HW 8

ankit gupta

July 2023

1 Fitting a Constant Function Using Least Squares

1.1 a

x_i	y_i	x_iy_i	x_i^2
-1	5/4	-5/4	1
2	4/3	8/3	4
3	5/12	5/4	9

x	y	xy	x^2
4	3	8/3	14

$$n = 3$$

$$C = \frac{y}{n} = \frac{3}{3} = 1$$

1.2 b

$$C = \frac{y}{n} = \frac{1.4 + 1.5 + 1.4}{3} = \frac{43}{30}$$

1.3 c

$$S = \Sigma(yk - C)^2$$

for k=0 to m
Taking the derivative of S with respect to C and setting it to zero (since we want to find the minimum), we get:

$$0 = \Sigma - 2(y_k - C)$$

for k=0 to m

$$0 = \Sigma - 2y_k + 2C$$

for k=0 to m

$$0 = -2\Sigma y_k + 2m * C$$

for k=0 to m
Solving for C, we find:

$$C = \frac{\Sigma y_k}{m}$$

for k=0 to m
yes I am surprised with the formula

2 Fitting Functions With One Parameter with Least Squares

2.1 a

To apply the least squares theory, we minimize the sum of the squares of the residuals, given by:

$$S = \sum_{i=0}^{m} (y_i - (x_i^2 - x_i + c))^2$$

Taking the derivative of S with respect to c and setting it to zero gives us:

$$0 = \sum_{i=0}^{m} -2(y_i - (x_i^2 - x_i + c))$$

This simplifies to:

$$0 = \sum_{i=0}^{m} (x_i^2 - x_i + c - y_i)$$

Solving for c, we find:

$$c = \frac{1}{m} \sum_{i=0}^{m} (y_i - x_i^2 + x_i)$$

This is the value of c that gives the best fit in the least squares sense.
Given the specific data set:

$$m = 3$$

$$\Sigma y = 3$$

$$\Sigma x = 4$$

$$\Sigma x^2 = 14$$

$$c = \frac{1}{m} \sum_{i=0}^{m} (y_i - x_i^2 + x_i)$$

$$c = \frac{3 - 14 + 4}{3} = \frac{-7}{3}$$

2.2 b

$$S = \sum_{i=0}^{m} (y_i - \kappa \log |x_i|)^2$$

Taking the derivative of S with respect to κ and setting it to zero gives us:

$$0 = \sum_{i=0}^{m} -2(y_i - \kappa \log |x_i|) \log |x_i|$$

Rearranging this gives:

$$0 = \sum_{i=0}^{m} (2\kappa \log^2 |x_i| - 2y_i \log |x_i|)$$

This simplifies to:

$$0 = 2\kappa \sum_{i=0}^{m} \log^2 |x_i| - 2 \sum_{i=0}^{m} y_i \log |x_i|$$

Solving for κ, we find:

$$\kappa = \frac{\sum_{i=0}^{m} y_i \log|x_i|}{\sum_{i=0}^{m} \log^2|x_i|}$$

Using matlab to solve this

$$\kappa = 0.8190$$

2.3 c

Given a model of the form $y = \log(\beta|x|)$, it is more straightforward to perform a linear regression on the logarithm of the data, transforming the equation into a linear form:

$$y = \log|x| + \log\beta$$

Similar to

$$y = mx + b$$

The formula for the slope m in simple linear regression is:

$$m = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

After calculating m, we can find the best-fit value of β by taking the exponential of m:

$$\beta = e^m$$

$$m = \frac{-2}{13}$$

$$\beta = 0.857404$$

3 The Method of Least Squares with Polynomial Regression

3.1 a

.

