Method of Steepest Descent for multivariate function minimization:

Minimize $f(X)$ with an initial guess X_0

where $f: R^n \to R$

[$R^n = \{(x_1, x_2, .., x_n) : x_1, x_2, .. x_n \in R\}$ is known as n-dimensional space over R.]

An important property of the gradient vector $\nabla f(X)$ is that it points in the direction of the most rapid increase in the function f, which is the direction of steepest ascent.

Note: If problem is of maximization, then convert it to minimization as $\text{Min} f(X) = -\text{Max} f(X)$ and proceed on the same way.

Conversely, $-\nabla f(X)$ points in the direction of the steepest descent.

On the basis of this steepest descent method for minimizing the multivariate function $f: R^n \to R$ can be described as:

For $k = 0, 1, 2, \ldots$

Step 1: Start with initial guess X_0 i.e, for $k=0$.

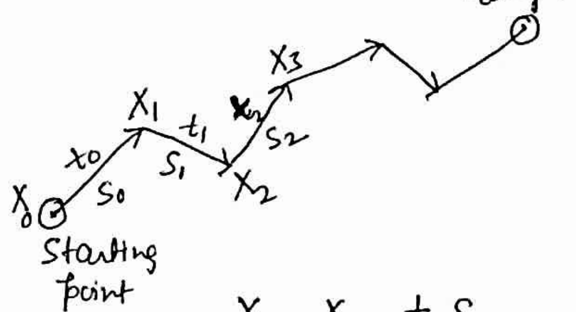
Step 2: Find the search direction S_k as

$$S_k = -\nabla f_k = -\nabla f(X_k)$$

Step 3: Determine the step length t_k by minimizing the function

$$\phi(t) = f(X_k + t S_k)$$

[Put $\phi'(t) = 0$ and find $t = t_k > 0$ such that $\phi''(t_k) > 0$]



$$X_1 - X_0 = t_0 S_0$$

Step 4: Find $X_{k+1} = X_k + t_k S_k$ and check the optimality for the new point X_{k+1} as

If $\nabla f(X_{k+1}) \cong 0$ or $t_k = 0$, then stop otherwise go to step (1) again for X_{k+1}.

Que By the method of steepest descent minimize the function $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 x_2 + x_2^2$ starting from the point $X_0 = (0, 0)$.

Sol Given $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

$\therefore \nabla f(x_1, x_2) = (1 + 4x_1 + 2x_2, -1 + 2x_1 + 2x_2)$ ——— (1)

$\left(\because \nabla f(x_1, x_2) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) \right)$

Iteration 1: $k = 0$, $X_0 = (0, 0)$

$\therefore S_0 = -\nabla f(X_0) = -\nabla f(0,0) = -(1, -1) = (-1, 1)$

We now minimize the function

$\phi(t) = f(X_0 + tS_0) = f(-t, t)$

{ To find $\phi'(t)$ we can find $\phi(t)$ in terms of t by def. diff and then $\phi'(t)$ but it is complicated

$\therefore \phi'(t) = \nabla f(-t, t) \cdot S_0 = (1 - 4t + 2t, -1 - 2t + 2t) \cdot (-1, 1)$

$= (1 - 2t, -1) \cdot (-1, 1)$

$= -1 + 2t - 1 = 2(t - 1)$

$\therefore \phi'(t) = 0 \Rightarrow t = 1$

and $\phi''(t) = 2 > 0$ for $t = 1$

$\therefore \phi(t)$ is minimum at $t = 1 = t_0$

$\therefore$ the new point is

$X_1 = X_0 + t_0 S_0 = (0, 0) + 1(-1, 1) = (-1, 1)$

$\therefore \nabla f(X_1) = \nabla f(-1, 1) = (1 - 4 + 2, -1 - 2 + 2) = (-1, -1)$ by (1) $\neq (0, 0)$

$\therefore$ We go to next iteration with $X_1 = (-1, 1)$ (i.e., for $k = 1$)

Iteration 2; $k = 1$, $X_1 = (-1, 1)$

$\therefore S_1 = -\nabla f(X_1) = (1, 1)$

We now minimize the function

$\phi(t) = f(X_1 + tS_1) = f(-1 + t, 1 + t)$

$\therefore \phi'(t) = \nabla f(-1+t, 1+t) \cdot S_1$

$= (1-4+4t+2+2t, -1+2t-2+2+2t) \cdot (1,1)$ by (1)

