Gradient descent algorithm

Prof. Junghyun Kim
Director, Engineering Systems Design Laboratory (ESDL)
Assistant Professor, School of Applied Artificial Intelligence
Handong Global University

1

Course objectives

- By the end of this module, you will be able to answer the following questions:
  - Where does the gradient descent algorithm come from?
  - What is the gradient descent algorithm?
  - What are the limitations of the gradient descent algorithm?

2

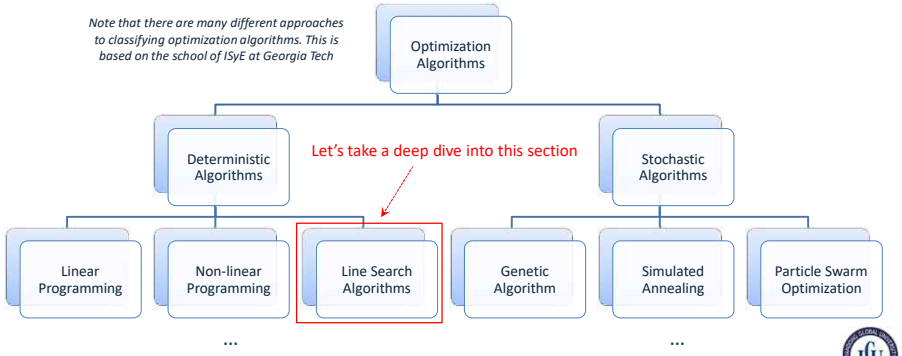
Where does the gradient descent algorithm come from?

3

Types of optimization algorithms

- Where does the gradient descent algorithm come from?

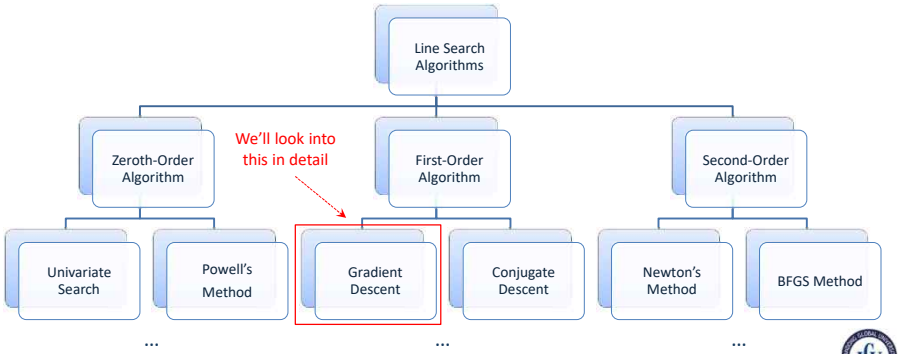
Note that there are many different approaches to classifying optimization algorithms. This is based on the school of ISyE at Georgia Tech



4

Types of line search algorithms

- The gradient descent algorithm is one of the first-order line search algorithms



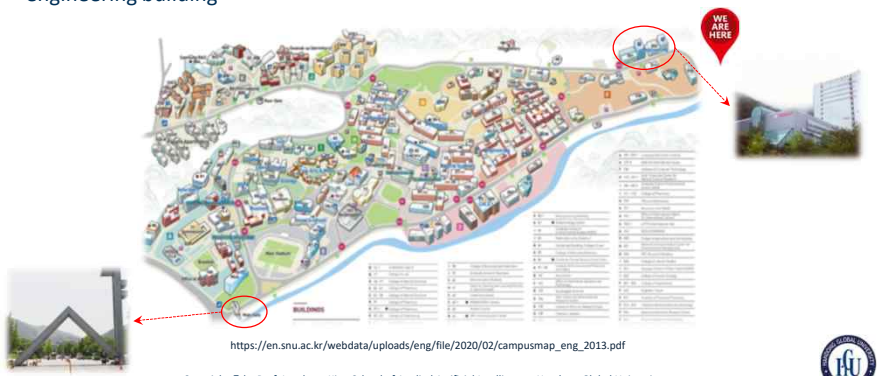
5

What is the gradient descent algorithm?

