Capítol 4

Teoria d'aproximació

4.1 Introducció

El problema que considerarem en aquest capítol és el d'aproximar una funció f per una altra funció f^* que sigui fàcil de calcular, sobre un interval $[a, b]$. Veiem-ne dos exemples per tenir en compte que volem fer.

Exemple 4.1.1 *Dades amb errors.*

Suposem que coneixem uns certs valors d'una funció f amb un cert error:

x	x_0	x_1	$\cdots$	x_n
f	$f_0 + \epsilon_0$	$f_1 + \epsilon_1$	$\cdots$	$f_n + \epsilon_n$

Per exemple, pot donar-se el cas que sabem que f és un cert tipus de funció (recta, polinomi de grau 2, polinomi trigonomètric, etc.), però degut als errors no podem usar interpolació per a trobar la funció.

Exemple 4.1.2 *Aproximació d'una funció.*

Volem una representació més simple d'una determinada funció f. Per exemple, quina és la recta que la representa millor?

Notem que els dos exemples corresponen a dos tipus diferents d'aproximació: El primer exemple correspon al cas que anomenarem **discret** (nombre finit de dades) mentre que el segon correspon al cas **continu** (la dada és una funció en un interval $[a, b]$).

4.2 Preliminars: espais de funcions, normes i productes escalars

Treballarem en espais de funcions. Un espai de funcions és un conjunt de funcions amb estructura d'espai vectorial sobre $\mathbb{R}$ (o sobre $\mathbb{C}$). Aquests espais poden ser de dimensió finita o infinita. Per exemple:

- Espai dels polinomis a coeficients reals de grau ≤ 2 (dimensió 3).
- Espai de les funcions contínues de $[0,1]$ a $\mathbb{R}$ (dimensió infinita).
- Espai dels polinomis a coeficients reals (dimensió infinita).

Per a mesurar distàncies en aquests espais usarem una norma. Recordem el concepte de norma:

Definició 4.2.1 *Sigui V un espai vectorial sobre $\mathbb{K}$, on $\mathbb{K}=\mathbb{R}$ o $\mathbb{K}=\mathbb{C}$. Diem que $(V, \|\,\|)$ és un* ***espai normat*** *si existeix una aplicació $\|\,\| : V \to \mathbb{R}^+$, que anomenem* ***norma****, tal que*

a) $\|v\| = 0$ *sii* $v = 0$.

b) $\|u+v\| \leq \|u\| + \|v\|$.

c) $\|\lambda v\| = |\lambda|\|v\|$, *per a tot* $\lambda \in \mathbb{K}$ *i* $v \in V$.

Nota 4.2.1 *Notem que $|\|u\| - \|v\|| \leq \|u-v\|$, i que si definim $d(u,v) = \|u-v\|$, d ens defineix una distància en V, i per tant (V,d) és un espai mètric.*

Veiem-ne uns quants exemples en el cas contínuo i en el cas discret:

a) Cas contínuo:

(a) Definim
$$V = \mathcal{C}^0([a,b]) = \{f : [a,b] \to \mathbb{R} \mid f \text{ és contínua}\},$$
amb la norma
$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)| < +\infty.$$

Proposició 4.2.1 $(\mathcal{C}^0([a,b]), \|\cdot\|_\infty)$ *és un espai normat.*

Demostració:

i. $\|f\|_\infty \geq 0$.
ii. $\|f\|_\infty = 0$ sii $\max_{x\in[a,b]} |f(x)| = 0$ sii $f(x) = 0$, $\forall\, x \in [a,b]$, sii $f \equiv 0$.
iii. $|\lambda f\|_\infty = \max_{x\in[a,b]} |\lambda f(x)| = |\lambda| \max_{x\in[a,b]} |f(x)| = |\lambda|\|f\|_\infty$.
iv. $\|f+g\|_\infty = \max_{x\in[a,b]} |f(x)+g(x)| \leq \max_{x\in[a,b]}(|f(x)|+|g(x)|) \leq \max_{x\in[a,b]} |f(x)| + \max_{x\in[a,b]} |g(x)| = \|f\|_\infty + \|g\|_\infty$.

□

Aquesta norma s'anomena norma del màxim, norma infinit, o a vegades sup-norm (norma del suprem).

(b) Un altre exemple és $(\mathcal{C}^0([a,b]), \|\cdot\|_2)$, on
$$\|f\|_2 = \left[\int_a^b |f(x)|^2\, dx\right]^{1/2}.$$

Aquesta s'anomena norma euclidana o norma L_2.

(c) També tenim els espais normats $(\mathcal{C}^0([a,b]), \|\cdot\|_p)$, $p \geq 1$, on

$$\|f\|_p = \left[\int_a^b |f(x)|^p \, dx\right]^{1/p}.$$

(d) Un altre espai normat rellevant és $(\mathcal{C}^0([a,b]), \|\cdot\|_{2,w})$, on

$$\|f\|_{2,w} = \left[\int_a^b |f(x)|^2 w(x) \, dx\right]^{1/p},$$

i $w : [a,b] \to \mathbb{R}$ és una funció contínua i positiva. A vegades aquesta norma s'anomena **norma ponderada**. També podem definir

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

b) Cas discret:

Podem considerar l'espai vectorial de polinomis de grau més petit o igual que m, $V = \mathbb{R}_m[x]$. Per a definim una norma en aquest espai vectorial, escollim uns valors de $x : x_0, \ldots, x_m$. Donada una funció $q \in \mathbb{R}_m[x]$, si $q_i = q(x_i)$, podem definir les normes:

(a) Norma del suprem:

$$\|q\|_\infty = \max_{0\leq i\leq m} |q_i|,$$

(b) Norma sub-p, $p \geq 1$:

$$\|q\|_p = \left[\sum_{i=0}^m |q_i|^p\right]^{1/p},$$

Notem que si tenim una taula de valors d'una funció f:

$$\begin{array}{c|ccc} x & x_0 & \cdots & x_m \\ \hline f & f_0 & \cdots & f_m \end{array},$$

li podem associar de manera unívoca el polinomi interpolador $q \in \mathbb{R}_m[x]$, tal que $q(x_i) = f_i$, $i = 0, \ldots, m$. Això ens permetrà interpretar el problema d'aproximació discreta de l'exemple 4.1.1 com un problema d'aproximació en un espai de funcions de dimensió finita.

Algunes de les normes que hem definit provenen de productes escalars. A continuació, definirem productes escalars en general i els espais associats:

Definició 4.2.2 *Sigui V un espai vectorial sobre $\mathbb{K}$. Diem que $(E, \langle,\rangle)$ és un **espai prehilbertià** si l'aplicació $\langle,\rangle : V \times V \to \mathbb{K}$, que anomenem **producte escalar**, satisfà, si $f, g, h \in V$ i $\lambda \in \mathbb{K}$,*

a) $\langle f,h\rangle = \overline{\langle h,f\rangle}$, *si* $\mathbb{K} = \mathbb{C}$ *i* $\langle f,h\rangle = \langle h,f\rangle$, *si* $\mathbb{K} = \mathbb{R}$.

b) $\langle f+g,h\rangle = \langle f,h\rangle + \langle g,h\rangle$,

c) $\langle \alpha f,g\rangle = \alpha\langle f,g\rangle$, *d)* $\langle f, f\rangle > 0$ *si* $f \neq 0$,

on $f, g, h \in V$ *i* $\alpha \in \mathbb{K}$.

Nota 4.2.2 *De la definició es dedueix que* $\langle f, g+h\rangle = \langle f, g\rangle + \langle f, h\rangle$ *i* $\langle f, \lambda g\rangle = \overline{\lambda}\langle f, g\rangle$ *(*$\mathbb{K} = \mathbb{C}$*) o* $\langle f, \lambda g\rangle = \lambda\langle f, g\rangle$ *(*$\mathbb{K} = \mathbb{R}$*).*

La propietat més rellevant dels espais prehilbertians és que també són espais normats:

Proposició 4.2.2 *Sigui* $(V, \langle,\rangle)$ *un espai prehilbertià sobre* $\mathbb{K}$. *Si definim* $\|f\| = \sqrt{\langle f, f\rangle}$, *per a tot* $f \in V$, *llavors* $(V, \|\cdot\|)$ *és un espai normat sobre* $\mathbb{K}$.

