# Interpolation vs. Best Approx.

Given $f \in H$ and dictionary $\{e_1, \dots, e_N\} \subset H$, we can find the best approximation to $f$ in $\mathrm{span}\{e_1, \dots, e_N\}$ via QR decomp:

**Algorithm:** Compute best approx. via QR.

Step 1. QR decomposition of dictionary:

$$\underset{\text{dictionary}}{\underbrace{\overset{E}{\begin{bmatrix} | & & | \\ e_1 & \cdots & e_N \\ | & & | \end{bmatrix}}}} = \underset{\text{ONB}}{\underbrace{\overset{Q}{\begin{bmatrix} | & & | \\ q_1 & \cdots & q_N \\ | & & | \end{bmatrix}}}} \underset{\text{change of basis}}{\underbrace{\overset{R}{\begin{bmatrix} \times & \cdots & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \times \end{bmatrix}}}}$$

Step 2. Solve upper triangular system

$$R\underline{c} = Q^* f$$

**Question:** What if we only have samples of $f : \underset{\subset \mathbb{R}^d}{\Omega} \to \mathbb{R}$ (continuous) at $\underbrace{x_1, \dots, x_M}_{\text{distinct}}$?

$$f_1 = f(x_1), \dots, f_M = f(x_M) \ ?$$

$\Rightarrow$ Interpolate or otherwise "fit" data with dict.

# Interpolation/Regression

How to choose combination of dictionary functions to interpolate/fit data?

$$f(x) \approx c_1 e_1(x) + \cdots + c_N e_N(x)$$

Solve $M \geq N$ equations for $c_1, \ldots, c_N$:

$$c_1 e_1(x_1) + c_2 e_2(x_1) + \cdots + c_N e_N(x_1) = f(x_1)$$

$$\vdots$$

$$c_1 e_1(x_M) + c_2 e_2(x_M) + \cdots + c_N e_N(x_M) = f(x_M)$$

Or in matrix notation, we have

Generalized Vandermonde

$$\begin{bmatrix} e_1(x_1) & e_2(x_1) & \cdots & e_N(x_1) \\ e_1(x_2) & e_2(x_2) & \cdots & e_N(x_2) \\ \vdots & \vdots & & \vdots \\ e_1(x_M) & e_2(x_M) & \cdots & e_N(x_M) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_M \end{bmatrix}$$

$$E_{N,M} \quad \underline{c} = \underline{f}$$

Case 1 If $N = M$ and $E_{N,M}$ is invertible, then $f_{N,M} = E\underline{c}$ will interpolate the data.

Case 2 If $N > M$ and $E_{N,M}$ has linearly indep. rows, then we can perform model regression on the data by solving in least-squares sense.

$$\Rightarrow E_{N,M} = Q_{N,M} R_{N,M} \quad \Rightarrow \underline{c} = R_{N,M}^{-1} Q_{N,M}^{*} \underline{f}$$

Case 3 If $N < M$ and $E_{N,M}$ has linearly indep. rows, then there are many solutions! We often look for a minimum-norm solution or impose additional constraints on model.

$\Rightarrow$ Could also impose sparsity, smoothness, etc.

Question: What happens if $x_1, \ldots, x_M$ are <u>not</u> all distinct points in $\Omega$?

$\Rightarrow E_{N,M}$ has repeated rows. Must modify. Taylor?

Similar to Best-Fit approximations, the condition number of $E_{N,M}$ as $N, M \to \infty$ is typically the key factor in convergence.

$N = M$ $$\kappa(E_N) = \|E_N\| \|E_N^{-1}\|$$

$N \geq M$ $$\kappa(E_{N,M}^* E_{N,M}) = \|E_{N,M}^* E_{N,M}\| \|(E_{N,M}^* E_{N,M})^{-1}\|$$

## Polynomial Interpolation

To illustrate the basic analysis framework consider $\{x_0, \dots, x_N\} \in [-1, 1]$ distinct and

$$E = \begin{bmatrix} | & | & & | \\ 1 & x & \cdots & x^N \\ | & | & & | \end{bmatrix}.$$

The Vandermonde matrix is invertible:

$$E_N = \begin{bmatrix} 1 & x_0 & \cdots & x_0^N \\ 1 & x_1 & \cdots & x_1^N \\ \vdots & \vdots & & \vdots \\ 1 & x_N & \cdots & x_N^N \end{bmatrix}.$$

$\Rightarrow$ Degree N polynomial interpolant through N+1 distinct points is unique

pf] If $p$ and $p'$ are degree N and interpolate $f$ at N+1 distinct points, then $p - p'$ is degree $\leq N$ and interpolates zero at N+1 points. The only deg $\leq N$ like this is the zero poly.

$$\Rightarrow \quad p = p'$$

For any $f : C[-1,1] \to \mathbb{R}$ and any set of distinct points $X = \{x_0, \dots, x_n\}$, we define

$$p = P_X f$$

as the unique degree $\leq N$ interpolant of $f$ on the set $X$.

The map $P_X : C[-1,1] \to \mathbb{P}_n$ is key to understanding the approximation qualities of polynomial interpolants on $X$:

$P_x$ is linear: $P_x(f + \beta g) = P_x f + \beta P_x g$

$P_x$ is onto $\mathbb{P}_N$: $\quad p = P_x f \in \mathbb{P}_N$

$P_x^2 = P_x$ : $\quad p = P_x f \qquad p = P_x p$

These three properties make $P_x$ a projection of $C[-1,1]$ onto $\mathbb{P}_n$. However, it is not, in general, an orthogonal projection! In particular,

$$\|P_x\| = \sup_{f \in C[-1,1]} \frac{\|P_x f\|}{\|f\|} > 1.$$

For "bad" interpolation points, $\|P_x\|$ can grow exponentially with $n$. For "good" point sets, $\|P_x\|$ grows modestly with $n$ and $P_x f$ behaves similarly to the "best" (orthogonal) projection onto $\mathbb{P}_n$.