High dimensional data analysis-revised version of HW3 105225019 陳怡方 2018/05/22

(P125) Exercise 14.

This problem focuses on the collinearity problem.

(a) Perform the following commands in R:

```
> set.seed(1)
> x1=runif(100)
> x2=0.5*x1+rnorm(100)/10
> y=2+2*x1+0.3*x2+rnorm(100)
```

The last line corresponds to creating a linear model in which y is a function of x1 and x2. Write out the form of the linear model. What are the regression coefficient?

(b) What is the correlation between x1 and x2? Create a scatterplot displaying the relationship between the variables.

(c) Using this data, fit a least squares regression to predict y using x1 and x2. Describe the results obtained. What are $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_2$? How do these relate to the true β_0, β_1, and β_2? Can you reject the null hypothesis $H_0 : \beta_1 = 0$? How about the null hypothesis $H_0 : \beta_2 = 0$?

(d) Now fit a least squares regression to predict y using only x1. Comment on your results. Can you reject the null hypothesis $H_0 : \beta_1 = 0$?

(e) Now fit a least squares regression to predict y using only x2. Comment on your results. Can you reject the null hypothesis $H_0 : \beta_1 = 0$?

(f) Do the results obtained in (c)–(e) contradict each other? Explain your answer.

(g) Now suppose we obtain one additional observation, which was unfortunately mismeasured.

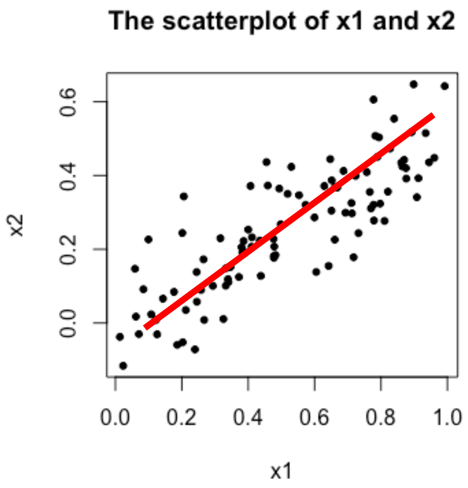
```
> x1=c(x1, 0.1)
> x2=c(x2, 0.8)
> y=c(y,6)
```

Re-fit the linear models from (c) to (e) using the new data. What effect does this new observation have on the each of the models? In each model, is this observation an outlier? A high-leverage point? Both? Explain your answers.

Solution:

(a) The linear model is $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$, where $\beta_0 = 2, \beta_1 = 2, \beta_2 = 0.3$.

(b) Correlation between x1 and x2 is $\rho_{12} = \frac{cov(x_1,x_2)}{\sigma_1\sigma_2} = \frac{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)(x_{i2}-\bar{x}_2)}{\sqrt{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2\sum_{i=1}^{n}(x_{i2}-\bar{x})^2}} = 0.8351$



Correlation is positive!

(c) model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$

```
Call:
lm(formula = y ~ x1 + x2)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.8311 -0.7273 -0.0537  0.6338  2.3359 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.1305     0.2319   9.188 7.61e-15 ***
x1            1.4396     0.7212   1.996   0.0487 *  
x2            1.0097     1.1337   0.891   0.3754    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.056 on 97 degrees of freedom
Multiple R-squared:  0.2088,	Adjusted R-squared:  0.1925 
F-statistic:  12.8 on 2 and 97 DF,  p-value: 1.164e-05
```

$\hat{\beta}_0 = 2.1305, \hat{\beta}_1 = 1.4396, \hat{\beta}_2 = 1.0097$

$\hat{y} = 2.1305 + 1.4396x_1 + 1.0097x_2$

The $\hat{\beta}_0$ and $\hat{\beta}_1$ close to the true value, but $\hat{\beta}_2$ is far from the true value.

