Chapter 6

Polynomial Regression Models

Overview

- Second order model with one predictor variable, x

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

  - Inference for second order polynomial with one predictor variable model

- Second order model with two predictor variables, x_1 and x_2

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon$$

  - Inference for second order polynomial with two predictor variables model

6.1 Introduction

- Relationships between variables are not always linear.
- Sometimes we have nonlinear relationships between variables.
- For example,

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon,$$

where β_0 = y-intercept;

β_1 = linear effect of x in y;

β_2 = quadratic effect of x in y and

ϵ = random effect in y.

Introduction (Continued)

- The above model is called a second-order model with one predictor variable.

- Note : The above model can be treated as a multiple regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

where $x_1 = x,\ x_2 = x^2$.

Introduction (Continued)

- A p-th order (or p th-degree polynomial) model with one predictor variable is given by

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_i^j + \epsilon_i$$

or $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p + \epsilon_i$

- Since a polynomial model can be considered as a multiple regression model, hence we can use the techniques in the multiple regression analysis to draw statistical inference for the polynomial model.

6.2 Second Order Model in One Predictor

- Let us consider a 2nd-order model in one predictor variable.

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, \qquad i = 1, \ldots, n,$$

where $\epsilon_i \sim N(0, \sigma^2)$ independently

- It can be represented in a matrix form as follows

$$\underline{y} = X\underline{\beta} + \underline{\epsilon}$$

where

$$\underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix}, \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \text{ and}$$

$$\underline{\epsilon} = (\epsilon_1 \quad \cdots \quad \epsilon_n)'$$

Second Order Model in One Predictor (Continued)

- Hence $\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{y}$ and

  the fitted regression equation is given by

$$y = \underline{x}'\underline{\hat{\beta}}$$

  where $\underline{x}' = (1 \quad x \quad x^2)$

- Is there any significant relationship between y and x?
  ($y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$ versus $y = \beta_0 + \epsilon$)

- Is there any significant difference between the 2nd order model ($y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$) and the 1st order model ($y = \beta_0 + \beta_1 x + \epsilon$)?

6.2.1 Inferences for Second Order Polynomial

- We want to determine whether there is a significant relationship between y and x.

- i.e. Test H_0: $\beta_1 = \beta_2 = 0$ (or $y = \beta_0 + \epsilon$)

  (There is no relationship between x and y.)

  against H_1: $\beta_1 \neq 0$ or $\beta_2 \neq 0$ or both $\beta_1 \neq 0$ and $\beta_2 \neq 0$.

  (or $y = \beta_0 + \beta_2 x^2 + \epsilon$, or $y = \beta_0 + \beta_1 x + \epsilon$, or $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$)

  (There is a relationship between x and y.)

6.2.2 Testing for relationship between *y* and x

$$SSE = \underline{y}'\underline{y} - \underline{\hat{\beta}}'X'\underline{y}$$ with n – (2 + 1) d.f.

$$SSR = \underline{\hat{\beta}}'X'\underline{y} - n\bar{y}^2$$ with 2 d.f.

- Let

$$F = \frac{SSR/2}{SSE/(n-3)}$$

- Reject H_0 at a significance level α if

$$F_{\text{obs}} > F_\alpha(2, n-3).$$

6.2.3 Testing for Quadratic Term

- Test H_0: $\beta_2 = 0$ (i.e. $y = \beta_0 + \beta_1 x + \epsilon$)

  (The 2nd-order model does not improve over the 1st-order model.)

  against H_1: $\beta_2 \neq 0$ (i.e. $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$)

  (The 2nd-order model is a better fit than the 1st-order model.)
- $SSR(x^2|x) = SSR(x,\ x^2) - SSR(x)$ with 1 d.f.
- Let

$$F = SSR(x^2|x)/[SSE/(n-3)]$$

- Reject H_0 at a significance level α if

$$F_{\text{obs}} > F_\alpha(1, n-3).$$

Testing for Quadratic Term (Continued)

