Section 2.2 Approximation in Hilbert Spaces

Fourier Series And Completeness:

The following discussion motivates the need for orthogonal polynomials.

- Let H be a Hilbert space.
- Use $\langle \cdot, \cdot \rangle$ to denote the inner product on H
- Use $\| \cdot \|$ to denote the induced norm on H.
  (so that $\|f\| = (\langle f, f\rangle)^{1/2}$ $\forall f \in H$)

Let $f \in H$, and choose the set $\{\phi_1, \phi_2, \dots, \phi_n\}$, $\phi_i \in H$.
Spse we want to approximate f as well as possible using only the functions $\{\phi_1, \phi_2, \dots \phi_n\}$.

(P1) We seek to find $\alpha_1, \alpha_2, \dots, \alpha_n$ so that the "cost function"

$$F(\alpha_1, \alpha_2, \dots, \alpha_n) = \left\| f - \sum_{i=1}^{n} \alpha_i \phi_i \right\|^2$$

is as small as possible.

(P1) is an unconstrained optimization problem on $\mathbb{R}^n$.
Note that

$$F(\alpha_1, \alpha_2, \dots, \alpha_n) = \left\langle f - \sum_{i=1}^{n} \alpha_i \phi_i , f - \sum_{i=1}^{n} \alpha_i \phi_i \right\rangle$$

$$= \langle f, f\rangle - \left\langle \sum_{i=1}^{n} \alpha_i \phi_i, f \right\rangle - \left\langle f, \sum_{i=1}^{n} \alpha_i \phi_i \right\rangle + \left\langle \sum_{i=1}^{n} \alpha_i \phi_i, \sum_{i=1}^{n} \alpha_i \phi_i \right\rangle$$

$$= \langle f, f\rangle - 2 \sum_{i=1}^{n} \alpha_i \langle f, \phi_i \rangle + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j \langle \phi_i, \phi_j \rangle$$

In order for $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ to minimize g, it is necessary that $\nabla F(\alpha_1, \alpha_2, \ldots, \alpha_n) = \vec{0}$.

Here,

$$\nabla F = \begin{bmatrix} \frac{\partial g}{\partial \alpha_1} \\ \frac{\partial g}{\partial \alpha_2} \\ \vdots \\ \frac{\partial g}{\partial \alpha_n} \end{bmatrix}$$

and for $k = 1, 2, \ldots n$

$$\frac{\partial E}{\partial \alpha_k} = -2\langle f, \phi_k \rangle + \sum_{\substack{i=1 \\ i \neq k}}^{n} \alpha_i \langle \phi_i, \phi_k \rangle + \sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j \langle \phi_k, \phi_j \rangle + 2\alpha_k \langle \phi_k, \phi_k \rangle$$

$$= -2\langle f, \phi_k \rangle + 2\sum_{\substack{j=1 \\ j \neq k}}^{n} \alpha_j \langle \phi_j, \phi_k \rangle + 2\alpha_k \langle \phi_k, \phi_k \rangle$$

$$\boxed{\frac{\partial E}{\partial \alpha_k} = -2\langle f, \phi_k \rangle + 2\sum_{j=1}^{n} \alpha_j \langle \phi_j, \phi_k \rangle}$$

Hence, $\nabla F = \vec{0} \Rightarrow \frac{\partial E}{\partial \alpha_k} = 0$ for $k = 1, 2, \ldots n$

This yield

$$2\sum_{j=1}^{n} \alpha_j \langle \phi_j, \phi_k \rangle = 2\langle f, \phi_k \rangle \quad \text{for } k = 1, 2, \ldots n$$

$$\sum_{j=1}^{n} \alpha_j \langle \phi_k, \phi_j \rangle = \langle f, \phi_k \rangle, \quad \text{for } k = 1, 2, \ldots, n$$

This gives us a Linear system of n eqns. for n unknowns $\alpha_1, \alpha_2, \ldots, \alpha_n$. The system can be written in the form

$$A\vec{\alpha} = \vec{\beta}$$

Where →

(I'm assuming scalar field is $\mathbb{R}$ above, no conjugates!)

$$\vec{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \vec{\beta} = \begin{bmatrix} \langle f, \phi_1 \rangle \\ \langle f, \phi_2 \rangle \\ \vdots \\ \vdots \\ \langle f, \phi_n \rangle \end{bmatrix}$$

and

$$A = \begin{bmatrix} \langle \phi_1, \phi_1 \rangle & \langle \phi_1, \phi_2 \rangle & \cdots & \langle \phi_1, \phi_n \rangle \\ \langle \phi_2, \phi_1 \rangle & \langle \phi_2, \phi_2 \rangle & \cdots & \langle \phi_2, \phi_n \rangle \\ \vdots & & & \\ \vdots & & & \\ \langle \phi_n, \phi_1 \rangle & \langle \phi_n, \phi_2 \rangle & \cdots & \langle \phi_n, \phi_n \rangle \end{bmatrix}$$

So, if $A = (a_{ij})$, then $a_{ij} = \langle \phi_i, \phi_j \rangle$.