$$m = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$c = \frac{\sum y_i - m\sum x_i}{n}$$

where n is the number of data points, and the sums are over all data points.
Given the specific data set:

$$\Sigma x = 10$$

$$n = 4$$

$$\Sigma x^2 = 30$$

$$\Sigma y = 4$$

$$\Sigma xy = 13$$

$$m = \frac{-6}{5}$$

$$c = 4$$

$$y = \frac{-6x}{5} + 4$$

3.2 b

$$\Sigma x = 0$$
$$n = 3$$
$$\Sigma x^2 = 0.02$$
$$\Sigma y = 4.3$$
$$\Sigma xy = 0$$

$$a = 0$$
$$b = 1.433$$

3.3 c

$$S = \sum (y_i - a x_i^2 - b)^2$$

Taking derivatives and setting them to zero gives:

$$\frac{\partial S}{\partial a} = -2 \sum (y_i - a x_i^2 - b) x_i^2 = 0$$
$$\frac{\partial S}{\partial b} = -2 \sum (y_i - a x_i^2 - b) = 0$$

Rearranging and simplifying gives:

$$a < x^2 > + b = < x^2 y >$$
$$a < x^2 > + bn = < y >$$

where n is the number of data points.
Solving this system for a and b gives:

$$a = \frac{< x^2 y > - < y >}{< x^2 > - 1}$$
$$b = < y > - a < x^2 >$$

Given the specific data set:

$$x_i = \{-1, 0, 1\}$$
$$y_i = \{3.1, 0.9, 2.9\}$$

Did the calculation in notebook. Answer

$$a = 2.1$$
$$b = -0.1$$
$$c = 0.9$$
$$y = 2.1x^2 - 0.1x + 0.9$$

4 The method of least squares with non-polynomial functions

4.1 a

In order to find a function of the form $g(x) = \alpha \sin x + \beta \cos x$ that best fits the data in a least squares sense, we want to minimize the sum of the squared residuals:

$$S = \sum (y_i - g(x_i))^2 = \sum (y_i - \alpha \sin(x_i) - \beta \cos(x_i))^2$$

Setting the derivatives of S with respect to α and β to zero gives us the following two normal equations:

$$\sum (y_i \sin(x_i)) = \alpha \sum (\sin^2(x_i)) + \beta \sum (\sin(x_i) \cos(x_i))$$
$$\sum (y_i \cos(x_i)) = \alpha \sum (\sin(x_i) \cos(x_i)) + \beta \sum (\cos^2(x_i))$$

This system of equations can be solved for α and β.
Given the specific data set we get:

$$\alpha = -1.996$$
$$\beta = -0.042$$

4.2 b

In order to find the constant c that minimizes the sum

$$S = \sum (f(x_k) - ce^{x_k})^2$$

we can differentiate S with respect to c and set the derivative equal to zero. The derivative is

$$\frac{dS}{dc} = \sum -2(f(x_k) - ce^{x_k})e^{x_k} = 0.$$

Rearranging this equation gives us the following:

$$\sum f(x_k)e^{x_k} = c \sum e^{2x_k}.$$

Solving for c gives the constant that minimizes the sum:

$$c = \frac{\sum f(x_k)e^{x_k}}{\sum e^{2x_k}}.$$

The denominator is the sum of the squares of the e^{x_k} values, and the numerator is the sum of the products of the $f(x_k)$ and e^{x_k} values. This solution makes sense intuitively: if the $f(x_k)$ values are larger for larger e^{x_k}, then we would expect c to be larger to minimize the sum of the squared differences.

5 Least Squares for Over-determined Systems

5.1 a

The error function $\psi(x_0, x_1, \ldots, x_n)$ is given by:

$$\psi(x_0, x_1, \ldots, x_n) = \sum_{k=0}^{m} \left(\left(\sum_{j=0}^{n} a_{kj} x_j \right) - b_k \right)^2$$

.

We want to find the set of $(x_0, x_1, \ldots, x_n)$ that minimizes this error. This corresponds to finding the set of x_j that makes the gradient of ψ equal to zero, i.e., the derivative of ψ with respect to each x_i for $i = 0, 1, \ldots, n$ should be zero.

