CS450: Numerical Analysis Interpolation

Howard Hao Yang

Assistant Professor, ZJU-UIUC Institute

20/12/2023

Announcement

Groups → TA

Jan. 12 → final exam

- Next week: presentation
- Dec. 22nd: review class
- Will release some sample problem sets later

Today’s Agenda

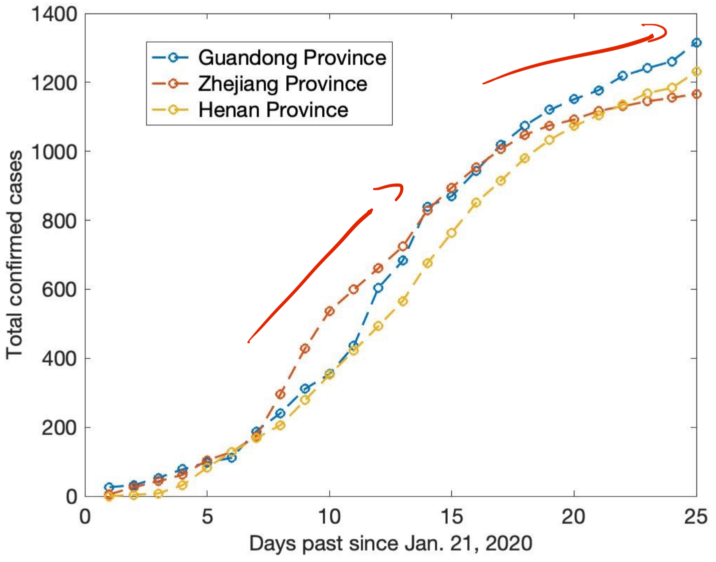
numerical

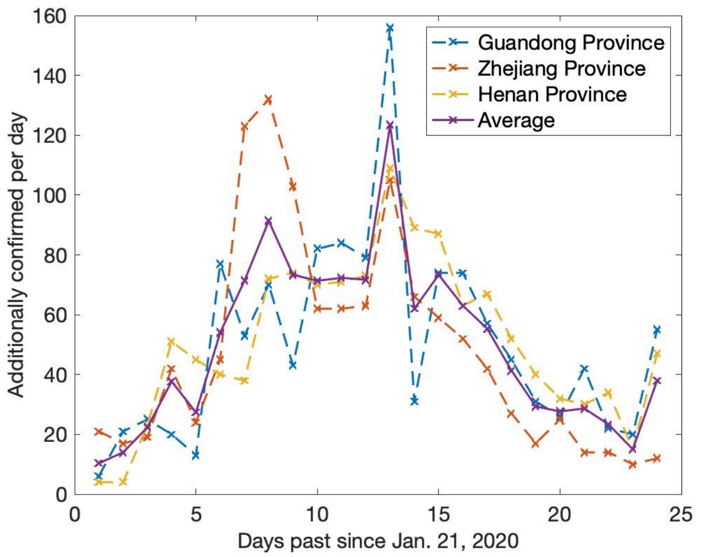
persp.

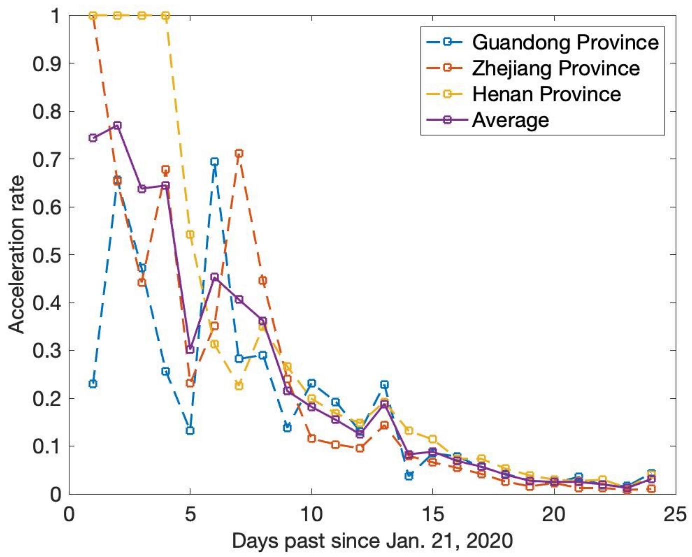
- Interpolation: Problem setup
- Choosing the basis
- Orthogonalizing functions and Chebyshev interpolation

