightarrow +\infty} \mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right) = \lim_{n \longrightarrow +\infty} B_n(x)$$

$\rightsquigarrow$ Nous allons maintenant préciser cette convergence.
Fixons $\delta \in]0,1[$ et désignons par $\|.\|_\infty$ la norme sup sur $[0,1]$. On a :

$$f(x) - B_n(x) = \mathbb{E}\left(f(x) - f\left(\frac{S_n}{n}\right)\right)$$

D'où :

$$|f(x) - B_n(x)| \leq \mathbb{E}\left(\left|f(x) - f\left(\frac{S_n}{n}\right)\right|\right) \leq \mathbb{E}\left(\omega\left(\left|x - \frac{S_n}{n}\right|\right)\right)$$

par définition du module de continuité ω.
De plus, d'après une propriété des modules de continuité, on sait que si $h, \lambda h \in [0,1]$, on a :

$$\omega(\lambda h) \leq (1+\lambda)\omega(h)$$

En particulier, cela nous donne :

$$\begin{aligned}\omega\left(\left|x-\frac{S_n}{n}\right|\right) &= \omega\left(\frac{1}{\sqrt{n}}\left|\sqrt{n}x-\frac{S_n}{\sqrt{n}}\right|\right)\\ &\leq \left(\sqrt{n}\left|x-\frac{S_n}{n}\right|+1\right)\omega\left(\frac{1}{\sqrt{n}}\right)\end{aligned}$$

car comme $n \geq 1$, alors $\frac{1}{\sqrt{n}} \in [0,1]$ et comme $x \in [0,1]$ et $S_n \in [0,n]$, alors $\frac{S_n}{n} \in [0,1]$ et donc $\left|x-\frac{S_n}{n}\right| \in [0,1]$.
Il en résulte que :

$$\begin{aligned}|f(x)-B_n(x)| &\leq \mathbb{E}\left(\omega\left(\left|x-\frac{S_n}{n}\right|\right)\right)\\ &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\mathbb{E}\left(\sqrt{n}\left(\left|x-\frac{S_n}{n}\right|+1\right)\right)\\ &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\left\|x-\frac{S_n}{n}\right\|_1\right) \quad \text{par définition de la norme } \|.\|_1\,;\\ &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\left\|x-\frac{S_n}{n}\right\|_2\right) \quad \text{par définition de } \|.\|_2 \text{ et l'inégalité de Cauchy-Schwarz}\,;\\ &= \omega\left(\frac{1}{\sqrt{n}}\right)(1+\sqrt{n})\\ &= \omega\left(\frac{1}{\sqrt{n}}\right)\left(1+\sqrt{n}\sqrt{\frac{x(1-x)}{n}}\right)\\ &\leq \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right)\end{aligned}$$

car :

$$\begin{aligned}\left\|x-\frac{S_n}{n}\right\|_2 &= \left(\mathbb{E}\left(\left|x-\frac{S_n}{n}\right|^2\right)\right)^{\frac{1}{2}}\\ &= \left(\sum_{k=0}^{n} C_n^k x^k(1-x)^{n-k}\left|x-\frac{k}{n}\right|^2\right)^{\frac{1}{2}}\\ &= \left(\sum_{k=0}^{n} C_n^k x^k(1-x)^{n-k}\left(x^2+\frac{k^2}{n^2}-2x\frac{k}{n}\right)\right)^{\frac{1}{2}}\\ &= \left(x^2\sum_{k=0}^{n} C_n^k x^k(1-x)^{n-k}+\frac{1}{n^2}\sum_{k=0}^{n} C_n^k x^k(1-x)^{n-k}k^2-\frac{2x}{n}\sum_{k=0}^{n} C_n^k x^k(1-x)^{n-k}k\right)^{\frac{1}{2}}\\ &= \left(x^2(x+(1-x))+\frac{1}{n^2}\mathbb{E}(S_n^2)-\frac{2x}{n}\mathbb{E}(S_n)\right)^{\frac{1}{2}}\end{aligned}$$

