Curve Fitting

Goal: Given a set of data (x_i, y_i) with $1 \le i \le n$, find a mathematical model, or function of $y(x)$, that would match as closely as possible, the given data.

Step 1: assume a model, or the format of the function $y(x)$, e.g., linear, exponential, log, …

Step 2: find the parameters of the function that minimize the squares of the errors/differences between the model prediction and the data.

1. Linear fitting

Step 1: the function is given by:

$$y(x) = ax + b$$

where a, b are constants to be determined.

Step 2: Find values of a, b

Approach: the sum of the squares of the differences between model and data is:

$$S = [y(x_1) - y_1]^2 + [y(x_2) - y_2]^2 + \cdots + [y(x_n) - y_n]^2 = \sum_{i=1}^{n} [y(x_i) - y_i]^2 = \sum_{i=1}^{n} [ax_i + b - y_i]^2$$

S is a function of a, b.

To minimize S,

$$\frac{\partial S}{\partial a} = 0$$

$$\frac{\partial S}{\partial a} = 2\sum_{i=1}^{n} [ax_i + b - y_i]\, x_i = 0$$

$$\frac{\partial S}{\partial b} = 2\sum_{i=1}^{n} [ax_i + b - y_i] \ = 0$$

After some calculation, you will get:

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{x=i}^{n} x_i\right)^2}$$

$$b = \frac{\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i}{n}$$

The coefficient of determination, a measure of how well model fits the data, is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} [y_i - y(x_i)]^2}{\sum_{i=1}^{n} \left[y_i - \frac{1}{n}\sum_{i=1}^{n} y_i\right]^2}$$

2. General theory of curve fitting

Find a function of the format:

$$y(x) = a_1 f_1(x) + a_2 f_2(x) + \cdots + a_m f_m(x) = \sum_{j=1}^{m} a_j f_j(x)$$

To fit the data of (x_i, y_i) with $1 \le i \le n$.

where $a_1, a_2, \ldots, a_m$ are parameters to be determined, $f_1(x), f_2(x), \ldots, f_m(x)$ are functions of x, e.g., 1, x, x^2, exp(x), sin(x), ….

$$S = [y(x_1) - y_1]^2 + [y(x_2) - y_2]^2 + \cdots + [y(x_n) - y_n]^2 = \sum_{i=1}^{n} [y(x_i) - y_i]^2 = \sum_{i=1}^{n} \left[\sum_{j=1}^{m} a_j f_j(x_i) - y_i \right]^2$$

Let $\frac{\partial S}{\partial a_1} = 0, \frac{\partial S}{\partial a_2} = 0, \ldots, \frac{\partial S}{\partial a_m} = 0$

We will get m equations, and we have m unknown of $(a_1, a_2, \ldots, a_m)$ -> determine a_j

$$\begin{aligned} S = [y(x_1) - y_1]^2 + [y(x_2) - y_2]^2 + \cdots + [y(x_n) - y_n]^2 \\ = [a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_m f_m(x_1) - y_1]^2 \\ + [a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_m f_m(x_2) - y_2]^2 \\ + \cdots \\ + [a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_m f_m(x_n) - y_n]^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial S}{\partial a_1} = \quad & 2[a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_m f_m(x_1) - y_1] f_1(x_1) \\ & + 2[a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_m f_m(x_2) - y_2] f_1(x_2) \\ & + \cdots \\ & + 2[a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_m f_m(x_n) - y_n] f_1(x_n) = 0 \end{aligned}$$

So,

$$\begin{aligned} & [a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_m f_m(x_1) - y_1] f_1(x_1) \\ & + [a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_m f_m(x_2) - y_2] f_1(x_2) \\ & + \cdots \\ & + [a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_m f_m(x_n) - y_n] f_1(x_n) = 0 \end{aligned}$$

So,

$$\begin{aligned} & f_1(x_1) f_1(x_1) a_1 + f_1(x_2) f_1(x_2) a_1 + \cdots + f_1(x_n) f_1(x_n) a_1 \\ & + f_1(x_1) f_2(x_1) a_2 + f_1(x_2) f_2(x_2) a_2 + \cdots + f_1(x_n) f_2(x_n) a_2 \\ & + \cdots \\ & + f_1(x_1) f_m(x_1) a_m + f_1(x_2) f_m(x_2) a_m + \cdots + f_1(x_n) f_m(x_n) a_m \\ & = y_1 f_1(x_1) + y_2 f_1(x_2) + \cdots + y_n f_1(x_n) \end{aligned}$$

So, if we let $\frac{\partial S}{\partial a_1} = 0$, we get an equation of:

$$\sum_{i=1}^{n} f_1(x_i)f_1(x_i)\, a_1 + \sum_{i=1}^{n} f_1(x_i)f_2(x_i)a_2 + \cdots + \sum_{i=1}^{n} f_1(x_i)f_m(x_i)\, a_m = \sum_{i=1}^{n} y_i f_1(x_i)$$

