Introduction to Machine Learning

TA Session - 1

14th Aug 2023

Some general info

- **Official Github Repository:**
  https://github.com/sarthakharne/Machine-Learning-TA-Material-Fall-2023

- **Where you can reach me:**
  Whatsapp or Mail

- **Find Good Resources?**
  Share it with the batch!

Some questions you might have:

- **Do I need a fancy computer? What interfaces are we going to use?**
  No!

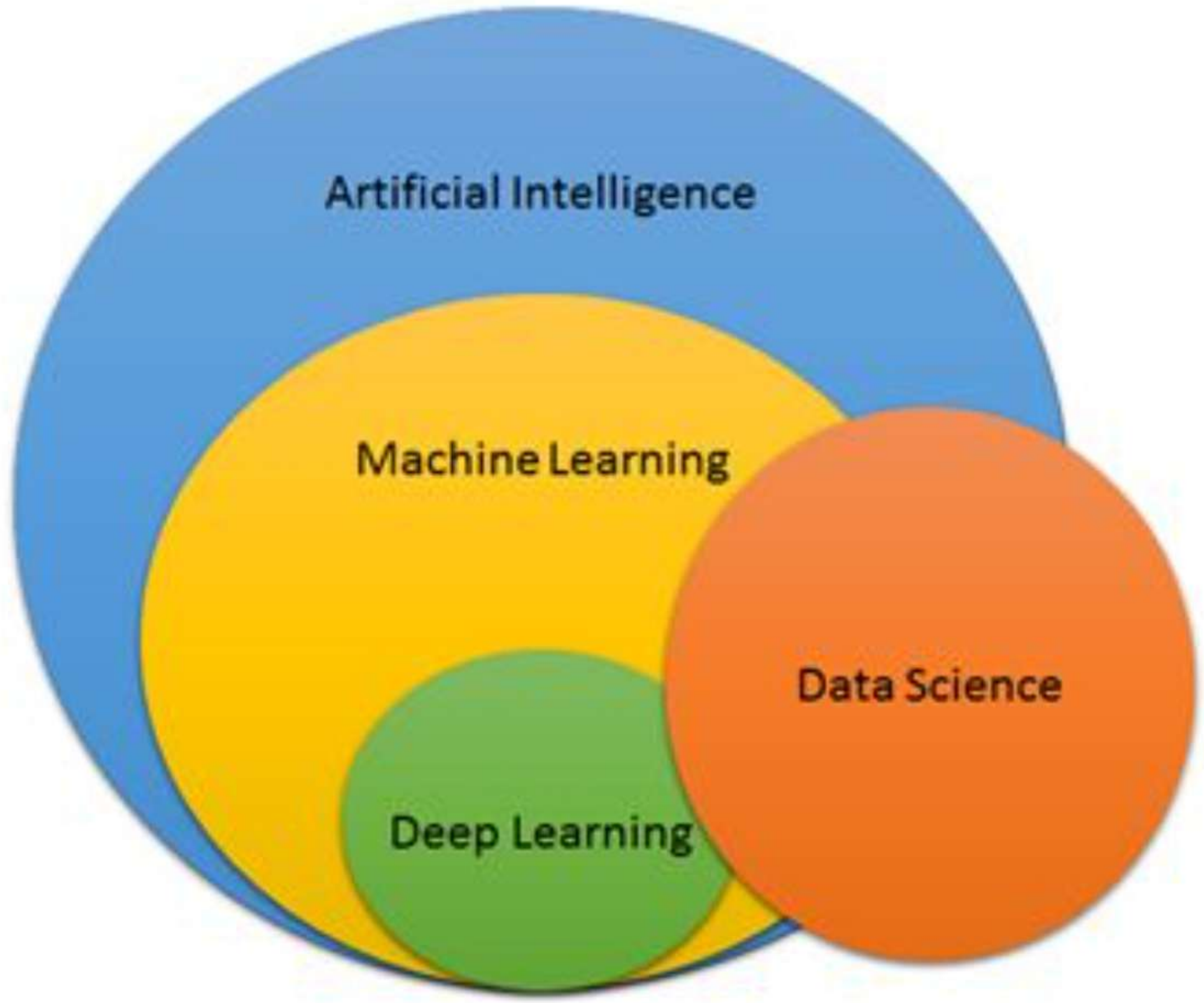
- **Do I need to be good at programming? (DSA trauma… I understand)**
  Don’t panic! Basics required.

- **Do I need to be a pro at maths?**
  Nope.

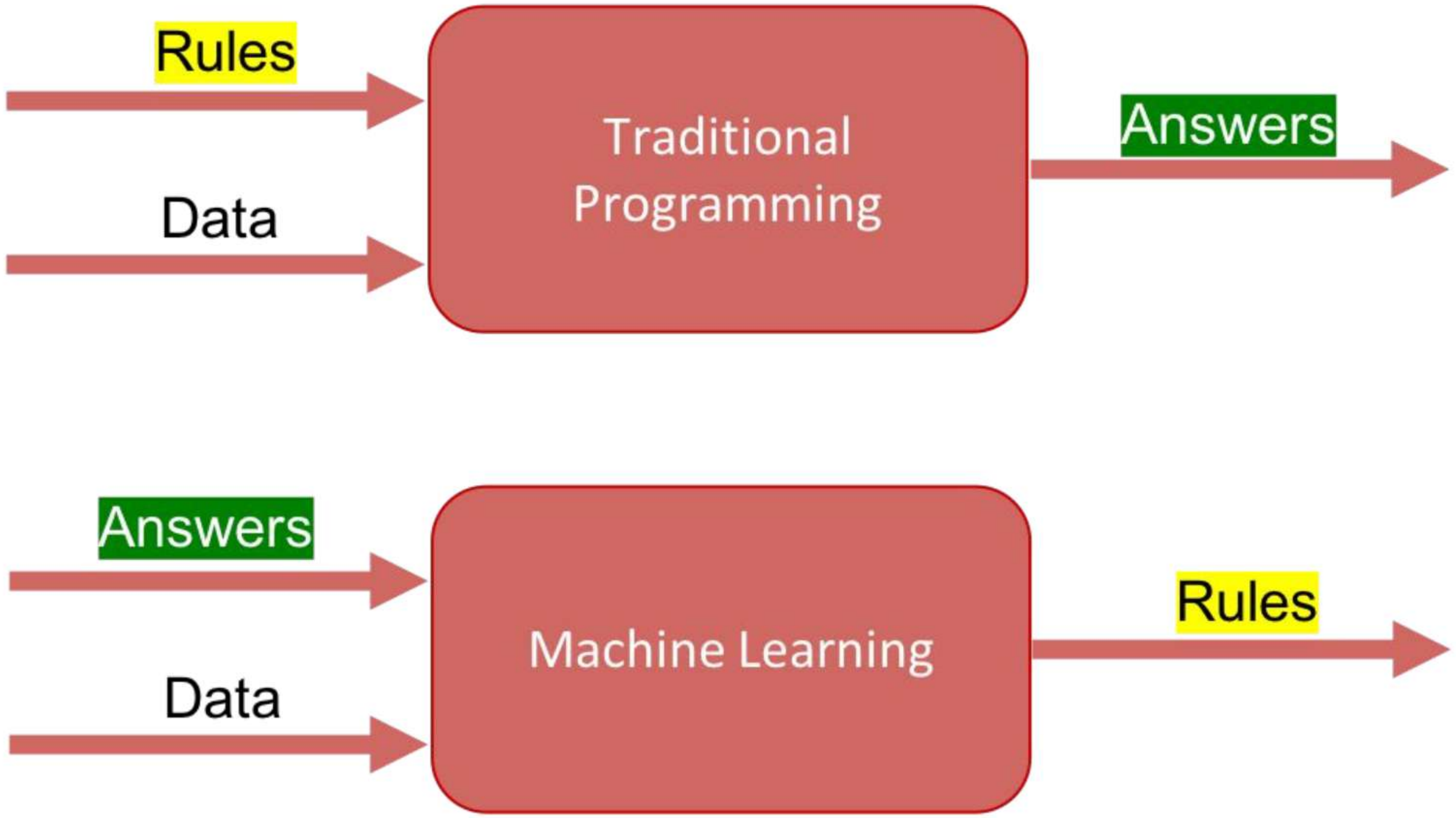
What is ML?

Another word you might've heard being thrown around...

AI?



