Online-Class 24-3-2021

Probability, Statistics and Reliability (MAT3003)

SLOT: B21 + B22 + B23

MODULE - 3

Topic: Nonlinear Regression: Fitting of a Quadratic Polynomial

Reference: Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

Nonlinear Regression: Fitting of a **Quadratic Polynomial**

- Recall the case of a Linear Regression
- Nonlinear Regression – Quadratic Case
- Fitting of a Quadratic Polynomial by the method of least squares
- Questions on Fitting of a Quadratic Function
- Practice Problems

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

➢ Recall the case of a Linear Regression

Straight Line: $\boldsymbol{y = a + bx}$

Normal Eqns: $\sum y = an + b\sum x \qquad (1)$

$\sum xy = a\sum x + b\sum x^2 \qquad (2)$

Straight Line: $\boldsymbol{y = ax + b}$

Normal Eqns: $\sum y = a\sum x + bn \qquad (3)$

$\sum xy = a\sum x^2 + b\sum x \qquad (4)$

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

Nonlinear Regression: Quadratic/Parabola/Second Degree Polynomial

Let the eq. of second degree polynomial be:

$$y = a + bx + cx^2 \qquad (1)$$

By the method of least square, the normal equations are:

(By taking Σ of both sides)

$$\sum y = an + b\sum x + c\sum x^2 \qquad (2)$$

(By multiplying both sides by x, and then taking Σ of both sides)

$$\sum xy = a\sum x + b\sum x^2 + c\sum x^3 \qquad (3)$$

(By multiplying both sides by x^2, and then taking Σ of both sides)

$$\sum x^2y = a\sum x^2 + b\sum x^3 + c\sum x^4 \qquad (4)$$

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

To fit the parabola: $y=a+bx+cx^2$:

- Form the normal equations
- $\sum y=na+b\sum x+c\sum x^2$, $\sum xy=a\sum x+b\sum x^2+c\sum x^3$ and $\sum x^2y=a\sum x^2+b\sum x^3+c\sum x^4$.

- Solve these as simultaneous equations for a,b,c.

- Substitute the values of a,b,c in $y=a+bx+cx^2$, which is the required parabola of the best fit.

Question 1

The following are data on the drying time of a certain varnish and the amount of an additive that is intended to reduce the drying time.

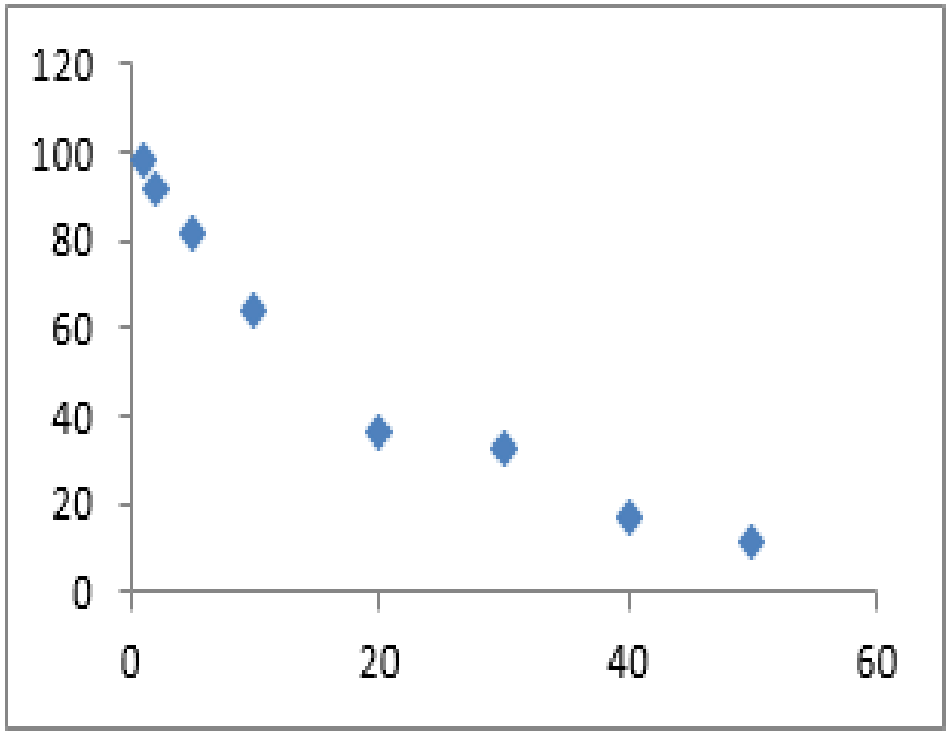
Amount of Varnish additive (grams) x	0	1	2	3	4	5	6	7	8
Drying time (hours) y	12.0	10.5	10.0	8.0	7.0	8.0	7.5	8.5	9.0

a)Draw a scatter plot to verify that it is reasonable to assume that the relationship is parabolic

b) Fit a second degree polynomial by the method of least squares

c) Use the result of (b) to predict the drying time of the varnish when 6.5 grams of the additive is used

Solution: a) It can be seen from the following figure overall pattern suggests fitting a second degree polynomial having one relative minimum.