Demostració:
Ho demostrarem per $\mathbb{K} = \mathbb{C}$. La demostració en el cas $\mathbb{K} = \mathbb{R}$ es dedueix fàcilment de l'anterior.

- Si $f = 0$ llavors $\|f\|^2 = \langle 0, 0\rangle = 0$.
- Si $\|f\| = 0$ llavors $\langle f, f\rangle = 0$ i per tant $f = 0$.
- $$\|\lambda f\|^2 = \langle \lambda f, \lambda f\rangle = \lambda\langle f, \lambda f\rangle = |\lambda|^2\langle f, f\rangle = |\lambda|^2\|f\|^2.$$
- Desigualtat triangular. Per demostrar-la usarem

Lema 4.2.1 *(desigualtat de Cauchy-Schwarz) Per a tot* $f, g \in V$, $|\langle f, g\rangle| \leq \|f\| \cdot \|g\|$.

Demostració:

Si $f = 0$ o $g = 0$ és trivial. Suposem doncs que $f \neq 0$ i $g \neq 0$. Sigui $\lambda \in \mathbb{C}$. Llavors

$$0 \leq \langle f - \lambda g, f - \lambda g\rangle = \langle f, f\rangle - \overline{\lambda}\langle f, g\rangle - \lambda\overline{\langle f, g\rangle} + |\lambda|^2\langle g, g\rangle.$$

Si agafem

$$\lambda = \frac{\langle f, g\rangle}{\langle g, g\rangle},$$

obtenim

$$0 \leq \langle f, f\rangle - \frac{|\langle f, g\rangle|^2}{\langle g, g\rangle},$$

d'on deduïm el lema. □

Usant el lema anterior:

$$\|f+g\|^2 = \langle f+g, f+g\rangle = \langle f, f\rangle + \langle f, g\rangle + \overline{\langle f, g\rangle} + \langle g, g\rangle \leq$$

$$\|f\|^2 + \|g\|^2 + 2|\langle f, g\rangle| \leq \|f\|^2 + \|g\|^2 + 2\|f\|\,\|g\| = (\|f\| + \|g\|)^2,$$

d'on $\|f+g\| \leq \|f\| + \|g\|$. □

Dos espais importants amb producte escalar són: $\mathbb{K}^n$ amb

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i},$$

(si $\mathbb{K} = \mathbb{C}$) o

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i,$$

si $\mathbb{K} = \mathbb{R}$, i $\mathcal{C}^0_{\mathbb{K}}([a, b])$ l'espai de funcions contínues de $[a, b]$ en $\mathbb{K}$ amb

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}w(x)\, dx,$$

si $\mathbb{K} = \mathbb{C}$ o

$$\langle f, g\rangle = \int_a^b f(x)g(x)w(x)\, dx,$$

si $\mathbb{K} = \mathbb{R}$, on w és una funció contínua i positiva fixada.

4.3 El problema general d'aproximació

La majoria de mètodes d'aproximació es basen en minimitzar alguna norma o seminorma de la funció error $\|f^* - f\|$, on f és la funció que volem aproximar, i que normalment té una forma determinada i f^* és l'aproximació, que pertany a un conjunt de funcions que habitualment comparteix alguna característica amb la funció f. Notem que la millor aproximació amb una norma no serà en general la millor aproximació amb una altra norma diferent. Per tant, és important tenir un criteri per decidir quina norma volem usar, o, el que és el mateix, saber quina mena d'aproximació volem. En aquest sentit, l'us de funcions pes pot ser rellevant.

De manera precisa, podem definir millor aproximació de la manera següent:

Definició 4.3.1 *Sigui $(V, \|\ \|)$ un espai normat i $T \subset V$ un subconjunt arbitrari. Donat un element $v \in V$ diem que $\tilde{u} \in T$ és una **millor aproximació** de v des de T si $\|v - \tilde{u}\| \leq \|v - u\|$, per a tot $u \in T$.*

Exemple 4.3.1 *Siguin $V = \mathbb{R}^2$, $\|\ \| = \|\ \|_2$ i $T = \{x \in V : \|x\| = 1\}$. En aquest cas, per a cada $v \in V$ existeix una única millor aproximació des de T. De la mateixa manera, si $T = \{x \in V : x^T h = 0\}$, on $h \in \mathbb{R}^2$ és un vector fixat, llavors també existeix una única millor aproximació per a tot $v \in V$.*

Exemple 4.3.2 *Sigui $T = \{u \in V |\, u(x) = e^{\beta x},\ \beta > 0\}$ un subconjunt de $(C([0, 1]), \|\ \|_\infty)$ i sigui $v : [0, 1] \to \mathbb{R}$ tal que $v(x) = 1/2$, per a tot $x \in [0, 1]$. Suposem que $\tilde{u} \in T$ és una millor aproximació de v des de T. Llavors $\tilde{u} = e^{\tilde{\beta}x}$, amb $\tilde{\beta} > 0$. Tenim que per tot $u \in T$, $u(x) = e^{\beta x}$ i*

$$\|v - u\|_\infty = \max_{0\leq x\leq 1} \left|\frac{1}{2} - e^{\beta x}\right| = e^\beta - \frac{1}{2} > \frac{1}{2},$$

però $\inf_{\beta>0} e^\beta - \frac{1}{2} = \frac{1}{2}$, el que implica que no existeix cap millor aproximació, ja que $1 \notin T$.

Observem que si $T \subset V$ i $\tilde{u} \in T$ és una millor aproximació de v llavors $\|v-\tilde{u}\| = \inf_{u\in T} \|v-u\| = \min_{u\in T} \|v-u\|$.

Definició 4.3.2 *El valor $E_T(v) = \inf_{u\in T} \|v-u\|$ s'anomena la* ***desviació minimal*** *de l'element v del conjunt T.*

Dels exemples anteriors deduïm que la millor aproximació no té perquè existir. En el cas d'existir, pot ser que no sigui única.

Nosaltres considererarem només el cas en el que el conjunt T és un subspai de dimensió finita. En canvi, l'espai V podrà tenir tenir dimensió finita o infinita. En aquest cas podem formular el problema de la millor aproximació de la següent manera:

Suposem que tenim un espai normat $(E, \|\cdot\|)$ i un subespai vectorial de E de dimensió finita, que anomenarem E^*. Sigui $\{\varphi_0, \ldots, \varphi_n\}$ una base de E^*. Aleshores el problema de trobar millor aproximació a $f \in E$ consisteix en trobar constants c_j, $j = 0, \ldots, n$ tals que $\| \sum_{j=0}^{n} c_j\varphi_j - f\|$ és el mínim possible, per a qualsevol altre elecció de les constants.

4.4 El problema de la millor aproximació en espais prehilbertians

Suposarem en aquesta secció que $(E, \langle, \rangle)$ és un espai prehilbertià i que E^* és un subespai vectorial de E de dimensió finita. Si no es diu el contrari, E és un espai prehilbertià real.

