COURSE 6

2.6. Least squares approximation

• It is an extension of the interpolation problem.

• More desirable when the data are contaminated by errors.

• To estimate values of parameters of a mathematical model from measured data, which are subject to errors.

When we know $f(x_i)$, $i = 0, ..., m$, an interpolation method can be used to determine an approximation φ of the function f, such that

$$\varphi\left(x_i\right) = f\left(x_i\right), \ i = 0, ..., m.$$

If only approximations of $f\left(x_i\right)$ are available or the number of interp. conditions is too large, instead of requiring that the approx. function reproduces $f(x_i)$ exactly, we ask only that it fits the data "as closely as possible".

The least squares approximation φ is determined such that:

- in the discrete case:

$$\left(\sum_{i=0}^{m}\left[f\left(x_i\right)-\varphi\left(x_i\right)\right]^2\right)^{1/2} \to \text{min},$$

- in the continuous case:

$$\left(\int_a^b \left[f\left(x\right)-\varphi\left(x\right)\right]^2 dx\right)^{1/2} \to \text{min},$$

Remark 1 *Notice that the interpolation is a particular case of the least squares approximation, with*

$$f\left(x_i\right)-\varphi\left(x_i\right)=0, \quad i=0,...,m.$$

Remark 2 *The first clear and concise exposition of the method of least squares was first published by Legendre in 1805. In 1809 Carl Friedrich Gauss applied the method in calculating the orbits of celestial bodies. In that work he claimed and proved that he have been in possession of the method since 1795.*

Linear least square. Consider the data

x	1	2	3	4	5
$f(x)$	1	1	2	2	4

The problem consists in finding a function φ that "best" represents the data.

Plot the data and try to recognize the shape of a "guess function φ" such that $f \approx \varphi$.

For this example, a resonable guess may be a linear one, $\varphi(x) = ax + b$. The problem: find a and b that makes φ the best function to fit the data. The least squares criterion consists in minimizing the sum

$$E(a,b)=\sum_{i=0}^{4}\left[f\left(x_i\right)-\varphi\left(x_i\right)\right]^2=\sum_{i=0}^{4}\left[f\left(x_i\right)-\left(ax_i+b\right)\right]^2.$$

The minimum of the sum is obtained when

$$\frac{\partial E(a,b)}{\partial a}=0$$
$$\frac{\partial E(a,b)}{\partial b}=0.$$

We get

$$55a+15b=37$$
$$15a+5b=10$$

and further $\varphi(x)=0.7x-0.1$.

Consider a more general problem with the data from the table

x	x_0	x_1	...	x_m
$f(x)$	y_0	y_1	...	y_m

and the approximating linear function $\varphi(x) = ax + b$. We have to find a and b.

We have to minimize the sum

$$E(a,b) = \sum_{i=0}^{m} \left[f\left(x_i\right) - \varphi\left(x_i\right)\right]^2 = \sum_{i=0}^{m} \left[f\left(x_i\right) - \left(ax_i + b\right)\right]^2. \qquad (1)$$

The minimum of the sum is obtained by

$$\frac{\partial E(a,b)}{\partial a} = 2 \sum_{i=0}^{m} \left[f\left(x_i\right) - \left(ax_i + b\right)\right] \cdot \left(-x_i\right) = 0$$
$$\frac{\partial E(a,b)}{\partial b} = 2 \sum_{i=0}^{m} \left[f\left(x_i\right) - \left(ax_i + b\right)\right] \cdot (-1) = 0$$

These are called **normal equations**. Further,

$$\sum_{i=0}^{m} x_i f\left(x_i\right) = a \sum_{i=0}^{m} x_i^2 + b \sum_{i=0}^{m} x_i$$
$$\sum_{i=0}^{m} f\left(x_i\right) = a \sum_{i=0}^{m} x_i + (m+1)b.$$

The solution is

$$a = \frac{(m+1)\sum\limits_{i=0}^{m} x_i f(x_i) - \sum\limits_{i=0}^{m} x_i \sum\limits_{i=0}^{m} f(x_i)}{(m+1)\sum\limits_{i=0}^{m} x_i^2 - (\sum\limits_{i=0}^{m} x_i)^2} \tag{2}$$

$$b = \frac{\sum\limits_{i=0}^{m} x_i^2 \sum\limits_{i=0}^{m} f(x_i) - \sum\limits_{i=0}^{m} x_i f(x_i) \sum\limits_{i=0}^{m} x_i}{(m+1)\sum\limits_{i=0}^{m} x_i^2 - (\sum\limits_{i=0}^{m} x_i)^2}.$$

Polynomial least squares. In many experimental results the data are not linear. Suppose that

$$\varphi(x) = \sum_{k=0}^{n} a_k x^k, \quad n \leq m$$

Find $a_i, i = 0, ..., n$, that minimize the sum

$$\begin{aligned} E(a_0, ..., a_n) &= \sum_{i=0}^{m} \left[f\left(x_i\right) - \varphi\left(x_i\right)\right]^2 \\ &= \sum_{i=0}^{m} \left[f\left(x_i\right) - \sum_{k=0}^{n} a_k x_i^k\right]^2 . \end{aligned} \tag{3}$$

The minimum is obtained when

$$\frac{\partial E(a_0, ..., a_n)}{\partial a_j} = 0, \quad j = 0, ...n,$$

which are **the normal equations** and have a unique solution.

General case. Solution of the least squares problem is

$$\varphi(x) = \sum_{i=1}^{n} a_i g_i(x),$$

where $\{g_i,\ i = 1, ..., n\}$ is a basis of the space and the coefficients a_i are obtained solving **the normal equations**:

$$\sum_{i=1}^{n} a_i \langle g_i, g_k \rangle = \langle f, g_k \rangle, \qquad k = 1, ..., n.$$

In the discrete case

$$\langle f, g \rangle = \sum_{k=0}^{m} w(x_k) f(x_k) g(x_k)$$

and in the continuous case

$$\langle f, g \rangle = \int_a^b w(x) f(x) g(x),$$

where w is a weight function.

Example 3 *Having the data*

x	0	1	2	3
$f(x)$	-4	0	4	-2

find the corresponding least squares polynomial of the first degree.

Sol. We have

$$E(a,b)=\sum_{i=0}^{3}\left[f\left(x_i\right)-\varphi\left(x_i\right)\right]^2=\sum_{i=0}^{3}\left[f\left(x_i\right)-\left(ax_i+b\right)\right]^2 \tag{4}$$

and we have to find a and b from the system

$$\begin{cases} \frac{\partial E(a,b)}{\partial a}=2\sum_{i=0}^{3}\left[f\left(x_i\right)-\left(ax_i+b\right)\right]\cdot x_i=0 \\ \frac{\partial E(a,b)}{\partial b}=2\sum_{i=0}^{3}\left[f\left(x_i\right)-\left(ax_i+b\right)\right]=0 \end{cases}$$

$$\begin{cases} \sum_{i=0}^{3}\left[f\left(x_i\right)-\left(ax_i+b\right)\right]\cdot x_i=0 \\ \sum_{i=0}^{3}\left[f\left(x_i\right)-\left(ax_i+b\right)\right]=0 \end{cases}$$

Example 4 *Fit the data in table*

x	-3	-1	2
$f(x)$	-4	-2	3

a) with the best least squares line;

b) with the best least squares polynomial of degree at most 2.