เฉลยแบบฝึกหัดบทที่ 5

5.1 จากตัวอย่างที่ 5.4 จง fit cubic และ quartic model กับ Rainfall data เพื่อหาค่าที่แสดงไว้ในตารางที่ 5.12 ซึ่งยังไม่ได้แสดงการหาไว้

Fit cubic model: $Y = X\gamma + e$, $p = 4$

$$X'X = \begin{bmatrix} 12 & 0 & 0 & 0 \\ 0 & 572 & 0 & 0 \\ 0 & 0 & 12012 & 0 \\ 0 & 0 & 0 & 5148 \end{bmatrix}, \quad X'Y = \begin{bmatrix} 387.9 \\ -202.3 \\ -1117.5 \\ 445.1 \end{bmatrix}$$

$$\begin{aligned} R(\gamma) &= \hat{\gamma}'X'Y = [(X'X)^{-1}X'Y]X'Y \\ &= R(\delta) + (445.1)^2/5148 \\ &= 12714.383 + 38.484 \\ &= 12752.867 \end{aligned}$$

$R(\gamma_3 | \gamma_0, \gamma_1, \gamma_2) = R(\gamma_3) = 38.484$

SSE(cubic) $= 12815.99 - 12752.867 = 63.123$,

df. $= 12 - 4 = 8$

Fit quartic model: $Y = X\lambda + e$, $p = 5$

$$X'X = \begin{bmatrix} 12 & 0 & 0 & 0 & 0 \\ 0 & 572 & 0 & 0 & 0 \\ 0 & 0 & 12012 & 0 & 0 \\ 0 & 0 & 0 & 5148 & 0 \\ 0 & 0 & 0 & 0 & 8008 \end{bmatrix}, \quad X'Y = \begin{bmatrix} 387.9 \\ -202.3 \\ -1117.5 \\ 445.1 \\ 525.1 \end{bmatrix}$$

$$R(\lambda) = \hat{\lambda}'X'Y = [(X'X)^{-1}X'Y]X'Y$$
$$= R(\gamma) + (525.1)^2/8008$$
$$= 12752.867 + 34.432$$
$$= 12787.299$$

$R(\lambda_4 | \lambda_0, \lambda_1, \lambda_2, \lambda_3) = R(\lambda_4) = 34.432$

SSE(quartic) = 12815.99 - 12787.299 = 28.691,

df. = 12 - 5 = 7

5.2 จงแสดงการหาค่า $\hat{\gamma}$ ในตัวอย่างที่ 5.3

ในตัวอย่างที่ 5.3 $\hat{\gamma}' = [-39.2812 \quad 11.888 \quad -0.78]$

$\hat{\delta}' = [3 \quad 0.9 \quad -0.07]$

จากการทำให้ X'X เป็นเมตริกซ์ทแยงมุม เรากำหนด $d_0 = 0$, $f_0 = -f_2(n+1)(n-1)/12$ และ $f_1 = 0$ เราจะหา d_1 และ f_2 จากสมการ (6)-(7) ในหัวข้อ 5.3.3 และค่า P_1 และ P_2

$$P_1[i-(n+1)/2] = d_1[i-(n+1)/2]$$
$$P_2[i-(n+1)/2] = f_2[-(n+1)(n-1)/12 + [i-(n+1)/2]^2]$$

$n = 5,\ i = 1,\ P_1[1-(n+1)/2] = -2 = d_1[1 - (5+1)/2]$

$-2 = d_1(-2)$

$d_1 = 1$

$= 5,\ i = 1,\ P_2[1-(n+1)/2] = 2 = f_2[-6(4)/12 + (-2)^2] = 2f_2$

$f_2 = 1$

$f_0 = -f_2(n+1)(n-1)/12 = -2$

$d_0 = 0,\ d_1 = 1,\ f_0 = -2,\ f_1 = 0\ ,\ f_2 = 1$

$$\delta_o+[d_o - d_1(n+1)/2]\delta_1 + \{f_o - f_1(n+1)/2 + f_2[(n+1)/2]^2\}\delta_2 = \gamma_o + a\gamma_1 + a^2\gamma_2 \qquad (1)$$

$$d_1\delta_1 + [f_1 - (n+1)f_2]\delta_2 = \gamma_1 h + 2ah\gamma_2 \qquad (2)$$

$$f_2\delta_2 = \gamma_2 h^2 \qquad (3)$$

ในที่นี้ $a = 4.8$, $h = 0.3$

จาก (3) $\quad 1(-0.07) = \hat{\gamma}_2(0.3)^2$

$$\hat{\gamma}_2 = -0.07/0.09 = -0.78$$

จาก (2)

$$\hat{\gamma}_1 = (1/h)\{d_1\hat{\delta}_1 + [f_1 - (n+1)f_2]\hat{\delta}_2 - 2ah\hat{\gamma}_2\}$$
$$= (1/0.3)[1(0.9)-6(-0.07)-2(4.8)(0.3)(-0.78)]$$
$$= 11.888$$

