Best approximation in the 2-norm

Given n, and $f \in L^2_w(a,b)$

find $p_n \in \mathcal{P}_n$ such that

$$(*) \quad \|f - p_n\|_2$$

is as small as possible.

Above, for $g \in L^2_w(a,b)$,

$$\|g\|_2 := \sqrt{\int_a^b w(x)\, g(x)^2\, dx}$$

and

$L^2_w(a,b)$ - vector space of functions $f : \mathbb{R} \to \mathbb{R}$ integrable on (a,b)

Remark

$$C[a,b] \subseteq L^2_w(a,b)$$

But there are integrable functions on (a,b) that are not continuous, e.g.

$$g(x) = \begin{cases} 1 & x \geq \frac{a+b}{2} \\ 0 & x < \frac{a+b}{2} \end{cases}$$

$g \in L^2_w(a,b)$, but $g \notin C[a,b]$

Today is about orthogonality, orthogonal polynomials.

Next lecture concerns minimization of $\|f-p_n\|_2$ over $p_n \in \mathcal{P}_n$.

Inner Product

The major difference between $\|\cdot\|_\infty$ & $\|\cdot\|_2$ is that latter is induced by an inner product

Defn

Let V be a vector space over $\mathbb{R}$.

A function $\langle \cdot , \cdot \rangle : V \times V \to \mathbb{R}$ is called an inner product (on V) if it satisfies

Positivity $\{$ (i) $\langle f, f \rangle > 0 \quad \forall f \in V,\ f \not\equiv 0$

Symmetry $\{$ (ii) $\langle f, g \rangle = \langle g, f \rangle \quad \forall f, g \in V$

Bilinearity $\begin{cases} \text{(iii) } \langle f, \alpha g \rangle = \alpha \langle f, g \rangle \ \forall f, g \in V,\ \forall \alpha \in \mathbb{R} \\ \text{(iv) } \langle f, g+h \rangle = \langle f, g \rangle + \langle f, h \rangle \ \forall f, g, h \in V \end{cases}$

A vector space with an inner product is called an inner product space.

Examples of inner product spaces

① $\mathbb{R}^n$ with

$$\langle v, u \rangle := v^T u = v_1 u_1 + \cdots + v_n u_n$$

for $v, u \in \mathbb{R}^n$.

② $L^2_w(a,b)$ with

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

for $f, g \in L^2_w(a,b)$.

Defn (Induced Norm)

Let $\langle \cdot, \cdot \rangle$ be an inner product on V.

Then

$$\|f\| := \sqrt{\langle f, f \rangle} \quad \forall f \in V$$

is called the norm induced by $\langle \cdot, \cdot \rangle$.

Examples

① $\|v\|_2 = \sqrt{v_1^2 + \cdots + v_n^2}$ on $\mathbb{R}^n$
is induced by $\langle v, u \rangle = v^T u$

② $\|f\|_2 = \sqrt{\int_a^b w(x) f(x)^2\,dx}$ on $L^2_w(a,b)$
is induced by $\langle f, g \rangle = \int_a^b w(x) f(x) g(x)\,dx$.

THM (Cauchy - Schwarz Inequality)

Let $\langle \circ, \circ \rangle$ be an inner product, and $\|\cdot\|$ be the corresponding induced norm on V. Then

$$|\langle f, g \rangle| \leqslant \|f\| \, \|g\| \quad \forall f, g \in V.$$

Proof - as before as we did on $\mathbb{R}^n$ by exploiting $\|f + \lambda g\|^2 \geqslant 0 \quad \forall \lambda \in \mathbb{R}$, in particular the non-~~negativity~~ positivity of the discriminant of

$$q(\lambda) := \langle f + \lambda g, f + \lambda g \rangle = \|f\|^2 + 2\lambda \langle f, g \rangle + \lambda^2 \|g\|^2$$

Defn (Orthogonality)

Two vectors f, g in an inner product space V are called orthogonal if $\langle f, g \rangle = 0$.