Taking the derivative of ψ with respect to x_i gives:

$$\frac{\partial \psi}{\partial x_i} = 2\sum_{k=0}^{m} a_{ki}\left(\left(\sum_{j=0}^{n} a_{kj}x_j\right) - b_k\right)$$

.

Setting this equal to zero gives the normal equations:

$$\sum_{k=0}^{m} a_{ki}\left(\left(\sum_{j=0}^{n} a_{kj}x_j\right) - b_k\right) = 0$$

.

This is a system of $n+1$ equations (one for each $i = 0, 1, \ldots, n$) for the $n+1$ unknowns $x_0, x_1, \ldots, x_n$.

5.2 b

The given system of equations is an over-determined system, i.e., it has more equations than unknowns. The system can be written in the matrix form as follows:

$$Ax = b,$$

where A is the matrix of coefficients, x is the column vector of variables (x and y), and b is the column vector of constants on the right side of the equations.

$$A = \begin{bmatrix} 2 & 3 \\ 1 & -4 \\ 2 & -1 \end{bmatrix}, x = \begin{bmatrix} x \\ y \end{bmatrix}, b = \begin{bmatrix} 1 \\ -9 \\ -1 \end{bmatrix}.$$

The normal equation for the least squares solution is $(A^T A)x = A^T b$, where A^T is the transpose of the matrix A.

$$A^T = \begin{bmatrix} 2 & 1 & 2 \\ 3 & -4 & -1 \end{bmatrix},$$

$$A^T A = \begin{bmatrix} 2*2+1*1+2*2 & 2*3+1*(-4)+2*(-1) \\ 3*2+(-4)*1+(-1)*2 & 3*3+(-4)*(-4)+(-1)*(-1) \end{bmatrix} = \begin{bmatrix} 9 & 2 \\ 2 & 26 \end{bmatrix},$$

$$A^T b = \begin{bmatrix} 2*1+1*(-9)+2*(-1) \\ 3*1+(-4)*(-9)+(-1)*(-1) \end{bmatrix} = \begin{bmatrix} -7 \\ 16 \end{bmatrix}.$$

So, the normal equation is:

$$\begin{bmatrix} 9 & 2 \\ 2 & 26 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -7 \\ 16 \end{bmatrix}.$$

By solving this 2x2 system of linear equations, we get

$$x = -0.930435$$

$$y = 0.686957$$

5.3 c

the system can be expressed in the form $Ax = b$, where A is a $m \times n$ matrix ($m > n$), x is a $n \times 1$ vector of unknowns, and b is a $m \times 1$ vector.