- If ϕ_i's are linearly independent, then $A\vec{\alpha} = \vec{\beta}$ is uniquely solvable for $\vec{\alpha}$.
- If ϕ_i's form an orthonormal set, then the A matrix reduces to the identity matrix, $A = I$, and
$$\alpha_i = \langle f, \phi_i \rangle$$
- If we choose to then increase the # of functions in our set $\{\phi_i\}$, then in general, all components of $\vec{\alpha}$ would change.
- However, if we begin with an orthonormal set $\{\phi_i\}$ and add functions to the set that preserves the "orthonormality", then the previous values of α_i do not change when ϕ_{n+1}, ϕ_{n+2}, etc. are added!

So, if we re-visit $g(\alpha)$ with stipulation that $\{\phi_i\}$ is an orthonormal set, then

$$\alpha_i = \langle f, \phi_i \rangle \quad \text{for } i=1,2,\dots n$$

and the cost function $g(\alpha)$ takes the form

$$F(\alpha) = \left\| f - \sum_{i=1}^{n} \langle f,\phi_i\rangle \phi_i \right\|^2$$

from pg. 1 of notes

$$= \langle f,f\rangle - 2\sum_{i=1}^{n} \alpha_i \langle f,\phi_i\rangle + \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \langle \phi_i,\phi_j\rangle$$

$$= \|f\|^2 - 2\sum_{i=1}^{n} \alpha_i \langle f,\phi_i\rangle + \sum_{i=1}^{n} \alpha_i^2$$

since $\{\phi_i\}$ is an orthonormal set

$$= \left[\sum_{i=1}^{n} \alpha_i^2 - 2\alpha_i\langle f,\phi_i\rangle + \langle f,\phi_i\rangle^2\right] \underbrace{- \sum_{i=1}^{n} \langle f,\phi_i\rangle^2}_{\text{adding "0" strategically}} + \|f\|^2$$

$$F(\alpha) = \underbrace{\sum_{i=1}^{n} \left(\alpha_i - \langle f,\phi_i\rangle\right)^2}_{①} + \|f\|^2 - \sum_{i=1}^{n} \langle f,\phi_i\rangle^2$$

Hence, we see how choosing $\alpha_i = \langle f,\phi_i\rangle$ for $i=1,2,\dots,n$ minimizes $g(\alpha)$. It makes term ① $=0$, and the remaining two terms of $g(\alpha)$ are actually independent of the values of α_i as they only depend on f and the ϕ_i's that are chosen.

Hence, if we choose $\alpha_i = \langle f,\phi_i\rangle$, then the error we incur by using $\{\phi_i\}_{i=1}^{n}$ to approximate f is given by

must be since it is $\|\cdot\|^2$

$$\left\| f - \sum_{i=1}^{n} \alpha_i \phi_i \right\|^2 = g(\alpha) = \|f\|^2 - \sum_{i=1}^{n} \langle f,\phi_i\rangle^2 \geq 0$$

But note $\sum_{i=1}^{n} \langle f, \phi_i \rangle^2 = \sum_{i=1}^{n} |\alpha_i|^2$, so that the previous inequality implies that

$$\sum_{i=1}^{n} |\alpha_i|^2 = \sum_{i=1}^{n} \langle f, \phi_i \rangle^2 \leq \underbrace{\|f\|^2 < \infty}_{\text{since } f \in H \text{ our Hilbert space}}$$

And this result is true for all n. So, we take the limit as $n \to \infty$ to obtain <u>Bessel's Inequality</u>:

$$\sum_{i=1}^{\infty} \langle f, \phi_i \rangle^2 \leq \|f\|^2$$

Note, $S_n = \sum_{i=1}^{n} \langle f, \phi_i \rangle^2$ forms a sequence of partial sums for $n = 1, 2, 3, \ldots$ This is a sequence of real #s, in particular, it forms a monotone increasing sequence that is bounded by $\|f\|^2 < \infty$. Hence, the sequence of partial sums converges; i.e.

$$\lim_{n \to \infty} \sum_{i=1}^{n} |\alpha_i|^2 = \lim_{n \to \infty} \sum_{i=1}^{n} \langle f, \phi_i \rangle^2$$

exists and is finite. Then we can examine the sequence of functions in H defined by

$$g_n = \sum_{i=1}^{n} \alpha_i \phi_i \quad , \text{ for } n = 1, 2, \ldots$$

Note here that we are still assuming that ϕ_i's form an orthonormal set.

