Fitting functions to data

1. Approximating a known function

- evaluating a function using basic arithmetic only
- computing the original function may be inefficient

The function is known exactly at some points. Other values are found by interpolation.

- the function and its derivatives known at one point
    - Taylor series
    - Padé approximation
- equally spaced points
    - interpolation polynomials
- arbitrary set of points
    - Lagrangian interpolation
    - spline functions
    - bezier curves

2. Least squares fit

Random errors in data, cannot be described exactly.

- linear fir (linear combination of arbitrary functions)
- nonlinear fit

Criteria of the fit

We try to minimize the "distance" $d(f,g)$ of the functions or the norm $\parallel f-g \parallel$. The norm can be defined in many ways

L_1 norm

$$\parallel f-g \parallel_1 = \int |f-g|\,dx$$

This allows large deviations if they occur in a narrow range.

L_2 norm corresponds to the distance in an Euclidean space:

$$\parallel f-g \parallel_2 = \sqrt{\int |f-g|^2\,dx}.$$

This is a special case of the more general L_p norm

$$\| f - g \|_p = \left(\int |f - g|^p \, dx \right)^{1/p}.$$

L_∞ norm or maximum norm

$$\| f - g \|_\infty = \sup |f - g|.$$

This prevents large deviations but the fit may not be very good anywhere.

Taylor series

Let f be a function $f : \mathbf{R} \to \mathbf{R}$. The equation of the tangent at x_0 is

$$y = f(x_0) + f'(x_0)(x - x_0),$$

where $f'(x_0)$ is the derivative df/dx at x_0. The function in the neighborhood of x_0 can be approximated by the tangent:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

The estimate is the worse the more the derivative f' varies in the interval $[x_0, x]$. This variation is described by the second derivative f'' etc.

At x the function has the value

$$\begin{aligned} f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + \ldots \\ + \frac{1}{n!} f^{(n)}(x_0)(x - x_0)^n + \ldots \end{aligned}$$

where $f^{(n)}(x_0)$ is the n:th derivative of f at x_0. This is the Taylor series of the function f at x_0.

Examples (in all these $x_0 = 0$):

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots \qquad \text{converges, when} |x| < 1$$

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \cdots$$

$$\sqrt{1-x} = 1 - \frac{1}{2}x - \frac{1}{8}x^2 - \frac{1}{16}x^3 - \cdots$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 - \cdots$$

$$\frac{1}{\sqrt{1-x}} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \cdots$$

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots \qquad \text{for all } x$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots \qquad x \in (-1, 1]$$

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots \qquad \text{for all } x$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots \qquad \text{for all } x$$

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots \qquad |x| < \frac{\pi}{2}$$

Very often we'll need linear approximations like

$$\sqrt{1+x} \approx 1 + \frac{1}{2}x,$$
$$\frac{1}{\sqrt{1+x}} \approx 1 - \frac{1}{2}x,$$

These can be used to linearize functions in a very small neighborhood of some point.

Rational approximations

Use a rational expression to approximate the function

$$\frac{a_0 + a_1 x + \ldots + a_n x^4}{1 + b_1 x + \ldots b_m x^m}.$$

Finding the coefficients will usually mean an optimization problem.

A simple and often used method is the Padé approximation.

Example: the Taylor saries of the exponential function at $x = 0$ is

$$f(x) = e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \ldots$$

Try to find a rational approximation of the form

$$R(x) = \frac{a + bx + cx^2}{1 + dx}.$$

We now require that at $x = 0$ this will give the same value as the Taylor series and the derivatives are equal

Consider the difference

$$f(x) - R(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 - \frac{a + bx + cx^2}{1 + dx}$$
$$= \frac{(1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3)(1 + dx) - (a + bx + cx^2)}{1 + dx}.$$

It is required that at the origin this difference anb its derivatives will vanish. The nominator must be zero for all x.

Expand the nominator and regroup the terms according to powers of x:

$$(1-a)+(1+d-b)x+\left(\frac{1}{2}+d-c\right)x^2+\left(\frac{1}{6}+\frac{1}{2}d\right)x^3+\frac{1}{6}x^4.$$

This will vanish for all x if all coefficients of the powers of x are zero. It is not possible to make the last term vanish, but we can drop it because we are searching for a third degree approximation.

We get a set of equations

$$\begin{aligned}
1-a&=0,\\
1+d-b&=0,\\
\frac{1}{2}+d-c&=0,\\
\frac{1}{6}+\frac{1}{2}d&=0,
\end{aligned}$$

whence

$$\begin{aligned} a &= 1, \\ b &= \frac{2}{3}, \\ c &= \frac{1}{6}, \\ d &= -\frac{1}{3}. \end{aligned}$$

Thus the approximation is

$$f(x) = \frac{1 + \frac{2}{3}x + \frac{1}{6}x^2}{1 - \frac{1}{3}x}.$$

E.g. at $x = 1$ a Taylor series of four terms would give $2\frac{2}{3} \approx 2.667$. The rational approximation will give $f(1) = \frac{11}{4} = 2.750$.

