SIAI
Swiss Institute of Artificial Intelligence

Name: Hye-Young, Park
Class: STA502: Math & Stat for MBA
Date: 2021/09/06
Problem Set 1 - Question 1

QUESTIONS	***NOTES***

Question 1.

1. Obtain formulae for $\hat{\beta}_1$ and $\hat{\beta}_2$, the least squares estimators, from the general result: $\hat{\beta} = (X'X)^{-1}X'y$

Least Squares estimator [최소자승법]

Aim : To create a straight line that minimizes them sum of the squares of the errors(ϵ_t)

즉, 잔차 제곱합(RSS, Residual sum of squares)인 $\Sigma(e_t)^2 = \Sigma(y_t - \beta_1 - \beta_2 x_t)^2$ 이 최소가 되도록 편미분 = 0을 만드는 β_1, β_2 를 찾는 방법

Source : https://www.investopedia.com/terms/l/least-squares-method.asp

Ordinary Least Squares [추정량]

1) $\sum_{i=1}^{n} \hat{e}_i = 0$: 잔차의 총합은 0 이다.

2) $Cov(x_i, \hat{e}_i) = \sum_{i=1}^{n} x_i \hat{e}_i = 0$: 설명 변수와 잔차는 무관하다.

Source :

$y_t = \beta_1 + \beta_2 x_t$ 을 행렬로 표현하면 다음과 같다.

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_2 \end{pmatrix}$$

최소자승법과 유사 역행렬(pseudo inverse)를 이용하여 β 을 구하면

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

Pseudo inverse [유사 역행렬]

역행렬은 정사각행렬에만 존재하는데, 정사각 행렬이 아닌 행렬에서 역행렬과 비슷한 역할을 하는 행렬이 필요할 때 쓴다. 어떤 mxn실수 행렬 X에 대해서 다음 4가지 조건을 만족하는 행렬 X^+를 무어-펜로즈 유사 역행렬이라고 한다. 유사역행렬은 전치 행렬과 모양이 같다.

(1) $X X^+ X = X$
(2) $X^+ X X^+ = X^+$
(3) $(X X^+)^T = X X^+$
(4) $(X^+ X)^T = X^+ X$

Source :

Transposed matrix [전치 행렬]

A = [aij] 일 때, A^T = [aji]

Source :

QUESTIONS	***NOTES***
	Solution.

$$X' = \begin{pmatrix} i' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_t \end{pmatrix} = X^T$$

$$X'X = \begin{pmatrix} i' \\ x' \end{pmatrix} \begin{pmatrix} i & x \end{pmatrix} = \begin{pmatrix} i'i & i'x \\ x'i & x'x \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_t \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_t \end{pmatrix} = \begin{pmatrix} T & \sum_{t=1}^{T} x_t \\ \sum_{t=1}^{T} x_t & \sum_{t=1}^{T} x_t^2 \end{pmatrix}$$

$$X'y = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_t \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_t \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} x_t y_t \end{pmatrix} = \begin{pmatrix} \sum_{t=1}^{T} y_t & \sum_{t=1}^{T} x_t y_t \end{pmatrix}^T$$

It follows that :

Inverse Matrix [역행렬]

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} (|A| = ad - bc)$$

$$(X'X)^{-1} = \frac{1}{|X'X|} \begin{pmatrix} \sum_{t=1}^{T} x_t^2 & -\sum_{t=1}^{T} x_t \\ -\sum_{t=1}^{T} x_t & T \end{pmatrix}$$

그래서 문제 1-1 을 풀어보면,

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \frac{1}{|X'X|} \begin{pmatrix} \sum_{t=1}^{T} x_t^2 & -\sum_{t=1}^{T} x_t \\ -\sum_{t=1}^{T} x_t & T \end{pmatrix} \begin{pmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} x_t y_t \end{pmatrix} = \frac{1}{|X'X|} \begin{pmatrix} \sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} x_t y_t \\ T \sum_{t=1}^{T} x_t y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} y_t \end{pmatrix}$$

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} x_t y_t}{T \sum_{t=1}^{T} x_t^2 - (\sum_{t=1}^{T} x_t)^2}, \quad \hat{\beta}_2 = \frac{T \sum_{t=1}^{T} x_t y_t - \sum_{t=1}^{T} x_t \sum_{t=1}^{T} y_t}{T \sum_{t=1}^{T} x_t^2 - (\sum_{t=1}^{T} x_t)^2}$$

