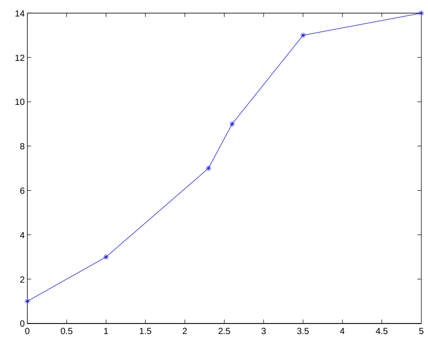
Numerical Analysis, FMN011

Carmen Arévalo
Lund University
carmen@maths.lth.se

Interpolation

Interpolation

To construct a function that has some specified values at certain points $\{x_0, x_1, \ldots, x_{n-1}\}$.



Function $y = P(x)$ interpolates the data points $\{(x_0, y_0), (x_1, y_1), \ldots, (x_{n-1}, y_{n-1})\}$ if $P(x_i) = y_i$ for each $i = 0, 1, 2, \ldots, n-1$.

Interpolating can be viewed as data compression.

Reasons for Interpolating

- construct a curve that passes through some discrete number of points: computer graphics
- evaluate a mathematical function easily and quickly: sines, cosines, log, exponentials, . . .
- substitute a “difficult” function by an “easy” one: simplifying a mathematical model for the weather report, integrating numerically
- extract information from a table of values: predict what the data would be at points where it wasn’t measured, or analyze the growth pattern of the data
- compress data: music, scanning (http://www.dtp-aus.com/interpol.htm)

How to choose the form of the interpolating function

Questions:

- are there relevant mathematical or physical considerations?
- how should the function behave between the data points?
- should the function have some properties like periodicity?
- should the graph be pleasant to the eye?
- do we need the mathematical description of the interpolating function or do we only need its graph?

Types of interpolating functions

Our choice must be based on ease of calculation and properties.

- polynomials
- piecewise polynomials
- trigonometric functions
- exponential functions
- rational functions

The family of interpolating functions

The chosen family is generated by the basis functions $\Phi_1, \Phi_2, \ldots, \Phi_n$.

$$f(x) = \sum_{j=1}^{n} y_j \Phi_j(x)$$

Each function in the family is uniquely determined by its coefficients y_j.

How to determine the coefficients: Solve

$$\begin{pmatrix} \Phi_1(x_0) & \Phi_2(x_0) & \cdots & \Phi_n(x_0) \\ \Phi_1(x_1) & \Phi_2(x_1) & \cdots & \Phi_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1(x_{n-1}) & \Phi_2(x_{n-1}) & \cdots & \Phi_n(x_{n-1}) \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{n-1}) \end{pmatrix}$$

Polynomial interpolation

There is a unique polynomial of degree $n-1$ that passes through n distinct points

How do we calculate this polynomial? Depends on the basis we choose for the space of polynomials of degree $n-1$, Π_{n-1}.

Some well-known bases:

- Monomials
- Lagrange
- Newton
- Orthogonal
- Bernstein

With the Monomial Basis

n nodes: $x_0, x_1, \ldots, x_{n-1}$

Basis functions: $1, x, x^2, x^3, \ldots, x^{n-1}$

Interpolating polynomial of degree $n-1$:

$$P(x) = y_{n-1}x^{n-1} + \cdots + y_2 x^2 + y_1 x + y_0$$

System matrix: a Vandermonde matrix,

$$\boxed{\Phi = \begin{pmatrix} x_0^{n-1} & \cdots & x_0 & 1 \\ x_1^{n-1} & \cdots & x_1 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{n-1}^{n-1} & \cdots & x_{n-1} & 1 \end{pmatrix}}$$

Example with Monomial Basis

Data points: $(-2,-27),(0,-1),(1,0)$

$$\begin{pmatrix} 4 & -2 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_2 \\ y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} -27 \\ -1 \\ 0 \end{pmatrix}$$

Solution: $y = (-4, 5, -1)^T$.

Unique polynomial of degree 2 passing through the three data points:

$$P(x) = -4x^2 + 5x - 1$$

The Lagrange Basis

Lagrange polynomials:

$$\boxed{L_j(x) = \frac{\prod_{k=0,k\neq j}^{n-1}(x - x_k)}{\prod_{k=0,k\neq j}^{n-1}(x_j - x_k)}}$$

Note that

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

Interpolating polynomial of degree n:

$$P(x) = y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x) + \cdots + y_{n-1} L_{n-1}(x)$$

System matrix: $\Phi = I$, an identity matrix!