Same problem as with the Taylor series: exactly correct value at one point but cannot minimize error elsewhwere.

Chebychev polynomials

Recurrence relation

$$\cos(n+1)\phi + \cos(n-1)\phi = 2\cos\phi\ \cos n\phi.$$

Proof:

$$\begin{aligned}
&\cos(n+1)\phi + \cos(n-1)\phi \\
&= \cos(n\phi+\phi) + \cos(n\phi-\phi) \\
&= \cos n\phi\cos\phi - \sin n\phi\sin\phi + \cos n\phi\cos\phi + \sin n\phi\sin\phi \\
&= 2\cos n\phi\cos\phi.
\end{aligned}$$

The recurrence relation gives

$$\begin{aligned}\cos 2\phi &= 2\cos^2\phi - 1, \\ \cos 3\phi &= 2\cos\phi\,\cos 2\phi - \cos\phi \\ &= 4\cos^3\phi - 3\cos\phi \\ &\vdots\end{aligned}$$

Cosines of the multiples of the angle ϕ can be expressed in a form containing powers of $\cos\phi$ only. We get expressions that are polynomials of $\cos\phi$.

We'll denote

$$x = \cos\phi,$$

Define the Chebychev polynomial T_n as

$$T_n(x) = \cos n\phi = \cos(n \arccos x).$$

$$\begin{aligned}
T_0(x) &= \cos 0 = 1, \\
T_1(x) &= \cos \arccos x = x. \\
T_2(x) &= \cos(2 \arccos x) = \cos 2\phi \\
&= 2\cos^2\phi - 1 \\
&= 2x^2 - 1, \\
T_3(x) &= 4x^3 - 3x, \\
T_4(x) &= 8x^4 - 8x^2 + 1 \\
T_5(x) &= 16x^5 - 20x^3 + 5x.
\end{aligned}$$

Assume that x is always in the range $-1 \leq x \leq +1$.

Properties

Symmetry:

$$T_n(-x) = (-1)^n T_n(x).$$

Zeros of the polynomials

Since the zeros of the unction $\cos\phi$ are $\phi = (2k+1)\pi/2$, $k = 0, 1, \ldots$, the zeros of the function $\cos n\phi$ are

$$\phi = \frac{2k+1}{n}\frac{\pi}{2}, \quad k = 0, 1, \ldots$$

Since $T_n(x) = \cos n\phi$, these are also zeros of the polynomial T_n. In terms of x the zeros are

$$x_k = \cos\left(\frac{2k+1}{n}\frac{\pi}{2}\right), \quad k = 0, 1, \ldots, n-1.$$

Orthogonality

Chebychev polynomials form a set of orthogonal functions, if we use a weight function $1/\sqrt{1-x^2}$:

$$\int_{-1}^{+1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}\,dx = \begin{cases} 0, & n \neq m, \\ \pi/2, & n = m \neq 0, \\ \pi, & n = m = 0. \end{cases}$$

For a finite set of points the orthogonality is valid in the form

$$\sum_{i=0}^{K} T_n(x_i)T_m(x_i) = \begin{cases} 0, & n \neq m, \\ (K+1)/2, & n = m \neq 0, \\ K+1, & n = m = 0, \end{cases}$$

where the points x_i are the zeros of the polynomial T_{K+1}, $n \leq K$, $m \leq K$.

Minimality

By definition, a Chebychev polynomial is a cosine of some angle. Thus its absolute value cannot exceed one. Its extrema are either 1 or -1. Thus its maximum norm is exactly 1.

Using the recurrence formula it can be shown that the coefficient of the highest power of the polynomial is 2^{n-1}. If the polynomial is multiplied by 2^{1-n}, we'll get a polynomial with one as the coefficient of the highest power.

Among all polynomials having one as the coefficient of the highest power, the polynomial $2^{1-n}T_n$ has the smallest maximum norm.

The polynomial $2^{1-n}T_n$ has $n+1$ extrema in the range $-1 \leq x \leq 1$. Of these, $n-1$ are points where the derivative will vanish. In addition to these, the polynomial has extrema also at the endpoints of the interval. Let the locations of the extrema be $x_0 = -1$, x_1, ..., $x_n = 1$. These are minima and maxima alternatively. Assume that x_0 is a maximum.

