MTH210: Lab 8

Linear Regression, MLE, Ridge, and Newton-Raphson

1. Load the `cars` dataset in R:

   ```
   data(cars)
   ```

   Fit a linear regression model using maximum likelihood with response y being the distance and x being speed. Remember to include an intercept term in X by making the first column as a column of 1s. Do not use inbuilt functions in R to fit the model.

2. Load the `fuel2001` dataset in R:

   ```
   fuel2001 <- read.csv("https://dvats.github.io/assets/fuel2001.csv", row.names = 1)
   ```

   a. Fit the linear regression model using maximum likelihood with response `FuelC`. Remember to include an intercept in X. What is your final estimate of β and σ^2?

   b. For $\lambda = 1$, what is the ridge regression estimator of β?

3. **Simulating data in R:** Let $X \in \mathbb{R}^{n \times p}$ be the design matrix, where all entries in its first column equal one (to form an intercept). Let $x_{i,j}$ be the (i, j)th element of X. For the i^{th} case, $x_{i1} = 1$ and $x_{i2}, \dots, x_{ip}$ are the values of the $p - 1$ predictors. Let y_i be the response for the ith case and define $y = (y_1, \dots, y_n)^T$. The model assumes that y is a realization of the random vector:

   $$Y \sim N_n(X\beta_*, \sigma_*^2 I_n)\,,$$

   where $\beta_* \in \mathbb{R}^p$ are unknown regression coefficients and $\sigma_*^2 > 0$ is the unknown variance. We would like to *generate data* that actually follows the following model. This is useful when building too methods and different estimation techniques for β.

   a. Study the code below and understand how the data is being generated according to the model:

   ```
   set.seed(1)
   n <- 50
   p <- 5
   sigma2.star <- 1/2
   beta.star <- rnorm(p)
   ```

```
X <- cbind(1, matrix(rnorm(n*(p-1)), nrow = n, ncol = (p-1)))
y <-  X %*% beta.star + rnorm(n, mean = 0, sd = sqrt(sigma2.star))
```

b. Having generated the above `(y,X)`, obtain the MLE of β. Is the MLE roughly close to β^*? What happens when you increase $n = 500$?

c. Repeat the data generation process, but now change $p = 100$ and keep $n = 50$. Can you find the traditional MLE in this case?

d. For the above data find the ridge regression estimator of β for $\lambda = 0.01, 0.1, 1, 10$.

4. Write a Newton-Raphson code to find the MLE of α for Gamma $(\alpha, 1)$ distribution. That is, suppose $X_1, X_2, \dots, X_n \overset{iid}{\sim} \text{Gamma}(\alpha, 1)$, then write a function to obtain $\hat{\alpha}_{MLE}$.

   In order to implement this, you will need data that is from $Gamma(\alpha, 1)$. You may use the following:

```
set.seed(100)
alpha <- 5 #true value of alpha
n <- 10 # actual data size is small first
dat <- rgamma(n, shape = alpha, rate = 1)
```

   The above generates $n = 10$ observations. Use `dat` to obtain the MLE of α. You will need the `pracma` library in R to calculate the derivatives of $\Gamma(\cdot)$ function.

```
library(pracma)  #for psi function
?psi
```

5. Using both Newton-Raphson and gradient ascent algorithm, maximize objective function

$$f(x) = \cos(x) \quad x \in [-\pi, 3\pi].$$