Interpolation by Splines

GOAL. •

Understand what splines are

• Why the spline is introduced

• Approximating functions by splines

We have seen in previous lecture that a function f(x) can be interpolated at n + 1 points in an interval [a, b] using a single polynomial pn(x) defined over the entire interval.

To draw smooth curves through data points, drafters once used thin flexible strips of wood, hard rubber, metal or plastic called mechanical splines. To use a mechanical spline, pins were placed at a judicious selection of points along a curve in a design, then the spline was bent so that it touched each of these pins. Clearly, with this construction, the spline interpolates the curve at these pins and could be used to reproduce the curve in other drawings. The location of the pins are called knots. We can change the shape of the curve defined by the spline by adjusting the location of the knots. For example, to interpolate the data {(xi , fi)} we can place knots at each of the nodes xi . Mathematically, a spline function consists of polynomial pieces on subintervals joined together with certain continuity conditions.

Cubic spline interpolation is a fast, efficient and stable method of function interpolation. The spline interpolation is an alternative for the [polynomial interpolation](http://www.alglib.net/interpolation/polynomial.php).

The spline interpolation is based on the following principle. The interpolation interval is divided into small subintervals. Each of these subintervals is interpolated by using the third-degree polynomial. The polynomial coefficients are chosen to satisfy certain conditions (these conditions depend on the interpolation method). General requirements are function continuity and, of course, passing through all given points. There could also be additional requirements: function linearity between nodes, continuity of higher derivatives and so on.

The main advantages of spline interpolation are its stability and calculation simplicity. Sets of linear equations which should be solved to construct splines are very well-conditioned, therefore, the polynomial coefficients are calculated precisely. As a result, the calculation scheme stays stable even for big N.

***Spline types***

***Linear spline***

The linear spline is just a piecewise linear function. The linear splines have low precision, it should also be noted that they do not even provide first derivative continuity. However, in some cases, piecewise linear approximation could be better than higher degree approximation. For example, the linear spline keeps the monotony of a set of points.

All splines considered on this page are **cubic splines** - they are all piecewise cubic functions. However, if someone says "cubic spline", they usually mean a special cubic spline with continuous first and second derivatives. The cubic spline is given by the function values in the nodes and derivative values on the edges of the interpolation interval (either of the first or second derivatives).

* If the exact values of the first derivative in both boundaries are known, such spline is called *clamped spline*, or *spline with exact boundary conditions*. This spline has interpolation error *O(h4)*.
* If the value of the first (or second) derivative is unknown, we can set the so-called natural boundary conditions*S''(A)=0*, *S''(B)=0*. Thus, we get a natural spline. The natural spline has interpolation error *O(h2)*. The closer to the boundary nodes the more the error becomes. In the inner nodes the interpolation accuracy is much better.
* One more boundary condition which we can use when boundary derivatives are unknown is the "parabolically terminated spline". In this case, the boundary interval is represented as the second (instead of the third) degree polynomial (for inner intervals, third-degree polynomials are still used). In a number of cases this provides better accuracy than natural boundary conditions.