	$H_0: \beta_1 = 0\ vs. H_1: \beta_1 \neq 0$	$H_0: \beta_2 = 0\ vs. H_1: \beta_2 \neq 0$
p-value	0.0487 < 0.05	0.3754 > 0.05
result	Reject H_0 at $\alpha = 0.05$	Do not reject H_0 at $\alpha = 0.05$

(d) model: $y = \beta_0 + \beta_1 x_1 + \epsilon$

```
Call:
lm(formula = y ~ x1)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.89495 -0.66874 -0.07785  0.59221  2.45560 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.1124     0.2307   9.155 8.27e-15 ***
x1            1.9759     0.3963   4.986 2.66e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.055 on 98 degrees of freedom
Multiple R-squared:  0.2024,	Adjusted R-squared:  0.1942 
F-statistic: 24.86 on 1 and 98 DF,  p-value: 2.661e-06
```

$\hat{\beta}_0 = 2.1124, \hat{\beta}_1 = 1.9758$

$\hat{y} = 2.1124 + 1.9758x_1$

	$H_0: \beta_1 = 0 \; vs. H_1: \beta_1 \neq 0$
p-value	2.66e-06 < 0.05
result	Reject H_0 at $\alpha = 0.05$

(e) model: $y = \beta_0 + \beta_1 x_2 + \epsilon$

```
Call:
lm(formula = y ~ x2)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.62687 -0.75156 -0.03598  0.72383  2.44890 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.3899     0.1949   12.26  < 2e-16 ***
x2            2.8996     0.6330    4.58 1.37e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.072 on 98 degrees of freedom
Multiple R-squared:  0.1763,	Adjusted R-squared:  0.1679 
F-statistic: 20.98 on 1 and 98 DF,  p-value: 1.366e-05
```

$\hat{\beta}_0 = 2.3899, \hat{\beta}_1 = 1.8996$

$\hat{y} = 2.3899 + 1.8996x_2$

	$H_0: \beta_1 = 0 \; vs. H_1: \beta_1 \neq 0$
p-value	1.37e-06 < 0.05
result	Reject H_0 at $\alpha = 0.05$

(f) When we use all of variables (x1 and x2) to fit the model, the coefficient of x2 is not significant.

When we use only x1 or only x2 to fit the model, the coefficients of x1 and x2 are all significant.

So, x1 and x2 maybe collinearity. And we have to check the VIF.

	$Var(\hat{\beta})$ of multiple regression	$Var(\hat{\beta})$ of simple regression
x_1	0.7212^2	0.3963^2
x_2	1.1337^2	0.6330^2

$$VIF_1 = \frac{Var(\hat{\beta}_1|Multi)}{Var(\hat{\beta}_1^{(0)}|Simple)} = \frac{0.7212^2}{0.3963^2} = 3.3118 < 5$$

$$VIF_2 = \frac{Var(\hat{\beta}_2|Multi)}{Var(\hat{\beta}_2^{(0)}|Simple)} = \frac{1.1337^2}{0.6330^2} = 3.2077 < 5$$

Hence, x1 and x2 are not collinearity.

(g) (c) model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$

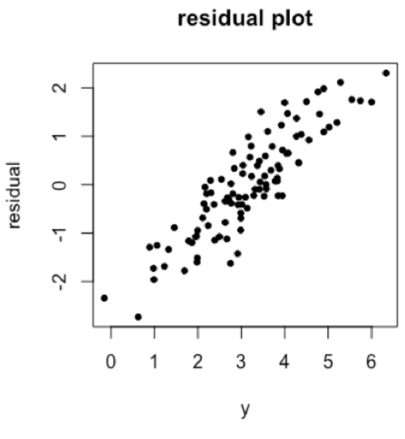
```
Call:
lm(formula = y2 ~ x11 + x21)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.73348 -0.69318 -0.05263  0.66385  2.30619 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.2267     0.2314   9.624 7.91e-16 ***
x11           0.5394     0.5922   0.911  0.36458    
x21           2.5146     0.8977   2.801  0.00614 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.075 on 98 degrees of freedom
Multiple R-squared:  0.2188,	Adjusted R-squared:  0.2029 
F-statistic: 13.72 on 2 and 98 DF,  p-value: 5.564e-06
```

$\hat{\beta}_0 = 2.2267, \hat{\beta}_1 = 0.5349, \hat{\beta}_2 = 2.5146$

$\hat{y} = 2.2267 + 0.5349x_1 + 2.5146x_2$

	$H_0: \beta_1 = 0\ vs. H_1: \beta_1 \neq 0$	$H_0: \beta_2 = 0\ vs. H_1: \beta_2 \neq 0$
p-value	0.36458 > 0.05	0.00614 < 0.05
result	Do not Reject H_0 at $\alpha = 0.05$	Reject H_0 at $\alpha = 0.05$

$\hat{\sigma}^2 = 1.1556$ are not larger than (c),and $R^2 = 0.2029$ are not smaller then (c).