Example with Lagrange Basis

Data points: $(-2, -27), (0, -1), (1, 0)$

Lagrange polynomials:

$$\frac{(x-0)(x-1)}{(-2-0)(-2-1)}, \frac{(x+2)(x-1)}{(0+2)(0-1)}, \frac{(x+2)(x-0)}{(1+2)(1-0)}$$

Unique polynomial of degree 2 passing through the three data points:

$$P(x) = -27\frac{x(x-1)}{6} - 1\frac{(x+2)(x-1)}{-2}$$

Extrapolation

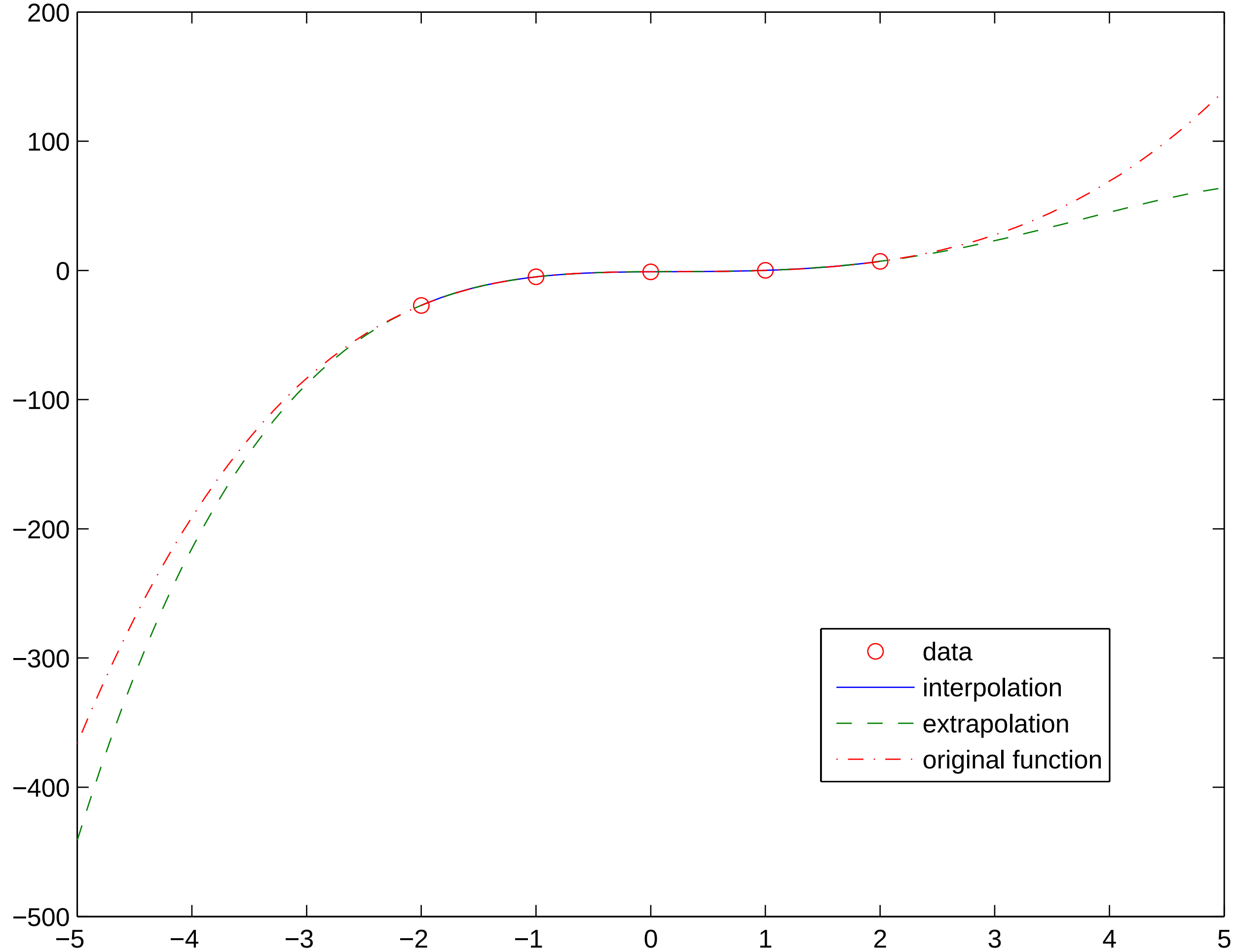
Suppose $P \in \Pi_N$ passes through nodes satisfying

