ST 495 Advanced computing for statistical methods
Homework problem set 7

March 2, 2022

No `R` packages are permitted for use in this assignment.

1. Write an `R` function that takes as input `( df, eta, x0, epsilon)`, where `df` is the gradient of some function $f : \mathbb{R}^n \to \mathbb{R}$ to minimize, `eta` is a positive-valued learning rate parameter, `x0` is an initialization point, and `epsilon` is a positive-valued convergence tolerance, and returns a minimizer of f along with the trace of the points $x^{(0)}, x^{(1)}, \ldots$. Follow the pseudocode for the gradient descent algorithm given during lecture.

2. Implement the gradient descent function you wrote for problem 1 to find minimizers of the following functions. Additionally, for each part be sure to plot the function, with vertical lines indicating each point in the trace of the algorithm. Denote the initializer $x^{(0)}$ with a red line, and the minimizer (i.e., the final point) with a green line. In a separate panel, also display the trace plot of the function evaluated at each iteration. Set the learning rate as $\eta = .1$ to start, and modify as necessary.

   (a) $f(x) = 4 + (x - .6)^2 + (x + 5.1)^4$.

   (b) $f(x) = \cos(x)$. Be sure to try at least two distinct initialization points that are separated by a distance of greater than 2π.

   (c) $f(x) = e^x$. What should you add to your gradient descent algorithm to handle this kind of function?

   (d) Repeat the analyses of parts (a), (b), and (c) with a learning rate of $\eta = 2$. What changes?

3. Implementing the gradient expression that we derived in lecture for the negative log-likelihood of the logistic regression model, use the gradient descent function from problem 1 to train a logistic regression model on synthetic data. Fixing some β vector, generate at least 100 synthetic data sets, and present histograms of the sampling distribution of

each component of the MLE of β. Be sure to set the sample size sufficiently large so that the histograms concentrate nicely around the true but unknown components of β.