Mathématiques assistées par ordinateur

Chapitre 7 : Approximation polynomiale

Michael Eisermann

Mat249, DLST L2S4, Année 2008-2009
www-fourier.ujf-grenoble.fr/˜eiserm/cours # mao
Document mis à jour le 6 juillet 2009

INSTITUT FOURIER

UNIVERSITE JOSEPH FOURIER
SCIENCES.TECHNOLOGIE.SANTÉ
GRENOBLE 1

Objectifs de ce chapitre

Ce chapitre initie aux questions d'approximation d'une fonction continue donnée par des polynômes. C'est une vaste théorie que nous n'esquisserons ici que superficiellement.

Par rapport à la norme uniforme, nous étudions l'interpolation de Lagrange, qui est analogue à l'approximation de Taylor. Dans les deux cas des phénomènes de non-convergence sont possibles et doivent être connus à titre d'avertissement.

Fort heureusement, le théorème de Weierstrass assure que toute fonction continue $f\colon [a,b] \to \mathbb{R}$ peut être uniformement approchée par des polynômes P_n, de sorte que $\|f - P\|_\infty \to 0$ pour $n \to \infty$. Nous énonçons ici la formulation constructive due à Bernstein.

Algorithmiquement, la norme quadratique s'avère plus avantageuse : elle provient d'un produit scalaire et permet des calculs très efficaces. Nous mentionnons ici la riche théorie des polynômes orthogonaux.

Sommaire

Retour sur l'interpolation de Lagrange

Théorème (interpolation de Lagrange)

Étant donnés des points distincts $x_0, \dots, x_n \in \mathbb{R}$ et des valeurs arbitraires $y_0, \dots, y_n \in \mathbb{R}$, il existe un unique polynôme $P \in \mathbb{R}[X]$ de degré $\leq n$ vérifiant $P(x_k) = y_k$ pour tout $k = 0, \dots, n$, à savoir

$$P = \sum_{k=0}^{n} y_k \prod_{j \neq k} \frac{X - x_j}{x_k - x_j}.$$

*On l'appelle le **polynôme interpolateur de Lagrange** donné par $x_0, \dots, x_n$ et $y_0, \dots, y_n$, ou passant par $(x_0, y_0), \dots, (x_n, y_n)$.* $\square$

Ce résultat est valable sur tout corps : même formule, même preuve.

Calcul pratique. Nous souhaitons une méthode de calcul efficace.

Approximation. Étant donnés $n+1$ points distincts $x_0, \dots, x_n \in [a, b]$, on peut approcher toute fonction $f \colon [a, b] \to \mathbb{R}$ par l'unique polynôme P de degré $\leq n$ donné par $x_0, \dots, x_n$ et $y_k = f(x_k)$ pour $k = 0, \dots, n$.

Par construction on a $P(x) = f(x)$ pour tout $x \in \{x_0, \dots, x_n\}$.
Pour $x \in [a, b]$ nous souhaitons majorer l'écart $|P(x) - f(x)|$.

Différences divisées : énoncé

Le polynôme interpolateur de f en $x_0, \dots, x_n$ peut s'écrire comme

$$P = a_0 \; + \; a_1(X - x_0) \; + \; \dots \; + \; a_n(X - x_0) \cdots (X - x_{n-1}).$$

Cette écriture permet de l'évaluer efficacement à la Horner :

$$P(x) = a_0 + (x - x_0)\big[a_1 + \dots (x - x_{n-2})\big[a_{n-1} + (x - x_{n-1})a_n\big] \dots\big].$$

Ceci ne nécessite que n soustractions, n additions, n multiplications.
Mais comment calculer efficacement les coefficients $a_0, a_1, \dots, a_n$?

On observe que $a_0 = f(x_0)$ et $a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$. Par récurrence on définit les ***différences divisées*** par $f[x_i] := f(x_i)$ puis

$$f[x_i, \dots, x_{i+k}] := \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}.$$

Théorème

Nous avons $a_k = f[x_0, \dots, x_k]$ *pour tout* $k = 0, \dots, n$.