The new observation in full is not outlier.

$h_{ii} = X_i^T(X^TX)^{-1}X_i = [0.1 \quad 0.8]\begin{bmatrix}0.2078 & -0.3523\\ -0.3523 & 0.6962\end{bmatrix}\begin{bmatrix}0.1\\0.8\end{bmatrix} = 0.3913 > \frac{2}{n} = 0.0198$. So the new observation is a leverage points.

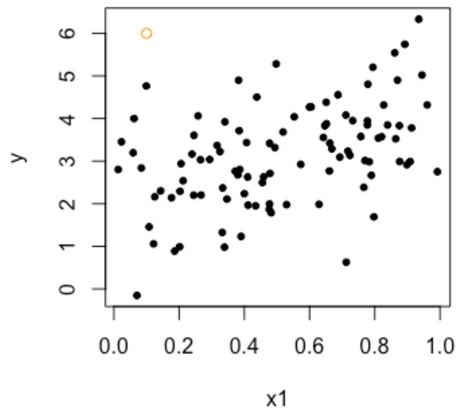
(d) model: $y = \beta_0 + \beta_1 x_1 + \epsilon$

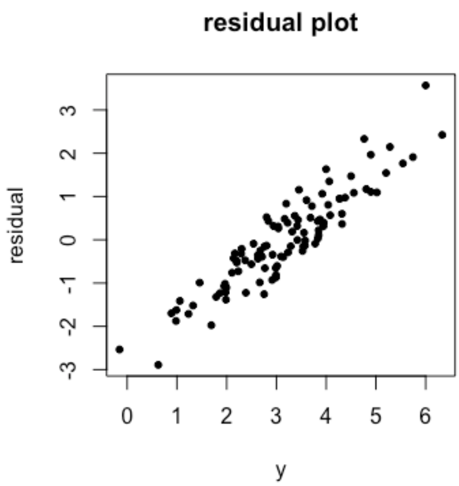
```
Call:
lm(formula = y2 ~ x11)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.8897 -0.6556 -0.0909  0.5682  3.5665 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.2569     0.2390   9.445 1.78e-15 ***
x11           1.7657     0.4124   4.282 4.29e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.111 on 99 degrees of freedom
Multiple R-squared:  0.1562,	Adjusted R-squared:  0.1477 
F-statistic: 18.33 on 1 and 99 DF,  p-value: 4.295e-05
```

$\hat{\beta}_0 = 2.2569, \hat{\beta}_1 = 1.7657$

$\hat{y} = 2.2569 + 1.7657x_1$

	$H_0: \beta_1 = 0 \ vs. H_1: \beta_1 \neq 0$
p-value	4.29e-05 < 0.05
result	Reject H_0 at $\alpha = 0.05$





$\hat{\sigma}^2 = 1.234$ is larger than (d),and $R^2 = 0.1477$ is smaller than (d).

The new observation in this model is outlier.

$$h_i = \frac{(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2} = \frac{(0.1 - 0.5137)^2}{7.2614} = 0.0236 > \frac{2}{n} = 0.0198$$

The new observation in this model is leverage point.