Assume that there is another polynomial P_n of degree n with a smaller maximum norm. Since x_0 is a maximum of the polynomial $2^{1-n}T_n$, we must have

$$P(x_0) < 2^{1-n}T_n(x_0).$$

Correspondingly x_1 is a minimum, where

$$P(x_1) > 2^{1-n}T_n(x_1).$$

Thus at the extrema P_n is alternatingly smaller or greater than $2^{1-n}T_n$.

$$\begin{aligned}
P_n(x_0) - 2^{1-n}T_n(x_0) &< 0, \\
P_n(x_1) - 2^{1-n}T_n(x_1) &> 0, \\
P_n(x_2) - 2^{1-n}T_n(x_2) &< 0, \\
&\vdots
\end{aligned}$$

The function $P_n(x) - 2^{1-n}T_n(x)$ will change sign n times. Since in both terms the coefficient of the highest power is one, these highest powers cancel out, and the function is a polynomial of degree $n-1$. Such a polynomial can have only $n-1$ zeros, and it cannot change sign n times. Contradiction $\Rightarrow$ our assumption is false.

How to place n points x_k, $k = 1, \ldots, n$ in the range $-1 \le x \le 1$ in such a way that the maximum norm of the polynomial $(x-x_1)(x-x_2)\cdots(x-x_n)$ will be as small as possible? Answer: the points must be zeros of the Chebychev polynomial T_n.

Interpolation

Equally spaced data

The function is known at the points x_i, $f(x_i) = y_i$.

i	x	$f(x)$	Δf	$\Delta^2 f$	$\Delta^3 f$
-1	x_{-1}	y_{i-1}			
			Δf_{-1}		
0	x_0	y_0		$\Delta^2 f_{-1}$	
			Δf_0		$\Delta^3 f_{-1}$
1	x_1	y_1		$\Delta^2 f_0$	
			Δf_1		
2	x_2	y_2			

Denote $h = x_1 - x_0$.

Step operator $Ef(x) = f(x+h)$.

Forward difference $\Delta f(x) = f(x+h) - f(x)$.

Backward difference $\nabla f(x) = f(x) - f(x-h)$.

All are linear operaators

The operators have e.g. the following obvious properties:

$$E^2 f(x) = E(Ef(x)) = E(f(x+h)) = f(x+2h).$$

$$E^n f(x) = f(x+nh).$$

$$\Delta f(x) = f(x+h) - f(x) = Ef(x) - f(x) = (E-1)f(x),$$

whence

$$\Delta = E - 1.$$

Newton-Gregory interpolation polynomial:

$$\begin{aligned} P(x_0 + sh) &= E^s f(x_0) = (1 + \Delta)^s f(x_0) \\ &= \left[1 + s\Delta + \binom{s}{2}\Delta^2 + \dots\right] f(x_0) \\ &= f_0 + s\Delta f_0 + \binom{s}{2}\Delta^2 f_0 + \dots \end{aligned}$$

$$P_2(x_0 + sh) = f_0 + s\Delta f_0 + \frac{s(s-1)}{2}\Delta^2 f_0.$$

When $s = 0, 1, 2$, this will go through the points (x_0, y_0), $(x_0 + h, y_1)$ and $(x_0 + 2h, y_2)$:

$$\begin{aligned} P_2(x_0) &= P_2(x_0 + 0h) = y_0, \\ P_2(x_1) &= P_2(x_0 + 1h) = y_0 + \Delta f_0 = y_1, \\ P_2(x_2) &= P_2(x_0 + 2h) = y_0 + 2\Delta f_0 + \Delta^2 f_0 = y_2 \end{aligned}$$

P_2 can be understood also as a function of s, where $s = (x - x_0)/h$ is an arbitrary real number.

In the same manner interpolation polynomials of higher degree can be derived.

Instead of forward differences it is possible to use backward differences or both together.

Methods using differences can be applied to equally spaced data only.

Example:

x	y
10	2.0
30	3.0
50	3.8
75	4.8
100	5.2

Since there are 5 points in the data, a fourth degree polynomial can describe it exactly:

$$y = P_4(n) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4.$$

Substituting the given values we get a set of equations

$$\begin{aligned}
a_0 + \; 10a_1 + \; 100a_2 + \; 1\,000a_3 + \; 10\,000a_4 &= 2.0,\\
a_0 + \; 30a_1 + \; 900a_2 + \; 27\,000a_3 + \; 810\,000a_4 &= 3.0,\\
a_0 + \; 50a_1 + \; 2\,500a_2 + \; 125\,000a_3 + \; 6250\,000a_4 &= 3.8,\\
a_0 + \; 75a_1 + \; 5\,625a_2 + \; 421\,875a_3 + \; 31640\,625a_4 &= 4.8,\\
a_0 + 100a_1 + 10\,000a_2 + 1000\,000a_3 + 100000\,000a_4 &= 5.2.
\end{aligned}$$

This is a set of linear equations, the solution of which is

$$\begin{aligned}
a_0 &= 1.234,\\
a_1 &= 0.09115,\\
a_2 &= -0.001672,\\
a_3 &= 0.00002347,\\
a_4 &= -0.0000001189.
\end{aligned}$$

Laborious! This is not the way to do it.