Effectuons quelques calculs intermédiaires :

$$\mathbb{E}(S_n) = \mathbb{E}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n}\mathbb{E}(X_i) = \sum_{i=1}^{n} x = nx$$

$$Var(S_n) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} x(1-x) = nx(1-x)$$

$$\mathbb{E}(S_n^2) = Var(S_n) + \mathbb{E}(S_n)^2 = nx(1-x) + n^2x^2$$

Ainsi, on trouve :

$$\left\| x - \frac{S_n}{n} \right\|_2 = \left(x^2 + \frac{x(1-x)}{n} + x^2 - 2x^2 \right)^{\frac{1}{2}} = \sqrt{\frac{x(1-x)}{n}}$$

De plus, comme $x \in [0,1]$, alors $x(1-x) \leq \frac{1}{4}$ d'où $1 + \sqrt{x(1-x)} \leq 1 + \frac{1}{2} = \frac{3}{2}$.
Par conséquent :

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

D'où l'assertion (1) avec $C = \frac{3}{2}$.

Deuxième assertion

Considérons la fonction $f(x) = \left| x - \frac{1}{2} \right|$, alors on observe que :

$$|f(u) - f(v)| = \left| \left| u - \frac{1}{2} \right| - \left| v - \frac{1}{2} \right| \right| \leq |u - v| \quad \text{d'après l'inégalité de Minkowski ;}$$

donc par définition de ω, on a $\omega(h) \leq h$.
On a également :

$$\begin{aligned}
\|f - B_n\|_\infty &\geq \left| f\left(\frac{1}{2}\right) - B_n\left(\frac{1}{2}\right) \right| \quad \text{par définition de } \|.\|_\infty \text{ ;} \\
&= \left| B_n\left(\frac{1}{2}\right) \right| \quad \text{par définition de } f \text{ ;} \\
&= \mathbb{E}\left| \frac{S_n}{n} - \frac{1}{2} \right| \quad \text{par définition de } B_n \text{ ;} \\
&= \frac{1}{2n}\mathbb{E}|2S_n - n| \\
&= \frac{1}{2n}\mathbb{E}|\epsilon_1 + \ldots + \epsilon_n|
\end{aligned}$$

où $S_n = \delta_1 + \ldots + \delta_n$ avec les (δ_j) une suite de variables aléatoires indépendantes et identiquement distribuées selon une variable de Bernoulli de paramètre $\frac{1}{2}$ et $(\epsilon_j)_j$ avec $\epsilon_j = 2\delta_j - 1$ étant donc une suite de variables indépendantes et identiquement distribuées selon une variable de Rademacher.
Ainsi d'après l'inégalité de Kintchine, on obtient :

$$\|f - B_n\|_\infty \geq \frac{1}{2n}\|\epsilon_1 + \ldots + \epsilon_n\|_1 \geq \frac{1}{2n\sqrt{e}}\|\epsilon_1 + \ldots + \epsilon_n\|_2 = \frac{1}{2\sqrt{e}\sqrt{n}} \geq \frac{1}{2\sqrt{e}}\omega\left(\frac{1}{\sqrt{n}}\right)$$

car :

$$\|\epsilon_1 + \ldots + \epsilon_n\|_2 = \mathbb{E}(|\epsilon_1 + \ldots + \epsilon_n|^2)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} \mathbb{E}(\epsilon_i^2) \right)^{\frac{1}{2}} = \sqrt{n}$$

car $\mathbb{E}(\epsilon_i) = 0$.
On ne peut donc pas améliorer la vitesse de convergence en $\omega\left(\frac{1}{\sqrt{n}}\right)$ de $B_n(f)$ vers f.

