Reading Quiz 1
Stochastic Gradient Descent
DATA 442/642

Exercise 1 (10 points)
Read pages 227-233 from the attached pdf file.

(a) Briefly explain what are the differences between gradient descent, gradient descent with momentum, and stochastic gradient descent.

(b) Using the given Python code for stochastic gradient descent (SGD), write a script to compare SGD and the gradient descent algorithm that was demonstrated during lecture. For both algorithms, use the same cost function and by starting with an initial guess from a Gaussian distribution, and by using 1000 iterations and a learning rate of 0.01, compare stochastic gradient descent and gradient descent in terms of convergence speed and accuracy. Try different learning rate and discuss your results. Demonstrate your results using figures like we did in class.