จาก (3) $\hat{\gamma}_o = \hat{\delta}_o - 3\hat{\delta}_1 + 7\hat{\delta}_2 - 4.8\hat{\gamma}_1 - (4.8)^2\hat{\gamma}_2$

$$= 3 - 3(0.9) + 7(-0.07) - 4.8(11.888) - (4.8)^2(-0.78)$$
$$= -39.2812$$

$$\hat{\gamma}' = [-39.2812 \quad 11.888 \quad -0.78]$$

5.3 เราทราบว่า cubic model นั้น fit กับข้อมูลต่อไปนี้

จงหา 95 % ช่วงความเชื่อมั่นของสัมประสิทธิ์แต่ละตัว คือ γ_o, γ_1, γ_2 และ γ_3

y	0.5	3.0	1.6	1.3	0.2	0.5	-.1	1.2
x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5

Cubic model: $y_i = \gamma_o + \gamma_1 x_i + \gamma_2 x_i^2 + \gamma_3 x_i^3 + e_i$, $i = 1,\ldots,8$

หรือ $\quad Y = X\gamma + e$

โดยที่ $Y' = [0.5 \quad 3.0 \ \ldots \ 1.2]$, $\gamma' = [\gamma_o \quad \gamma_1 \quad \gamma_2 \quad \gamma_3]$

$$X = \begin{bmatrix} 1 & 1.0 & 1 & 1 \\ 1 & 1.5 & 2.25 & 3.375 \\ 1 & 2.0 & 4 & 8 \\ & & \cdots & \\ 1 & 4.5 & 20.25 & 91.125 \end{bmatrix}$$

$$X'X = \begin{bmatrix} 8 & 22 & 71 & 253 \\ 22 & 71 & 253 & 958.25 \\ 71 & 253 & 958.25 & 3775.75 \\ 253 & 958.25 & 3775.75 & 15287.56 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 8.2 \\ 18.8 \\ 52.4 \\ 173.525 \end{bmatrix}$$

<u>ขั้นตอนการทำ</u> (ให้นักศึกษาทำเอง)

จาก Normal equations: $X'X\hat{\beta} = X'Y$

เราจะหา $\hat{\beta}$ และ $(X'X)^{-1}$ โดยวิธีดูลิตเติลอย่างย่อ

แล้วหา $R(\beta) = \hat{\beta}'X'Y$

ในการหา Confidence limits ของ β_0, β_1, β_2 และ β_3 เราต้องหา $\hat{\sigma}^2$ ก่อน

โดยที่ $\hat{\sigma}^2 = [Y'Y - R(\beta)]/(n-p)$, $n = 8$, $p = 4$

95% C.I. ของ β_i คือ $\hat{\beta}_i + t_{4,.025}\,(c_{ii}\,\hat{\sigma}^2)^{1/2}$,

c_{ii} คือสมาชิกบน main diagonal ของ $C = (X'X)^{-1}$

ถ้าใช้ Orthogonal Polynomial เพื่อหา $\hat{\gamma}$

แบบจำลอง: $Y = X\delta + e$

โดยที่

$$X = \begin{bmatrix} 1 & -7 & 7 & -7 \\ 1 & -5 & 1 & 5 \\ 1 & -3 & -3 & 7 \\ & \cdots & & \\ 1 & 7 & 7 & 7 \end{bmatrix}$$

$$X'X = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 168 & 0 & 0 \\ 0 & 0 & 168 & 0 \\ 0 & 0 & 0 & 264 \end{bmatrix}, \quad X'Y = \begin{bmatrix} 8.2 \\ -15 \\ 1 \\ 31.4 \end{bmatrix}$$

$$\hat{\delta}' = [\hat{\delta}_0 \quad \hat{\delta}_1 \quad \hat{\delta}_2 \quad \hat{\delta}_3]$$

$$= [8.2/8 \quad -15/168 \quad 1/168 \quad 31.4/264]$$

$$= [1.025 \quad -0.0892857 \quad 0.0059523 \quad 0.1189393]$$

แล้วหา $\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3$

<u>หมายเหตุ</u> ข้อ 5.3 นี้ ควรใช้วิธีดูลิตเติลอย่างย่อกับแบบจำลองเริ่มต้น เพราะเราต้องการ $\hat{\gamma}$, $(X'X)^{-1}$ และ $\hat{\sigma}^2$

5.4 ในการทดลองเกี่ยวกับการสึกกร่อนของโลหะ ได้ข้อมูลดังต่อไปนี้ (กำหนด α =.05)