Ex

Monomials $x^j, x^k \in P_n$ with $j, k \leqslant n$ are not orthogonal, w.r.t.

$$\langle f, g \rangle = \int_0^1 f(x) g(x) \, dx$$

i.e.,

$$\int_0^1 x^j x^k \, dx = \frac{x^{j+k+1}}{j+k+1} \Bigg|_0^1 = \frac{1}{j+k+1} > 0$$

$\{1, x, x^2, \ldots, x^n\}$ is a basis for P_n, but with elements that are not orthogonal in $L^2_w(0,1)$ with $w(x) \equiv 1$.

<u>Defn</u>

A system of orthogonal polynomials $\varphi_0(x), \varphi_1(x), \ldots, \varphi_n(x)$ in $L^2_w(a,b)$ is such that $\varphi_j(x)$ is exactly of degree j, and

$$\langle \varphi_j, \varphi_k \rangle \begin{cases} = 0 & k \neq j \\ \neq 0 & k = j \end{cases}$$

for $j = 0, 1, \ldots, n$.

A procedure (Gram-Schmidt procedure) to form an orthogonal system

$$\varphi_0(x) \equiv 1$$

$$\varphi_j(x) = x^j - \beta_0 \varphi_0(x) - \ldots - \beta_{j-1} \varphi_{j-1}(x)$$

Multiply with $\varphi_k(x)$ to obtain β_k that is (by the requirement of orthogonality)

$$(\times) \quad \beta_k = \frac{\langle \varphi_k, x^j \rangle}{\langle \varphi_k, \varphi_k \rangle} \qquad k = 0, 1, \ldots, j-1$$

Now verify that $\langle \varphi_j, \varphi_k \rangle = 0$ for $k = 0, \ldots, j-1$ indeed holds for $\beta_0, \beta_1, \ldots, \beta_{j-1}$ as in (x).

Ex.

① Orthogonal polynomials $\varphi_0, \varphi_1, \varphi_2$ on $L^2_w(0,1)$ with $w(x) \equiv 1$.

$$\varphi_0(x) \equiv 1$$

$$\varphi_1(x) = x - \beta_0 \varphi_0(x)$$

with $\beta_0 = \dfrac{\langle \varphi_0, x \rangle}{\langle \varphi_0, \varphi_0 \rangle}$

$$\langle \varphi_0, x \rangle = \int_0^1 x \, dx = \frac{x^2}{2} \Big|_0^1 = 1/2$$

$$\langle \varphi_0, \varphi_0 \rangle = \int_0^1 1 \, dx = 1$$

$$\varphi_1(x) = x - \frac{1}{2}$$

$$\varphi_2(x) = x^2 - \tilde{\beta}_0 \varphi_0(x) - \tilde{\beta}_1 \varphi_1(x)$$

with $\tilde{\beta}_0 = \dfrac{\langle \varphi_0, x^2 \rangle}{\langle \varphi_0, \varphi_0 \rangle}$ $\quad \tilde{\beta}_1 = \dfrac{\langle \varphi_1, x^2 \rangle}{\langle \varphi_1, \varphi_1 \rangle}$

$\langle \varphi_0, x^2 \rangle = \int_0^1 x^2 dx = 1/3$ $\quad \langle \varphi_1, x^2 \rangle = \int_0^1 x^3 - \frac{1}{2} x^2 dx = \frac{1}{12}$

$$\langle \varphi_1, \varphi_1 \rangle = \int_0^1 (x - \tfrac{1}{2})^2 dx = \tfrac{1}{3}(x - \tfrac{1}{2})^3 = \tfrac{1}{12}$$

$$\varphi_2(x) = x^2 - \frac{1}{3} - (x - 1/2) = x^2 - x + 1/6$$

② Legendre polynomials, orthogonal polynomials on $\mathbf{L}^2_w(-1,1)$ with $w(x) \equiv 1$

