Jake Stephens

Assignment #4

August 3, 2021

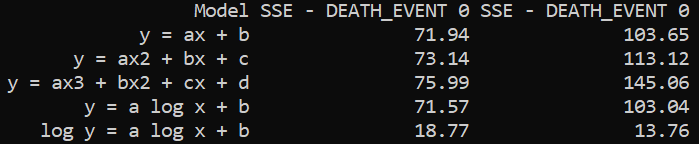
**Question #1:**

1. Done
2. Done (see M\_0.pdf and M\_1.pdf)
3. 1. Out of the surviving patients, serum\_sodium and serum\_creatine have the highest correlation.
   2. Out of the surviving patients, platelets and serum\_sodium have the lowest correlation.
   3. Out of deceased patients, serum\_sodium and creatinine\_phosphokinase have the highest correlation.
   4. Out of deceased patients, serum\_creatine and platelets have the lowest correlation.
   5. For both deceased and surviving patients, different features have the lowest correlation. For instance, for surviving patients, serum\_sodium and serum creatine have the highest correlation, where as for deceased patients, serum\_sodium and creatinine\_phosphokinase have the highest correlation. For lowest correlation, surviving patients and deceased patients, platelets had the lowest correlation, but with serum\_sodium for surviving patients and with serum\_creatine for deceased patients. This was not surprising to me as I would expect there to be a difference in values for deceased versus surviving patients.

**Question #2:**

See console output of q2.py for the values of the weights and the loss functions.

**Question #3:**



1. I believe the best model is the last one of log(y) = a log(x) + b, since it has significantly lower SSE’s than the other algorithms. This algorithm was the best for both surviving patients and deceased patients.
2. The worst model is the cubic algorithm. It was also the worst algorithm for both surviving patients and deceased patients. I believe this has to do with that fact that depending on the dataset, different algorithms tent to work better than others.