Tutorial: Gradient Descent

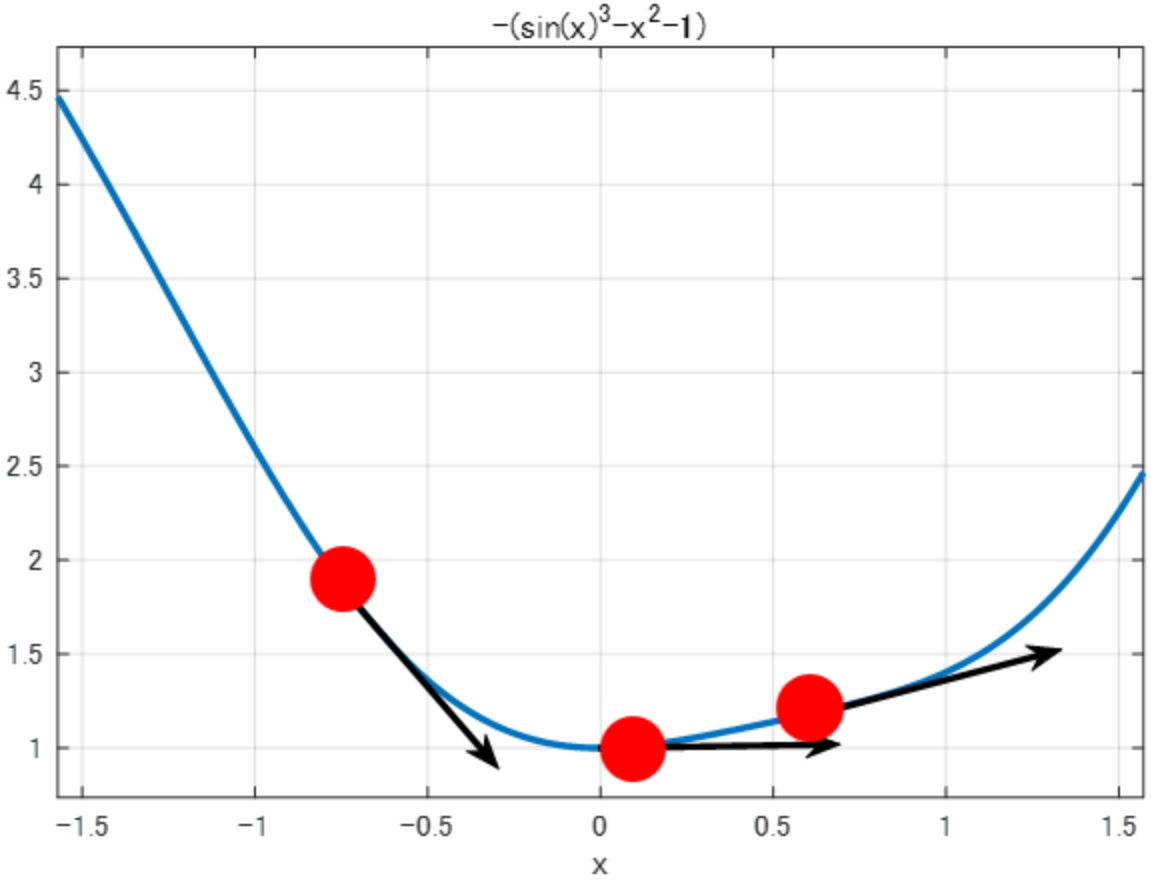
Song Liu (song.liu@bristol.ac.uk)
GA 18, Fry Building,
Microsoft Teams (search "song liu").

Finding the Minimum of a Function

- Given a function $f(x)$, finding its minimum value.
  - Easy when $f(x)$ takes some simple form (quadratic).
  - Difficult when $f(x)$ is more complicated.
    - $f(x) := \sin(x)^3 + x^2 + 1, x \in [-\pi/2, \pi/2]$
- This is a fundamental question in the field **optimization** and has many applications in data science.

Idea: Gradient Descent

- The derivative of a function evaluated at x is positive when the function is ascending at x and vice versa.
- The derivative is 0 when the function is (locally) minimum



Algorithm: Gradient Descent

- Start from an arbitrary location x.
- Repeat the following, until $\left|f'(x)\right| \leq \epsilon$
  - Subtract the derivative from x:
    - $x \leftarrow x - \eta \cdot f'(x)$
    - η is a small value, say 0.1.
- If f is ascending at x, subtract the derivative will move "backwards" and vice versa.
- It stops when $f'(x) \approx 0$.

Algorithm: Gradient Descent

- Write a R program to find the local minimum of
  - $f(x) := \sin(x)^3 + x^2 + 1, x \in [-\pi/2, \pi/2]$.
- $f'(x)$ is given below

```
df <- function(x){
    return(3*sin(x)^2*cos(x) + 2*x)
}
```

- Write **no more than 5 lines** of R code.
  - Hint: use `while` loop.