Illustrating Example

- The data I collected during the early stage of COVID-19…

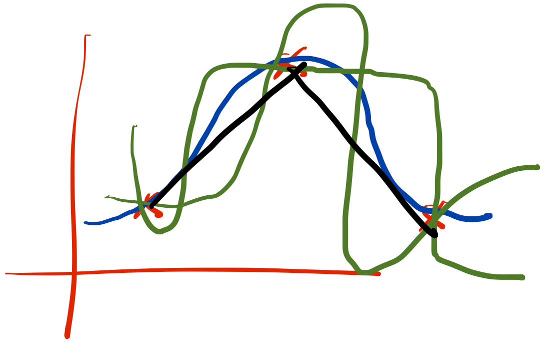






Problem Setup

- What we are given: a set of data points $(x_i, y_i)_{i=1}^{m}$
- What we want: a function f that satisfies $f(x_i) = y_i$
  - represent discrete data by relatively simple functions that are easily manipulated
  - such a function is termed an interpolant
- Can we do this?
- Does the problem have a unique answer? Why?



Making the Problem Unique

10

- Given a set of data points $(x_i, y_i)_{i=1}^{m}$, an interpolation is chosen from the space of functions spanned by a suitable set of basis functions $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$, whereas the interpolating function f is defined as

Fourier series.

find

$$f(x) = \sum_{j}^{n} \alpha_j \phi_j(x)$$

$f(x_i) = y_i$

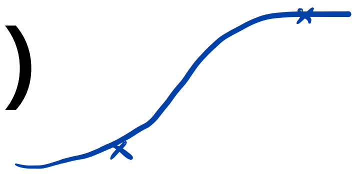
where the parameters α_j are to be determined from the data points $(x_i, y_i)_{i=1}^{m}$

- Two main questions to ask:
  - What basis shall we choose?
  - How many coefficients shall we assign?

ideally → cos(x). sin(x) e^x log x x^2

Making the Problem Unique (Cont'd)

$(x_1, y_1) \quad (x_2, y_2)$



- What shall n be if we want unique solution (suppose $m = 2$)?
- By imposing $f(x_i) = y_i$, we can write the problem as a system of equations:

$f(x) = \sum \alpha_i \phi_i(x)$

$f(x_1) = y_1$

$f(x_2) = y_2$

$\Rightarrow$

$$\begin{cases} \alpha_1 \phi_1(x_1) + \alpha_2 \phi_2(x_1) + \cdots + \alpha_n \phi_n(x_1) = y_1 \\ \alpha_1 \phi_1(x_2) + \alpha_2 \phi_2(x_2) + \cdots + \alpha_n \phi_n(x_2) = y_2 \end{cases}$$

$$\underbrace{\begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_n(x_2) \end{bmatrix}}_{V} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$= ?$

$n > m$

$n = m.$

α_1, α_2

$(x_1, y_1) \cdots (x_m, y_m)$

$\rightarrow \phi_1, \cdots, \phi_m$

Choosing the Basis

$\phi_i(x) = x^{i-1}$

- The choice of basis functions has a dramatic effect on the cost of computing and manipulating the interpolant

$\phi_n(x)$ $\phi_1(x)$ $\phi_0(x)$

- One natural choice for interpolation is via the algebraic polynomials

$$P_n(x) = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \cdots + \alpha_1 x + \alpha_0$$

- Why? – **Weierstrass Approximation Theorem**: Suppose f is defined and continuous on a closed interval $[\mathrm{a}, \mathrm{b}]$. For each $\epsilon > 0$, there exists a polynomial P with the property that

$$|f(x) - P(x)| < \epsilon, \text{ for all } x \in [a, b]$$

Illustrative Exercise

$$\begin{cases} \phi_0(x) = x^0 \\ \phi_1(x) = x^1 \\ \phi_2(x) = x^2 \end{cases}$$

f

- Determine a polynomial of degree two that interpolates the three data points $(-2,-27)$, $(0,-1)$, and $(1,0)$