QUESTIONS	*NOTES*
Question 1. **2.** Obtain formuale for $var(\hat{\beta}_1), var(\hat{\beta}_2)$ and $covar(\hat{\beta}_1, \hat{\beta}_2)$ from the general OLS result: $V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$	Solution. $(X'X)^{-1} = \frac{1}{\lvert X'X \rvert} \begin{pmatrix} \sum_{t=1}^{T} x_t^{\,2} & -\sum_{t=1}^{T} x_t \\ -\sum_{t=1}^{T} x_t & T \end{pmatrix}$ $V(\hat{\beta}) = \frac{\sigma_\varepsilon^{\,2}}{\lvert X'X \rvert} \begin{pmatrix} \sum_{t=1}^{T} x_t^{\,2} & -\sum_{t=1}^{T} x_t \\ -\sum_{t=1}^{T} x_t & T \end{pmatrix} = \frac{\sigma_\varepsilon^{\,2}}{T \sum_{t=1}^{T} (x_t - \bar{x})^2} \begin{pmatrix} \sum_{t=1}^{T} x_t^{\,2} & -\sum_{t=1}^{T} x_t \\ -\sum_{t=1}^{T} x_t & T \end{pmatrix}$ $= \frac{\sigma_\varepsilon^{\,2}}{\sum_{t=1}^{T} (x_t - \bar{x})^2} \begin{pmatrix} \frac{1}{T} \sum_{t=1}^{T} x_t^{\,2} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}$

Name: Hye-Young, Park
Class: STA502: Math & Stat for MBA
Date: 2021/09/06
Problem Set 1 - Question 2

QUESTIONS	*NOTES*

Question 2.

1. Test the hypothesis that the electricity industry displays constant returns to scale against the alternative that it displays increasing returns at the 5% significance level.

Returns to Scale [규모에 대한 수익(보수)]

"모든 생산요소를 똑같은 비율로 변동시킬 때, 생산량이 어떻게 변하는가?"에 대한 생산함수

$$q = f(K, L) = AK^{\alpha} L^{\beta} e^{\varepsilon}$$

$$\ln q = \ln A + \alpha \ln K + \beta \ln L + \varepsilon \ln e$$

q는 생산량, K(자본), L(노동)

Source : https://support.sas.com/rnd/app/ets/examples/cobbdoug/index.html

---

Significance level [유의 수준,알파]

A measure of the strength of the evidence that must be present in your sample before you will reject the H_0 and conclude that the effect is statically significant.

실험(test) 전에 유의 수준을 정해야 한다. 유의 수준이란, 참 임에도 불구하고 귀무 가설(H_0)을 기각하는 식의 잘못된 판단을 내릴 위험성을 의미한다.

Source : https://statisticsbyjim.com/glossary/significance-level/

문제의 Cost function 을 위의 생산함수를 바탕으로 이해해보면,

$$Cost = e^{\beta_1} A^{\beta_2} B^{\beta_3} C^{\beta_4} D^{\beta_5}$$

$$\log(Cost) = \beta_1 + \beta_2 \log A + \beta_3 \log B + \beta_4 \log C + \beta_5 \log D$$

$$y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + \beta_4 x_{3i} + \beta_5 x_{4i} + \hat{\varepsilon}$$

A: Output, B: Price of Labor, C: Cost of Capital, D: Price of Fuel

---

Constant return : 각 요소가 n 배 커지면 비용도 n 배 커짐 $\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 = 0$

Increasing return : 각 요소가 n 배 커지면 비용도 n 배 보다 더 커짐

$$\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 > 0$$

Decreasing return : 각 요소가 n 배 커지면 비용도 n 배 보다 더 작아짐

$$\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 < 0$$

즉 2-1 문제는 위의 회귀 모델이 constant return 이 맞는지 검증(test)해보는 문제라고 이해할 수 있다. 그래서 $\tau < significance - level = 0.05$ 이라면, Output 에 대해선 constant return 임을 뜻하는 귀무 가설 H0: $\beta_2 = 1$ 이 기각되면서 non-constant return 이 된다.

QUESTIONS

Covariance matrix [공분산 행렬]

데이터 구조를 설명해준다. 특징 쌍(feature pairs)들의 변동이 얼마나 닮았는가, 즉 각 변수가 얼마만큼 함께 변하는가를 알 수 있다.

$$\sum = \frac{1}{n} X^T X$$

Source :

$$\tau = \frac{(\hat{\beta}_2 - 1)}{SE(\hat{\beta}_2)} = \frac{(\hat{\beta}_2 - 1)}{\sqrt{\hat{V}_2}} = \frac{(0.720 - 1)}{\sqrt{0.305e^{-3}}} = -16.0327...$$

유의 수준보다 작으므로 H0 이 기각된다, 즉 constant return 이 아니다.

(임계값 $t_{(n=5,\alpha)} = t*_{5\%}(140) = -1.655$ 는 아직 이해를 다 못했습니다.)

