

Chapitre II : approximation polynomiale

Introduction

Soit I un intervalle de $\mathbb{R}$ et $f: I \rightarrow \mathbb{R}$ une fonction. L'approximation polynomiale consiste à approcher la fonction f par une fonction polynôme P . Sur le plan numérique, il est en effet bien plus simple de manipuler des fonctions polynômes que des fonctions quelconques. Plusieurs des méthodes que nous introduisons pour calculer de manière approchée des intégrales ou des solutions d'équations différentielles ordinaires reposent sur ce constat.

Dans ce chapitre, nous étudions deux moyens d'approcher la fonction f par une fonction polynôme P :

- l'interpolation polynomiale qui cherche à déterminer un polynôme P de degré le plus bas possible qui prend les mêmes valeurs que la fonction f en un nombre fini de points où sa valeur est connue,
- l'approximation polynomiale proprement dite qui cherche à déterminer un polynôme P le plus proche possible de la fonction f pour une norme associée à une classe de fonctions à laquelle appartient la fonction f .

Nous nous intéresserons d'abord à la méthode d'interpolation polynomiale de Lagrange. Étant donné $n+1$ points $(x_0, \dots, x_n) \in \mathbb{R}^n$, il s'agit de déterminer un polynôme $P \in \Pi_n(\mathbb{R})$ tel que:

$$\forall 0 \leq i \leq n, P(x_i) = f(x_i).$$

Nous étudierons l'existence et l'unicité d'un tel polynôme, détaillerons une manière efficace de le construire, puis nous analyserons l'erreur commise en remplaçant la fonction f par ce polynôme, lorsque cette fonction est suffisamment régulière.

On étudiera ensuite la notion de meilleure approximation polynomiale. Étant donnée une classe de fonctions E , typiquement l'espace des fonctions continues sur I , $C^0(I, \mathbb{R})$, et une norme $\|\cdot\|_E$ sur cette classe de fonctions, il s'agit de déterminer pour un degré $N \geq 0$ fixé le polynôme P_N qui minimise la distance à la fonction $f \in E$ parmi tous les polynômes de degré inférieur ou égal à N de $\mathbb{R}_N[X]$:

$$\|f - P_N\|_E = \min_{Q \in \mathbb{R}_N[X]} \|f - Q\|_E.$$

Nous établirons l'existence de ce polynôme de meilleure approximation, puis nous concentrerons sur le cas de la norme uniforme en relation avec le théorème de Weierstrass de densité des fonctions polynomiales dans l'espace $C^0([a, b], \mathbb{R})$ muni de la norme uniforme, et enfin, sur le cas des normes hilbertiennes en relation avec l'étude des familles de polynômes orthogonaux.

I Méthode d'interpolation de Lagrange

Pour $N \in \mathbb{N}$, nous introduisons l'espace des polynômes de degrés inférieurs ou égaux à N :

$$\mathbb{R}_N[X] = \{P \in \mathbb{R}[X] \text{ t.q. } \deg(P) \leq N\}.$$

Étant donnée une fonction $f: I \rightarrow \mathbb{R}$ et $N+1$ points $x_0, \dots, x_N$ de I , deux à deux distincts, le problème d'interpolation de Lagrange consiste à déterminer un polynôme $P \in \mathbb{R}_N[X]$ tel que:

$$\forall 0 \leq i \leq N, P(x_i) = f(x_i).$$

1. Existence et unicité du polynôme d'interpolation de Lagrange

Thm: Considérons une fonction $f: I \rightarrow \mathbb{R}$ et $N+1$ points $(x_0, \dots, x_N) \in I^{N+1}$ tels que:

$$\forall 0 \leq i \neq j \leq N, x_i \neq x_j.$$

Il existe un unique polynôme $P_N \in \mathbb{R}_N[X]$ tel que:

$$\forall 0 \leq i \leq N, P_N(x_i) = f(x_i).$$

De plus, ce polynôme est donné par l'expression:

$$P_N(X) = \sum_{i=0}^N f(x_i) l_i(X),$$

où les polynômes de base de Lagrange $(l_0(X), \dots, l_N(X)) \in \mathbb{R}_N(X)^{N+1}$ sont définies par:

$$\forall 0 \leq i \leq N, l_i(X) = \prod_{j \neq i} \frac{(X - x_j)}{(x_i - x_j)}.$$

Le polynôme P_N est appelé polynôme d'interpolation de Lagrange de la fonction f aux points $x_0, \dots, x_N$.

Dm: (i) Existence: Elle découle de la formule ci-dessus dans laquelle les polynômes $l_0, \dots, l_N$ sont de degré inférieur ou égal à N , et satisfont: $\forall 0 \leq i, j \leq N, l_i(x_j) = \delta_{ij}$.

(ii) Unicité: Supposons que P et Q sont solutions du problème. Le polynôme $P - Q$ a un degré inférieur ou égal à N et possède au moins $N+1$ racines distinctes $x_0, \dots, x_N$. Il est donc nul, d'où l'unicité de P_N .

Ex: (i) Pour $N=0$, le polynôme d'interpolation de Lagrange l_0 est égal à:

$$l_0(X) = 1$$

(ii) Pour $N=1$, le polynôme d'interpolation de Lagrange l_0 est égal à:

$$l_0(X) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} X + \frac{f(x_0)x_1 - x_0 f(x_1)}{x_1 - x_0}.$$

Prop: (i) Les polynômes de base de Lagrange $(l_0(X), \dots, l_N(X))$ forment une base de $\mathbb{R}_N(X)$.

Cette famille est libre, car s'il existe des nombres $(\lambda_0, \dots, \lambda_N) \in \mathbb{R}^{N+1}$ tels que $\sum_{i=0}^N \lambda_i l_i(X) = 0$, alors: $\forall 0 \leq j \leq N, \lambda_j = \sum_{i=0}^N \lambda_i l_i(x_j) = 0$. Comme $\mathbb{R}_N(X)$ est de dimension finie égale à $N+1$ et $\text{Card} \{l_0(X), \dots, l_N(X)\} = N+1$, il s'agit bien d'une base de $\mathbb{R}_N(X)$.

(ii) La méthode d'interpolation de Hermite est une généralisation de la méthode d'interpolation de Lagrange qui prend en compte les dérivées de la fonction f . Choisissons $N+1$ points distincts $(x_0, \dots, x_N) \in I^{N+1}$ et $N+1$ entiers $(d_0, \dots, d_N) \in \mathbb{N}^{N+1}$, et notons $D = N + d_0 + \dots + d_N$ et $d = \max_{0 \leq i \leq N} d_i$. Étant donnée une fonction f de classe $\mathcal{C}^d$ sur I , il existe un unique polynôme $P_D \in \mathbb{R}_D(X)$ tel que:

$$\forall 0 \leq i \leq N, \forall 0 \leq h \leq d_i, P_D^{(h)}(x_i) = f^{(h)}(x_i).$$

Le polynôme P_D est appelé polynôme d'interpolation de Hermite de la fonction f aux points $x_0, \dots, x_N$ et aux entiers $d_0, \dots, d_N$.

2. Méthode de construction des différences divisées

Considérons une fonction $f: I \rightarrow \mathbb{R}$ et $N+1$ points $(x_0, \dots, x_N) \in I^{N+1}$, deux à deux distincts.

Def: La différence divisée d'ordre N de la fonction f aux points $(x_0, \dots, x_N)$ est le coefficient noté $f[x_0, \dots, x_N]$ du monôme x^N du polynôme d'interpolation de Lagrange L_N de la fonction f aux points $(x_0, \dots, x_N)$.