$$x_1 < x_2 < \cdots < x_{N-1} < x_N$$

Consider the curve $y = P(x)$. When

- $x_1 < x < x_N$, $P(x)$ is called an interpolated value

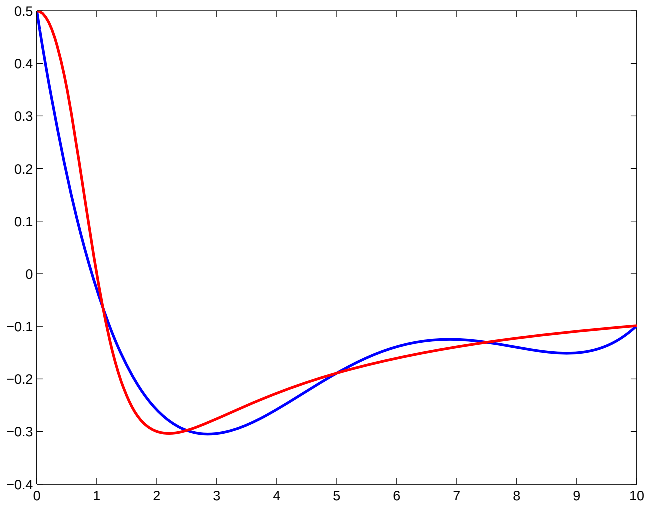
- $x < x_1$ or $x_N < x$, $P(x)$ is called an extrapolated value



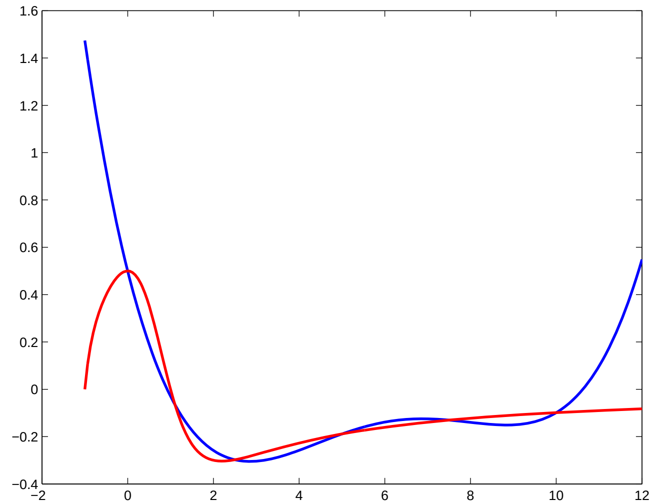
200
100
0
−100
−200
−300
−400
−500
−5
−4
−3
−2
−1
0
1
2
3
4
5
data
interpolation
extrapolation
original function

Exercise (interpolation vs extrapolation)

$$P(x) = 0.5000 - 0.3191t + 0.0725x(x-2.5)- \\ 0.0102x(x-2.5)(x-5) + 0.0010x(x-2.5)(x-5)(x-7.5)$$



Function (red) and interpolating polynomial (blue) on $[0, 10]$



The red curve is the original function,
the blue curve is the interpolating polynomial on $[-1, 12]$.

Accuracy of Interpolation

When data points $(x_1, f(x_1)), \ldots, (x_n, f(x_n))$ are a sample of a continuous function f on the interval $[x_1, x_n]$, the interpolating error for the interpolating polynomial of degree $n-1$ is

$$\boxed{f(x) - P(x) = \frac{f^{(n)}(\theta)}{n!}(x - x_1)(x - x_2)\ldots(x - x_n)}$$

where $\theta \in [x_1, x_n]$ is unknown.

This error cannot, in general be evaluated, but it provides an error bound.

