HISTORY

Numerical values of analytic functions (trigonometric, exponential, logarithmic, etc.) are compiled and presented to users in tabular forms. Quite often, the exact answers are not listed and users need to read between the lines. That is when "Interpolation" comes to the rescue. In simple terms, interpolation helps the estimation of data between two known values. In Ancient Babylon and Greece times, astronomy was developed to keep track of time and to predict astronomical events about the sun, the moon and planets. The astronomical records though regularly made would contain gaps due to the observation being hampered or celestial bodies being invisible. Ancient mathematicians were able to fill the gaps by using simple, linear interpolation. As time went by, higher order interpolation had been established and employed.

The current page is to highlight the major formulations of interpolation in history. Interpolation is an approximation theory and its goal for one dimensional case is to construct analytically a function f(x) which is based on given values of f(x) at N locations ($x_1, x_2, \ldots x_N$) in the domain interval. Two interpolation formulas (interpolants) are given first. (1) Newton Interpolant (one dimension); (2) Lagrange Interpolant (one dimension).

With function values given at three X locations:

$$f_A(x) = a_1 + a_2\ (x-x_1) + a_3\ (x-x_1)\ (x-x_2) \qquad \text{(Newton's)}$$

$$\text{where,}\quad a_1 = f(x_1),\quad a_2 = \frac{f(x_2)-f(x_1)}{x_2-x_1}\ ,\quad a_3 = \frac{\frac{f(x_3)-f(x_2)}{x_3-x_2}-\frac{f(x_2)-f(x_1)}{x_2-x_1}}{x_3-x_1}$$

$$f_A(x) = \frac{(x-x_2)(x-x_3)f(x_1)}{(x_1-x_2)(x_1-x_3)} + \frac{(x-x_3)(x-x_1)f(x_2)}{(x_2-x_3)(x_2-x_1)} + \frac{(x-x_1)(x-x_2)f(x_3)}{(x_3-x_1)(x_3-x_2)} \qquad \text{(Lagrange's)}$$

$f_A(x)$ is the interpolant. Note that these two interpolants are polynomial-based and are exactly the same but with different appearance. Sample locations X_1, X_2, X_3 can be equally spaced or scattered at random. Computationally speaking, Newton's form is easier to accommodate more/less data values. However, Lagrange's form provides the notion of a "global weighting scheme" for interpolation. Next, the interpolant of Whittaker-Shannon (also called Cardinal Series) has the form,

$$f_A(t) = \sum_{n=-\infty}^{+\infty} a_n\ \mathrm{SINC}\left(\frac{t-nt_0}{t_0}\right)$$

Note that $f(nt_0) = a_n$. The normalized SINC function is defined as,

$$\text{SINC}\left(\frac{t-nt_0}{t_0}\right) = \frac{\text{SIN}\left[\pi\left(\frac{t-nt_0}{t_0}\right)\right]}{\left[\pi\left(\frac{t-nt_0}{t_0}\right)\right]}$$

$$\textit{and} \qquad \int_{-\infty}^{+\infty} \text{SINC}\left(\frac{t-nt_0}{t_0}\right) dt = \int_{-\infty}^{+\infty} \frac{\text{SIN}\left[\pi\left(\frac{t-nt_0}{t_0}\right)\right]}{\left[\pi\left(\frac{t-nt_0}{t_0}\right)\right]} dt = t_0$$

The formula of $f_A(t)$ above, which is the interpolant, leads to the well-known Shannon's sampling theorem in information theory (1948). The variable t is the time variable and t_0 is the sampling interval. The term a_n is the sample collected at $t = nt_0$. Given all the "infinite, discrete" samples from n=-∞ to n=+∞, the original "continuous" signal $f(t)$ can be interpolated and perfectly reconstructed for all "t" values by using the formula. That is, $f_A(t) = f(t)$ in this case. The quantity $\frac{1}{t_0}$ is the sample rate (or sampling frequency) and the quantity $\frac{1}{2t_0}$ is the Nyquist frequency which is the highest frequency contained in the signal. Samples are periodically collected.

The Shepard's interpolation formula (1968) is an "inverse distance weighting" method. The formula can be used in 1-dimension or multi-dimension case (2-Dimension or higher). Shepard introduced the notion of inverse distance in the formula. The idea is compatible to Tobler's First Law of Geography, namely, "Everything is related to everything else. But near things are more related than distant things." It can be seen that Shepard's formula is originally established for data interpolation in "physical" geographical space. In Shepard's formula, the position vector is $\vec{x} = (x_1, x_2, \ldots x_m)$ where m is the dimensionality. Sample location $\vec{x_i} = (x_{i1}, x_{i2}, \ldots x_{im})$, where i = 1,2,3.....N, and N is the sample number.