Ex: (i) $f[x_0] = f(x_0)$

$$(ii) f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

Thm: Les différences divisées successives de la fonction f aux points $(x_0, \dots, x_N)$ sont données par les formules:

$$f[x_0] = f(x_0), \text{ et } \forall 1 \leq j \leq N, f[x_0, \dots, x_j] = \frac{f[x_0, \dots, x_{j-1}] - f[x_0, \dots, x_{j-2}]}{x_j - x_0}$$

Dém: Les formules est vraie pour $j=0$ par définition de $f[x_0]$. Pour $1 \leq j \leq N$, considérons les polynômes d'interpolation de Lagrange de f , $L_{j-1}^{(i)}$ et $L_{j-2}^{(i)}$, aux points $(x_0, \dots, x_{j-2})$, respectivement $(x_0, \dots, x_{j-1})$. Les polynômes

$$R(X) = \frac{1}{x_j - x_0} [(X - x_0) L_{j-2}^{(i)}(X) - (X - x_j) L_{j-1}^{(i)}(X)]$$

a un degré au plus égal à j , est satisfait:

$$R(x_0) = L_{j-2}^{(i)}(x_0) = f(x_0), \quad \forall 1 \leq 2 \leq j-1, \quad R(x_0) = f(x_0) \text{ et } R(x_j) = L_{j-1}^{(i)}(x_j) = f(x_j)$$

D'où $R(X) = L_j(X)$, et l'identification du terme de plus haut degré conduit à la formule du théorème.

Formule de Newton: Le polynôme d'interpolation de Lagrange L_N de la fonction f aux points $(x_0, \dots, x_N)$ est égal à:

$$L_N = \sum_{k=0}^N f[x_0, \dots, x_k] \prod_{j=0}^{k-1} (X - x_j)$$

Dém: Par récurrence sur $n \in \mathbb{N}$. Au rang 0, $L_0 = f[x_0]$. Supposons que la formule est vraie jusqu'au rang $N-1$. Au rang N , $L_N - L_{N-1}$ a un degré au plus égal à

$$N \text{ et s'annule en } x_0, \dots, \text{ et } x_{N-1} \Rightarrow \text{il existe } \lambda \in \mathbb{R} \text{ tel que: } L_N - L_{N-1} =$$

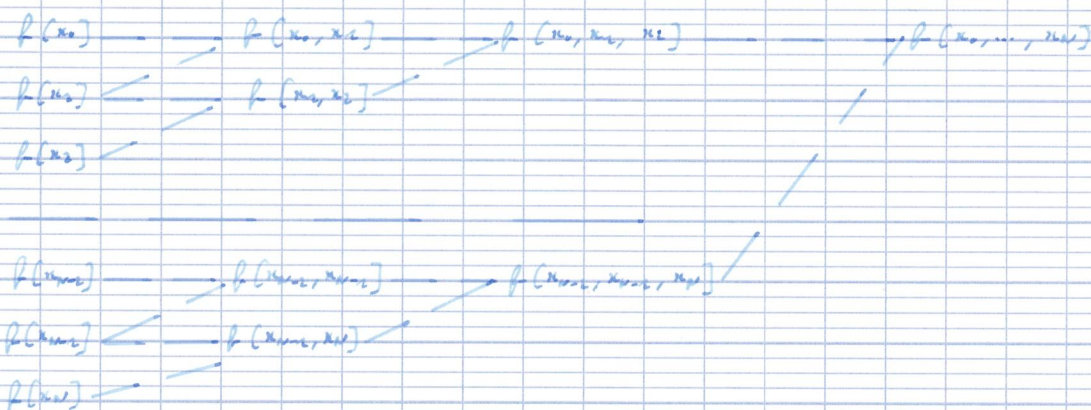
$$\lambda \prod_{j=0}^{N-1} (X - x_j). \text{ Comme le terme dominant des polynômes } L_N - L_{N-1} \text{ est égal à } f[x_0, \dots, x_N] X^N, \lambda = f[x_0, \dots, x_N], \text{ et: } L_N = \sum_{k=0}^N f[x_0, \dots, x_k] \prod_{j=0}^{k-1} (X - x_j).$$

Cette formule permet un calcul rapide du polynôme d'interpolation de Lagrange L_N .

Le calcul repose sur l'algorithme suivant:

2^{ème} étape: Calcul itératif des différences divisées successives de la fonction f .

Au rang 0, sont calculées les différences divisées $f[x_0], \dots$ et $f[x_n]$, qui permettent le calcul des différences divisées $f[x_0, x_1], f[x_1, x_2], \dots$ et $f[x_{n-1}, x_n]$, et ainsi de suite, suivant le tableau ci-dessous:



3^{ème} étape: Calcul des polynômes Q_j par l'algorithme de Horner.

Cet algorithme consiste à poser $Q_n = f[x_0, \dots, x_n]$, puis à déterminer de façon itérative les polynômes $Q_{n-1}, \dots, Q_1$ et Q_0 suivant la formule:

$$\forall 0 \leq j \leq n-1, Q_j = f[x_0, \dots, x_j] + (x - x_{j+1}) Q_{j+1}.$$

Par la formule de Newton, le polynôme Q_0 est alors égal au polynôme d'interpolation de Lagrange L_n .

L'algorithme présente l'intérêt d'être peu coûteux en temps comme en mémoire. Son caractère récursif facilite aussi l'ajout d'un ou plusieurs points d'interpolation.

3. Estimation de l'erreur d'interpolation

Considérons une fonction $f: I \rightarrow \mathbb{R}$ et $N+1$ points $(x_0, \dots, x_N) \in I^{N+1}$ deux à deux distincts.

L'estimation de l'erreur d'interpolation entre la fonction f et son polynôme d'interpolation de Lagrange L_n aux points $(x_0, \dots, x_n)$ repose sur le corollaire suivant du théorème de Rolle.

Chm: Soit $p \geq 2$. Supposons que:

- (i) La fonction f est p -fois dérivable sur l'intervalle I .
- (ii) Il existe des points $c_0 < c_1 < \dots < c_p$ tels que: $\forall 0 \leq i \leq p, f(c_i) = 0$.

alors il existe un point $c_0 < \xi < c_p$ tel que: $f^{(p)}(\xi) = 0$.

Dém: Vérifions par récurrence sur $0 \leq j \leq p$ qu'il existe des points $(c_0^{(j)}, \dots, c_{p-j}^{(j)}) \in I$ tels que: $\forall 0 \leq i \leq p-j, c_i^{(j-1)} < c_i^{(j)} < c_{i+1}^{(j-1)}$ et $f^{(j)}(c_i^{(j)}) = 0$. Cette propriété au rang $j=0$. Supposons qu'elle soit vraie jusqu'au rang $j-1$. Au rang j , la fonction $f^{(j-1)}$ est continue sur $[c_{i-1}^{(j-1)}, c_{i+1}^{(j-1)}]$ et dérivable sur $]c_{i-1}^{(j-1)}, c_{i+1}^{(j-1)}[$. Comme $f^{(j-1)}(c_i^{(j-1)}) = f^{(j-1)}(c_{i+1}^{(j-1)})$, il existe par le théorème de Rolle un point $c_i^{(j)} \in]c_{i-1}^{(j-1)}, c_{i+1}^{(j-1)}[$ tel que $f^{(j)}(c_i^{(j)}) = 0$, d'où la construction des points $c_i^{(j)}$ pour tout $0 \leq j \leq p$ et $0 \leq i \leq p-j$. Pour $j=p$, le point $\xi = c_0^{(p)}$ vérifie: $c_0 < \xi < c_p$ et $f^{(p)}(\xi) = 0$.

Grâce à ce théorème, l'erreur d'interpolation $E_p(f) = f - L_p$ entre la fonction f et son polynôme d'interpolation de Lagrange L_p s'écrit de la façon suivante.

Thm: Supposons que la fonction f est $N+1$ -fois dérivable sur l'intervalle I . Alors:

$$\forall x \in I, \exists \min\{x_0, x\} < \xi_x < \max\{x_N, x\} \text{ t.q. } f(x) - L_p(x) = \frac{f^{(N+1)}(\xi_x)}{(N+1)!} \prod_{i=0}^N (x - x_i).$$

Dém: Observons que, lorsque $x \in]x_0, \dots, x_N]$, tout point $\xi_x \in]\min\{x_0, x\}, \max\{x_N, x\}]$ convient. Considérons alors le polynôme d'interpolation de Lagrange L_x de la fonction f aux points $(x, x_0, \dots, x_N)$. Comme la fonction $f - L_x$ s'annule en $x, x_0, \dots, x_N$, il existe un point $\min\{x_0, x\} < \xi_x < \max\{x_N, x\}$ tel que $f^{(N+1)}(\xi_x) = L_x^{(N+1)}(\xi_x)$.