Source: https://medium.com/@dilip.rajani/comparing-ds-ml-dl-and-ai-65627109e67a

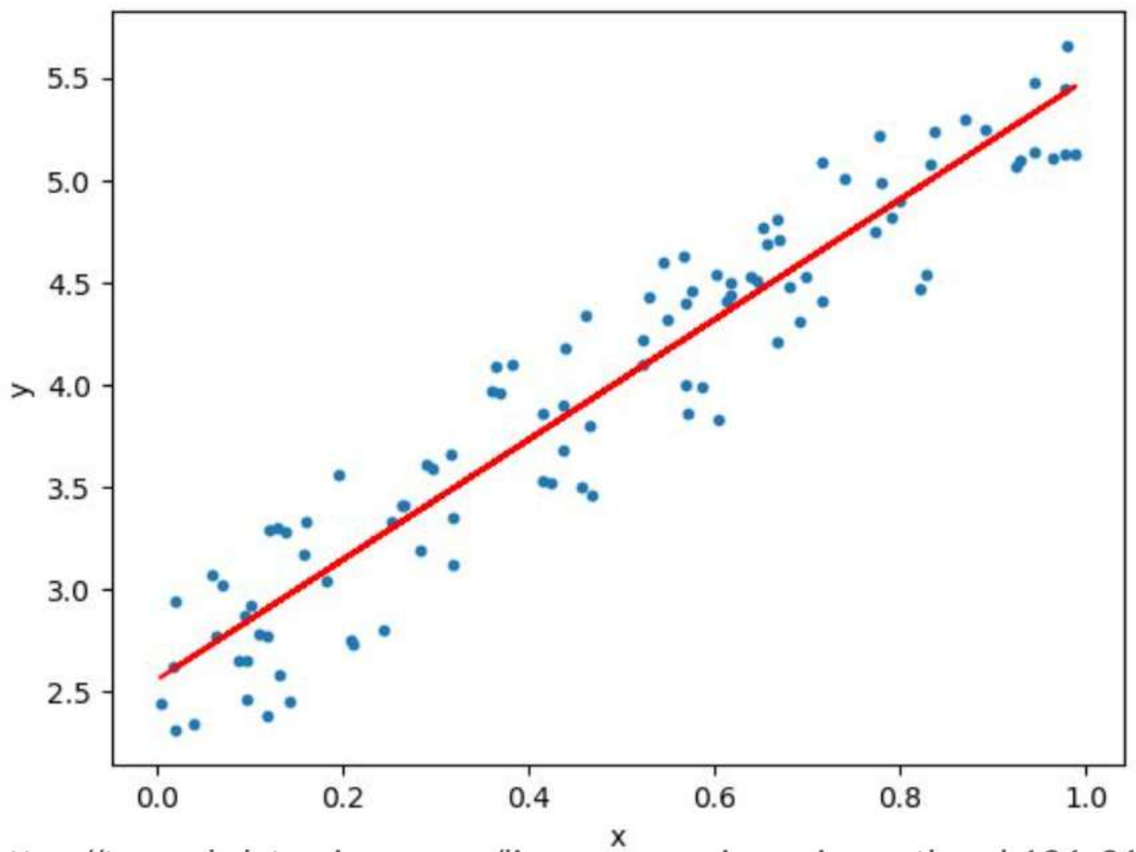


https://www.congrelate.com/15-programming-without-machine-learning-images/

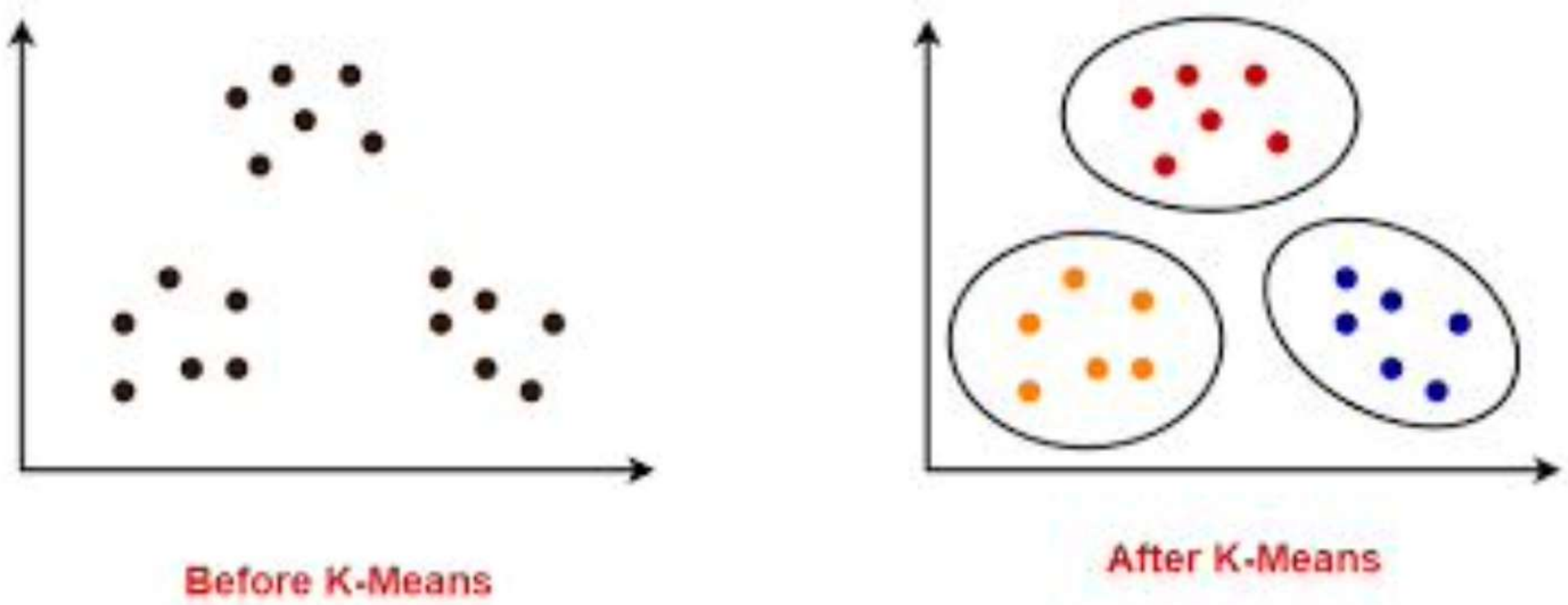
Supervised v/s Unsupervised Learning

- Where your data has labels, it is supervised learning i.e. someone is telling you what the true label is for the training data.
- Simple linear regression is an example.

- Unsupervised learning, there are no labels at all. No ‘teacher’. No ground rules. No true labels.
- Clustering is an example.



https://towardsdatascience.com/linear-regression-using-python-b136c91bf0a2



https://www.gatevidyalay.com/k-means-clustering-algorithm-example/

Jargon to keep in mind

- Tasks
- Datasets
- Features
- Models

Formal definitions

- **Tasks** - Type of prediction being made, based on the question that is being asked and the available data.
- **Datasets** - The raw data available.
- **Features** - The factors or characteristics taken into consideration.
- **Models** - Almost like a function. Algorithms are used to tune this model.

Linear Regression Intro

Q1: Statement

Minimize the expression $\frac{x^2}{4} - 2x + 5$ with respect to x using

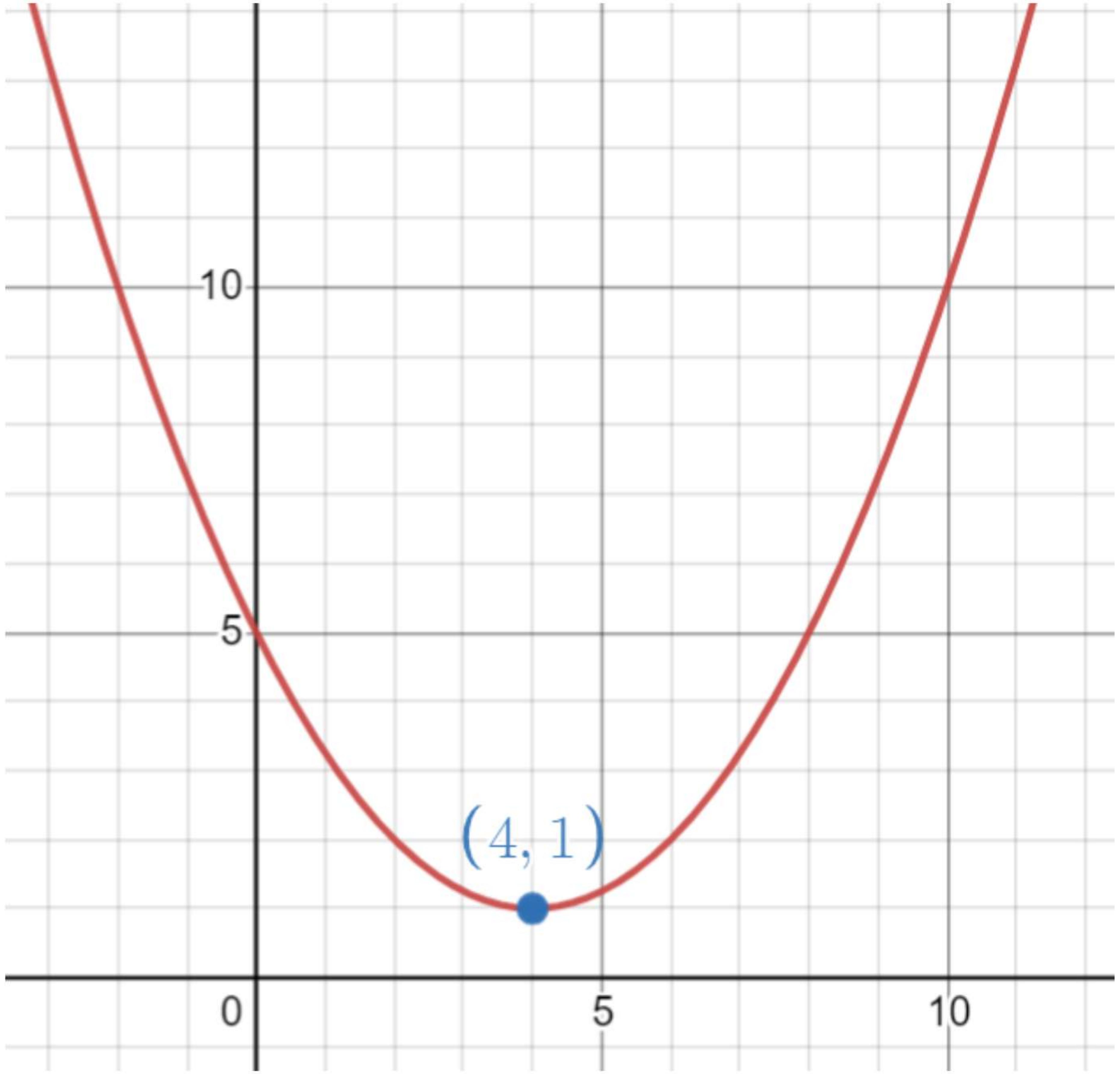
1. The closed-form solution
2. Gradient descent

