Part 2

Questions

1 Introduction

In this homework you will implement linear and ridge regressions using gradient descent and stochastic gradient descent. We've provided a lot of support Python code to get you started on the right track. Take advantage of this being a relatively straightforward assignment (it's not as long as it looks!) by improving your programming skills and/or pursuing independent investigations. For example:

- Study up on numpy's "broadcasting" to see if you can simplify and/or speed up your code: http://docs.scipy.org/doc/numpy/user/basics.broadcasting.html
- Think about how you could make the code more modular so that you could easily try different loss functions and step size methods.
- Experiment with more sophisticated approaches to setting the step sizes for SGD (e.g. try out the recommendations in "Bottou's SGD Tricks" on the website)
- Instead of taking 1 data point at a time, as in SGD, try "mini-batch gradient descent" where you only use randomly selected subsets of data points to determine each step. How does this effect convergence speed? Are you getting computational speedup as well by using vectorized code?

Now have fun, and go learn!