To map $[0,1]$ into $[a,b]$ use the substitution

$$\hat{x} = (b-a)x + a \qquad \left(\begin{array}{l} \hat{x} = a \Leftrightarrow x = 0 \\ \hat{x} = b \Leftrightarrow x = 1 \end{array} \right.$$

specificially with $a=-1$, $b=1$

$$\hat{x} = 2x - 1$$

$$\Longleftrightarrow \quad x = (\hat{x}+1)/2$$

Choose

$$\hat{\varphi}_0(\hat{x}) = 1$$

$$0 = \int_0^1 \varphi_0(x)\, \varphi_1(x)\, dx \qquad \left(\begin{array}{l} \varphi_0, \varphi_1 \\ \text{orthogonal on } [0,1] \\ \text{as in page ⑥} \end{array} \right.$$

$$= \int_{-1}^{1} \underbrace{\varphi_0\left(\frac{\hat{x}+1}{2}\right)}_{\hat{\varphi}_0(\hat{x}) :=} \underbrace{\varphi_1\left(\frac{\hat{x}+1}{2}\right)}_{\hat{\varphi}_1(\hat{x}) :=} \frac{1}{2}\, d\hat{x}$$

$$\hat{\varphi}_1(\hat{x}) = \varphi_1\left(\frac{\hat{x}+1}{2}\right) = \frac{\hat{x}+1}{2} - \frac{1}{2} = \frac{\hat{x}}{2}$$

Similarly,

$$\hat{\varphi}_2(\hat{x}) = \varphi_2\left(\frac{\hat{x}+1}{2}\right) = \left(\frac{\hat{x}+1}{2}\right)^2 - \left(\frac{\hat{x}+1}{2}\right) + \frac{1}{6} = \frac{\hat{x}^2}{4} - \frac{1}{12}$$

is orthogonal to $\hat{\varphi}_1(\hat{x}), \hat{\varphi}_0(\hat{x})$

Normalized versions $\tilde{\varphi}_0, \tilde{\varphi}_1, \tilde{\varphi}_2$ so that $\tilde{\varphi}_0(1) = \tilde{\varphi}_1(1) = \tilde{\varphi}_2(1) = 1$

$$\tilde{\varphi}_0(\hat{x}) \equiv 1$$

$$\tilde{\varphi}_1(\hat{x}) = \hat{x}$$

$$\tilde{\varphi}_2(\hat{x}) = \frac{3}{2}\hat{x}^2 - \frac{1}{2}$$

(3) Chebyshev polynomials, orthogonal on $L^2_w(-1,1)$ with $w(x) = 1/\sqrt{1-x^2}$.

$$T_n(x) = \cos(n \arccos(x))$$

$$\langle T_n, T_m \rangle = \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} \cos(n \arccos(x)) \cos(m \arccos(x))\, dx$$

Letting θ s.t $x = \cos\theta$

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

$$= \int_0^{\pi} \cos(n\theta)\cos(m\theta)\, d\theta$$

$$= \int_0^{\pi} \frac{\{\cos((n+m)\theta) + \cos((n-m)\theta)\}}{2}\, d\theta$$

$$= \begin{cases} 0 & n \neq m \\ \pi/2 & n = m \end{cases}$$

THM (Roots of Orthogonal Polynomials)

Let $\varphi_0, \varphi_1, \ldots, \varphi_n$ be a system of orthogonal polynomials in $L^2_w(a,b)$. Then all roots of φ_i, $i \geq 1$ are simple, and lie in (a,b).

Proof

Suppose $r_1, \ldots, r_k$ are the (distinct) points in (a,b) where φ_j changes sign.

If $k=0$ (i.e., φ_j does not change sign at all),

$$0 = \langle \varphi_j, 1 \rangle$$
$$= \int_a^b w(x)\, \varphi_j(x)\, dx$$

yields a contradiction, as the last integral cannot be zero due to the fact that $w(x)\varphi_j(x) > 0$ OR $w(x)\varphi_j(x) < 0$ $\forall x \in (a,b)$ except possibly the roots of φ_j.

Now suppose $k < j$. But then letting $\phi(x) = \prod_{l=1}^{k} (x - r_l)$

$$0 = \langle \varphi_j, \phi \rangle$$
$$= \int_a^b w(x)\, \varphi_j(x)\, \phi(x)\, dx.$$

But again the last integral cannot be zero, because either $w(x)\varphi_j(x)\phi(x) > 0$ $\forall x \in (a,b)$ (except roots of φ_j, ϕ) OR $w(x)\varphi_j(x)\phi(x) < 0$ $\forall x \in (a,b)$ (except roots of φ_j, ϕ).