(b) Let the eq. of second degree polynomial be:

$$y = a + bx + cx^2 \qquad (1)$$

By the method of least square, the normal equations are:

(By taking Σ of both sides)

$$\sum y = an + b\sum x + c\sum x^2 \qquad (2)$$

(By multiplying both sides by x, and then taking Σ of both sides)

$$\sum xy = a\sum x + b\sum x^2 + c\sum x^3 \qquad (3)$$

(By multiplying both sides by x^2, and then taking Σ of both sides)

$$\sum x^2y = a\sum x^2 + b\sum x^3 + c\sum x^4 \qquad (4)$$

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

x	y	xy	x^2	x^3	x^4	x^2y
0	12					
1	10.5					
2	10					
3	8					
$\sum x=$	$\sum y=$	$\sum xy=$	$\sum x^2=$	$\sum x^3=$	$\sum x^4=$	$\sum x^2y=$

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

From the Table,

$$\sum x = 36, \qquad \sum x^2 = 204,$$

$$\sum x^3 = 1296, \sum x^4 = 8772, \quad \sum y = 80.5,$$

$$\sum xy = 299, \qquad \sum x^2 y = 1697,$$

Normal equations are:

$$80.5=9a+36b+204c$$

$$299=36a+204b+1296c$$

$$1697=204a+1296b+8772c$$

Solving, a = 12.2, b = - 1.85 and c = 0.183, we find that the equation of the least squares polynomial is

$$y = 12.2 - 1.85x + 0.183x^2$$

c) Substituting x=6.5 into this equation, we get

$$y = 12.2 - 1.85(6.5) + 0.183(6.5)^2 = 7.9.$$

That is, predicted drying time or 7.9 hours.

Question 2 (For Students)

An experiment was conducted on a new model of a particular make of an automobile to determine the stopping distance at various speeds. The following data were recorded.

Speed, x(km/h)	:	35	50	65	80	95	110
Stopping distance y	:	16	26	41	62	88	119

a) Fit a regression curve of the form $y=a+bx+cx^2$.

b) Estimate the stopping distance when the car is travelling at 70 km/h.

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

Solution:

The equation is $y = a + bx + cx^2$ and the normal equations are

$$\sum y = an + b\sum x + c\sum x^2$$

$$\sum xy = a\sum x + b\sum x^2 + c\sum x^3$$

$$\sum x^2y = a\sum x^2 + b\sum x^3 + c\sum x^4$$

The values are calculated using the following table

x	y	x^2	x^3	x^4	$x \cdot y$	$x^2 \cdot y$
35	16	1225	42875	1500625	560	19600
50	26	2500	125000	6250000	1300	65000
65	41	4225	274625	17850625	2665	173225
80	62	6400	512000	40960000	4960	396800
95	88	9025	857375	81450625	8360	794200
110	119	12100	1331000	146410000	13090	1439900
---	---	---	---	---	---	---
$\sum x = 435$	$\sum y = 352$	$\sum x^2 = 35475$	$\sum x^3 = 3142875$	$\sum x^4 = 294421875$	$\sum x \cdot y = 30935$	$\sum x^2 \cdot y = 2888725$

Reference: Probability and Statistics for Engineers & Scientists
by R.E.Walpole, Pearson (2012).

Substituting these values in the normal equations

$6a + 435b + 35475c = 352$

$435a + 35475b + 3142875c = 30935$

$35475a + 3142875b + 294421875c = 2888725$

Solution using Elimination method.

$a = 13.3587, b = -0.3394, c = 0.0118$

Now substituting this values in the equation is $y = a + bx + cx^2$, we get

$y = 13.3587 - 0.3394x + 0.0118x^2$

(b) Substituting x=70, we have

$y = 13.35875 - 0.33944(70) + 0.01183\ (70)^2$

$= 47.5420$

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

Question -3 (For Students)

The following data pertain to the number of computer jobs per day and the central processing unit (CPU) time required

Number of jobs x :	1	2	3	4	5
CPU time y :	2	5	4	9	10

a) Fit a straight line to the given data by the method of least squares

b) Use the equation of the least squares line to estimate the mean CPU time at x=3.5.

Reference: Text Book: 1.Probability and Statistics for Engineers & Scientists by R.E.Walpole, Pearson (2012).

Practice Questions

THANK YOU

Reference: Probability and Statistics for Engineers & Scientists
by R.E.Walpole, Pearson (2012).