3 data samples

2-degree

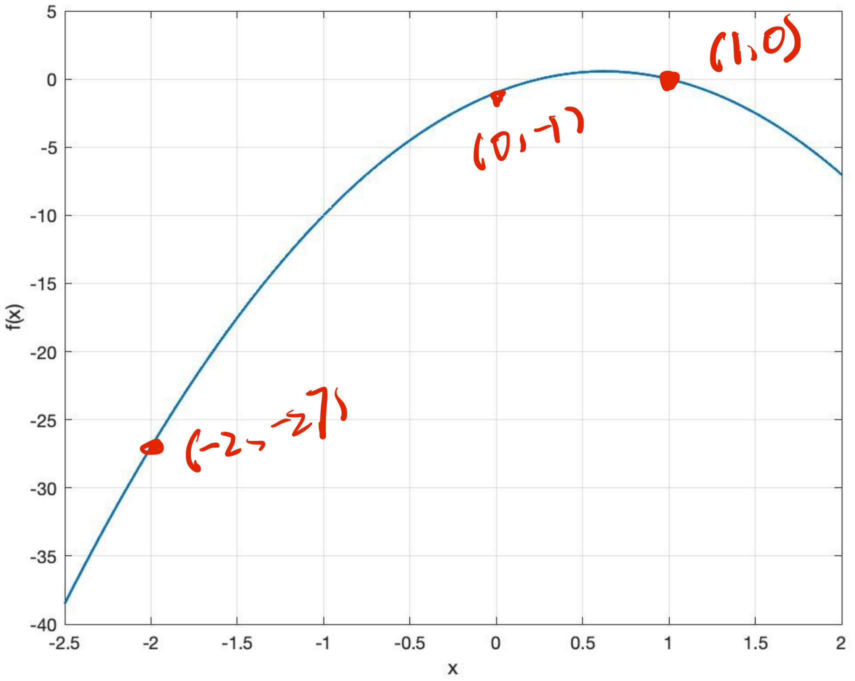
- Step 1: suppose the polynomial takes the form $f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2$
- Step 2: assign $f(-2) = -27$, $f(0) = -1$, $f(1) = 0$; which can be written as

$$\begin{bmatrix} 1 & -2 & 4 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} -27 \\ -1 \\ 0 \end{bmatrix}$$

- Step 3: solve the linear system of equations, obtain $\alpha_0 = -1$, $\alpha_1 = 5$, $\alpha_2 = -4$; hence $\boxed{f(x) = -1 + 5x - 4x^2}$

Illustrative Exercise (Cont'ed)

- Determine a polynomial of degree two that interpolates the three data points $(-2,-27)$, $(0,-1)$, and $(1,0)$
- An interpolation via degree-two polynomial $f(x) = -1 + 5x - 4x^2$



Monomial Basis

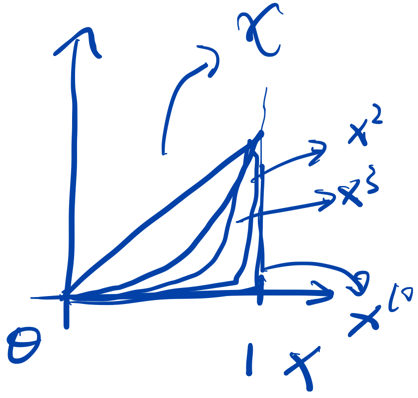
$\{(x_i, y_i)\}_{i=1}^{m}$

$\rightarrow f(x) = \alpha_m x^{m-1} + \alpha_{m-1} x^{m-2} + \cdots + \alpha_0$

$x_i \neq x_j$

- Choose $\phi_j(x) = x^{j-1}$ for $j = 1, \ldots, m$

$$\underline{\Phi(\boldsymbol{x})\boldsymbol{\alpha}} = \underbrace{\begin{bmatrix} 1 & x_1 & \cdots & x_1^{m-1} \\ \vdots & & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{m-1} \end{bmatrix}}_{m} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \underline{\boldsymbol{y}}$$



- The matrix $\Phi(\boldsymbol{x})$ is generally called Vandermonde matrix
- The problem is now equivalent to solving a linear system of equations
- Is $\Phi(\boldsymbol{x})$ invertable?

$\det(\Phi(x)) = \prod_{j>i} (x_j - x_i) \neq 0 \rightarrow \alpha = \Phi(x)^{-1} y$

- What is the (potential) issue here?

$x_i \in [0, 1]$

$x^n \approx x^{n+1}$

Lagrange Basis

ZJU-UIUC INSTITUTE

- Underlying philosophy: Given a set of data points $(x_i, y_i)_{i=1}^{m}$, find basis ① functions $L_1(x), L_2(x), \ldots, L_m(x)$ to be ② polynomial functions of degree $m-1$, such that ③

?

$$L_i(x_j) = \begin{cases} 1, & \text{if} \quad j = i \\ 0, & \text{otherwise} \end{cases}$$

$x = x_1$

$L_1(x_1) = 1$

$L_i(x_1) = 0, \; i \geq 2$

$f(x_1) = y_1 \cdot L_1(x_1) = y_1$

- Then, the interpolation is given by

$$f(x) = y_1 L_1(x) + y_2 L_2(x) + \cdots + y_m L_m(x)$$

- Q: How to construct $L_i(x)$? $\longrightarrow$ $\{(x_i, y_i)\}_{i=1}^{m}$

Illustrative Exercise

(x_2, y_2)

