DISCRETE FUNCTION APPROXIMATION BY LEAST SQUARES

Ref. : « Méthode de calcul numérique – Tome 2 – Programmes en Basic et en Pascal By Claude Nowakowski, Editions du P.S.1 ?, Paris 1984, p. 23 - 26 »

Translation in English By J-P Moreau, Paris.

This function f approximation by least squares is based on norm :

$$|| f || = < f, f >^{1/2}$$

where < f, g > is the scalar product, i. e. if g is defined by its components f0, f1, …, fn and g by g0, g1, …, gn then < f, g > = f0.g0 + f1.g1 + … + fn.gn is a scalar (number).

Let us suppose the f function be given by n+1 $y_i = f(x_i)$ ordinate values at n+1 distinct abscissa xi values within interval [a, b], then :

$$F^* = \sum_{k=0}^{m} a_k^* \, phi_k$$ is the best discreet least squares approximation of f, only if :

$$\sum_{i=0}^{N} [f(x_i) - F^*(x_i)]^2 < \sum_{i=0}^{N} [f(x_i) - F(x_i)]^2$$

for any F function belonging to F_{m+1}, subspace of dimension m+1 of continuous functions in interval [a, b], with vectorial base phi_0, phi_1, …, phi_m, such as :

$$F = \sum_{k=0}^{m} a_k \; phi_k$$

The problem is to calculate the m+1 a_k^* coefficients.

Let be r_i the error, here called residual, for each i point :

$$r_i = F(x_i) - y_i \quad | \; i=0, n$$

The F(x) function that gives the best least squares approximation, for the given set of data, is the linear combination $a_0\, phi_0(x) + a_1\, phi_1(x) + \dots + a_m\, phi_m(x)$ that gives the smallest sum of residual squares :

$$Q = \sum_{i=0}^{n} r_i^2$$

$$= \sum_{i=0}^{n} [a_0 phi_0(x_i) + a_1 phi_1(x_i) + \dots + a_m phi_m(x_i) - y_i]^2$$

So Q is a function of parameters a_0, a_1, …, a_m (if one of the parameters varies, Q varies) and if one considers a_k parameters as independent variables for Q. The problem is then to minimize this Q function. The minimum is obtained when the m+1 partial derivatives of $Q(a_0,a_1,\dots,a_m)$ for a_k are simultaneously null :

$$\frac{\partial dQ}{\partial a_k} = 2 \sum [\, F(x_i) - y_i \,] \bullet \frac{\partial F(xi)}{\partial a_k} = 0 \quad | \; k=0, m$$

Thus we obtain a linear system with m +1 equations the unknowns of which are $a_0, a_1, \ldots, a_m$. Let us calculate the coefficients of the system matrix: We can easily see that

$$\frac{\partial F(xi)}{\partial a_k} = \phi_k(x_i)$$

Hence (for k = 0, m):

$$\frac{\partial Q}{\partial a_k} = 2 \sum_{i=0}^{n} [a_0\, \phi_0(x_i) + \ldots + a_m\, \phi_m(x_i) - y_i]\, \phi_k(x_i) = 0$$

or else (for k = 0, m):

$$\sum [\, a_0\, \phi_k(x_i)\, \phi_0(x_i) + a_1\, \phi_k(x_i)\, \phi_1(x_i) + \ldots + a_m\, \phi_k(x_i)\, \phi_m(x_i) \,] = \sum \phi_k(x_i)\, y_i$$

or in a matrix form defining the normal equations system:

$$\begin{bmatrix} \sum [\varphi_0(x_i)]^2 & \sum \varphi_0(x_i)\varphi_1(x_i) \text{----} & \sum \varphi_0(x_i)\varphi_m(x_i) \\ \sum \varphi_1(x_i)\varphi_0(x_i) & \sum [\varphi_1(x_i)]^2 \text{-------} & \sum \varphi_1(x_i)\varphi_n(x_i) \\ \text{------} & \text{------} & \text{------} \\ \sum \varphi_m(x_i)\varphi_0(x_i) & \sum \varphi_m(x_i)\varphi_1(x_i) & \sum [\varphi_m(x_i)]^2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \text{--} \\ \alpha_m \end{bmatrix} = \begin{bmatrix} \sum \varphi_0(x_i).y_i \\ \sum \varphi_1(x_i).\, y_i \\ \text{------} \\ \sum \varphi_m(x_i)\; y_i \end{bmatrix}$$

We shall retain as $\phi_k(x)$ the particular case of polynomials having the form :

$\phi_k(x) = x^k$ i.e. $F = P_m(x)$ with

$P_m(x) = \alpha_0\, x^0 + \alpha_1\, x^1 + \alpha_2\, x^2 + \ldots + \alpha_m\, x^m$

Warning : This polynomial $P_m(x)$ must not be confused with the interpolation polynomial !

We have seen that the α_k parameters must be determined such as:

$$Q = \sum_{i=0}^{n} r_i^2$$

$$= \sum_{i=0}^{n} [\, P_m(x_i) - y_i \,]^2$$

be mimimum.

This approximation technique is often called *polynomial smoothing*. The normal equations to calculate the α_k coefficients for this particular case can easily be obtained by substituting x_I^k in $\phi_k(x_i)$. Hence:

$$
\begin{bmatrix}
\sum_i (x_0^0)^2 & \sum (x_i^0 x_i^1) & \cdots & \sum (x_i^0 x_i^m) \\
\sum (x_i^1 x_i^0) & \sum (x_i^1)^2 & \cdots & \sum (x_i^1 x_i^m) \\
\cdots & \cdots & \cdots & \cdots \\
\sum (x_i^m x_i^0) & \cdots & \cdots & \sum (x_i^m)^2
\end{bmatrix}
\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}
=
\begin{bmatrix} \sum x_i^0 y_i \\ \sum x_i^1 y_i \\ \vdots \\ \sum x_i^m y_i \end{bmatrix}
$$

et en effectuant :

$$
\begin{bmatrix}
n+1 & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^m \\
\sum x_i & \sum x_i^2 & \cdots & \cdots & \sum x_i^{m+1} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\sum x_i^m & \sum x_i^{m+1} & \cdots & \cdots & \sum x_i^{2m}
\end{bmatrix}
\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}
=
\begin{bmatrix} \sum x_i^0 y_i \\ \sum x_i^1 y_i \\ \vdots \\ \sum x_i^m y_i \end{bmatrix}
$$

So the matrix is symetrical and one can notice that the determinant is very often near zero. In such a case, the matrix is ill conditioned and the solution (the α_k coefficients) strongly varies for small changes in the matrix coefficient.

We can show that, for $x \varepsilon [0, 1]$ and regularly spaced points, the matrix is, to a scaling factor, an Hilbert matrix of order n the determinant of which is given by:

$$H_n = \frac{[1!\,2!\,3!\ldots(n-1)!]^3}{n!\,(n+1)!\ldots(2n-1)!}$$

The matrix coefficients are:

$$H_{ij} = \frac{1}{i+j-1}$$