Convex function and Gradients

Dr. Pritam Anand.

Assistant Professor,

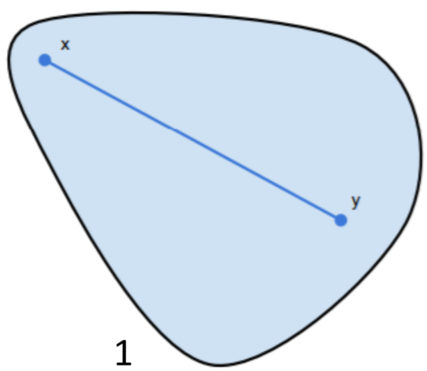
DA-IICT, Gandhinagar.

Convex Sets

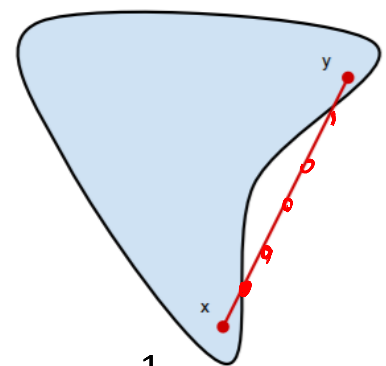
- $C \subseteq R^n$ is convex $f: C \rightarrow R$,

  if $\lambda x + (1 - \lambda)y \in C$ for any $x, y \in C$ and $0 \leq \lambda \leq 1$.

  that is, a set is convex if the line connecting any two points in the set is entirely inside the set



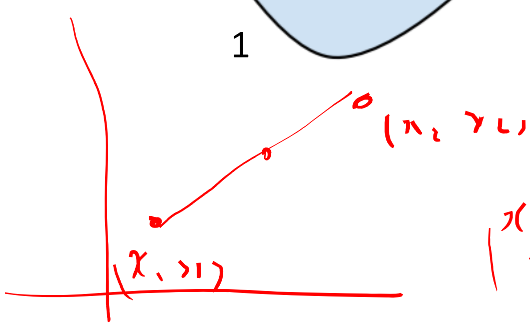
1



1

$$\frac{x+y}{2}$$

$$\lambda x + (1-\lambda) y$$



$$\left(\frac{x_1 + x_2}{2} , \frac{y_1 + y_2}{2} \right)$$

$$\frac{m x_1 + (1-m) x_2}{}$$

1. Nick Henderson et al., Convex set functions and Problems , 2019 .

Convexity and Gradient Descent methods

- $f : C \rightarrow R$ is convex

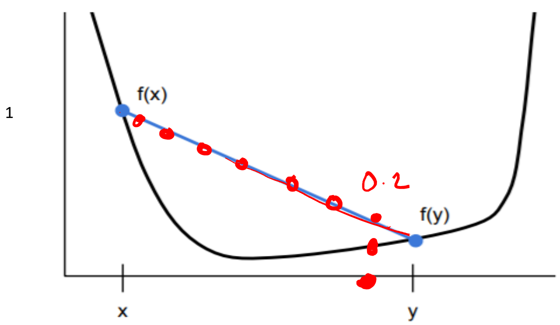
  if dom (f) (the domain of f) is a convex set,

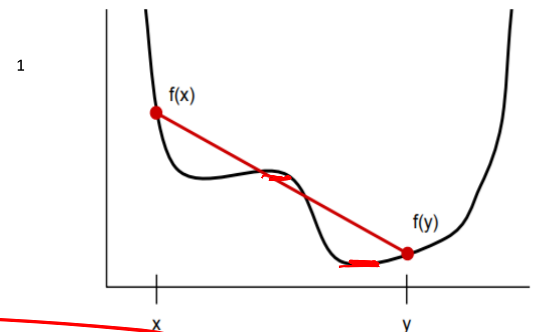
  and if $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for any $x, y \in \text{dom}(f)$ and $0 \leq \lambda \leq 1$.

  that is, the line connecting any two points on the graph of the function stays above the graph.