- Alternatively, we may use the t-test, where

$$t = \frac{\hat{\beta}_2}{s.e.\left(\hat{\beta}_2\right)}$$

- Reject H_0 at a significance level α if

$$t_{\text{obs}} > t_{\alpha/2}(n-3).$$

6.2.4 Testing for Linear Term

- H_0: $\beta_1 = 0$ (i.e. $y = \beta_0 + \beta_2 x^2 + \varepsilon$)
  (Including the linear effect term does not improve the quadratic effect model) against
- H_1: $\beta_1 \neq 0$ (i.e. $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$)
  (Including the linear effect term improves the quadratic effect model.)
- $SSR(x|x^2) = SSR(x,\ x^2) - SSR(x^2)$ with 1 d.f.
- Let $F = SSR(x|x^2)/[SSE/(n-3)]$
- Reject H_0 at a sig level α if $F_{\text{obs}} > F_\alpha(1, n-3)$.
- We may also use the t-test, where $t = \dfrac{\hat{\beta}_1}{s.e.(\hat{\beta}_1)}$

6.3 Example 1

- The cloud point of a liquid is a measure of the degree of crystallization in a stock that can be measured by the refractive index.

- It has been suggested that the percentage of I-8 in the base stock is an excellent predictor of cloud point using the 2nd-order model

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

- The following data were collected on stocks with known percentage of I-8.

Example 1 (Continued)

I-8, x	0	1	2	3	4	5	6	7	8	0
Cloud point, y	22.1	24.5	26.0	26.8	28.2	28.9	30	30.4	31.4	21.9

I-8, x	2	4	6	8	10	0	3	6	9
Cloud point, y	26.1	28.5	30.3	31.5	33.1	22.8	27.3	29.8	31.8

- By using statistical software, we obtain the following results.
- The fitted regression equation is given by

$$\hat{y} = 22.5612 + 1.6680x - 0.6796x^2$$

Example 1 (Continued)

ANOVA Table

Source	SS	df	MS	F	p-value
Regression	201.9944	2	100.9972	649.87	< 0.0001
Error	2.4866	16	0.1554		
Total	204.4811	18			

- Since $F_{\text{obs}} = 649.87 > F_{0.05}(2, 16) = 3.63$ (or p-value < 0.05), therefore we reject the null hypothesis that there is no significant model at the 5% level of significance. (i.e. We reject H_0: $\beta_1 = \beta_2 = 0$ at the 5% significance level.)

Example 1 (Continued)

- Note:

$$R^2 \left(= r^2_{y\cdot\{x,x^2\}}\right) = \frac{SSR}{SST} = 0.9878$$

- Thus 98.8% of the variation in cloud point can be explained by the 2nd-degree polynomial relationship between percentage of I-8 and the cloud point.

Example 1 (Continued)

- Next, we want to test

$$H_0: \beta_2 = 0 \text{ against } H_1: \beta_2 \neq 0.$$

or $H_0: y = \beta_0 + \beta_1 x + \varepsilon$ against

$H_1: y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$

- We have $SSR(x^2 \mid x) = 6.7516$ with 1 d.f.
- $F = SSR(x^2 \mid x)/MSE = 6.7516/0.1554 = 43.44$.
- Since $F_{\text{obs}} = 43.44 > F_{0.05}(1, 16) = 4.49$ (or p-value $= 6.2088(10)^{-6} < 0.05$), we reject H_0 and conclude that the 2nd-order model is significantly better than the 1st-order model.

Example 1 (Continued)

- We may also consider to test

$$\text{H}_0: \beta_1 = 0 \text{ against } \text{H}_1: \beta_1 \neq 0$$

or $\text{H}_0: y = \beta_0 + \beta_2 x^2 + \epsilon$ against

$$\text{H}_1: y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

- We have $SSR(x|x^2) = 44.1607$ with 1 d.f.
- $F = SSR(x|x^2)/MSE = 44.1607/.01554 = 284.17.$
- Since $F_{\text{obs}} = 284.17 > F_{0.05}(1, 16) = 4.49$ (or p-value $= 1.3112(10)^{-11} < 0.05$), we reject H_0 and conclude that the polynomial model $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$ is a significantly better fit than the one which includes only quadratic effect ($y = \beta_0 + \beta_2 x^2 + \epsilon$).

