Approximation in $\mathcal{L}_2$

(P) Given $f \in C^0([a,b])$, find $P_n \in \mathbb{P}_n$ such that

$$\|f - P_n\|_{\mathcal{L}_2([a,b])} \triangleq \left(\int_a^b (f(x) - P_n(x))^2 \right)^{1/2}$$

is minimal.

Note: we have seen $\mathcal{L}_\infty([a,b])$ error bounds!
Now we want the optimal $\mathcal{L}_2$ bound.

Def: Let V be a linear 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:

(1) $\langle f+g, h \rangle = \langle f, h \rangle + \langle g, h \rangle \quad \forall f, g, h \in V$

(2) $\langle \lambda f, g \rangle = \lambda \langle f, g \rangle \quad \forall \lambda \in \mathbb{R},\ f, g \in V$

(3) $\langle f, g \rangle = \langle g, f \rangle \quad \forall f, g \in V$

(4) $\langle f, f \rangle > 0 \quad \forall f \neq 0 \in V$.

We call the pair $(V, \langle \cdot, \cdot \rangle)$ an inner product space.

Ex: $\mathbb{R}^n$, $\langle \underline{x}, \underline{y} \rangle = \underline{x}^T \underline{y} = \underline{x} \cdot \underline{y}$

Def: If $\langle f, g \rangle = 0$, we say f and g are orthogonal.

Note: on $\mathbb{R}^n$, $\|\underline{x}\|_2^2 = \sum_{i=1}^{n} x_i^2 = \langle \underline{x}, \underline{x} \rangle$.

Def: Let $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ denote an inner product space.

We define $\|f\| = \langle f, f \rangle^{1/2}$ for $f \in \mathcal{V}$.

It turns out $\|f\|$ is a norm, as we now show.

Lemma: (Cauchy-Schwarz)

$$\forall f, g \in \mathcal{V}, \quad |\langle f, g \rangle| \leq \|f\| \, \|g\|$$

Pf: See book.

Hence, we have the triangle inequality,

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

Cauchy-Schwarz

$$\leq \|f\|^2 + \|g\|^2 + 2\|f\| \, \|g\|$$

$$= (\|f\| + \|g\|)^2$$

$$\Rightarrow \|f+g\| \leq \|f\| + \|g\|.$$

and we already have $\|f\| \geq 0$, $\|f\| = 0 \Leftrightarrow f = 0$,

$$\|\lambda f\| = |\lambda| \, \|f\| \quad \forall \lambda \in \mathbb{R}, \ f \in \mathcal{V}$$

from the definition of $\langle \cdot, \cdot \rangle$. This gives the theorem

<u>Thm</u>: The function $\|\cdot\| : \mathcal{V} \to \mathbb{R}_{\geq 0}$ defined by

$$\|f\| = \langle f, f \rangle^{1/2}$$ for $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ an inner-product space is a norm.

<u>EX</u>: Consider the set of functions $C^0([a,b])$ equipped with the inner product

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

with $w(x) \geq 0$ on $[a,b]$ a <u>weight function</u>.

The norm

$$\|f\|_2 = \left(\int_a^b w(x) f(x)^2 \, dx \right)^{1/2}$$

is called the 2-norm on $C^0([a,b])$.

Note: f need not be continuous to have $\|f\|_2 < \infty$!

Def: we define $L^2_w(a,b) \triangleq \{ f : [a,b] \to \mathbb{R} \mid \|f\|_2 < \infty \}$.

Note: $C^0([a,b]) \subsetneq L^2_w(a,b)$.

i.e., all continuous functions are in $L^2_w(a,b)$.

But there are more! e.g. $f(x) = \operatorname{sign}\left(x - \frac{a+b}{2}\right)$.

Let us now return to the problem (P).

Ex: $\varepsilon > 0$, $f(x) = 1 - \exp(-x/\varepsilon)$, $x \in [0,1]$

Consider polynomials of degree 0 (constants).

$$\min_c \int_0^1 (f(x) - c)^2 dx = \min_c \int_0^1 f(x)^2 dx - \int_0^1 2c f(x) dx + c^2.$$

$$\Rightarrow c = \int_0^1 f(x) dx = 1 - \varepsilon + \varepsilon \exp(-1/\varepsilon)$$

$$\Rightarrow p_0^{(2)}(x) = 1 - \varepsilon + \varepsilon \exp(-1/\varepsilon)$$

Now, consider the "minimax" degree 0 polynomial

$$p_0^{(\infty)}(x) \triangleq \arg\min_{c} \max_{x} | f(x) - c |$$

Because f is continuous, monotonically increasing, this is simply the mean of the value at 0 and at 1:

$$p_0^{(\infty)}(x) = 1/2 \left(1 - \exp(-1/\varepsilon)\right)$$

<u>Note:</u> For $\varepsilon \ll 1$, we have:

$$p_0^{(2)}(x) \gtrsim 1$$

$$p_0^{(\infty)}(x) \gtrsim 1/2$$

Changing the norm has a large effect on the resulting polynomial! $\triangle$

For now, assume that for every $f \in L_w^2(a,b)$, the best approximating polynomial (in 2-norm) exists.

For now, set $w = 1$.

We write that

$$p_n(x) = c_0 + c_1 x + \dots + c_n x^n.$$

Goal: Choose the $\{c_i\}$ so that

$$\|e_n\|_2 = \|f - p_n\|_2 = \left(\int_0^1 |f(x) - p_n(x)|^2 dx \right)^{1/2}$$

is minimized.

Note: Equivalent to minimizing

$$\|e_n\|_2^2 = \int_0^1 \left(f(x) - \sum_{j=0}^{n} c_j x^j \right)^2 dx$$

$$= \int_0^1 f(x)^2 dx - \sum_{j=0}^{n} c_j \int_0^1 f(x) x^j dx + \sum_{j,l=0}^{n} c_j c_l \int_0^1 x^{j+l} dx$$

View as a function of $\underline{c} \in \mathbb{R}^{n+1}$, and take gradient.

$$\frac{\partial}{\partial c_k} = 0 = \int_0^1 f(x) x^k dx + \sum_{l=0}^{n} c_l \int_0^1 x^{k+l} dx$$

$$\Rightarrow \underline{\underline{M}}\,\underline{c} = \underline{b}, \quad M_{ij} = \int_0^1 x^{i+j} dx = \frac{1}{i+j+1}$$

$$b_i = \int_0^1 f(x) x^i dx$$

or, using the definition of the inner product,

$$M_{ij} = \langle x^i, x^j \rangle, \quad b_i = \langle f, x^i \rangle$$

To obtain the polynomial, we solve this linear system for $\underline{c} \in \mathbb{R}^{n+1}$.

Note: By changing the def. of $\langle \cdot, \cdot \rangle$, works on any $[a, b]$ and for any weight function w.

Note: The matrix $\underline{\underline{M}}$ is the Hilbert matrix, which we have seen is very poorly conditioned!