$f(x_0) \leq f(x)$

$\forall x \in \text{dom}(f)$





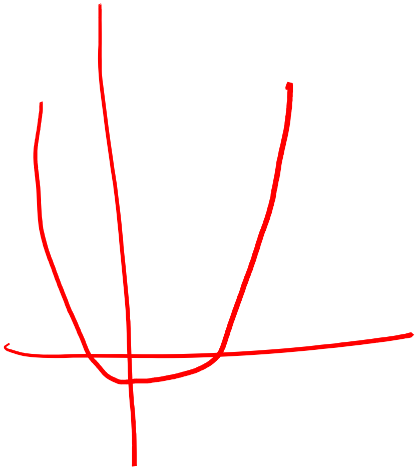
$$\boxed{\lambda f(x) + (1-\lambda) f(y) \geq f((\lambda x + (1-\lambda) y))}$$

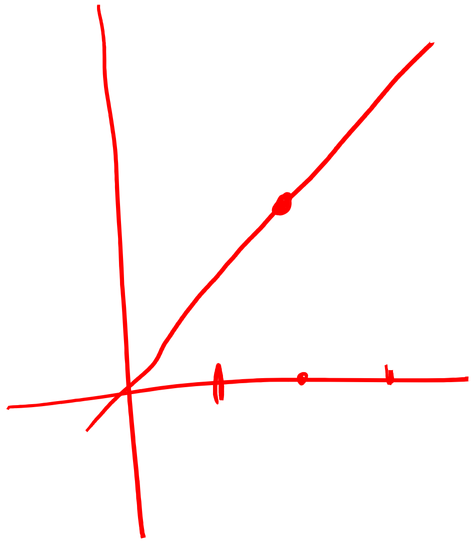
$f(x_0) \leq f(x)$

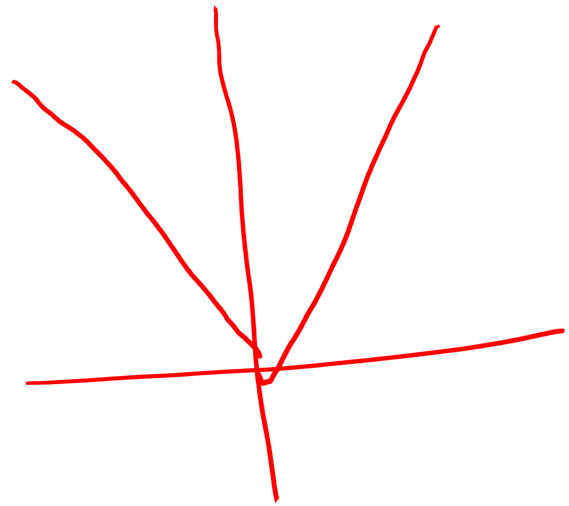
$\forall x \in N(x_0)$

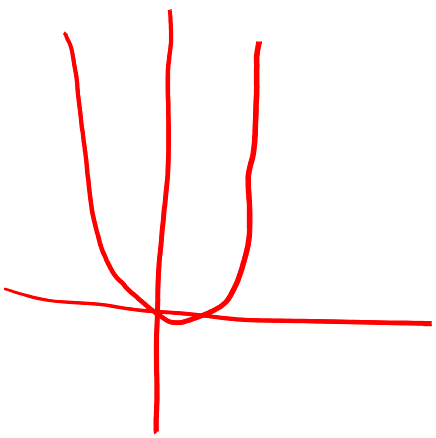
1. Nick Henderson et al., Convex set functions and Problems , 2019 .

