Descent Method :- Steepest Descent method, Newton's method, DFP (Quasi Newton's) method.

Newton method :-

Taylor's Series :- Taylor series of a real valued function $f(x)$, which is infinitely differentiable function at a real number 'a' is a power series

$$\boxed{f(x) \cong f(a) + \frac{1}{1!} f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2} + \frac{1}{3!} f'''(a)(x-a)^3 + \cdots$$

where $n! = 1 \times 2 \times \cdots \times n$

$f'(a)$ = first derivative of function at a

$f''(a)$ = second derivative of function at a

and so on.

$$f'(x) = f'(a) + \frac{1}{2} f''(a) \cdot 2(x-a)$$

To find maxima or minima $f'(x) = 0$

$$0 = f'(a) + f''(a)(x-a)$$

$$x - a = \frac{-f'(a)}{f''(a)} \Rightarrow \boxed{x = a - \frac{f'(a)}{f''(a)}}$$

Algorithm

- Objective function $f(x)$
- initial point a_1
- stopping condition $\rightarrow |f'(x)| < \varepsilon$

- compute $f'(x)$ and $f''(x)$
- compute $f'(a_1)$ and $f''(a_1)$

- compute $f'(x)$ ...
- compute $f'(a_1)$ and $f''(a_1)$
- $$a_2 = a_1 - \left[\frac{f'(a_1)}{f''(a_1)}\right]$$

$|f'(a_2)| < \varepsilon$: a_2 is local minima

$|f'(a_2)| \geq \varepsilon$: a_2 is your new initial point.

Example:- $\min(x^2+2x)$

$a = 0$ (100) starting point

$\varepsilon = 0.1$

$|f'(x)| < \varepsilon$ → stopping condition.

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

(1) $f'(x) = 2x+2$ $\qquad f''(x) = 2$

(2) $f'(0) = 2$, $f'(100) = 202$ $\qquad f''(0) = 2$, $f''(100) = 2$

(3) $$a_2 = a_1 - \frac{f'(a_1)}{f''(a_1)} = 0 - \frac{2}{2} = -1$$

$\boxed{a_2 = -1}$ $\qquad \boxed{a_2 = 100 - \frac{202}{2} = -1}$

(4) $f'(-1) = (2\times -1)+2 = 0$ $\qquad |f'(-1)| = 0 < 0.1$

$\boxed{a_2 = -1}$ is a local minima.

Newton Method :- Multivariable function

single variable $\qquad$ multivariable function

$x = (x_1, x_2, \ldots, x_n)$

Newton Method :- Multivariable function

single variable

1. $f'(x)$ and $f''(x)$
2. $f'(a)$ and $f''(a)$ at initial point a
3. $a_2 = a - \dfrac{f'(a)}{f''(a)}$

multivariable function

$f(X)$, $X = (x_1, x_2, \ldots, x_n)$

1. $\nabla f(X)$ and $H_f(X)$
2. $\nabla f(a_1, a_2, \ldots, a_n)$ and $H_f(X)\Big|_{(a_1, a_2, \ldots, a_n)}$
3. $(a_1', a_2', \ldots, a_n') = (a_1, a_2, \ldots, a_n) - \dfrac{\nabla f(a_1, a_2, \ldots, a_n)}{H_f(a_1, \ldots, a_n)}$

stopping condition

$$(a_1', a_2', \ldots, a_n') = (a_1, a_2, \ldots, a_n) - [H_f(a_1, a_2, \ldots, a_n)]^{-1} \nabla f(a_1, \ldots, a_n)$$

$n \times 1$ column wise. | $n \times n$ matrix $n>>>$ | $n \times 1$ column wise

- Identity matrix $\rightarrow \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ $n \times n$ matrix.
- $A \rightarrow n \times n$ matrix. B is called inverse of A if $BA = I = AB$
- A is 2x2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Example :- Minimize

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

Example :- minimize

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

starting point $x_1 = (0,0)$

stopping condition $|\nabla f(x_k)| = 0.$

① $f(x,y) = x - y + 2x^2 + 2xy + y^2$

$$\nabla f = \begin{pmatrix} \partial f/\partial x \\ \partial f/\partial y \end{pmatrix} = \begin{pmatrix} 1 + 4x + 2y \\ -1 + 2x + 2y \end{pmatrix}$$

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\nabla f\big|_{(0,0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} , \quad H\big|_{(0,0)} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

$$x_2 = (0,0) - \left(H\big|_{(0,0)}\right)^{-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$x_2 = -\frac{1}{4} \underbrace{\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{2\times 1} = -\frac{1}{4} \begin{bmatrix} 4 \\ -6 \end{bmatrix} = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$$

check stopping condition

$$\nabla f\big|_{(-1,\,3/2)} = \begin{pmatrix} \overset{1-4+3}{1 + 4x + 2y} \\ \underset{-1-2+3}{-1 + 2x + 2y} \end{pmatrix}\Bigg|_{(-1,\,3/2)}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- $(-1, \frac{3}{2})$ is local minima.