Q1: Closed-form solution

$$f(x) = \frac{x^2}{4} - 2x + 5$$

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2 = 0$$

$$x = 4$$



Q1: Run Through Gradient Descent

Let us set $x = 10$ initially. Let $\alpha = 0.1$

$$f(x) = \frac{x^2}{4} - 2x + 5$$

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

$$\begin{aligned} f(10) &= \frac{10^2}{4} - 2 * 10 + 5 \\ &= 25 - 20 + 5 \\ &= 10 \end{aligned}$$

x	$f(x)$
10	10

Q1: Run Through Gradient Descent

Let us set $x = 10$ initially. Let $\alpha = 0.1$

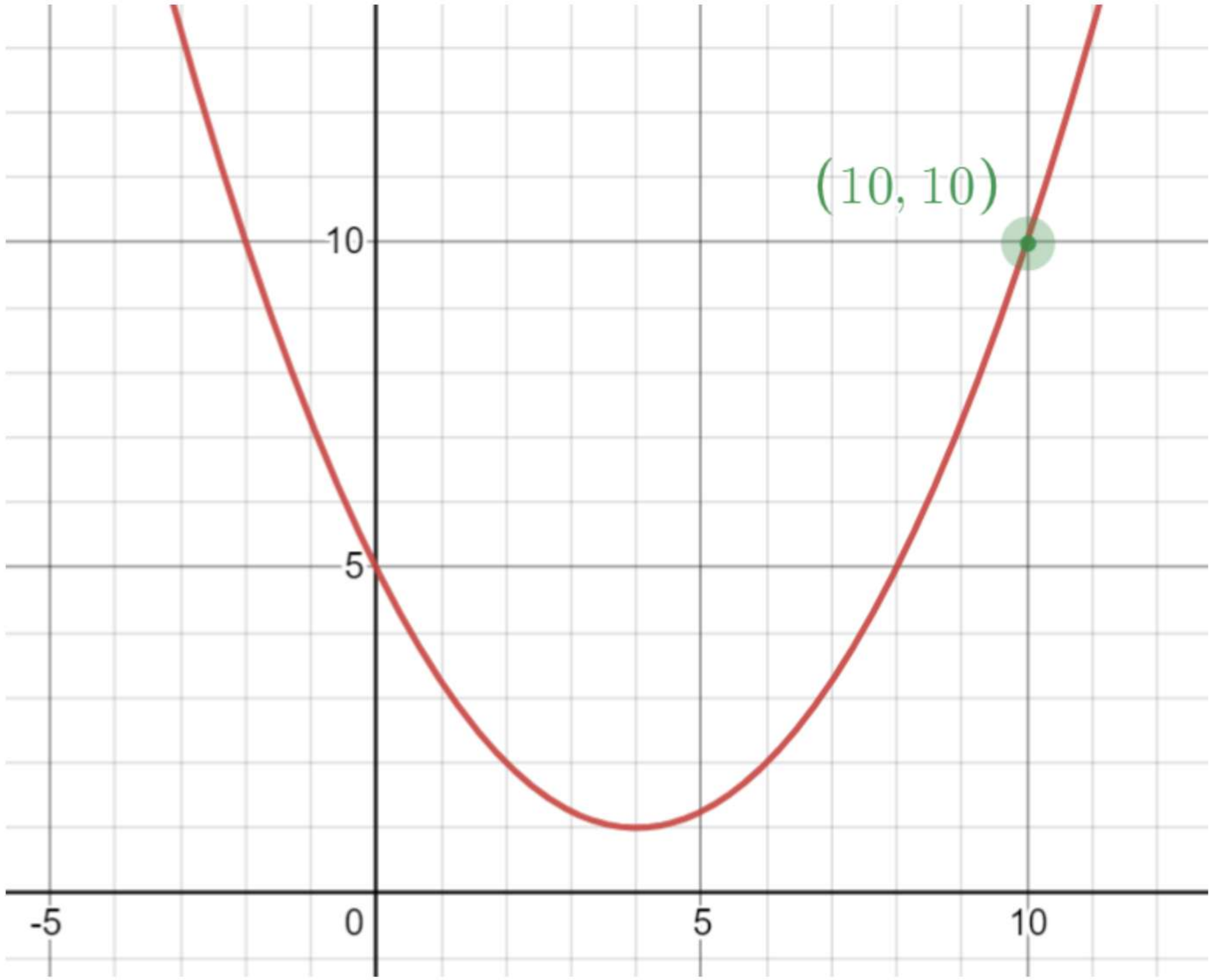
x	$f(x)$
10	10

$$f(x) = \frac{x^2}{4} - 2x + 5$$

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

$$\begin{aligned} f(10) &= \frac{10^2}{4} - 2 * 10 + 5 \\ &= 25 - 20 + 5 \\ &= 10 \end{aligned}$$

Q1: Run Through Gradient Descent



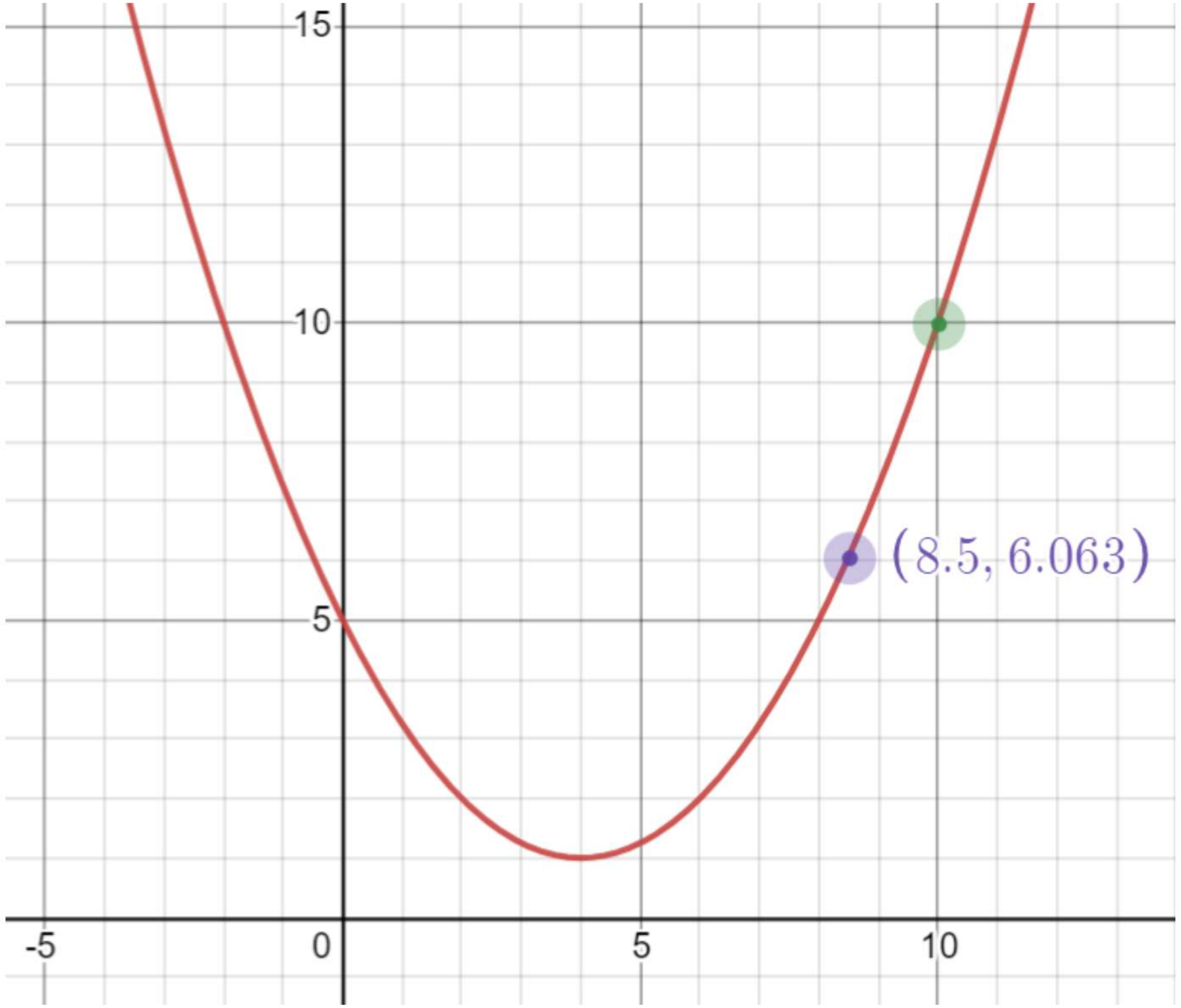
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

x	$f(x)$
10	10
8.5	6.025

$$\begin{aligned} f'(10) &= \frac{10}{2} - 2 \\ &= 3 \\ x_{next} &= x - \alpha f'(x) \\ &= 10 - 0.5 * 3 \\ &= 8.5 \\ f(8.5) &= \frac{8.5^2}{4} - 2 * 8.5 + 5 \\ &= 6.0625 \end{aligned}$$

