

EEB 581, Problem Set Five

Solutions

1 : Consider quadratic regression forced through the origin, $y_i = \beta_1 x_i + \beta_2 x_i^2 + e$.

(a) For n observations, write this in matrix form.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \text{where } \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_1 & x_1^2 \\ \vdots & \vdots \\ x_n & x_n^2 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

(b) What is the OLS estimator for β_1 and β_2 .

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

$$\mathbf{X}^T \mathbf{X} = \begin{pmatrix} \sum x_i^2 & \sum x_i^3 \\ \sum x_i^3 & \sum x_i^4 \end{pmatrix}, \quad (\mathbf{X}^T \mathbf{X})^{-1} = \frac{1}{(\sum x_i^2)(\sum x_i^4) - (\sum x_i^3)^2} \begin{pmatrix} \sum x_i^4 & -\sum x_i^3 \\ -\sum x_i^3 & \sum x_i^2 \end{pmatrix}$$

$$\mathbf{X}^T \mathbf{y} = \begin{pmatrix} \sum x_i y_i \\ \sum x_i^2 y_i \end{pmatrix}$$

putting these together,

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = \frac{1}{(\sum x_i^2)(\sum x_i^4) - (\sum x_i^3)^2} \begin{pmatrix} \sum x_i^4 & -\sum x_i^3 \\ -\sum x_i^3 & \sum x_i^2 \end{pmatrix} \begin{pmatrix} \sum x_i y_i \\ \sum x_i^2 y_i \end{pmatrix}$$

$$= \frac{1}{(\sum x_i^2)(\sum x_i^4) - (\sum x_i^3)^2} \begin{pmatrix} (\sum x_i^4)(\sum x_i y_i) - (\sum x_i^3)(\sum x_i^2 y_i) \\ -(\sum x_i^3)(\sum x_i y_i) + (\sum x_i^2)(\sum x_i^2 y_i) \end{pmatrix}$$

(c) What is the variance-covariance matrix for these estimates?

$$(\mathbf{X}^T \mathbf{X})^{-1} = \frac{\sigma_e^2}{(\sum x_i^2)(\sum x_i^4) - (\sum x_i^3)^2} \begin{pmatrix} \sum x_i^4 & -\sum x_i^3 \\ -\sum x_i^3 & \sum x_i^2 \end{pmatrix}, \quad \sigma_e^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n-2}$$

2 : Suppose you have the following data for 50 observations

$$\sum_i x_i^2 = 300, \quad \sum_i x_i^3 = 100, \quad \sum_i x_i^4 = 12000, \quad \sum_i x_i y_i = -200, \quad \sum_i x_i^2 y_i = 600$$

(a) Compute the OLS estimate of β_1 and β_2 .

```
> XY <- matrix(c(-200,600),nrow=2)
> X <- matrix(c(300,100,100,12000),nrow=2)
> solve(X)%*%XY
```

R returns

```
      [,1]
[1,] -0.68523677
[2,]  0.05571031
```

(b) Suppose $\sum (y_i - \hat{y}_i)^2 = 400$. Estimate σ_e^2

$$\sigma_e^2 = 400/(50 - 2) = 8.33$$

(c) Compute $\sigma^2(\hat{\beta}_1)$.

$$\text{cov}(\beta) = \sigma_e^2 (\mathbf{X}^T \mathbf{X})^{-1}$$

In R

```
> 8.33* solve(X)
```

R returns

```
          [,1]          [,2]
[1,] 0.0278440111 -0.0002320334
[2,] -0.0002320334  0.0006961003
```

Hence $\sigma^2(\hat{\beta}_1) = 90.0278$

(d) Compute $\sigma^2(\hat{\beta}_2)$.

$$\sigma^2(\hat{\beta}_1) = 0.000696$$

(e) Compute $\sigma(\hat{\beta}_1, \hat{\beta}_2)$

$$\sigma(\hat{\beta}_1, \hat{\beta}_2) = -0.000232$$