STAT 151A: Lab 7
Midterm Review

Stephanie DeGraaf

October 5, 2018

1. **Normal Regression Theory**
   Suppose $X \in \mathbb{R}^{n \times p}$ and $X^T X$ is invertible. Consider the normal linear model $Y = X\beta + e$, where $\beta \in \mathbb{R}^p$ and $e \sim N(0, \sigma^2 I_n)$. Let $\hat{\beta}$ be the least squares solution.

   (a) Give the explicit form of $\hat{\beta}$. Show that $\hat{\beta}$ is a multivariate normal with $\mathbb{E}(\hat{\beta}) = \beta$ and $cov(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$.

   (b) Let $\hat{Y} = X\hat{\beta}$ and $\hat{e} = Y - \hat{Y}$. Show that $\hat{Y} \sim N(X\beta, \sigma^2 H)$ and $\hat{e} \sim N(0, \sigma^2(I - H))$, where $H = X(X^T X)^{-1} X^T$, the hat matrix.

   (c) Show that H and $I - H$ are idempotent.

   (d) Show that $\hat{Y}$ and $\hat{e}$ are independent.

2. **Simple Linear Regression**
   Suppose we want to predict the longevity of some insects based on their thorax size. We observe 10 insects, measure their thorax sizes, and record their life span. Then we fit a simple linear regression of their longevities $y = (y_1, ..., y_{10})$ on their thorax sizes $x = (x_1, ..., x_{10})$. We computed the following information:

   b hat = r Sy/Sx
   y hat = a hat + b hatX
   y bar = b hat * x bar + a hat

$$\bar{x} = 0.752 \qquad s_x = 0.064$$
$$\bar{y} = 49.2 \qquad s_y = 9.4.$$
$$r = 0.7067.$$

   (a) Report the least squares estimates of the intercept and slope.

   (b) Report the total sum of squares, TSS.

   (c) Report the regression sum of squares, RegSS.

   (d) Report the residual standard error.

   (e) Explain the assumption of homoskedastic errors and how you would assess the validity of this assumption.

   (f) Predict the average life span of these insects with thorax size of 0.85mm.

Under assumption of homoskedastic, the least square estimator is the best unbiased linear estimator among all estimators

if lambdaT*beta is estimable iff lambda is in the row space of X or column space of X transpose

3. **Estimating $\hat{\beta}$**

   (a) State the Gauss-Markov Theorem, including all required assumptions.

   (b) Write the definition of estimability.

   (c) Suppose $\lambda^T\beta$ is estimable.

      i. If (X^TX) is not invertible, what is a solution for β? Is this solution unique?

      ii. If (X^TX) is invertible, what is a solution for β? Is this solution unique?

      iii. What is an estimate for $\lambda^T\beta$? Is this estimate unique?

i, not unique,
ii, unique, xTxbeta = xTy
iii