Q1: Run Through Gradient Descent



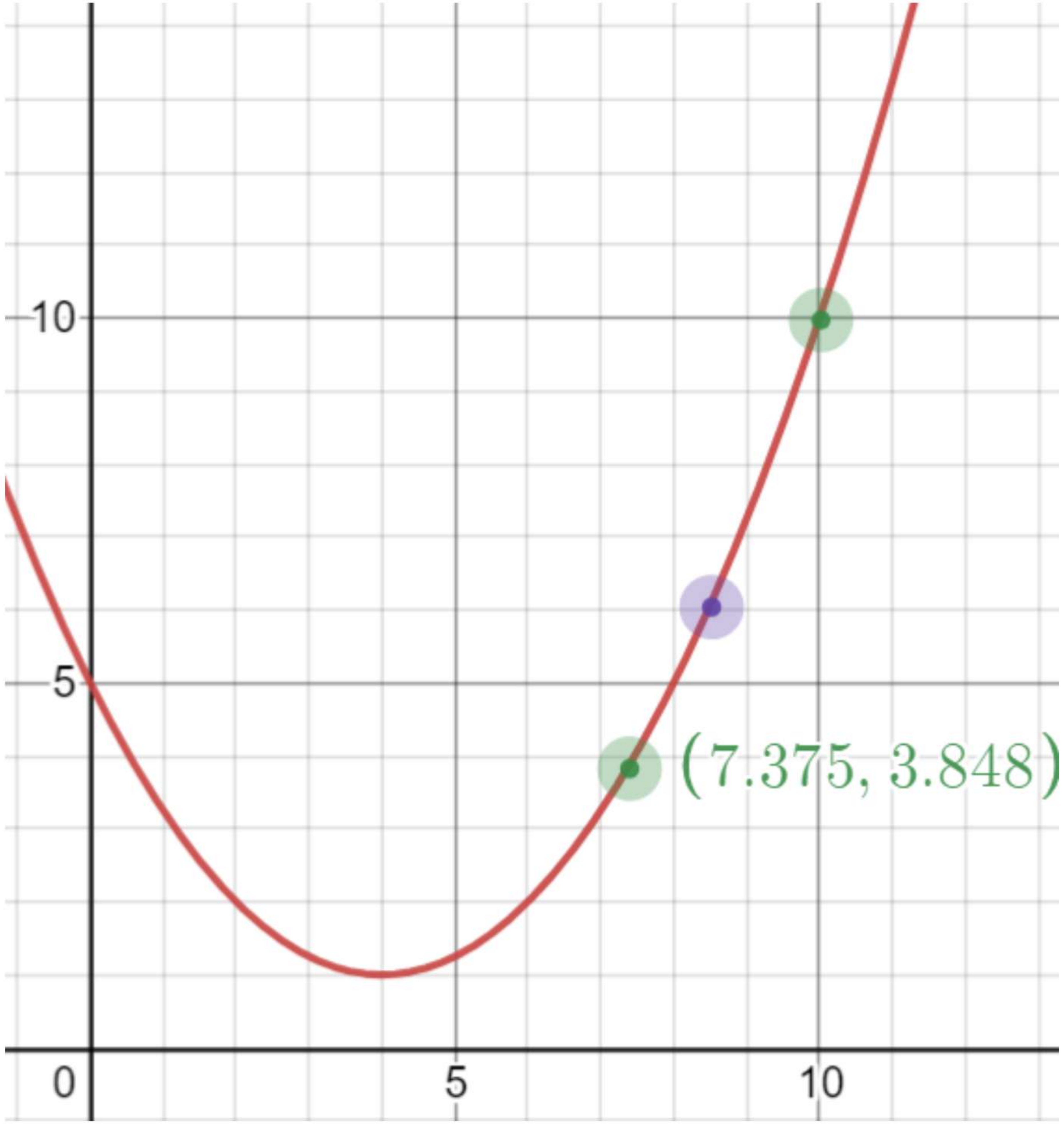
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

x	$f(x)$
10	10
8.5	6.025
7.375	3.848

$$\begin{aligned} f'(8.5) &= \frac{8.5}{2} - 2 \\ &= 2.25 \\ x_{next} &= x - \alpha f'(x) \\ &= 8.5 - 0.5 * 2.25 \\ &= 7.375 \\ f(7.375) &= \frac{7.375^2}{4} - 2 * 7.375 + 5 \\ &= 3.848 \end{aligned}$$

Q1: Run Through Gradient Descent



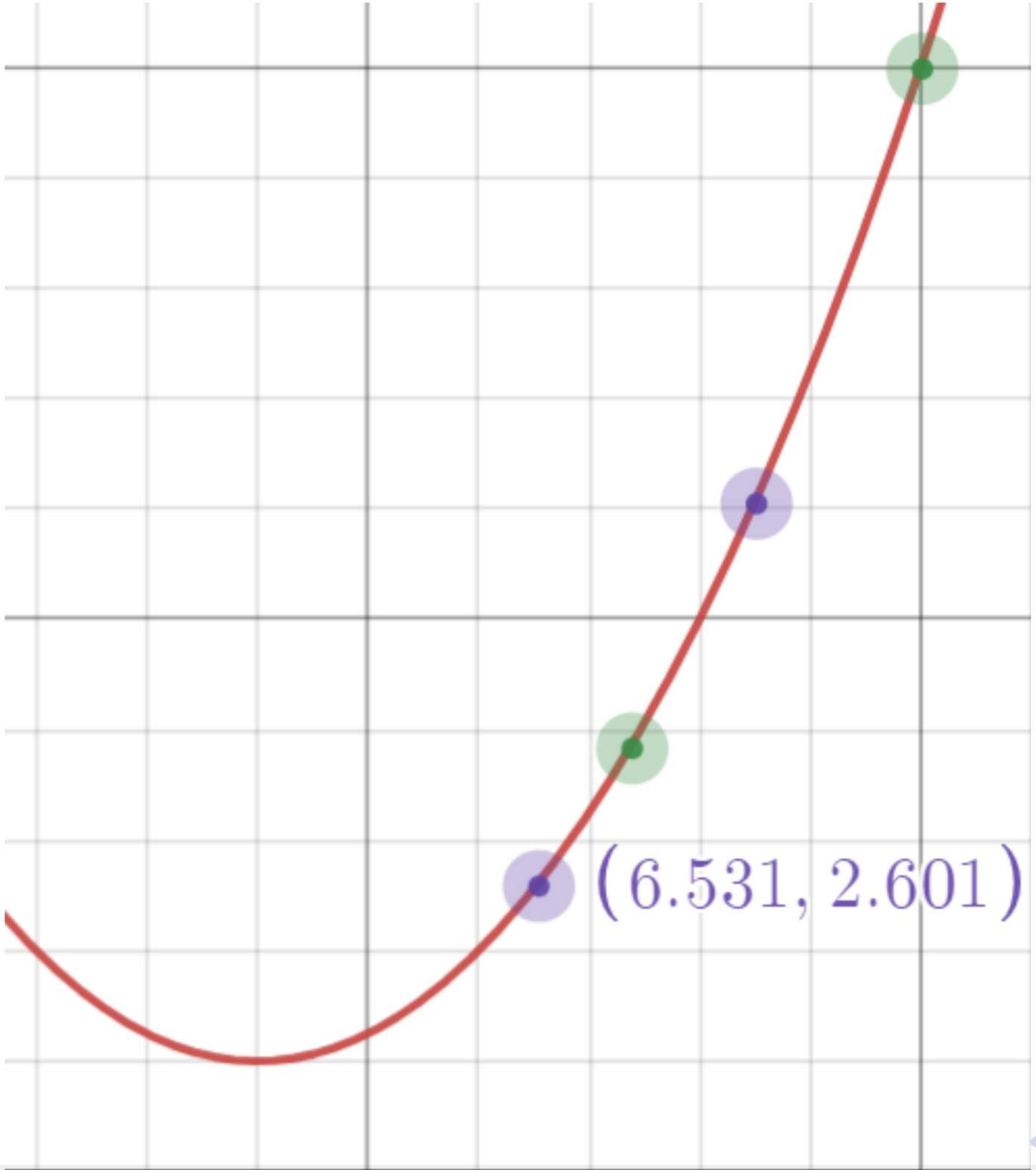
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

x	$f(x)$
10	10
8.5	6.025
7.375	3.848
6.531	2.601

$$\begin{aligned}
f'(7.375) &= \frac{7.375}{2} - 2 \\
&= 1.688 \\
x_{next} &= x - \alpha f'(x) \\
&= 7.375 - 0.5 * 1.688 \\
&= 6.531 \\
f(6.531) &= \frac{6.531^2}{4} - 2 * 6.531 + 5 \\
&= 2.601
\end{aligned}$$

Q1: Run Through Gradient Descent



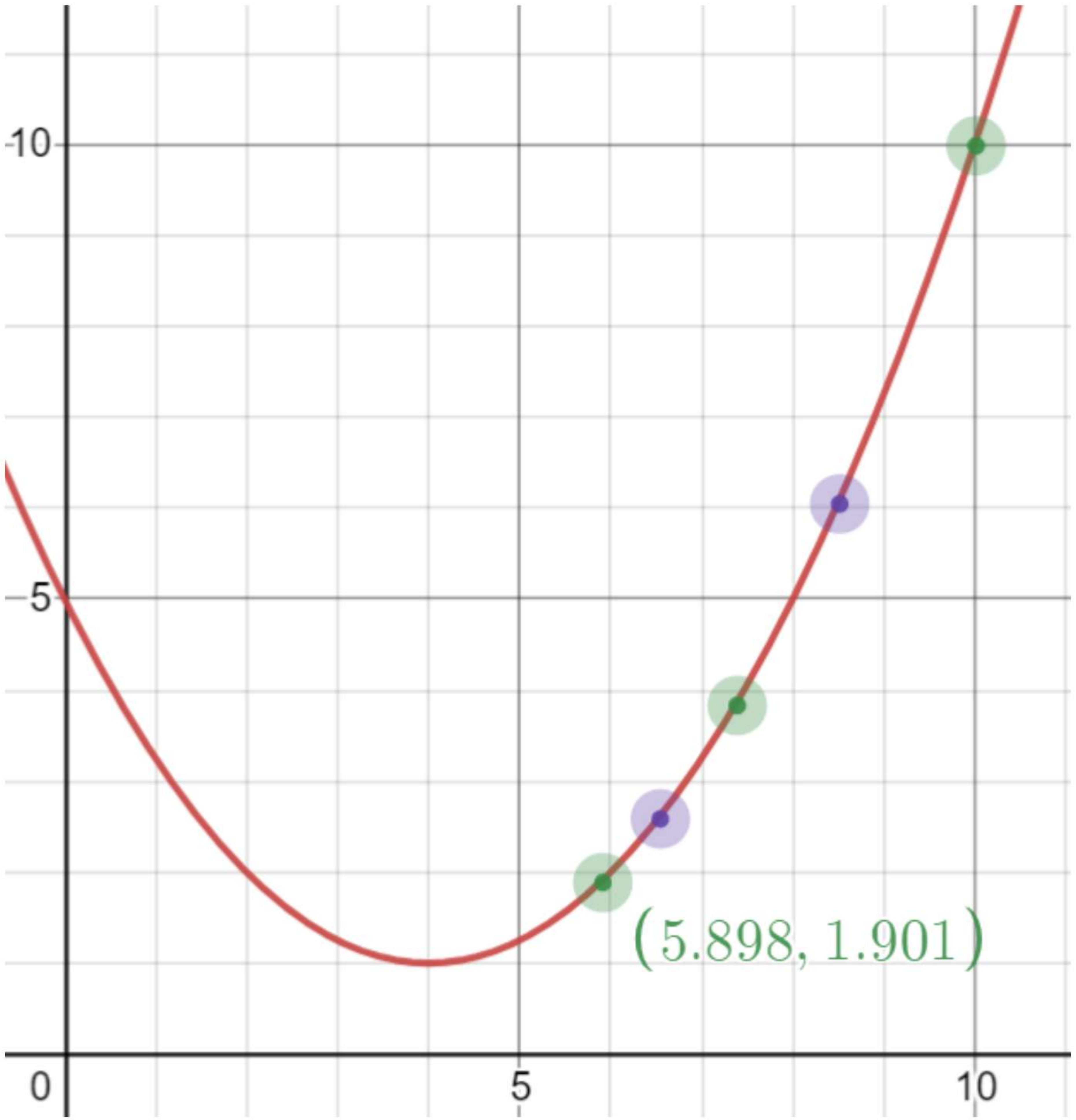
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{df(x)}{dx} = \frac{x}{2} - 2$$