Lagrangian interpolation

First, find a set of cardinal functions; they are polynomials, whose values at the given points are only zeros or ones.

If there are n points, we'll need n cardinal functions, L_i, $i = 1, \ldots, n$. They are chosen so that $L_i(x_i) = 1$, $L_i(x_j) = 0$, $i \neq j$.

In the previous example $n = 4$, and the cardinal functions are polynomials of degree four. Acoording to the fundamental theorem of algebra they can be expressed as products of four factors:

$$\begin{aligned} L_1(x) &= A_1(x - x_2)(x - x_3)(x - x_4)(x - x_5), \\ &\vdots \\ L_5(x) &= A_5(x - x_1)(x - x_2)(x - x_3)(x - x_4). \end{aligned}$$

The constants A_i are given by the condition $L_i(x_i) = 1$; for example

$$A(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_1 - x_5) = 1.$$

Thus the first cardinal function is

$$L_1(x) = \frac{x - x_2}{x_1 - x_2} \frac{x - x_3}{x_1 - x_3} \frac{x - x_4}{x_1 - x_4} \frac{x - x_5}{x_1 - x_5}.$$

The cardinal functions of the example are

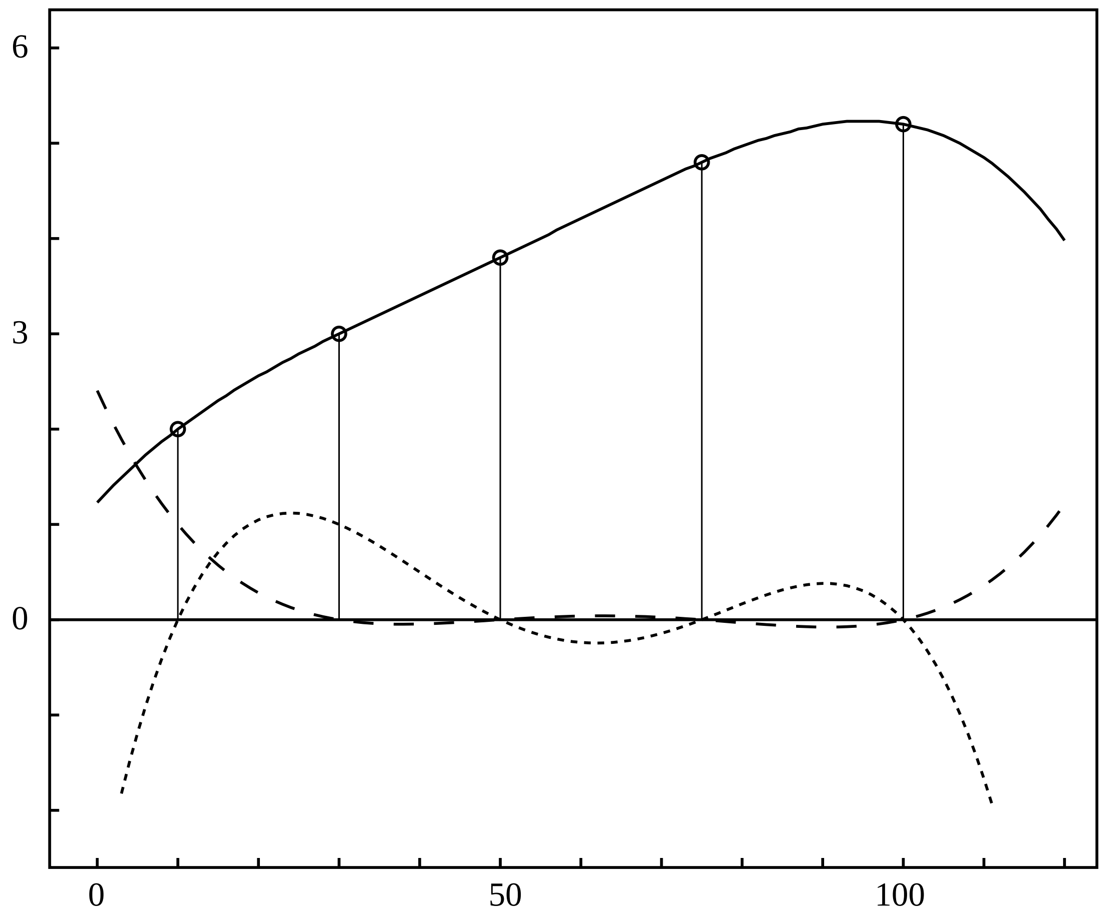
$$\begin{aligned}
L_1(x) &= \frac{1}{4680000}(x - 30)(x - 50)(x - 75)(x - 100),\\
L_2(x) &= \frac{1}{-800000}(x - 10)(x - 50)(x - 75)(x - 100),\\
L_3(x) &= \frac{1}{700000}(x - 10)(x - 30)(x - 75)(x - 100),\\
L_4(x) &= \frac{1}{-1421875}(x - 10)(x - 30)(x - 50)(x - 100),\\
L_5(x) &= \frac{1}{7875000}(x - 10)(x - 30)(x - 50)(x - 75).
\end{aligned}$$

The interpolation polynomial can be expressed as a linear combination of the cardinal functions:

$$P(x) = y_1 L_1(x) + y_2 L_2(x) + \cdots + y_n L_n(x).$$

In the example we have

$$P(x) = 2L_1(x) + 3L_2(x) + 3.8L_3(x) + 4.8L_4(x) + 5.2L_5(x).$$



Polynomials are good for interpolation, not for extrapolation.

