STAT 598Z: Course Project

Due: 23rd April 2013

1 Project Details

1. This project will contribute 15 points towards your final score.
2. Hand in your Project (including print outs of your source code) at the beginning of the class on 23rd April 2013. Additionally your source code should be emailed to `stat598z@gmail.com` **before** the project is submitted in the class. No late submissions will be accepted!

2 Visualizing Optimization Algorithms

One of the best ways to learn about optimization is to visualize how the algorithms work on simple problems. In this project we will learn about gradient descent by visualizing its output first on a set of simple quadratic problems and then on a two dimensional linear regression problem. For this you will need to familiarize yourself with `mplot3d` which is the 3d plotting toolkit of matplotlib. Then do the following:

- Plot and visualize the following functions on three separate plots:

$$J(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (1)$$

$$J(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (2)$$

$$J(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \quad (3)$$

- With your initial starting point as $x = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$ plot the trajectory of points produced by gradient descent on the above three functions.
- Comment on what you observe. I am expecting a neat half page write up explaining and summarizing your observations.

- Draw 1000 samples each uniformly at random from two different 2-d Gaussian distributions with $\mu_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mu_2 = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$, and $\Sigma_1 = \Sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Label the points from the first Gaussian as $+1$ and from the second Gaussian as -1. Plot the following objective function:

$$J(w) = \frac{\lambda}{2}\|w\|^2 + \sum_i (w^\top x_i - y_i)^2 \tag{4}$$

  with $\lambda = 0.001$.

- Plot the trajectory of points produced by gradient descent on the above function. Vary λ and repeat.

- Comment on what you observe. I am expecting a neat half page write up explaining and summarizing your observations.