Hence, $k = j$ meaning all roots of φ_j are simple and contained in (a,b). $\square$

<u>**Ex**</u>

$\varphi_0(x) = 1,\quad \varphi_1(x) = x - \frac{1}{2}\quad \varphi_2(x) = x^2 - x + 1/6$

root $x_0 = \frac{1}{2} \in (0,1)$ roots $x_{0,1} = \frac{1 \pm \sqrt{1/3}}{2} \in (0,1)$

on $L^2_w(0,1)$ with $w(x) \equiv 1$

Chebyshev polynomials

$\varphi_0(x) = 1\quad \varphi_1(x) = x\quad \varphi_2(x) = 2x^2 - 1$

root $x_0 = 0 \in (-1,1)$ roots $x_{0,1} = \pm\sqrt{\frac{1}{2}} \in (-1,1)$

$\varphi_3(x) = 4x^3 - 3x$

roots $x_{0,1,2} = \frac{-\sqrt{3}}{2}, 0, \frac{\sqrt{3}}{2} \in (-1,1)$

on $L^2_w(-1,1)$ with $w(x) = 1/\sqrt{1-x^2}$

Optimal Polynomial for the Best Approximation Problem in the 2-Norm

Objective to be minimized

$$\| f - p_n \|_2$$

over $p_n \in \mathcal{P}_n$.

$\varphi_0, \varphi_1, \ldots, \varphi_n$ - a system of orthogonal polynomials on $L^2_w(a,b)$.

$$p_n = \beta_0 \varphi_0 + \beta_1 \varphi_1 + \ldots + \beta_n \varphi_n$$

where

$$\beta_j = \frac{\langle \varphi_j, p_n \rangle}{\langle \varphi_j, \varphi_j \rangle} \qquad j = 0, 1, \ldots, n.$$

Let us have a closer look at

$$\| f - p_n \|_2^2 = \langle f - p_n, f - p_n \rangle$$

$$= \| f \|_2^2 - 2 \langle f, p_n \rangle + \| p_n \|_2^2$$

$$= \| f \|_2^2 - 2 \sum_{j=0}^{n} \beta_j \langle f, \varphi_j \rangle + \sum_{j=0}^{n} \beta_j^2 \langle \varphi_j, \varphi_j \rangle$$

Letting $E : \mathbb{R}^{n+1} \to \mathbb{R}$,

$$E(\beta_0, \ldots, \beta_n) := \|f - p_n\|_2^2$$

and assuming $\langle \varphi_j, \varphi_j \rangle = 1$ (i.e. $\|\varphi_j\|_2 = 1$) for $j = 0, 1, \ldots, n$, we have

$$E(\beta_0, \ldots, \beta_n) = \sum_{j=0}^{n} \{\langle f, \varphi_j \rangle - \beta_j\}^2$$

$$+ \left\{ \|f\|_2^2 - \sum_{j=0}^{n} \langle f, \varphi_j \rangle^2 \right\}$$

Hence, for $\beta_{*,j} := \langle f, \varphi_j \rangle$, we have

$$E(\beta_{*,0}, \ldots, \beta_{*,n}) = \left\{ \|f\|_2^2 - \sum_{j=0}^{n} \langle f, \varphi_j \rangle^2 \right\}$$

$$\leqslant E(\beta_0, \ldots, \beta_n) \quad \forall \beta_0, \ldots, \beta_n \in \mathbb{R}$$

Equivalently, for $p_{*,n} := \beta_{*,0}\,\varphi_0(x) + \beta_{*,1}\,\varphi_1(x) + \ldots + \beta_{*,n}\,\varphi_n(x)$, we have

$$\|f - p_{*,n}\|_2^2 \leqslant \|f - p_n\|_2^2 \quad \forall p_n \in \mathcal{P}_n .$$

<u>THM</u>

There exists a unique $p_* \in \mathcal{P}_n$ such that

$$\|f - p_*\|_2 = \min_{p_n \in \mathcal{P}_n} \|f - p_n\|_2$$

Ex

Solve

$$\min_{p_1 \in P_1} \| e^{-x} - p_1 \|_2$$

over $L^2_w(0,1)$ with $w(x) \equiv 1$.

