MATH 531 Homework 5
Least squares and maximum likelihood
February 23, 2024

Each subsection counts for 10 points – 60 points total and the extra credit counts for 5.

1. Assume a linear model
$$\mathbf{y} = X\beta + \mathbf{e}$$
where $\mathbf{y}$ is an vector of length n and X is a full rank, $n \times p$ matrix. $\mathbf{e}_i$ are independent $N(0, \sigma^2)$. Also suppose that β_T is "true" value for β and $\hat{\beta}$ the OLS/MLE estimate and $\hat{\sigma}$ the MLE for σ

   (a) Let $RSS(\beta) = (\mathbf{y} - X\beta)^T(\mathbf{y} - X\beta)$ (aka the residual sums of squares). Show the regression *fun fact* that for *any* β
   $$RSS(\beta) - RSS(\hat{\beta}) = (\beta - \hat{\beta})^T(X^T X)(\beta - \hat{\beta})$$

   (b) Let $\ell(\mathbf{y}, \beta, \sigma^2)$ be the log likelihood for this model. Consider the difference
   $$D(\beta_T) = 2\left[\ell(\mathbf{y}, \hat{\beta}, \hat{\sigma}^2) - \ell(\mathbf{y}, \beta_T, \hat{\sigma}^2)\right]$$
   Explain why this difference is always nonnegative.

   (c) Use the fun fact above to simplify as much as you can $D(\beta_T)$.

   (d) Let $\mathcal{I}(\beta_T, \sigma^2)$ be the Fisher information for this model and let A be the submatrix that pertains to just β and where the MLE's are substituted for the true parameters. Simplify the expression
   $$U(\beta_T) = (\hat{\beta} - \beta_T)^T(A)(\hat{\beta} - \beta_T)$$
   and compare your result $D(\beta_T)$.

2. This problem is a continuation of problem 4 from HW04.
Let $\mathbf{y} = \{Y_k\}$ be an independent random sample of size n each value distributed according to the Laplace distribution with location and scale parameters μ and b. Look up the density and details of this distribution on wikipedia.

   (a) Report the log likelihood for $\mathbf{y}$ and provide a complete derivation of the MLEs for μ and b. (There are many references for this problem find a solution that you like best! You can assume that n is odd if that makes your argument related to the median of $\mathbf{y}$ simpler.)

   (b) Assume a linear model but let the errors be distributed according to the Laplace distribution.
   $$\mathbf{y} = X\beta + \mathbf{e}$$
   where $\mathbf{y}$ is an vector of length n and X is a full rank, $n \times p$ matrix. $\mathbf{e}_i$ are independent Laplace, mean zero and scale equal to b. Setup up the log likelihood to find the MLE for β and b. β must be found numerically, however, explain why given the MLE $\hat{\beta}$ the MLE for b has a closed form.

   (c) EXTRA CREDIT. Identify an R package that will allow you to find the MLE for β for this model.