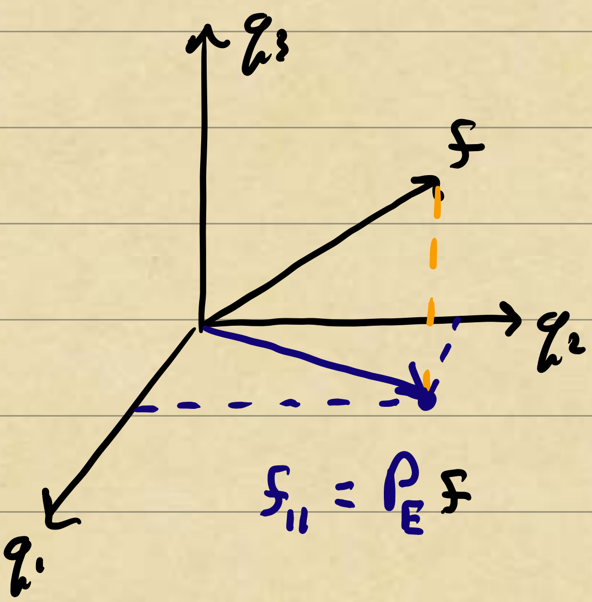
Best Approx. in Hilbert Space

Question: How to choose $c_1, c_2, \ldots, c_n$ so that $f - f_n$ is as "small" as possible?

$$(*) \qquad \underbrace{\begin{bmatrix} | \\ f(x) \\ | \end{bmatrix}}_{f} \approx \underbrace{\begin{bmatrix} | & | & & | \\ e_1 & e_2 & \cdots & e_n \\ | & | & & | \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}}_{\underline{c}}$$

Goal: Minimize $\|f - E\underline{c}\|$ in Hilbert norm.



Two step procedure:

$\Rightarrow$ Project $\quad f_{\parallel} = P_E f$

$\Rightarrow$ Solve $\quad E\underline{c} = f_{\parallel}$

Today: Orthogonal Projection in Hilbert Spaces.

Orthogonal Residual Criterion

We want to establish that $\|f - E\underline{c}\|$ is minimized when $f - E\underline{c} \perp \mathrm{col}(E)$. The key property we need is that $\mathrm{col}(E)$ is closed.

A closed subspace $V \subset H$, where H is a Hilbert space, has $f \in V$ whenever $\{f_n\} \subset V$ converges to f, i.e., $\lim_{n\to\infty} \|f - f_n\| = 0$.

Example: The space $\mathbb{P}_n = \mathrm{span}\{1, x, \ldots, x^n\}$ is a closed subspace of $L^2([-1,1])$, but the space $\mathbb{P} = \bigcup_{n=1}^{\infty} \mathbb{P}_n$ is not closed in $L^2([-1,1])$.

Theorem: Suppose V is a closed subspace of H and $f \in H$. Then,

i) There is a unique $g_* \in V$ s.t.

$$\|f - g_*\| = \inf_{g \in V} \|f - g\|.$$

ii) The element $f - g_*$ is $\perp$ to V, i.e.

$$\langle f - g_*, g \rangle = 0 \quad \text{for all } g \in V$$

pf| If $f \in V$, then $g_* = f$. Otherwise,

$$d = \inf_{g \in V} \|f - g\| > 0 \quad \text{(since } V \text{ closed and } f \notin V\text{)}.$$

Consider a sequence $\{g_n\}_{n=1}^{\infty} \subset V$ s.t.

$$\lim_{n \to \infty} \|f - g_n\| = d.$$

The idea is to show that $\{g_n\}_{n=1}^{\infty}$ is Cauchy and, therefore, has a limit in V. We will use the Parallelogram Law, which states

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2) \quad u, v \in H.$$

(derive by expanding norms in inner products)

Now, set $u = f - g_n$ and $v = f - g_m$, so

$$\|2f - (g_n + g_m)\|^2 + \|g_m - g_n\|^2 = 2(\|f - g_n\|^2 + \|f - g_m\|^2)$$

Since $g_n, g_m \in V$, $g_n + g_m \in V$, and

$$\|2f-(g_n+g_m)\| = 2\|f-\tfrac{1}{2}(g_n+g_m)\| \geq 2d.$$

Therefore, we can bound $\|g_n-g_m\|$ above:

$$\|g_n-g_m\|^2 \leq 2(\|f-g_n\|^2+\|f-g_m\|^2) - 4d^2.$$

Since $\|f-g_n\| \to d$ and $\|f-g_m\| \to d$ by construction, we can for any $\varepsilon > 0$ find $N > 0$ s.t. $n,m > N$ implies $\|g_n-g_m\| < \varepsilon$. Consequently $\{g_n\}_{n\geq 1}^{\infty}$ is Cauchy and has a limit g_* in $\mathcal{H}$. Moreover, since $\{g_n\} \subset V$ and V is closed,

i) $\Rightarrow$ $\lim_{n\to\infty} g_n = g_* \in V.$

(uniqueness later)

To prove (ii), take $g \in V$ and $\varepsilon \in \mathbb{R}$, then

$$\|f-(g_*-\varepsilon g)\|^2 \geq \|f-g_*\|^2$$

Expanding the norms gives

(+) $\quad 2\varepsilon \operatorname{Re}\langle f-g_*, g\rangle + \varepsilon^2\|g\|^2 \geq 0$

If $\mathrm{Re}\langle f-g_*, g\rangle < 0$, then taking $\varepsilon > 0$ suff. small contradicts (+). If $\mathrm{Re}\langle f-g_*, g\rangle > 0$, then taking $\varepsilon < 0$ suff. small contradicts (+). Therefore, $\mathrm{Re}\langle f-g_*, g\rangle = 0$. A similar argument with $g_* \pm i\varepsilon g$ shows $\mathrm{Im}\langle f-g_*, g\rangle = 0$.

ii) $\Rightarrow$ $\langle f-g_*, g\rangle = 0$ for all $g \in V$.