Différences divisées : algorithme

$$
\begin{array}{cccccccc}
\text{étape 0} & & \text{étape 1} & & \text{étape 2} & \dots & \text{étape n} & \text{résultat} \\
f(x_0) & & & & & & & a_0 \\
f(x_1) & \overset{\searrow}{\rightarrow} & f[x_0, x_1] & & & & & a_1 \\
f(x_2) & \overset{\searrow}{\rightarrow} & f[x_1, x_2] & \overset{\searrow}{\rightarrow} & f[x_0, x_1, x_2] & & & a_2 \\
\vdots & & \vdots & & \vdots & & & \\
f(x_{n-2}) & \overset{\searrow}{\rightarrow} & f[x_{n-3}, x_{n-2}] & & & & & a_{n-2} \\
f(x_{n-1}) & \overset{\searrow}{\rightarrow} & f[x_{n-2}, x_{n-1}] & \overset{\searrow}{\rightarrow} & f[x_{n-3}, x_{n-2}, x_{n-1}] & & & a_{n-1} \\
f(x_n) & \overset{\searrow}{\rightarrow} & f[x_{n-1}, x_n] & \overset{\searrow}{\rightarrow} & f[x_{n-2}, x_{n-1}, x_n] & \overset{\searrow}{\rightarrow} & f[x_0, \dots, x_n] & a_n
\end{array}
$$

Algorithme 1 calcul des différences divisées

Entrée: les points $x_0, \dots, x_n$ et les valeurs $y_0, \dots, y_n$ où $y_k = f(x_k)$

Sortie: les coefficients $a_0, \dots, a_n$ comme spécifiés ci-dessus

pour k **de** 0 **à** n **faire** $a_k \leftarrow y_k$ **fin pour**
pour i **de** 1 **à** n **faire**
  pour k **de** n **à** i **faire** $a_k \leftarrow \frac{a_k - a_{k-1}}{x_k - x_{k-i}}$ **fin pour**
fin pour
retourner $a_0, \dots, a_n$

Différences divisées : démonstration

On va prouver la formule $a_k = f[x_0, \dots, x_k]$ par récurrence sur k.
L'énoncé est vrai au rang $k = 0$ car $a_0 = f(x_0) = f[x_0]$.

Supposons l'énoncé vrai au rang $k-1$ et prouvons-le au rang k.
Soit P_{k-1} le polynôme interpolateur de f en $x_0, \dots, x_{k-1}$ et soit Q_{k-1} le polynôme interpolateur de f en $x_1, \dots, x_k$. Alors

$$P_k = \frac{(X - x_0)Q_{k-1} - (X - x_k)P_{k-1}}{x_k - x_0}$$

vérifie $P_k(x_0) = f(x_0)$, ..., $P_k(x_k) = f(x_k)$. C'est donc le polynôme interpolateur de f en $x_0, \dots, x_k$. On en déduit que

$$\begin{aligned} a_k = \operatorname{dom} P_k &= \frac{\operatorname{dom} Q_{k-1} - \operatorname{dom} P_{k-1}}{x_k - x_0} \\ &= \frac{f[x_1, \dots, x_k] - f[x_0, \dots, x_{k-1}]}{x_k - x_0} = f[x_0, \dots, x_k]. \end{aligned}$$

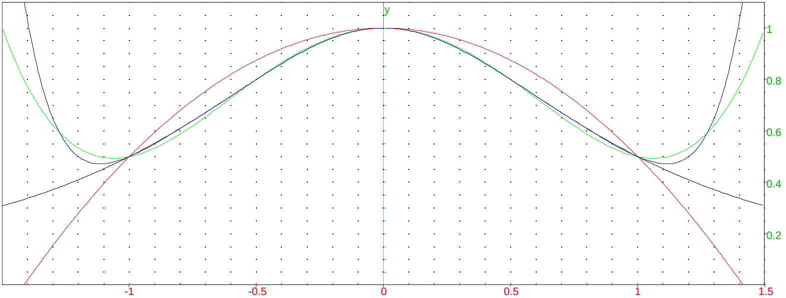
Ceci prouve la formule souhaitée. $\square$

Remarque. Puisque P_k est unique et $\operatorname{dom} P_k = f[x_0, \dots, x_k]$, cette valeur est indépendante de l'ordre des points $x_0, \dots, x_k$.

Interpolation et approximation : exemple

On considère la fonction $f\colon [-1,+1] \to \mathbb{R}$ donnée par $f(x) = \frac{1}{1+x^2}$.

