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".

It seems that the least squares method was first introduced by C. F. Gauss in 1795. In 1801 he used it for making the best prediction for the orbital position of the planet Ceres (dwarf planet, lies between Mars and Jupiter, first considered planet and further reclassified as an asteroid) using the measurements of G. Piazzi (who was the first that discovered it).

The first clear and concise exposition of the method of least squares was first published by A. M. Legendre in 1805. P. S. Laplace and R. Adrain have also contributed to the development of this theory.

In 1809 C. F. Gauss applied the method in calculating the orbits of some celestial bodies. In that work he claimed and proved that he have been in possession of the method since 1795. 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*

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

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

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

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,

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

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

The solution is

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

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

Example 2 *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

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

and we have to find a and b from the system

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

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

Polynomial least squares. In many experimental results the data are not linear or can be better estimated by a polynomial. Suppose that

$$\varphi(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Consider $m + 1$ points (x_i, y_i), $i = 0, \ldots, m$.

We have to find $a_i, i = 0, ..., n$, that minimize the sum

$$\begin{aligned} E(a_0, ..., a_n) &= \sum_{i=0}^{m} \left[f\left(x_i\right) - \varphi\left(x_i\right) \right]^2 \qquad (4) \\ &= \sum_{i=0}^{m} \left[f\left(x_i\right) - \sum_{k=0}^{n} a_k x_i^k \right]^2 . \end{aligned}$$

Denoting $y_i = f(x_i)$ we have

$$
\begin{aligned}
E(a_0, ..., a_n) &= \sum_{i=0}^{m} y_i^2 - 2\sum_{i=0}^{m} y_i \varphi(x_i) + \sum_{i=0}^{m} (\varphi(x_i))^2 \\
&= \sum_{i=0}^{m} y_i^2 - 2\sum_{i=0}^{m} \left(\sum_{j=0}^{n} a_j x_i^j\right) y_i + \sum_{i=0}^{m} \left(\sum_{j=0}^{n} a_j x_i^j\right)^2 \\
&= \sum_{i=0}^{m} y_i^2 - 2\sum_{j=0}^{n} a_j \left(\sum_{i=0}^{m} x_i^j y_i\right) + \sum_{j=0}^{n} \sum_{k=0}^{n} a_j a_k \left(\sum_{i=0}^{m} x_i^{j+k}\right).
\end{aligned}
$$

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.

It is obtain

$$
\frac{\partial E}{\partial a_j} = -2\sum_{i=0}^{m} x_i^j y_i + 2\sum_{k=0}^{n} a_k \left(\sum_{i=0}^{m} x_i^{j+k}\right) = 0
$$

which gives $n+1$ unknowns $a_j,\ j \in \{0,1,\dots,n\}$ and $n+1$ equations

$$\sum_{i=0}^{m} x_i^j y_i = \sum_{k=0}^{n} a_k \left(\sum_{i=0}^{m} x_i^{j+k} \right), \quad \text{for each } j \in \{0,1,\dots,n\}. \tag{5}$$

We have the system

$$
\begin{aligned}
&a_0 \sum_{i=0}^{m} x_i^0 + a_1 \sum_{i=0}^{m} x_i^1 + a_2 \sum_{i=0}^{m} x_i^2 + \cdots + a_n \sum_{i=0}^{m} x_i^n = \sum_{i=0}^{m} x_i^0 y_i,\\
&a_0 \sum_{i=0}^{m} x_i^1 + a_1 \sum_{i=0}^{m} x_i^2 + a_2 \sum_{i=0}^{m} x_i^3 + \cdots + a_n \sum_{i=0}^{m} x_i^{n+1} = \sum_{i=0}^{m} x_i^1 y_i\\
&\cdots\\
&a_0 \sum_{i=0}^{m} x_i^n + a_1 \sum_{i=0}^{m} x_i^{n+1} + a_2 \sum_{i=0}^{m} x_i^{n+2} + \cdots + a_n \sum_{i=0}^{m} x_i^{2n} = \sum_{i=0}^{m} x_i^n y_i
\end{aligned}
$$

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 *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.