Comme les polynômes $L_x - L_p$ s'annule en $x_0, \dots, x_N$ et a un degré inférieur au degré $N+1$, il existe un nombre $\lambda \in \mathbb{R}$ tel que:

$$L_x = L_p + \lambda \prod_{j=0}^N (x - x_j)$$

$$\text{Puisque que } d^0(L_p) \leq N, L_x^{(N+1)}(\xi) = (N+1)! \lambda, \text{ d'où la formule: } L_x = L_p + \frac{f^{(N+1)}(\xi_x)}{(N+1)!} \prod_{j=0}^N (x - x_j) \Rightarrow f(x) = L_p(x) = L_p(x) + \frac{f^{(N+1)}(\xi_x)}{(N+1)!} \prod_{j=0}^N (x - x_j).$$

Exe: Supposons que:

(i) $I = [a, b]$ est un segment (non vide).

(ii) Une fonction f est de classe C^{N+1} sur le segment $[a, b]$.

$$\text{alors: } \|f - L_p\|_\infty = \sup_{x \in [a, b]} |f(x) - L_p(x)| \leq \frac{\|f^{(N+1)}\|_\infty}{(N+1)!} \left\| \prod_{i=0}^N (x - x_i) \right\|_\infty$$

Dém: Par le théorème précédent, $\forall x \in [a, b], |f(x) - L_p(x)| \leq \frac{|f^{(N+1)}(\xi_x)|}{(N+1)!} \left| \prod_{i=0}^N (x - x_i) \right|$; comme la fonction f et sa dérivée $f^{(N+1)}$ sont continues sur $[a, b]$, elles sont

connus sur $[a, b]$, ce qui autorise à écrire l'inégalité recherchée.

Rem: (i) Les résultats certifient que l'erreur d'interpolation ne dépend que de :

- la taille des oscillations de la fonction f à travers la norme uniforme $\|f^{(N+1)}\|_\infty$,
- la répartition des points $(x_0, \dots, x_N)$ à travers la norme $\|\prod_{j=0}^N (x - x_j)\|_\infty$.

(ii) Il n'est pas toujours vrai que l'erreur diminue lorsque le nombre de points augmente.

Ex: (i) Supposons que $I = [-1, 1]$ et que les points $(x_0, \dots, x_N)$ sont équirépartis :

$$\forall 0 \leq i \leq N, x_i = -1 + \frac{2i}{N}.$$

Il est possible de vérifier que'il existe un nombre $C \geq 0$ tel que :

$$\forall x \in I, \left| \prod_{i=0}^N (x - x_i) \right| \leq C \left(\frac{2}{e} \right)^N.$$

L'estimation de l'erreur d'interpolation devient alors :

$$\|f - L_N\|_\infty \leq C \frac{2^N}{2^N (N+1)!} \|f^{(N+1)}\|_\infty.$$

(ii) Supposons que $I = [-1, 1]$ et considérons les points de Chebyshev :

$$\forall 0 \leq i \leq N, x_i = \cos\left(\frac{2i+1}{2(N+1)} \pi\right).$$

Il est possible de vérifier que :

$$\forall x \in I, \left| \prod_{i=0}^N (x - x_i) \right| \leq \frac{1}{2^{N+1}}.$$

L'estimation de l'erreur d'interpolation devient alors :

$$\|f - L_N\|_\infty \leq \frac{1}{2^{N+1} (N+1)!} \|f^{(N+1)}\|_\infty.$$

Cette erreur est bien meilleure que pour les points équirépartis.

4. Stabilité numérique de la méthode d'interpolation de Lagrange

Supposons que l'intervalle I est un segment (a, b) (non réduit à un point) et que la fonction f est continue sur ce segment. En pratique, nous ne disposons souvent que d'une valeur approchée de la fonction f que nous souhaitons interpoler. La méthode de Lagrange aux points $(x_0, \dots, x_N) \in (a, b)^{N+1}$. Il est naturel de chercher à contrôler la façon dont cette méthode va amplifier cette erreur dans le calcul du polynôme d'interpolation de Lagrange L_N . Ce contrôle repose sur l'introduction de l'opérateur d'interpolation de Lagrange L_N .

Def: L'opérateur L_n d'interpolation de Lagrange aux points $(x_0, \dots, x_N)$ est défini par:

$$\forall f \in \mathcal{C}^0([a, b], \mathbb{R}), \quad L_n(f) = L_n,$$

où L_n désigne le polynôme d'interpolation de Lagrange de la fonction f aux points $(x_0, \dots, x_N)$.

lem: L'opérateur L_n d'interpolation de Lagrange aux points $(x_0, \dots, x_N)$ est linéaire.

Dém: Si $L_n(f)$ et $L_n(g)$ désignent les polynômes d'interpolation de Lagrange des fonctions f , respectivement g , aux points $(x_0, \dots, x_N)$, et si $\lambda \in \mathbb{R}$, alors, $\lambda L_n(f) + L_n(g)$ est le polynôme d'interpolation de Lagrange de la fonction $\lambda f + g$ aux points $(x_0, \dots, x_N)$, d'où la formule: $L_n(\lambda f + g) = \lambda L_n(f) + L_n(g)$.

Chm: Considérons les polynômes de base de Lagrange $l_0(x), \dots$, et $l_N(x)$ définis par:

$$\forall 0 \leq i \leq N, \quad l_i(x) = \prod_{j \neq i} \frac{(x - x_j)}{(x_i - x_j)}.$$

L'opérateur L_n d'interpolation de Lagrange aux points $(x_0, \dots, x_N)$ est continu de $(\mathcal{C}^0([a, b], \mathbb{R}), \|\cdot\|_\infty)$ dans $(\mathcal{C}^0([a, b], \mathbb{R}), \|\cdot\|_\infty)$. De plus, il vérifie:

$$\forall f \in \mathcal{C}^0([a, b], \mathbb{R}), \quad \|L_n(f)\|_\infty \leq \Lambda_n \|f\|_\infty,$$

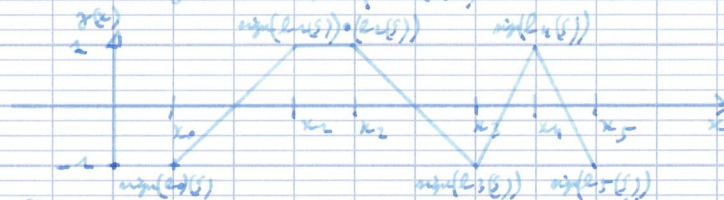
où la constante de Lebesgue Λ_n est donnée par la formule:

$$\Lambda_n = \sup \{ \|L_n(f)\|_\infty, f \in \mathcal{C}^0([a, b], \mathbb{R}) \text{ t.q. } \|f\|_\infty \leq 1 \} = \left\| \sum_{i=0}^N |l_i| \right\|_\infty.$$

Dém: (i) Continuité: Comme $\forall x \in [a, b], L_n(x) = \sum_{i=0}^N f(x_i) l_i(x)$, $\|L_n(f)\|_\infty \leq \|f\|_\infty \left\| \sum_{i=0}^N |l_i| \right\|_\infty$; par linéarité, $\|L_n(f_1) - L_n(f_2)\|_\infty = \|L_n(f_1 - f_2)\|_\infty \leq \left\| \sum_{i=0}^N |l_i| \right\|_\infty \|f_1 - f_2\|_\infty$, ce qui garantit la continuité de l'opérateur L_n .