Example 1 (Continued)

- Since there were repeat measurements for some of the predictor values, we may like to perform a lack of fit test.

H_0: There is no lack of fit against

H_1: There is lack of fit

Example 1 (Continued)

X	Y	$\sum_{k=1}^{n_j}(y_{jk} - \bar{y}_j)^2$	d.f.
0	22.1, 21.9,22.8	0.44667	2
2	26.0, 26.1	0.005	1
3	26.8, 27.3	0.125	1
4	28.2, 28.5	0.045	1
6	30.0, 03.3, 29.8	0.12667	2
8	31.4, 31.5	0.005	1
	SSPE =	0.75334	8

Example 1 (Continued)

- $SSLF = SSE - SSPE = 1.733325$ with 8 d.f.

- $F_L = (SSLF/8)/(SSPE/8) = 2.30$.

- Since $F_L = 2.30 < F_{0.05}(8, 8) = 3.44$ (or p-value = 0.1300 > 0.05), hence we do not reject the hypothesis that there is no lack of fit and conclude that the quadratic model is sufficient for the predictive purpose.

6.4 Second Order Model in Two Predictors

- Consider the model
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon$$

- This model contains two linear effect terms, $\beta_1 x_1$ and $\beta_2 x_2$; two quadratic effect terms $\beta_{11} x_1^2$ and $\beta_{22} x_2^2$ and an interaction effect term $\beta_{12} x_1 x$.

- Note: The model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon$ is a 2nd-order polynomial with $\beta_{11} = \beta_{22} = 0$.

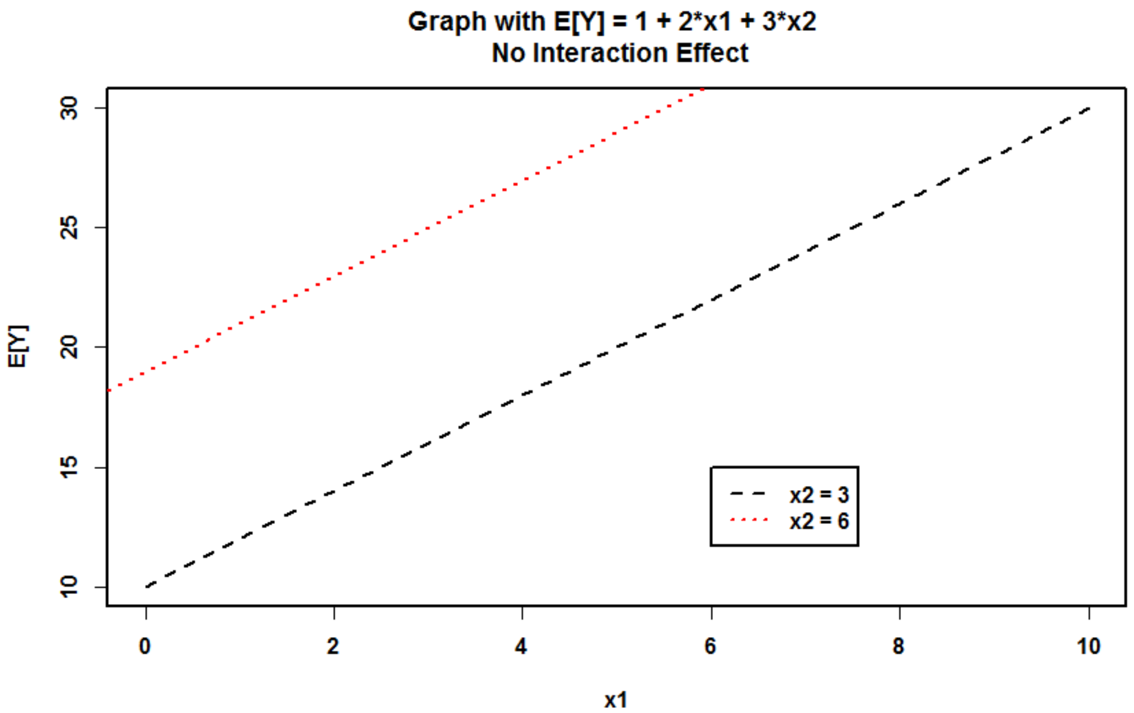
<u>6.4.1 Interpretation of the Interaction Effect</u>

- If y varies in a similar way with respect to x_1 **regardless of the levels of $\boldsymbol{x_2}$**, i.e. the relationship between y and x_1 does not in any way depend on x_2, then we say that there is **no interaction** between x_1 and x_2, (and $\beta_{12} = 0$).

