



CS 403/725
Tutorial - 2
Spring 2016
IIT Bombay

Q1 Gradient descent is a very helpful algorithm. But it is not guaranteed to converge to global minima always. Give an example of a continuous function and initial point for which gradient descent converges to a value which is not global minima?

Q2 (a) Consider the function

$$f(x) = x_1^2 - 4x_1 + 2x_1x_2 + 2x_2^2 + 2x_2 + 14$$

This function has a minimum at $\mathbf{x} = (5, -3)^T$. Suppose you are at a point $(4, -4)^T$ after few iterations, using the **exact line search** algorithm discussed in the class find the point for the next iteration.

(b) Now consider solving the Least Squares Linear Regression problem using the gradient descent algorithm. And let us say $w^{(0)} = 0$ and that the step length $t^{(k)}$ is computed using exact line search for each value of k . In how many steps will the gradient descent algorithm converge? What would be your answer if we had a different initialization for $w^{(0)}$?

Q3 Suppose you are solving the equation $Ax=b$ using gradient descent on least squares solution. How do you think the Eigenvalues of the matrix affect the convergence?

(Hint: Consider a 2x2 diagonal matrix for A what do you observe?) [Source : Quora]