S631 HW5

Erik Parker

September 23, 2017

1. Show the following equalities:

a) $\sum_{i=1}^{n}(y_i - x_i^T\hat{\beta})^2 = (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$

From lecture, we know that y_i represents a single observed value of a random variable, and there are 1:n of these observed values. Furthermore, we know that x_i^T is the transpose of all regressors used to construct a linear model in relation to our observed y_i, with 1:p different regressors for the 1:n observations. Finally, $\hat{\beta}$ is a vector of unknown regression coefficients ($\boldsymbol{\beta} = \beta_0, \beta_1, ..., \beta_p$) related to the regression of our x_i's on y_i. With the summation of these terms, we are able to combine all of the individual values into matrices of dimension $nX1$ (for $\boldsymbol{Y}$), $nX(p+1)$ (for $\boldsymbol{X}$), and $(p+1)X1$ (for $\boldsymbol{\beta}$) as we are now dealing not with the individual values but their summations.

With this, we can then see that the RHS of this equality, when expressed with the dimensions of the matrices, becomes

$$(\boldsymbol{Y}_{nX1} - \boldsymbol{X}_{nX(p+1)}\hat{\boldsymbol{\beta}}_{(p+1)Xn})^T(\boldsymbol{Y}_{nX1} - \boldsymbol{X}_{nX(p+1)}\hat{\boldsymbol{\beta}}_{(p+1)Xn})$$

which when simplified becomes

$$(\boldsymbol{Y}_{1Xn}^T - \hat{\boldsymbol{\beta}}\boldsymbol{X}_{1Xn})(\boldsymbol{Y}_{nX1} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{nX1})$$

and then

$$(\boldsymbol{Y}^T - \hat{\boldsymbol{\beta}}\boldsymbol{X})_{1Xn}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})_{nX1}$$

From this we can see that the final product of the RHS is a 1X1 scalar. This means that the final product of the LHS also must be a 1X1 scalar.

Furthermore, from chapter 3.4.2 of ALR and our class notes, we know that $y_i = x_i^T\beta + e_i$ and so, $y_i - x_i^T\hat{\beta}$ is equal to the random vector of errors e_i and that the summation of these errors is a $nX1$ matrix where $\boldsymbol{e} = (e_1, ..., e_n)^T$.
This means

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

can be written as

$$\sum_{i=1}^{n}{e_i}^2 = \boldsymbol{e}^2$$

and in order to conclude with a 1X1 scalar as the final product, we can then express this as

$$\boldsymbol{e}^2 = \boldsymbol{e}_{1Xn}^T\boldsymbol{e}_{nX1} \quad and \ \ so \quad \boldsymbol{e}^T\boldsymbol{e} = (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$$

therefore

$$\sum_{i=1}^{n}(y_i - x_i^T\hat{\beta})^2 = (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$$

b. $\boldsymbol{H}$ is symmetric and idempotent

First, recall that $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X^T X})^{-1}\boldsymbol{X}^T$. Also, for $\boldsymbol{H}$ to be symmetric means $\boldsymbol{H} = \boldsymbol{H}^T$, and to be idempotent means $\boldsymbol{HH} = \boldsymbol{H}$.

Now, to show $\boldsymbol{H}$ is symmetric:

$$\begin{aligned}
\boldsymbol{H}^T &= \boldsymbol{H} \\
(\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T)^T &= \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \\
proving\ \ the\ \ left\ \ side & \\
\boldsymbol{X}((\boldsymbol{X}^T\boldsymbol{X})^{-1})^T\boldsymbol{X}^T & \\
\boldsymbol{X}(\boldsymbol{X}^{-1}(\boldsymbol{X}^{-1})^T)^T\boldsymbol{X}^T & \\
\boldsymbol{X}(\boldsymbol{X}^{-1}(\boldsymbol{X}^T)^{-1})\boldsymbol{X}^T & \\
\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T &= \boldsymbol{H} \\
\blacksquare &
\end{aligned}$$

Now, to show $\boldsymbol{H}$ is idempotent:

$$\begin{aligned}
\boldsymbol{H}^2 &= \boldsymbol{H} \\
proving\ \ the\ \ left\ \ side & \\
(\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T_{p+1Xn})(\boldsymbol{X}_{nXp+1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T) & \\
\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{p+1Xp+1}(\boldsymbol{X}^T\boldsymbol{X})_{p+1Xp+1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T & \\
using\ \ \ AA^{-1} = I\ \ \ from\ \ matrix\ \ handout & \\
\boldsymbol{X}_{nXp+1}I_{p+1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T & \\
using\ \ \ A_{nXm}I_m = A_{nXm} & \\
\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T &= \boldsymbol{H} \\
\blacksquare &
\end{aligned}$$

c. $\boldsymbol{I} - \boldsymbol{H}$ is symmetric and idempotent

Now, to show $\boldsymbol{I} - \boldsymbol{H}$ is symmetric:

$$\begin{aligned}
(\boldsymbol{I} - \boldsymbol{H})^T &= \boldsymbol{I} - \boldsymbol{H} \\
\boldsymbol{I}^T - \boldsymbol{H}^T &= \boldsymbol{I} - \boldsymbol{H} \\
from\ \ part\ \ b\ \ we\ \ know\ \ \ \boldsymbol{H}^T &= \boldsymbol{H} \\
we\ \ also\ \ know\ \ \ \boldsymbol{I}^T &= \boldsymbol{I} \\
so\ \ \ \boldsymbol{I} - \boldsymbol{H} &= \boldsymbol{I} - \boldsymbol{H} \\
\blacksquare &
\end{aligned}$$