ความดันไฟฟ้าที่ใช้	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
เปอร์เซ็นต์การสึกกร่อน	1.10	1.43	2.11	3.12	2.50	2.21	2.50	5.90

(a) จงหา degree ของโพลิโนเมียลที่เหมาะสมกับข้อมูล

(b) จงประมาณค่าของสัมประสิทธิ์ของโพลิโนเมียลในข้อ (a) โดยการใช้วิธี Abbreviated Doolittle (วิธีที่อธิบายในบทที่ 4)

(a) ใช้ Orthogonal Polynomial a = 1.0, h = 0.5 , n = 8

y	P_1	P_2	P_3	P_4	yP_1	yP_2	yP_3	yP_4	P_5	yP_5
1.10	-7	7	-7	7	-7.7	7.7	-7.7	7.7	-7	-7.7
1.43	-5	1	5	-13	-7.15	1.43	7.15	-18.59	23	32.89
2.11	-3	-3	7	-3	-6.33	-6.33	14.77	-6.33	-17	-35.87
3.12	-1	-5	3	9	-3.12	-15.6	9.36	28.08	-15	-46.8
2.50	1	-5	-3	9	2.5	-12.5	-7.5	22.5	15	37.5
2.21	3	-3	-7	-3	6.63	-6.63	-15.47	-6.63	17	37.57
2.50	5	1	-5	-13	12.5	2.5	-12.5	-32.5	-23	-57.5
5.9	7	7	7	7	41.3	41.3	41.3	41.3	7	41.3
ΣP^2	168	168	264	616	38.63	11.87	29.41	35.53	2184	1.39

$\Sigma y = 20.87$, $\Sigma y^2 = 69.6355$

ขั้นตอนการทำนั้นให้ปฏิบัติตามตัวอย่างที่ 5.4 คือเริ่ม fit linear, quadratic, cubic, quartic,... ในแต่ละครั้งของการ fit ให้ทดสอบว่าส.ป.ส. ของเทอมกำลังสูงสุดมีนัยสำคัญหรือไม่ แต่เนื่องจากการใช้ Orthogonal Polynomial ถ้าเรา fit polynomial degree สูง เราอาจหาสิ่งที่ต้องการของแบบจำลอง degree ต่ำกว่าได้

พิจารณาจาก Quartic model: $Y = X\lambda + e$, $p = 5$

$$\lambda' = [\lambda_0 \quad \lambda_1 \quad \lambda_2 \quad \lambda_3 \quad \lambda_4]$$

$$X'X = \begin{bmatrix} 8 & 0 & 0 & 0 & 0 \\ 0 & 168 & 0 & 0 & 0 \\ 0 & 0 & 168 & 0 & 0 \\ 0 & 0 & 0 & 264 & 0 \\ 0 & 0 & 0 & 0 & 616 \end{bmatrix}, \quad X'Y = \begin{bmatrix} 20.87 \\ 38.63 \\ 11.87 \\ 29.41 \\ 35.53 \end{bmatrix}$$

$$\begin{aligned} R(\lambda) &= \hat{\lambda}'X'Y = [(X'X)^{-1}X'Y]X'Y \\ &= (20.87)^2/8 + (38.63)^2/168 + (11.87)^2/168 \\ &\quad + (29.41)^2/264 + (35.53)^2/616 \\ &= 54.444612+8.8826005+0.838672+3.2763185+2.0493196 \\ &= 69.491521 \end{aligned}$$

$Y'Y = \Sigma\, y_i^2 = 69.6355$

SSE(quartic) = 0.143979, df. = 8 - 5 = 3

MSE = 0.047993

$R(\lambda_4 | \lambda_0, \lambda_1, \lambda_2, \lambda_3) = 2.0493196$

ถ้า fit linear model: $R(\alpha) = 54.444612 + 8.8826005$

$= 63.327212$

$R(\alpha_1 | \alpha_0) = 8.8826005$

SSE(linear) = 6.308288, df. = 8 - 2 = 6

MSE = 1.0513813

ถ้า fit quadratic model:

$R(\delta) = 54.444612 + 8.8826005 + 0.838672$

$= 64.165884$

$R(\delta_2 | \delta_0, \delta_1) = 0.838672$

SSE(quadratic) = 5.469616, df. = 8 - 3 = 5

MSE = 1.0939232

ถ้า fit cubic model:

$$R(\gamma) = 54.444612 + 8.8826005 + 0.838672 + 3.2763185$$
$$= 67.442202$$

$R(\gamma_3 | \gamma_0, \gamma_1, \gamma_2) = 3.2763185$

SSE(cubic) = 2.193298, df. = 8 - 4 = 4

MSE = 0.5483245

Reduction for mean = $(\Sigma y)^2/n = (20.87)^2/8 = 54.444612$

ANOVA

S.V.	d.f.	SS	MS	f_c
Total	8	69.6355		
Reduction for mean	1	54.444612		
Linear	1	8.8826005	8.8826005	8.45*
Error	6	6.308288	1.0513813	
Quadratic	1	0.838672	0.838672	0.77(n.s.)
Error	5	5.469616	1.0939232	
Cubic	1	3.2763185	3.2763185	5.98(n.s.)
Error	4	2.193298	0.5483245	

