5 Finite sample tests of linear hypotheses

5.1 Linear hypotheses

5.2 The joint distribution of $\hat{\sigma}^2$ and $\hat{\beta}$

Recall the definition of the variance estimator:

$$\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-k}$$

To express this in terms of the population $\boldsymbol{\varepsilon}$'s examine the following, where we denote the residual maker matrix by $\mathbf{M_X} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}$:

$$\begin{aligned}
(n-k)\hat{\sigma}^2 &= \hat{\varepsilon}'\hat{\varepsilon} \\
&= (\mathbf{M_X y})'\mathbf{M_X y} \\
&= (\mathbf{M_X}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}))'\mathbf{M_X}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\
&= \boldsymbol{\varepsilon}'\mathbf{M_X'}\mathbf{M_X}\boldsymbol{\varepsilon} \qquad (\text{since } \mathbf{M_X X} = \mathbf{0}) \\
&= \boldsymbol{\varepsilon}'\mathbf{M_X}\boldsymbol{\varepsilon} \qquad (\text{since } \mathbf{M_X'}\mathbf{M_X} = \mathbf{M_X}\mathbf{M_X} = \mathbf{M_X})
\end{aligned}$$

Since $\mathbf{M_X}$ is symmetric, it is positive definite when all eigenvalues are positive. Since it is also idempotent, $\mathbf{M_X^2} = \mathbf{M_X}$, all eigenvalues are either zero or one, meaning $\mathbf{M_X}$ is positive semi-definite.[1]

> **Theorem 5.2.1 (Spectral decomposition).** For every $n \times n$ real symmetric matrix, the eigenvalues are real and the eigenvectors can be chosen real and orthonormal. Thus a real symmetric matrix $\mathbf{A}$ can be decomposed as
>
> $$\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}'$$
>
> where $\mathbf{Q}$ is an orthogonal matrix whose columns are the real, orthonormal eigenvectors of $\mathbf{A}$, and $\boldsymbol{\Lambda}$ is a diagonal matrix whose entries are the eigenvalues of $\mathbf{A}$.

The spectral decomposition of $\mathbf{M_X}$ is $\mathbf{M_X} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}'$ where $\mathbf{H}\mathbf{H}' = \mathbf{I_n}$ and $\boldsymbol{\Lambda}$ is diagonal with the eigenvalues of $\mathbf{M_X}$ along the diagonal. Since $\mathbf{M_X}$ is idempotent with rank $n-k$, it has $n-k$ eigenvalues equalling 1 and k eigenvalues equalling 0, so:

$$\boldsymbol{\Lambda} = \begin{bmatrix} \mathbf{I}_{n-k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_k \end{bmatrix}$$

In the normal regression $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I_n}\sigma^2)$, we want to find the distribution of $\mathbf{H}'\boldsymbol{\varepsilon}$. A linear combination of normals is also normal, meaning $\mathbf{H}'\boldsymbol{\varepsilon}$ is normal with mean $\mathbb{E}[\mathbf{H}'\boldsymbol{\varepsilon}] = \mathbf{H}'\mathbb{E}[\boldsymbol{\varepsilon}] = 0$ and variance $\mathrm{Var}(\mathbf{H}'\boldsymbol{e}) = \mathbf{H}'\mathbf{I_n}\sigma^2\mathbf{H} = \sigma^2\mathbf{H}'\mathbf{H} = \mathbf{I_n}\sigma^2$. Thus $\mathbf{H}'\boldsymbol{\varepsilon} \sim N(0, \mathbf{I_n}\sigma^2)$.

Let $\mathbf{u} = \mathbf{H}'\boldsymbol{\varepsilon}$, and partition $\underset{n\times 1}{\mathbf{u}} = \begin{bmatrix} \underset{(n-k)\times 1}{\mathbf{u_1}} \\ \underset{k\times 1}{\mathbf{u_2}} \end{bmatrix}$ where $\mathbf{u_1} \sim N(0, \mathbf{I_n}\sigma^2)$, then we have

[1] Alternatively since $\mathbf{M_X^2} = \mathbf{M_X}$ and $\mathbf{M_X}' = \mathbf{M_X}$, note that $\mathbf{v}'\mathbf{M_X v} = \mathbf{v}'\mathbf{M_X^2 v} = \mathbf{v}'\mathbf{M_X}'\mathbf{M_X v} = (\mathbf{v}'\mathbf{M_X})'(\mathbf{M_X v}) = \|\mathbf{M_X v}\|^2$ for all $\mathbf{v} \in \mathbb{R}^n$.

$$
\begin{aligned}
(n-k)\hat{\sigma}^2 &= \boldsymbol{\varepsilon}'\mathbf{M}_{\mathbf{X}}\boldsymbol{\varepsilon} \\
&= \boldsymbol{\varepsilon}'\mathbf{H}\boldsymbol{\Lambda}\mathbf{H}'\boldsymbol{\varepsilon} \\
&= \mathbf{u}'\begin{bmatrix} \mathbf{I}_{n-k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_k \end{bmatrix}\mathbf{u} \\
&= [\mathbf{u_1}' \;\; \mathbf{u_2}'] \begin{bmatrix} \mathbf{I}_{n-k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_k \end{bmatrix} \begin{bmatrix} \mathbf{u_1} \\ \mathbf{u_2} \end{bmatrix} \\
&= \mathbf{u_1}'\mathbf{u_1}
\end{aligned}
$$

where $\mathbf{u_1}'\mathbf{u_1}$ is the product of two standard normals with dimension $n-k$, thus it is distributed χ^2_{n-k}. Since $\boldsymbol{\varepsilon}$ is independent of $\hat{\boldsymbol{\beta}}$ it follows that $\hat{\sigma}^2$ is independent of $\hat{\boldsymbol{\beta}}$ as well.

Theorem 5.2.2. In normal regression,

$$
\frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k}
$$

and is independent of $\hat{\boldsymbol{\beta}}$.