Exercise

Bound the errors for the approximation to $f(x) = e^x$ at $x = 0.1$ for the following interpolation polynomials:

Nodes: 0 and 0.5 $\Rightarrow P_1(x) = 1 + 2(e^{0.5} - 1)x$
Nodes: 0, 0.5 and 1 $\Rightarrow P_2(x) = 1 + 2(e^{0.5} - 1)x + 2(e - 2e^{0.5} + 1)x(x - 0.5)$
Nodes: 0 and 0.25 $\Rightarrow P_3(x) = 1 + 4(e^{0.25} - 1)x$

$|f^{(n)}(x)| = e^x \Rightarrow \max|f^{(n)}(x)| = e^{\max x}$ for $x \in [0, \infty]$

$$\begin{aligned}
|E_1(x)| &\leq \frac{e^{0.5}}{2}|x(x-0.5)| \approx 0.0330 \; (= 0.0246) \\
|E_2(x)| &\leq \leq \frac{e}{6}|x(x-0.5)(x-1)| \approx 0.0163 \; (= 0.0091) \\
|E_3(x)| &\leq \frac{e^{0.25}}{2}|x(x-0.25)| \approx 0.0096 \; (= 0.0084)
\end{aligned}$$

Placement of Interpolation Points

Does the error decrease as the degree of the polynomial increases?

- **YES:** if the points are well-chosen
- **NO:** in some cases, the error gets larger as $n \to \infty$

Recall the error formula:

$$f(x) - P(x) = \frac{f^{(n)}(\theta)}{n!}(x - x_1)(x - x_2)\ldots(x - x_n)$$

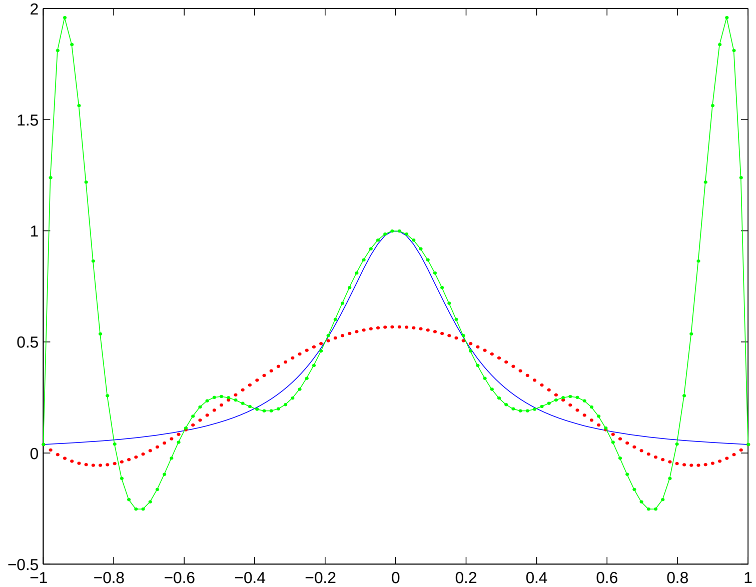
where $\theta \in [x_1, x_n]$ is unknown.

The error is reduced by choosing $\{x_1, x_2, \ldots, x_n\}$
to minimize $|(x - x_1)(x - x_2)\ldots(x - x_n)|$.