Spline functions

Example:

x	y
0.0	0
1.2	6
2.0	11
3.5	9
4.1	17
5.0	24

The data cannot be described well using one function only.

Use a piecewise fit: the data is divided into appropriate parts, and a different function is used for each part.

Spline functions are third degree polynomial, whose coefficients are chosen so that at the boundaries of the parts second derivatives are continuous.

In the example there are 6 points and 5 intervals. We use five third degree polynomials to represent the data.

$$S_i(x) = a_i + b_i\left(\frac{x - x_i}{h_i}\right) + c_i\left(\frac{x - x_i}{h_i}\right)^2 + d_i\left(\frac{x - x_i}{h_i}\right)^3, \quad i = 1, \ldots, 5,$$

where

$$h_i = x_{i+1} - x_i.$$

The first and second derivatives are:

$$S_i'(x) = \frac{1}{h_i}\left(b_i + 2c_i\left(\frac{x - x_i}{h_i}\right) + 3d_i\left(\frac{x - x_i}{h_i}\right)^2\right),$$

$$S_i''(x) = \frac{1}{(h_i)^2}\left(2c_i + 6d_i\left(\frac{x - x_i}{h_i}\right)\right).$$

The polynomial representing the first interval must pass through both endpoints:

$$
\begin{aligned}
S_1(0) &= a_1 = 0, \\
S_1(1.2) &= a_1 + b_1 + c_1 + d_1 = 6.
\end{aligned}
$$

At the final point of the first interval the first and second derivative of the polynomial S_1 must be equal to the derivatives of the polynomial S_2 at the beginning of the second interval:

$$
\begin{aligned}
S_1'(1.2) &= S_2'(1.2), \\
S_1''(1.2) &= S_2''(1.2),
\end{aligned}
$$

or

$$
\begin{aligned}
\frac{1}{1.2}(b_1 + 2c_1 + 3d_1) &= \frac{1}{0.8}b_2, \\
\frac{1}{1.2^2}(2c_1 + 6d_1] &= \frac{1}{0.8^2}2c_2.
\end{aligned}
$$

In the example there are 5 intervals, and for each interval we get four equations. Thus, in principle, there are 20 equations. However, the equations for the derivatives cannot be formed for the last interval. Thus there are 18 equations but 20 constants to ve determined.

There are many ways to select the two missing constants.

Natural spline: second derivatives zeros at the first and last points.

$$\begin{aligned} S_1''(0) &= 2c_1 = 0, \\ S_5''(5) &= 2c_5 + 6d_5 = 0. \end{aligned}$$

