# Problem Set 8

Will Miller

April 1, 2025

## 1 Discussion of optimization techniques

Overall, most of these techniques were very similar. I much prefer the syntax of the nloptr approach to the gradient descent produced using a while loop, and I also found the OLS objective and cost functions easier to understand/work with than the maximum likelihood approach. As I understand it, though, maximum likelihood is about as ubiquitous as it gets in statistics, so I will keep reading/practicing so I can understand it better. As for comparing coefficients, I found that all of the methods we used were essentially identical, at least to the number of digitis that R is capable of displaying them to. For each result, I got coefficients of 1.50029, -0.9978091, -0.2493249, 0.7485952 3.500883, -1.999718, 0.5005669, 0.99937, 1.251415, and 1.999692. In terms of computational efficiency, I found that both L-BFGS methods (typical gradient descent and maximum likelihood) ran quickly when compared to the Nelder-Mead algorithm, with the gradient descent L-BFGS taking 10 iterations to reach the optimal solution, maximum likelihood L-BFGS taking 76 iterations to reach the optimal solution, and the Nelder-Mead simplex method taking 1,725, a quite substantial increase. I think this drastic difference is likely due to the relative computational simplicity of the ordinary least squares estimator. Its cost function is quite simple, so something like Nelder-Mead is just overkill. I also have to assume that the methods relying on the closed form solution (the linear algebra approach and the lm() approach) were by far the most efficient because they directly produce analytic solutions rather than relying on numerical approximation which is always more costly.

Table 1: Coefficients estimated using lm()

| | (1) |
|---|---|
| X1 | 1.50*** |
| | (0.00) |
| X2 | −1.00*** |
| | (0.00) |
| X3 | −0.25*** |
| | (0.00) |
| X4 | 0.75*** |
| | (0.00) |
| X5 | 3.50*** |
| | (0.00) |
| X6 | −2.00*** |
| | (0.00) |
| X7 | 0.50*** |
| | (0.00) |
| X8 | 1.00*** |
| | (0.00) |
| X9 | 1.25*** |
| | (0.00) |
| X10 | 2.00*** |
| | (0.00) |

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$