# MATH 531 Homework 4

## Generalized least squares and variance estimation

## February 20, 2024

Consider a linear model for data as

$$\mathbf{y} = X\beta + \mathbf{e}$$

where $\mathbf{y}$ is an vector of length $n$ and $X$ is a full rank, $n \times p$ matrix. Assume that $cov(\mathbf{e}) = \Sigma$ ($n \times n$ is known only up to a constant. That is we are given a matrix, $\Omega$, that has full rank, where $\Sigma = \sigma^2\Omega$ but we don't know the scalar parameter $\sigma^2$.

Although this may seem a strange setup it is exactly how we think about the OLS model : $\Sigma = \sigma^2 I_n$ with $I_n$ the identity matrix.

Now suppose you create the "star" model. $\mathbf{y}^* = \Omega^{-1/2}\mathbf{y}$, $X^* = \Omega^{-1/2}X$, and $\mathbf{e}^* = \Omega^{-1/2}\mathbf{e}$. Here $\Omega^{-1/2}$ is the symmetric version so we don't have to worry about keeping track of transposes.

$$\mathbf{y}^* = X^*\beta + \mathbf{e}^*$$

Note that in a (practical) data analysis both $\mathbf{y}^*$ and $X^*$ can be directly computed from the data without estimating any unknown parameters. Often $\Omega$ will depend on other parameters that we don't know – but let's not go there!

1. (a) Explain how to construct $\Omega^{-1/2}$ based on the singular value decomposition of $\Omega$.
   (b) Show that $E(\mathbf{e}^*) = 0$, $E(\mathbf{y}^*) = X^*\beta$ and $COV(\mathbf{e}^*) = COV(\mathbf{y}^*) = \sigma^2 I_n$
   (c) Show that if $\mathbf{c}^T\beta$ is estimable in the original model it is also estimable in the "star" model.
2. (a) Let $\hat{\beta}$ be the OLS estimate based on the "star" model. Use the Gauss-Markov theorem to argue the optimality of this estimate.
   (b) Explain how to use the star model and some moment conditions to guarantee an unbiased and minimum variance estimate for $\sigma^2$.
   BTW: Will the square root of your estimate also be unbiased for $\sigma$ (YES,NO) ?)
3. Writing the results in terms of the original data – your previous answers are in the form of the "star" variables $\mathbf{y}^*$ and $X^*$.
   (a) Express your estimate of $\beta$ in terms of the original $\mathbf{y}$ and $\mathbf{X}$.
   (b) Express your estimate of $\sigma^2$ in terms of the original data and covariates.
   (c) Is your estimate for $X\hat{\beta}$ based on the projection of $\mathbf{y}$ on the column space of $X$? (YES,NO).

4. This problem explores the optimality of the estimate of the variance motivated by least squares fitting. The idea is to use Monte Carlo simulation in R and we will focus on just a simple random sample instead of a linear model to make things easier.

   (a) Let $\{y_i\}$ be a random sample from a $N(\mu, \sigma^2)$. Of course from intro stats we know that $\mu$ is estimated by the sample mean and $\sigma^2$ estimated by the sample average. We also know that $\frac{(n-1)\hat{\sigma}^2}{\sigma^2}$ will have a Chi-squared distribution with $n-1$ degrees of freedom. (We will derive more general results later in the course.)

      - Explain how we can also find these estimates using a simple linear regression model.
      - Report the variance of $\hat{\sigma}^2$

   (b) Now consider the estimate for $\sigma$ based on least absolute deviations.

   $$\hat{\gamma} = (1/(nC)) \sum_{i=1}^{n} |y_i - \bar{y}|$$

   and estimate $\sigma^2$ by squaring this: $\hat{\gamma}^2$.

      - Derive the constant $C$ to make this closer to unbaised using $C = E(|Z|)$ where $Z \sim N(0, 1)$.
      ( I think it is $2/\sqrt{2\pi}$ – how would you check this?)
      - Generate $10^5$ random samples of size n=30, from a $N(10, (2)^2)$. For each sample compute the estimate $\hat{\gamma}$ from the Monte Carlo results determine the bais, variance and root mean squared error of the estimator $\hat{\gamma}^2$ for $\sigma^2$.
      - Compare your results to the estimator $\hat{\sigma}^2$ – which is better according to these metrics?

   (c) Compare your results using both $\hat{\gamma}$ and $\hat{\sigma}$ as estimators for $\sigma$.

   (d) Under what circumstances might you prefer to use $\hat{\gamma}$ as an estimate for $\sigma$?