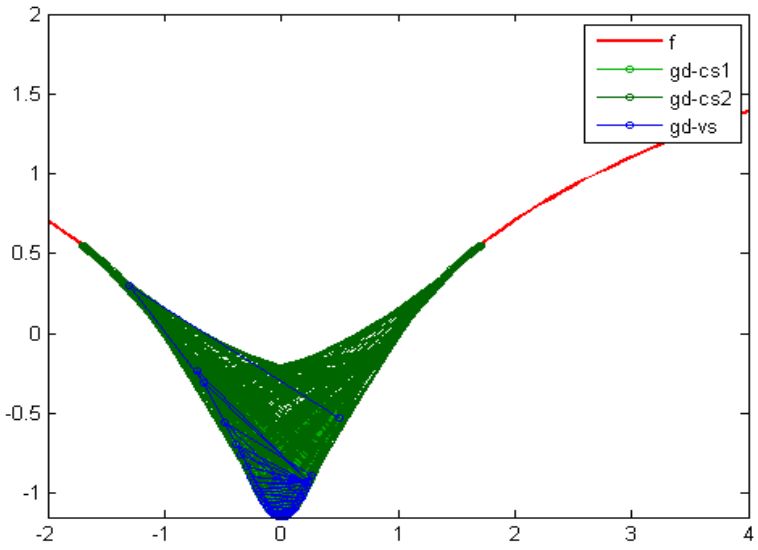
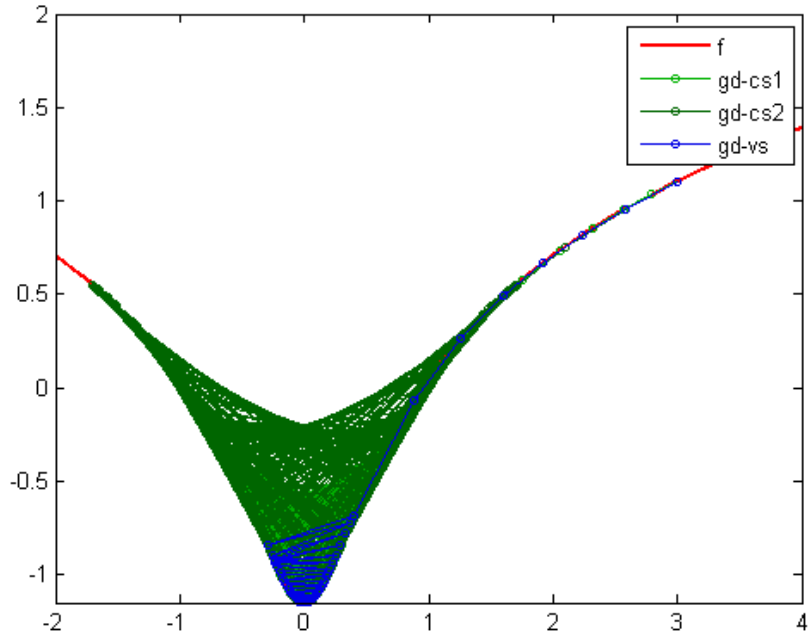
Exercise 4, e)

When x0 = 0.5, the plot of the objective function and iterations of the Gradient Descent looks like this:



When x0 = 3, it looks like this:



The gradient descent doesn't converge when the alphas of the descent are constant (when x0 = 0.5 and when x0 = 3). Using dynamic alphas (third method), the gradient descent converges. In this case, choosing x0=0.5 or x0=3 doesn't really change the convergence rate –both converge in roughly equal amount of iterations.