Claim: $\{g_n\}$ is a Cauchy sequence in H.

Justification: Note that

$$\|g_n - g_m\|^2 = \|\sum_{i=1}^{n} \alpha_i \phi_i - \sum_{i=1}^{m} \alpha_i \phi_i\|^2$$

$$= \|\sum_{i=m}^{n} \alpha_i \phi_i\|^2$$ WLOG, assume $n \geq m$

$$= \sum_{i=m}^{n} \|\alpha_i \phi_i\|^2$$ by using orthogonality of ϕ_i's and applying Parallelogram Law repeatedly

$$= \sum_{i=m}^{n} |\alpha_i|^2 \|\phi_i\|^2$$

$$= \sum_{i=m}^{n} |\alpha_i|^2$$ since $\|\phi_i\| = 1$ for all i

$$= \sum_{i=1}^{n} |\alpha_i|^2 - \sum_{i=1}^{n} |\alpha_i|^2$$

$$= |S_n - S_m|$$ $\longrightarrow$ Since $S_n = \sum_{i=1}^{n} \langle f, \phi_i \rangle^2 = \sum_{i=1}^{n} |\alpha_i|^2$

Since S_n is a convergent sequence in $\mathbb{R}$, then S_n is a Cauchy Sequence in $\mathbb{R}$. But by our relation above, then $\{g_n\}$ must be a Cauchy sequence in H.

This follows because for any $\varepsilon > 0$, we can choose N so that for $m, n \geq N$, $|S_n - S_m| < \varepsilon^2$ since $\{S_n\}$ is Cauchy. Then for $m, n \geq N$, we also have

$$\|g_n - g_m\|^2 = |S_n - S_m| < \varepsilon^2 \implies \|g_n - g_m\| < \varepsilon$$

Hence, the sequence $\{g_n\}$ is Cauchy in H. //

Since H is a Hilbert space and since $\{g_n\}$ is a Cauchy sequence in H, then $\exists\ g \in H$ so that

$$g = \sum_{i=1}^{\infty} \alpha_i \phi_i$$

i.e.

$$g = \sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i$$

Observations:

1. The identification $\alpha_i = \langle f, \phi_i \rangle$ (with $\{\phi_i\}$ orthonormal set) induces a correspondence between $f \in L^2$ and the sequence $\{\alpha_i\}$ in the sequence space ℓ^2.
In particular,

$$g = \sum_{i=1}^{\infty} \underbrace{\langle f, \phi_i \rangle}_{\alpha_i} \phi_i \in L^2 \iff \text{the sequence } \{\alpha_i\} \in \ell^2$$

i.e. $\sum_{i=1}^{\infty} |\alpha_i|^2 < \infty$

(This should make sense after our Cauchy sequence argument on the previous page)

2. We refer to $g \in H$ as the projection of f onto the space spanned by the set $\{\phi_i\}$.

3. For any choice of the orthonormal set $\{\phi_i\}$, the function g defined by

$$g = \sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i = \sum_{i=1}^{\infty} \alpha_i \phi_i$$

is the "Best Approximation" of f. (in the sense that this choice of α_i's minimizes the error $\|f - \sum_{i=1}^{\infty} \alpha_i \phi_i\|^2$, where the $\langle \cdot, \cdot \rangle$ & $\|\cdot\|$ are those associated with the particular Hilbert space H.

When is $g = \sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i = f$?

Definition 2.4

An orthonormal set $\{\phi_i\}_{i=1}^{\infty}$ is complete if

$$\sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i = f$$

for every f in the Hilbert space H.

- Note: "Complete Set" is different than a "Complete Vectorspace"

Theorem 2.2

A set $\{\phi_i\}_{i=1}^{\infty}$ is complete in H if and only if any of the following (equivalent) statements hold:

1. $f = \sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i$ for all $f \in H$

2. For $\varepsilon > 0$, there is an $N < \infty$ so that for all $n \geq N$,
$$\| f - \sum_{i=1}^{\infty} \langle f, \phi_i \rangle \phi_i \| < \varepsilon$$

3. $\|f\|^2 = \sum_{i=1}^{\infty} \langle f, \phi_i \rangle^2$ for all $f \in H$ (Parseval's Equality)

4. If $\langle f, \phi_i \rangle = 0$ for all i, then $f = 0$.

5. There is no function $\psi \neq 0$ for which the set $\{\phi_i\}_{i=1}^{\infty} \cup \psi$ forms an orthogonal set.