즉, decreasing returns 이라고 할 수 있다.

2. We would expect the cost function to be homogeneous of degree one in prices. What does this imply for the coefficients of the cost function? Test the hypothesis that the cost function is homogeneous of degree one (i.e. produce just as much input invested) against the alternative that it is not at the 5% significance level

공분산 행렬을 살펴보면,

$$\sum(X) = Cov(X) = E(X-\mu)^T (X-\mu)$$

$$= \frac{1}{n} \begin{pmatrix} E_1 - \mu_1 \\ E_2 - \mu_2 \\ E_3 - \mu_3 \\ E_4 - \mu_4 \\ E_5 - \mu_5 \end{pmatrix} \begin{pmatrix} E_1 - \mu_1 & E_2 - \mu_2 & E_3 - \mu_3 & E_4 - \mu_4 & E_5 - \mu_5 \end{pmatrix}$$

$$= \begin{pmatrix} V_1 & & & & \\ V_{12} & V_2 & & & \\ V_{13} & V_{23} & V_3 & & \\ V_{14} & V_{24} & V_{34} & V_4 & \\ V_{15} & V_{25} & V_{35} & V_{45} & V_5 \end{pmatrix}$$

$$= \begin{pmatrix} 3.148 & & & & \\ -0.437e^{-2} & 0.305e^{-3} & & & \\ -0.141 & -0.455e^{-3} & 0.0847 & & \\ 0.591 & 0.323e^{-3} & 0.0237 & 0.115 & \\ 0.715e^{-2} & 0.315e^{-3} & -0.0109 & -0.663e^{-2} & 0.0101 \end{pmatrix}$$

Prices 와 관련된 설명 변수인 B: Price of Labor, C: Cost of Capital, D: Price of Fuel 에 관련해서, 동차함수라는 가정을 검증해보고자 한다.

	Homogeneous of degree 0 [동차함수] A property of an equation the exists if independent variables are increased by a constant value, then the dependent variable is increased by the value raised to the power of 0. Source : https://ko.wikipedia.org/wiki/%EB%8F%99%EC%B0%A8%ED%95%A8%EC%88%98 $\lvert\tau\rvert = \frac{\lvert\hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5 - 1\rvert}{SE(\hat{\beta}_3 + \hat{\beta}_4 + \hat{\beta}_5)} = \frac{(0.436 + 0.220 + 0.427 - 1)}{\sqrt{\hat{V}_3 + \hat{V}_4 + \hat{V}_5 + 2(\hat{V}_{34} + \hat{V}_{35} + \hat{V}_{45})}} = -0.7575... < 1.977$ 단, 임계값 $t_{(n-5,\frac{\alpha}{2})} = t*_{2.5\%(140)} = 1.977$ 이므로, H0 는 기각되지 않는다. 즉 constant returns 이라고 할 수 있다.

Name: Hye-Young, Park
Class: STA502: Math & Stat for MBA
Date: 2021/09/06
Problem Set 1 - Question 3

QUESTIONS	*NOTES*

Question 2.

1. Construct 80% prediction intervals for the dependent variable *y* for observations 12 and 13.

Out-of-sample prediction(forecast) [기간 외 샘플 예측]

- in-sample: 기간 내 샘플로, 초기 매개변수 추정과 모델 선택에 이용됨.
- out-of-sample: 기간 외 샘플로, 예측성능평가에 이용됨

Prediction interval [예측구간]

An estimate of an interval in which afuture obsercation will fall with acertain probability.

예측구간이 95%이면, 새 관측치가 이 범위 안에 속살 것이라고 95% 확신한다는

$\varepsilon_t \sim iid, N(0, \sigma^2_{\ \varepsilon})$ 이면, h단계 예측값에 대한 95% 예측구간은

$\hat{y}_{t+h|t} \pm 1.96\hat{\sigma}_h$

$\hat{\sigma}_h$ 는 h단계 예측분포(forecast distribution) 표준편차추정값이다.

예를 들어, 관측한 시계열의 마지막 y값이 531.48 이고, 표준편차가 6.21이면

95% 예측구간은 [531.48-1.96*6.21, 531.48+1.96*6.21]=[519.3, 543.6] 이다.