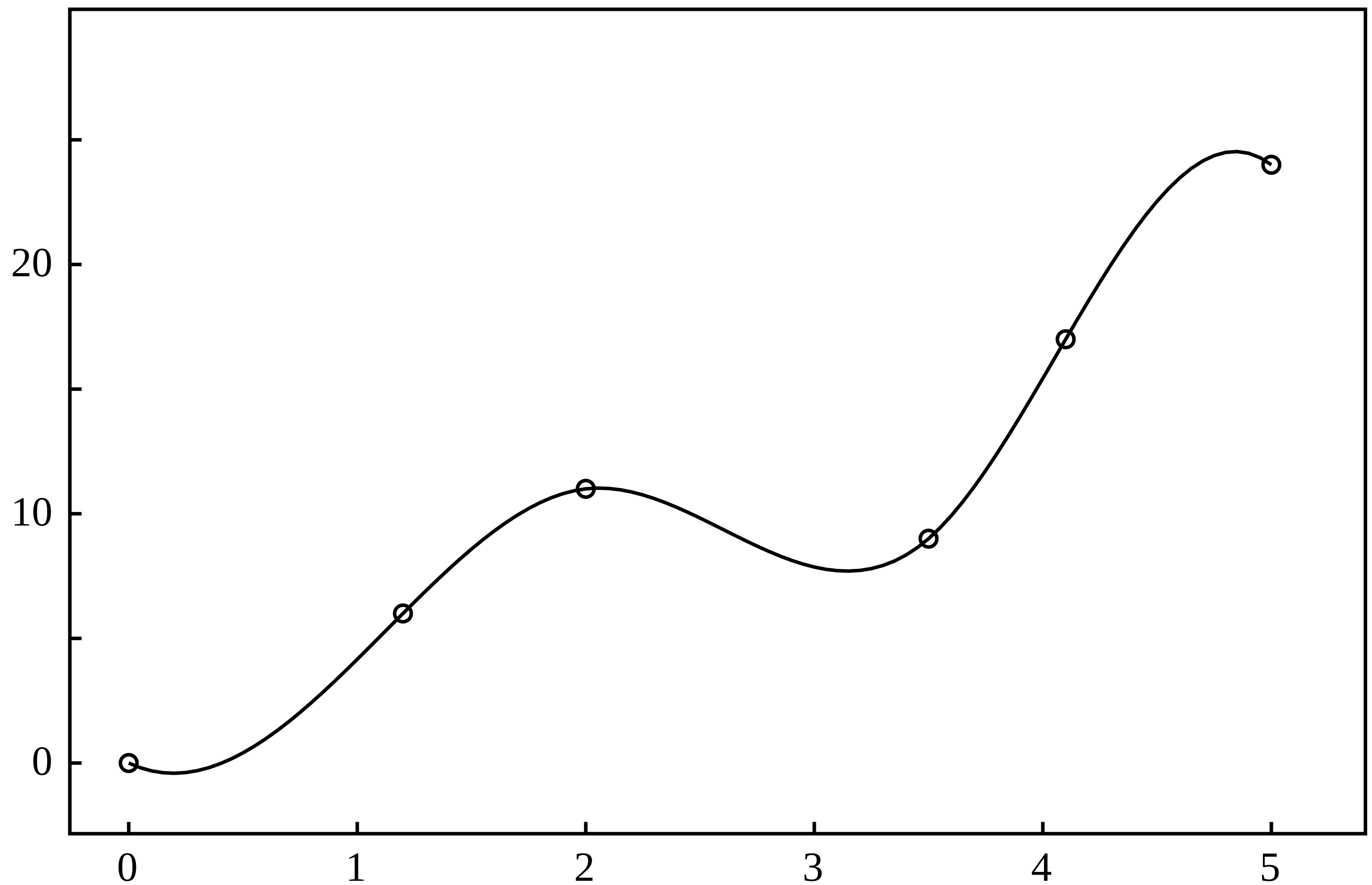
We get a set of linear equations:

$$\begin{array}{ccccccccccl}
a_1 & & & & & & & & & & = 0, \\
a_1 & + & b_1 & + & c_1 & + & d_1 & & & & = 6, \\
b_1 & + & 2c_1 & + & 3d_1 & & & - & 1.500b_2 & & = 0, \\
 & & 2c_1 & + & 6d_1 & & & - & 4.5002c_2 & & = 0, \\
 & & 2c_1 & & & & & & & & = 0, \\
a_2 & & & & & & & & & & = 6, \\
a_2 & + & b_2 & + & c_2 & + & d_2 & & & & = 11, \\
b_2 & + & 2c_2 & + & 3d_2 & & & - & 0.533b_3 & & = 0, \\
 & & 2c_2 & + & 6d_2 & & & - & 0.5692c_3 & & = 0, \\
 & & & & & & & & & & \vdots \\
a_5 & & & & & & & & & & = 17, \\
a_5 & + & b_5 & + & c_5 & + & d_5 & & & & = 24, \\
 & & 2c_5 & + & 6d_5 & & & & & & = 0,
\end{array}$$

The coefficients a_i are obtained directly and can be substituted to other equations.

The solution is

i	a_i	b_i	c_i	d_i
1	0.0	4.54	0.00	1.46
2	6.0	5.94	1.94	-2.89
3	11.0	2.19	-23.68	19.50
4	9.0	5.32	5.57	-2.89
5	17.0	11.67	-7.00	2.33



The set of equations obtained is not in the most convenient form.

$$
\begin{aligned}
y &= a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3, \\
y_i &= a_i, \\
y_{i+1} &= a_i + b_i h_i + c_i h_i^2 + d_i h_i^3, \\
y' &= b_i + 2c_i(x - x_i) + 3d_i(x - x_i)^2, \\
y_i' &= b_i, \\
y_{i+1}' &= b_i + 2c_i h_i + 3d_i h_i^2, \\
y'' &= 2c_i + 6d_i(x - x_i), \\
y_i'' &= 2c_i, \\
y_{i+1}'' &= 2c_i + 6d_i h_i
\end{aligned}
$$

Let's take the values of the second derivatives $D_i = y_i''$ as the new variables. For the unknown variables we get the expressions

$$\begin{aligned}
a_i &= y_i, \\
c_i &= D_i/2, \\
d_i &= (D_{i+1} - D_i)/6h_i, \\
b_i &= \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i D_i + h_i D_{i+1}}{6}.
\end{aligned}$$