x	$f(x)$
10	10
8.5	6.025
7.375	3.848
6.531	2.601
5.898	1.901

$$\begin{aligned} f'(6.531) &= \frac{6.531}{2} - 2 \\ &= 1.266 \\ x_{next} &= x - \alpha f'(x) \\ &= 6.531 - 0.5 * 1.266 \\ &= 5.898 \\ f(5.898) &= \frac{5.898^2}{4} - 2 * 5.898 + 5 \\ &= 1.901 \end{aligned}$$

Q1: Run Through Gradient Descent



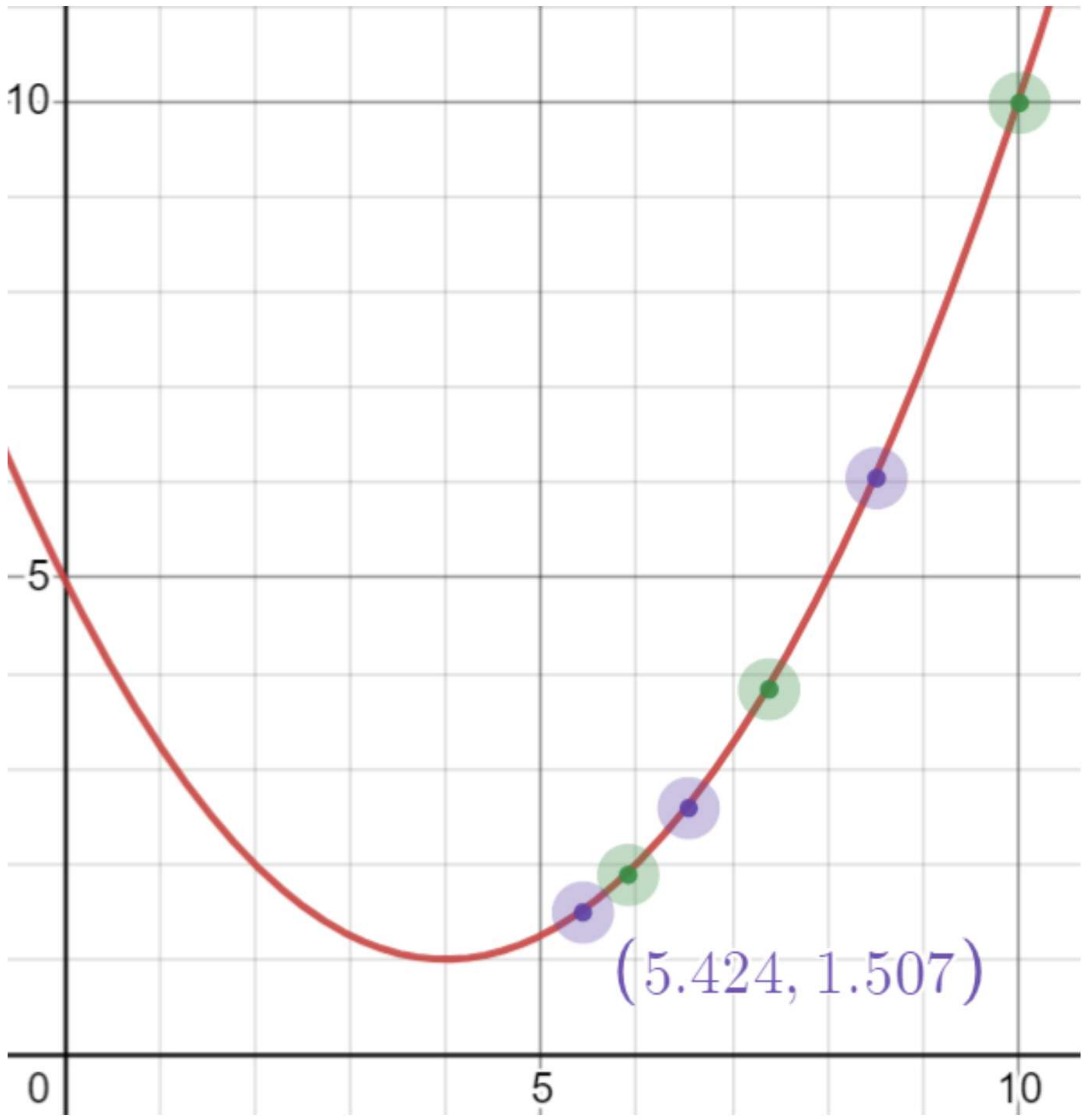
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{df(x)}{dx} = \frac{x}{2} - 2$$

$$\begin{aligned}
f'(5.898) &= \frac{5.898}{2} - 2 \\
&= 0.949 \\
x_{next} &= x - \alpha f'(x) \\
&= 5.898 - 0.5 * 0.949 \\
&= 5.424 \\
f(5.424) &= \frac{5.424^2}{4} - 2 * 5.424 + 5 \\
&= 1.507
\end{aligned}$$

x	$f(x)$
10	10
8.5	6.025
7.375	3.848
6.531	2.601
5.898	1.901
5.424	1.507

Q1: Run Through Gradient Descent



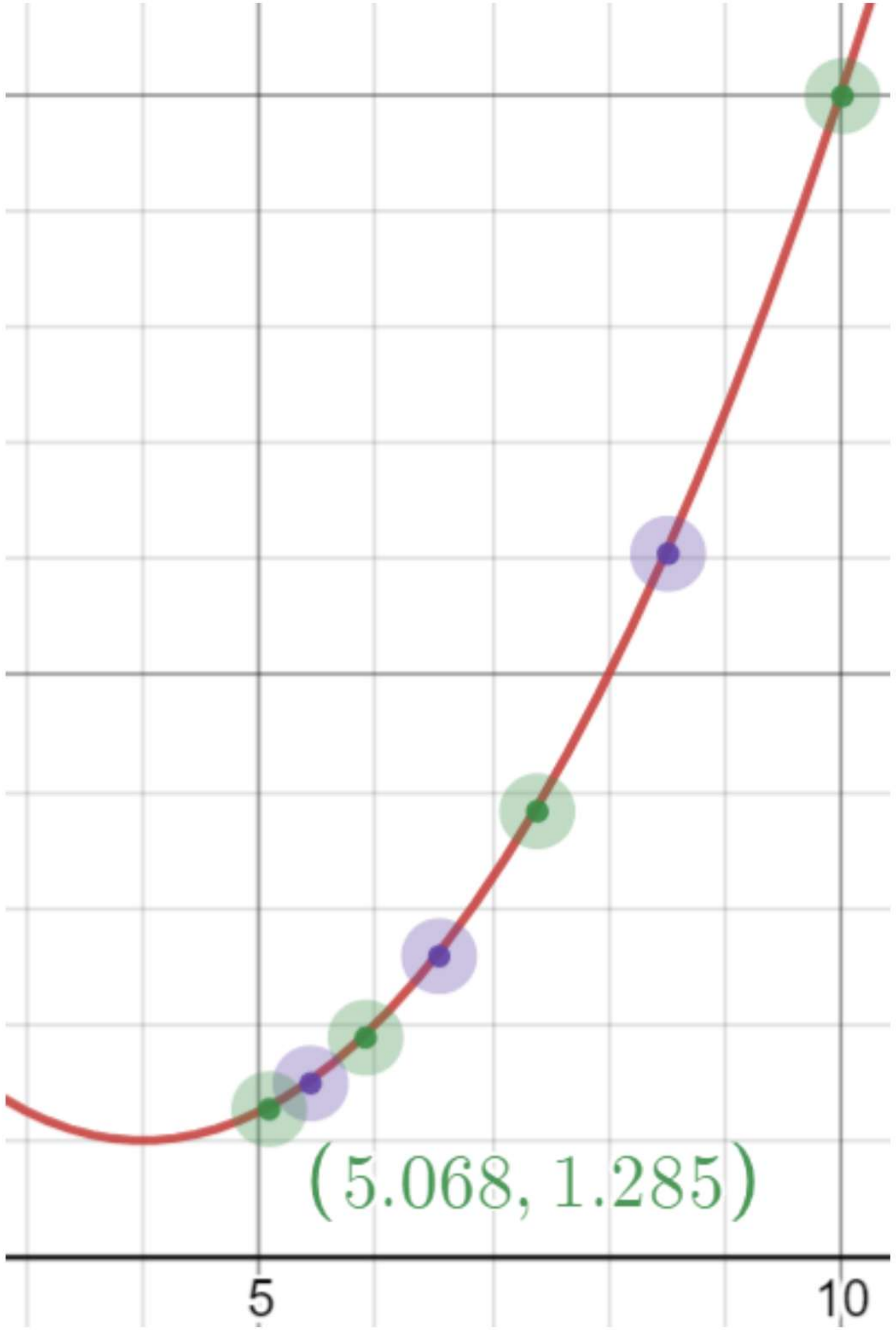
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

$$\begin{aligned} f'(5.424) &= \frac{5.424}{2} - 2 \\ &= 0.712 \\ x_{next} &= x - \alpha f'(x) \\ &= 5.424 - 0.5 * 0.712 \\ &= 5.068 \\ f(5.068) &= \frac{5.068^2}{4} - 2 * 5.068 + 5 \\ &= 1.285 \end{aligned}$$

x	$f(x)$
10	10
8.5	6.025
7.375	3.848
6.531	2.601
5.898	1.901
5.424	1.507
5.068	1.285

