2. Polynomial Regression

Suppose that we observe ten points $(X_1, Y_1), \ldots, (X_{10}, Y_{10})$ where $X_1 = 1, X_2 = 2, \ldots, X_{10} = 10$. We believe that the data is governed by a polynomial relationship:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i$$

where ϵ_i are i.i.d. $\mathcal{N}(0, \sigma^2)$, and $\sigma^2 = 0.1$.

(a)

1.0/1 point (graded)

Treat the expression for Y_i on the right hand side as a linear function of 1, X_i and X_i^2, plus the noise variable ϵ.

What is the design matrix $\mathbb{X}$? Recall that the desired setup for linear regression in this course is $\mathbf{Y} = \mathbb{X}\beta + \epsilon$, where both β and ϵ are column vectors, so carefully consider what the size of the matrix $\mathbb{X}$ is.

(Enter your answer as a matrix. For instance, to enter the 3×2 matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, type **[[1,2],[3,4],[5,6]]**. Your answer may be a large matrix.)

$\mathbb{X} =$ [[1,1,1],[1,2,4],[1,3,9],[1,4,16],[1,5,25],[1,6,36],[1,7,49],[1,8,64],[1,9,81],[1,10,100]]

Answer: [[1,1,1],[1,2,4],[1,3,9],[1,4,16],[1,5,25],[1,6,36],[1,7,49],[1,8,64],[1,9,81],[1,10,100]]

STANDARD NOTATION

Solution:

By the conventions specified in the problem (setup of linear regression) the ith row of $\mathbb{X}$ are the entries 1, X_i and X_i^2. To verify, notice that this yields a 10×3 matrix, which is compatible with multiplication by β, a 3×1 matrix, from the right.

Submit You have used 1 of 4 attempts

Answers are displayed within the problem

(b)

1.0/1 point (graded)

Calculate the matrix $\mathbb{X}^T\mathbb{X}$. Since the values of each X_i happen to be integers, your answer should also have integer entries.

(Enter your answer as a matrix. For instance, to enter the 3×2 matrix $\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, type **[[1,2],[3,4],[5,6]]**.)

$\mathbb{X}^T\mathbb{X} =$ [[10,55,385],[55,385,3025],[385,3025,25333]]

Answer: [[10,55,385],[55,385,3025],[385,3025,25333]]

STANDARD NOTATION

Solution:

One ought to directly plug in the answer from the previous part into a calculator or compute by hand. In either case, one ends up with the 3×3 matrix

$$\mathbb{X}^T\mathbb{X} = \begin{pmatrix} 10 & 55 & 385 \\ 55 & 385 & 3025 \\ 385 & 3025 & 25333 \end{pmatrix}$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

(c)

3/3 points (graded)

Calculate the least squares estimator $\hat{\beta}$ for $\beta = (\beta_0, \beta_1, \beta_2)$ given the data:

x	1	2	3	4	5	6	7	8	9	10
y	1	3	5	8	11	14	18	21	25	28

Round each entry of your final answer to the nearest 0.01.

$\hat{\beta}_0 =$ -1.433333 ✔ **Answer:** -1.43

$\hat{\beta}_1 =$ 2.007576 ✔ **Answer:** 2.01

$\hat{\beta}_2 =$ 0.098485 ✔ **Answer:** 0.10

Solution:

By plugging directly into the formula $\left(\mathbb{X}^T\mathbb{X}\right)^{-1}\mathbb{X}^T\mathbf{Y}$, one obtains the approximate answer

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} -1.43 \\ 2.01 \\ 0.10 \end{pmatrix}$$

Submit You have used 2 of 3 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 6 Linear Regression:Homework 10 Linear regression / 2. Polynomial Regression