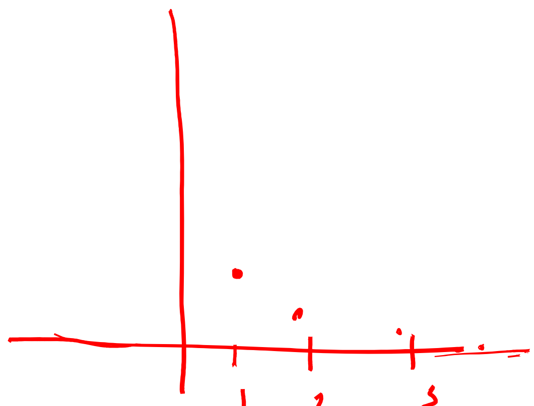
$f(x)$











x^2

x^4

x^6

x^8

$$f(\lambda x_0 + (1-\lambda) y) \geq \lambda f(x) + (1-\lambda) f(y)$$

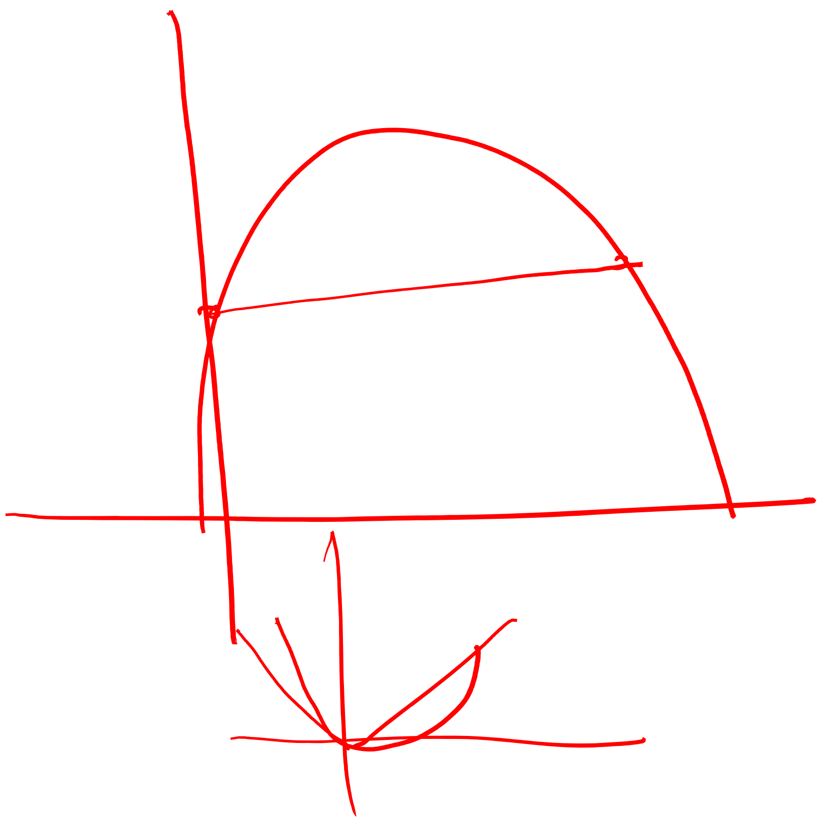
Concave functions

f, g convex func

$f+g$ would be convex func

$\alpha f(x)$ would be convex for $\alpha > 0$

$\max(f(x), g(x))$ would be also convex



Convexity and Gradient Descent methods

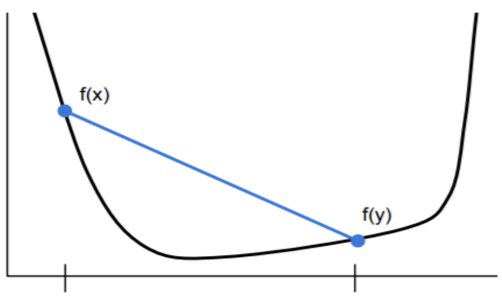
- f : C → R is convex

  if dom (f) (the domain of f) is a convex set,

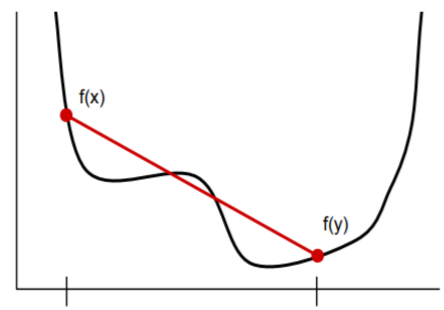
  and if f(tx + (1 − t)y) ≤ tf(x) + (1 − t)f(y) for any x, y ∈ dom(f) and 0 ≤ t ≤ 1 .

