

Homework Set Three
ECE 251B Spring 2022

In this homework, we will investigate various optimization algorithms discussed in the class.

1. Consider the cost function

$$J(\mathbf{w}) = (\mathbf{w} - \mathbf{w}_o)^T \mathbf{A}(\mathbf{w} - \mathbf{w}_o),$$

where $\mathbf{w} \in \mathbb{R}^2$, $\mathbf{w}_o = (-2, 2)^T$ and $\mathbf{A} = \mathbf{I}$ (the identity matrix).

- a) Make a contour plot showing the iso-contours of $J(\mathbf{w})$ and a surface plot of the function.
- b) Compute the expression of the gradient of $J(\mathbf{w})$ and make a quiver plot of the gradient.
- c) Write down the expression for the parameter updates under the gradient descent algorithm. Using as initialization the vector $\mathbf{w}^{(0)} = (5, 15)^T$ and a step size $\mu = 0.09$, run the gradient descent algorithm until

$$\|\mathbf{w}^{(n)} - \mathbf{w}_o\|^2 < 0.001.$$

Plot the values of $\mathbf{w}^{(n)}$ obtained after each iteration on the contour plot of $J(\mathbf{w})$. How many iterations were required for convergence?

- d) Repeat c) for step sizes $\mu = 5$ and μ found by a line search. Comment on the ability to converge and convergence speed of the different algorithms. How do you explain the different behaviors?
- e) Write down the expression for the parameter updates under Newton's method. This is of the form

$$\mathbf{w}^{(n+1)} = \mathbf{w}^{(n)} + F(\mathbf{w}^{(n)}).$$

Make a quiver plot of the function $F(\mathbf{w})$. Comment on the differences between this plot and that of b).

- f) Using as initialization the vector $\mathbf{w}^{(0)} = (5, 15)^T$, run Newton's method until

$$\|\mathbf{w}^{(n)} - \mathbf{w}_o\|^2 < 0.001.$$

Plot the values of $\mathbf{w}^{(n)}$ obtained after each iteration on the contour plot of $J(\mathbf{w})$. How many iterations were required for convergence?

- g) Comment on the ability to converge and convergence speed of the algorithm as compared with the algorithms of c) and d). How do you explain the different behaviors?

2. Repeat Problem 1. with $\mathbf{A} = \text{diag}(2, 1)$. Compare your answers to all questions to those obtained in Problem 1. and discuss any differences.

3. Repeat Problem **1.** with $\mathbf{A} = \text{diag}(20, 1)$. Compare your answers to all questions to those obtained in Problems **1.** and **2.** and discuss any differences.

4. It is known that the contours of $J(\mathbf{w})$ are ellipses whose principal directions have angle $\pi/4$ and $-\pi/4$ with the horizontal axis. The eigenvalues of \mathbf{A} are $\lambda_1 = 2$ and $\lambda_2 = 1$. What is the matrix \mathbf{A} ? Repeat Problem **1.** for this \mathbf{A} . Compare your answers to all questions to those obtained in Problems **1.**, **2.**, and **3.** and discuss any differences.

5. Repeat Problem **4.** for $\lambda_1 = 20$ and $\lambda_2 = 1$.