The Runge function

$$f(x) = \frac{1}{1+25x^2}, \qquad x \in [-1, 1]$$

Interpolation with 6 (red) and 11 (green) equally spaced points



Chebyshev Nodes

The Chebyshev polynomials in $x \in [-1, 1]$ are

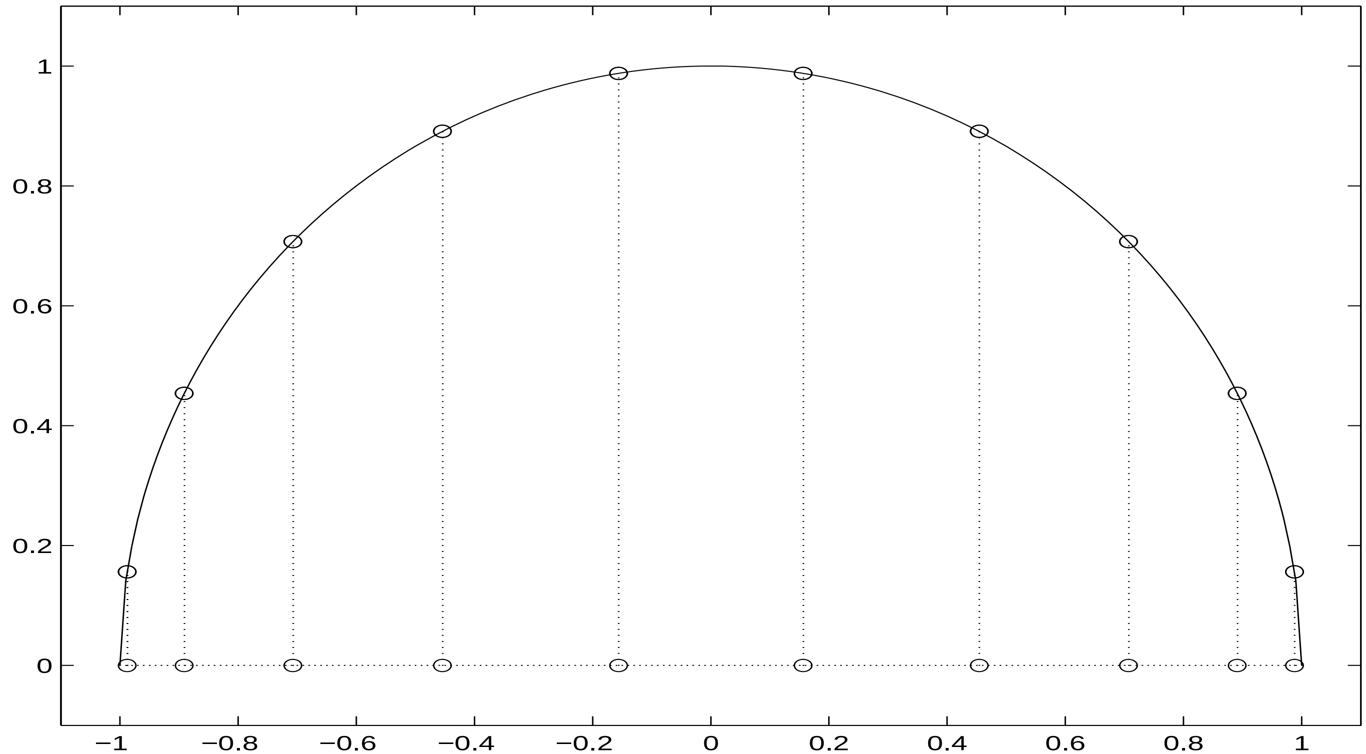
$$\boxed{\begin{array}{l} T_n(x) = \cos(n \cdot \arccos x), \text{ or recursively,} \\ T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \end{array}}$$

The zeros of T_n are

$$\boxed{x_k = \cos\frac{(2k-1)\pi}{2n}, \quad k = 1 : n.}$$

These nodes minimize $|(x - x_1)(x - x_2)\dots(x - x_n)|$ and the error is distributed evenly in $[-1, 1]$.

Chebyshev nodes



$$z_j = e^{i\theta},\ \theta \in [0, \pi],\ x_j = \mathrm{Re}(z_j)$$

The points are evenly distributed along the semicircle.