Example $\quad \text{Min } f(x) = 100\,(y - x^2)^2 + (1-x)^2$

starting point $(0,0)$

$$\left|\nabla f\right|\Big|_{x_k} = (0,0)$$

- $$\nabla f = \begin{pmatrix} \partial f/\partial x \\ \partial f/\partial y \end{pmatrix} = \begin{pmatrix} 100 \cdot 2\,(y-x^2)(-2x) + 2\,(1-x)(-1) \\ 100 \cdot 2\,(y-x^2) \end{pmatrix}$$

- $$H_f = \begin{pmatrix} \overset{+400\ +800+2\ =\ 1202}{-400\,(y-x^2) - 400(x)(-2x) + 2} & \overset{-400}{-400x} \\ \underset{-400}{-400x} & \underset{200}{200} \end{pmatrix}$$

- $$\nabla f\Big|_{(0,0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \qquad H_f\Big|_{(0,0)} = \begin{pmatrix} 2 & 0 \\ 0 & 200 \end{pmatrix}$$

$$\left(H_f\Big|_{(0,0)}\right)^{-1} = \frac{1}{400}\begin{pmatrix} 200 & 0 \\ 0 & 2 \end{pmatrix}$$

$$X_2 = \overset{(0,0)}{\overset{\shortparallel}{X_1}} - \begin{pmatrix} 1/2 & 0 \\ 0 & \frac{1}{200} \end{pmatrix}\begin{pmatrix} -2 \\ 0 \end{pmatrix} = -\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\boxed{X_2 = (1,0)}$$

$$\nabla f\Big|_{(1,0)} = (400, -200) \neq (0,0)$$

$X_2 = (1,0)$ is new initial point.

Second Iteration:-

$$\nabla f(1,0) = (400, -200)$$

$$H_f(1,0) = \begin{pmatrix} 1202 & -400 \\ -400 & 200 \end{pmatrix}$$

$$\left(H_f(1,0)\right)^{-1} = \frac{1}{80400}\begin{pmatrix} 200 & 400 \\ 400 & 1202 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{80400}\begin{pmatrix} 200 & 400 \\ 400 & 1202 \end{pmatrix}\begin{pmatrix} 400 \\ -200 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{80400}\begin{pmatrix} 8000 & -80000 \\ 160000 & -240400 \end{pmatrix}$$

$$X_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{80400}\begin{pmatrix} 0 \\ -80400 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\nabla f\big|_{(1,1)} = (0,0).$$

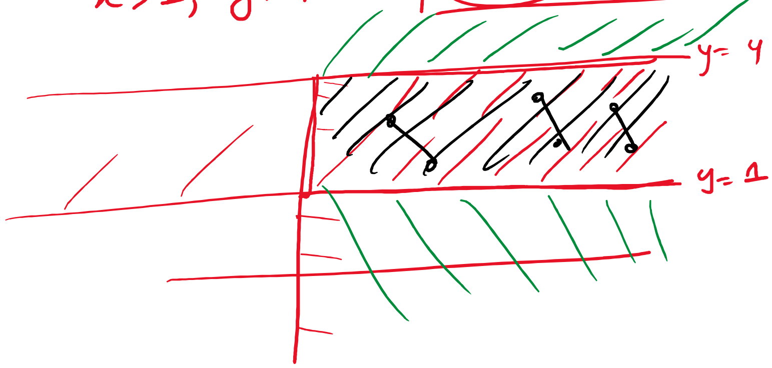
(1,1) is local minima.

$x \geq 0$, $y \leq 4$, $y \geq 1$

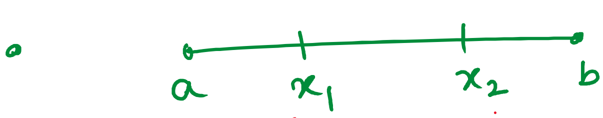
$x > 1$, $y > 4$

$x \geq 0$, $y \geq 4$, $y \leq 1$

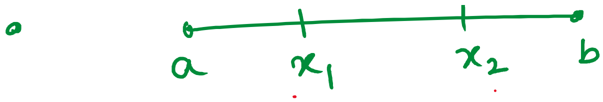
$\boxed{\phi}$



$\varepsilon \longrightarrow$ an iteration $x_1 = x_2$ will happen only at $(n-1)^{th}$ iteration.



will happen at $(n-1)^{th}$ iteration.

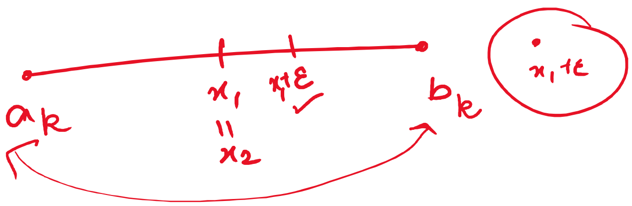


$\boxed{\text{Not } \varepsilon} \rightarrow \varphi(x_1) = \varphi(x_2)$

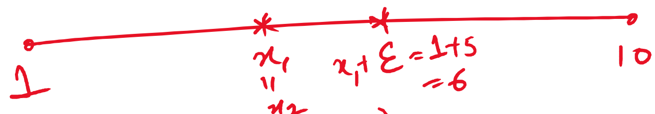
$\varphi(x_1) \geq \varphi(x_2)$

$\varphi(x_1) \leq \varphi(x_2)$

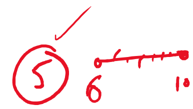
How to choose ε



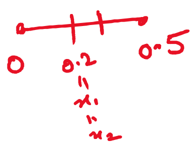
for Example



9



0.3 0.1 0.05



1.2

0.1

10%