



MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 0. Brief Prerequisite Reviews, Homework 0, and Project 0](#) / [Home](#)[← Previous](#)

9 of 15 ▾

9. Gradients and Optimization

Bookmark this page

Homework0 due Feb 8, 2023 08:59 -03 Completed

Multivariable Calculus Review: Gradient

1.0/1 point (graded)

Let

$$f: \mathbb{R}^d \rightarrow \mathbb{R}$$

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{pmatrix} \mapsto f(\theta).$$

denote a **differentiable** function. The **gradient** of f is the vector-valued function

$$\nabla_{\theta} f: \mathbb{R}^d \rightarrow \mathbb{R}^d$$

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{pmatrix} \mapsto \left(\begin{pmatrix} \frac{\partial f}{\partial \theta_1} \\ \frac{\partial f}{\partial \theta_2} \\ \vdots \\ \frac{\partial f}{\partial \theta_d} \end{pmatrix} \right) \bigg|_{\theta}.$$

Consider

$$f(\theta) = \theta_1^2 + \theta_2^2.$$

Here, θ has dimension **2**. Compute the gradient ∇f .

(Enter your answer as a vector, e.g., type **[2,x]** for the vector $\begin{pmatrix} 2 \\ x \end{pmatrix}$. Note the square brackets and the use of x as a variable. Enter **theta_i** for θ_i .)

$$\nabla_{\theta} f(\theta) =$$

[2*theta_1, 2*theta_2]

? STANDARD NOTATION

Submit

You have used 1 of 3 attempts

☐ line

Consider how the level curves change as increases from to . Do they have a global maximum, or global minimum, or neither?

☐ global maximum☐ global minimum☒ neither

At each point in the xy -plane, $f(x, y)$ decreases in the direction of...



You have used 2 of 2 attempts

Gradient ascent/descent methods are typical tools for maximizing/minimizing functions $f(x, y)$ where x and y are fixed. Our goal is to select x and y to maximize/minimize the value of $f(x, y)$ while keeping x and y fixed.

Compute the Gradient

1/1 point (graded)

The gradient $\nabla f(x, y)$ is a vector with 2 components:



You have used 1 of 1 attempt

Gradient Ascent or Descent

0/1 point (graded)

The direction of the derivative of a function gives us the direction of the largest change in the function as the independent variables vary.

In gradient ascent/descent methods, we make an educated guess about the next value of the parameters and update them with updates that will hopefully eventually converge to the global minimum of the cost function (if it exists).

If

where ϵ is a small positive real number, Which of the following is true?

[< Previous](#)[Next >](#)[? STANDARD NOTATION](#)

🗨 [new error message](#)

[Couldn't execute jailed code: stdout: b'', stderr: b'' with status code: -9](#)

🗨 [i want ans of question guys...](#)

[i try from 5-6 hr...i didnt get any ans...so plz help me](#)

🗨 [Review material](#)

[I haven't seen this kind of calculus before... I got them right from intuition on other calculus but is there a tex](#)

🗨 [in the last problem, ' is not a derivative!](#)

🗨 [Is the first question, \(gradient\) complete?](#)

[I honestly think the question is incomplete. Without knowing the dimensions of angle theta, it becomes hard](#)

Legal

[Terms of Service & Honor Code](#)

[Privacy Policy](#)

[Accessibility Policy](#)

[Trademark Policy](#)

[Sitemap](#)

[Cookie Policy](#)

[Do Not Sell My Personal Information](#)

Connect

[Blog](#)

[Contact Us](#)

[Help Center](#)

[Security](#)

[Media Kit](#)