(ii) Calcul de la constante Λ_n : D'après (i), $\Lambda_n \leq \left\| \sum_{i=0}^N |l_i| \right\|_\infty$. Comme la fonction $\sum_{i=0}^N |l_i|$ est continue sur $[a, b]$, il existe un point $\xi \in [a, b]$ tel que $\sum_{i=0}^N |l_i(\xi)| = \left\| \sum_{i=0}^N |l_i| \right\|_\infty$. Construisons alors une fonction affine par morceaux $g \in \mathcal{C}^0([a, b], \mathbb{R})$ telle que:

$$\forall 0 \leq i \leq N, \quad g(x_i) = \text{sign}(l_i(\xi)) \text{ et } \|g\|_\infty \leq 1.$$



$$\text{et vérifions que: } \|L_n(g)\|_\infty \geq L_n(g)(\xi) = \sum_{i=0}^N |l_i(\xi)| = \left\| \sum_{i=0}^N |l_i| \right\|_\infty.$$

Comme $\|f\|_\infty = 1$, $\|\sum_{i=0}^N |L_i|\|_\infty \leq \Lambda_N$, d'où l'égalité.

Rem. (i) En pratique, le calcul du polynôme d'interpolation de Lagrange L_N s'effectue à partir d'une valeur approchée $\tilde{f} = f + \varepsilon$ de la fonction f , où la fonction ε prend en compte les erreurs d'arrondis. L'erreur numérique sur le calcul du polynôme L_N est égale à la fonction $L_N(\tilde{f}) - L_N(f) = L_N(\varepsilon)$. La taille maximale est contrôlée par le nombre $\Lambda_N \| \varepsilon \|_\infty$. La constante de Lebesgue Λ_N détermine l'accumulation maximale des erreurs. Plus elle est grande, moins la méthode sera numériquement stable.

(ii) Il est possible de vérifier que la constante de Lebesgue Λ_N tend vers $+\infty$ lorsque $N \rightarrow +\infty$, quel que soit le choix des points d'interpolation $(x_0, \dots, x_N)$.

Ex. (i) Supposons que $I = [-1, 1]$ et que les points $(x_0, \dots, x_N)$ sont équirépartis:

$$\forall 0 \leq i \leq N, x_i = -1 + \frac{2i}{N}.$$

Il est possible de vérifier que:

$$\Lambda_N \underset{N \rightarrow +\infty}{\sim} \frac{2^{N+1}}{2N \ln(N)} \rightarrow +\infty$$

(ii) Supposons que $I = [-1, 1]$ et considérons les points de Chebyshev:

$$\forall 0 \leq i \leq N, x_i = \cos\left(\frac{2i+1}{2(N+1)}\pi\right).$$

Il est possible de vérifier que:

$$\Lambda_N \underset{N \rightarrow +\infty}{\sim} \frac{2}{\pi} \ln(N).$$

Cet équirépart est le meilleur parmi tous les choix de points $(x_0, \dots, x_N)$ possibles.

5. Phénomène de Runge et divergence de l'interpolation de Lagrange

Supposons que l'intervalle I est un segment $[a, b]$ (non réduit à un point) et que la fonction f est continue sur ce segment. Il est naturel de se demander à quelles conditions sur les propriétés de la fonction f et sur le choix des points $(x_0, \dots, x_N)$ le polynôme d'interpolation de Lagrange L_N de la fonction f aux points $(x_0, \dots, x_N)$ converge (uniformément) vers la fonction f lorsque le nombre de points N tend vers $+\infty$. Pour une fonction f quelconque et des points $(x_0, \dots, x_N)$ arbitraires, la réponse à cette question est en réalité négative. Cette affirmation est liée au phé-

nomme de Runge que nous allons maintenant décrire (et peut se démontrer de façon rigoureuse à l'aide du théorème de Banach - Steinhaus).

Ex: Phénomène de Runge.

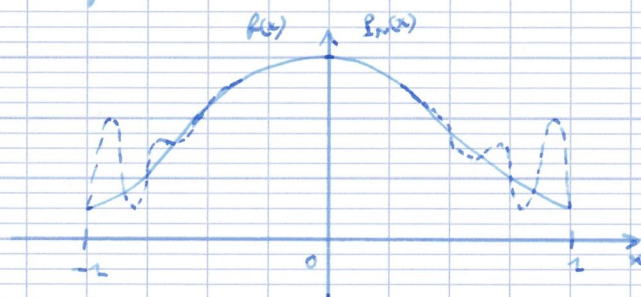
Supposons que $I = [-1, 1]$ et que les points $(x_0, \dots, x_N)$ sont équirépartis:

$$\forall 0 \leq i \leq N, x_i = -1 + \frac{2i}{N}.$$

Considérons la fonction f définie par:

$$\forall x \in [-1, 1], f(x) = \frac{1}{1+25x^2}.$$

Pour N grand, le graphe du polynôme d'interpolation de Lagrange présente des oscillations de la forme suivante:



Ce phénomène mis en évidence par Runge en 1901 induit que la distance uniforme $\|f - L_N\|_\infty$ diverge vers $+\infty$ lorsque $N \rightarrow +\infty$. Il s'explique par le fait que :

- Les constantes de Lebesgue Λ_N augmentent rapidement lorsque $N \rightarrow +\infty$,
- Les dérivées $f^{(k)}$ de la fonction f croissent également très vite lorsque $N \rightarrow +\infty$.

Observons cependant que les polynômes d'interpolation de Lagrange L_N convergent uniformément vers la fonction f sur un intervalle strict de la forme $[-p, p]$, avec $p < 1$. Cette propriété s'explique par l'analyticité de la fonction f sur l'intervalle ouvert $] -\frac{1}{5}, \frac{1}{5} [$ (qui permet de contrôler ces dérivées successives sur cet intervalle). Les oscillations n'apparaissent donc que sur les bords du segment $[-1, 1]$, où elles s'intensifient lorsque N augmente.

Même si le polynôme d'interpolation de Lagrange L_N fournit la valeur exacte de la fonction f aux points $(x_0, \dots, x_N)$, il ne constitue en général donc pas une bonne approximation de cette fonction, question que nous allons désormais étudier.

II Polynôme de meilleure approximation

Supposons que l'intervalle I est un segment $[a, b]$ (non réduit à un point) et considérons un sous-espace E de l'espace des fonctions définies sur I et à valeurs réelles. Supposons que le sous-espace E contient l'espace des fonctions polynômes identifié à $\mathbb{R}[X]$ et qu'il est de plus muni d'une norme $\| \cdot \|_E$. Pour un degré $N \geq 0$ fixé, le problème du polynôme de meilleure approximation consiste à déterminer un polynôme $P_N \in \mathbb{R}_N[X]$ qui minimise la distance d'une fonction $f \in E$ à l'espace $\mathbb{R}_N[X]$ des polynômes de degré inférieur ou égal à N .

1. Existence (et unicité) du polynôme de meilleure approximation

La construction du polynôme de meilleure approximation repose sur la propriété d'équivalence des normes en dimension finie.

Déf. Soit E un $\mathbb{R}$ -espace vectoriel. Une norme $\| \cdot \|_E$ sur l'espace E est une application de E dans $\mathbb{R}_+$ qui satisfait les propriétés suivantes :

- (i) $\forall x \in E, \|x\|_E = 0 \Leftrightarrow x = 0$
- (ii) $\forall \lambda \in \mathbb{R}, \forall x \in E, \|\lambda x\|_E = |\lambda| \|x\|_E$
- (iii) $\forall (x, y) \in E^2, \|x+y\|_E \leq \|x\|_E + \|y\|_E$

Ex. (i) Soit $\forall x \in \mathbb{R}^n, \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$. L'application $\| \cdot \|_\infty$ est la norme uniforme sur $\mathbb{R}^n$.

(ii) Soit $\forall f \in \mathcal{C}^0([a, b], \mathbb{R}), \|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{\frac{1}{2}}$. L'application $\| \cdot \|_2$ est une norme (hillbertienne) sur $\mathcal{C}^0([a, b], \mathbb{R})$.

Thm. Supposons que :

- (i) L'espace vectoriel E est de dimension finie.
- (ii) Les applications $\| \cdot \|_1$ et $\| \cdot \|_2$ sont des normes sur l'espace E .