- <u>This does not mean that y and x_2 are uncorrelated</u>, but that the relationship between y and x_1 does not vary as a function of x_2.

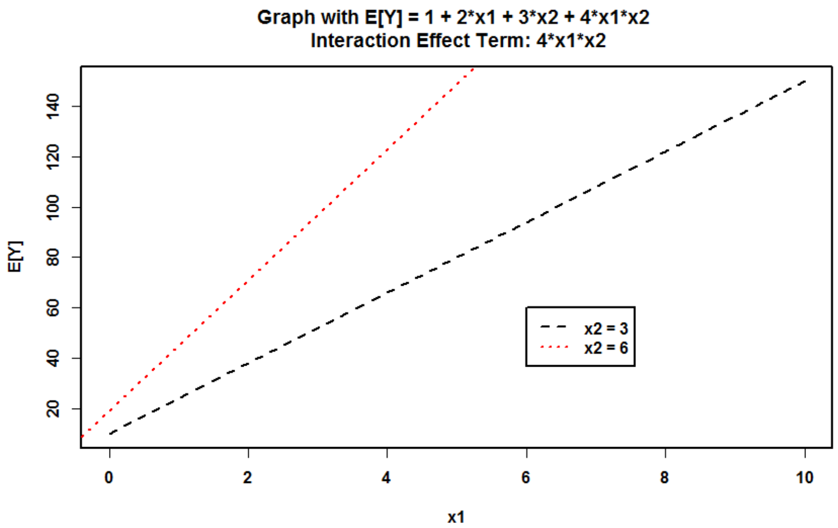
6.4.1 Interpretation of the Interaction Effect

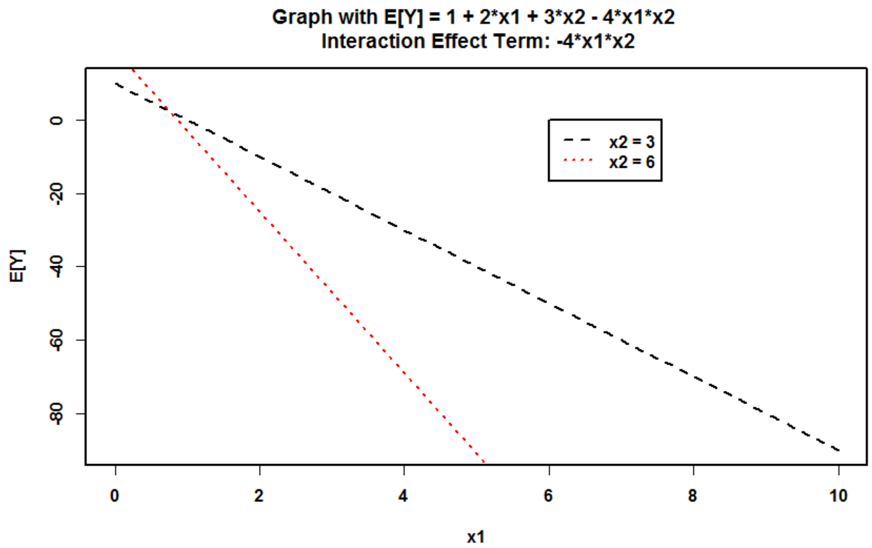
- In general, if x_1 does not interact with x_2, the regression equations $E(y)$ against x_1 are 'parallel' for different values of x_2.



6.4.1 Interpretation of the Interaction Effect

- If x_1 interacts with x_2, the regression equations $E(y)$ against x_1 are not 'parallel' for different values of x_2.





6.5 Example 2

- A chemical engineer is investigating the influence of two variables, reaction time and temperature, on process yield.

- Twenty-four observations were collected.

- A second order polynomial model is used to fit the data.

- Let y = yield, x_1 = reaction time and x_2 = temperature.