Orthogonal polynomials on $L^2_w(0,1)$

$$\hat{\varphi}_0(x) \equiv 1, \qquad (\|\hat{\varphi}_0\|_2 = 1)$$

$$\hat{\varphi}_1(x) = x - 1/2 \qquad (\|\hat{\varphi}_1\|_2 = \sqrt{12})$$

Normalize them so that they have unit norm

$$\varphi_0(x) = \frac{\hat{\varphi}_0(x)}{\|\hat{\varphi}_0(x)\|_2} = 1$$

$$\varphi_1(x) = \frac{\hat{\varphi}_1(x)}{\|\hat{\varphi}_1(x)\|_2} = 2\sqrt{3}\, x - \sqrt{3}$$

Compute the coordinates of optimal p_*

$$\beta_{*,0} := \langle e^{-x}, \varphi_0 \rangle = \int_0^1 e^{-x} dx = 1 - \frac{1}{e}$$

$$\beta_{*,1} := \langle e^{-x}, \varphi_1 \rangle = \int_0^1 e^{-x}(2\sqrt{3}x - \sqrt{3})\, dx$$

$$= 2\sqrt{3} \underbrace{\int_0^1 e^{-x} x \, dx}_{\text{integration by parts}} - \sqrt{3} \int_0^1 e^{-x} dx$$

$$= 2\sqrt{3} \underbrace{\{-e^{-x}(x+1)\}\Big|_0^1}_{1 - 2/e} - \sqrt{3}\{1 - 1/e\} = \sqrt{3} - \frac{3\sqrt{3}}{e}$$

Hence,

$$p_*(x) = \{1 - 1/e\}\, 1 + \{\sqrt{3} - 3\sqrt{3}/e\}\,(2\sqrt{3}x - \sqrt{3})$$
$$= (6 - 18/e)\,x + (-2 + 8/e)$$

with error

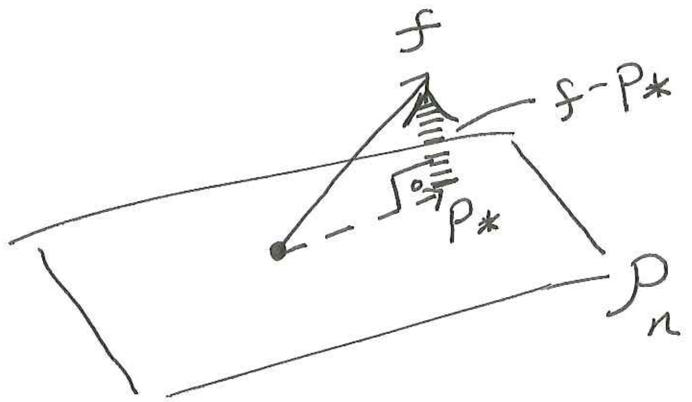
$$\|e^{-x} - p_*\|_2 = \int_0^1 \{e^{-x} + (18/e - 6)x + (2 - 8/e)\}^2\, dx.$$

<u>THM</u>

The following are equivalent:

(1) $p_* \in \mathcal{P}_n$ is such that
$$\|f - p_*\|_2 = \min_{p \in \mathcal{P}_n} \|f - p\|_2 .$$

(2) $p_* \in \mathcal{P}_n$ is such that $\langle f - p_*, q \rangle = 0 \quad \forall q \in \mathcal{P}_n .$



e.g. $f(x) = e^{-x}$

$p_*(x) = (6 - 18/e)x + (-2 + 8/e)$

is s.t. $\|f - p_*\|_2 = \min_{p \in \mathcal{P}_1} \|f - p\|_2$

$$\langle f - p_*, 1 \rangle = \int_0^1 e^{-x} + (18/e - 6)x + (2 - 8/e)\, dx = 0$$

$$\langle f - p_*, x \rangle = \int_0^1 \{ e^{-x} + (18/e - 6)x + (2 - 8/e) \} \, x \, dx = 0$$

To prove the thm on page ④, we will employ the Pythagorean thm.