Lemmes utilisés

Définition 0.1 (Échantillon) *Un échantillon d'une variable aléatoire X est une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes de même loi que X (on dit aussi que les X_n sont i.i.d.).*

Définition 0.2 (Rademacher) *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. Soit X une variable aléatoire. On dit que X est une variable de Rademacher si elle vaut $+1$ et -1 avec une probabilité de $\frac{1}{2}$, ie $X \sim \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$.*

Proposition 0.2 (Inégalité de Kintchine) *Soit $(\Omega, \mathbb{A}, \mathbb{P})$ un espace de probabilité. Soient $r_1, \ldots, r_n$ n variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$ de Rademacher. Soit F_n un $\mathbb{R}$-espace vectoriel de dimension n engendré par $r_1, \ldots, r_n$. On définit deux normes sur F_n :*

$$\|f\|_1 = \mathbb{E}(|f|) = \int_\Omega |f(\omega)| d\mathbb{P}(\omega)$$

$$\|f\|_2 = \mathbb{E}(|f|^2)^{\frac{1}{2}} = \left(\int_\Omega |f(\omega)|^2 d\mathbb{P}(\omega)\right)^{\frac{1}{2}}$$

Alors :

$$\|f\|_1 \leq \|f\|_2 \leq \sqrt{e}\|f\|_1$$

Démonstration $\rightsquigarrow$ L'inégalité de gauche est évidente, elle découle de l'inégalité de Cauchy-Schwarz et elle est même optimale (prendre $f = r_1$).
$\rightsquigarrow$ Pour l'inégalité de droite, posons :

$$f = \sum_{i=1}^n a_i r_i \in F_n$$

Par homogénéité, on peut supposer que $\|f\|_2 = 1$, ie $\sum_{i=1}^n a_i^2 = 1$ (car $r_1, \ldots, r_n$ forment un système orthonormé de $(\Omega, \mathcal{A}, \mathbb{P})$ car $|r_j| = 1$ et $\mathbb{E}(r_i r_j) = \mathbb{E}(r_i)\mathbb{E}(r_j) = 0$ si $i \neq j$).
On va maintenant minorer $\|f\|_1$ par dualité, ie on cherche $g \in L^\infty$, car ainsi on pourra écrire :

$$|\mathbb{E}(fg)| \leq \|f\|_1 \|g\|_\infty$$

donc si $g \neq 0$, on aura :

$$\|f\|_1 \geq \frac{|\mathbb{E}(fg)|}{\|g\|_\infty}$$

Considérons :

$$g = \prod_{j=1}^n (1 + i a_j r_j)$$

Pour presque tout ω, on a :

$$\begin{aligned}
|g(\omega)| &= \prod_{j=1}^n \sqrt{1 + a_j^2 r_j^2(\omega)} \\
&= \prod_{j=1}^n \sqrt{1 + a_j^2} \quad \text{car } r_j(\omega) \text{ vaut } \pm 1\,; \\
&\leq \prod_{j=1}^n \sqrt{\exp(a_j^2)} \quad \text{car } 1 + a \leq e^a\,; \\
&= \exp\left(\frac{1}{2}\sum_{j=1}^n a_j^2\right) = \sqrt{e}
\end{aligned}$$

Donc $\|g\|_\infty \leq \sqrt{e}$.
D'autre part, si $j \in \{1, \ldots, n\}$, l'indépendance des r_j entraîne :

$$\begin{aligned}
\mathbb{E}(r_j g) &= \mathbb{E}(r_j(1 + i a_j r_j) \prod_{k \neq j} (1 + i a_k r_k)) \\
&= \mathbb{E}(r_j(1 + i a_j r_j))\mathbb{E}(\prod_{k \neq j} (1 + i a_k r_k)) \\
&= \mathbb{E}(r_j + i a_j r_j^2)) \prod_{k \neq j} \mathbb{E}(1 + i a_k r_k) \\
&= i a_j \prod_{k \neq j} 1 \quad \text{car } \mathbb{E}(r_j) = 0\,; \\
&= i a_j
\end{aligned}$$

Donc puisque l'espérance est linéaire, on a :

$$|\mathbb{E}(fg)| = |\sum_{j=1}^{n} a_j \mathbb{E}(r_j g)| = |i\sum_{j=1}^{n} a_j^2| = 1$$

Ainsi :

$$\|f\|_1 \geq \frac{1}{\sqrt{e}}$$

D'où le résultat.