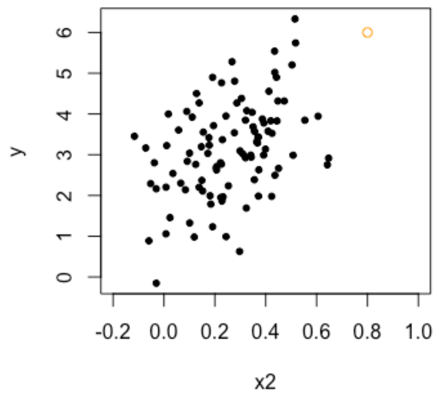
(e) model: $y = \beta_0 + \beta_1 x_2 + \epsilon$

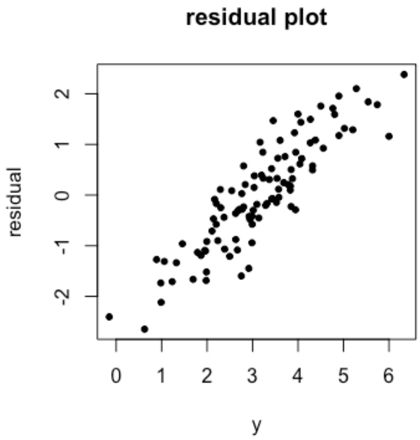
```
Call:
lm(formula = y2 ~ x21)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.64729 -0.71021 -0.06899  0.72699  2.38074 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2.3451     0.1912  12.264  < 2e-16 ***
x21           3.1190     0.6040   5.164 1.25e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.074 on 99 degrees of freedom
Multiple R-squared:  0.2122,	Adjusted R-squared:  0.2042 
F-statistic: 26.66 on 1 and 99 DF,  p-value: 1.253e-06
```

$\hat{\beta}_0 = 2.4351, \hat{\beta}_1 = 3.1190$

$\hat{y} = 2.3451 + 3.1190x_2$

	$H_0: \beta_1 = 0 \; vs. H_1: \beta_1 \neq 0$
p-value	1.25e-06 < 0.05
result	Reject H_0 at $\alpha = 0.05$





$\hat{\sigma}^2 = 1.1535$ not larger than (d),and $R^2 = 0.2042$ not smaller than (d).

The new observation in this model is not outlier.

$$h_i = \frac{(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2} = \frac{(0.8 - 0.2625)^2}{3.1601} = 0.0914 > \frac{2}{n} = 0.0198$$

The new observation in this model is leverage point.

(P298) Exercise 2.

Suppose that the curve $\hat{g}$ is computed to smoothly fit a set of n points, using the following formula:

$$\hat{g} = argmin\left[\sum_{i=1}^{n}(y_i - g(x_i))^2 + \lambda\int\left[g^{(m)}(x)\right]^2 dx\right],$$

where $g^{(m)}$ represents the mth derivative of g $\left(\text{and } g^{(0)} = g\right)$. Provide example sketches of $\hat{g}$ in each of the following scenarios.

(e) $\lambda = 0, m = 3$

solution:

$g^{(3)} = g'''(x)$, for $\lambda = 0$

case1. $\int\left(g'''(x)\right)^2 dx = 0$

$g'''(x) = 0 \Longrightarrow g(x) = ax^2 + bx + c$

Find $\hat{g}$ such that $\sum_{i=1}^{n}(y_i - \hat{g}(x_i))^2 = 0$

$$\text{Let } X = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_3^2 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} c \\ b \\ a \end{bmatrix}$$

$$X^TX = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \end{bmatrix}\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 \end{bmatrix}$$

$(X^TX)^{-1} =$

$$\frac{1}{det(X^TX)}\begin{bmatrix} \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 - (\sum_{i=1}^{n} x_i^3)^2 & -(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^3) & \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - (\sum_{i=1}^{n} x_i^2 \\ -(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i^2) & n\sum_{i=1}^{n} x_i^4 - (\sum_{i=1}^{n} x_i^2)^2 & -(n\sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} \\ \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - (\sum_{i=1}^{n} x_i^2)^2 & -(n\sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2) & n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2 \end{bmatrix}$$

$$X^Ty = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i^2 y_i \end{bmatrix}$$

$\hat{\beta} = (X^TX)^{-1}X^Ty =$

$$\frac{1}{det(X^TX)}\begin{bmatrix} \left(\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 - (\sum_{i=1}^{n} x_i^3)^2\right)\sum_{i=1}^{n} y_i - (\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^3)\sum_{i=1}^{n} x_i y_i + \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - (\sum_{i=1}^{n} x_i^2)^2\right)\sum_{i=1}^{n} x_i^2 \\ -(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i^2)\sum_{i=1}^{n} y_i + \left(n\sum_{i=1}^{n} x_i^4 - (\sum_{i=1}^{n} x_i^2)^2\right)\sum_{i=1}^{n} x_i y_i - (n\sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2)\sum_{i=1}^{n} x_i^2 y_i \\ \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - (\sum_{i=1}^{n} x_i^2)^2\right)\sum_{i=1}^{n} y_i - (n\sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2)\sum_{i=1}^{n} x_i y_i + (n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)\sum_{i=1}^{n} x_i^2 y_i \end{bmatrix}$$