La graphique trace f (noir) ainsi que trois polynômes interpolateurs :

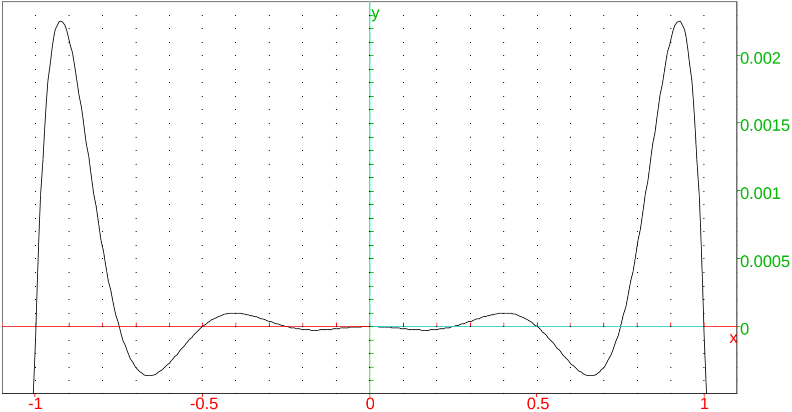


P_2 (rouge) interpole en $-1, 0, +1$,
P_4 (vert) interpole en $-1, -\frac{1}{2}, 0, +\frac{1}{2}, +1$,
P_8 (bleu) interpole en $-1, -\frac{3}{4}, -\frac{1}{2}, -\frac{1}{4}, 0, +\frac{1}{4}, +\frac{1}{2}, +\frac{3}{4}, +1$.

Si P_2 est encore loin, P_4 approche f déjà assez bien sur $[-1,+1]$.
Les graphes de f et P_8 semblent coïncider (pour cette résolution).
À noter toutefois que l'erreur explose en dehors de $[-1,+1]$.

Interpolation et approximation : exemple

Pour mieux visualiser l'écart entre f et P_8 il est beaucoup plus informatif d'afficher la différence $f - P_8$ pour pouvoir « zoomer » :



Graphiquement on voit que $|f - P_8| \leq 0.0025$ sur $[-1, +1]$.
On remarque que l'erreur est la plus grande proche des extrémités.
Dans la suite nous allons prouver et quantifier ces observations.

Écart entre une fonction et son polynôme interpolateur

Théorème

Soit $f\colon [a,b] \to \mathbb{R}$ *une fonction* $n+1$ *fois dérivable.*
Soient $x_0, \dots, x_n \in [a,b]$ *des points distincts de cet intervalle.*
Soit P *le polynôme interpolateur donné par les* x_k *et* $y_k = f(x_k)$*.*
Pour tout $x \in [a,b]$ *il existe* $\xi \in [a,b]$ *(qui dépend de* x*) tel que :*

$$f(x) - P(x) \quad = \quad \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot \prod_{k=0}^{n} (x - x_k).$$

Ainsi l'erreur $|f(x) - P(x)|$ dépend d'une majoration de $|f^{(n+1)}|$ mais aussi de la disposition des points x_j par rapport à x.

Par exemple, si les points $x_0, \dots, x_n$ sont ***équidistants***, le terme $|\prod_{k=0}^{n}(x - x_k)|$ est plus grand près du bord de I qu'au centre de I.

Un choix judicieux des points $x_0, \dots, x_n$, plus dense aux extrémités, donnera une meilleure approximation : ***points de Tchebychev***.

📖 J.-P. Demailly, *Analyse numérique et équations différentielles*, EDP Sciences, Grenoble 2006 (3e édition), chapitre 2

Démonstration du théorème

Pour $x \in \{x_0, \ldots, x_n\}$ l'égalité est vraie, indépendamment de ξ.

$$f(x) - P(x) \quad = \quad \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot \prod_{k=0}^{n} (x - x_k).$$

Nous fixons $x \notin \{x_0, \ldots, x_n\}$ et posons $C := \frac{f(x)-P(x)}{\prod_{k=0}^{n}(x-x_k)}$.
Considérons la fonction :

$$g(t) = f(t) - P(t) - C \prod_{k=0}^{n} (t - x_k).$$

Elle s'annule pour $t \in \{x_0, \ldots, x_n, x\}$, donc en au moins $n+2$ points.

Par conséquent g' s'annule en au moins $n+1$ points de $[a,b]$.

Ensuite g'' s'annule en au moins n points de $[a,b]$, etc. . .

