





MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 1. Linear Classifiers and Generali...](#) / [Lecture 4. Linear Classificat](#)[< Previous](#)

4. Gradient Descent

[Bookmark this page](#)

Exercises due Feb 22, 2023 08:59 -03 Completed

Gradient Descent**Video** [Download video file](#)**Transcripts** [Download SubRip \(.srt\) file](#) [Download Text \(.txt\) file](#)**Gradient Descent: Geometrically Revisited**

2/2 points (graded)

Assume $\theta \in \mathbb{R}$. Our goal is to find θ that minimizes

$$J(\theta, \theta_0) = \frac{1}{n} \sum_{i=1}^n \text{Loss}_h(y^{(i)}(\theta \cdot x^{(i)} + \theta_0)) + \frac{\lambda}{2} \|\theta\|^2$$

through gradient descent. In other words, we will

1. Start θ at an arbitrary location: $\theta \leftarrow \theta_{start}$ 2. Update θ repeatedly with $\theta \leftarrow \theta - \eta \frac{\partial J(\theta, \theta_0)}{\partial \theta}$ until θ does not change significant

☐ away from the origin

☒ towards the origin

☐ upwards

☐ downwards

< Previous

Next >



edX

[About](#)

[Affiliates](#)

[edX for Business](#)

[Open edX](#)

[Careers](#)

[News](#)

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](#)