Optimal error

$$E(x) = \frac{f^{(n)}(\theta(x))}{n!}(x - x_1)(x - x_2)\cdots(x - x_n)$$

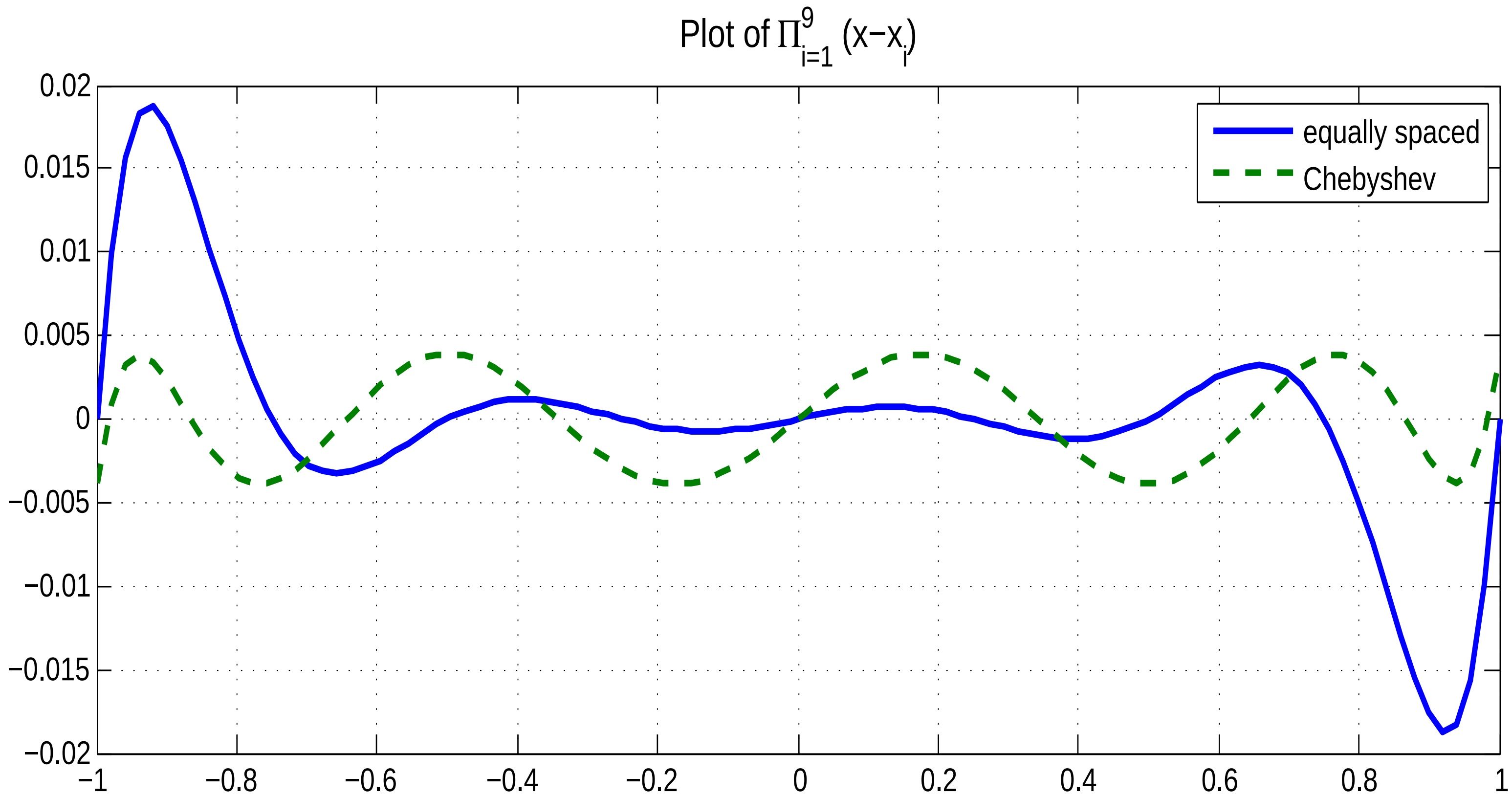
If $x_1, x_2, \ldots, x_n$ are the Chebyshev nodes, then $\max_{-1 \le x \le 1} |(x - x_1)(x - x_2)\cdots(x - x_n)|$ attains its minimum possible value, and