Example 2 (Continued)

y	50.95	47.35	50.99	44.96	41.89	41.44	51.79	50.78
x_1	76.0	80.5	78.0	89.0	93.0	92.1	77.8	84.0
x_2	170	165	182	185	180	172	170	180

y	42.48	49.80	48.74	46.20	50.49	52.78	49.71	52.75
x_1	87.3	75.0	85.0	90.0	85.0	79.2	83.0	82.0
x_2	165	172	185	176	178	174	168	179

y	39.41	43.63	38.19	50.92	46.55	44.28	48.72	49.13
x_1	94.0	91.4	95.0	81.1	88.8	91.0	87.0	86.0
x_2	181	184	173	169	183	178	175	175

Example 2 (Continued)

- The following results are obtained using either SAS or R.

- The programs are shown at the end of this chapter.

- The regression equation is given by

$$\hat{y} = 1317.6309 + 7.2660x_1 + 12.2765x_2 - 0.0597x_1^2 - 0.0377x_2^2 + 0.0126x_1x_2$$

Example 2 (Continued)

- ANOVA table

Source	SS	df	MS	F	p-value
Regression	416.3111	5	83.2622	206.29	< 0.0001
Error	7.2654	18	0.4036		
Total	423.5765	23			

- Since $F_{obs} = 206.29 > F_{0.05}(5, 18) = 2.77$ (or p-value < 0.05), therefore we reject the null hypothesis that there is no significant model at the 5% level of significance (i.e. we reject H_0: $\beta_1 = \beta_2 = \beta_{11} = \beta_{22} = \beta_{12} = 0$.) and conclude that there is a significant relationship between yield and the 2 predictors, reaction time and temperature.

Example 2 (Continued)

- Note:

$$R^2 \left(= r^2_{y\cdot\{1,2,11,22,12\}}\right) = \frac{SSR}{SST} = 0.9828$$

- Thus 98.3% of the variation in yield can be explained by the 2nd-order polynomial model with reaction time and temperature.

Example 2 (Continued)

- Test the contribution of the quadratic terms to the model.

- Test H_0: $\beta_{11} = \beta_{22} = \beta_{12} = 0$
(i.e. the 2nd-order model does not improve over the 1st order model.)
against H_1: at least one $\beta_{jk} \neq 0$.

- Under H_0, the reduced model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

Example 2 (Continued)

- From the computer printout of fitting the above reduced model, we have
- $SSE_{\mathrm{H}} = 90.4884$ with 21 d.f.
- $F_1 = [(SSE_{\mathrm{H}} - SSE)/3]/MSE$

  $= [(90.4884 - 7.2654)/3]/0.4036 \approx 68.73$.
- Since $F_{1,\,\mathrm{obs}} = 68.73 > F_{0.05}(3, 18) = 3.16$ (or p-value $\approx 4.7(10)^{-10} < 0.05$), we reject H_0 and conclude that at least one of the quadratic terms is necessary.
- Note: $SSE_{\mathrm{H}} = SSE(x_1, x_2)$

Example 2 (Continued)

- Test for the interaction effect.
- Test H_0: $\beta_{12} = 0$ (no interaction effect.) against H_1: $\beta_{12} \neq 0$.
- Using the partial F-test:

$$F_2 = \frac{SSR(x_1 x_2 | x_1, x_2, x_1^2, x_2^2)}{MSE}$$

$$= \frac{SSR(x_1, x_2, x_1^2, x_2^2, x_1 x_2) - SSR(x_1, x_2, x_1^2, x_2^2)}{MSE}$$

$$= \frac{416.3111 - 413.9780}{0.4036} \approx 5.78$$

Example 2 (Continued)

- Since $F_{2,\text{ obs}} = 5.78 > F_{0.05}(1, 18) = 4.41$ (or p-value ≈ 0.0272 < 0.05), we reject H_0 at the 5% significance level and conclude that there is a significant interaction effect between reaction time and temperature.

