Gradient Descent: This is a method used to find the minima of $f : \mathbb{R}^n \to \mathbb{R}$, for $n \geq 1$. We will discuss the method for $n = 1, 2$ but it will be easy to see how the method generalises for arbitrary n. For $z \in R^n$,

$$\text{if } n = 1, \text{ let } \nabla f(z) = f'(z) \text{ and if } n = 2 \text{ let } \nabla f(z) := \left(\frac{\partial f}{\partial x}(z), \frac{\partial f}{\partial y}(z)\right).$$

The algorithm is as follows for $n = 1$ or $n = 2$:

Step 1: Choose an initial point $z^{(0)}$ in $\mathbb{R}^n$.

Step 2: Define $z^{(k)} = z^{(k-1)} - t_k \nabla f(z^{(k-1)})$ for some suitable t_k and for $k = 1, 2, \ldots T$.

Step 3: Choose T “appropriately” or take best possible answer from the above.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$ for $x \in \mathbb{R}$.
   (a) Find $f'(x)$
   (b) Draw a sketch of the curve.
   (c) Let $z^{(0)} = 3$ and $t_k = \frac{1}{4}$. Compute $z^{(k)}$ from gradient algorithm for $k = 1, 2, 3, 4$.
   (d) Mark $(z^{(k)}, f(z^{(k)}))$ on your graph of f.
   (e) Take $z^{(0)} = -4$ and repeat the above.
   (f) In either case, Can you explain what the Gradient Descent algorithm is doing ? Further decide if $z^{(k)}$ will converge to 0 as $k \to \infty$.
2. Let $\gamma > 1$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = \frac{1}{2}(x^2 + \gamma y^2)$.
   (a) Can you guess a minima for f ? item Draw the level curves of f at levels $4, 1, \frac{1}{100}, \frac{1}{10000}$.
   (b) Let $z^{(0)} = (\gamma, 1)$ and $t_k = \frac{2}{\gamma+1}$ for $k \geq 1$. Calculate $z^{(k)}$ using Gradient Descent algorithm for $k = 1, 2, 3, 4$.
   (c) Mark $(z^{(k)}, f(z^{(k)}))$ on your graph that you drew for the level curves.
   (d) Does $z^{(k)}$ converge and if so where ?
3. Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^4 - 4x^2$.
   (a) Using curve sketching techniques discussed earlier draw a rough sketch of the Curve and identify its global minima.
   (b) Write out the steps of the Gradient descent algorithm for an arbitrary $z^{(0)}, t_k$.
   (c) Decide (pictorially), for $z^{(0)} \in \{\frac{1}{2}, -\frac{1}{2}\}$ and $t_k = \frac{1}{4}$ where the gradient algorithm will converge to with the help of sketch that you have drawn.

Some Questions :

(i) How does one choose T and t_k, $1 \leq k \leq T$?

(ii) Suppose $T = \infty$, will the sequence $\{z^{(k)}\}_{k \geq 1}$ converge to a minima $z^\star$ of f ?

(iii) Will we be able to locate a global minima if there was one ?