Q1: Run Through Gradient Descent



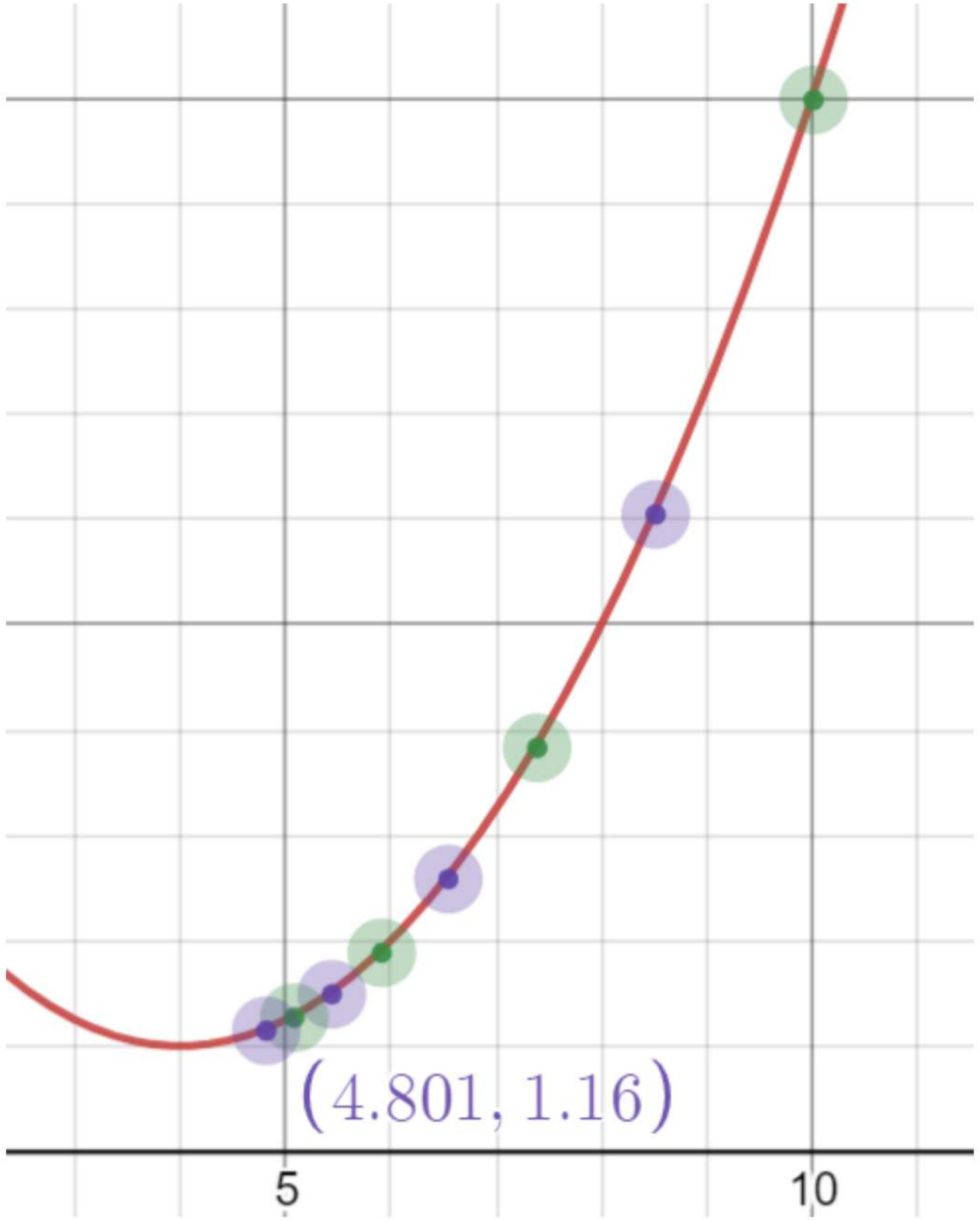
Q1: Run Through Gradient Descent

$$\alpha = 0.5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

$$\begin{aligned}
f'(5.068) &= \frac{5.068}{2} - 2 \\
&= 0.534 \\
x_{next} &= x - \alpha f'(x) \\
&= 5.068 - 0.5 * 0.534 \\
&= 4.801 \\
f(4.801) &= \frac{4.801^2}{4} - 2 * 4.801 + 5 \\
&= 1.16
\end{aligned}$$

x	$f(x)$
10	10
8.5	6.025
7.375	3.848
6.531	2.601
5.898	1.901
5.424	1.507
5.068	1.285
4.801	1.16

Q1: Run Through Gradient Descent



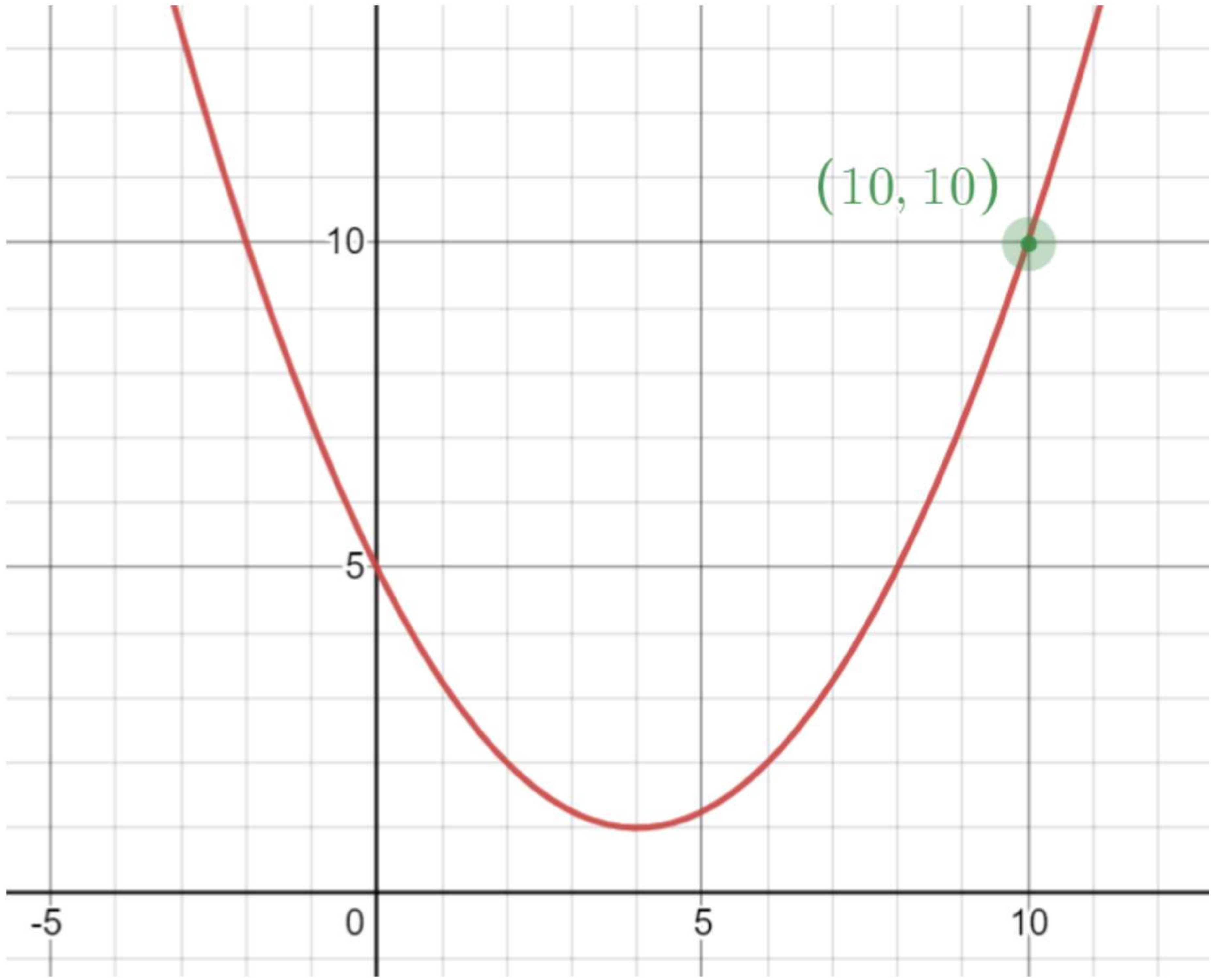
Q1: Gradient Descent With Excess Learning Rate