Hence,

$a = \frac{1}{det(X^TX)}[(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - (\sum_{i=1}^{n} x_i^2)^2)\sum_{i=1}^{n} y_i - (n\sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2)\sum_{i=1}^{n} x_i y_i +$

$(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)\sum_{i=1}^{n} x_i^2 y_i]$

$b = \frac{1}{det(X^TX)}[-(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i^2)\sum_{i=1}^{n} y_i + (n\sum_{i=1}^{n} x_i^4 - (\sum_{i=1}^{n} x_i^2)^2)\sum_{i=1}^{n} x_i y_i -$

$(n\sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2)\sum_{i=1}^{n} x_i^2 y_i]$

$c = \frac{1}{det(X^TX)}[(\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 - (\sum_{i=1}^{n} x_i^3)^2)\sum_{i=1}^{n} y_i - (\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^3)\sum_{i=1}^{n} x_i y_i +$

$\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \left(\sum_{i=1}^{n} x_i^2\right)^2\right) \sum_{i=1}^{n} x_i^2 y_i \Big]$

Therefore, $\hat{g} = X\hat{\beta}$

case2. $\int \left(g'''(x)\right)^2 dx = 1$

$g'''(x) = 1 \Longrightarrow g(x) = \frac{1}{6}x^3 + ax^2 + bx + c$

Find $\hat{g}$ such that $\sum_{i=1}^{n}(y_i - \hat{g}(x_i))^2 = 0$

$$Let\ X = \begin{bmatrix} 1 & x_1 & x_2^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}, y = \begin{bmatrix} y_1 - \frac{1}{6}x_1^3 \\ y_2 - \frac{1}{6}x_2^3 \\ \vdots \\ y_n - \frac{1}{6}x_n^3 \end{bmatrix}, \beta = \begin{bmatrix} c \\ b \\ a \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_2^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 \end{bmatrix}$$

$(X^T X)^{-1} =$

$$\frac{1}{det(X^T X)} \begin{bmatrix} \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^3\right)^2 & -\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^3\right) & \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \left(\sum_{i=1}^{n} x_i^2 \right. \\ -\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i^2\right) & n \sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^2\right)^2 & -\left(n \sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_1 \sum_{i=1}^{n} \right. \\ \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \left(\sum_{i=1}^{n} x_i^2\right)^2 & -\left(n \sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2\right) & n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \end{bmatrix} \begin{bmatrix} y_1 - \frac{1}{6}x_1^3 \\ y_2 - \frac{1}{6}x_2^3 \\ \vdots \\ y_n - \frac{1}{6}x_n^3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \\ \sum_{i=1}^{n} x_i \left(y_i - \frac{1}{6}x_i^3\right) \\ \sum_{i=1}^{n} x_i^2 \left(y_i - \frac{1}{6}x_i^3\right) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \\ \sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \\ \sum_{i=1}^{n} \left(x_i^2 y_i - \frac{1}{6}x_i^5\right) \end{bmatrix}$$

$\hat{\beta} = (X^T X)^{-1} X^T y =$

$$\frac{1}{det(X^T X)} \begin{bmatrix} \sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \left(\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^3\right)^2\right) - \sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^3\right) + \sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \left(\sum_{i=1}^{n} x_i^2\right)^2\right) \\ -\sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i^2\right) + \sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \left(n \sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^2\right)^2\right) - \sum_{i=1}^{n} \left(x_i^2 y_i - \frac{1}{6}x_i^5\right) - \left(n \sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_1 \sum_{i=1}^{n} x_i^2\right) \\ \sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \left(\sum_{i=1}^{n} x_i^2\right)^2\right) - \sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \left(n \sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2\right) + \sum_{i=1}^{n} \left(x_i^2 y_i - \frac{1}{6}x_i^5\right) \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right) \end{bmatrix}$$