$$|(x - x_1)(x - x_2)\cdots(x - x_n)| \le \frac{1}{2^{n-1}}$$

that is,

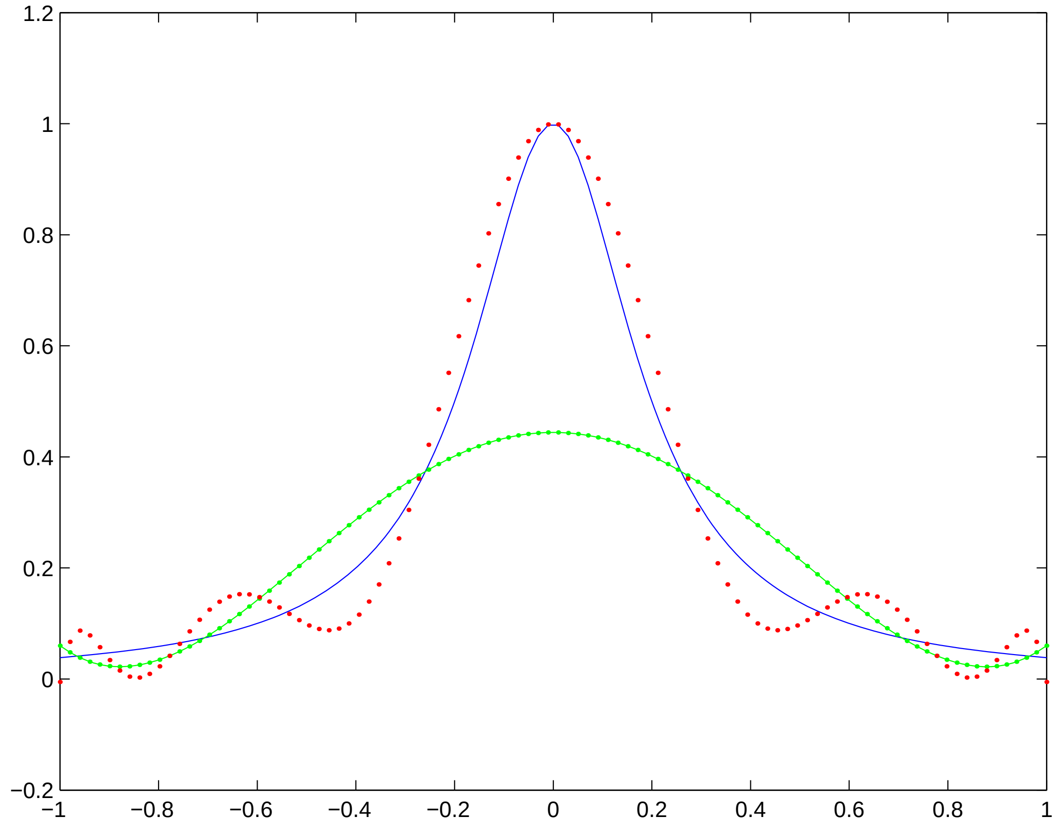
$$\boxed{|E(x)| \le \frac{\max_{-1 \le \theta \le 1}\{|f^{(n)}(\theta)|\}}{2^{n-1} \cdot n!}}$$

Comparison of error size and distribution



Runge function revisited

Interpolation with 6 (green) and 11 (red) Chebyshev points



Example

Approximate $f(x) = \sin x,\ x \in [-1, 1]$ by a polynomial of degree 5.

Equally spaced nodes: $x_i = -1 + 2(i-1)/5 \quad i = 1, \ldots, 6$
$\{\pm 1, \pm 0.6, \pm 0.2\}$