Definició 4.4.1 *Dues funcions $f, g \in E$ s'anomenen* ***ortogonals*** *si $\langle f, g\rangle = 0$. Un conjunt de funcions $\{\varphi_0, \varphi_1, \ldots, \varphi_n\}$ s'anomena un* ***sistema ortogonal*** *si $\varphi_i \neq 0$, $\forall\, i$ i $\langle \varphi_i, \varphi_j\rangle = 0$ per a tot $i \neq j$. Si, a més, $\|\varphi_i\| = 1$, $\forall\, i$, diem que és un* ***sistema ortonormal****. Anomenem al conjunt de funcions base ortogonal (o ortonormal) de l'espai que generen.*

Proposició 4.4.1 *Si $\{\varphi_0, \varphi_1, \ldots, \varphi_n\}$ és un sistema ortogonal llavors els vectors $\varphi_0, \varphi_1, \ldots, \varphi_n$ són linealment independents.*

Demostració:

Siguin $c_i \in \mathbb{R}$, $i = 0, \ldots, n$ tals que

$$c_0\varphi_0 + c_1\varphi_1 + \cdots + c_n\varphi_n = 0.$$

Això implica que

$$\langle c_0\varphi_0 + c_1\varphi_1 + \cdots + c_n\varphi_n, \varphi_j\rangle = c_j\langle \varphi_j, \varphi_j\rangle = 0,$$

i per tant $c_i = 0$, $i = 0, \ldots, n$, el que implica la independència lineal. $\square$

Teorema 4.4.1 *(de Pitàgoras). Si $\langle f, g\rangle = 0$, llavors*

$$\|f+g\|^2 = \|f\|^2 + \|g\|^2.$$

Demostració:

$$\|f+g\|^2 = \langle f+g, f+g\rangle = \langle f,f\rangle + 2\langle f,g\rangle + \langle g,g\rangle = \langle f,f\rangle + \langle g,g\rangle = \|f\|^2 + \|g\|^2.$$

□

Es pot donar una versió més general:

Teorema 4.4.2 *Si $\{\varphi_0, \varphi_1, \ldots, \varphi_n\}$ és un sistema ortogonal, llavors:*

$$\left\|\sum_{i=0}^{n} c_i\varphi_i\right\|^2 = \sum_{i=0}^{n} c_i^2\|\varphi_i\|^2.$$

Tornem al problema d'aproximació de partida:

Teorema 4.4.3 *(de la projecció ortogonal) Sigui E un espai vectorial, $\langle\cdot,\cdot\rangle$ un producte escalar, E^* un subespai vectorial de E, i $f \in E$. Llavors $f^* \in E^*$ és una millor aproximació a f per E^* sii*

$$\langle f - f^*, g\rangle = 0, \quad \forall\, g \in E^*.$$

A més, en el cas d'existir una millor aproximació, és única.

Demostració:

- Si $\langle f - f^*, g\rangle = 0$, per a tot $g \in E^*$ llavors

  $$\|f-g\|^2 = \|f - f^* + f^* - g\|^2 = \|f-f^*\|^2 + \|f^*-g\|^2,$$

  on hem usat el teorema de Pitàgoras, tenint en compte que $f^* - g \in E^*$ si $g \in E^*$. Per tant,

  $$\|f-f^*\| \le \|f-g\|, \quad \forall\, g \in E^*.$$

- Sigui f^* una millor aproximació i suposem que existeix $g^* \in E^*$ tal que $\langle f - f^*, g^*\rangle = c \neq 0$. Sempre podem suposar que $\|g^*\| = 1$, ja que si no és cert, podem substituir g^* per $g^*/\|g^*\|$. Sigui

  $$h = f^* + cg^* \in E^*.$$

  Llavors

  $$\|f-h\|^2 = \|f-f^*-cg^*\|^2 = \langle f-f^*-cg^*, f-f^*-cg^*\rangle =$$
  $$\langle f-f^*, f-f^*\rangle - 2c\langle g^*, f-f^*\rangle + c^2\langle g^*, g^*\rangle = \|f-f^*\|^2 - 2c^2 + c^2 = \|f-f^*\|^2 - c^2.$$

  Per tant, $\|f-h\| < \|f-f^*\|$ el que contradiu que f^* sigui una millor aproximació.

  Per a demostrar la unicitat, suposem que $f^*, g^* \in E^*$ són dues millors aproximacions. Llavors, $\|f-f^*\| = \|f-g^*\|$ i, com abans,

  $$\|f-g^*\|^2 = \|f-f^*+f^*-g^*\|^2 = \|f-f^*\|^2 + \|f^*-g^*\|^2.$$

  Per tant, $\|f^*-g^*\| = 0$, el que implica que $f^* = g^*$. □

□

Nota 4.4.1 *Notem que la dimensió de E^* no té perquè ser finita, però el teorema no demostra l'existència de la millor aproximació.*

Corol·lari 4.4.1 *L'error $\|f - f^*\|$ satisfà $\|f - f^*\|^2 = \|f\|^2 - \|f^*\|^2$.*

Demostració:

$$\|f\|^2 = \|f - f^* + f^*\|^2 = \|f - f^*\|^2 + \|f^*\|^2,$$

on hem usat el teorema de Pitàgoras, tenint en compte que $\langle f^*, f - f^*\rangle = 0$. Per tant $\|f - f^*\|^2 = \|f\|^2 - \|f^*\|^2$. □

4.4.1 Les equacions normals i el teorema de la projecció ortogonal

Falta demostrar l'existència de la millor aproximació. Per a poder demostrar-la, suposarem que E^* és de dimensió finita. Per tant, existeix una base $\{\varphi_0, \varphi_1, \ldots, \varphi_n\}$ de E^*. Suposem que f^* és una millor aproximació. Llavors, podem escriure

$$f^* = \sum_{i=0}^{n} c_i \varphi_i.$$

La condició d'ortogonalitat del teorema de la projecció ortogonal diu que

$$\langle f - f^*, g\rangle = 0, \quad \forall\, g \in E^*,$$

o equivalentment

$$\langle f^*, g\rangle = \langle f, g\rangle, \quad \forall\, g \in E^*.$$

En particular

$$\langle f^*, \varphi_j\rangle = \langle f, \varphi_j\rangle, \quad \forall\, j = 0, 1, \ldots, n.$$

Si ara usem que $f^* = \sum_{i=0}^{n} c_i \varphi_i$, tenim que

$$\sum_{i=0}^{n} c_i \langle \varphi_i, \varphi_j\rangle = \langle f, \varphi_j\rangle, \qquad j = 0, 1, \ldots, n.$$

A aquest sistema lineal $(n+1) \times (n+1)$ l'anomenem **sistema de les equacions normals**, o simplement equacions normals.

Proposició 4.4.2 *Les equacions normals tenen solució única.*

Demostració:

Per veure que existeix solució i és única, n'hi ha prou en demostrar que l'única solució del sistema lineal homogeni amb incògnites $c_0, \ldots, c_n \in \mathbb{R}$,

$$\sum_{i=0}^{n} \langle \varphi_i, \varphi_j\rangle c_i = 0, \quad j = 0, \ldots, n$$

és $c_0 = c_1 = \cdots = c_n = 0$. En efecte, suposem que $c_0, \ldots, c_n$ és solució del sistema homogeni. Llavors

$$\left\| \sum_{i=0}^{n} c_i \varphi_i \right\|^2 = \langle \sum_{i=0}^{n} c_i \varphi_i, \sum_{j=0}^{n} c_j \varphi_j \rangle = \sum_{j=0}^{n} \left[\sum_{i=0}^{n} \langle \varphi_i, \varphi_j \rangle c_i \right] c_j = 0,$$

el que implica que $\sum_{i=0}^{n} c_i \varphi_i = 0$. Com que els φ_i són linealment independents, tenim que $c_0 = \cdots = c_n = 0$, que és el que voliem demostrar. □

Ara podem reformular el teorema de projecció ortogonal en el cas en el que E^* és de dimensió finita:

Teorema 4.4.4 *(caracterització de la millor aproximació) Sigui E un espai vectorial, $\langle \cdot, \cdot \rangle$ un producte escalar, E^* un subespai vectorial de E de dimensió finita, $\{\varphi_0, \varphi_1, \ldots, \varphi_n\}$ una base de E^*, i $f \in E$. Llavors $f^* = \sum_{i=0}^{n} c_i \varphi_i \in E^*$ és la millor aproximació de f des de E^*, on $(c_0, \ldots, c_n)$ és la solució del sistema de les equacions normals*

$$\sum_{i=0}^{n} c_i \langle \varphi_i, \varphi_j \rangle = \langle f, \varphi_j \rangle, \qquad j = 0, 1, \ldots, n.$$