- Note:

$$SSR(x_1, x_2, x_1^2, x_2^2, x_1x_2) - SSR(x_1, x_2, x_1^2, x_2^2) = SSE_H - SSE$$

6.5 Programs

6.5.1 SAS Program

```
data ch6ex2;
  infile "d:\ST3131\Lecture\ch6ex2.txt" firstobs=2;
  input y x1 x2;
  x3=x1**2; x4=x2**2; x5=x1*x2;
proc glm data=ch6ex2;
  model y = x1 x2 x3 x4 x5;
/* model y = x1 x2 x1*x1 x2*x2 x1*x2; */
run;
proc glm data=ch6ex2;
  model y = x1 x2; /* Fit the reduced model */
run;
proc reg data=ch6ex2;
  model y = x1 x2 x3 x4 x5;
  hypo 1: test x3=0, x4=0, x5=0;
run;
```

Partial SAS Output

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	416.3111251	83.2622250	206.28	<.0001
Error	18	7.2653707	0.4036317		
Corrected Total	23	423.5764958			

R-Square	Coeff Var	Root MSE	y Mean
0.982848	1.344676	0.635320	47.24708

Source	DF	Type I SS	Mean Square	F Value	Pr > F
x1	1	312.00718	312.007188	773.00	<.0001
x2	1	21.0808622	21.0808622	52.23	<.0001
x1*x1	1	49.0279411	49.0279411	121.47	<.0001
x2*x2	1	31.8626357	31.8626357	78.94	<.0001
x1*x2	1	2.3324982	2.3324982	5.78	0.0272

$SSR(X_1)$ $SSR(X_2 | X_1)$ $SSR(X_1^2 | X_1, X_2)$ $SSR(X_2^2 | X_1, X_2, X_1^2)$

$SSR(X_1X_2 | X_1, X_2, X_1^2, X_2^2)$

Partial SAS Output (Continued)

Source	DF	Type III SS	Mean Square	F Value	Pr > F
x1	1	22.01233071	22.01233071	54.54	<.0001
x2	1	32.78318758	32.78318758	81.22	<.0001
x1*x1	1	69.20633577	69.20633577	171.46	<.0001
x2*x2	1	34.06268226	34.06268226	84.39	<.0001
x1*x2	1	2.33249815	2.33249815	5.78	0.0272

$SSR(X_1X_2 \mid X_1, X_2, X_1^2, X_2^2)$

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	-1317.630871	33.8001455	-9.85	<.0001
x1	7.265961	0.9839040	7.38	<.0001
x2	12.276522	1.3622047	9.01	<.0001
x1*x1	-0.059653	0.0045557	-13.09	<.0001
x2*x2	-0.037676	0.0041012	-9.19	<.0001
x1*x2	0.012577	0.0052319	2.40	0.0272

$SSR(X_1 \mid X_2, X_1^2, X_2^2, X_1X_2)$

$SSR(X_2 \mid X_1, X_1^2, X_2^2, X_1X_2)$

$SSR(X_1^2 \mid X_1, X_2, X_2^2, X_1X_2)$

$SSR(X_2^2 \mid X_1, X_2, X_1^2, X_1X_2)$

Partial SAS Output (Continued)

Reduced Model:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	333.08805	66.54403	38.65	<.0001
Error	21	90.48845	4.30897		
Corrected Total	23	423.57650			

Source	DF	Type I SS	Mean Square	F value	Pr > F
x1	1	312.00719	312.00719	72.41	<.0001
x2	1	21.08086	21.08086	4.89	0.038

Source	DF	Type III SS	Mean Square	F value	Pr > F
x1	1	28.05010	328.05010	76.13	<.0001
x2	1	21.08086	21.08086	4.89	0.0382

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	76.56114	12.59687	6.08	<.0001
x1	-0.68984	0.07906	-8.73	<.0001
x2	0.16863	0.07624	2.21	0.0382

Partial SAS Output (Continued)

Output from "Proc reg" for testing $\beta_3 = \beta_4 = \beta_5 = 0$

(SAS statement "Hypo 1: test x3=0, x4=0, x5=0;")

Source	DF	Mean Square	F Value	Pr > F
Numerator	3	27.74102	68.73	<.0001
Denominator	18	0.40363		