Finalement $g^{(n+1)}$ s'annule au moins une fois dans $[a,b]$. Or,

$$g^{(n+1)}(t) = f^{(n+1)}(t) - C \cdot (n+1)!$$

On conclut qu'il existe $\xi \in [a,b]$ tel que $C = \frac{f^{(n+1)}(\xi)}{(n+1)!}$. $\square$

Application aux points équidistants

Corollaire

Soit $f\colon [a,b] \to \mathbb{R}$ *une fonction* $n+1$ *fois dérivable. Soit* P *le polynôme interpolateur de* f *en les* $n+1$ *points équidistants* $x_k = a + kh$ *où* $h = \frac{b-a}{n}$ *et* $k = 0, \dots, n$*. Alors pour* $x \in [a,b]$ *on a*

$$\big|f(x) - P(x)\big| \quad \leq \quad \frac{h^{n+1}}{n+1} \cdot \max_{[a,b]} \big|f^{(n+1)}\big|.$$

Démonstration. Pour $x = a + th$ avec $t \in [0,n]$ nous avons

$$\prod_{k=0}^{n} (x - x_k) = h^{n+1} \prod_{k=0}^{n} (t-k).$$

La fonction $\pi(t) := |t(t-1)\cdots(t-n)|$ est symétrique, $\pi(n-t) = \pi(t)$. Pour $1 \leq t \leq n/2$ on voit que $\frac{\pi(t-1)}{\pi(t)} = \frac{n+1-t}{t} > 1$, d'où $\pi(t-1) > \pi(t)$. Par conséquent, π atteint son maximum dans $[0,1]$, d'où

$$\max_{[0,n]} \pi = \max_{[0,1]} \pi \leq n!$$

Ainsi le corollaire découle du théorème précédent. $\square$

Comparaison entre Taylor et Lagrange

Les approximations de Taylor et de Lagrange se ressemblent. Considérons une fonction $f\colon [a,b] \to \mathbb{R}$ de classe C^{n+1}.

Le polynôme de Taylor $T_n \in \mathbb{R}[X]_{\le n}$ est construit pour coïncider avec les valeurs $f(x_0), f'(x_0), \dots, f^{(n)}(x_0)$. On a la majoration d'erreur

$$\big|f(x) - T_n(x)\big| \le \frac{|x-x_0|^{n+1}}{(n+1)!} \max_{[a,b]}\big|f^{(n+1)}\big| \le \frac{(b-a)^{n+1}}{(n+1)!} \max_{[a,b]}\big|f^{(n+1)}\big|.$$

Cette majoration ne dit pas que $|f(x) - T_n(x)| \to 0$ pour $n \to \infty$, car $|f^{(n+1)}|$ peut exploser : problème du rayon de convergence !

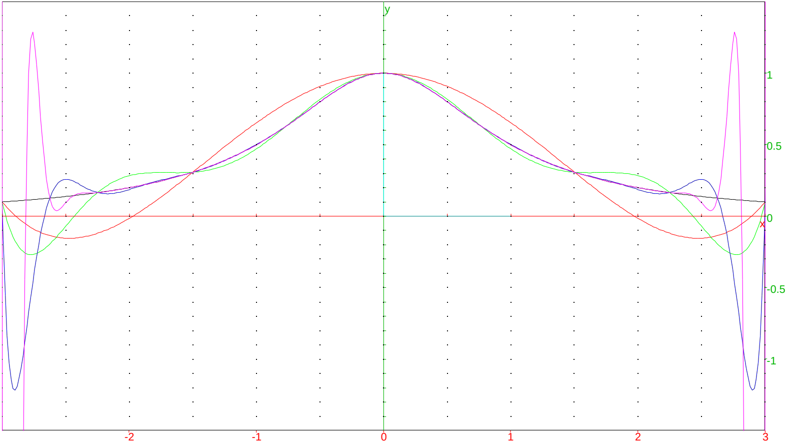
Le polynôme de Lagrange $L_n \in \mathbb{R}[X]_{\le n}$ est construit pour coïncider avec les valeurs $f(x_0), f(x_1), \dots, f(x_n)$. On a la majoration d'erreur

$$\big|f(x) - L_n(x)\big| \le \frac{\prod_{k=0}^{n}|x-x_k|}{(n+1)!} \max_{[a,b]}\big|f^{(n+1)}\big| \le \frac{(b-a)^{n+1}}{(n+1)!} \max_{[a,b]}\big|f^{(n+1)}\big|.$$

Cette majoration ne dit pas que $|f(x) - L_n(x)| \to 0$ pour $n \to \infty$, car $|f^{(n+1)}|$ peut exploser : phénomène de Runge !