Alors les normes $\| \cdot \|_1$ et $\| \cdot \|_2$ sont équivalentes, c'est-à-dire qu'il existe deux nombres $\alpha_1 > 0$ et $\alpha_2 > 0$ tels que :

$$\forall x \in E, \alpha_1 \|x\|_2 \leq \|x\|_1 \leq \alpha_2 \|x\|_2.$$

Dém: Considérons une base $B = (e_1, \dots, e_N)$ de l'espace vectoriel E et la norme uniforme: $\forall x = \sum_{i=1}^N x_i e_i, \|x\|_\infty = \max_{1 \leq i \leq N} |x_i|$. Par définition d'une norme: $\forall x \in E$, $\|x\|_1 \leq \sum_{i=1}^N |x_i| \|e_i\|_1 \leq \left(\sum_{i=1}^N \|e_i\|_1\right) \|x\|_\infty \Rightarrow \forall (x, y) \in E^2, \|\|x\|_1 - \|y\|_1\| \leq \|x - y\|_1 \leq \left(\sum_{i=1}^N \|e_i\|_1\right) \|x - y\|_\infty \Rightarrow \|\| \cdot \|_1$ est continue dans l'espace vectoriel normé $(E, \|\cdot\|_\infty)$; comme $X_\infty = \{x \in E \text{ t.q. } \|x\|_\infty = 1\}$ est borné et fermé dans cet espace de dimension finie, X_∞ est compact $\Rightarrow \exists a \in X_\infty$ t.q. $\forall u \in X_\infty, \|u\|_1 \geq \|a\|_1 \Rightarrow \forall x \in E, \|x\|_1 \geq \|a\|_1 \|x\|_\infty \Rightarrow$ Les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sont équivalentes; par transitivité de la notion d'équivalence, les normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sont équivalentes.

Ex: Soit $\forall x \in \mathbb{R}^N, \|x\|_1 = \sum_{i=1}^N |x_i|$ et $\|x\|_2 = \left(\sum_{i=1}^N x_i^2\right)^{\frac{1}{2}}$. Les normes $\|\cdot\|_1, \|\cdot\|_2$ et $\|\cdot\|_\infty$ sont équivalentes.

Chm: Supposons que E est un sous-espace de l'espace des fonctions définies sur I et à valeurs réelles, qui contient le sous-espace des fonctions polynômes $\mathbb{R}[X]$. Considérons une norme $\|\cdot\|_E$ sur E et un entier $N \geq 0$. Soit la fonction $f \in E$, il existe un polynôme $P_N \in \mathbb{R}_N[X]$ tel que la fonction polynôme associée satisfait:

$$\|f - P_N\|_E = \min_{Q \in \mathbb{R}_N[X]} \|f - Q\|_E.$$

Le polynôme P_N est appelé polynôme de meilleure approximation (d'ordre N) de la fonction f dans E .

Dém: Considérons la norme $\|\cdot\|_\infty$ sur le sous-espace $\mathbb{R}_N[X]$ définie par:

$\forall P = \sum_{h=0}^N a_h X^h, \|P\|_\infty = \max_{0 \leq h \leq N} |a_h|$. Par définition de $I = \inf_{Q \in \mathbb{R}_N[X]} \|f - Q\|_E$, il existe une suite $(Q_m)_{m \in \mathbb{N}}$ tel que: $\forall m \in \mathbb{N}, Q_m \in \mathbb{R}_N[X]$ et $\|f - Q_m\|_E \rightarrow I$. En particulier, $(Q_m)_{m \in \mathbb{N}}$ est bornée pour la norme $\|\cdot\|_E$, donc par équivalence des normes, pour la norme $\|\cdot\|_\infty$. Si $\forall m \in \mathbb{N}, Q_m = \sum_{h=0}^N b_h^m X^h$, alors les suites $(b_h^m)_{m \in \mathbb{N}}$ sont bornées pour tout $0 \leq h \leq N$. Soit α un réel, il existe des coefficients $(a_h)_{0 \leq h \leq N}$ tels que: $\forall 0 \leq h \leq N, b_h^m \rightarrow a_h$. Si $P_N = \sum_{h=0}^N a_h X^h$, alors $\|Q_m - P_N\|_\infty \xrightarrow{m \rightarrow +\infty} 0 \Rightarrow \|Q_m - P_N\|_E \xrightarrow{m \rightarrow +\infty} 0$, d'où $\|f - P_N\|_E \leq \lim_{m \rightarrow +\infty} \|Q_m - P\|_E + \|Q_m - P_N\|_E = I \Rightarrow \|f - P_N\|_E = I = \min_{Q \in \mathbb{R}_N[X]} \|f - Q\|_E$, puisque $P_N \in \mathbb{R}_N[X]$.

Def: (i) Un polynôme de meilleure approximation p_N dépend de l'entier $N \geq 0$, de la fonction $f \in E$, mais aussi de la norme $\|\cdot\|_E$ choisie sur l'espace E .

Ex: Soit $E = \mathcal{C}^0([a, b], \mathbb{R})$, et $\forall f \in E$, $\|f\|_2 = \left(\int_a^b f(t)^2 dt \right)^{\frac{1}{2}}$ et $\|f\|_\infty = \max_{t \in [a, b]} |f(t)|$.

Un polynôme de meilleure approximation p_N pour la norme $\|\cdot\|_2$ est appelé polynôme de meilleure approximation quadratique ou hilbertienne; un polynôme de meilleure approximation p_N pour la norme $\|\cdot\|_\infty$ est appelé polynôme de meilleure approximation uniforme.

(ii) Un polynôme de meilleure approximation p_N n'est pas en général unique, mais le polynôme de meilleure approximation quadratique et le polynôme de meilleure approximation uniforme sont uniques.

Ex: Soit $E = L^2([-1, 1], \mathbb{R})$ et $\forall x \in [-1, 1]$, $f(x) = -1$ si $x < 0$, 0 si $x = 0$ et 1 si $x > 0$. Alors quelque soit $a \in [-1, 1]$, $p_0 = a$ est un polynôme de meilleure approximation de la fonction f pour la norme: $\forall f \in L^2([-1, 1], \mathbb{R})$, $\|f\|_2 = \left(\int_{-1}^1 |f(t)|^2 dt \right)^{\frac{1}{2}}$.

Il est naturel de se demander si une suite de polynômes de meilleure approximation $(p_N)_{N \in \mathbb{N}}$ converge vers la fonction f lorsque l'entier N tend vers $+\infty$. Le théorème de Weierstrass répond de façon positive à cette question dans le cas de la norme uniforme $\|\cdot\|_\infty$ sur un segment $[a, b]$.

2. Théorème de Weierstrass

Nous supposons dans un premier temps que l'intervalle I est égal au segment $[0, 1]$.

Def: Soit $f \in \mathcal{C}^0([0, 1], \mathbb{R})$. Le polynôme d'approximation de Bernstein d'ordre $N \in \mathbb{N}^*$ de la fonction f est défini par:

$$B_N = \sum_{k=0}^N \binom{N}{k} f\left(\frac{k}{N}\right) x^k (1-x)^{N-k}.$$

Ex: Soit $f \in \mathcal{C}^0([0, 1], \mathbb{R})$, alors: $B_1 = f(0)(1-x) + f(1)x$ et $B_2 = f(0)(1-x)^2 + 2f\left(\frac{1}{2}\right)x(1-x) + f(1)x^2$.

Théorème de Bernstein: Supposons que $f \in \mathcal{C}^0([0, 1], \mathbb{R})$. La suite des polynômes d'approximation de Bernstein $(B_N)_{N \geq 1}$ de la fonction f converge uniformément vers la

fonction f sur le segment $[0, 1]$.

Dém: (i) Rappelons que: $\forall x \in [0, 1], \sum_{k=0}^N \binom{N}{k} x^k (1-x)^{N-k} = (x+1-x)^N = 1$

(ii) Soit $x \in [0, 1]$. Par définition, $B_N(x) - f(x) = \sum_{k=0}^N \binom{N}{k} f\left(\frac{k}{N}\right) x^k (1-x)^{N-k} - f(x)$
 $= \sum_{k=0}^N \binom{N}{k} \left(f\left(\frac{k}{N}\right) - f(x)\right) x^k (1-x)^{N-k}$.