$$f_A(\vec{x}) = \frac{\sum_{i=1}^{N} w_i(\vec{x}, \vec{x_i}) f(\vec{x_i})}{\sum_{i=1}^{N} w_i(\vec{x}, \vec{x_i})} \qquad \text{for } \vec{x} \neq \vec{x_i}$$

$$f_A(\vec{x}) = f(\vec{x_i}) \qquad \text{for } \vec{x} = \vec{x_i}$$

$w_i(\vec{x}, \vec{x_i})$ is the weighting function which is defined as,

$$w_i(\vec{x}, \vec{x_i}) = \frac{1}{d(\vec{x}, \vec{x_i})^p} \qquad \text{(Inverse distance)}$$

and $d(\vec{x}, \vec{x_i})^p$ is the distance function and p is the positive power parameter (for example, p=1 for the Euclidean distance) . The distance function is defined as,

$$d(\overrightarrow{x}, \overrightarrow{x_i})^p = [\sqrt[2]{(x_1 - x_{i1})^2 + (x_2 - x_{i2})^2 + \cdots .. + (x_m - x_{im})^2}\ ]^p$$

The Hardy's multiquadric formula (1971) is one type of RBF (radial basis function) interpolation method. The formula is generally regarded as one of the best for application. The interpolant takes the form of a weighted sum of radial basis functions. It has the following form,

$$f_A(\overrightarrow{x}) = \sum_{i=1}^{N} w_i\ H_i(\vec{x}, \overrightarrow{x_i})$$

w_i are the weights and need to be solved, and $H_i(\vec{x}, \overrightarrow{x_i})$ are the radial basis functions. N is the sample number.

$$H_i(\vec{x}, \overrightarrow{x_i}) = \sqrt[2]{(x_1 - x_{i1})^2 + (x_2 - x_{i2})^2 + \cdots + (x_m - x_{im})^2 + \Delta^2}$$

$H_i(\vec{x}, \overrightarrow{x_i})$ is Hardy's RBF and Δ is the chosen smoothing parameter. Another popular RBF is the Gaussian function,

$$H_i(\vec{x}, \overrightarrow{x_i}) = [\frac{1}{\Delta\sqrt{2\pi}}]\, e^{\frac{-[(x_1 - x_{i1})^2 + (x_2 - x_{i2})^2 + \cdots + (x_m - x_{im})^2]}{2\Delta^2}}$$

Note that $H_i(\vec{x}, \overrightarrow{x_i}) = \left[\frac{1}{\Delta\sqrt{2\pi}}\right]$ when $\vec{x} = \overrightarrow{x_i}$.

In mathematical science, there is a parallel development of approximation theory in statistics, which is called regression analysis. Given N sample values, the regression formula can interpolate as well as extrapolate to predict data values. In "nonparametric, kernel" regression analysis, the kernel acts like the weighting function in Shepard's formula. The well-known formula (or called estimator), "Nadaraya-Watson" (1964) has the form,

$$f_A(\overrightarrow{x}) = \frac{\sum_{i=1}^{N} w_i(\overrightarrow{x}, \overrightarrow{x_i})\, f(\overrightarrow{x_i})}{\sum_{i=1}^{N} w_i(\overrightarrow{x}, \overrightarrow{x_i})}$$

$$w_i(\overrightarrow{x}, \overrightarrow{x_i}) = k_\Delta(\vec{x} - \overrightarrow{x_i}) = k_\Delta(x_1 - x_{i1})k_\Delta(x_2 - x_{i2}) \ldots . k_\Delta(x_m - x_{im})$$

$k_\Delta(\vec{x} - \overrightarrow{x_i})$ is the kernel in m-dimensional space and is the composite product of the kernel of one-dimensional space , assuming the same Δ for all dimensions. For example, a popular one-dimensional kernel function is the Gaussian kernel,

$$k_{\Delta}(x_j - x_{ij}) == [\frac{1}{\Delta\sqrt{2\pi}}]\, e^{\frac{-(x_j - x_{ij})^2}{2\Delta^2}}$$

One more powerful statistical estimator method is called "Kriging" (1951). It is originated from geostatistics (mining). The method depends on "variogram modeling". Variogram modeling requires understanding of statistics as well as the underlying processes from which the measurement data are drawn. In essence, Kriging has the formula like Shepard's,

$$f_A(\vec{x}) = \sum_{i=1}^{N} \lambda_i(\vec{x}, \vec{x_i})\, f(\vec{x_i}) \qquad \text{for } \vec{x} \neq \vec{x_i}$$

$$f_A(\vec{x}) = f(\vec{x_i}) \qquad \text{for } \vec{x} = \vec{x_i}$$

$\lambda_i(\vec{x}, \vec{x_i})$ is the weight. However, unlike Shepard's but similar to Hardy's, the weight needs to be solved through algebraic process.