Avertissement : phénomène de Runge

On considère $f \colon [-3, 3] \to \mathbb{R}$, $f(x) = \frac{1}{x^2+1}$. La graphique montre f et ses polynômes de Lagrange d'ordre 4, 8, 16, 32 à points équidistants.



Sur $[-2, +2]$ ces polynômes semblent bien approcher la fonction f. Vers le bord, par contre, ils oscillent de plus en plus violemment !

Théorème de Weierstrass, polynômes de Bernstein

Peut-on approcher toute fonction continue par des polynômes ?

Théorème (Weierstrass 1885)

Soit $f\colon [a,b] \to \mathbb{R}$ *une fonction continue sur un intervalle compact. Pour tout* $\varepsilon > 0$ *il existe un polynôme* P *tel que* $|f(x) - P(x)| \le \varepsilon$ *pour tout* $x \in [a,b]$*. Autrement dit, on assure que* $\|f - P\|_\infty \le \varepsilon$*.* □

Le théorème, tel qu'il est énoncé ici, n'établit que l'existence de P. Dans la pratique nous souhaitons une construction effective :

Théorème (Bernstein 1912)

Pour $n \in \mathbb{N}$ *et* $0 \le k \le n$ *le polynôme de base de Bernstein est*

$$B_n^k(x) := \tbinom{n}{k} x^k (1-x)^{n-k}.$$

Étant donnée une fonction continue $f\colon [0,1] \to \mathbb{R}$ *on pose*

$$P_n := \textstyle\sum_{k=0}^{n} f(k/n) B_n^k(x).$$

Alors on a convergence uniforme $\|f - P\|_\infty \to 0$ *pour* $n \to \infty$*.* □

Polynômes de Bernstein : illustration

On constate que

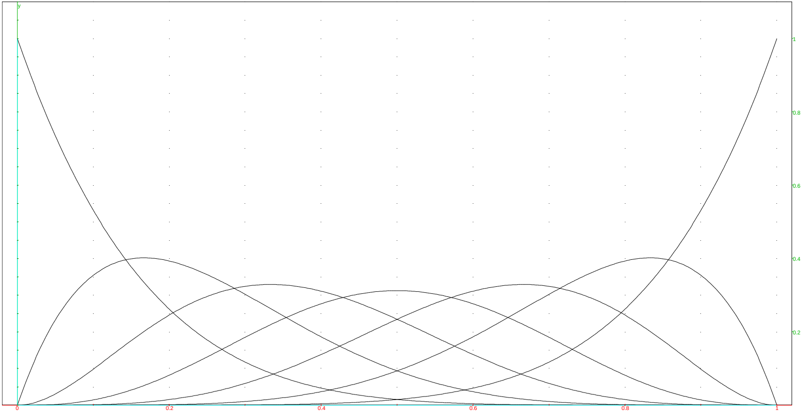
$$1 = (x + 1 - x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}.$$

Pour tout $n \in \mathbb{N}$ les polynômes de Bernstein

$$B_n^k(x) := \binom{n}{k} x^k (1-x)^{n-k}$$

où $k = 0, \ldots, n$ forment une base de $\mathbb{R}[X]_{\leq n}$.

Voici la graphique pour $n = 6$ et $k = 0, \ldots, 6$:



Approximations uniforme, en moyenne, et quadratique

Quand on approche $f\colon [a,b] \to \mathbb{R}$ par un polynôme P
on s'intéresse à l'erreur, c'est-à-dire à la distance de P à f.

Jusqu'ici nous avons regardé la distance uniforme :

$$\|f - P\|_\infty \quad := \quad \max_{x\in[a,b]} |f(x) - P(x)|.$$

On pourrait s'intéresser à la distance en moyenne :

$$\|f - P\|_1 \quad := \quad \frac{1}{b-a}\int_a^b |f(x) - P(x)| dx.$$

Ou bien la moyenne quadratique :

$$\|f - P\|_2 \quad := \quad \left(\frac{1}{b-a}\int_a^b |f(x) - P(x)|^2 dx\right)^{\frac{1}{2}}.$$

Chacune est utile, mais la dernière est particulièrement élégante.
Elle correspond à la distance euclidienne bien connue de $\mathbb{R}^n$.