(x_1, y_1)

- Determine the Lagrange basis that interpolates the two data points (x_1, y_1) ~~and~~ , (x_2, y_2) , (x_3, y_3) → $L_1(x)$, $L_2(x)$, $L_3(x)$ 2-order ⇒ $L_1(x_1)=1$, $L_1(x_2)=L_1(x_3)=0$
- Step 1: How would the basis look like?

$$L_1(x) = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)}$$

$$\boxed{L_1(x) = \frac{x-x_2}{x_1-x_2}} \quad \text{and} \quad L_2(x) = \frac{x-x_1}{x_2-x_1}$$

$$L_2(x) = \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)}$$

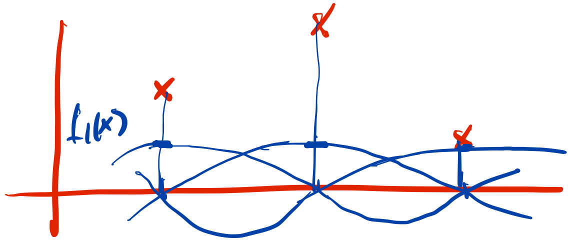
- Check: $L_1(x_1) = 1$, $L_1(x_2) = 0$, $L_2(x_1) = 0$, $L_2(x_2) = 1$
- Step 2: The interpolation is

$f(x_1) = y_1$

$f(x_2) = y_2$

$$f(x) = y_1 L_1(x) + y_2 L_2(x) = y_1 \cdot \frac{x-x_2}{x_1-x_2} + y_2 \cdot \frac{x-x_1}{x_2-x_1}$$

Lagrange Basis (Cont'd)



- General setting: Given a set of data points $(x_i, y_i)_{i=1}^{m}$
- Form $L(x)$ as $\boxed{L(x) = \prod_{i=1}^{m}(x - x_i)}$
- Form the barycentric weight at the j-th node

$$w_j = \frac{1}{\prod_{i=1}^{m}(x_j - x_i)}$$

$(x - x_j)$

$$\frac{(x-x_1)(x-x_2)\cdots(x-x_m)}{(x_j-x_1)(x_j-x_2)\cdots(x_j-x_m)}$$

- Form fundamental polynomials as

$$L_j(x) = \frac{L(x) \cdot w_j}{x - x_j}$$

$$L_j(x_i) = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases}$$

- What is the order of $L_j(x)$?

Lagrange Basis (Cont'd)

- The interpolation is termed fundamental polynomial, expressed as

$$P_{m-1}(x) = \sum_{j=1}^{m} y_j \cdot L_j(x) = L(x) \sum_{j=1}^{m} \frac{y_j w_j}{x - x_j}$$

- How much is the approximation error?
- Suppose $x_1, \ldots, x_m$ are distinct numbers in [a, b] and $f \in C^m[a, b]$. Then, for each $x \in$ [a, b], a number $\min\{x_i\} \le \xi(x) \le \max\{x_i\}$ exists such that

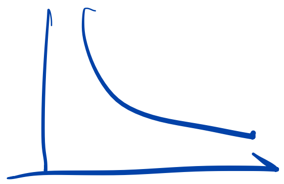
$$f(x) = P_{m-1}(x) + \boxed{\frac{f^{(m)}(\xi(x))}{m!} \prod_i (x - x_i)}$$

Lagrange Basis (Cont'd)

ZJU-UIUC INSTITUTE

- Suppose $P(x)$ is the Lagrange interpolant for the m distinct data points $(x_1, y_1), \dots, (x_m, y_m)$; then $P(x)$ is a polynomial of order $m-1$
- Suppose $\hat{P}(x)$ is another polynomial that has order $n < m-1$ but also interpolate the same set of data points
- Key fact: a polynomial of order n has at most n zeros or is identically zero
- Observation: $P(x) - \hat{P}(x)$ is of order $m-1$ but has m zeros
- Conclusion: $P(x) = \hat{P}(x)$, namely, the Lagrange interpolation finds the lowest order polynomial that interpolates the m distinct data points $(x_1, y_1), \dots, (x_m, y_m)$

Working Exercise



- Use the nodes $x_1 = 2$, $x_2 = 2.75$, and $x_3 = 4$ to determine the Lagrange interpolating polynomial for $\tilde{f}(x) = \frac{1}{x}$

$\left(2, \frac{1}{2}\right), \left(2.75, \frac{1}{2.75}\right), \left(4, \frac{1}{4}\right)$

degree 2

- Use the interpolation to approximate $\tilde{f}(3) = \frac{1}{3}$
- How large is the error?