Now, to show $\boldsymbol{I} - \boldsymbol{H}$ is idempotent:

$$
\begin{aligned}
(\boldsymbol{I}-\boldsymbol{H})^2 &= \boldsymbol{I}-\boldsymbol{H} \\
proving\ \ the\ \ left\ \ side& \\
(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{I}-\boldsymbol{H})& \\
\boldsymbol{I}^2-\boldsymbol{IH}-\boldsymbol{HI}+\boldsymbol{H}^2& \\
we\ \ know\ \ \ \boldsymbol{I}^2=\boldsymbol{I}\ ,\ \ \boldsymbol{IH}\ \ and\ \ \boldsymbol{HI}=\boldsymbol{H}\ \ as\ \ \boldsymbol{H}_n\ \ and\ \ \boldsymbol{I}_n\ ,\ \ and\ \ from\ \ b\ \ \boldsymbol{H}^2=\boldsymbol{H}& \\
so& \\
\boldsymbol{I}-2\boldsymbol{H}+\boldsymbol{H}& \\
\boldsymbol{I}-\boldsymbol{H} &= \boldsymbol{I}-\boldsymbol{H} \\
&\blacksquare
\end{aligned}
$$

d. $\boldsymbol{HX}=\boldsymbol{X}$**.**

First, recall from part b that $AA^{-1}=I$ and $A_{nXm}I_m=A_{nXm}$.

$$
\begin{aligned}
\boldsymbol{HX} &= \boldsymbol{X} \\
proving\ \ left\ side& \\
(\boldsymbol{X}(\boldsymbol{X^T}\boldsymbol{X})^{-1}\boldsymbol{X}^T)\boldsymbol{X}& \\
\boldsymbol{X}(\boldsymbol{X^T}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{X})& \\
\boldsymbol{XI} &= \boldsymbol{X} \\
&\blacksquare
\end{aligned}
$$

e. $(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}$

For this proof, first recall from class notes, and ALR chapter 9.1 that $\boldsymbol{X}\hat{\boldsymbol{\beta}}=\hat{\boldsymbol{Y}}$. Also recall from problem d that $\boldsymbol{HX}=\boldsymbol{X}$. Finally, recall again that$A_{nXm}I_m=A_{nXm}$.

$$
\begin{aligned}
(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}) &= (\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} \\
proving\ \ the\ \ left\ \ side& \\
\boldsymbol{IY}-\boldsymbol{IX}\hat{\boldsymbol{\beta}}-\boldsymbol{HY}+\boldsymbol{HX}\hat{\boldsymbol{\beta}}& \\
\boldsymbol{IY}-\boldsymbol{I}_n\hat{\boldsymbol{Y}}_{nX1}-\boldsymbol{HY}+\boldsymbol{X}\hat{\boldsymbol{\beta}}& \\
\boldsymbol{IY}-\hat{\boldsymbol{Y}}-\boldsymbol{HY}+\hat{\boldsymbol{Y}}& \\
\boldsymbol{IY}-\boldsymbol{HY}& \\
(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}& \\
&\blacksquare
\end{aligned}
$$

f. $(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}$

First, recall from problem d that $\boldsymbol{HX}=\boldsymbol{X}$, also from problem e that $(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})=(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}$, also from part b $A_{nXm}I_m=A_{nXm}$, finally from ALR chapter 9.1, if $\boldsymbol{X}\hat{\boldsymbol{\beta}}=\hat{\boldsymbol{Y}}$ then $(\boldsymbol{X}\hat{\boldsymbol{\beta}})^T=\hat{\boldsymbol{Y}}^T$.

$$
\begin{aligned}
(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}) &= \boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} \\
\textit{proving LHS} & \\
(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} & \\
(\boldsymbol{Y^T I}-\boldsymbol{Y^T H}-\hat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X^T I}+\hat{\boldsymbol{\beta}}^{\boldsymbol{T}}\boldsymbol{X^T H})\boldsymbol{Y} & \\
(\boldsymbol{Y^T I}-\boldsymbol{Y^T H}-\hat{\boldsymbol{Y}}^T_{1Xn}\boldsymbol{I}_n+\hat{\boldsymbol{\beta}}^T\boldsymbol{X}^T)\boldsymbol{Y} & \\
(\boldsymbol{Y}^T\boldsymbol{I}-\boldsymbol{Y}^T\boldsymbol{H}-\hat{\boldsymbol{Y}}^T+\hat{\boldsymbol{Y}}^T)\boldsymbol{Y} & \\
\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} & \\
\blacksquare &
\end{aligned}
$$

g. $RSS(\hat{\beta}) = \boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}$

For this proof, recall that $RSS(\hat{\beta}) = (\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})$, also from ALR 9.1 that $\boldsymbol{Y}-\hat{\boldsymbol{Y}} = (\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y}$, from problem c that $(\boldsymbol{I}-\boldsymbol{H})$ is symmetric and idempotent, and finally from problem f that $\boldsymbol{X}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{Y}}$.

$$
\begin{aligned}
RSS(\hat{\beta}) &= \boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} \\
\textit{proving LHS} & \\
(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}) & \\
((\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y})^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} & \\
\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} & \\
\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} & \\
\boldsymbol{Y}^T(\boldsymbol{I}-\boldsymbol{H})\boldsymbol{Y} & \\
\blacksquare &
\end{aligned}
$$