DATA SCIENCE FOR ENGINEERS

Optimization-FAQ's

Multivariate Optimization

1. In the lecture on Multivariate Optimization with Equality Constraints, In example 3x1+2x2-12=0, how to get the following values for x1 and x2?

Answer:

$$-4x_1 - 3\lambda = 0 \quad \rightarrow (1)$$
$$-8x_2 - 2\lambda = 0 \quad \rightarrow (2)$$
$$3x_1 + 2x_2 - 12 = 0 \rightarrow (3)$$

$$(1) * 3 \Rightarrow -12x_1 + 0x_2 - 3\lambda = 0 \rightarrow (4)$$
$$(3) * 4 \Rightarrow 12x_1 + 8x_2 + 0\lambda = 48 \rightarrow (5)$$

$$(4) + (5) \Rightarrow 8x_2 - 9\lambda = 48 \rightarrow (6)$$

$$\textit{Solving } (6) \textit{ and } (2)$$
$$(6) + (2) \Rightarrow -11\lambda = 48$$
$$\lambda = -\frac{48}{11}$$

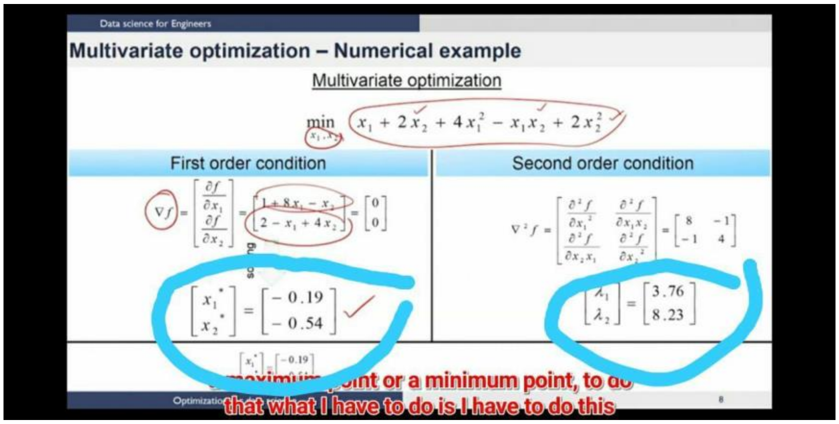
$\textit{Substituting } \lambda \textit{ value in } (1) \textit{ to find } x_1$

$$-4x_1 - 3\left(-\frac{48}{11}\right) = 0$$
$$-4x_1 = -\frac{144}{11}$$
$$4x_1 = 13.09$$
$$x_1 = \frac{13.09}{4} = 3.2725$$

$\textit{Substituting } \lambda \textit{ value in } (2) \textit{ to find } x_2$

$$-8x_2 - 2\left(-\frac{48}{11}\right) = 0$$
$$-8x_2 = -\frac{96}{11}$$
$$8x_2 = 8.72$$
$$x_2 = 1.09$$

2. How will we get the values of x1* and x2* and the second derivative lamda1 and lamda2 in the below slide (Multivariate optimization)?



Answer:

Solve the two equations

8x1- x2=-1 ->(1)

-x1+4x2= -2 ->(2)

Multiply the (2nd) equation by 8

x2= -17/31= -0.54

Substitute the value of x2 in any one of the equation and get the x1 value.

Lambda 1 and lambda 2 are the eigen values of the hessian matrix.

Enter the 2*2 matrix in R and use eigen(matrix_name), you will be able to get the values of lambda.

Stationary points

1.What is a stationary point for a function?

Answer:

For f(x),

Solution of grad (f) =0 in case of multivariate optimization problem

Solution of f'(x) =0 in case of univariate optimization problem

Unconstrained Multivariate Optimization

1.In the lectures on Unconstrained Multivariate Optimization, why have we taken the negative of gradient for search direction while calculating the f(x).

Answer:

Searching requires an objective. Lets assume this objective is minimization of F. Then, from where we are (epoch n), we need to get a lower value (epoch n+1). On the other hand, numerically, the gradient can be thought of as the ratio of change in F when x is increased. This ratio can be positive (F increases with x) or negative (F decreases with increase in x).

For simplicity, assume the positive gradient case with learning rate of unity. A lower value of F is required in order to minimize. One then needs to subtract the gradient from current epoch value so that the value is minimized. Adding the gradient will only make F larger and is of no use for minimization. Hence the negative sign for the gradient.

If the gradient is zero, there are no more epochs to improve upon, so the search is stopped.