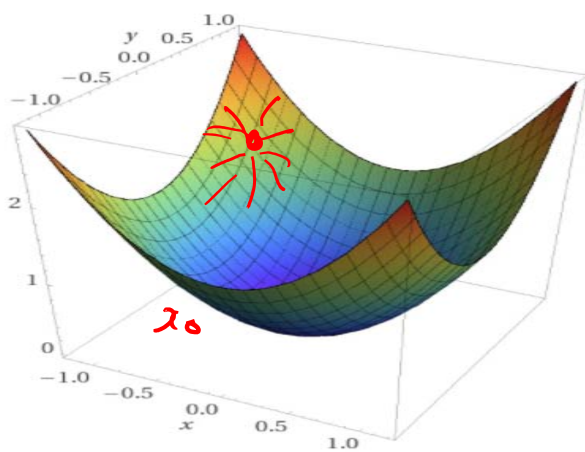
  that is, the line connecting any two points on the graph of the function stays above the graph.

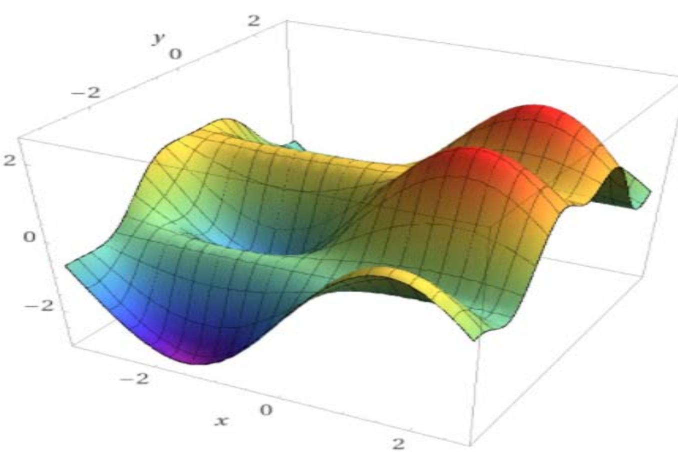
[1]



[1]