$$L_1(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}$$

$$= \frac{(x - 2.75)(x - 4)}{(2 - 2.75)(2 - 4)}$$

Working Exercise (Cont'd)

ZJU-UIUC INSTITUTE

- Use the nodes $x_1 = 2$, $x_2 = 2.75$, and $x_3 = 4$ to determine the Lagrange interpolating polynomial for $\tilde{f}(x) = \frac{1}{x}$
- Step 1: form the Lagrange basis

$$L_1(x) = \frac{(x-2.75)(x-4)}{(2-2.75)(2-4)} = \frac{2}{3}(x-2.75)(x-4)$$

$$L_2(x) = \frac{(x-2)(x-4)}{(2.75-2)(2.75-4)} = -\frac{16}{15}(x-2)(x-4)$$

$$L_3(x) = \frac{(x-2)(x-2.75)}{(4-2)(4-2.75)} = \frac{2}{5}(x-2)(x-2.75)$$

- Step 2: calculate $y_1 = \tilde{f}(x_1) = \frac{1}{2}$, $y_2 = \tilde{f}(x_2) = \frac{1}{2.75} = \frac{4}{11}$, $y_3 = \tilde{f}(x_3) = \frac{1}{4}$

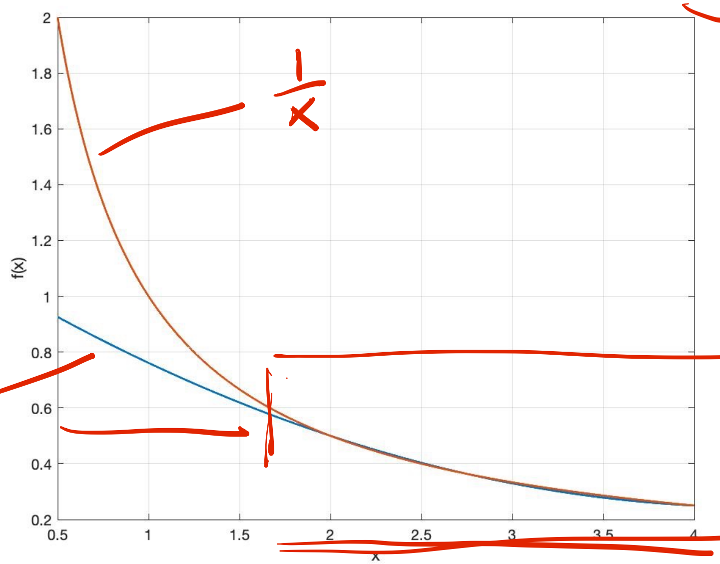
Working Exercise (Cont'd)

- Use the nodes $x_1 = 2$, $x_2 = 2.75$, and $x_3 = 4$ to determine the Lagrange interpolating polynomial for $\tilde{f}(x) = \frac{1}{x}$

$\sum_{n=0}^{\infty} (-1)^n (x-1)^n$

- Step 3: calculate the interpolation as

$$f(x) = y_1 L_1(x) + y_2 L_2(x) + y_3 L_3(x) = \frac{x^2}{22} - \frac{35}{88}x + \frac{49}{44}$$



$f(x)$

Working Exercise (Cont'd)

- Use the nodes $x_1 = 2$, $x_2 = 2.75$, and $x_3 = 4$ to determine the Lagrange interpolating polynomial for $\tilde{f}(x) = \frac{1}{x}$
- The interpolation is given as

$$f(x) = y_1 L_1(x) + y_2 L_2(x) + y_3 L_3(x) = \frac{x^2}{22} - \frac{35}{88}x + \frac{49}{44}$$

- The evaluation at $x = 3$ is $f(3) = \frac{9}{22} - \frac{105}{88} + \frac{49}{44} \approx 0.32955$
- The error is $\left|\tilde{f}(3) - f(3)\right| = 0.33333 - 0.32955 = 0.00378$

Issues? And How to Address...

- For monomial basis, possibly ill-conditioned
  - Possibly some terms become closely to linearly dependent
- For Lagrange basis, cheap to form but expensive to evaluate
- How to strike a balance between these two?