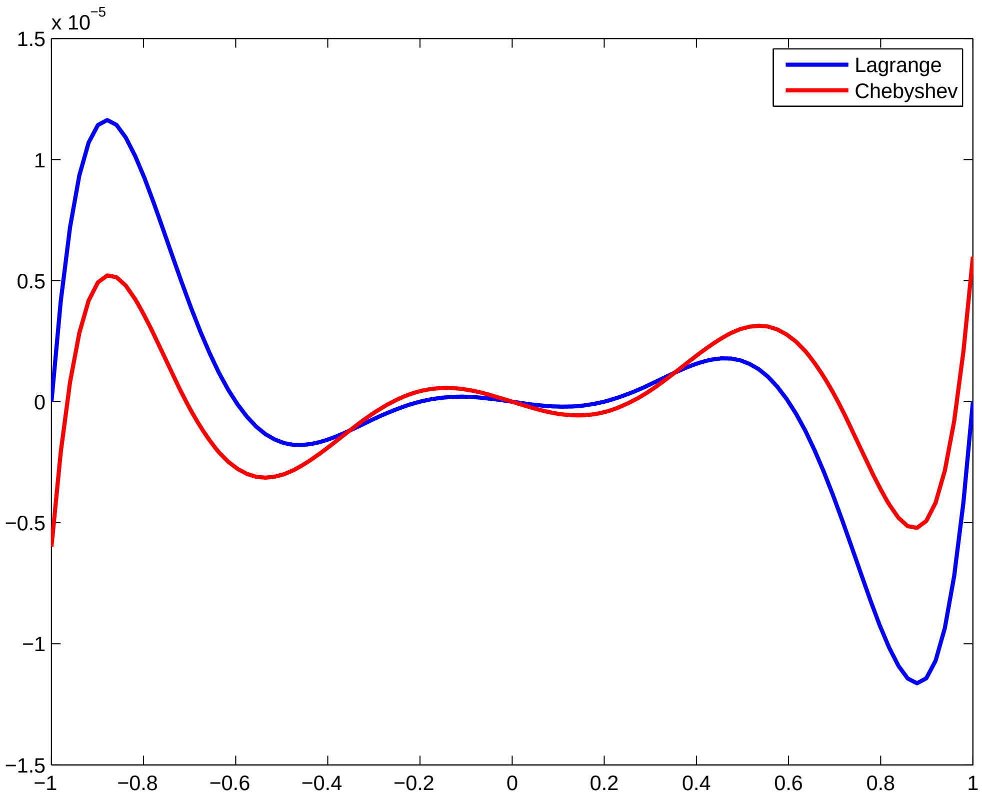
$\max|Pl(x) - \sin(x)| = 0.000012$

Chebyshev nodes: $x_i = \cos\frac{(2i-1)\pi}{12} \quad i = 1, \ldots, 6$

$\{\pm 0.9659258, \pm 0.7071067, \pm 0.258819\}$

$\max|Pc(x) - \sin(x)| = 0.000006$

Absolute error for equally spaced (blue) and Chebyshev (red) points

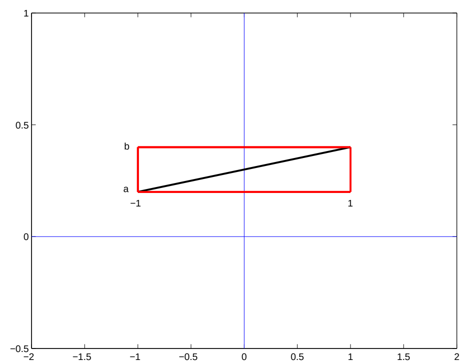


Interval transformation

To interpolate on [a,b], calculate the Chebyshev nodes on $[-1, 1]$ and use the transformation

$$\boxed{x = \frac{b-a}{2}t + \frac{a+b}{2}, \quad t \in [-1, 1],}$$

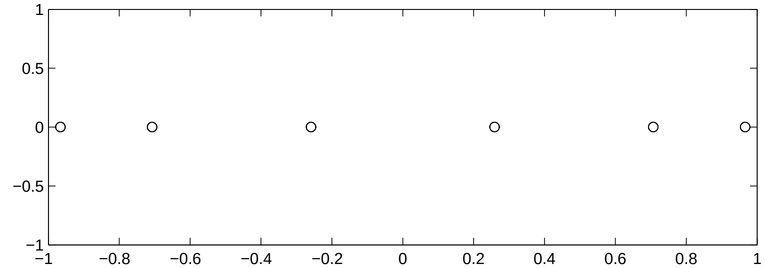
to get the nodes on $[a, b]$.

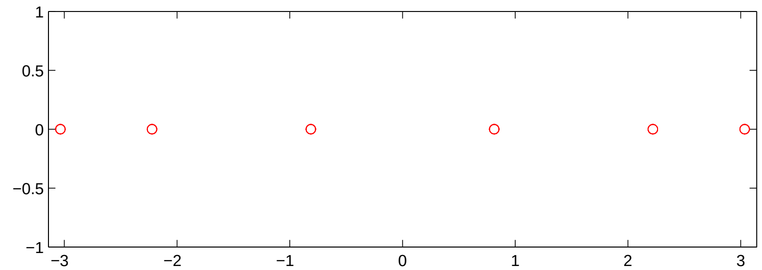


Example

$x \in [-\pi, \pi]$; polynomial of degree 5. $\{t_i = \pm 0.9659258, \pm 0.7071067, \pm 0.258819\}$
$\{x_i = \pi t_i\}$

Chebyshev nodes in $[-1, 1]$ and in $[-\pi, \pi]$





Example

$x \in [2\pi, 4\pi]$; polynomial of degree 5.

Chebyshev nodes on $[-1, 1]$: $\{t_i = \pm 0.9659258, \pm 0.7071067, \pm 0.258819\}$

Chebyshev nodes on $[2\pi, 4\pi]$: $x_i = \pi t_i + 3\pi =$
$\{2.0341\pi, 2.2929\pi, 2.7412\pi, 3.2588\pi, 3.7071\pi, 3.9659\pi\}$

Equally space nodes on $[2\pi, 4\pi]$: $\{2\pi, 12\pi/5, 14\pi/5, 16\pi/5, 18\pi/5, 4\pi\}$