3. Numerical integration of functions

The need: for evaluating definite integrals of functions that has no explicit antiderivatives or whose antiderivatives are not easy to obtain.

Let $f : [a, b] \to \mathbb{R}$ be an integrable function, x_k, $k = 0, ..., m$, distinct nodes from $[a, b]$.

Definition 4 *A formula of the form*

$$\int_a^b f(x)dx = \sum_{k=0}^{m} A_k f(x_k) + R(f),$$

is **a numerical integration formula** *or* **a quadrature formula**.

A_k - **the coefficients**; x_k–**the nodes**; $R(f)$ - **the remainder (the error)**.

Definition 5 **Degree of exactness (degree of precision)** *of a quadrature formula is r if and only if the error is zero for all the polynomials of degree $k = 0, 1, ..., r$, but is not zero for at least one polynomial of degree $r + 1$.*

From the linearity of R we have that the degree of exactness is r if and only if $R(e_i) = 0$, $i = 0, ..., r$ and $R(e_{r+1}) \neq 0$, where $e_i(x) = x^i$, $\forall i \in \mathbb{N}$.

3.1. Interpolatory quadrature formulas

Definition 6 *A quadrature formula*

$$\int_a^b f(x)dx = \sum_{k=0}^{m} A_k f(x_k) + R(f),$$

is **an interpolatory quadrature formula** *if it is obtained by integrating each member of an interpolation formula regarding the function f and the nodes x_k.*

Remark 7 *An interpolatory quadrature formula has its degree of exactness at least the degree of the corresponding interpolation polynomial.*

Consider Lagrange interpolation formula regarding the nodes $x_k \in [a,b]$, $k = 0, ..., m$:

$$f(x) = \sum_{k=0}^{m} \ell_k(x) f(x_k) + (R_m f)(x).$$

Integrating the two parts of this formula one obtains

$$\int_a^b f(x)dx = \sum_{k=0}^{m} A_k f(x_k) + R_m(f), \tag{6}$$

where

$$A_k = \int_a^b \ell_k(x)dx$$

and

$$R_m(f) = \int_a^b (R_m f)(x)dx. \tag{7}$$

If the nodes are equidistant, i.e., $x_k = a + kh$, $h = \frac{b-a}{m}$ then

$$A_k = (-1)^{m-k} \frac{h}{k!(m-k)!} \int_0^m \frac{t(t-1)...(t-m)}{(t-k)} dt, \;\; k = 0, ..., m. \tag{8}$$

The remainder from the Lagrange interpolation formula can be written as:

$$(R_m f)(x) = \frac{u(x)}{(m+1)!} f^{(m+1)}(\xi(x)),$$

where $u(x) = \prod_{k=0}^{m} (x - x_k)$, so the remainder of the quadrature formula may be written as

$$R_m(f) = \frac{1}{(m+1)!} \int_a^b u(x) f^{(m+1)}(\xi(x)) dx. \tag{9}$$

Definition 8 *The quadrature formulas with equidistant nodes are called* **Newton-Cotes formulas.**

Consider the case $m = 1$ $(x_0 = a, x_1 = b, h = b - a)$.

Lagrange polynomial is

$$(L_1 f)(x) = \frac{x-b}{a-b} f(a) + \frac{x-a}{b-a} f(b),$$

and integrating, one obtains

$$\begin{aligned}\int_a^b f(x)dx &\simeq \int_a^b \left[\frac{x-b}{a-b} f(a) + \frac{x-a}{b-a} f(b)\right] dx \\ &= \left[\frac{(x-b)^2}{2(a-b)} f(a) + \frac{(x-a)^2}{2(b-a)} f(b)\right]\Bigg|_a^b \\ &= \frac{b-a}{2}[f(a) + f(b)]\end{aligned}$$