Définition 0.3 (Module de continuité) *Un module de continuité est une application non nulle $\phi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ telle que :*

1. *$\phi(0) = 0$;*
2. *ϕ est continue et croissante ;*
3. *ϕ est sous-additive, ie $\forall t_1, t_2 \in \mathbb{R}_+$:*

$$\phi(t_1 + t_2) \leq \phi(t_1) + \phi(t_2)$$

Proposition 0.3 *Soit ϕ un module de continuité. Alors :*

1. *$\forall r \in \mathbb{N}$ et $\forall t \in \mathbb{R}_+$, $\phi(rt) \leq r\phi(t)$;*
2. *$\forall \lambda \in \mathbb{R}_+$ et $\forall t \in \mathbb{R}_+$, $\phi(\lambda t) \leq (\lambda + 1)\phi(t)$.*

Démonstration $\rightsquigarrow$ (1) Par récurrence sur $r \in \mathbb{N}$ et en utilisant la sous-additivité, on obtient le résultat, en effet :

- Si $r = 0$, alors on a bien $\phi(0t) = 0 \leq 0\phi(t)$ pour tout $t \in \mathbb{R}_+$;
- Supposons que la relation soit vraie au rang r. On a :

$$\phi((r+1)t) = \phi(rt + t) \leq \phi(rt) + \phi(t) \leq r\phi(t) + \phi(t) = (r+1)\phi(t)$$

D'où le résultat.

$\rightsquigarrow$ (2) Soit q l'entier positif tel que $q \leq \lambda < q+1$. La croissance de ϕ et l'assertion (1) nous donne :

$$\phi(\lambda t) \leq \phi((q+1)t) \leq (q+1)\phi(t) \leq (\lambda+1)\phi(t)$$

Théorème 0.4 (Loi faible des grands nombres) *Soit $X \in L^2$. Soit $(X_n)_{n\geq 1}$ un échantillon de X. Soit $S_n = X_1 + \ldots + X_n$. Alors pour tout $t > 0$:*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}(X)\right| > t\right) \leq \frac{Var(X)}{nt^2}$$

En particulier $\frac{S_n}{n}$ converge en probabilité vers la constante $\mathbb{E}(X)$.

Démonstration D'après l'inégalité de Tchebychev, on a :

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}(X)\right| > t\right) = \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| > t\right) \leq \frac{1}{t^2} Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2 t^2}\sum_{i=1}^{n} Var(X_i) = \frac{Var(X)}{nt^2}$$

Remarque : Pour une variable aléatoire X de Bernoulli de paramètre p, on a : $\mathbb{E}(X) = p$ et $Var(X) = p(1-p)$.

Proposition 0.5 (Inégalité de Tchebychev) *Soit $X \in L^2$, alors pour tout $t > 0$:*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{Var(X)}{t^2}$$

Démonstration On utilise l'inégalité de Markov :

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) = \mathbb{P}((X - \mathbb{E}(X))^2 \geq t^2) \leq \frac{1}{t^2}\mathbb{E}((X - \mathbb{E}(X))^2) = \frac{Var(X)}{t^2}$$

Proposition 0.6 (Inégalité de Markov) *Soit $X \in L^1$. Alors pour tout $t > 0$:*

$$\mathbb{P}(|X| \geq t) \leq \frac{\mathbb{E}(X)}{t}$$

Démonstration On observe que, en notant $\mathbb{I}$ la fonction indicatrice :

$$\mathbb{I}_{[t,+\infty[}(X) \leq \frac{X}{t}\mathbb{I}_{[t,+\infty[}(X) \leq \frac{X^+}{t} \leq \frac{|X|}{t}$$

puis on intègre par rapport à $\mathbb{P}$ et on obtient le résultat (car $X = X^+ - X^-$ et $|X| = X^+ + X^-$).

Références

[QZ06] Hervé Queffelec and Claude Zuily. *Analyse pour l'agrégation.* Dunod, 2006.