13.1.4 Gradient descent algorithm in one dimension

Bookmark this page

MO2.11

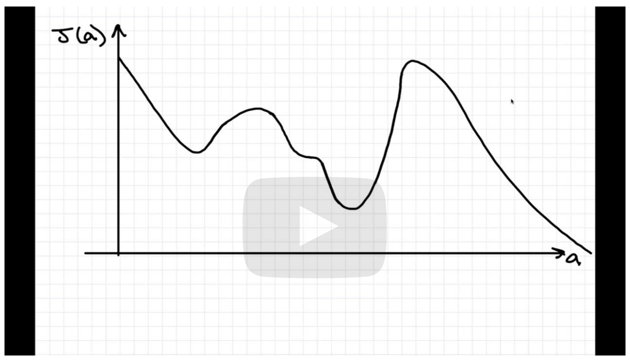
To find the value of a where $J\left(a\right)$ is minimum, the simple idea is to set up an iterative scheme based on the slope of the function. If the slope is negative, then a should be increased, while if the slope is positive, then a should be decreased. We can write this as

$$a^{k+1} = a^{k} - \alpha J'\left(a^{k}\right), \tag{13.13}$$

for some adequate number $\alpha > 0$. We are using the shorthand derivative notation in this one dimensional case where $J'\left(a\right) = \mathrm{d}J/\mathrm{d}a\left(a\right)$. We might call this algorithm “slope descent", in analogy with the gradient descent algorithm that we will encounter later. It requires the computation of the derivative in the a variable. When $J'\left(a^{k}\right) = 0$ then a no longer changes, and we have converged to a local minimum. We have no guarantee that this will be the global minimum though. If this is an issue, special care should be exercised to choose the point from which to initialize the iteration.

The number α should be chosen carefully: large enough that the algorithm converges fast to the minimum, but small enough to stay in the neighborhood of the minimum and not overshoot it. Some trial and error helps. The reliable way of choosing α is to perform what is called a line search, but it is outside of the scope of the class.

Video discussing 1D gradient descent algorithm



Start of transcript. Skip to the end.

PROFESSOR: Now we're going to

< Previous Next >

Discussions

Posting in discu

All posts sorted by recent activity

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]
Your Privacy Choices

Connect

[Idea Hub]
[Contact Us]
[Help Center]
[Security]
[Media Kit]

GET IT ON Google play
Download on the App Store