Polynomials of Best Fit

Last Lecture

- We derived the method of least squares by looking at an example of a scientist trying to find an m^{th}-degree polynomial which best fits a set of experimental data.

In This Lecture

- We will look at some more examples after proving under which condition we get a unique polynomial of best fit.

Polynomials of Best Fit

Theorem 9.6.2

Let n data points $(x_1, y_1), \dots, (x_n, y_n)$ be given, let

$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix}$$

If $\vec{a} = \begin{bmatrix} a_0 \\ \vdots \\ a_m \end{bmatrix}$ is any solution to the normal system

$$X^T X \vec{a} = X^T \vec{y}$$

then the polynomial

$$p(x) = a_0 + a_1 x + \cdots + a_m x^m$$

is a best fitting polynomial of degree m for the given data. Moreover, if at least $m + 1$ of the numbers $x_1, \dots, x_n$ are distinct, then the matrix $X^T X$ is invertible and hence $\vec{a}$ is unique with

$$\vec{a} = (X^T X)^{-1} X^T \vec{y}$$

Polynomials of Best Fit

Proof of Theorem 9.6.2

Putting the data points into $p(x)$ we get the system of n equations in $m+1$ unknowns

$$\begin{aligned} a_0 + a_1 x_1 + \cdots + a_m x_1^m &= y_1 \\ &\vdots \\ a_0 + a_1 x_n + \cdots + a_m x_n^m &= y_n \end{aligned}$$

This is the system $X\vec{a} = \vec{y}$. Thus, from our work last class, we get that any solution $\vec{a}$ of the normal system $X^T X\vec{a} = X^T\vec{y}$ gives the coefficients of a polynomial of best fit.

Assume that at least $m+1$ of the numbers $x_1, \ldots, x_n$ are distinct.
Consider a linear combination of the columns of X.

$$\begin{aligned} \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} &= c_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + c_1 \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \cdots + c_m \begin{bmatrix} x_1^m \\ \vdots \\ x_n^m \end{bmatrix} \\ &= \begin{bmatrix} c_0 + c_1 x_1 + \cdots + c_m x_1^m \\ \vdots \\ c_0 + c_1 x_n + \cdots + c_m x_n^m \end{bmatrix} \end{aligned}$$

This implies that $x_1, \ldots, x_n$ are roots of the polynomial $q(x) = c_0 + c_1 x + \cdots + c_m x^m$. Therefore, $q(x)$ has at least $m+1$ distinct roots. But, by the Fundamental Theorem of Algebra, the only m-th degree polynomial with greater than m roots is the zero polynomial. Hence,

$$0 = q(x) = c_0 + c_1 x + \cdots + c_m x^m$$

Therefore, $c_0 = \cdots = c_m = 0$. Consequently, the columns of X are linearly independent.

Polynomials of Best Fit

Proof of Theorem 9.6.2

Let $\vec{v}$ be such that $X^T X\vec{v} = \vec{0}$. We wish to show that $\vec{v} = \vec{0}$.
We have

$$\|X\vec{v}\|^2 = (X\vec{v}) \cdot (X\vec{v}) = (X\vec{v})^T (X\vec{v}) = \vec{v}^T X^T X\vec{v} = \vec{v}^T \vec{0} = 0$$

Thus, $X\vec{v} = \vec{0}$. Since the columns of X are linearly independent, $\vec{v} = \vec{0}$ as required.
Therefore, $X^T X$ is invertible and so we get

$$\vec{a} = (X^T X)^{-1} X^T \vec{y}$$

□

Note: The matrix X is called the design matrix. It is derived from the form of the desired polynomial.

Polynomials of Best Fit

Example

Find a, b, c to obtain the best fitting equation of the form $y = a + bx + cx^2$ for the given data.

x	-3	-1	0	1	3
y	3	1	1	2	4

Solution

We have $\vec{y} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ 2 \\ 4 \end{bmatrix}$ and $X = \begin{bmatrix} 1 & -3 & 9 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \end{bmatrix}$. Therefore, by Theorem 9.6.2

$$\vec{a} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (X^T X)^{-1} X^T \vec{y} = \begin{bmatrix} 121/105 \\ 1/5 \\ 11/42 \end{bmatrix}$$

So, the best fit parabola is $p(x) = \frac{121}{105} + \frac{1}{5}x + \frac{11}{42}x^2$.

Polynomials of Best Fit

Example

Find the best fitting polynomial of the form $p(x) = ax^2 + bx$ for the following data:

x	-1	0	1
y	4	1	1

Solution

We observe that we want to pick X to be of the form

$$X = \begin{bmatrix} x_1^2 & x_1 \\ x_2^2 & x_2 \\ x_3^2 & x_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

We let $\vec{y} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$, and then get

$$\begin{bmatrix} a \\ b \end{bmatrix} = (X^T X)^{-1} X^T \vec{y} = \begin{bmatrix} 5/2 \\ -3/2 \end{bmatrix}$$

Therefore, $p(x) = \frac{5}{2}x^2 - \frac{3}{2}x$.