Nota 4.4.2 *Notem que si la base és ortogonal ($\langle \varphi_i, \varphi_j \rangle = 0$, si $i \neq j$), llavors el sistema (anomenat també sistema de Gram) és diagonal, amb el que la solució és immediata:*

$$c_i = \frac{\langle f, \varphi_i \rangle}{\langle \varphi_i, \varphi_i \rangle}, \qquad i = 0, 1, \ldots, n.$$

L'estratègia de resolució del problema de la millor aproximació suggerida per aquest resultat és la següent: si es desitja aproximar elements de E mitjançant elements d'un subspai E^*, s'obté primer una base ortonormal $\{\varphi_0, , \ldots, \varphi_n\}$ per a E^*. En aquest cas la millor aproximació a f és

$$f^* = \sum_{i=0}^{n} \langle f, \varphi_i \rangle \varphi_i.$$

Aquesta estratègia té l'avantatge de la permanència, és a dir, si afegim un vector ortonormal més φ_{n+1} llavors la millor aproximació a f des de $\langle \varphi_0, \ldots, \varphi_{n+1} \rangle$ és $\tilde{f} = f^* + \langle f, \varphi_{n+1} \rangle \varphi_{n+1}$.

Exemple 4.4.1 *Usis el teorema de caracterització per a determinar la millor aproximació de la funció $f(x) = \sin x$ mitjançant un polinomi $g(x) = c_1 x + c_2 x^3 + c_3 x^5$ en l'interval $[-1, 1]$, i un s'usa la norma:*

$$\|f\| = \left\{ \int_{-1}^{1} [f(x)]^2 \, dx \right\}^{1/2}.$$

Resposta:

La funció òptima g posseeix la propietat $f - g$ ortogonal a G, on G és l'espai generat per $g_1(x) = x$, $g_2(x) = x^3$ i $g_3(x) = x^5$. Per tant, hem de demanar que $\langle g - f, g_i \rangle = 0$ per $i = 1, 2, 3$. Aquestes equacions es poden escriure com

$$c_1 \langle g_1, g_i \rangle + c_2 \langle g_2, g_i \rangle + c_3 \langle g_3, g_i \rangle = \langle f, g_i \rangle, \qquad (i = 1, 2, 3)$$

i s'anomenen **equacions normals** en aquest problema. Més explícitament obtenim

$$c_1\int_{-1}^{1} x^2\,dx + c_2\int_{-1}^{1} x^4\,dx + c_3\int_{-1}^{1} x^6\,dx = \int_{-1}^{1} x\sin x\,dx,$$
$$c_1\int_{-1}^{1} x^4\,dx + c_2\int_{-1}^{1} x^6\,dx + c_3\int_{-1}^{1} x^8\,dx = \int_{-1}^{1} x^3\sin x\,dx,$$
$$c_1\int_{-1}^{1} x^6\,dx + c_2\int_{-1}^{1} x^8\,dx + c_3\int_{-1}^{1} x^{10}\,dx = \int_{-1}^{1} x^5\sin x\,dx.$$

Després d'integrar totes les expressions obtenim el sistema

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{5} & \frac{1}{7} \\ \frac{1}{5} & \frac{1}{7} & \frac{1}{9} \\ \frac{1}{7} & \frac{1}{9} & \frac{1}{11} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_2 \end{bmatrix} = \begin{bmatrix} \alpha-\beta \\ -3\alpha+5\beta \\ 65\alpha-101\beta \end{bmatrix},$$

on $\alpha = \sin 1$ i $\beta = \cos 1$. Resolent, obtenim $c_1 \approx 9.999842124453209e{-}01$, $c_2 \approx -1.665241810658541e{-}01$, $c_3 \approx 8.018110364744436e-03$. Observem tanmateix que el sistema és mal condicionat ($\kappa_\infty \approx 1251$), i per tant, la base que hem considerat és inapropiada.

Finalment, podem calcular l'error de l'aproximació: Tenim que

$$\int_{-1}^{1} \sin^2 x\,dx = 1 - \frac{1}{2}\sin 2 \approx 0.54535128658715915231,$$

$$\int_{-1}^{1} (c_1x + c_2x^3 + c_3x^5)^2\,dx = \frac{2}{3}c_1^2 + \frac{2}{7}c_2^2 + \frac{2}{11}c_3^2 + \frac{4}{5}c_1c_2 + \frac{4}{7}c_1c_3 + \frac{4}{9}c_2c_3 \approx 0.54530058542127386676$$

Per tant

$$\|f-f^*\|_2^2 = \|f\|_2^2 - \|f^*\|_2^2 \approx 0.00005070116588528555,$$

i per tant

$$\|f-f^*\| \approx 0.00712047511654142008.$$

Del teorema de la projecció ortogonal, podem deduir els següents resultats relacionats amb les sèries de Fourier:

Corol·lari 4.4.2 *(Desigualtat de Bessel) Si $\varphi_0, \ldots, \varphi_n$ són vectors ortonormals d'un espai prehilibertià E llavors, per a tot $f \in E$:*

$$\sum_{i=0}^{n} |\langle f, \varphi_i\rangle|^2 \le \|f\|^2.$$

Demostració:

Pel teorema anterior sabem que la millor aproximació g a f satisfà

$$f^* = \sum_{i=0}^{n} \langle f, \varphi_i\rangle \varphi_i.$$

Pel corol·lari 4.4.1,

$$\|f-f^*\|^2 = \|f\|^2 - \|f^*\|^2 = \|f\|^2 - \sum_{i=0}^{n} |\langle f, \varphi_i\rangle|^2 \ge 0,$$

d'on es deduiex el resultat cercat. □

Corol·lari 4.4.3 *Si* $(\varphi_n)_{n\geq 0}$ *és una successió infinita de vectors ortonormals aleshores*

$$\sum_{i=0}^{\infty} |\langle f, \varphi_i\rangle|^2 \leq \|f\|^2.$$

Corol·lari 4.4.4 *Si* $(\varphi_n)_{n\geq 0}$ *és una successió infinita de vectors ortonormals aleshores*

$$\lim_{i\to\infty} \langle f, \varphi_i\rangle = 0.$$

Definició 4.4.2 *Un sistema finit o infinit d'elements* $\varphi_0, \varphi_1, \ldots$ *en un espai normat* V *s'anomena* ***tancat*** *(o total o complet) si donat* $v \in V$*, existeix una successió* $(v_n)_{n\geq 0}$ *tal que* $v_n \in \langle \varphi_0, \varphi_1, \ldots, \varphi_n\rangle$ *i* $\lim_{i\to\infty} \|v - \varphi_n\| = 0$.

Definició 4.4.3 *Sigui* $(E, \langle\cdot,\cdot\rangle)$ *un espai normat i* $\varphi_0, \varphi_1, \ldots$ *un sistema ortogonal tancat d'elements de* E*. Donada* $f \in E$*, la sèrie*

$$\sum_{k=0}^{\infty} \frac{\langle f, \varphi_k\rangle}{\langle \varphi_k, \varphi_k\rangle}\varphi_k$$

s'anomena ***sèrie de Fourier*** *de* f*, associada al sistema ortogonal donat.*

Teorema 4.4.5 *(teorema fonamental de desenvolupaments ortogonals) Sigui* $\varphi_0, \varphi_1, \ldots$ *una successió d'elements ortonormals en un espai prehilbertià* E*. Les següents afirmacions són equivalents*

a) La successió $(\varphi_i)_i$ *és un sistema ortonormal tancat en* E*.*

b) La sèrie de Fourier de qualsevol element $f \in E$ *convergeix en norma a* f *és a dir*

$$\lim_{n\to\infty} \left\| f - \sum_{k=0}^{n} \langle f, \varphi_k\rangle\varphi_k \right\| = 0.$$

c) Identitat de Parseval:

$$\|f\|^2 = \sum_{n=0}^{\infty} |\langle f, \varphi_n\rangle|^2.$$

Demostració:

a) $\Rightarrow$ b) Definim, per a cada $n \geq 0$, $E_n^* = \langle \varphi_0, \varphi_1, \ldots, \varphi_n\rangle$. Com que $(\varphi_i)_i$ és tancat, existeix una successió $(f_i)_i$, $f_i \in E_i^*$, tal que $\lim_{i\to\infty} \|f - f_i\| = 0$. Ara sigui $\tilde{f}_i$ la millor aproximació de f des de E_i^*. Pel teorema de la projecció ortogonal sabem que $\tilde{f}_n = \sum_{i=0}^{n} \langle f, \varphi_i\rangle\varphi_i$. A més, $\|f - \tilde{f}\| \leq \|f - f_n\|$, per la propietat de millor aproximació. Això prova b).