80% 예측구간은 $\hat{y}_{t+h|t} \pm 1.28\hat{\sigma}_h$

Source1 : https://statisticsbyjim.com/glossary/prediction-intervals/

Source2 : https://otexts.com/fppkr/prediction-intervals.html

---

11 번째까지의 관측값이 있고, 12 번째 및 13 번째 관측값(observation)에 대한 예측구간 80%를 설정하라는 문제이므로,

12 번째 및 13 번째 관측값에 대한 80% 예측구간 : $[\hat{y} - 1.28\hat{\sigma}, \hat{y} + 1.28\hat{\sigma}]$

$$y'y = Y^T Y = [y_1 \quad \cdots \quad y_{11}] \begin{bmatrix} y_1 \\ \vdots \\ y_{11} \end{bmatrix} = \sum_{t=1}^{11} y_t^{\ 2} = \frac{7}{3}$$

정의에 의해서,

$$(y_i - \bar{Y}) = (\hat{y}_i - \bar{Y}) + e_i$$

총 제곱합(SST) = 회귀 제곱합(SSR) + 잔차 제곱합(SSE) 이므로,

$$\sum (y_i - \bar{Y})^2 = \sum (\hat{y}_i - \bar{Y})^2 + \sum e_i^{\ 2}$$

$$SSE = \sum_{i=1}^{11} e_i^{\ 2} = \sum_{i=1}^{11} (y_i - \hat{y})^2$$

첫번째 문제에서

$y_i = \beta_1 x_{1i} + \beta_2 x_{2i}$ 을 행렬로 표현하면 다음과 같다.

$$\begin{pmatrix} x_{11} & x_{21} \\ \vdots & \vdots \\ x_{1i} & x_{2i} \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_2 \end{pmatrix}$$

최소자승법과 유사 역행렬(pseudo inverse)를 이용하여 β 을 구하면

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

$$X^T X = \begin{pmatrix} x_{11} & \cdots & x_{1i} \\ x_{21} & \cdots & x_{2i} \end{pmatrix}\begin{pmatrix} x_{11} & x_{21} \\ \vdots & \vdots \\ x_{1i} & x_{2i} \end{pmatrix} = \begin{pmatrix} \sum x_{1i}{}^2 & \sum x_{1i}x_{2i} \\ \sum x_{1i}x_{2i} & \sum x_{2i}{}^2 \end{pmatrix}$$

$$= \begin{pmatrix} x_1'x_1 & x_1'x_2 \\ x_1'x_2 & x_2'x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$X'y = \begin{pmatrix} x_{11} & \cdots & x_{1i} \\ x_{21} & \cdots & x_{2i} \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_i \end{pmatrix} = \begin{pmatrix} \sum_i x_{1i} y_t \\ \sum_i x_{2i} y_t \end{pmatrix} = \begin{pmatrix} x_1'y_1 \\ x_2'y_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$(X^T X)^{-1} = \frac{1}{2\cdot 2 - 1\cdot 1}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = (X^T X)^{-1} X^T y = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\therefore \beta_1 = 1, \beta_2 = 0$$

$$\hat{y} = \beta_1 x_1 + \beta_2 x_2 = x_1(\beta_1 = 1, \beta_2 = 0))$$

12 번째 y 의 80% 예측구간

$$\hat{y} = x_1 = 5$$

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum e_i{}^2 = \frac{1}{9}(\frac{7}{3} + 11\times 25) = 30.81481$$

$$\hat{\sigma} = 5.551109 = 5.55$$

$$[\hat{y} - 1.28\hat{\sigma}, \hat{y} + 1.28\hat{\sigma}] = [-2.11, 12.11]$$

13 번째 y 의 80% 예측구간

$$\hat{y} = x_1 = 3$$

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum e_i{}^2 = \frac{1}{n-2}\sum_{i=1}^{11}(y_i - \hat{y})^2 = \frac{1}{9}(\frac{7}{3} + 11\times 9) = 11.25926$$

$$\hat{\sigma} = 3.355482 = 3.36$$

$$[\hat{y} - 1.28\hat{\sigma}, \hat{y} + 1.28\hat{\sigma}] = [-1.30, 7.30]$$

2. Construct 80% prediction intervals for the expected value of y_{12} and y_{13}	12 번째와 13 번째 y 의 expected value = 4 $\hat{\sigma}^2 = \frac{1}{n-2}\sum e_i^2 = \frac{1}{9}(\frac{7}{3}+11\times 16) = 19.81481$ $\hat{\sigma} = 4.451383 = 4.45$ $[\hat{y}-1.28\hat{\sigma}, \hat{y}+1.28\hat{\sigma}] = [-1.70, 8.30]$
3. We would expect the cost function to be homogeneous of degree one in prices. What does this imply for the coefficients of the cost function? Test the hypothesis that the cost function is homogeneous of degree one (i.e. produce just as much input invested) against the alternative that it is not at the 5% significance level	다르다. (Differ) $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = (X^T X)^{-1} X^T y = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\therefore \beta_1 = 1, \beta_2 = 0$ $\beta_2 = 0$ 인 것을 보아, 설명 변수를 잘못 설정했던 게 오류로 이어진 것으로 보인다.