Let us set $x = 10$ initially. Let $\alpha = 5$

$$
\begin{aligned}
f(x) &= \frac{x^2}{4} - 2x + 5 \\
\frac{df(x)}{dx} &= \frac{x}{2} - 2 \\
f(10) &= \frac{10^2}{4} - 2 * 10 + 5 \\
&= 25 - 20 + 5 \\
&= 10
\end{aligned}
$$

x	$f(x)$
10	10

Q1: Gradient Descent With Excess Learning Rate



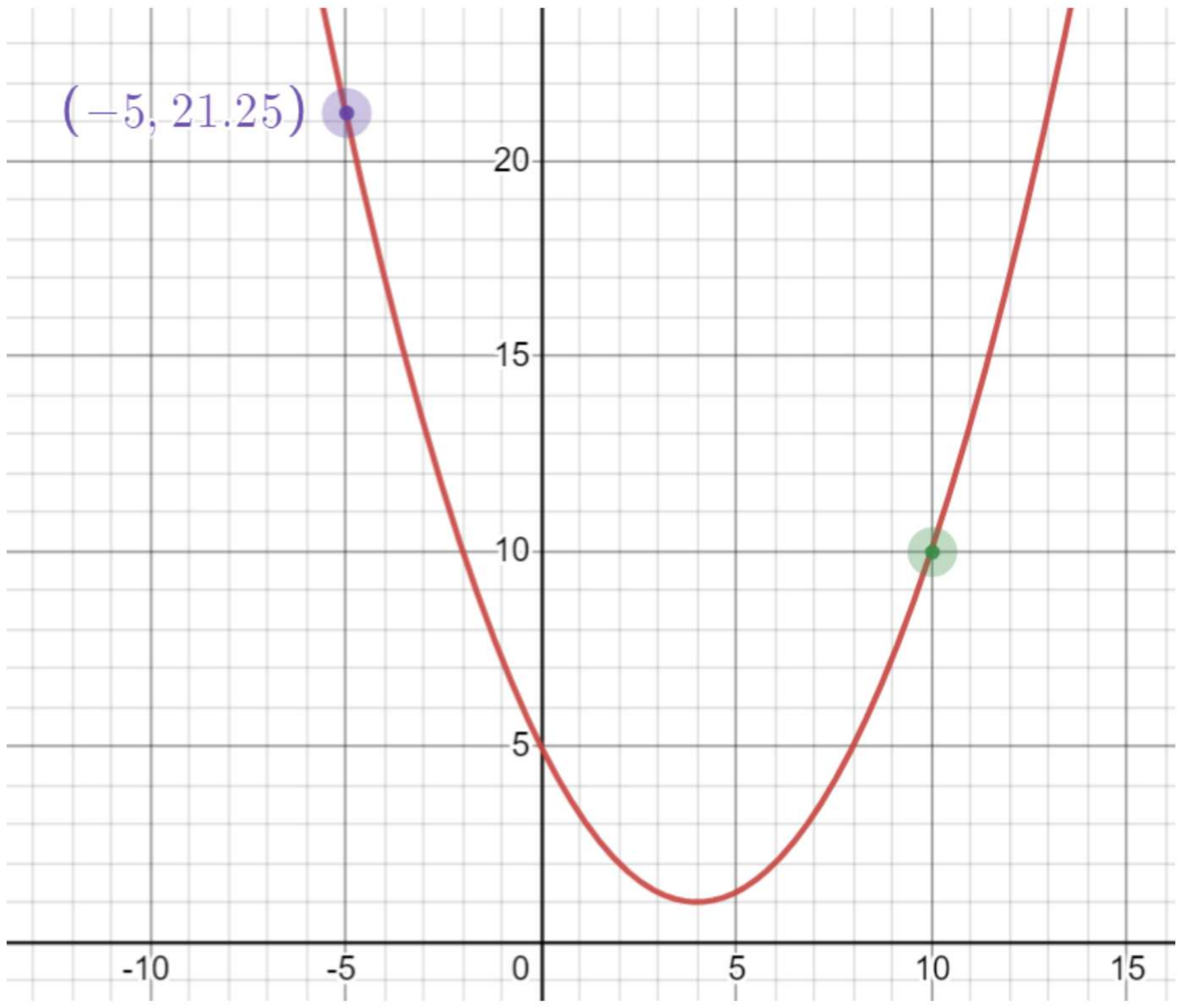
Q1: Run Through Gradient Descent

$$\alpha = 5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

x	$f(x)$
10	10
-5	21.25

$$\begin{aligned}
f'(10) &= \frac{10}{2} - 2 \\
&= 3 \\
x_{next} &= x - \alpha f'(x) \\
&= 10 - 5 * 3 \\
&= -5 \\
f(-5) &= \frac{(-5)^2}{4} - 2 * (-5) + 5 \\
&= 21.25
\end{aligned}$$

Q1: Gradient Descent With Excess Learning Rate



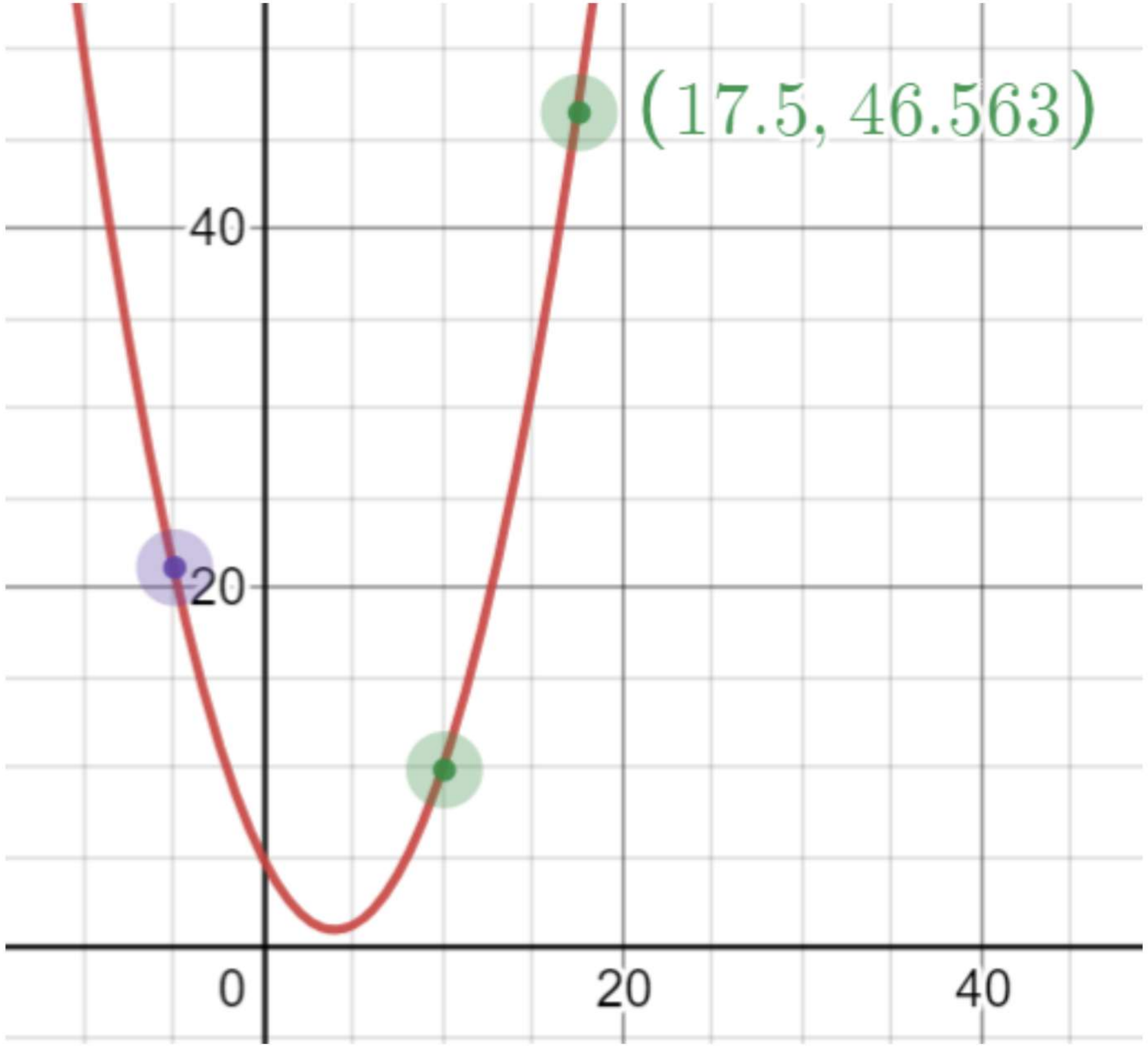
Q1: Run Through Gradient Descent

$$\alpha = 5, f(x) = \frac{x^2}{4} - 2x + 5, \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{x}{2} - 2$$