b) $\Rightarrow$ c) Sabem que $\lim_{n\to\infty} \|f - \tilde{f}_n\|^2 = \lim_{n\to\infty}(\|f\|^2 - \sum_{i=0}^{n} |\langle f, \varphi_i\rangle|^2) = 0$.

c) $\Rightarrow$ a) Si definim $f_n = \sum_{i=0}^{n} \langle f, \varphi_i\rangle\varphi_i$, llavors $\|f - f_n\|^2 = \|f\|^2 - \sum_{i=0}^{n} |\langle f, \varphi_i\rangle|^2) \to 0$ quan $n \to \infty$.

$\square$

Finalment, el següent teorema ens dona una manera d'ortogonalitzar vectors linealment independents en un espai prehilbertià, per tal de trobar vectors ortonormals:

Teorema 4.4.6 *(ortogonalització de Gram-Schmidt) Sigui $\{v_1, v_2, \ldots, v_n\}$ una base d'un subspai U en un espai prehilbertià. Definim recursivament*

$$u_i = \frac{1}{\left\| v_i - \sum_{j=1}^{i-1} \langle v_i, u_j \rangle u_j \right\|} \left(v_i - \sum_{j=1}^{i-1} \langle v_i, u_j \rangle u_j \right) \qquad i = 1, 2, \ldots, n.$$

Aleshores $\{u_1, u_2, \ldots, u_n\}$ és una base ortonormal de U.

4.5 Aproximació de funcions contínues: Polinomis ortogonals

Si considerem el problema d'aproximació de funcions contínues en un interval amb un cert producte escalar, és natural considerar com a aproximació, un polinomi de grau més petit o igual que n, per una certa n fixada, perquè és una funció fàcil d'avaluar. Per tant, si $E = C^0[(a, b])$, considerem $E^* = \langle 1, x, x^2, \ldots, x^n \rangle$. Pel que hem vist a la secció anterior, per a trobar la millor aproximació és convenient usar una base ortogonal de E^*. Desgraciadament, la base canònica no és ortogonal, i cal ortogonalitzar-la. Tanmateix, una simplificació sorprenent s'esdevé quan el procés de Gram-Schmidt s'aplica als monomis $1, x, x^2, \ldots$ (en el seu ordre natural) considerats com a funcions en un espai prehilbertià $C[a, b]$ amb un producte escalar que satisfaci la propietat que per a tres funcions qualsevol

$$\langle fg, h \rangle = \langle f, gh \rangle.$$

És obvi que això és vàlid per el producte escalar (o intern) que s'usa comunment

$$\langle f, g \rangle = \int_a^b f(x) g(x) w(x) \, dx,$$

on $w(x) \geq 0$, $\int_a^b w(x)\, dx > 0$ i $\int_a^b |x|^n w(x)\, dx < \infty$, per a tot $n \in \mathbb{N}$.

Teorema 4.5.1 *Sigui $(E, \langle \cdot \rangle)$ un espai prehilbertià de funcions reals contínues definides en un interval, tal que $\langle fg, h \rangle = \langle f, gh \rangle$, $\forall\, f, g, h \in E$. Aleshores existeix una única família de polinomis $p_0, \ldots, p_n, \ldots$ a coeficients reals tal que $p_0 \equiv 1$ i, per tot $i \geq 1$:*

a) $\operatorname{grau} p_i = i$, $i = 0, 1, \ldots, n$.

b) p_i *és mònic.*

c) $\langle p_i, p_j \rangle = 0$, $\forall\, 0 \leq j < i$.

A més, $(p_i)_{i \geq 0}$ satisfà la recurrència:

$$p_i(x) = (x - \delta_i) p_{i-1}(x) - \gamma_i^2 p_{i-2}(x),$$

on $p_{-1}(x) = 0$ *i*

$$\gamma_i^2 = \begin{cases} 0 & si \quad i = 1, \\ \dfrac{\langle p_{i-1}, p_{i-1}\rangle}{\langle p_{i-2}, p_{i-2}\rangle} & si \quad i \geq 2 \end{cases}$$

$$\delta_i = \frac{\langle xp_{i-1}, p_{i-1}\rangle}{\langle p_{i-1}, p_{i-1}\rangle}.$$

Demostració:

Demostrarem el teorema per inducció. Volem demostrar que si $p_0, \ldots, p_i$ són polinomis mònics ortogonals de graus $0, 1, \ldots, i$ respectivament, satisfent la recurrència donada, llavors existeix un únic polinomi mònic p_{i+1} de grau $i+1$, que és ortogonal a $p_0, \ldots, p_i$ i que satisfà la recurrència donada.

a) Cas $i = 1$: Agafem un polinomi mònic de grau 1 $p_1(x) = x - a$. Veurem $a = \delta_1$: En efecte, si imposem que $p_1(x) = x - a$ satisfà $\langle p_0, p_1\rangle = 0$ llavors

$$\langle 1, x - a\rangle = \langle 1, x\rangle - a\langle 1, 1\rangle = 0,$$

i per tant $a = \delta_1$, on δ_1 és el valor donat, i $\gamma_1 = 0$.

b) Suposem que és cert fins a i. Llavors, com que $p_0, \ldots, p_i$ és una base del subespai de polinomis de grau $\leq i$ i suposem que p_{i+1} és mònic, podem escriure

$$p_{i+1}(x) - xp_i(x) = -c_i p_i(x) + c_{i-1}p_{i-1}(x) + c_{i-2}p_{i-2}(x) + \cdots + c_0 p_0(x),$$

o

$$p_{i+1}(x) = (x - c_i)p_i(x) + c_{i-1}p_{i-1}(x) + c_{i-2}p_{i-2}(x) + \cdots + c_0 p_0(x).$$

Sabem que $\langle p_j, p_k\rangle = 0$, per a $j, k \leq i$ i $j \neq k$. Si imposem que $\langle p_{i+1}, p_j\rangle = 0$, $j \leq i$ tenim

$$0 = \langle p_{i+1}, p_i\rangle = \langle xp_i, p_i\rangle - c_i\langle p_i, p_i\rangle,$$

$$0 = \langle p_{i+1}, p_{j-1}\rangle = \langle xp_{j-1}, p_i\rangle + c_{j-1}\langle p_{j-1}, p_{j-1}\rangle, \quad j \leq i.$$

De la primera igualtat tenim que $c_i = \delta_{i+1}$. Per altra banda, per hipòtesi d'inducció

$$p_j(x) = (x - \delta_j)p_{j-1}(x) - \gamma_j^2 p_{j-2}(x), \qquad j \leq i.$$

Fent el producte escalar amb p_i obtenim

$$\langle p_j, p_i\rangle = \langle xp_{j-1}, p_i\rangle - \gamma_j^2\langle p_{j-2}, p_i\rangle = \langle xp_{j-1}, p_i\rangle, \qquad j \leq i.$$

Per tant

$$c_{j-1} = -\frac{\langle p_j, p_i\rangle}{\langle p_{j-1}, p_{j-1}\rangle} = \begin{cases} -\gamma_{i+1}^2 & \text{si} \quad j = i \\ 0 & \text{si} \quad j < i. \end{cases}$$

Això acaba la demostració. □

Definició 4.5.1 *Als polinomis obtinguts en el teorema anterior els anomenem* ***polinomis ortogonals (mònics)*** *associats al producte escalar* $\langle \cdot, \cdot \rangle$.