① $(x_i, y_i)_{i=1}^{m}$

$f(x) = \alpha_m x^{m-1} + \alpha_{m-1} x^{m-2} + \cdots + \alpha_0$

$\Phi(x)\alpha = y$

$\vec{\alpha} = \Phi^{-1}(x)\, y$

② $L_i(x) = \prod_j \frac{(x - x_j)}{(x_i - x_j)}$

Newton Interpolation

$\pi_0(x) = 1$

$\pi_1(x) = (x - x_1)$

$\pi_2(x) = (x - x_1)(x - x_2)$

ZJU-UIUC INSTITUTE

- Given a set of data points $(x_i, y_i)_{i=1}^{m}$, form the following basis

$$\pi_j(x) = \prod_{k=1}^{j}(x - x_k), \quad j = 0, \cdots, m-1$$

  - Only requires multiplication, no division is needed
  - If $i < j$, what is the value of $\pi_j(x_i)$? $\longrightarrow 0$ $\quad \pi_2(x_1) = 0$

- The interpolation is given by

$$f(x) = \alpha_0\pi_0(x) + \alpha_1\pi_1(x) + \cdots + \alpha_{m-1}\pi_{m-1}(x)$$

- What's the advantage of this form?

Illustrative Example

- Use the Newton interpolation to interpolate (x_1, y_1), (x_2, y_2), (x_3, y_3)
- The Newton basis are $\pi_0(x) = 1$, $\pi_1(x) = (x - x_1)$, $\pi_2(x) = (x - x_1)(x - x_2)$
- The targeted interpolating polynomial is

$f(x_1) = \alpha_0 \underbrace{\pi_0(x_1)}_{1} + \alpha_1 \pi_1(x_1) = 0$

$+ \alpha_2 \pi_2(x_1) = 0$

$$f(x) = \alpha_0\pi_0(x) + \alpha_1\pi_1(x) + \alpha_2\pi_2(x)$$

- By assigning $f(x_1) = y_1$, $f(x_2) = y_2$, $f(x_3) = y_3$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & \pi_1(x_2) & 0 \\ 1 & \pi_1(x_3) & \pi_2(x_3) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

?

- The matrix is a lower-triangular matrix, complexity in solving the linear equation system is $O(n^2)$

Newton Interpolation (Cont'd)

- Given a set of data points $(x_i, y_i)_{i=1}^{m}$, form the following basis

$$\pi_j(x) = \prod_{k=1}^{j}(x - x_k), \quad j = 0, \cdots, m-1$$

- The interpolation of a $m-1$ degree polynomial is given by

$$f_{m-1}(x) = \alpha_0\pi_0(x) + \alpha_1\pi_1(x) + \cdots + \alpha_{m-1}\pi_{m-1}(x)$$

- If we add a new data point (x_{m+1}, y_{m+1}) to the collection, it follows that

$$f_m(x) = f_{m-1}(x) + \alpha_m\pi_m(x)$$

$$\alpha_m = \frac{y_{m+1} - f_{m-1}(x_{m+1})}{\pi_m(x_{m+1})}$$

  - This new polynomial $f_m(x)$ is of degree m
  - The polynomial $f_m(x)$ interpolates the same m data points

Orthogonality of Functions

- Given two functions f and g, how to define orthogonality
  - If discrete: $\langle f, g \rangle = \sum_i f_i \cdot g_i = 0$
  - If continuous: $\langle f, g \rangle = \int f(x) g(x) dx = 0$
- How can we practically obtain a set of orthogonal functions?
- A very elegant theory about the three-term recurrence:

$$\phi_{k+1}(x) = (\alpha_k x + \beta_k)\phi_k(x) - \gamma_k \phi_{k-1}(x)$$

Chebyshev Polynomials: Definitions

- Three term recurrence:

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$$

where $T_1(x) = x$ and $T_0(x) = 1$

- An analytical solution

$$T_k(x) = \cos(k \cos^{-1}(x))$$

Chebyshev Polynomials

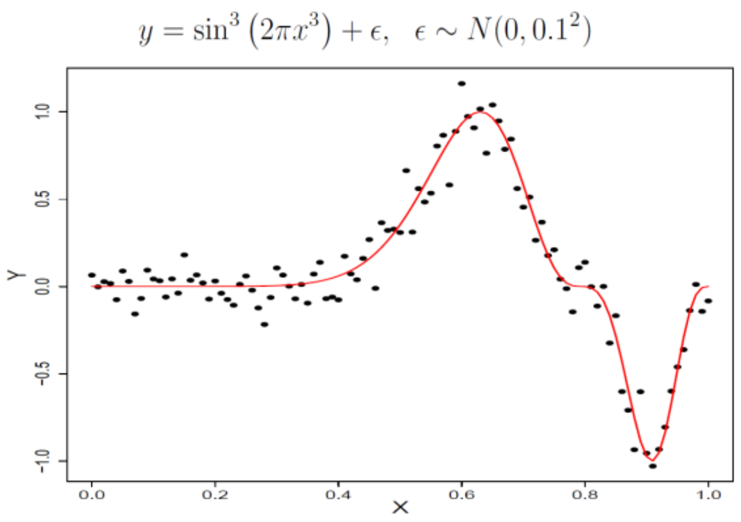
- If x_i is chosen as $x_i = \cos(\frac{i}{k}\pi)$
- Then the basis matrix has the following form

$$\Phi_{i,j}(x) = \cos(j\cos^{-1}(\cos(\frac{i}{k}\pi))) = \cos(j\frac{i}{k}\pi)$$