(iii) Soit $\varepsilon > 0$. Comme f est continue sur $[0, 1]$, par le théorème de Heine, elle est uniformément continue sur $[0, 1] \Rightarrow$ Il existe un nombre $\delta > 0$ tel que

$$\forall (p, q) \in [0, 1]^2, |x - y| \leq \delta \Rightarrow |f(y) - f(x)| \leq \varepsilon \Rightarrow \left| \sum_{0 \leq k \leq N} \binom{N}{k} \left(f\left(\frac{k}{N}\right) - f(x)\right) x^k (1-x)^{N-k} \right| \leq \sum_{0 \leq k \leq N} \binom{N}{k} x^k (1-x)^{N-k} \leq \sum_{\left|\frac{k}{N} - x\right| \leq \delta} \binom{N}{k} x^k (1-x)^{N-k} \leq \varepsilon.$$

(iv) Comme f est continue sur $[0, 1]$, il existe un nombre $M \geq 0$ tel que: $\forall y \in [0, 1], |f(y)| \leq M \Rightarrow \left| \sum_{\substack{0 \leq k \leq N \\ \left|\frac{k}{N} - x\right| \geq \delta}} \binom{N}{k} \left(f\left(\frac{k}{N}\right) - f(x)\right) x^k (1-x)^{N-k} \right| \leq 2M \sum_{\substack{0 \leq k \leq N \\ \left|\frac{k}{N} - x\right| \geq \delta}} \binom{N}{k} x^k (1-x)^{N-k}$.

(v) Calculons: $\forall x \in [0, 1], \sum_{k=0}^N \left(\frac{k}{N} - x\right)^2 \binom{N}{k} x^k (1-x)^{N-k} = \sum_{k=0}^N \binom{N}{k} \left(\frac{k^2}{N^2} - 2x\frac{k}{N} + x^2\right) x^k (1-x)^{N-k}$; par la formule du binôme de Newton, $\forall y \in [0, 1]$,
 $\sum_{k=0}^N \binom{N}{k} y^k (1-x)^{N-k} = (y+1-x)^N \Rightarrow \sum_{k=0}^N \binom{N}{k} k y^{k-1} (1-x)^{N-k} = N(y+1-x)^{N-1}$,
 $\sum_{k=0}^N \binom{N}{k} k(k-1) y^{k-2} (1-x)^{N-k} = N(N-1)(y+1-x)^{N-2}$,
d'où $\sum_{k=0}^N \binom{N}{k} k^2 x^{k-2} (1-x)^{N-k} = N(N-1)x + N$ et $\sum_{k=0}^N \binom{N}{k} (k^2 - k) x^{k-1} (1-x)^{N-k} = (N^2 - N)x^2$,
 $\Rightarrow \sum_{k=0}^N \left(\frac{k}{N} - x\right)^2 \binom{N}{k} x^k (1-x)^{N-k} = \frac{x(1-x)}{N}$, d'où $\sum_{\substack{0 \leq k \leq N \\ \left|\frac{k}{N} - x\right| \geq \delta}} \binom{N}{k} x^k (1-x)^{N-k} \leq \frac{x(1-x)}{N\delta^2} \leq \frac{1}{4N\delta^2}$.

(vi) Nous aurons à: $|B_N(x) - f(x)| \leq \varepsilon + \frac{M}{2N\delta^2}$. Nous fixons $N_\varepsilon \leq N$ tel que $N_\varepsilon \geq \frac{M}{2\delta^2\varepsilon}$ et concluons que: $\forall N \geq N_\varepsilon, |B_N(x) - f(x)| \leq 2\varepsilon$, ce qui assure la convergence uniforme de la suite $(B_N)_{N \geq 1}$ vers la fonction f sur le segment $[0, 1]$.

Nous pouvons ensuite étendre ce résultat de densité des fonctions polynômes parmi les fonctions continues au cas d'un segment $[a, b]$ quelconque.

Théorème de Weierstrass: L'ensemble des fonctions polynômes $\mathbb{R}[X]$ est dense dans l'espace

$\mathcal{C}^0([a, b], \mathbb{R})$ pour la norme uniforme. Autrement dit, quelle que soit la

fonction $f \in \mathcal{C}^0([a, b], \mathbb{R})$, il existe une suite de fonctions polynômes $(p_n)_{n \in \mathbb{N}}$ telle que:

$$\|P_n - f\|_\infty = \max_{x \in (a,b)} |P_n(x) - f(x)| \xrightarrow{n \rightarrow +\infty} 0.$$

Dém: La preuve repose sur le théorème de Bernstein. Soit $\forall t \in [0,1]$, $g(t) = f(a+t(b-a))$. Lorsque f est continue sur $[a,b]$, la fonction g est continue sur $[0,1]$. Par le théorème de Bernstein, les polynômes de Bernstein $(B_n)_{n \in \mathbb{N}^+}$ de la fonction g satisfont: $\|B_n - g\|_\infty \xrightarrow{n \rightarrow +\infty} 0$. D'où: $\forall n \in \mathbb{N}$, $\forall x \in (a,b)$, $P_n(x) = B_{n+1}\left(\frac{x-a}{b-a}\right)$. Les fonctions P_n sont bien définies et sont des fonctions polynômes. De plus, le changement de variable $x = a+t(b-a)$ assure que: $\|P_n - f\|_\infty = \|B_{n+1} - g\|_\infty \xrightarrow{n \rightarrow +\infty} 0$, d'où le théorème de Weierstrass.

Dém: La preuve du théorème de Weierstrass assure qu'il est possible de choisir les fonctions polynômes P_n (quitte à les renumérotter) de façon à ce que: $\forall n \in \mathbb{N}$, $d^0(P_n) \leq n$.

Le théorème de Weierstrass assure la convergence uniforme de la suite des polynômes $(P_n)_{n \geq 0}$ de meilleure approximation uniforme d'une fonction $f \in \mathcal{C}^0([a,b], \mathbb{R})$ vers cette fonction f .

Cor: Étant donnée une fonction $f \in \mathcal{C}^0([a,b], \mathbb{R})$, considérons la suite $(P_n)_{n \geq 0}$ des polynômes de meilleure approximation uniforme de la fonction f . Alors:

$$\|f - P_n\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

Dém: Par le théorème de Weierstrass, il existe des polynômes $(Q_n)_{n \geq 0}$ tels que: $\forall n \geq 0$, $d^0(Q_n) \leq n$ et $\|f - Q_n\|_\infty \xrightarrow{n \rightarrow +\infty} 0$. Par définition, $\forall n \geq 0$, $\|f - P_n\|_\infty \leq \|f - Q_n\|_\infty \Rightarrow \|f - P_n\|_\infty \xrightarrow{n \rightarrow +\infty} 0$.

3. Polynômes orthogonaux

a. Définition et propriétés

Nous revenons au cas général où I désigne un intervalle quelconque de $\mathbb{R}$. Nous considérons une fonction poids $w \in \mathcal{C}^0(I, \mathbb{R})$ telle que:

$$\forall x \in I, w(x) > 0,$$

et nous introduisons l'espace E défini par:

$$E = \left\{ f \in \mathcal{C}^0(I, \mathbb{R}) \text{ telle que } \int_I f(x)^2 w(x) dx < +\infty \right\}.$$

Nous vérifions que cet espace E est un espace euclidien ou pré-hilbertien pour la forme bilinéaire :

$$\forall (f, g) \in E^2, \langle f, g \rangle_E = \int_I f(x) g(x) w(x) dx.$$

Dém: La forme bilinéaire $\langle, \rangle_E$ est un produit scalaire sur l'espace E .

Dém: (i) Comme $\forall (f, g) \in E^2, \forall x \in \mathbb{R}, |f(x)g(x)| \leq \frac{1}{2}(f(x)^2 + g(x)^2)$, $\int_I |f(x)g(x)| w(x) dx \leq \frac{1}{2} \int_I (f(x)^2 + g(x)^2) w(x) dx < +\infty \Rightarrow \langle, \rangle_E$ est bien définie sur E^2 .