Proposició 4.5.1 (propietat minimal dels polinomis ortogonals) *Sigui M_n el conjunt de polinomis mònics de grau n. Llavors els polinomis p_n descrits en el teorema 4.5.1 satisfàn $\|p_n\| = \min_{q\in M_n} \|q\|$.*

Demostració:
Sigui $\tilde{p}$ és la millor aproximació a $f(x) = x^n$ entre els polinomis de grau $\leq n-1$. Pel teorema de la projecció ortogonal tenim que

$$\langle f - \tilde{p}, p_i\rangle = 0, \qquad 0 \leq i \leq n-1.$$

Com que $f - \tilde{p}$ és un polinomi mònic de grau n, pel teorema 4.5.1 tenim que $f - \tilde{p} = p_n$. Per altra banda, si q és un altre polinomi mònic de grau n el podem escriure com $f - q$, on q és un polinomi de grau $\leq n-1$. Per tant p_n és el que té la norma mínima entre tots els polinomis mònics de grau n. □

Nota 4.5.1 *La proposició anterior es pot usar per a demostrar la unicitat dels polinomis ortogonals que satisfan a), b) i c) del teorema anterior, per a un producte escalar arbitrari.*

Teorema 4.5.2 (zeros dels polinomis ortogonals) *Sigui p_n el polinomi de grau n pertanyent a una família de polinomis ortogonals com la descrita anteriorment, en l'espai prehilbertià $(C^0([a,b]), \langle\cdot,\cdot\rangle)$, on*

$$\langle f, g\rangle = \int_a^b w(x)f(x)g(x)\,dx, \qquad w \text{ contínua i positiva.}$$

Llavors p_n té n zeros reals simples en l'interval obert (a,b).

Demostració:
Suposem que $p_n(x)$ té k canvis de signe, $0 \leq k < n$ en (a,b). Siguin $t_1, \ldots, t_k$ els corresponents zeros (de multiplicitat senar). Llavors, si

$$p(x) = \prod_{i=1}^{k}(x - t_i)$$

tenim que $p_n(x)p(x)$ és un polinomi sense canvis de signe, i per tant:

$$\langle p_n, p\rangle = \int_a^b p_n(x)p(x)w(x)\,dx \neq 0.$$

Per altra banda, $\operatorname{grau} p = k < n$, i p_n és ortogonal a tots els polinomis de grau menor que n. Absurd. Per tant p_n té n canvis de signe en $[a,b]$, el que implica el que voliem. □

4.5.1 Polinomis de Txebixev

És un dels exemples importants de polinomis ortogonals.

Definició 4.5.2 *Definim els **polinomis de Txebixev** com $T_n(x) = \cos(n \arccos x)$, on $\arccos : [-1,1] \to [-\pi, 0]$.*

Proposició 4.5.2 *Els polinomis de Txevixev verifiquen la recurrència:*

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

amb $T_0 \equiv 1$, $T_1(x) = x$. *Les funcions* $\hat{T}_n(x) = 2^{-(n-1)}T_n(x)$ *són polinomis mònics de grau* n.

Demostració:
Ho demostrarem per inducció sobre n.
Per $n = 1$ tenim que $T_1(x) = \hat{T}_1(x) = x$.
Suposem que és cert per $i \leq n-1 \geq 0$. Sigui $\theta = \arccos x$. Llavors $T_n(x) = \cos(n\theta)$. A més, sabem que

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(\theta)\cos(n\theta).$$

Per tant,

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Utilitzant la recurrència tenim que T_n és un polinomi i $T_n(x) = 2^{n-1}x^n + \cdots$. □

Proposició 4.5.3 *Per a tot* n, $T_n(-x) = (-1)^n T_n(x)$.

Demostració:
Notem que és cert per T_0 i T_1. Suposem que és cer per $k < n$. Llavors,

$$T_n(-x) = 2(-x)T_{n-1}(-x) - T_{n-2}(-x) =$$

$$-2x(-1)^{n-1}T_{n-1}(x) - (-1)^{n-2}T_{n-2}(x) = (-1)^n[2xT_{n-1}(x) - T_{n-2}(x)] = (-1)^n T_n(x).$$

Per tant, acabem de demostrar la proposició per inducció sobre n. □

Proposició 4.5.4 $T_n(x)$ *té* n *zeros reals a l'interval* $[-1, 1]$, *donats per*

$$x_k = \cos\left(\frac{2k+1}{n}\frac{\pi}{2}\right), \qquad k = 0, 1, \ldots, n-1.$$

Demostració:
Sabem que $T_n(x_k) = 0$, és a dir $\cos(n \arccos x_k) = 0$. Recordem que $\cos(n\alpha) = 0$ sii $\alpha = \frac{2k+1}{n}\frac{\pi}{2}$, $k = 0, 1, \ldots, n-1$. Per tant,

$$\arccos x_k = \frac{2k+1}{n}\frac{\pi}{2},$$

el que implica el que voliem. □

Definició 4.5.3 *Els zeros* $\{x_k\}_{0 \leq k \leq n-1}$ *s'anomenen* ***abscisses de Txebitxev.***

Proposició 4.5.5 *Per a* $x \in [-1, 1]$, $|T_n(x)| \leq 1$ *i assoleix els valors* ± 1 *en*

$$x'_k = \cos\left(\frac{k\pi}{n}\right), \quad T_n(x'_k) = (-1)^k, \quad k = 0, 1, \ldots, n.$$

Demostració:
És immediata usant que $|\cos n\alpha| = 1$, per $\alpha = k\pi/n$. □

Proposició 4.5.6 *Els polinomis de Txebixev formen un sistema ortogonal en* $[-1,1]$ *quan s'usa el producte escalar:*

$$\langle f,g\rangle = \int_{-1}^{1} f(x)g(x)\frac{dx}{\sqrt{1-x^2}}.$$

A més

$$\langle T_i,T_i\rangle = \begin{cases} \frac{\pi}{2} & si \quad i\neq 0 \\ \pi & si \quad i=0 \end{cases}$$

Demostració:
Si fem el canvi de variable $x=\cos\theta$, $\theta\in[-\pi,0]$ tenim que

$$\langle f,g\rangle = \int_{-\pi}^{0} f(\cos\theta)g(\cos\theta)\,d\theta.$$

Llavors, si $n\neq m$,

$$\langle T_n,T_m\rangle = \int_{-\pi}^{0}\cos n\theta\cos m\theta\,d\theta = \frac{1}{2}\int_{-\pi}^{0}[\cos(n+m)\theta+\cos(n-m)\theta]\,d\theta =$$

$$= \frac{1}{2}\left[\frac{\sin(n+m)\theta}{n+m}+\frac{\sin(n-m)\theta}{n-m}\right]_{-\pi}^{0} = 0.$$

Si considerem $n=m\neq 1$:

$$\langle T_n,T_n\rangle = \int_{-\pi}^{0}\cos(n\theta)^2\,d\theta = \frac{1}{2}\int_{-\pi}^{0}(\cos 2n\theta+1)\,d\theta = \frac{1}{2}\left[\frac{\sin 2n\theta}{2n}+\theta\right]_{\theta=-\pi}^{0} = \frac{\pi}{2}.$$

Quan $n=0$:

$$\langle T_0,T_0\rangle = \int_{-\pi}^{0} dx = \pi.$$

□

Proposició 4.5.7 *(propietat minimax) Sigui* M_n *el conjunt de polinomis mònics de grau* n*. L'element de* M_n *que té menor norma infinit a* $[-1,1]$ *és* $2^{1-n}T_n$*. El valor de la norma és* 2^{1-n}*.*

Demostració:
És clar que

$$\max_{x\in[-1,1]}|2^{1-n}T_n(x)| = 2^{1-n}.$$

Suposem que existeix un polinomi $p_n\in M_n$ tal que

$$|p_n(x)|<2^{1-n}, \qquad \forall\, x\in[-1,1].$$

Siguin x'_k, $k=0,1,\ldots,n$ els llocs on $|T_n(x'_k)|=1$. Recordem que $T_n(x'_k)=(-1)^k$. Llavors

$$\begin{array}{rcl} p_n(x'_0) & < & 2^{1-n}T_n(x'_0) \\ p_n(x'_1) & > & 2^{1-n}T_n(x'_1) \\ p_n(x'_2) & < & 2^{1-n}T_n(x'_2) \\ & \vdots & \end{array}$$

fins x'_n. Això impica que el polinomi $r(x)=p_n(x)-2^{1-n}T_n(x)$ que té grau més petit o igual que $n-1$. té n canvis de signe. Però llavors tindria més de $n-1$ zeros, el que és impossible! D'aquesta manera queda demostrada la proposició per reducció a l'absurd. □

Corol·lari 4.5.1 *L'element que té menor norma infinit en $[a,b]$ és*

$$\frac{(b-a)^n}{2^{2n-1}} T_n\left(\frac{2}{b-a}(x-a)-1\right)$$

i val $(b-a)^n/2^{2n-1}$.