$= (-1+6t, -1+4t) \cdot (1,1)$

$= -1+6t-1+4t$

$= 2(5t-1)$

$\therefore \phi'(t) = 0 \Rightarrow t = \frac{1}{5}$ and $\phi''(t) = 10 > 0$

$\therefore \phi(t)$ is minimum at $t = \frac{1}{5} = t_1$

$\therefore$ the new point is

$X_2 = X_1 + t_1 S_1 = (-1,1) + \frac{1}{5}(1,1) = \left(-\frac{4}{5}, \frac{6}{5}\right)$

$\therefore \nabla f(X_2) = \nabla f\left(-\frac{4}{5}, \frac{6}{5}\right)$

$= \frac{1}{5}(5-16+12, -5-8+12) = \left(\frac{1}{5}, -\frac{1}{5}\right) \neq (0,0)$

$\therefore X_2$ is not optimum. So, we go to next iteration with X_2 (i.e, for $k=2$)

Iteration 3: $k=2$, $X_2 = \left(-\frac{4}{5}, \frac{6}{5}\right)$

$\therefore S_2 = -\nabla f(X_2) = \left(-\frac{1}{5}, \frac{1}{5}\right)$

We now minimize the function

$\phi(t) = f(X_2 + tS_2) = f\left(\frac{-4-t}{5}, \frac{6+t}{5}\right)$

$\therefore \phi'(t) = \nabla f\left(\frac{-4-t}{5}, \frac{6+t}{5}\right) \cdot S_2$

$= \left(\frac{5-16-4t+12+2t}{5}, \frac{-5-8-2t+12+2t}{5}\right) \cdot \left(-\frac{1}{5}, \frac{1}{5}\right)$

$= \frac{1}{25}(1-2t, -1) \cdot (-1,1) = \frac{1}{25}(-1+2t-1) = \frac{2}{25}(t-1)$

$\therefore \phi(t)$ is minimum at $t = 1 = t_2$ $\quad \left(\because \phi''(t) = \frac{2}{25} > 0\right)$

$\therefore$ the new point is $X_3 = X_2 + t_2 S_2$

$= \left(-\frac{4}{5}, \frac{6}{5}\right) + \left(-\frac{1}{5}, \frac{1}{5}\right) = \left(-1, \frac{7}{5}\right)$

$\therefore \nabla f(X_3) = \nabla f(-1, \frac{7}{5})$

$= \left(1-4+\frac{14}{5}, -1-2+\frac{14}{5}\right)$ by (1)

$= \left(-\frac{1}{5}, -\frac{1}{5}\right) \neq (0,0)$

$\therefore$ We go to next iteration with $X_3 = (-1, \frac{7}{5})$ (i.e, for $k=3$)

<u>Iteration 4</u>: $k=3$, $X_3 = (-1, \frac{7}{5}) = (-\frac{5}{5}, \frac{7}{5})$

$\therefore S_3 = -\nabla f(X_3) = \left(\frac{1}{5}, \frac{1}{5}\right)$

We now minimize the function

$\phi(t) = f(X_3 + tS_3)$

$= f\left(\frac{-5+t}{5}, \frac{7+t}{5}\right)$

$\nabla f\left(\frac{-5+t}{5}, \frac{7+t}{5}\right) = \frac{1}{5}\left(5-20+4t+14+2t, -5-10+2t+14+2t\right)$

$= \frac{1}{5}\left(-1+6t, -1+4t\right)$

$\therefore \phi'(t) = \nabla f\left(\frac{-5+t}{5}, \frac{7+t}{5}\right) \cdot S_3$

$= \frac{1}{5}\left(-1+6t, -1+4t\right) \cdot \left(\frac{1}{5}, \frac{1}{5}\right)$

$= \frac{1}{25}\left(-1+6t-1+4t\right) = \frac{2}{25}(5t-1)$

$\therefore \phi'(t) = 0 \Rightarrow t = \frac{1}{5}$ and $\phi''(t) = \frac{2}{5} > 0$

$\therefore \phi(t)$ is minimum at $t = \frac{1}{5} = t_3$

$\therefore$ the new point is $X_4 = X_3 + t_3 S_3$

$= \left(-\frac{5}{5}, \frac{7}{5}\right) + \frac{1}{5}\left(\frac{1}{5}, \frac{1}{5}\right)$

$= \left(-\frac{24}{25}, \frac{36}{25}\right)$

$\therefore \nabla f(X_4) = \nabla f\left(-\frac{24}{25}, \frac{36}{25}\right) = \left(\frac{25-96+72}{25}, \frac{-25-48+72}{25}\right)$