6

Understanding the gradient descent algorithm

- Suppose that we need to go to the main gate of Seoul National University from the engineering building



https://en.snu.ac.kr/webdata/uploads/eng/file/2020/02/campusmap_eng_2013.pdf

7

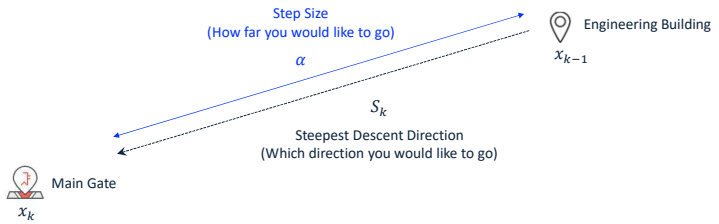
Understanding the gradient descent algorithm

- Discussion
  - How can we get to the main gate as quickly as possible?

8

Understanding the gradient descent algorithm

- One good strategy to arrive at the main gate quickly is to go downhill in the direction of the steepest slope (i.e., steepest descent direction)
- This is exactly what the gradient descent algorithm does!
  - Note that the size of the steps is a user-defined parameter in the gradient descent algorithm



Ingredients of the gradient descent algorithm

- The gradient descent algorithm is based on an iterative procedure of the following form:

$$x_k = x_{k-1} + \alpha s_k$$

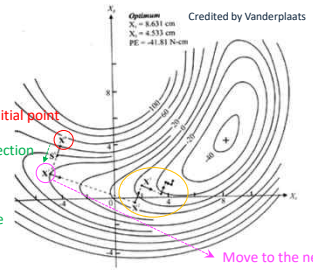
where:

$x_k \equiv$ new design variable vector
$x_{k-1} \equiv$ previous design variable vector
$s_k \equiv$ a vector in a descent direction that defines the line being searched (e.g., $s_k = -\nabla f(x_{k-1})$)
$\alpha \equiv$ a scalar indicating the distance we should move to minimize the function along the line

The general idea is to tweak parameters iteratively to minimize a function (e.g., error)

Steps of the gradient descent algorithm

- Steps in detail
  - Step 1. Choose an initial point in the design space, x_0
  - Step 2. Select a descent direction, s_k (i.e., a direction along which $f(x)$ decreases)
  - Step 3. Determine a step size, α minimizing $f(x)$ along the line by s_k, and move to the next point
  - Step 4. Repeat the process and stop when the process is converged to an acceptable solution



Credited by Vanderplaats

Step 1. Choose an initial point

Step 2. Select a descent direction

$s_k = -\nabla f(x_{k-1})$

Note that the direction with the maximum rate of decreasing is perpendicular to the isoline

Step 3. Determine a step size

Note that it tends to use a fixed step size for simplicity but it can implement an algorithm determining a step size minimizing $f(x)$ along the line defined by s_k