$SSR(X_1^2, X_2^2, X_1{*}X_2 \mid X_1, X_2) / 3$

$MSE(X_1, X_2, X_1^2, X_2^2, X_1{*}X_2)$

6.5.2 R Program

```
> ch6ex2=read.table("d:/ST3131/ch6ex2.txt",header=T)
> attach(ch6ex2)
> #Full model y = x1 + x2 + x1^2 + x2^2 +x1*x2
> x11=x1^2; x22=x2^2; x12=x1*x2
> model1=lm(y~x1+x2+x11+x22+x12)
> summary(model1)
Coefficients:
              Estimate  Std. Error  t value  Pr(>|t|)
(Intercept) -1.318e+03  1.338e+02    -9.848  1.13e-08 ***
x1           7.266e+00  9.839e-01     7.385  7.51e-07 ***
x2           1.228e+01  1.362e+00     9.012  4.32e-08 ***
x11         -5.965e-02  4.556e-03   -13.094  1.22e-10 ***
x22         -3.768e-02  4.101e-03    -9.186  3.24e-08 ***
x12          1.258e-02  5.232e-03     2.404    0.0272 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
   ' ' 1
```

R Program (continued)

```
Residual standard error: 0.6353 on 18 degrees of freedom
Multiple R-squared: 0.9828,     Adjusted R-squared: 0.9781
F-statistic: 206.3 on 5 and 18 DF,  p-value: 3.097e-15
> anova(model1)
Analysis of Variance Table

Response: y
           Df  Sum Sq  Mean Sq   F value      Pr(>F)
x1          1 312.007  312.007  772.9997   3.066e-16 ***
x2          1  21.081   21.081   52.2280   1.010e-06 ***
x11         1  49.028   49.028  121.4670   1.960e-09 ***
x22         1  31.863   31.863   78.9399   5.333e-08 ***
x12         1   2.332    2.332    5.7788     0.02721 *
Residuals  18   7.265    0.404
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

SSE

R Program (continued)

```
> #Reduced model y = x1 + x2
> model2=lm(y~x1+x2)
> anova(model2,model1)
Analysis of Variance Table

Model 1: y ~ x1 + x2
Model 2: y ~ x1 + x2 + x11 + x22 + x12
  Res. Df      RSS     Df   Sum of Sq        F      Pr(>F)
1       21   90.488
2       18    7.265      3      83.223   68.729   4.707e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

SSE_{H0}

SSE

$SSE_{H0} - SSE$

R Program (continued)

```
> #Reduced model y = x1 + x2 + x1^2 + x2^2
> model3=lm(y~x1+x2+x11+x22)
> anova(model3,model1)
Analysis of Variance Table

Model 1: y ~ x1 + x2 + x11 + x22
Model 2: y ~ x1 + x2 + x11 + x22 + x12
     Res.Df      RSS    Df   Sum of Sq         F     Pr(>F)
1        19   9.5979
2        18   7.2654     1      2.3325    5.7788    0.02721 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Recap

- Second order model with one predictor variable, *x*.

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

- Inference for second order polynomial model
  - Test H_0: $\beta_1 = \beta_2 = 0$ (Overall significance)
    - Full model: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$
    - Reduced model: $y = \beta_0 + \epsilon$
  - Test H_0: $\beta_2 = 0$
    - Full model: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$
    - Reduced model: $y = \beta_0 + \beta_1 x + \epsilon$
  - Test H_0: $\beta_1 = 0$
    - Full model: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$
    - Reduced model: $y = \beta_0 + \beta_2 x^2 + \epsilon$

Recap (Continued)

- Second order model with two predictor variables, x_1 and x_2. Full model is given by:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon$$

- Inference for second order polynomial model
  - Test H_0: $\beta_1 = \beta_2 = \beta_{11} = \beta_{22} = \beta_{12} = 0$ (Overall significance)
    - Reduced model under H_0: $y = \beta_0 + \epsilon$
  - Test H_0: $\beta_{11} = \beta_{22} = \beta_{12} = 0$
    - Reduced model under H_0: $y = \beta_0 + \beta_1 x + \beta_2 x_2 + \epsilon$
  - Test H_0: $\beta_{12} = 0$
    - Reduced model under H_0: $y = \beta_0 + \beta_1 x + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \epsilon$