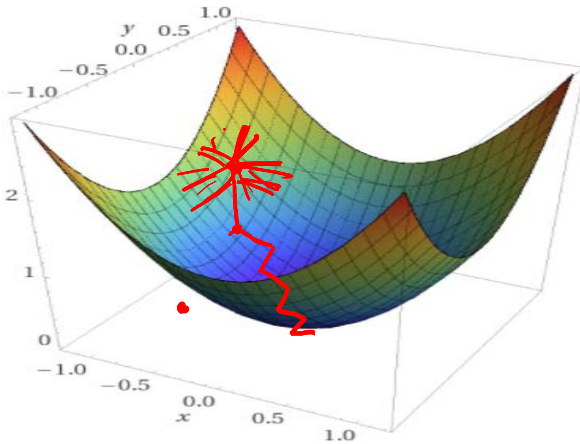


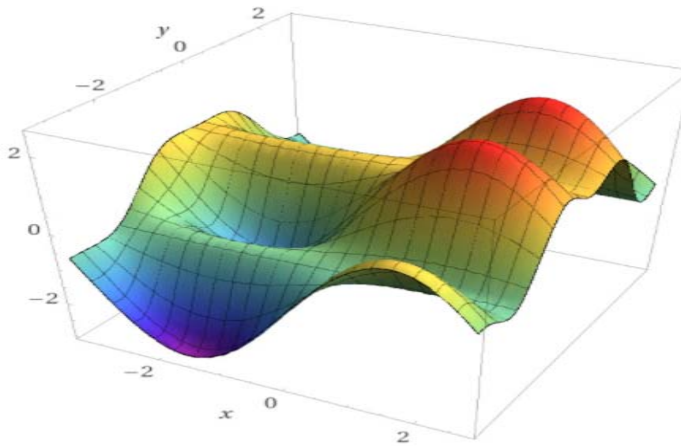


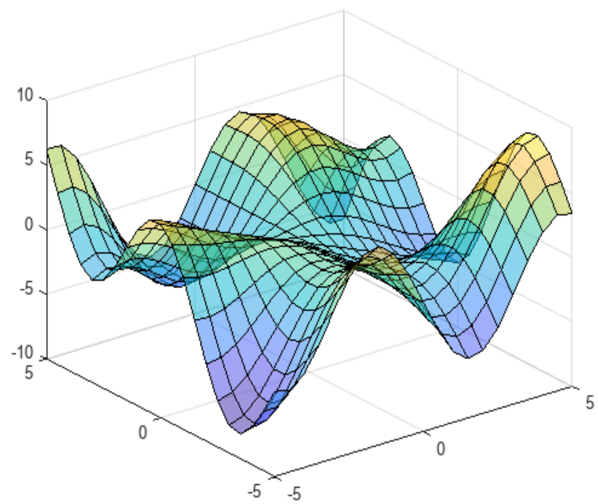


1. Nick Henderson et al., Convex set functions and Problems , 2019 .

More examples of convex functions







1. Nick Henderson et al., Convex set functions and Problems , 2019 .

Convexity and Gradient Descent methods

- f : C → R is concave

  if dom (f) (the domain of f) is a convex set,

  and if $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for any $x, y \in dom(f)$ and $0 \leq \lambda \leq 1$.

  that is, the line connecting any two points on the graph of the function stays above the graph.

[1]

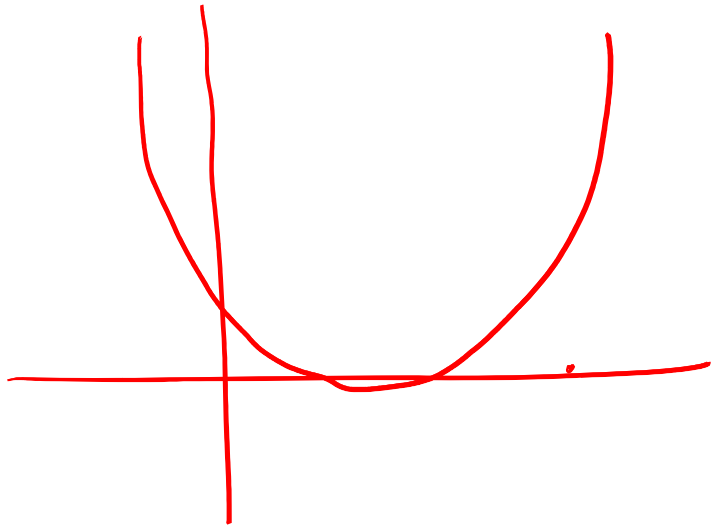
[1]

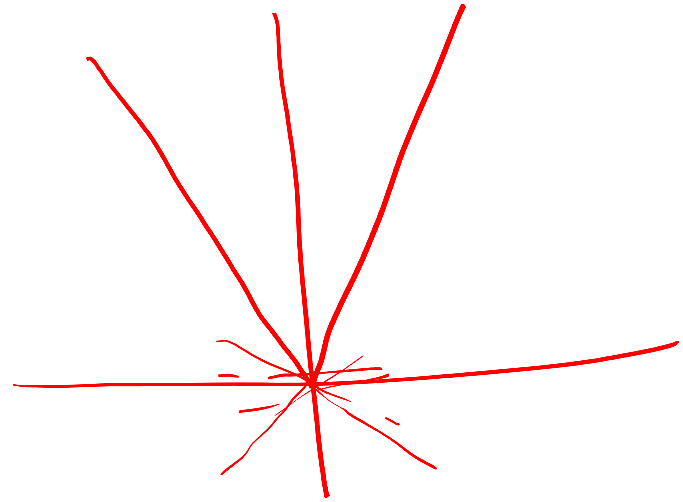
1. Nick Henderson et al., Convex set functions and Problems , 2019 .

Properties of convex functions

- if f , g C → R are convex , then *f+g* is also a convex function.
- if f: C → R is convex and α ≥ 0 , then αf is also a convex function.
- Every linear function is convex function.
- if f , g C → R are convex , then *max (f,g)* are also a convex function.

