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?

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

Idea: Develop "linear algebra for functions" to solve $(*)$ in a "least-squares" sense.

Recall: A real or complex Hilbert space H is a complete inner product space over $\mathbb{R}$ or $\mathbb{C}$.

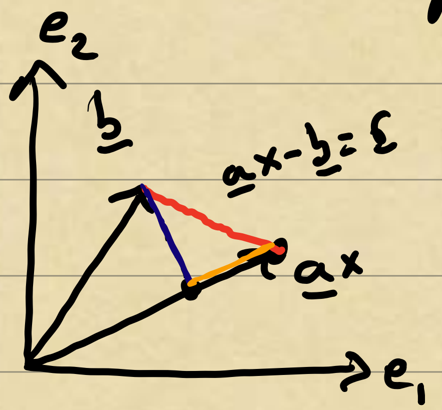
Today: The solution to $(*)$ is obtained by the orthogonal residual condition

$$\underline{f} - \underline{E}\,\underline{c} \perp \overbrace{\text{span}\{e_1, e_2, \ldots, e_n\}}^{\text{col}(E)},$$

and we can calculate $\underline{c}$ via ONB for $\text{col}(E)$.

Orthogonal Residual Criterion

The basic intuition behind least-squares approximation is captured in a picture:



$$\begin{bmatrix} | \\ \underline{a} \\ | \end{bmatrix} x = \begin{bmatrix} | \\ \underline{b} \\ | \end{bmatrix}$$

$$\|\underline{r}\|^2 = \|\underline{a}x - \underline{b}\|^2 = \|\underline{r}_\perp\|^2 + \|\underline{r}_\parallel\|^2$$

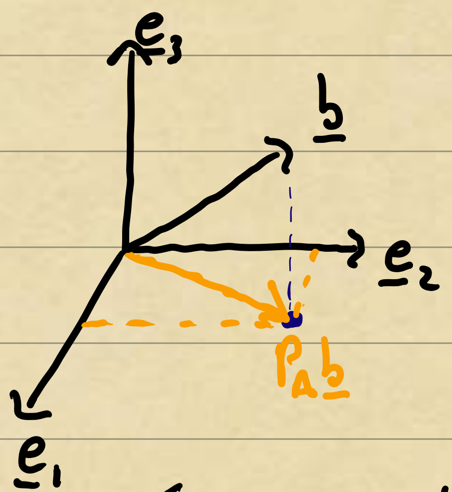
We can choose x to make $\underline{r}_\parallel = 0$, but we can never change $\underline{r}_\perp$ by changing x.

$$\underline{r}_\parallel = \underline{a}\,\frac{\underline{a}^T\underline{r}}{\underline{a}^T\underline{a}} = \frac{\underline{a}}{\underline{a}^T\underline{a}}\left[\underline{a}^T(\underline{a}x - \underline{b})\right]$$

So we choose x to make $\underline{r}_\parallel = 0$:

$$x = \frac{\underline{a}^T\underline{b}}{\underline{a}^T\underline{a}}$$

The basic picture remains the same in higher, even infinite, dimensions. However, its useful to formulate the solution via ONB.



$$\min_{\underline{x} \in \mathbb{R}^2} \left\| \underbrace{\begin{bmatrix} | & | \\ \underline{a}_1 & \underline{a}_2 \\ | & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\underline{x}} - \underbrace{\begin{bmatrix} | \\ \underline{b} \\ | \end{bmatrix}}_{\underline{b}} \right\|$$

Suppose that $\{e_1, e_2\}$ is an ONB for col(A), i.e.,

$$\Rightarrow \quad \text{span}\{e_1, e_2\} = \text{span}\{\underline{a}_1, \underline{a}_2\}.$$

$$\Rightarrow \quad \langle e_i, e_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

We can solve the least-squares problem by projecting $\underline{b}$ orthogonally onto col(A) and solving the new linear system for $\underline{x}$.

Project $\Rightarrow$ $P_A \underline{b} = \langle \underline{b}, \underline{e}_1 \rangle \underline{e}_1 + \langle \underline{b}, \underline{e}_2 \rangle \underline{e}_2$

Solve $\Rightarrow$ $A \underline{x} = P_A \underline{b}$

We can solve for $\underline{x}$ by expanding the columns of A in the basis $\{e_1, e_2\}$.

$$\underline{a}_1 = c_1^1 \underline{e}_1 + c_2^1 \underline{e}_2 , \qquad \underline{a}_2 = c_1^2 \underline{e}_1 + c_2^2 \underline{e}_2$$

$$\Rightarrow \quad \underbrace{A}_{} \underline{x} = \underbrace{\begin{bmatrix} | & | \\ \underline{a}_1 & \underline{a}_2 \\ | & | \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ e_1 & e_2 \\ | & | \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} c_1^1 & c_1^2 \\ c_2^1 & c_2^2 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\underline{x}}$$

Now, to solve $A\underline{x} = \underline{b}$, we multiply both sides by E^T and use orthonormality ($E^T E = I$)

$$A\underline{x} = \underline{b} \quad \Rightarrow \quad EC\underline{x} = E\hat{b} \qquad \left(\hat{b} = \begin{bmatrix} \langle b, e_1 \rangle \\ \langle b, e_2 \rangle \end{bmatrix}\right)$$

$$(**) \qquad \Rightarrow \quad C\underline{x} = \hat{b}$$

Note that even if the columns of A and $\underline{b}$ are elements of an abstract Hilbert space (e.g., could be functions instead of vectors in $\mathbb{R}^n$), working with an ONB for $\text{col}(A) = \text{span}\{\text{cols of } A\}$ allows us to solve $(*)$ by solving a finite linear system $(**)$.