(ii) Par définition, $\langle, \rangle_E$ est une forme bilinéaire symétrique sur E et $\forall f \in E, \|f\|_E^2 = \langle f, f \rangle_E = \int_I f(x)^2 w(x) dx \geq 0$. De plus, comme f et w sont continues sur I , cette intégrale est nullessi $f^2 w = 0$, soitssi $f = 0$, puisque w est à valeurs strictement positives.

Nous supposons de plus que l'espace des fonctions polynômes $\mathbb{R}[X]$ est un sous-espace de E .

Elm: Il existe une suite unique $(l_n)_{n \in \mathbb{N}}$ de fonctions polynômes telle que :

$$(i) \forall n \in \mathbb{N}, d^0(l_n) = n$$

(ii) $\forall n \in \mathbb{N}, l_n$ est unitaire, autrement dit, son coefficient de plus haut degré est égal à 1.

$$(iii) \forall (m, n) \in \mathbb{N}^2, m \neq n \Rightarrow \langle l_m, l_n \rangle_E = 0.$$

Les polynômes $(l_n)_{n \in \mathbb{N}}$ sont appelés polynômes orthogonaux (pour la fonction poids w).

Dém: La preuve repose sur une récurrence sur l'entier $n \in \mathbb{N}$ qui utilise le procédé d'orthogonalisation de Gram-Schmidt. Au rang $n=0$, sachant que l_0 est de degré 0 et unitaire, il est défini de manière unique par $l_0 = 1$. Supposons avoir construit les polynômes $(l_h)_{0 \leq h \leq n}$ jusqu'au rang, de façon unique. Au rang $n+1$, comme $\forall 0 \leq h \leq n, d^0(l_h) = h$, la famille $(l_h)_{0 \leq h \leq n}$ forme une base de $\mathbb{R}_n[X]$. Sachant que l_{n+1} est de degré $n+1$ et unitaire,

il existe des nombres $(\lambda_h)_{0 \leq h \leq n}$ tels que : $L_{n+1} = X^{n+1} + \sum_{h=0}^n \lambda_h L_h$. Comme les polynômes $(L_h)_{0 \leq h \leq n}$ sont deux à deux orthogonaux et orthogonaux à L_{n+1} , $\forall 0 \leq h \leq n$, $\lambda_h \|L_h\|_E^2 = -\langle X^{n+1}, L_h \rangle_E$. Comme les polynômes $(L_h)_{0 \leq h \leq n}$ sont de degré h , donc non nuls, nous concluons que : $L_{n+1} = X^{n+1} - \sum_{h=0}^n \frac{\langle X^{n+1}, L_h \rangle_E}{\|L_h\|_E^2} \times L_h$, d'où l'existence et l'unicité de la suite $(L_n)_{n \in \mathbb{N}}$ par récurrence sur l'entier $n \in \mathbb{N}$.

Dém: (i) Les polynômes orthogonaux $(L_n)_{n \in \mathbb{N}}$ sont échelonnés donc forment une base de $\mathbb{R}(X)$.
(ii) Les polynômes orthogonaux ne sont pas orthonormés en général, mais la famille $\left(\frac{L_n}{\|L_n\|_E}\right)_{n \in \mathbb{N}}$ est orthonormée, donc forme une base orthonormée de $\mathbb{R}(X)$.

Lem: Soit $n \in \mathbb{N}^*$. Le polynôme orthogonal L_n satisfait :

$$\forall Q \in \mathbb{R}_{n-1}[X], \langle L_n, Q \rangle_E = 0.$$

Dém: Soit $Q \in \mathbb{R}_{n-1}[X]$. Comme $\forall 0 \leq h \leq n-1$, $d^0(L_h) = h$, $(L_h)_{0 \leq h \leq n-1}$ est une base de $\mathbb{R}_{n-1}[X] \Rightarrow$ il existe des nombres $(\lambda_h)_{0 \leq h \leq n-1} \in \mathbb{R}^n$ tels que $Q = \sum_{h=0}^{n-1} \lambda_h L_h$
 $\Rightarrow \langle L_n, Q \rangle_E = \sum_{h=0}^{n-1} \lambda_h \langle L_n, L_h \rangle_E = 0$.

Lem: Soit $\forall n \geq 2$, $\lambda_n = \frac{\langle X L_{n-1}, L_{n-1} \rangle_E}{\|L_{n-1}\|_E^2}$ et $\mu_n = \frac{\|L_{n-1}\|_E^2}{\|L_{n-2}\|_E^2}$. Les polynômes orthogonaux $(L_n)_{n \in \mathbb{N}}$ satisfont la relation de récurrence :

$$\forall n \geq 2, L_n = (X - \lambda_n) L_{n-1} - \mu_n L_{n-2}.$$

Dém: Soit $n \geq 2$. Rappelons que $(L_h)_{0 \leq h \leq n}$ forme une base de $\mathbb{R}_n[X]$. Comme $d^0(L_{n-1}) = n-1$, $d^0(X L_{n-1}) = n \Rightarrow \exists (\alpha_h)_{0 \leq h \leq n} \in \mathbb{R}^{n+1}$ t.q. $X L_{n-1} = \sum_{h=0}^n \alpha_h L_h$; sachant que $X L_{n-1}$ et L_n sont unitaires, et que $\forall 0 \leq h \leq n-1$, $d^0(L_h) = h < n$, $\alpha_n = 1$;
 $\forall 0 \leq h \leq n-1$, $\alpha_h \|L_h\|_E^2 = \langle X L_{n-1}, L_h \rangle_E = \langle L_{n-1}, X L_h \rangle_E$; si $0 \leq h \leq n-3$, alors, $d^0(X L_h) = h+1 \leq n-1 \Rightarrow \langle L_{n-1}, X L_h \rangle_E = 0 \Rightarrow \alpha_h = 0$; si $h = n-1$, alors $\alpha_{n-1} = \lambda_n = \frac{\langle X L_{n-1}, L_{n-1} \rangle_E}{\|L_{n-1}\|_E^2}$; si $h = n-2$, alors $X L_{n-2}$ est unitaire et de degré $n-1 \Rightarrow \exists Q \in \mathbb{R}_{n-1}[X]$ t.q. $X L_{n-2} = L_{n-1} + Q \Rightarrow \langle L_{n-1}, X L_{n-2} \rangle_E = \|L_{n-1}\|_E^2 \Rightarrow \alpha_{n-2} = \mu_n = \frac{\|L_{n-1}\|_E^2}{\|L_{n-2}\|_E^2}$.

Dém: Cette formule de récurrence permet un calcul rapide des polynômes orthogonaux $(L_n)_{n \in \mathbb{N}}$ lorsque les coefficients $(\lambda_n)_{n \geq 2}$ et $(\mu_n)_{n \geq 2}$ sont explicites, ce qui est le cas des exemples classiques ci-dessous. Il est aussi possible de calculer de façon approchée ces coefficients par des méthodes de quadrature adaptées.

Ex: (i) Pour $I = [-1, 1]$ et $\forall x \in I$, $w(x) = 1$, les polynômes orthogonaux $(L_n)_{n \in \mathbb{N}}$ ont

appelés polynômes de Legendre.

(ii) Pour $I =]-1, 1[$ et $\forall x \in I$, $w(x) = \frac{1}{\sqrt{1-x^2}}$, les polynômes orthogonaux $(P_n)_{n \in \mathbb{N}}$ sont appelés polynômes de Tchebychev.

(iii) Pour $I = [0, +\infty[$ et $\forall x \in I$, $w(x) = e^{-x}$, les polynômes orthogonaux $(P_n)_{n \in \mathbb{N}}$ sont appelés polynômes de Laguerre.

(iv) Pour $I = \mathbb{R}$ et $\forall x \in I$, $w(x) = e^{-x^2}$, les polynômes orthogonaux $(P_n)_{n \in \mathbb{N}}$ sont appelés polynômes de Hermite.

Les racines des polynômes orthogonaux $(P_n)_{n \in \mathbb{N}}$ sont à la base des méthodes de quadrature de Gauss pour le calcul approché d'intégrales. Les racines du polynôme P_n sont (pour $n \geq 1$) sont réelles, distinctes et intérieures à l'intervalle I .