So, if we let $\frac{\partial S}{\partial a_2} = 0$, we get an equation of:

$$\begin{aligned} S = [y(x_1) - y_1]^2 + [y(x_2) - y_2]^2 + \cdots + [y(x_n) - y_n]^2 \\ = [a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_m f_m(x_1) - y_1]^2 \\ + [a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_m f_m(x_2) - y_2]^2 \\ + \cdots \\ + [a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_m f_m(x_n) - y_n]^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial S}{\partial a_2} = \quad & 2[a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_m f_m(x_1) - y_1] f_2(x_1) \\ & + 2[a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_m f_m(x_2) - y_2] f_2(x_2) \\ & + \cdots \\ & + 2[a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_m f_m(x_n) - y_n] f_2(x_n) = 0 \end{aligned}$$

So,

$$\begin{aligned} & [a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_m f_m(x_1) - y_1] f_2(x_1) \\ & + [a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_m f_m(x_2) - y_2] f_2(x_2) \\ & + \cdots \\ & + [a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_m f_m(x_n) - y_n] f_2(x_n) = 0 \end{aligned}$$

So,

$$\begin{aligned} & f_2(x_1)f_1(x_1)a_1 + f_2(x_2)f_1(x_2)a_1 + \cdots + f_2(x_n)f_1(x_n)a_1 \\ & + f_2(x_1)f_2(x_1)a_2 + f_2(x_2)f_2(x_2)a_2 + \cdots + f_2(x_n)f_2(x_n)a_2 \\ & + \cdots \\ & + f_2(x_1)f_m(x_1)a_m + f_2(x_2)f_m(x_2)a_m + \cdots + f_2(x_n)f_m(x_n)a_m \\ & = y_1 f_2(x_1) + y_2 f_2(x_2) + \cdots + y_n f_2(x_n) \end{aligned}$$

So, if we let $\frac{\partial S}{\partial a_2} = 0$, we get an equation of:

$$\sum_{i=1}^{n} f_2(x_i)f_1(x_i)\, a_1 + \sum_{i=1}^{n} f_2(x_i)f_2(x_i)a_2 + \cdots + \sum_{i=1}^{n} f_2(x_i)f_m(x_i)\, a_m = \sum_{i=1}^{n} y_i f_2(x_i)$$

…

So, if we let $\frac{\partial S}{\partial a_m} = 0$, we get an equation of:

$$\sum_{i=1}^{n} f_m(x_i)f_1(x_i)\, a_1 + \sum_{i=1}^{n} f_m(x_i)f_2(x_i)a_2 + \cdots + \sum_{i=1}^{n} f_m(x_i)f_m(x_i)\, a_m = \sum_{i=1}^{n} y_i f_m(x_i)$$

$$\sum_{i=1}^{n} f_1(x_i)f_1(x_i)\,a_1 + \sum_{i=1}^{n} f_1(x_i)f_2(x_i)a_2 + \cdots + \sum_{i=1}^{n} f_1(x_i)f_m(x_i)\,a_m = \sum_{i=1}^{n} y_i f_1(x_i)$$

$$\sum_{i=1}^{n} f_2(x_i)f_1(x_i)\,a_1 + \sum_{i=1}^{n} f_2(x_i)f_2(x_i)a_2 + \cdots + \sum_{i=1}^{n} f_2(x_i)f_m(x_i)\,a_m = \sum_{i=1}^{n} y_i f_2(x_i)$$

$$\dots\dots$$

$$\sum_{i=1}^{n} f_m(x_i)f_1(x_i)\,a_1 + \sum_{i=1}^{n} f_m(x_i)f_2(x_i)a_2 + \cdots + \sum_{i=1}^{n} f_m(x_i)f_m(x_i)\,a_m = \sum_{i=1}^{n} y_i f_m(x_i)$$

$$\Rightarrow \boldsymbol{A} \cdot a = \boldsymbol{B}$$

Here,

$$\boldsymbol{A} = \begin{bmatrix} \sum_{i=1}^{n} f_1(x_i)f_1(x_i) & \sum_{i=1}^{n} f_1(x_i)f_2(x_i) & \dots & \sum_{i=1}^{n} f_1(x_i)f_m(x_i) \\ \sum_{i=1}^{n} f_2(x_i)f_1(x_i) & \sum_{i=1}^{n} f_2(x_i)f_2(x_i) & \dots & \sum_{i=1}^{n} f_2(x_i)f_m(x_i) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} f_m(x_i)f_1(x_i) & \sum_{i=1}^{n} f_m(x_i)f_2(x_i) & \dots & \sum_{i=1}^{n} f_m(x_i)f_m(x_i) \end{bmatrix}$$

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} \sum_{i=1}^{n} y_i f_1(x_i) \\ \sum_{i=1}^{n} y_i f_2(x_i) \\ \dots \\ \sum_{i=1}^{n} y_i f_m(x_i) \end{bmatrix}$$

If the matrix is simple, then we can do:

$$a = \boldsymbol{A}^{-1} \cdot \boldsymbol{B}$$

Let's talk about Gauss-seidel method to solve the matrix.

