STAT 460/560- Assignments

Week V:

1. Assume the linear regression model as in lecture notes. The hat matrix is defined to be
$$\mathbf{H}_n = \mathbf{Z}_n(\mathbf{Z}_n^\tau \mathbf{Z}_n)^{-1}\mathbf{Z}_n^\tau,$$
and the residual of the fit is $\hat{\boldsymbol{\epsilon}}_n = (\mathbb{I}_n - \mathbf{H}_n)\mathbf{y}_n$.

   (a) Prove that
$$\mathbf{y}_n^\tau(\mathbb{I}_n - \mathbf{H}_n)\mathbf{y}_n = \hat{\boldsymbol{\epsilon}}_n^\tau \hat{\boldsymbol{\epsilon}}_n.$$

   (b) The fitted value of $\mathbf{y}_n$ is defined to be $\hat{\mathbf{y}}_n = \mathbf{Z}_n\hat{\boldsymbol{\beta}}_n$. Prove that
$$\mathbf{H}_n\hat{\mathbf{y}}_n = \mathbf{y}_n, \quad \mathbf{H}_n\hat{\boldsymbol{\epsilon}}_n = 0.$$

   (c) Denote by h_{ij} the (i,j)-th entry in the hat matrix $\mathbf{H}_n$, prove that
$$h_{ii} = \sum_{j=1}^{n} h_{ij}^2,$$
for all $i = 1, 2, ..., n$.

2. Recall that the least squares estimate of $\boldsymbol{\beta}$ is given by
$$\hat{\boldsymbol{\beta}}_n = (\mathbf{Z}_n^\tau \mathbf{Z}_n)^{-1}\mathbf{Z}_n^\tau \mathbf{y}_n,$$
when $\mathbf{Z}_n^\tau \mathbf{Z}_n$ is of full rank.

   Directly verify that this $\hat{\boldsymbol{\beta}}_n$ solves the least squares problem as the solutions to
$$\frac{\partial\, M_n(\boldsymbol{\beta})}{\partial\, \boldsymbol{\beta}} = \mathbf{0},$$
where
$$M_n(\boldsymbol{\beta}) = (\mathbf{y}_n - \mathbf{Z}_n\boldsymbol{\beta})^\tau(\mathbf{y}_n - \mathbf{Z}_n\boldsymbol{\beta}).$$

3. Let $\hat{\boldsymbol{\beta}}_n$ be the least squares estimator and $\tilde{\boldsymbol{\beta}}_n$ be the least absolute deviation estimator under the linear regression model with $p = 0$. Namely, there are no covariates but a constant term in the regression.

   (a) If the error distribution is double exponential, what is the asymptotic efficiency of $\hat{\boldsymbol{\beta}}_n$ compared to $\tilde{\boldsymbol{\beta}}_n$?

   (b) If the error distribution is normal, what is the asymptotic efficiency of $\hat{\boldsymbol{\beta}}_n$ compared to $\tilde{\boldsymbol{\beta}}_n$?

   Remark: Search internet to find the limiting distributions of the $\tilde{\boldsymbol{\beta}}_n$. You are not asked to derive it.

4. Suppose the error distribution in the linear regression model is normal.

   (a) Show that $\hat{\mathbf{y}}_n$ and $\hat{\boldsymbol{\epsilon}}_n$ are independent.

   (b) Derive the distribution of $\hat{\boldsymbol{\epsilon}}_n^{\tau}\hat{\boldsymbol{\epsilon}}_n/\sigma^2$.