Lecture #15

02/13/2023

Artificial oscillation in polynomial interpolation

→ especially near the endpoints of the data set

$\underbrace{\hspace{10em}}$

This is called Runge Phenomenon.

↖ Overfitting due to high degree of the polynomial

Mathematically, the $\underbrace{\text{coefficient matrix } X}_{\text{Vandermonde matrix}}$

can have large condition number in practice

A theoretical lower bound on $\kappa_p(A)$:

<u>Claim:</u> Consider any nonsingular matrix

$$A = \begin{bmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

$$\kappa_p(A) \geq \frac{\max\limits_{i=1,\ldots,n} \|\underline{a}_i\|_p}{\min\limits_{i=1,\ldots,n} \|\underline{a}_i\|_p}$$

Proof: See CANVAS file Section: folder: Supplementary Notes

Application example for the above claim:

Let us apply the above result for nonsingular Vandermonde matrix

$$X = \underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}}_{n \times n}$$

with

$x_1 = 1, \ x_2 = 2, \ \ldots, \ x_n = n$

$$K_2(X) \geq \frac{\max_{i=1,\ldots,n} \| i^{th} \text{ column of } X \|_2}{\min_{i=1,\ldots,n} \| i^{th} \text{ column of } X \|_2} = \frac{?}{\sqrt{n}}$$

$$? = \sqrt{(1^{n-1})^2 + (2^{n-1})^2 + \cdots + (n^{n-1})^2}$$

$$= \sqrt{1^{2(n-1)} + 2^{2(n-1)} + \cdots + n^{2(n-1)}}$$

$$\geqslant \sqrt{n^{2(n-1)}} = n^{n-1}$$

$$\boxed{\therefore\ \kappa_2(X) \geqslant \frac{n^{n-1}}{\sqrt{n}} = n^{n-1-1/2} = n^{n-3/2}}$$

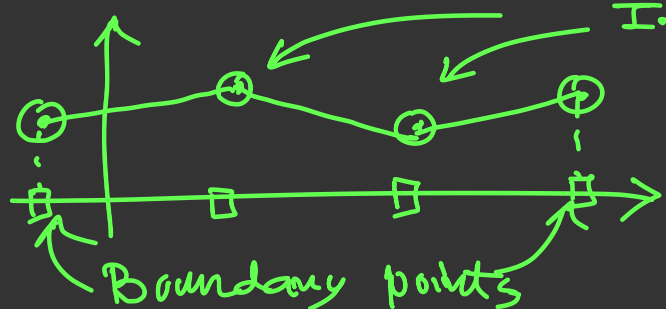
very large condition number even for moderately large n (e.g., $n = 10$).

How to fix / counter this issue for interpolation?

Idea # 1 : $\underbrace{\text{Spline interpolation}}_{\text{piecewise polynomial interpolation}}$

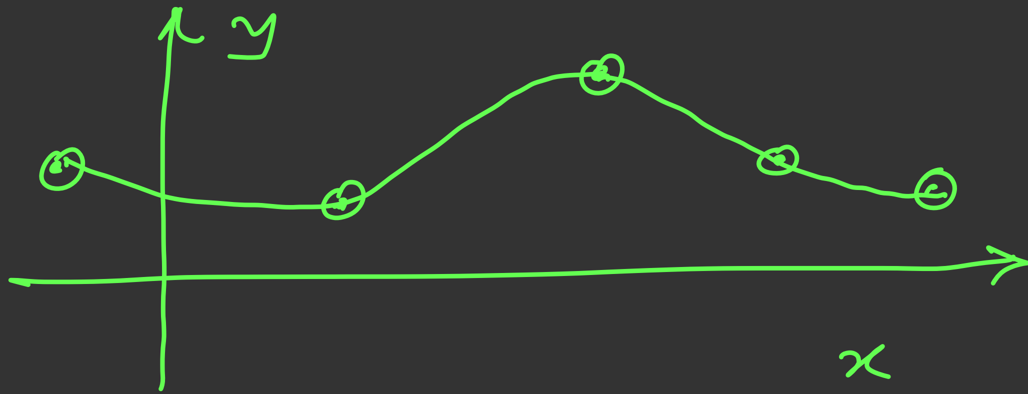
$\Leftrightarrow$ "fit different polynomials @ different parts of the given dataset"

Example: (linear spline)



We want to enforce continuity @ breakpoints

Cubic spline interpolation:



Enforce:

1. continuity / function value match
2. slope / first derivative value match
3. Curvature / second derivative value match