Demostració:
Sigui $p \in M_n$. Llavors

$$\|p\|_\infty = \max_{x\in[a,b]} |p(x)| = \max_{t\in[-1,1]} \left| p\left(\frac{b-a}{2}(t+1)+a\right)\right|.$$

Aleshores,

$$q(t) = \frac{2^n}{(b-a)^n} p\left(\frac{b-a}{2}(t+1)+a\right)$$

és un polinomi mònic, i per tant,

$$\|p\|_\infty = \frac{(b-a)^n}{2^n} \max_{t\in[-1,1]} |q(t)| \geq \max_{t\in[-1,1]} \frac{(b-a)^n}{2^{2n-1}} |T_n(t)| = \frac{(b-a)^n}{2^{2n-1}}.$$

A més,

$$\frac{(b-a)^n}{2^{2n-1}} \max_{t\in[-1,1]} |T_n(t)| = \frac{(b-a)^n}{2^{2n-1}} \max_{x\in[a,b]} \left| T_n\left(\frac{2}{b-a}(x-a)-1\right)\right|,$$

el que prova el que voliem. □

Aplicació al problema d'interpolació

Suposem que tenim $m+1$ punts $x_0, x_1, \ldots, x_m \in [a,b]$, $x_i \neq x_j$ si $i \neq j$. Sabem que existeix un únic polinomi de grau m, p_m, tal que $p_m(x_i) = f_i$, per tot i. L'error d'interpolació ve donada per la fórmula

$$f(x) - p_m(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!}(x-x_0)(x-x_1)\cdots(x-x_m).$$

Suposem que la funció f és coneguda i que podem triar les abscisses $\{x_i\}_0^m$ on realitzar la interpolació. De cara a minimitzar l'error d'interpolació, quina és la millor elecció d'aquestes abscisses?

Hem vist al corol·lari que el polinomi que minimimitza la norma infinit en $[a,b]$ té arrels

$$x_i = \frac{b-a}{2}\left[\cos\left(\frac{2i+1}{m+1}\frac{\pi}{2}\right)+1\right]+a, \qquad i = 0, 1 \ldots, m.$$

Per tant aquesta és la millor elecció. Es pot demostrar que l'us d'aquestes abscisses granteix la convergència de la interpolació si f és una funció analítica. És a dir, el fenòmen de Runge no es dona per a aquestes abscisses.

Definició 4.5.4 *S'anomena polinomi d'interpolació de Txebixev al polinomi d'interpolació sobre les abscisses de Txebixev.*

4.5.2 Polinomis de Legendre

Es poden definir de moltes maneres. Per exemple:

Definició 4.5.5 *Els polinomis de Legendre són*

$$P_0(x) = 1,$$

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n], \qquad n = 1, 2, \dots$$

Teorema 4.5.3 *Els polinomis de Legendre P_n són ortogonals respecte del producte escalar*

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx,$$

més concretament:

$$\langle P_j, P_k\rangle = \begin{cases} 0 & si \quad j \neq k, \\ \frac{2}{2j+1} & si \quad j = k \end{cases}$$

Proposició 4.5.8 $P_n(-x) = (-1)^n P_n(x)$.

Proposició 4.5.9 *Es verifica la següent recurrència:*

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x),$$

$$P_0(x) = 1, \qquad P_1(x) = x, \qquad P_2(x) = \frac{1}{2}(3x^2 - 1).$$

Proposició 4.5.10 $|P_n(x)| \leq 1$, $\forall x \in [-1, 1]$.

4.6 Aproximació de funcions periòdiques: polinomis trigonomètrics

Si volem aproximar una funció periòdica, és natural agafar la seva aproximació dins d'un subespai E^* de funcions periòdiques. Les funcions periòdiques en $\mathbb{R}$ de periode 2π venen representades per l'espai $C^0([0, 2\pi])$, si admetem que puguin ser discontínues en 0 i 2π (si no caldria restrigir-se a les continues tals que a 0 i 2π tenen el mateix valor). En efecte, només cal definir, per $x \in [2k\pi, 2(k+1)\pi]$, $f(x) = f(x - 2k\pi)$. Una elecció natural de E^* és l'espai P_{2n+1} generat per les funcions $\psi_0(\theta) = \frac{1}{2}$, $\psi_1(\theta) = \cos\theta$, $\psi_2(\theta) = \sin\theta, \dots,$ $\psi_{2n-1}(\theta) = \cos n\theta$, $\psi_{2n}(\theta) = \sin n\theta$.

Definició 4.6.1 *A qualsevol combinació lineal de funcions ψ_i l'anomenem* ***polinomi trogonomètric***

Nota 4.6.1 *En general, si considerem funcions periòdiques de periode $T > 0$, l'espai E^* estarà generat per les funcions $\psi_0(\theta) = \frac{1}{2}$, $\psi_1(\theta) = \cos(2\pi\theta/T)$, $\psi_2(\theta) = \sin(2\pi\theta/T), \dots,$ $\psi_{2n-1}(\theta) = \cos(2\pi n\theta/T)$, $\psi_{2n}(\theta) = \sin(2\pi n\theta/T)$.*

Si usem el producte escalar:

$$\langle f, g\rangle = \int_0^{2\pi} f(\theta)g(\theta)\, d\theta$$

veiem que les funcions ψ_i són ortogonals:

Proposició 4.6.1 *Si $\{\psi_j\}_{j=0}^{2n}$ és la família introduida abans, tenim que*

$$\langle \psi_i, \psi_j\rangle = \begin{cases} 0 & si \quad i \neq j \\ \frac{\pi}{2} & si \quad i = j = 0, \\ \pi & si \quad i = j > 0. \end{cases}$$

Demostració:
Cal usar les identitats:

$$\cos A \cos B = \frac{1}{2}[\cos(A-B) + \cos(A+B)],$$

$$\sin A \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)],$$

$$\sin A \cos B = \frac{1}{2}[\sin(A-B) + \sin(A+B)].$$

□

Per tant, $\{\psi_j\}_{j=0}^{2n}$ és una família ortogonal i base de P_{2n+1}. Si tornem al problema d'aproximació, donada una funció periòdica i contínua $f : \mathbb{R} \to \mathbb{R}$ llavors la millor aproximació f^* des de P_{2n+1} és

$$f^*(\theta) = \frac{a_0}{2} + \sum_{j=1}^{n} a_j \cos j\theta + \sum_{j=1}^{n} b_j \sin\ j\theta,$$

on

$$a_0 = \frac{\langle \psi_0, f\rangle}{\langle \psi_0, \psi_0\rangle} = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\, d\theta,$$

$$a_j = \frac{\langle \psi_{2j-1}, f\rangle}{\langle \psi_{2j-1}, \psi_{2j-1}\rangle} = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\cos j\theta\, d\theta,$$

$$b_j = \frac{\langle \psi_{2j}, f\rangle}{\langle \psi_{2j}, \psi_{2j}\rangle} = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin j\theta\, d\theta.$$

A aquesta aproximació se l'anomena **aproximació de Fourier**, i els coeficients $a_0, a_1, \ldots, b_1, \ldots, b_n$, **coeficients de Fourier**.