At the beginning of the interval i we have $y_i' = b_i$. Using the previous interval the derivative is

$$\begin{aligned} y_i' &= b_{i-1} + 2c_{i-1}(x_i - x_{i-1}) + 3d_{i-1}(x_i - x_{i-1})^2 \\ &= b_{i-1} + 2c_{i-1}h_{i-1} + 3d_{i-1}h_{i-1}^2. \end{aligned}$$

These expressions must be equal. We'll then express the constants in terms of the derivatives D_i and the y values:

$$\begin{aligned} y_i' = \frac{y_{i+1} - y_i}{h_i} - \frac{2h_iD_i + h_iD_{i+1}}{6} \\ = 3\left(\frac{D_i - D_{i+1}}{6h_{i-1}}\right)h_{i-1}^2 + 2\left(\frac{D_{i-11}}{2}\right)h_{i-1} \\ + \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2h_{i-1}D_{i-1} + h_{i-1}D_i}{6}, \end{aligned}$$

which yields

$$\begin{aligned}
&h_{i-1}D_{i-1} + (2h_{i-1} + 2h_i)D_i + h_i D_{i+1} \\
&= 6\left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right), \\
&i = 2, \ldots, n-1.
\end{aligned}$$

Here we have $n-2$ equations and n unknowns D_i. Additionally, we can take, for example, $D_1 = D_n = 0$, in which case the coefficient matrix is

$$\begin{pmatrix}
2(h_1 + h_2) & h_2 & 0 & 0 \cdots & 0 & 0 \\
h_2 & 2(h_2 + h_3) & h_3 & 0 \cdots & 0 & 0 \\
0 & h_3 & 2(h_3 + h_4) & \cdots & 0 & 0 \\
\vdots & & & & & \\
0 & 0 & 0 & \cdots & h_{n-2} & 2(h_{n-2} + h_{n-1})
\end{pmatrix}.$$

This is a tridiagonal set of equations, which is easy to solve.

```
subroutine cubicspline(n, x, y, a, b, c, d)
 integer, intent(in) :: n
 real, intent(in), dimension(maxpoint) :: x, y
 real, intent(out), dimension(maxpoint) :: a, b, c, d

 integer i
 real, dimension(maxpoint,4) :: u
 real, dimension(maxpoint) :: s, h
 real t

 do i=1,n-1          ! h = step in x direction
   h(i) = x(i+1)-x(i)
 end do

 do i=1,n-2          ! calculate the coefficient matrix
  u(i, 1) = h(i)
  u(i, 2) = 2*(h(i)+h(i+1))
  u(i, 3) = h(i+1)   ! right hand side vector
  u(i, 4) = 6.0* ((y(i+2)-y(i+1))/h(i+1) &
          - (y(i+1)-y(i))/h(i))
 end do
```

```
u(1,1) = 0.0
u(n-2,3) = 0.0

do i=2,n-2        ! elimination
 t = u(i, 1)/u(i-1, 2)
 u(i, 2) = u(i,2) - t * u(i-1, 3)
 u(i, 4) = u(i,4) - t * u(i-1, 4)
end do

u(n-2,4) = u(n-2,4) / u(n-2,2)  ! back substitution
do i=n-3,1,-1
 u(i,4) = (u(i, 4) - u(i, 3) * u(i+1, 4))/u(i,2)
end do

s(1) = 0.0        ! second derivatives
do i=2,n-1
 s(i)=u(i-1,4)
end do
s(n) = 0.0

do i=1,n-1        ! spline coefficients
 a(i) = y(i)
 b(i) = (y(i+1)-y(i))/h(i) - (2.0*h(i)*s(i) + h(i)*s(i+1))/6.0
```

```
    c(i) = s(i) / 2.0
    d(i) = (s(i+1)-s(i))/(6*h(i))
  end do

end subroutine
```

Splines can be ”stiffened” to avoid sharp bends, but then they do not describe the data exactly.

Splines are good for interpolation, not for extrapolation. (They are polynomials!)