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ \dots \\ B_m \end{bmatrix}$$

$$A_{kl} = \sum_{i=1}^{n} f_k(x_i) f_l(x_i)$$

$$B_k = \sum_{i=1}^{n} y_i f_k(x_i)$$

In Gauss-Seidel method, we write A in the following form:

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{bmatrix} = \begin{bmatrix} 0 & A_{12} & \dots & A_{1m} \\ 0 & 0 & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} A_{11} & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{bmatrix} = U + L$$

$$\boldsymbol{U} = \begin{bmatrix} 0 & A_{12} & \dots & A_{1m} \\ 0 & 0 & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

$$\boldsymbol{L} = \begin{bmatrix} A_{11} & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{bmatrix}$$

$$\boldsymbol{A} \cdot a = \boldsymbol{B}$$

Becomes:

$$(\boldsymbol{U} + \boldsymbol{L}) \cdot a = \boldsymbol{B}$$

$$\boldsymbol{L} \cdot a = \boldsymbol{B} - \boldsymbol{U} \cdot a$$

$$\begin{bmatrix} A_{11} & 0 & \dots & 0 \\ A_{21} & A_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ \dots \\ B_m \end{bmatrix} - \begin{bmatrix} 0 & A_{12} & \dots & A_{1m} \\ 0 & 0 & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix}$$

Then, we have

$$A_{11} a_1 = B_1 - (0a_1 + A_{12} a_2 + \cdots + A_{1m} a_m), (1)$$

$$A_{21} a_1 + A_{22} a_2 \ = B_2 - (0a_1 + 0a_2 + A_{23} a_3 \dots + A_{2m} a_m), (2)$$

$$\dots$$

$$A_{m1} a_1 + A_{m2} a_2 + \cdots + A_{mm} a_m = B_m - (0a_1 + 0a_2 + \cdots + 0a_m), (3)$$

Next is fun: from eq. (1), we get:

$$a_1 = \frac{B_1 - (0a_1 + A_{12} a_2 + \cdots + A_{1m} a_m)}{A_{11}} = \frac{B_1 - \sum_{j=2}^{m} A_{1j} a_j}{A_{11}}$$

From Eq. (2), we get:

$$a_2 = \frac{B_2 - (0a_1 + 0a_2 + A_{23}a_3 \dots + A_{2m}a_m) - A_{21}a_1}{A_{22}} = \frac{B_2 - \sum_{j=3}^{m} A_{2j}a_j - \sum_{j=1}^{1} A_{2j}a_j}{A_{22}}$$

$$a_3 = \frac{B_3 - (0a_1 + 0a_2 + 0a_3 + A_{34}a_4 \dots + A_{3m}a_m) - (A_{31}a_1 + A_{32}a_2)}{A_{33}}$$

$$= \frac{B_3 - \sum_{j=4}^{m} A_{3j}a_j - \sum_{j=1}^{2} A_{3j}a_j}{A_{33}}$$

We can summarize and see the pattern:

$$a_1 = \frac{B_1 - \sum_{j=2}^{m} A_{1j}a_j}{A_{11}}$$

$$a_2 = \frac{B_2 - \sum_{j=3}^{m} A_{2j}a_j - \sum_{j=1}^{1} A_{2j}a_j}{A_{22}}$$

$$a_3 = \frac{B_3 - \sum_{j=4}^{m} A_{3j}a_j - \sum_{j=1}^{2} A_{3j}a_j}{A_{33}}$$

$$\dots$$

$$a_i = \frac{B_i - \sum_{j=i+1}^{m} A_{ij}a_j - \sum_{j=1}^{i-1} A_{ij}a_j}{A_{ii}}$$

For Gauss-Seidel method, we write:

$$a_i^{(k+1)} = \frac{B_i - \sum_{j=i+1}^{m} A_{ij}a_j^{(k)} - \sum_{j=1}^{i-1} A_{ij}a_j^{(k+1)}}{A_{ii}}, for\ 1 \le i \le m$$

$$A_{ij} = \sum_{l=1}^{n} f_i(x_l) f_j(x_l)$$

$$B_i = \sum_{l=1}^{n} y_l f_i(x_l)$$

These are the equations that we use to find the parameters of a_i.