LEMMA

Let $f, g \in L^2_w(a,b)$ such that $\langle f, g \rangle = 0$.
Then

$$\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2$$

Proof

$$\begin{aligned} \|f+g\|_2^2 &= \langle f+g, f+g \rangle \\ &= \|f\|_2^2 + 2\langle f, g\rangle + \|g\|_2^2 \\ &= \|f\|_2^2 + \|g\|_2^2 . \end{aligned}$$

□

Proof of thm on page ④

(2) $\Longrightarrow$ (1)

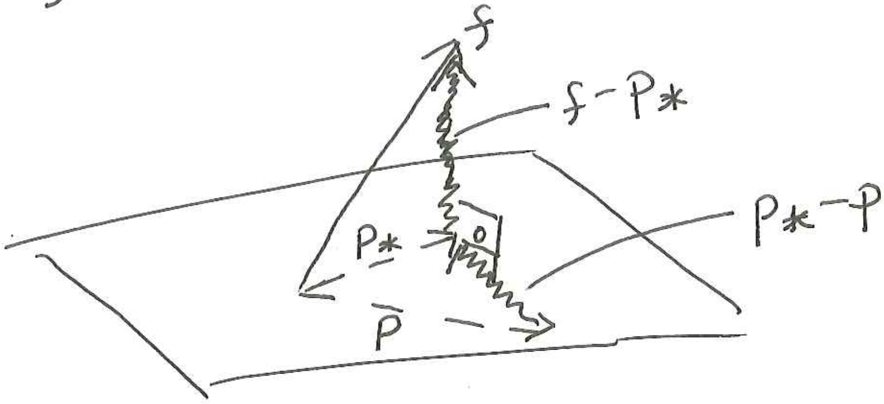
Suppose (2) holds. Consider any $p \in \mathcal{P}_n$, we shall show $\|f - p_*\|_2 \leq \|f - p\|_2$. To this end, by the pythagorean thm as $p_* - p \in \mathcal{P}_n$, we have

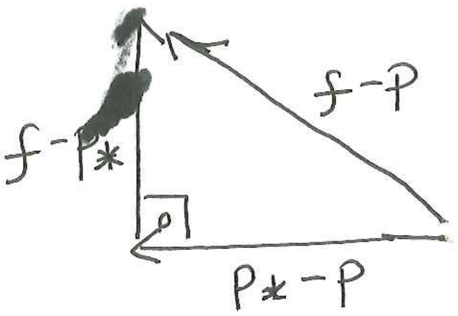
$$\begin{aligned} \|f-p\|_2^2 &= \|(f-p_*) + (p_* - p)\|_2^2 \\ &= \|f - p_*\|_2^2 + \|p_* - p\|_2^2 \end{aligned}$$

Note
$\langle f - p_*, p_* - p \rangle = 0$

$$\Longrightarrow \quad \|f-p\|_2 \geqslant \|f-p_*\|_2$$

hence (1) follows.





(1) $\Longrightarrow$ (2)

As already shown p_* is uniquely given by

$$p_*(x) = \beta_{*,0}\varphi_0(x) + \dots + \beta_{*,n}\varphi_n(x)$$

$$\beta_{*,j} = \langle f, \varphi_j\rangle \cancel{/ \langle \varphi_j, \varphi_j\rangle}$$

for any system $\varphi_0, \varphi_1, \dots, \varphi_n$ of orthogonal polynomials such that $\|\varphi_j\|_2 = 1$ $j = 0, 1, \dots, n$.

It suffices to show $\langle f - p_*, \varphi_j\rangle = 0$ for $j = 0, \dots, n$, as $\varphi_0, \dots, \varphi_n$ form ~~an orthonormal~~ a basis for P_n. To this end,

due to $\langle \varphi_k, \varphi_j\rangle = 0$ for $j \neq k$

$$\langle f - p_*, \varphi_j\rangle = \langle f, \varphi_j\rangle - \beta_{*,j}\underbrace{\langle \varphi_j, \varphi_j\rangle}_{1}$$

$$= 0$$

for $j = 0, 1, \dots, n$. $\square$