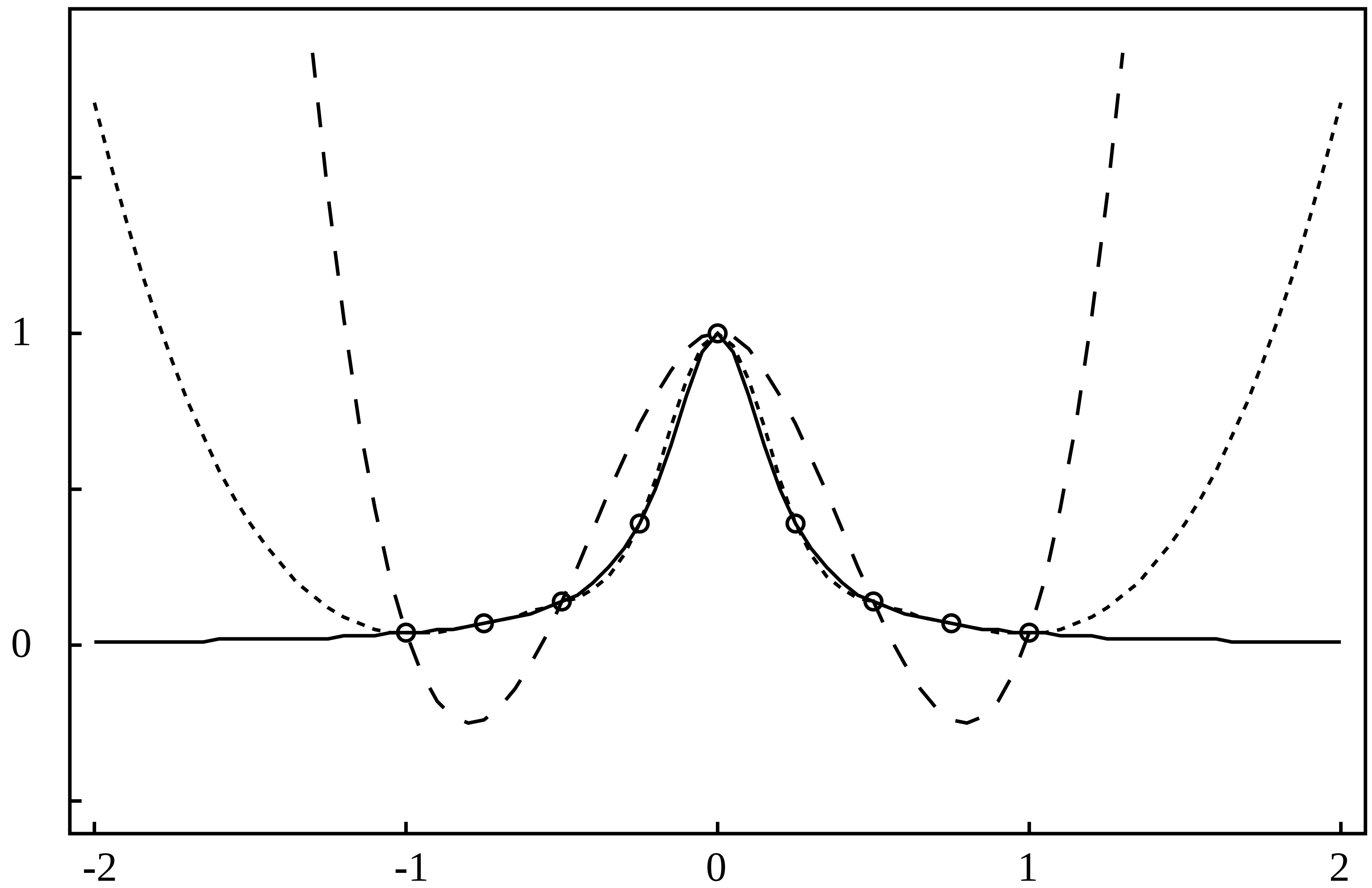
If the data contains sharp turns or long empty gaps, extra bends or big oscillations may appear.

Problems can be fixed by adding more data points. It is importanta that the data set is dense enough near sharp bends.

If one point of the data set is changed, the whole solution must be recomputed. The coefficient matrix is a band matrix, and the disturbance due to changing one point will usually diminish quickly when moving further from that point.

Example: Runge's function

$$y = \frac{1}{1 + 25x^2}.$$



Twodimensional curves

This far we have assumed that the x values are monotonously increasing.

If the data represents a general plane curve, the x coordinate cannot be used as the independent variable.

If both coordinates depend on some paramater t, that can be used as the independent variable. We get two functions $x = x(t)$ and $y = y(t)$ that can be described separately using spline.

Otherwise we have to select some parametrisation. For example

$$
\begin{aligned}
x_i(t) &= a_i + b_i t + c_i t^2 + d_i t^3, \\
y_i(t) &= a'_i + b'_i t + c'_i t^2 + d'_i t^3.
\end{aligned}
$$

where $0 \le t \le 1$.

Bézier curves

Bernštein interpolation polynomial is a curve in the complex plane

$$z(t) = (1-t)^3 z_1 + 3t(1-t)^2 z_2 + 3t^2(1-t) z_3 + t^3 z_4,$$

whwre $0 \leq t \leq 1$ and the points z_i are the four control points of the curve.

Pierre Bézier introduced these in the 1960's to computer aided design.

Bézier curve is defined in terms of two endpoints and two control poinys. The curve will always pass through the endpoints. The control points give the directions of the tangent at the endpoints.

The curve will start from the endpoint in the direction of the control point. The further away the control point is the more straight the curve is.

Convenient in interactive applications.

The PostScript language contains an operator (`curveto`) for drawing Bézier curves.