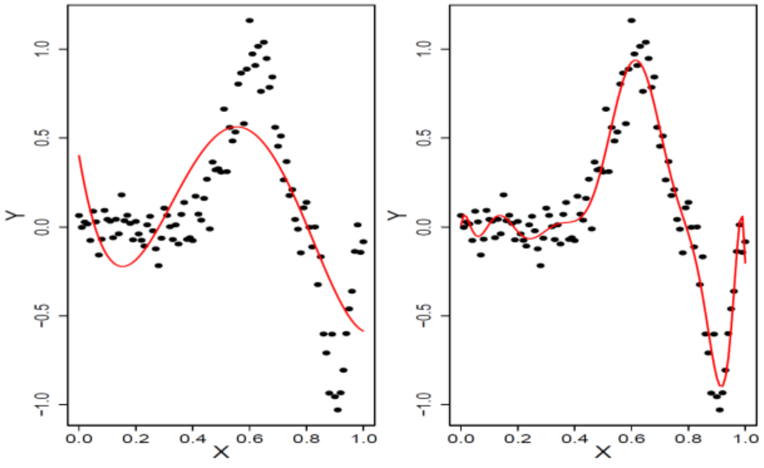
- This is termed the Discrete Cosine Transform (DCT)

Polynomial Regression

- Disadvantages
  - Remote part of the function is very sensitive to outliers
  - Less flexibility due to global functional structure



Example from Ji Zhou, 2011

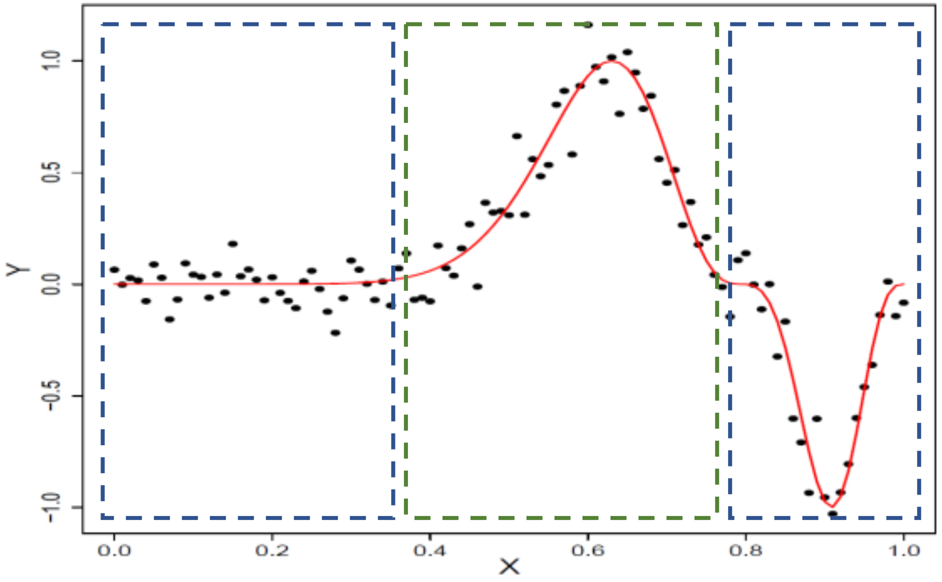


Estimated using polynomials

Splines

- Unique features
  - Linear combination of piecewise polynomial functions under continuity assumption
  - Partition the domain of x into continuous intervals and fit polynomials in each interval
  - Provides flexibility and local fitting

$$y = \sin^3\left(2\pi x^3\right) + \epsilon, \quad \epsilon \sim N(0, 0.1^2)$$



Splines (cont'd)

- Setting
  - Suppose $x \in [a, b]$, partition the interval using the following points (a.k.a. knots)