Existence et unicité d'une meilleure approximation

Proposition

Soit $\|\cdot\|$ une norme sur l'espace $C([a,b])$ des fonctions continues $[a,b] \to \mathbb{R}$. Pour toute fonction $f \in C([a,b])$ il existe un unique polynôme $P_n \in \mathbb{R}[X]_{\le n}$ qui minimise la distance $\|f - P_n\|$.

Démonstration. (On ne montre ici que l'existence.)

La fonction $d\colon \mathbb{R}[X]_{\le n} \to \mathbb{R}$, $d(P) = \|f - P\|$ est continue.

Le polynôme 0 réalise une approximation avec $\|f - 0\| = \|f\|$.

Si $\|P\| > 2\|f\|$ alors $\|P - f\| \ge \|P\| - \|f\| > \|f\|$ est pire.

Il suffit donc de minimiser d sur la boule fermée

$$\bar{B}(0, 2\|f\|) = \{P \in \mathbb{R}[X]_{\le n} : \|P\| \le 2\|f\|\}.$$

Puisque $\mathbb{R}[X]_{\le n} \cong \mathbb{R}^{n+1}$ est de dimension finie, cette boule est compacte. La fonction d atteint donc son minimum. □

⚠ Après l'existence, nous souhaitons une construction effective. Or, en général il est difficile de déterminer la meilleure approximation.

◆ Ici la norme $\|\cdot\|_2$ se révélera particulièrement efficace.

Produit scalaire, norme, distance

Pour $f, g\colon [a,b] \to \mathbb{R}$ continues on pose $\langle f|g\rangle := \frac{1}{b-a}\int_a^b f(x)g(x)dx$.

C'est un ***produit scalaire*** dans le sens qu'il est

symétrique : $\langle f|g\rangle = \langle g|f\rangle,$

bilinéaire : $\langle f|\lambda g + \mu h\rangle = \lambda\langle f|g\rangle + \mu\langle f|h\rangle,$

positif-défini : $\langle f|f\rangle \geq 0,$ et $\langle f|f\rangle = 0 \Leftrightarrow f = 0.$

Ceci entraîne ***l'inégalité de Cauchy-Schwarz*** : $\langle f|g\rangle^2 \leq \langle f|f\rangle\langle g|g\rangle$.

Le produit scalaire induit une ***norme***, $\|f\|_2 := \sqrt{\langle f|f\rangle}$:

homogénéité : $\|\lambda f\|_2 = |\lambda| \cdot \|f\|_2,$

inégalité triangulaire : $\|f+g\|_2 \leq \|f\|_2 + \|g\|_2,$

positivité : $\|f\|_2 \geq 0,$ et $\|f\|_2 = 0 \Leftrightarrow f = 0.$

La norme induit une ***distance***, $\operatorname{dist}(f,g) := \|f-g\|_2$:

symétrie : $\operatorname{dist}(f,g) = \operatorname{dist}(g,f),$

inégalité triangulaire : $\operatorname{dist}(f,h) \leq \operatorname{dist}(f,g) + \operatorname{dist}(g,h),$

positivité : $\operatorname{dist}(f,g) \geq 0,$ et $\operatorname{dist}(f,g) = 0 \Leftrightarrow f = g.$

Polynômes orthogonaux : construction

On note $\mathbb{R}[X]_{\leq n}$ le $\mathbb{R}$-espace vectoriel des polynômes de degré $\leq n$.

Tout polynôme $P \in \mathbb{R}[X]_{\leq n}$ s'écrit de manière unique comme $P = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n$ où $a_0, a_1, a_2, \ldots, a_n \in \mathbb{R}$:
Les monômes $1, X, X^2, \ldots, X^n$ forment une ***base*** de $\mathbb{R}[X]_{\leq n}$

Muni du produit scalaire $\langle | \rangle$, l'espace $\mathbb{R}[X]_{\leq n}$ est un espace euclidien.
Malheureusement la base $1, X, X^2, \ldots, X$ n'est pas ***orthonormée***.
On peut en produire une selon l'algorithme de Gram–Schmidt :

Algorithme 2 orthonormalisation selon Gram–Schmidt

Entrée: une base $u_0, \ldots, u_n$ d'un espace euclidien E
Sortie: une base $v_0, \ldots, v_n$ qui soit orthonormée

