A NOTE ON THE PROOF OF THE RELATION

\[ X'(X'X)^{-1}X = 1/n \]

BY HEE-KYUN CHOI

1. Introduction.

It is well known that in the multiple regression model

\[ Y_i = B_1 + B_2X_{2i} + \cdots + B_kX_{ki} + u_i \quad i = 1, \ldots, n \]  (1)

\[ E(u_i) = 0 \quad \text{for all} \ i \]  (2)

\[ E(u_iu_j) = \begin{cases} \sigma^2 & i = j: i, j = 1, \ldots, n \\ 0 & i \neq j: i, j = 1, \ldots, n \end{cases} \]  (3)

the error variance of predicting the expected value of \( Y \) associated with \( X_0 \) is given by

\[ \text{Var}(\hat{Y}_0 - E(Y_0)) = \sigma^2 \ (X_0'(X'X)^{-1}X_0) \]  (4)

where \( \hat{Y}_0 \) is the prediction of \( Y \) at \( X_0 \) and

\[ X_0 = \begin{pmatrix} X_{20} \\ X_{30} \\ \vdots \\ X_{k0} \end{pmatrix}, \quad X = \begin{pmatrix} 1 & X_{21} & \cdots & X_{2k} \\ 1 & X_{31} & \cdots & X_{3k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{k1} & \cdots & X_{kn} \end{pmatrix} \]

The matrix \( X \) on \( n \) sample observations on the \( k \) independent variables with \( X_{1i} = 1 \) has rank \( k < n \) and the vector \( X_0 \) is the particular values of the independent variables to predict \( Y \).

Desalvo [1] has shown that error variance at the sample mean of \( X \) is given by

\[ \text{Var}(\hat{Y}_0 - E(Y_0))_{X_0=x} = \sigma^2(X'(X'X)^{-1}X) = \sigma^2/n \]  (5)

as proving the relation

\[ X'(X'X)^{-1}X = 1/n \]  (6)

without statistical property as will be shown below.
<A summary of Desalvo's proof>

\[ X'(X'X)^{-1}X = \frac{1}{\|X'X\|^2} \sum_{i=1}^{n} \sum_{j=1}^{n} X_{ri} X_{rj} c_{ij} \]  

\[ (C \text{ is the cofactor matrix}) \]

When \( r=1 \) in the right side of (7)

\[ \sum_{i=1}^{n} X_{ri} \sum_{j=1}^{n} \left( \sum_{i=1}^{n} X_{ri} \right) c_{ij} = n \|X'X\| \]  

while \( r \neq 1 \),

\[ \sum_{i=1}^{n} X_{ri} \sum_{j=1}^{n} \left( \sum_{i=1}^{n} X_{ri} \right) c_{ij} = 0, \]

hence

\[ X'(X'X)^{-1}X = 1/n. \]

Now we will show a simpler method for the proof of the relation (6) with such statistical property as the least-square method.

2. An alternative proof.

A consequence of the least-squares fit is that the sum of the residuals is zero.

That is

\[ \sum_{i=1}^{n} e_i = 1'(I_n - X(X'X)^{-1}X')U = \sum_{i=1}^{n} a_i u_i = 0, \]

where \( u_i \) and \( a_i \) are the \( i \)th component of the disturbance vector \( U \) and \( A' \)

\( (=I_n - X(X'X)^{-1}X') \) respectively and \( \mathbf{1} \) denotes the vector of \( n \) unities.

Then it is clear that \( u_1, u_2, \ldots, u_n \) are linearly independent and every \( a_i = 0 \) by (3) and (9) respectively. i.e. \( A' = 0 \).

Since \( I_n - X(X'X)^{-1}X' \) is symmetric and idempotent [2] and \( X = \frac{X'1}{n} \)

we obtain that

\[ 0 = 1'(I_n - X(X'X)^{-1}X') (I_n - X(X'X)^{-1}X')'1 \]
\[ = 1'(I_n - X(X'X)^{-1}X')1 \]
\[ = 1'I_n - 1'X(X'X)^{-1}X'1 \]
\[ = n - nX(X'X)^{-1}Xn \]
A note on the proof of the relation $\bar{X}'(X'X)^{-1}\bar{X} = 1/n$

hence

$\bar{X}'(X'X)^{-1}\bar{X} = 1/n$.

This final result is the same as that of Desalvo.

References


Ajou Institute of Technology