Step 4. Repeat the process and stop when convergence criteria are met

Move to the next "better" point $x_k = x_{k-1} + \alpha s_k$

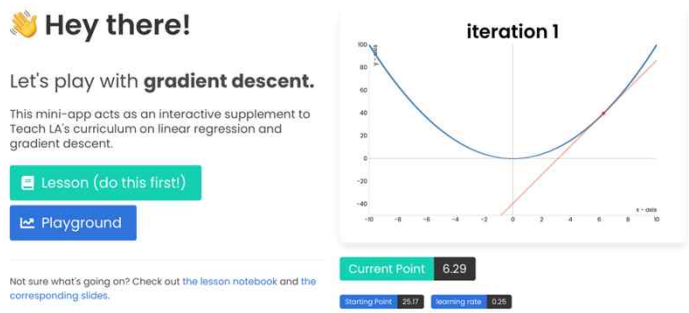
YouTube video clip

- The process of the gradient descent algorithm

https://www.youtube.com/watch?v=qg4PchTECck

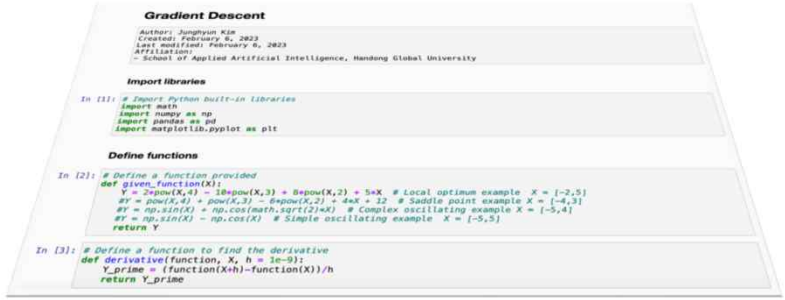
Demo

- https://uclaacm.github.io/gradient-descent-visualiser/



Classroom activity

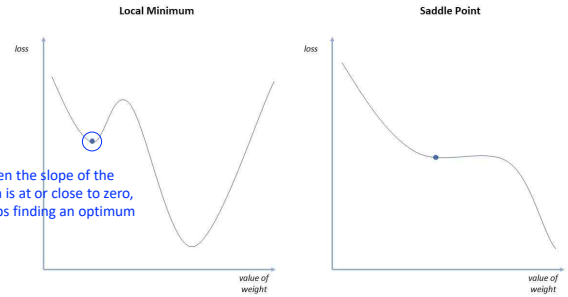
- Python code – Gradient descent algorithm



What are the limitations of the gradient descent algorithm?

Limitations of the gradient descent algorithm

- For convex problems, the gradient descent algorithm can easily find a global minimum
- For non-convex problems, the algorithm may be struggling to find a global minimum

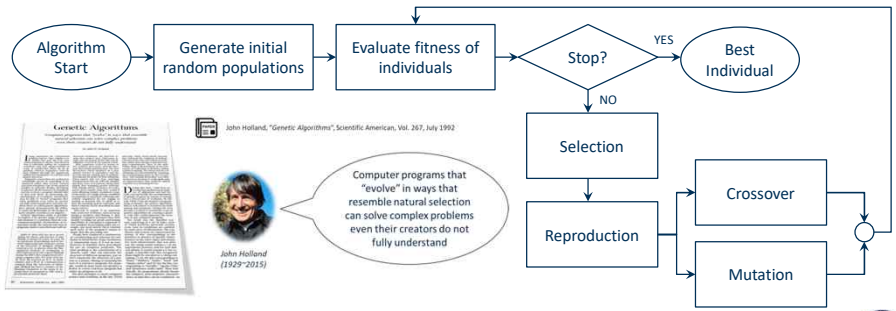


Recall that when the slope of the objective function is at or close to zero, the algorithm stops finding an optimum

https://www.ibm.com/cloud/learn/gradient-descent

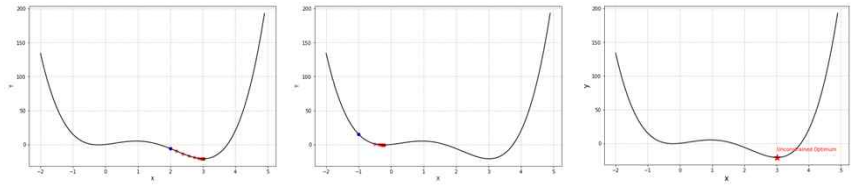
Genetic algorithm

- The Genetic Algorithm (GA), inspired by biological principles, may address the limitations of the gradient descent algorithm



Genetic algorithm

- Gradient descent algorithm vs. Genetic algorithm
  - The gradient descent algorithm is effective only when the initial point is well-defined
  - The genetic algorithm provides a global optimum (or near-optimum) regardless of the initial point



Gradient descent algorithm with proper initial point location

Gradient descent algorithm with improper initial point location

The genetic algorithm is covered in the Algorithm (AIX30006) class

Course summary

- Throughout this module, you have learned:
  - Where does the gradient descent algorithm come from?
  - What is the gradient descent algorithm?
  - What are the limitations of the gradient descent algorithm?

THANK YOU

For more information, please reach out to Prof. Junghyun Kim at
junghyun.kim@handong.edu