To establish uniqueness of g_*, suppose $\hat{g}_* \in V$ also achieves $\|f-\hat{g}_*\| = d$. Then, by (ii)

$$\langle f-g_*, \overbrace{g_*-\hat{g}_*}^{\in V}\rangle = 0.$$

Pythagorean Thm. $\Rightarrow$ $\|f-\hat{g}_*\|^2 = \|f-g_*\|^2 + \|g-g_*\|^2$

Since $\|f-g_*\| = \|f-\hat{g}_*\| = d$ $\Rightarrow$ $\|g_*-\hat{g}_*\|^2 = 0$ $\square$

Theorem 1 establishes our geometric intuition, that the distance between f and $g \in V$ is minimized when $f-g \perp V$.

In the setting of best approximation by a dictionary, as in (*), Thm 1 tells us there is a unique point $f_{II} \in \mathrm{col}(E)$. This is the projection step (step I) in our two-step framework to solve (*).

Orthogonal Projections

How do we actually compute g_* in Thm 1? We need the machinery of orthogonal projection.

Given a subspace $V \subset H$, the orthogonal complement of V is

$$V^{\perp} = \{ f \in H \,/\, \langle f, g \rangle = 0 \text{ for all } g \in V \}.$$

$V^{\perp}$ is a closed subspace of H and

$$H = V \oplus V^{\perp},$$

meaning that every $f \in H$ has $f = \overset{\in V}{g} + \overset{\in V^{\perp}}{h}$.

This decomposition is a direct consequence of Thm 1, since we can choose $g = g_*$ from i) and by part ii) $h = f - g \perp V$ so $h \in V^\perp$. Note the decomposition is unique!

The map $P_V : H \to V$ defined by $P_V f = g_*$ (from i) in Thm 1) is called the orthogonal projection onto V. It is

i) $f \to P_V f$ is linear
ii) $P_V f = f$ when $f \in V$
iii) $P_V f = 0$ when $f \in V^\perp$
iv) $\|P_V f\| \leq \|f\|$ for all $f \in H$.

<u>Example:</u> Fourier Series : Best Approx. in $L^2([-1,1])$.

A continuous, periodic function f on $[-1,1]$ has

$$f(x) = \frac{1}{\sqrt{2}} \sum_{k=-\infty}^{+\infty} \hat{f}_k e^{i\pi k x}, \qquad \hat{f}_k = \frac{1}{\sqrt{2}} \int_{-1}^{+1} f(x) e^{-i\pi k x} \, dx.$$

Consider the best approximation of f in the subspace $V_N = \{e^{-i\pi N x}, \ldots, e^{+i\pi N x}\}$ of

$$L^2([-1,1]) = \left\{ f : [-1,1] \to \mathbb{R} \,/\, \int_{-1}^{+1} |f(x)|^2 dx < \infty \right\},$$

which is a Hilbert space w/inner product

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

Since $\{e^{ikx}\}_{k=-\infty}^{+\infty}$ are pairwise orthogonal,

$$f(x) = \underbrace{\frac{1}{\sqrt{2}} \sum_{k=-N}^{+N} \hat{f}_k e^{i\pi k x}}_{f_N} + \underbrace{\frac{1}{\sqrt{2}} \sum_{|k|>N} \hat{f}_k e^{i\pi k x}}_{f_N^{\perp}}$$

where $f_N \in V_N$ and $f_N^{\perp} \in V_N^{\perp}$. Therefore, $\langle f - f_N, g \rangle = 0$ for all $g \in V_N$ and f_N is the best approximation of f in V_N. Moreover, if P_N is the orthogonal projection onto V_N,

$$P_N f = \frac{1}{\sqrt{2}} \sum_{k=-N}^{N} \hat{f}_k e^{i\pi k x}, \qquad \hat{f}_k = \frac{1}{\sqrt{2}} \int_{-1}^{+1} f(x) e^{-i\pi k x}.$$

Fourier modes $\{e^{-i\pi N x}, \ldots, e^{+i\pi N x}\}$ are an ONB for V_N.

Orthonormal Basis

An orthonormal basis for a subspace $V_N \subset H$ with dimension $N < \infty$ is a basis $\{q_1, \dots, q_N\} \subset V_N$ such that $\langle q_i, q_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$.

Claim: The orthogonal projection onto V_N can be computed explicitly via

$$P_N f = \sum_{k=1}^{N} \langle f, q_k \rangle q_k .$$

pf/ For any $f \in H$, we have that

$$f = f_N + f_N^{\perp}, \quad \text{where } f_N \in V_N,\ f_N^{\perp} \in V_N^{\perp},$$

and $P_N f = f_N$ by definition. Since $\{q_1, \dots, q_N\}$ is a basis for V_N, there are unique scalars $\alpha_1, \dots, \alpha_N$ s.t.

$$f_N = \alpha_1 q_1 + \cdots + \alpha_N q_N .$$

Using the orthonormality of $\{\varphi_1, \dots, \varphi_N\}$, and the linearity of $\langle \cdot, \cdot \rangle$, we have

$$\langle f_N, \varphi_j \rangle = \alpha_1 \overset{=0}{\langle \varphi_1, \varphi_j \rangle} + \cdots + \alpha_j \overset{=1}{\langle \varphi_j, \varphi_j \rangle} + \cdots + \alpha_N \overset{=0}{\langle \varphi_N, \varphi_j \rangle}$$

$$= \alpha_j$$

Therefore, $f_N = \sum_{k=1}^{N} \langle f, \varphi_k \rangle \varphi_k$ as claimed.