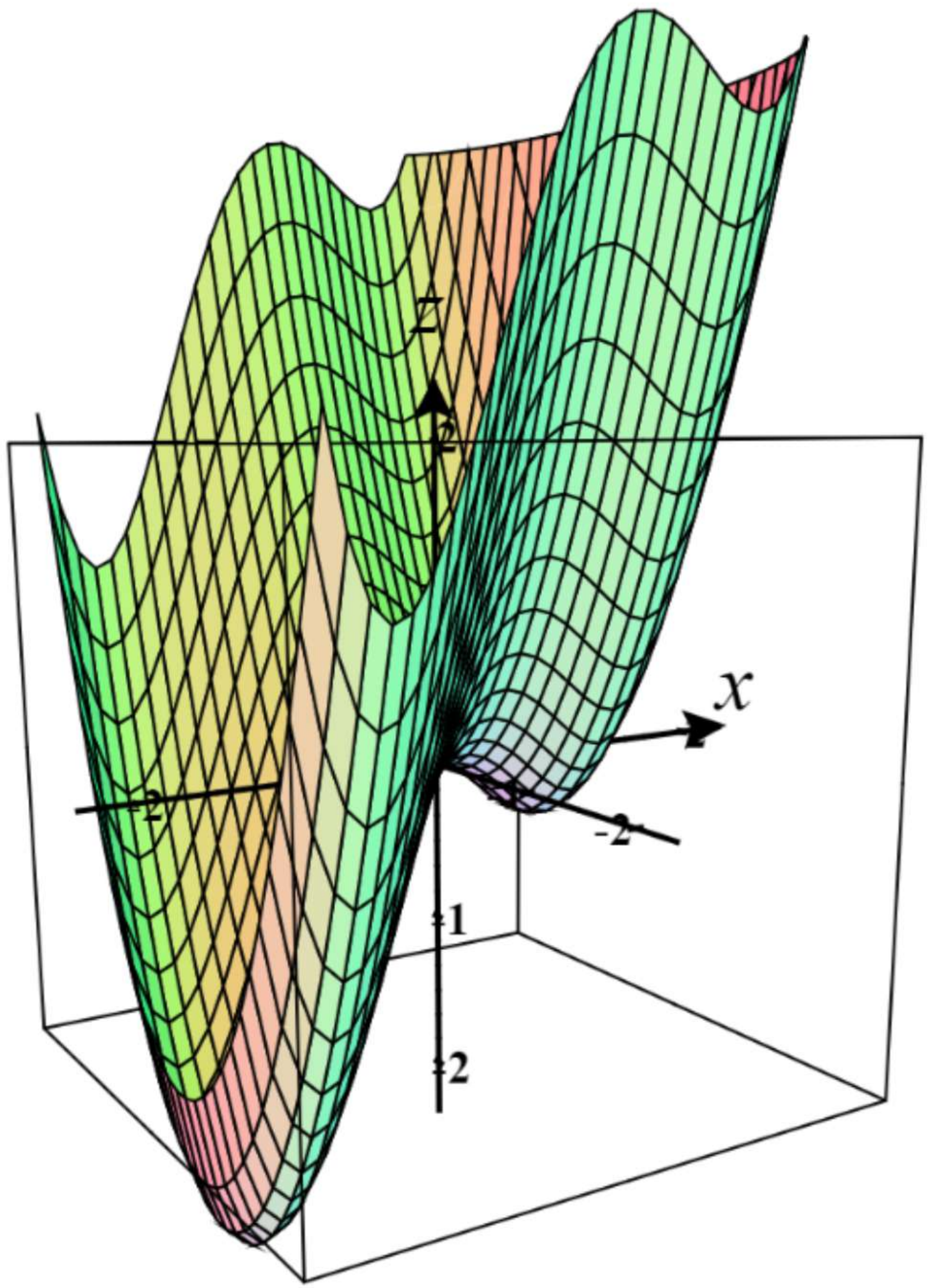
x	$f(x)$
10	10
-5	21.25
17.5	46.563

$$\begin{aligned} f'(-5) &= \frac{-5}{2} - 2 \\ &= -4.5 \\ x_{next} &= x - \alpha f'(x) \\ &= -5 - 5 * (-4.5) \\ &= 17.5 \\ f(17.5) &= \frac{17.5^2}{4} - 2 * 17.5 + 5 \\ &= 46.563 \end{aligned}$$

Q1: Gradient Descent With Excess Learning Rate



Q2: Statement



Find the minimum of $f(x, y) = x^4 + x^3 - 2x^2 + y^2$, using the closed-form solution and using gradient descent. Interact with the 3D plot here.

Q2: Closed-form solution

$$f(x,y) = x^4 + x^3 - 2x^2 + y^2$$

$$\frac{\partial f(x,y)}{\partial x} = 4x^3 + 3x^2 - 4x = 0$$

$$\therefore x = -1.443, 0, 0.693$$

$$\frac{\partial f(x,y)}{\partial y} = 2y = 0$$

$$\therefore y = 0$$

$$f(-1.443, 0) = -2.833$$

$$f(0,0) = 0$$

$$f(0.693, 0) = -0.39$$

$\therefore$ the minimum of $f(x,y)$ is at$(-1.443, 0)$

Q2: Run Through Gradient Descent

Let's start from $(-0.5, 4.5)$ with $\alpha = 0.1$.

x	y	$f(x,y)$
-0.5	4.5	19.688
-0.725	3.6	11.80

$$\alpha = 0.1$$
$$f(x,y) = x^4 + x^3 - 2x^2 + y^2$$
$$\frac{\partial f(x,y)}{\partial x} = 4x^3 + 3x^2 - 4x$$
$$\frac{\partial f(x,y)}{\partial y} = 2y$$

$$\begin{aligned} f(-0.5, 4.5) &= 19.688 \\ (\nabla_{x,y} f)(-0.5, 4.5) &= (2.25, 9) \\ (x_{next}, y_{next}) &= (-0.5, 4.5) - \alpha(\nabla_{x,y} f)(-0.5, 4.5) \\ &= (-0.725, 3.6) \\ f(-0.725, 3.6) &= 11.80 \end{aligned}$$

Q2: Run Through Gradient Descent

$$\alpha = 0.1$$
$$f(x,y) = x^4 + x^3 - 2x^2 + y^2$$
$$\frac{\partial f(x,y)}{\partial x} = 4x^3 + 3x^2 - 4x$$
$$\frac{\partial f(x,y)}{\partial y} = 2y$$

x	y	$f(x,y)$
-0.5	4.5	19.688
-0.725	3.6	11.80
-1.02	2.88	6.23

$$f(-0.725, 3.6) = 11.80$$
$$(\nabla_{x,y} f)(-0.725, 3.6) = (2.952, 7.2)$$
$$(x_{next}, y_{next}) = (-0.725, 3.6) - \alpha(\nabla_{x,y} f)(-0.725, 3.6)$$
$$= (-1.02, 2.88)$$
$$f(-1.02, 2.88) = 6.23$$

Q2: Run Through Gradient Descent

$$\alpha = 0.1$$
$$f(x, y) = x^4 + x^3 - 2x^2 + y^2$$
$$\frac{\partial f(x, y)}{\partial x} = 4x^3 + 3x^2 - 4x$$
$$\frac{\partial f(x, y)}{\partial y} = 2y$$

x	y	$f(x, y)$
-0.5	4.5	19.688
-0.725	3.6	11.80
-1.02	2.88	6.23
-1.31	2.30	2.56

$$f(-1.02, 2.88) = 6.23$$
$$(\nabla_{x,y} f)(-1.02, 2.88) = (2.96, 5.76)$$
$$(x_{next}, y_{next}) = (-1.02, 2.88) - \alpha(\nabla_{x,y} f)(-1.02, 2.88)$$
$$= (-1.31, 2.30)$$
$$f(-1.31, 2.30) = 2.56$$

Q2: Run Through Gradient Descent

$$\alpha = 0.1$$
$$f(x, y) = x^4 + x^3 - 2x^2 + y^2$$
$$\frac{\partial f(x, y)}{\partial x} = 4x^3 + 3x^2 - 4x$$
$$\frac{\partial f(x, y)}{\partial y} = 2y$$

x	y	$f(x, y)$
-0.5	4.5	19.688
-0.725	3.6	11.80
-1.02	2.88	6.23
-1.31	2.30	2.56
-1.45	1.84	0.55

$$f(-1.31, 2.30) = 2.56$$
$$(\nabla_{x,y} f)(-1.31, 2.30) = (1.39, 4.6)$$
$$(x_{next}, y_{next}) = (-1.31, 2.30) - \alpha(\nabla_{x,y} f)(-1.31, 2.30)$$
$$= (-1.45, 1.84)$$
$$f(-1.45, 1.84) = 0.55$$

Q2: Run Through Gradient Descent

$$\alpha = 0.1$$
$$f(x,y) = x^4 + x^3 - 2x^2 + y^2$$
$$\frac{\partial f(x,y)}{\partial x} = 4x^3 + 3x^2 - 4x$$
$$\frac{\partial f(x,y)}{\partial y} = 2y$$

x	y	$f(x,y)$
-0.5	4.5	19.688
-0.725	3.6	11.80
-1.02	2.88	6.23
-1.31	2.30	2.56
-1.45	1.84	0.55
-1.44	1.47	-0.67

$$f(-1.45, 1.84) = 0.55$$
$$(\nabla_{x,y} f)(-1.45, 1.84) = (-0.087, 3.68)$$
$$(x_{next}, y_{next}) = (-1.45, 1.84) - \alpha(\nabla_{x,y} f)(-1.45, 1.84)$$
$$= (-1.44, 1.47)$$
$$f(-1.44, 1.47) = -0.67$$

Q2: Run Through Gradient Descent

$$
\begin{aligned}
\alpha &= 0.1 \\
f(x,y) &= x^4 + x^3 - 2x^2 + y^2 \\
\frac{\partial f(x,y)}{\partial x} &= 4x^3 + 3x^2 - 4x \\
\frac{\partial f(x,y)}{\partial y} &= 2y
\end{aligned}
$$

x	y	$f(x,y)$
-0.5	4.5	19.688
-0.725	3.6	11.80
-1.02	2.88	6.23
-1.31	2.30	2.56
-1.45	1.84	0.55
-1.44	1.47	-0.67
-1.44	1.17	-1.45

$$
\begin{aligned}
f(-1.44, 1.47) &= -0.67 \\
(\nabla_{x,y} f)(-1.44, 1.47) &= (0.037, 2.94) \\
(x_{next}, y_{next}) &= (-1.44, 1.47) - \alpha(\nabla_{x,y} f)(-1.44, 1.47) \\
&= (-1.44, 1.17) \\
f(-1.44, 1.17) &= -1.45
\end{aligned}
$$