1. Nick Henderson et al., Convex set functions and Problems , 2019 .





Gradient and derivatives

- For a function $f(x) = f(x_1, x_2, \ldots, x_n)$, and a unit vector $u = (u_1, u_2, \ldots, u_n)$, then the directional derivative is defined as $\nabla_u f(x) = \lim_{h \to 0} \left(\frac{f(x+hu) - f(x)}{h} \right)$. [1]

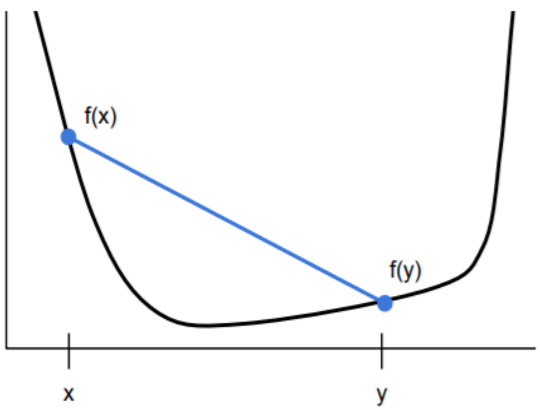
- $f(x)$ is differentiable implies that $\nabla_u f(x)$ is well-defined for all x and u.

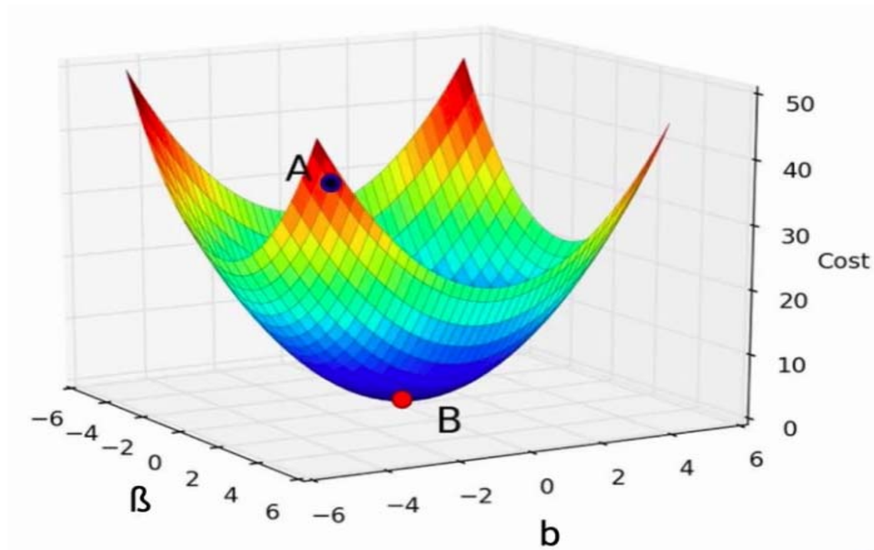
- $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ . \\ . \\ \frac{\partial f}{\partial xn} \end{bmatrix}$ and the directional derivative of f at point x, for any direction (unit vector) u can be obtained as $\nabla_u f(x) =. u^T \nabla f$.

https://www.cs.utah.edu/~jeffp/IDABook/T6-GD.pdf

Gradient and convex function

- *Consider the optimization problem* $\min_{x} f(x)$, where f : $R^n \rightarrow R$ is convex and smooth, then the necessary and sufficient condition for optimal solution x_0 is $\nabla f(x) = 0 \quad at\ x = x_0.$





$f(x,y) = x^2 + y^2$

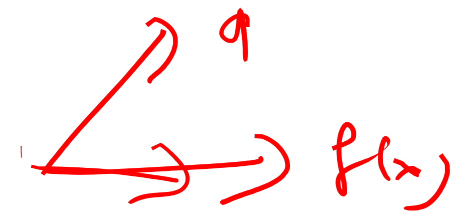
$\begin{bmatrix} 2x \\ 2y \end{bmatrix} = 0$

$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

https://www.cs.utah.edu/~jeffp/IDABook/T6-GD.pdf

Find the direction d such that rate of change in the function is maximum

$$\boxed{d^T \nabla f(x)}$$



$$d = \frac{\nabla f(x)}{\|\nabla f(x)\|}$$ direction of ascent

$$d = \frac{-\nabla f(x)}{\|\nabla f(x)\|}$$ direction of descent

Gradient descent algorithm

An iterative algorithm

The negative Gradient direction is the direction of steepest descent.