Thm: Soit J l'intervalle intérieur à I . Quel que soit l'entier $n \geq 1$, il existe des nombres $(x_i)_{1 \leq i \leq n} \in J^n$, deux à deux distincts, tels que:

$$\forall 1 \leq i \leq n, P_n(x_i) = 0$$

Autrement dit, le polynôme P_n est scindé à racines simples dans l'intervalle J .

Dém: Soit $(x_1, \dots, x_h) \in J^h$ les racines de multiplicité impaire du polynôme P_n dans l'intervalle J . Soit $Q = \prod_{j=1}^h (x - x_j)$ (avec la convention que $Q = 1$ si le polynôme P_n n'a pas de racines de multiplicité impaire dans l'intervalle J).

Toutes les racines du polynôme $P_n Q$ dans l'intervalle J sont de multiplicité paire; comme $P_n Q \neq 0$ et que $P_n Q$ a un signe constant sur J , $\langle P_n, Q \rangle_E = \int_J P_n(x) Q(x) w(x) dx \neq 0$; sachant que $d^\circ(P_n) = n$, $d^\circ(Q) = h \leq n$; si $h < n$, alors, $\langle P_n, Q \rangle_E = 0$, ce qui est absurde! $\Rightarrow d^\circ(Q) = n \Rightarrow P_n$ a n racines de multiplicité impaire dans l'intervalle J ; comme la somme de ces multiplicités est égale à $d^\circ(P_n) = n$, ces racines sont toutes simples, d'où le théorème.

Ex: Pour $I =]-1, 1[$ et $\forall x \in I$, $w(x) = \frac{1}{\sqrt{1-x^2}}$, les polynômes de Tchebychev $(P_n)_{n \geq 0}$ ont (pour $n \geq 1$) les n racines réelles:

$$\forall 1 \leq j \leq n, x_j = \cos\left(\frac{(2j-1)\pi}{2n}\right).$$

b. Application au polynôme de meilleure approximation quadratique

Nous revenons au cas où I désigne un segment $[a, b]$ (non réduit à un point). Nous considérons une fonction poids $w \in \mathcal{C}^0([a, b], \mathbb{R})$ telle que:

$$\forall x \in [a, b], w(x) > 0,$$

et nous rappelons que l'espace E associé est défini par:

$$E = \{ f \in \mathcal{C}^0([a, b], \mathbb{R}) \text{ telle que } \int_a^b f(x)^2 w(x) dx < +\infty \},$$

et qu'il s'agit d'un espace euclidien ou pré-hilbertien pour le produit scalaire:

$$\forall (f, g) \in E^2, \langle f, g \rangle_E = \int_a^b f(x)g(x)w(x)dx.$$

Observons aussi que l'espace des fonctions polynômes est, puisque w est continue sur $[a, b]$, un sous-espace de E .

Blm: Considérons la suite des polynômes orthogonaux $(P_n)_{n \in \mathbb{N}}$ associés à l'espace E . Soit f la fonction f de E , et l'entier $N \geq 0$, le polynôme Q_N de meilleure approximation quadratique de la fonction f défini par:

$$\|f - Q_N\|_E = \min_{Q \in \mathbb{R}_N[X]} \|f - Q\|_E,$$

est donné de manière unique par la formule:

$$Q_N = \sum_{k=0}^N \frac{\langle f, P_k \rangle_E}{\|P_k\|_E^2} P_k.$$

De plus, il est caractérisé par le fait d'être l'unique polynôme $Q \in \mathbb{R}_N[X]$

tel que:

$$\forall R \in \mathbb{R}_N[X], \langle Q, R \rangle_E = \langle f, R \rangle_E.$$

Dém: Soit $\tilde{Q}_N = \sum_{k=0}^N \frac{\langle f, P_k \rangle_E}{\|P_k\|_E^2} P_k$.

(i) Caractérisation de $\tilde{Q}_N$: $\forall 0 \leq k \leq N, \langle \tilde{Q}_N, P_k \rangle_E = \langle f, P_k \rangle_E$ par orthogonalité des polynômes $(P_k)_{0 \leq k \leq N}$; comme $(P_k)_{0 \leq k \leq N}$ est une base de $\mathbb{R}_N[X]$, par linéarité, $\forall Q \in \mathbb{R}_N[X], \langle \tilde{Q}_N, Q \rangle_E = \langle f, Q \rangle_E$; s'il existe $\check{Q}_N \in \mathbb{R}_N[X]$ tel que $\forall Q \in \mathbb{R}_N[X], \langle \check{Q}_N, Q \rangle_E = \langle f, Q \rangle_E$, alors $\langle \tilde{Q}_N - \check{Q}_N, Q \rangle_E = 0 \Rightarrow \|\tilde{Q}_N - \check{Q}_N\|_E^2 = 0 \Rightarrow \tilde{Q}_N = \check{Q}_N$.

(ii) Minimalité de $\tilde{Q}_N$: $\forall Q \in \mathbb{R}_N[X], \|f - Q\|_E^2 = \|f - \tilde{Q}_N + \tilde{Q}_N - Q\|_E^2 = \|f - \tilde{Q}_N\|_E^2 + 2 \langle f - \tilde{Q}_N, \tilde{Q}_N - Q \rangle_E + \|\tilde{Q}_N - Q\|_E^2$; comme $\tilde{Q}_N - Q \in \mathbb{R}_N[X], \langle f, \tilde{Q}_N - Q \rangle_E = \langle \tilde{Q}_N, \tilde{Q}_N - Q \rangle_E \Rightarrow \|f - Q\|_E^2 = \|f - \tilde{Q}_N\|_E^2 + \|\tilde{Q}_N - Q\|_E^2 \geq \|f - \tilde{Q}_N\|_E^2 \Rightarrow$

$\tilde{Q}_N$ est un polynôme de meilleure approximation quadratique de la fonction f ; si Q_N est un autre polynôme de meilleure approximation quadratique de f , alors

$$\|f - Q_N\|_E^2 = \|f - \tilde{Q}_N\|_E^2 + \|\tilde{Q}_N - Q_N\|_E^2 = \|f - \tilde{Q}_N\|_E^2 \Rightarrow \|Q_N - \tilde{Q}_N\|_E^2 = 0$$

$\Rightarrow Q_N = \tilde{Q}_N$, d'où l'unicité de ce polynôme de meilleure approximation quadratique.

Dém: La formule $Q_N = \sum_{k=0}^N \frac{\langle f, P_k \rangle_E}{\|P_k\|_E^2} P_k$ signifie que le polynôme de meilleure approximation quadratique de la fonction f est sa projection orthogonale sur le sous-espace $\mathcal{R}_N(x)$. En particulier, l'application qui, à la fonction f associe son polynôme de meilleure approximation quadratique Q_N , est linéaire, ce qui n'est pas le cas en général.

Nous achèverons ce chapitre par le résultat de convergence des polynômes de meilleure approximation quadratique Q_N vers la fonction f associée lorsque $N \rightarrow +\infty$.

Chm: Soit $f \in E$. Considérons la suite $(Q_N)_{N \in \mathbb{N}}$ des polynômes de meilleure approximation quadratique de la fonction f . Alors:

$$\|Q_N - f\|_E \xrightarrow{N \rightarrow +\infty} 0.$$

Dém: Nous invoquons le théorème de Weierstrass qui fournit une suite de polynômes $(L_N)_{N \in \mathbb{N}}$ telle que: $\forall N \in \mathbb{N}$, $\deg(L_N) = N$ et $\|L_N - f\|_{\infty} \xrightarrow{N \rightarrow +\infty} 0$. Nous avons alors: $\forall N \in \mathbb{N}$, $\|Q_N - f\|_E^2 \leq \|L_N - f\|_E^2 = \int_a^b (L_N(x) - f(x))^2 w(x) dx \leq \left(\int_a^b w(x) dx \right) \|L_N - f\|_{\infty}^2 \xrightarrow{N \rightarrow +\infty} 0$, d'où le théorème.

Dém: Ce théorème n'est plus nécessairement valable lorsque l'intervalle I n'est plus un segment.