$f_{(1,6),.05} = 5.59$, $f_{(1,5),.05} = 6.61$,

$f_{(1,4),.05} = 7.71$, $f_{(1,3),.05} = 10.13$

สรุปว่าแบบจำลองเส้นตรงพอดีกับข้อมูล

นั่นคือแบบจำลอง $y = \beta_o + \beta_1 x + e$

$\hat{y} = -0.38048 + 0.91976\ x$

(ดูผลลัพธ์จากการใช้โปรแกรม SPSS/PC)

SPSS/PC commands สำหรับแบบฝึกหัดข้อ 5.4

```
SET DISK=ON.
DATA LIST FREE/Y X.
BEGIN DATA.
1.10 1.5 1.43 2.0 2.11 2.5 3.12 3.0
2.5 3.5 2.21 4.0 2.5 4.5 5.9 5.0
END DATA.
COMPUTE X2=X*X.
COMPUTE X3=X2*X.
COMPUTE X4=X2*X2.
REGRESSION VARIABLES=X X2 X3 X4 Y
          /CRITERIA=PIN(.05)
          /DEPENDENT=Y
          /METHOD=ENTER X
          /METHOD=ENTER X2
          /METHOD=ENTER X3
          /METHOD=ENTER X4.
```

Output สำหรับแบบฝึกหัดข้อ 5.4

```
This procedure was completed at 17:28:13
The raw data or transformation pass is proceeding
     8 cases are written to the uncompressed active file.
------------------------------------------------------------------------------
Page   2                         SPSS/PC+                              1/25/81

       * * * *   M U L T I P L E   R E G R E S S I O N   * * * *


Listwise Deletion of Missing Data

Equation Number 1    Dependent Variable..   Y

Beginning Block Number  1.  Method:  Enter      X
------------------------------------------------------------------------------
```

Page 3 SPSS/PC+ 1/25/81

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number
1.. X

Multiple R	.76468
R Square	.58473
Adjusted R Square	.51552
Standard Error	1.02537

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	8.88260	8.88260
Residual	6	6.30829	1.05138

F = 8.44851 Signif F = .0271

Page 4 SPSS/PC+ 1/25/81

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

------------------ Variables in the Equation ------------------

Variable	B	SE B	Beta	T	Sig T
X	.91976	.31644	.76468	2.907	.0271
(Constant)	-.38048	1.09044		-.349	.7391

------------- Variables not in the Equation -------------

Variable	Beta In	Partial	Min Toler	T	Sig T
X2	1.54525	.36462	.02312	.876	.4213
X3	1.00586	.43211	.07664	1.071	.3330
X4	.85191	.49564	.14056	1.276	.2580

End Block Number 1 All requested variables entered.

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 2. Method: Enter X2

Variable(s) Entered on Step Number
2.. X2

Multiple R	.79996
R Square	.63994
Adjusted R Square	.49592
Standard Error	1.04591

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	9.72127	4.86064
Residual	5	5.46961	1.09392

F = 4.44331 Signif F = .0778

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

------------------ Variables in the Equation ------------------

Variable	B	SE B	Beta	T	Sig T
X	-.91726	2.12272	-.76260	-.432	.6837
X2	.28262	.32277	1.54525	.876	.4213
(Constant)	2.23375	3.18612		.701	.5145

------------- Variables not in the Equation -------------

Variable	Beta In	Partial	Min Toler	T	Sig T
X3	17.47776	.77395	2.1301E-04	2.444	.0709
X4	7.19201	.81322	7.5724E-04	2.795	.0491

End Block Number 2 All requested variables entered.

```
Page  7                    SPSS/PC+                    1/25/81

        **** M U L T I P L E   R E G R E S S I O N ****

Equation Number 1    Dependent Variable..  Y

Beginning Block Number  3.  Method:  Enter     X3


End Block Number   3   Tolerance =     .010 Limits reached.
No variables entered for this block.
------------------------------------------------------------------------------
Page  8                    SPSS/PC+                    1/25/81

        **** M U L T I P L E   R E G R E S S I O N ****

Equation Number 1    Dependent Variable..  Y

Beginning Block Number  4.  Method:  Enter     X4


End Block Number   4   Tolerance =     .010 Limits reached.
No variables entered for this block.
------------------------------------------------------------------------------
Page  9                    SPSS/PC+                    1/25/81

This procedure was completed at 17:29:00
```