$= \left(\frac{1}{25}, -\frac{1}{25}\right) = (0.04, -0.04)$

$\therefore\ \nabla f(X_4) = (0.04, -0.04) \cong (0,0)$ (upto one decimal place)

$\therefore$ At this step optimum value of X is

$$X^* = \left(\frac{-24}{25}, \frac{36}{25}\right) = (-0.96, 1.44)\ \text{Ans}$$

Note: If we want to find more exact answer then we can proceed further for next iteration as:

Iteration 5: $k=4,\ X_4 = \left(\frac{-24}{25}, \frac{36}{25}\right)$

$\therefore\ S_4 = -\nabla f(X_4) = \left(\frac{-1}{25}, \frac{1}{25}\right)$

We now minimize the function

$$\phi(t) = f(X_4 + tS_4) = f\left(\frac{-24-t}{25}, \frac{36+t}{25}\right)$$

$$\therefore\ \phi'(t) = \nabla f\left(\frac{-24-t}{25}, \frac{36+t}{25}\right)\cdot S_4$$

$$= \left(\frac{25-96-4t+72+2t}{25}, \frac{-25-48-2t+72+2t}{25}\right)\cdot\left(\frac{-1}{25}, \frac{1}{25}\right)$$

$$= \left(\frac{1-2t}{25}, \frac{-1}{25}\right)\cdot\left(\frac{-1}{25}, \frac{1}{25}\right)$$

$$= \frac{-1+2t-1}{625} = \frac{2(t-1)}{625}$$

$\therefore\ \phi'(t) = 0 \Rightarrow t = 1$ and $\phi''(t) = \frac{2}{625} > 0$ for $t=1$

$\therefore\ \phi(t)$ is minimum at $t = 1 = t_4$

$\therefore$ the new point is $X_5 = X_4 + t_4 S_4$

$$= \left(\frac{-24}{25}, \frac{36}{25}\right) + \left(\frac{-1}{25}, \frac{1}{25}\right) = \left(-1, \frac{37}{25}\right)$$

$$\nabla f(X_5) = \nabla f\left(-1, \frac{37}{25}\right) = \left(1-4+\frac{74}{25}, -1-2+\frac{74}{25}\right)$$

$$= \left(-\frac{1}{25}, -\frac{1}{25}\right) = (-0.04, -0.04)$$

$$\neq (0,0)$$

$\therefore$ We go to next iteration with X_5 i.e, for $k=5$

Iteration 6: $k=5$, $X_5 = \left(-1, \frac{37}{25}\right)$

$\therefore S_5 = -\nabla f(X_5) = \left(\frac{1}{25}, \frac{1}{25}\right)$

We now minimize the function

$\phi(t) = f(X_5 + tS_5) = f\left(\frac{-25+t}{25}, \frac{37+t}{25}\right)$

$\therefore \phi'(t) = \nabla f\left(\frac{-25+t}{25}, \frac{37+t}{25}\right) \cdot S_5$

$= \left(\frac{25-100+4t+74+2t}{25}, \frac{-25-50+2t+74+2t}{25}\right) \cdot \left(\frac{1}{25}, \frac{1}{25}\right)$ by (1)

$= \frac{1}{625}\left(-1+6t, -1+4t\right) \cdot (1,1)$

$= \frac{1}{625}\left(-1+6t-1+4t\right) = \frac{2}{625}(5t-1)$

$\therefore \phi'(t) = 0 \Rightarrow t = \frac{1}{5}$ and $\phi''(t) > 0$

$\therefore \phi(t)$ is minimum at $t = \frac{1}{5} = t_5$

$\therefore$ the new point is $X_6 = X_5 + t_5 S_5 = \left(\frac{-25}{25}, \frac{37}{25}\right) + \frac{1}{5}\left(\frac{1}{25}, \frac{1}{25}\right)$

$= \left(\frac{-124}{125}, \frac{186}{125}\right)$

$= (0.992, 1.46)$

and $\nabla f(X_6) = \nabla f\left(\frac{-124}{125}, \frac{186}{125}\right)$

$= \left(\frac{125-496+372}{125}, \frac{-125-248+372}{125}\right)$

$= \left(\frac{1}{125}, \frac{1}{125}\right)$

$= (0.008, 0.008) \cong (0,0)$

$\therefore$ We stop here and hence $X^* = X_6 \simeq (1.0, 1.5)$ Ans

(correct to one decimal places)

$f(x_1, x_2) = 1 - 1.5 + 2 + 3 + 2.25$

$= 6.75$