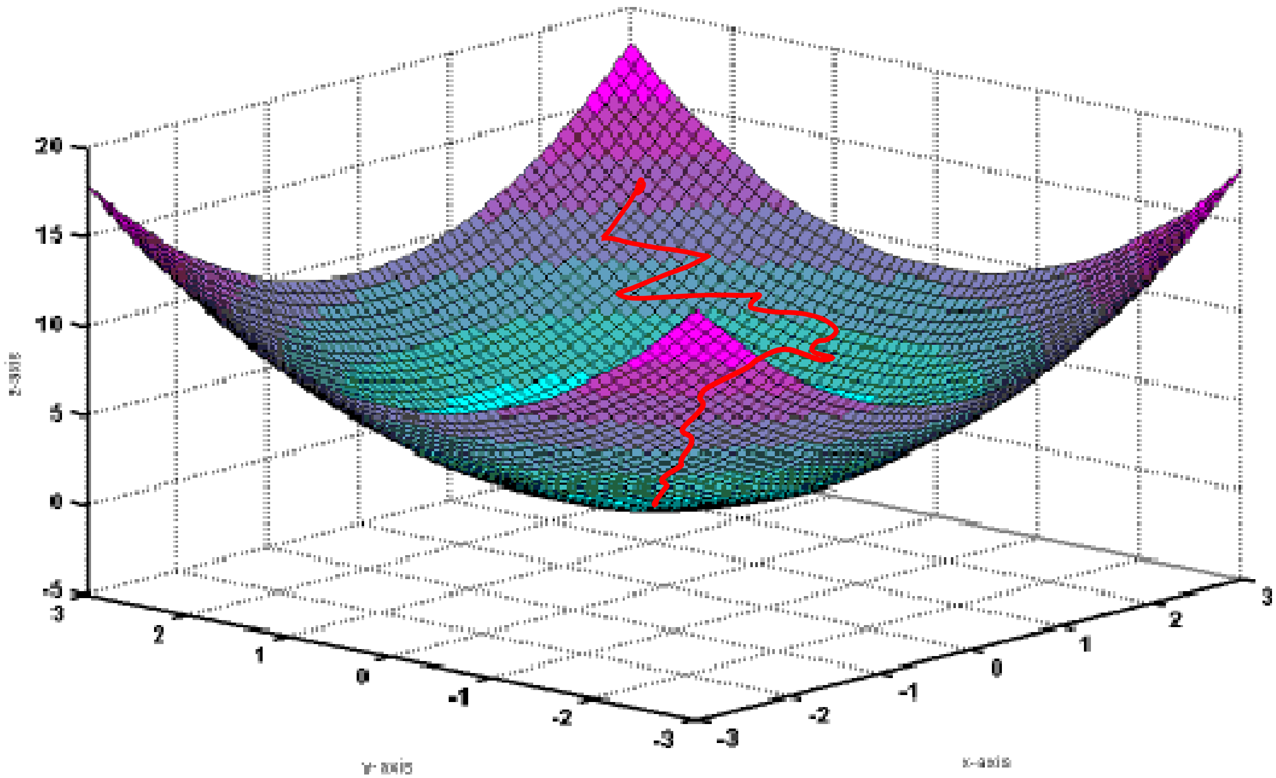
Algorithm:- Gradient descent method

Initialize $x^0 = x^{start} \in R^n$

repeat $x^{(k+1)} := x^{(k)} - \gamma_k \nabla f(x^{(k)})$.

$\gamma_{n_k} \| \nabla f(x^k) \|$

Until $|| \nabla f(x^{(k)}) || \leq \varepsilon$


20
15
10
5
0
-5
z-axis
3
2
1
0
-1
-2
-3
y-axis
-3
-2
-1
0
1
2
3
x-axis