BIOSTAT 653 Homework #2

Due Wednesday October 11th, 3:10pm, in class.

Problem 1

Consider the general linear model $y = X\beta + \epsilon$, $E(\epsilon) = 0$, $V(\epsilon) = \Sigma$, where y is an n-vector of responses, X is an n by p matrix of covariates, ϵ is an n-vector of residual errors, and Σ is a known n by n symmetric positive definite matrix. Let W be an arbitrary, symmetric positive definite matrix. Let $\hat{\beta}_W$ denote a solution to the following optimization problem

$$\min_{\beta}(y - X\beta)^T W(y - X\beta)$$

The weighted/generalized least square estimator $\hat{\beta}_{\Sigma^{-1}}$ corresponds to $W = \Sigma^{-1}$. Let c be a p-vector (i.e. contrast matrix) and $c^T\beta$ be the corresponding transformed parameter.

(1) Show that $E\left(c^T\hat{\beta}_W\right) = E\left(c^T\hat{\beta}_{\Sigma^{-1}}\right)$

(2) Show that $V\left(c^T\hat{\beta}_W\right) \geq V\left(c^T\hat{\beta}_{\Sigma^{-1}}\right)$

Note that, above, for simplicity, we effectively set the sample size N=1. Also, when M=I, the corresponding estimator $\hat{\beta}_I$ (by setting $W = \mathrm{I}^{-1}$) is the ordinal least squares estimator.

Hint: note that, for question (2), we have $V\left(c^T\hat{\beta}_W\right) - V\left(c^T\hat{\beta}_{\Sigma^{-1}}\right) = c^T((X^TWX)^{-1}X^TW - (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1})\Sigma(WX(X^TWX)^{-1} - \Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1})c$.

Problem 2

We simulate a data set with N=1,000 individuals. We assume that for each individual we obtain n=2 repeated measurements. These measurements are simulated from a multivariate normal distribution with mean $(0, 0)^{\mathrm{T}}$ and a covariance matrix $\begin{pmatrix} 2.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$. Our goal is to use the general linear model to fit the data and obtain parameter estimates. In particular, we consider the model $Y_i = 1_2\mu + \epsilon_i$, for $i = 1, \cdots, N$, where Y_i is a 2-vector of outcomes, 1_2 is a 2-vector of 1's, μ is the intercept, ϵ_i is a 2-vector of residual errors. Note that we do not use the notation X here.

(1) Write down code and simulate a data.

(2) Under a normality assumption that $\epsilon_i \sim MVN(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma)$, write down the log-likelihood and the MLE algorithm to estimate μ and Σ. How do you compute V($\hat{\mu}_{MLE}$)?

Hint: direct copy the equations from slides. Ensure dimensionality matches, and replace X with a proper vector.

(3) Implement the algorithm in question (2) to obtain $\hat{\mu}_{MLE}$ and its variance V($\hat{\mu}_{MLE}$). Also obtain $\hat{\Sigma}_{MLE}$.

Hint: Either based on the previous question (2), or modify based on the draft code provided in question (6).

(4) Is $\hat{\Sigma}_{MLE}$ close to what you expect? Is $\hat{\mu}_{MLE}$ close to what you expect? What is the p-value for testing $H_0: \mu = 0$?

(5) Now we take a step back. Instead of using the normality assumption, we will use the generalized estimating equation (a.k.a weighted least squares) to perform estimation. Assume that our working covariate matrix is of this form: $\Sigma = \begin{pmatrix} \sigma^2 & \rho \\ \rho & \sigma^2 \end{pmatrix}$, with only two parameters instead of three parameters used in question (2). Write down the generalized estimating equation and an algorithm similar to what we describe in class to estimate (μ, σ^2, ρ).

Hint: Either copy and modify the equations from lecture slides, or modify based on the draft code in (6).

(6) Implement the algorithm in question (5) to obtain $\hat{\mu}_{WLS}$. Obtain the robust variance estimate V($\hat{\mu}_{WLS}$). Also, obtain the estimate $\hat{\Sigma}_{WLS}$.

Hind: A draft R code for question (6) is here:

```
# WLS Algorithm:
Si=diag(2); nidv=1000
for (i in 1:10) {
        mu=sum(apply(Si%*%t(Y), 1, sum)/(nidv*sum(Si)))
        S=t(Y-mu)%*%(Y-mu)/nidv
        sigma2=mean(diag(S))
        rho=S[1,2]
        S=matrix(rho, 2, 2)
        diag(S)=sigma2
        Si=solve(S)
}
se_model=sqrt(1/(sum(Si)*nidv))
se_robust=sqrt(sum(Si%*%t(Y-mu)%*%(Y-mu)%*%Si)/(sum(Si)*nidv)^2)
```

(7) Is $\hat{\Sigma}_{WLS}$ close to what you expect? Is $\hat{\mu}_{WLS}$ close to what you expect? Which of the variances are bigger: V($\hat{\mu}_{MLE}$) or V($\hat{\mu}_{WLS}$)? Does the comparison between V($\hat{\mu}_{MLE}$) and V($\hat{\mu}_{WLS}$) fit your expectation?

(8) Now compute the model-based variance estimate V'($\hat{\mu}_{WLS}$) from the WLS algorithm. Assuming that V'($\hat{\mu}_{WLS}$) is different from V($\hat{\mu}_{WLS}$), which variance would you use in practice, and why?

Problem 3

Problem 5.1 on the textbook (page 140-141).