Curve Fitting

CE 311 K - Introduction to Computer Methods

Daene C. McKinney

Polynomial Regression

- Minimize the residual between the data points and the curve -- least-squares regression
  - Linear $y_i = a_0 + a_1 x_i$
  - Quadratic $y_i = a_0 + a_1 x_i + a_2 x_i^2$
  - Cubic $y_i = a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3$
  - General $y_i = a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + \cdots + a_m x_i^m$

Must estimate values of a_0 , a_1, a_2, … a_m

Polynomial Regression

- Residual

$$e_i = y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + \cdots + a_m x_i^m)$$

- Sum of squared residuals

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} [y - (a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_m x^m)]^2$$

- Minimize by taking derivatives

Polynomial Regression

- Normal Equations

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \cdots & \sum_{i=1}^{n} x_i^m \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \cdots & \sum_{i=1}^{n} x_i^{m+1} \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 & \cdots & \sum_{i=1}^{n} x_i^{m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & \sum_{i=1}^{n} x_i^{m+2} & \cdots & \sum_{i=1}^{n} x_i^{2m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i^2 y_i \\ \vdots \\ \sum_{i=1}^{n} x_i^m y_i \end{bmatrix}$$

Quadratic Functional Form

- Functional Form $y_i = a_0 + a_1 x_i + a_2 x_i^2$
- Normal Equations

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i^2 y_i \end{bmatrix}$$

Must estimate values of a_0 , a_1, a_2

Cubic Functional Form

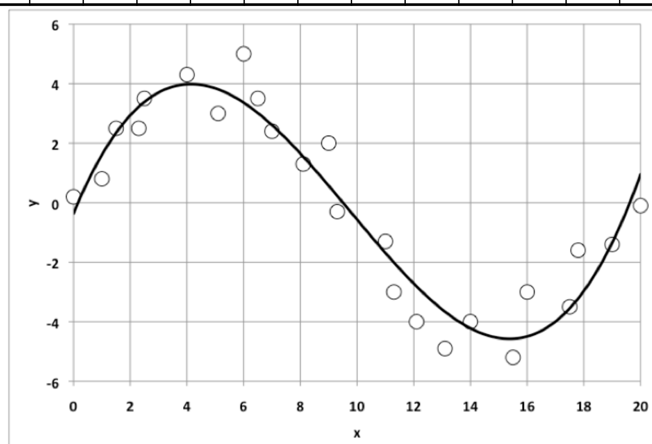
- Functional Form $y_i = a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3$
- Normal Equations

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 & \sum_{i=1}^{n} x_i^5 \\ \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 & \sum_{i=1}^{n} x_i^5 & \sum_{i=1}^{n} x_i^6 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i^2 y_i \\ \sum_{i=1}^{n} x_i^3 y_i \end{bmatrix}$$

Must estimate values of a_0 , a_1, a_2, a_3

Example – Cubic Functional Form

x	0	1.0	1.5	2.3	2.5	4.0	5.1	6.0	6.5	7.0	8.1	9.0
y	0.2	0.8	2.5	2.5	3.5	4.3	3.0	5.0	3.5	2.4	1.3	2.0
x	9.3	11.0	11.3	12.1	13.1	14.0	15.5	16.0	17.5	17.8	19.0	20.0
y	-0.3	-1.3	-3.0	-4.0	-4.9	-4.0	-5.2	-3.0	-3.5	-1.6	-1.4	-0.1



Example

- Need $\sum_{i=1}^{n} x_i \quad \sum_{i=1}^{n} x_i^2 \quad \sum_{i=1}^{n} x_i^3 \quad \sum_{i=1}^{n} x_i^4 \quad \sum_{i=1}^{n} x_i^5 \quad \sum_{i=1}^{n} x_i^6$

  $\sum_{i=1}^{n} y_i \quad \sum_{i=1}^{n} x_i y_i \quad \sum_{i=1}^{n} x_i^2 y_i \quad \sum_{i=1}^{n} x_i^3 y_i$

- Normal Equations

$$\begin{bmatrix} 24 & 229.6 & 3060.2 & 46342.8 \\ 229.6 & 3060.2 & 46342.8 & 752835.2 \\ 3060.2 & 46342.8 & 752835.2 & 12780147.7 \\ 46342.8 & 752835.2 & 12780147.7 & 223518116.8 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -1.30 \\ -316.9 \\ -6037.2 \\ -9943.36 \end{bmatrix}$$

- Solution

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -0.3593 \\ 2.3051 \\ -0.3532 \\ 0.0121 \end{bmatrix}$$

Example

- Regression Equation

$y = -0.359 + 2.305x - 0.353x^2 + 0.012x^3$