Hence,

$$a = \frac{1}{det(X^T X)} \Big[\sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^3 - \left(\sum_{i=1}^{n} x_i^2\right)^2\right) - \sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \left(n \sum_{i=1}^{n} x_i^3 -\right.$$

$$\left. \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^2\right) + \sum_{i=1}^{n} \left(x_i^2 y_i - \frac{1}{6}x_i^5\right) \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\Big]$$

$$b = \frac{1}{det(X^T X)} \Big[-\sum_{i=1}^{n} \left(x_i y_i - \frac{1}{6}x_i^4\right) \left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i^4 - \sum_{i=1}^{n} x_i^3 \sum_{i=1}^{n} x_i^2\right) + \sum_{i=1}^{n} \left(x_i y_i -\right.$$

$$\left. \frac{1}{6}x_i^4\right) \left(n \sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^2\right)^2\right) - \sum_{i=1}^{n} \left(x_i^2 y_i - \frac{1}{6}x_i^5\right) - \left(n \sum_{i=1}^{n} x_i^3 - \sum_{i=1}^{n} x_1 \sum_{i=1}^{n} x_i^2\right)\Big]$$

$$c = \frac{1}{det(X^T X)} \Big[\sum_{i=1}^{n} \left(y_i - \frac{1}{6}x_i^3\right) \left(\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} x_i^4 - \left(\sum_{i=1}^{n} x_i^3\right)^2\right) - \sum_{i=1}^{n} \left(x_i y_i -\right.$$

$$\frac{1}{6}x_i^4\right)\left(\sum_{i=1}^n x_i \sum_{i=1}^n x_i^4 - \sum_{i=1}^n x_i^2 \sum_{i=1}^n x_i^3\right) + \sum_{i=1}^n \left(y_i - \frac{1}{6}x_i^3\right)\left(\sum_{i=1}^n x_i \sum_{i=1}^n x_i^3 - \left(\sum_{i=1}^n x_i^2\right)^2\right)\Bigg]$$

Therefore, $\hat{g} = X\hat{\beta}$

Appendix (code)

```
set.seed(1)
x1 = runif(100)
x2 = 0.5*x1 + rnorm(100)/10
y = 2 + 2*x1 + 0.3*x2 + rnorm(100)

plot(x1,x2, main ='The scatterplot of x1 and x2', xlab = 'x1', ylab = 'x2',pch = 20)
leastsquare = lm(y~x1+x2)
summary(leastsquare)

onlyx1 = lm(y~x1)
summary(onlyx1)

onlyx2 = lm(y~x2)
summary(onlyx2)

x11 = c(x1, 0.1)
x21 = c(x2, 0.8)
y2 = c(y, 6)

newleast = lm(y2~x11+x21)
summary(newleast)
resall = resid(newleast)
plot(y2,resall,xlab = 'y', ylab = 'residual', main = 'residual plot', pch = 20)
X = cbind(x11,x21)
XTX_1 = solve(t(X)%*%X)
xi = matrix(c(0.1,0.8),ncol = 1,nrow =2 )
t(xi)%*%XTX_1%*%xi

newx1 = lm(y2~x11)
summary(newx1)
y_hat = 2.2569 + 1.7657*x1
plot(x1,y,pch = 20, xlab='x1',ylab='y')
points(x = 0.1, y =6, type = 'p',col ='orange')
resx1 = resid(newx1)
plot(y2,resx1,xlab='y',ylab='residual',main = 'residual plot',pch = 20)
xbar1 = mean(x11)
Sxx1 = sum((x11-xbar1)^2)

newx2 = lm(y2~x21)
summary(newx2)
plot(x2,y,pch = 20,xlab='x2',ylab='y',xlim = c(-0.2,1))
points(x = 0.8,y=6, type= 'p',col ='orange')
resx2 = resid(newx2)
plot(y2,resx2,xlab ='y',ylab='residual',main ='residual plot',pch = 20)
xbar2 = mean(x21)
Sxx2 = sum((x21-xbar2)^2)
```