Lecture 11

Wednesday, February 12, 2025 1:22 AM

Gradient descent algorithm for finding minimum:

- Guess x_0 (location of minimum)
- Update: $x_{n+1} = x_n - \alpha f'(x_n)$

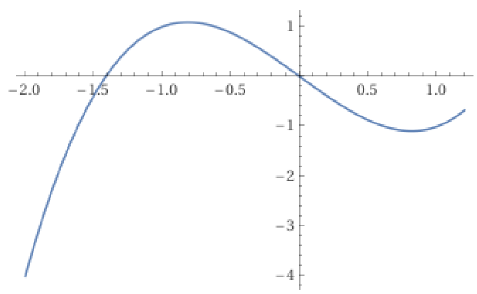
Gradient ascent algorithm for finding maximum:

- Guess x_0 (location of maximum)
- Update: $x_{n+1} = x_n + \alpha f'(x_n)$

Gradient descent/ascent algorithms are *short-sighted* in nature. They don't see the big picture (the graph) of the function. They only give you a *local* minimum and *local* maximum.

Example: Let $f(x) = x^3 - 2x$
Find the minimum and maximum value of $f(x)$ on the interval $[-2, 1.2]$ using the Gradient descent/ascent method. Pick an initial guess x_0 and a learning rate α.

$f'(x) = 3x^2 - 2$
Gradient descent: $x_{n+1} = x_n - \alpha(3x_n^2 - 2)$
Gradient ascent: $x_{n+1} = x_n + \alpha(3x_n^2 - 2)$



If $x_0 = 0$, the Gradient descent method will give you a local minimum at about 0.8.
If $x_0 = 1$, the Gradient ascent method will give you a local maximum at about 1.2.
If x_n falls out of bound, i.e. the interval $[-2, 1.2]$, then stop. If your algorithm stops too early, perhaps that is because α is too large. You can reduce α and try again.

Bisection method: a root-finding algorithm relying on Intermediate Value Theorem.
Motivating example: I have a natural number between 1 and 100 in mind. I will tell you whether my number is bigger than or smaller than the number you guess. How can you guess it with the least number of guesses? Guess 50 first. Then guess either 25 or 75 based on my answer to your first guess. And so on. This is the key idea of bisection method. Its mathematical basis is the *Intermediate Value Theorem*:

If f is continuous on $[a, b]$ and $f(a), f(b)$ have different signs, then there exists $c \in (a, b)$ such that $f(c) = 0$.

Example: find $\sqrt{2}$ using bisection method.
Note that $\sqrt{2}$ is a root of $f(x) = x^2 - 2$. We have $f(1) = -1 < 0$ and $f(2) = 2 > 0$.