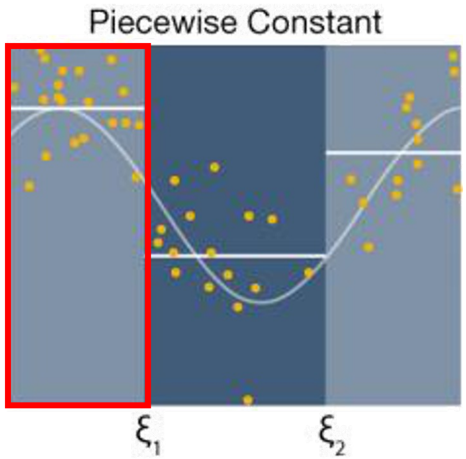
$$a < \xi_1 < \xi_2 < \cdots < \xi_K < b \quad \xi_0 = a, \xi_{K+1} = b$$

  - Fit a polynomial in each interval under the continuity conditions and integrate them by

$$f(X) = \sum_{m=1}^{K} \beta_m h_m(X)$$

  - Look for the best interpretator from a dictionary of polynomials
  - What is the potential issue here?

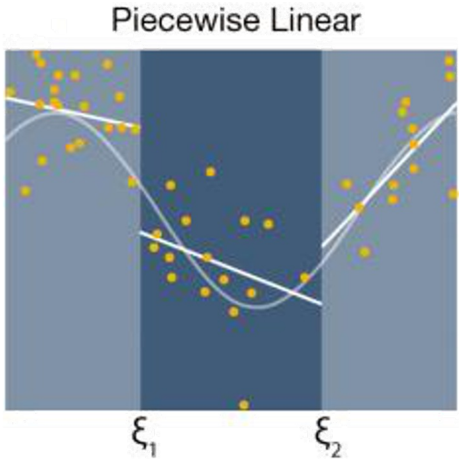
Splines: Simple Examples



Works like a filter

$$\boxed{h_1(X) = I(X < \xi_1),} \quad h_2(X) = I(\xi_1 \leq X < \xi_2), \quad h_3(X) = I(\xi_2 \leq X).$$

$$f(X) = \sum_{m=1}^{3} \beta_m h_m(X) \xrightarrow{\text{LSE}} \hat{\beta}_m = \bar{Y}_m$$

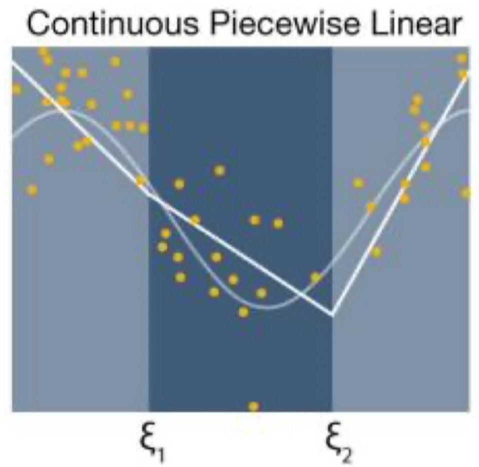


$$h_{m+3} = h_m(X)X, \;\; m = 1, \ldots, 3.$$

$$f(X) = \sum_{m=1}^{6} \beta_m h_m(X)$$

Question: We only have three knots, why there are six summons?

Splines: Simple Examples (cont'd)

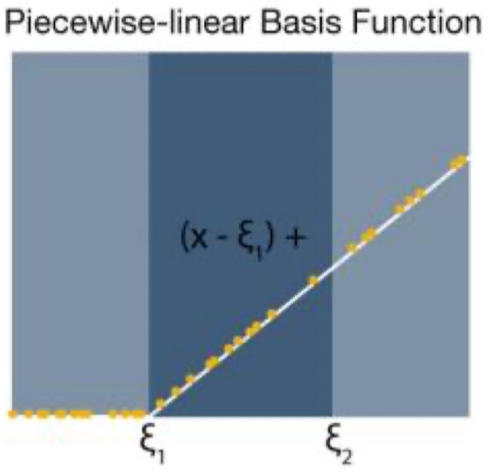


$$f(X)=\sum_{m=1}^{6}\beta_m h_m(X)$$

Impose continuity constraint for each knot:

$$f(\xi_1^-)=f(\xi_1^+) \quad \Rightarrow \quad \beta_1+\xi_1\beta_4=\beta_2+\xi_1\beta_5$$

Total number of free parameters (degrees of freedom) is 6-2=4



Alternatively, one could incorporate the constraints into the basis functions:

$$h_1(X)=1,\quad h_2(X)=X,\quad h_3(X)=(X-\xi_1)_+,\quad h_4(X)=(X-\xi_2)_+,$$

This basis is known as truncated power basis

Learning Objectives

- Interpolation