To find the least squares solution,

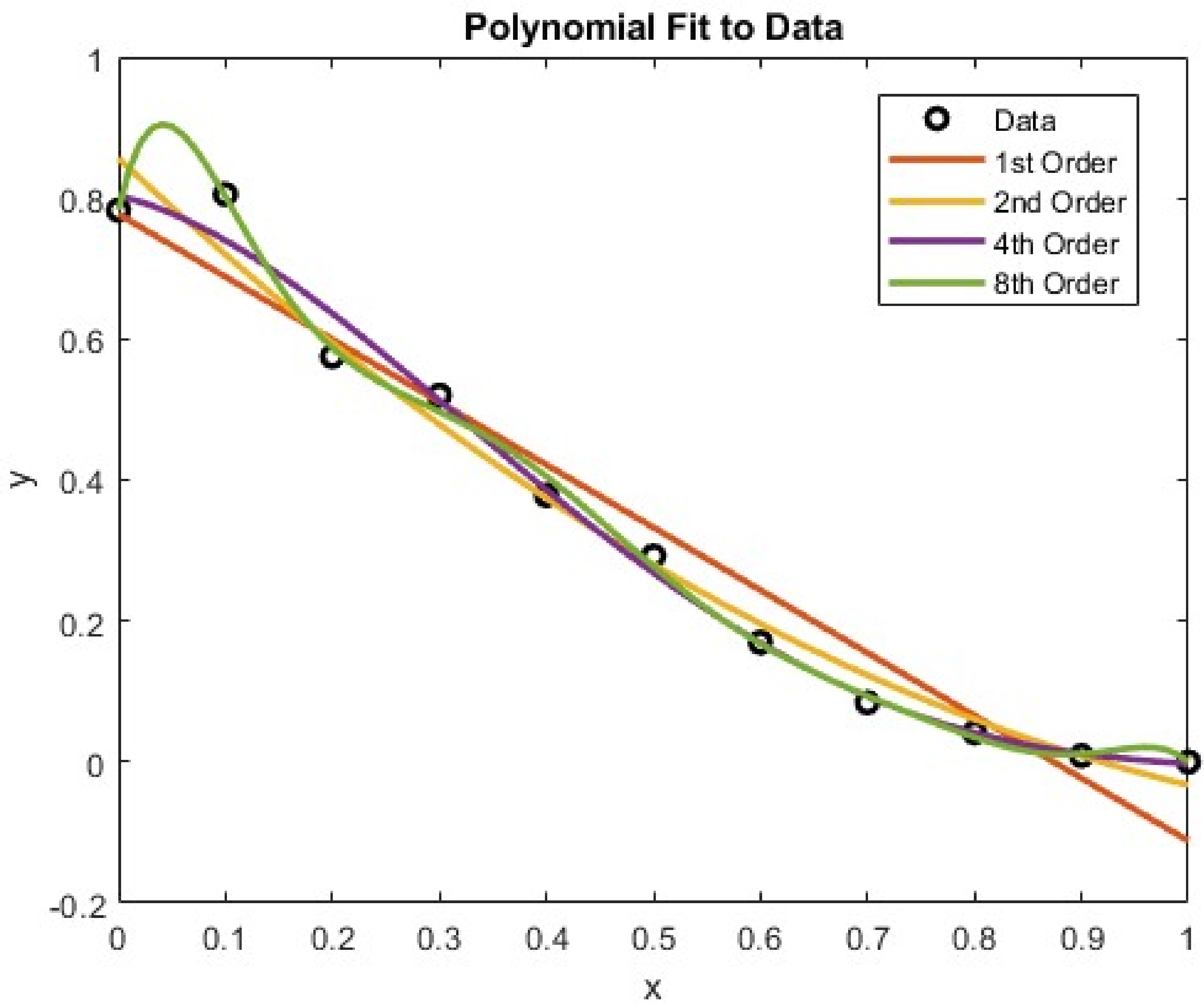
$$A^T(Ax) = A^T b$$

Since the product of a scalar and a matrix is commutative, $A^T(Ax)$ simplifies to $(A^T A)x$:

$$(A^T A)x = A^T b$$

.

6 The method of least squares in Matlab



The resulting plot shows how well each polynomial fits the data. As the order of the polynomial increases, the fit generally becomes better, but it's important to be cautious about overfitting. Overfitting can occur when a high-order polynomial is used to model a data set with a small number of points, resulting in a model that fits the data very well but may not accurately represent the underlying relationship between the variables.

7 Least squares approximation of functions

We want to approximate the function $f(x)$ defined on the interval $[-1, 1]$ by a function of the form $g(x) = a\cos(\pi x) + b\sin(\pi x)$. To find the best possible constants a and b using the method of least squares, we minimize the integral:

$$\int_{-1}^{1} [f(x) - g(x)]^2 \, dx$$

To find the values of a and b that minimize this integral, we take the derivatives with respect to a and b and set them equal to zero:

$$\frac{\partial}{\partial a} \int_{-1}^{1} [f(x) - a\cos(\pi x) - b\sin(\pi x)]^2 \, dx = 0$$

$$\frac{\partial}{\partial b} \int_{-1}^{1} [f(x) - a\cos(\pi x) - b\sin(\pi x)]^2 \, dx = 0$$

Solving these equations will give us the best possible constants a and b. Using Matlab to solve this

$$a = 4.0716e - 05$$

$$b = 1.2731$$