4.7 Aproximació discreta

L'aproximació discreta consisteix en aproximar una funció f definida en una taula de valors. Més concretament tenim la següent definició:

Definició 4.7.1 *Donades una funció $f \in C^0([a,b])$, un conjunt de punts $\{x_0,\dots,x_m\}$, $x_i \in [a,b]$, $0 \le i \le m$, i un subespai $E^* \subset C^0([a,b])$ de dimensió finita, al problema de trobar la funció $f^* \in E^*$ tal que*

$$\|f - f^*\|_2^2 = \sum_{i=0}^{m} |f(x_i) - f^*(x_i)|^2 = \min_{g \in E^*} \|f - g\|_2^2$$

se l'anomena ***problema d'aproximació per mínims quadrats*** *de la funció f des de el subespai E^* en els nodes $x_0,\dots,x_m$.*

El problema de l'aproximació per mínims quadrats es pot generalitzar si permetem l'existència de pesos. En general, definim el producte escalar associat:

$$\langle f, g \rangle = \sum_{i=0}^{m} f(x_i) g(x_i) w_i, \tag{4.1}$$

on $x_i \in [a,b]$ i $w_i > 0$, per $0 \le i \le m$, i $x_i \neq x_j$ si $i \neq j$ L'aplicació $\langle \cdot, \cdot \rangle$ compleix totes les condicions de producte escalar tret d'una: $\langle f, f \rangle = 0$ no implica $f \equiv 0$. A aquest tipus de producte escalar l'anomenarem producte escalar degenerat.

EL nostre objectiu serà reformular el problema d'aproximació discreta a un altre problema en el que poguem aplicar la teoria d'aproximació en espais prehilbertians. Per això, substituirem l'espai vectorial $C^0([a,b])$ per un subespai en el que (4.1) sigui un producte escalar.

Ens interessarem per dos problemes d'aproximació discreta. El primer en el que la funció aproximant és un polinomi i el segon en el que la funció a aproximar és periòdica i la funció aproximant és un polinomi trigonomètric.

4.7.1 Aproximació polinomial discreta

En aquest cas substituïm $C^0([a,b])$ per $E = \mathbb{R}_m[x]$ (espai de polinomis de grau més petit o igual que m). Tenim

Proposició 4.7.1 *Sigui $E = \mathbb{R}_m[x]$ (polinomis de grau $\le m$) i considerem la xarxa $\{x_0, x_1, \dots, x_m\} \subset [a,b]$. Llavors el producte escalar degenerat (4.1) és un producte escalar en aquest espai.*

Demostració: Es conseqüència immediata de que un polinomi no nul de grau m té com a màxim m arrels reals. □

Com que volem aproximar f per un polinomi, definirem $E^* = \mathbb{R}_n[x]$, on $n < m$ (excloem el cas $n = m$ perquè el problema es redueix al problema d'interpolació polinomial).

Ara, per a trobar la millor aproximació podem aplicar el teorema de la projecció ortogonal i usar les equacions normals.

Si considerem el cas particular en el que les abscisses són equiespaiades a $[-1,1]$ i els pesos són tots 1:

$$x_i = -1 + \frac{2i}{m}, \qquad i = 0, 1, \dots, m$$

els **polinomis de Gram** són polinomis ortogonals respecte d'aquest producte escalar. Els anomenem $\{P_{n,m}\}_{n=0}^{m}$, i verifiquen

$$\langle P_{i,m}, P_{j,m}\rangle = \begin{cases} 0 & \text{si} \quad i \neq j \\ 1 & \text{si} \quad i = j. \end{cases}$$

La fórmula recurrent és:

$$P_{n+1,m}(x) = \alpha_{n,m} x P_{n,m}(x) - \gamma_{n,m} P_{n-1,m}(x),$$

on

$$\alpha_{n,m} = \frac{m}{n+1}\left(\frac{4(n+1)^2-1}{(m+1)^2-(n+1)^2}\right)^{1/2}, \quad \gamma_{n,m} = \frac{\alpha_{n,m}}{\alpha_{n-1,m}},$$

$P_{0,m}(x) = (m+1)^{-1/2}$, $P_{-1,m}(x) = 0$.

4.7.2 Aproximació trigonomètrica discreta

En aquest cas és natural suposar que la funció $f \in C^0([a,b])$ és periòdica de període $T = b-a$. Per simplificar, suposarem que $a = 0$, $b = 2\pi$, i per tant $T = 2\pi$. Definim, com abans

$$\psi_0 = 1/2, \psi_1(\theta) = \cos\theta, \psi_2(\theta) = \sin\theta, \ldots, , \psi_{2k-1}(\theta) = \cos k\theta, \psi_{2k}(\theta) = \sin k\theta, \ldots, \psi_m(\theta) = \sin(\frac{m}{2}\theta).$$

Sigui $\{\theta_0, \ldots, \theta_m\} \subset [0, 2\pi)$ on m és un nombre parell, i el producte escalar (4.1).

Proposició 4.7.2 *Sigui $E = \langle\psi_0, \ldots, \psi_m\rangle$, on $m = 2m'$, i considerem la xarxa $\{\theta_0, \theta_1, \ldots, \theta_m\} \subset [0, 2\pi)$. Llavors el producte escalar degenerat (4.1) és un producte escalar en aquest espai.*

Demostració:

Suposem que $\langle f, f\rangle = 0$, on $f = \sum_{k=0}^{m} c_k\psi_k$. Això implica que $f(\theta_k) = 0$, $k = 0, \ldots, m$. Però si definim $\xi = e^{i\theta}$, tenim que $f(\theta) = \xi^{-m'}\sum_{k=-m'}^{m'} d_k\xi^{m'+k} = \xi^{-m'}p_m(\xi)$, on p_m és un polinomi de grau més petit o igual que m. Com que $f(\theta_k) = 0$ implica que $p_m(\xi_k) = 0$, i $e^{i\theta_k} \neq e^{i\theta_j}$, si $k \neq j$, llavors p_m té $m+1$ zeros, el que implica que $p \equiv 0$, i per tant $f \equiv 0$. $\square$

Ara, ja podem aplicar el teorema de la projecció ortogonal i usar les equacions normals. Com que volem aproximar f per un polinomi trigonomètric, definirem $E^* = \langle\psi_0, \psi_1, \ldots, \psi_{2n}\rangle$, on $2n \leq m$.

Si considerem el cas particular en el que els pesos són iguals a 1, és a dir que tenim el producte escalar

$$\langle f, g\rangle_m = \sum_{k=0}^{m} f(\theta_k)g(\theta_k),$$

i les abscisses són

$$\theta_k = \frac{2\pi k}{m+1}, \qquad k = 0, 1, \ldots, m,$$

llavors

$$\langle\psi_j, \psi_k\rangle_m = \begin{cases} 0 & \text{si} \quad j \neq k \\ \frac{m+1}{4} & \text{si} \quad j = k = 0 \\ \frac{m+1}{2} & \text{si} \quad j = k > 0. \end{cases}$$

Per tant, la millor aproximació en aquest cas ve donada per:

$$f^*(\theta) = \frac{a_0}{2} + \sum_{j=1}^{n} a_j \cos j\theta + \sum_{j=1}^{n} b_j \sin\ j\theta,$$

on

$$a_0 = \frac{\langle \psi_0, f \rangle_m}{\langle \psi_0, \psi_0 \rangle_m} = \frac{2}{m+1} \sum_{k=0}^{m} f(\theta_k),$$

$$a_j = \frac{2}{m+1} \sum_{k=0}^{m} f(\theta_k) \cos j\theta_k,$$

$$b_j = \frac{2}{m+1} \sum_{k=0}^{m} f(\theta_k) \sin j\theta_k.$$