---

pour k **de** 0 **à** n **faire**
  $v_k \leftarrow u_k$
  pour j **de** 0 **à** $k-1$ **faire** $v_k \leftarrow v_k - \langle u_k | v_j \rangle v_j$ **fin pour**
  $v_k \leftarrow v_k / \|v_k\|_2$
fin pour
retourner $v_0, \ldots, v_n$

Orthonormalisation selon Gram–Schmidt

Théorème (Gram–Schmidt)

Soit $u_0, \dots, u_n$ *une base d'un espace euclidien* E.
Pour $k = 0, \dots, n$ *on pose*

$$\tilde{v}_k := u_k - \sum_{j=0}^{k-1} \langle u_k | v_j \rangle v_j \qquad \textit{puis} \qquad v_k := \tilde{v}_k / \|\tilde{v}_k\|_2.$$

Alors $v_0, \dots, v_n$ *est une base orthonormée de* E*, c'est-à-dire*

$$\langle v_k | v_k \rangle = 1 \quad \textit{et} \quad \langle v_k | v_j \rangle = 0 \quad \textit{pour } j \neq k.$$

Démonstration. Par récurrence sur k on prouve que

$$\langle v_k | v_k \rangle = 1 \quad \text{et} \quad \langle v_k | v_j \rangle = 0 \quad \text{pour } j < k.$$

C'est vrai pour $k = 0$ car on a $v_0 = u_0 / \|u_0\|_2$.
Au rang k on trouve $\langle \tilde{v}_k | v_j \rangle = 0$ pour tout $j < k$.
Si l'on avait $\tilde{v}_k = 0$ alors $u_0, \dots, u_k$ seraient linéairement dépendants.
On peut donc passer à $v_k = \tilde{v}_k / \|\tilde{v}_k\|_2$ pour assurer $\langle v_k | v_k \rangle = 1$. $\square$

Meilleure approximation pour la distance $\|\cdot\|_2$

Proposition

Soit $P_0, \ldots, P_n$ une base de $\mathbb{R}[X]_{\leq n}$ orthonormée par rapport à $\langle | \rangle$.
Soit $f\colon [a,b] \to \mathbb{R}$ une fonction continue quelconque.
Posons $a_k := \langle f | P_k \rangle$ puis $P = \sum_{k=0}^{n} a_k P_k$. Alors P est l'unique polynôme de $\mathbb{R}[X]_{\leq n}$ qui minimise la distance $\|f - P\|_2$.

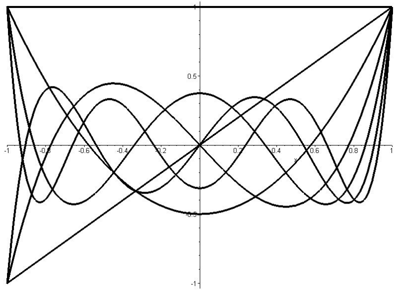
Démonstration.
Tout $Q \in \mathbb{R}[X]_{\leq n}$ s'écrit de manière unique comme $Q = \sum_{k=0}^{n} b_k P_k$.
Minimiser la distance $\|f - Q\|_2$ c'est minimiser

$$\begin{aligned} \|f - Q\|_2^2 &= \langle f - Q | f - Q \rangle = \langle f | f \rangle - 2\langle f | Q \rangle + \langle Q | Q \rangle \\ &= \|f\|_2 - 2\sum_{k=0}^{n} a_k b_k + \sum_{k=0}^{n} b_k^2 = \|f\|_2 - \sum_{k=0}^{n} a_k^2 + \sum_{k=0}^{n} (a_k - b_k)^2. \end{aligned}$$

Le minimum est atteint si et seulement si $b_k = a_k$ pour tout k. $\square$

Exemple : polynômes de Legendre

Orthonormalisons $1, X, \ldots, X^n$ sur $[-1, 1]$ par Gram–Schmidt.
À des facteurs constants près on obtient les polynômes de Legendre :



Exercice

Calculez puis affichez les premiers termes avec votre calculatrice et/ou votre logiciel préféré. Établir une récurrence

$$P_n = (X - \lambda_n)P_{n-1} - \mu_n P_{n-2}$$

avec des constantes réelles λ_n, μ_n que l'on déterminera.
Est-ce une suite de Sturm ? Que dire alors des racines de P_n ?