Name: Hamed Vaheb

High Dimensional Statistics: Exercise Sheet 1

The exercises that I ask to be corrected are: 3, 4

3. Show that the estimator $\hat{\sigma}^2$ is unbiased (i.e., $\mathbb{E}[\hat{\sigma}^2] = \sigma^2$), where

$$\hat{\sigma}^2 = \frac{1}{n-k}\|Y - X\hat{b}\|_1$$

and

$$\hat{b} = (X^T X)^{-1} X^T Y$$

Solution. Approach I. We can express $\hat{\sigma}^2$ as function of a projection matrix R as the following:

$$\begin{aligned}
\hat{\sigma}^2 = \frac{1}{n-k}\|Y - X\hat{b}\|_2^2 &= \|Y - X(X^T X)^{-1} X^T Y\|^2 \\
&= \frac{1}{n-k}\|\underbrace{(I_n - X(X^T X)^{-1} X^T)}_{R} Y\|^2 \\
&= \frac{1}{n-k}\|RY\|^2 = \frac{1}{n-k} Y^T R^T R Y
\end{aligned}$$

Now prove that the matrix $R = I_n - X(X^T X)^{-1} X^T$ is a projection matrix, i.e.,

1. $R^T = R$
2. $R^2 = R$

(a):

$$R^T = (I_n - X(X^T X)^{-1} X^T)^T = I_n - X((X^T X)^{-1})^T X^T = I_n - X((X^T X)^T)^{-1} X^T = I_n - X(X^T X)^{-1} X^T = R$$

(b):

$$\begin{aligned}
R^2 &= (I_n - X(X^T X)^{-1} X^T) \cdot (I_n - X(X^T X)^{-1} X^T) \\
&= I_n - X(X^T X)^{-1} X^T - X(X^T X)^{-1} X^T + X\underbrace{(X^T X)^{-1} X^T X}_{I_n}(X^T X)^{-1} X^T \\
&= I_n - X(X^T X)^{-1} X^T - X(X^T X)^{-1} X^T + X(X^T X)^{-1} X^T \\
&= I_n - X(X^T X)^{-1} X^T = R
\end{aligned}$$

Now we have the following:

$$\hat{\sigma}^2 = \frac{1}{n-k}\|RY\|_2^2 = \frac{1}{n-k} Y^T \underbrace{R^T}_{R} R Y = \frac{1}{n-k} Y^T R^2 Y = \frac{1}{n-k} Y^T R Y$$

Note that $rank(X) = k$ and $rank(R) = n - k$

Now we remind this theorem presented in the lecture:

Theorem. Let $Y \sim \mathcal{N}_n(\mu, \Sigma)$, with $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$. and P, a symmetric matrix with $rank(P) = r$. Then, P is an projection matrix (i.e., $P^2 = P$) $\leftrightarrow Q = \frac{(Y-\mu)^T P (Y-\mu)}{\sigma^2} \sim \chi^2_{(r)}$

Now based on proof of the theorem, we have

$$(n-k)\hat{\sigma}^2/\sigma^2 = \frac{Y^TRY}{\sigma^2} = \frac{(Y-\mu)^TR(Y-\mu)}{\sigma^2} \sim \chi^2_{(n-k)},$$

where $\mu = Xb$.
The last equality comes from the following equations:

$$\begin{aligned}(Y-\mu)^TR(Y-\mu) &= Y^TRY - Y^TR\mu - \mu^TRY + \mu^TR\mu \\ &= Y^TRY - Y^T(I_n - X(X^TX)^{-1}X^T)\mu - \mu^T(I_n - X(X^TX)^{-1}X^T)Y + \mu^TR\mu\end{aligned}$$

For second and third term we have

$$Y^T(I_n-X(X^TX)^{-1}X^T)\mu-\mu^T(I_n-X(X^TX)^{-1}X^T)Y = Y^T(I_n-X(X^TX)^{-1}X^T)Xb-Xb^T(I_n-X(X^TX)^{-1}X^T)Y =$$

$$Y^T(I_n - X(X^TX)^{-1}X^T)Xb - b^TX^T(I_n - X(X^TX)^{-1}X^T)Y = Y^T(Xb - Xb) - (b^TX^T - b^TX^T)Y = 0j$$

And for final term we have,

$$(Y-\mu)^TR(Y-\mu) = Y^TRY - Y^TR\mu - \mu^TRY + \mu^TR\mu$$

$$\mu^TR\mu = b^TX^T(I_n - X(X^TX)^{-1}X^T)Xb = b^T(X^TX - X^TX(X^TX)^{-1}X^TX)b = b^T(X^TX - X^TX)b = 0$$

Consequently,

$$(Y-\mu)^TR(Y-\mu) = Y^TRY$$

As $\hat{\sigma}^2/\sigma^2$ has the Chi-squared distribution, its expectation equates it degree of freedom, i.e.,

$$\mathbb{E}[(n-k)\hat{\sigma}^2/\sigma^2] = n-k$$

which implies

$$\mathbb{E}[\hat{\sigma}^2] = \sigma^2$$

Approach II.

We already showed in previous approach that

$$\hat{\sigma}^2 = \frac{1}{n-k}Y^TRY$$

and that

$$\mu^TR\mu = 0$$

Then if we use exercise 1.1, we can deduce that

$$\mathbb{E}[\hat{\sigma}^2] = \frac{1}{n-k}\mathbb{E}[Y^TRY] = \frac{1}{n-k}(\mu^TR\mu + tr(R\Sigma^2)) = \frac{1}{n-k}tr(R\Sigma^2)$$

$$tr(R\Sigma) = tr(R\sigma^2I_n) = \sigma^2tr(RI_n) = \sigma^2tr(I_n - X(X^TX)^{-1}X^T) = \sigma^2(\underbrace{tr(I_n)}_{\text{n}} - tr(\underbrace{X(X^TX)^{-1}X^T}_{B})$$

B is an idempotnet matrix, i.e., multiplying by itself yields itself. Therefore its rank equates its trace, i.e., $tr(B) = rank(B) = k$. Using this, we will have the following:

$$tr(R\sigma^2) = \sigma^2(n-k)$$

We can conclude that

$$\mathbb{E}[\hat{\sigma^2}] = \frac{1}{n-k}(\sigma^2)(n-k) = \sigma^2$$