Optimization

Fourth Class

2020 - 2021

By

Dr. Jawad Mahmoud Jassim

Dept. of Math.

Education College for Pure Sciences

University of Basrah

Iraq

Chapter Three

جامعة البصرة * كلية التربية للعلوم
الدكتور
جواد
محمود
جاسم

The Steepest Descent and Newton Methods

Lecture 1

First: The Steepest Descent Method

The steepest descent method is one of the simplest and most fundamental minimization methods for unconstrained optimization.

Since it uses the negative gradient as its descent direction, it is called the gradient method.

Suppose that $f(X)$ is continuously differentiable function and the gradient vector of $f(X)$ at X_k is $g_k = \nabla f(X_k) \neq 0$.

From the Taylor expansion

$$f(X) = f(X_k) + (X - X_k)^T g_k + O(\|X - X_k\|) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

***We know that, if we write $X - X_k = \alpha d_k$*, then the direction d_k satisfy $d_k{}^T g_k < 0$ is called a *descent direction that is such that* $f(X) < f(X_k)$.**

***By the Cauchy – Schwartz inequality, we have*:**

$$\left|d_k^T g_k\right| \leq \|d_k\| \|g_k\| \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots . \quad (2)$$

The value of $d_k^T g_k$ is the smallest if and only if $d_k = -g_k$. Therefore $-g_k$ is the steepest descent direction. The iterative scheme of the steepest descent method is $X_{k+1} = X_k - \alpha_k g_k$.

In the following we give the algorithm.

Algorithm (1): (The Steepest Descent Method)

Step 0:

Given an initial point $X_0 \epsilon R^n$ and $0 < \varepsilon < 1$ be the termination tolerance. Set $k = 0$.

Step 1:

If $\|g_k\| < \varepsilon$, stop. Otherwise $let\ d_k = -g_k$, where is the gradient vector of the function $f(X)$ which be minimized and d_k is the descent direction.

Step 2:

Find the step length factor α_k $such\ that$

$$f(X_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(X_k + \alpha d_k).$$

Step 3:

Compute $X_{k+1} = X_k + \alpha_k d_k$.

Step 4:

Set $k = k + 1$ and go to step 1.

<u>Example:</u>

Compute the first two iterates of the method of steepest descent applied to the objective function $f(X) = x_1^2 + 2x_2^2$ *with initial point* $X_0 = [1, 1]^T$.

<u>Solution:</u>

First, we find the gradient vector $g(X)$ as:

$$g(X) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}]^T = [2x_1, 4x_2]^T.$$

$\therefore$The gradient vector at $X_0 = [1, 1]^T$is:

$$g_0 = g(X_0) = [2, 4]^T.$$

$$\therefore X_1 = X_0 - \alpha g_0 = [1 - 2\alpha, 1 - 4\alpha]^T.$$

$$f(X_1) = (1 - 2\alpha)^2 + 2(1 - 4\alpha)^2.$$

$$\frac{df}{d\alpha} = 2(1 - 2\alpha)(-2) + 4(1 - 4\alpha)(-4) = -4(1 - 2\alpha) - 16(1 - 4\alpha)$$

$$= -4 + 8\alpha - 16 + 64\alpha = 72\alpha - 20.$$

Set $\frac{df}{d\alpha} = 0 \rightarrow 72\alpha - 20 = 0 \rightarrow 72\alpha = 20 \rightarrow \alpha = \frac{20}{72} = \frac{5}{18}$.

$$\frac{d^2 f}{d\alpha^2} = 72 > 0 .$$

$\therefore \alpha = \frac{5}{18}$ represents a minimizer of $f(X_1)$.

$$\therefore X_1 = [1 - \frac{10}{18}, 1 - \frac{20}{18}]^T = [\frac{8}{18}, \frac{-2}{18}]^T = [\frac{4}{9}, -\frac{1}{9}]^T .$$

$$g_1 = g(X_1) = [\frac{8}{9}, -\frac{4}{9}]^T .$$

$$\therefore X_2 = X_1 - \alpha g_1 = \left[\frac{4}{9} - \frac{8}{9}\alpha , -\frac{1}{9} + \frac{4}{9}\alpha\right]^T .$$

$$f(X_2) = (\frac{4}{9} - \frac{8}{9}\alpha)^2 + 2(-\frac{1}{9} + \frac{4}{9}\alpha)^2 .$$

$$\frac{df}{d\alpha} = 2\left(\frac{4}{9} - \frac{8}{9}\alpha\right)\left(-\frac{8}{9}\right) + 4\left(-\frac{1}{9} + \frac{4}{9}\alpha\right)\left(\frac{4}{9}\right)$$
$$= -\frac{16}{9}\left(\frac{4}{9} - \frac{8}{9}\alpha\right) + \frac{16}{9}(-\frac{1}{9} + \frac{4}{9}\alpha)$$

Set $\frac{df}{d\alpha} = 0 \rightarrow -\frac{16}{9}\left(\frac{4}{9} - \frac{8}{9}\alpha\right) + \frac{16}{9}\left(-\frac{1}{9} + \frac{4}{9}\alpha\right) = 0 \rightarrow$

$-\left(\frac{4}{9} - \frac{8}{9}\alpha\right) + \left(-\frac{1}{9} + \frac{4}{9}\alpha\right) = 0 \rightarrow -\frac{5}{9} + \frac{12}{9}\alpha = 0 \rightarrow \frac{12}{9}\alpha = \frac{5}{9} \rightarrow$

$12\alpha = 5 \rightarrow \alpha = \frac{5}{12}$.

? $\frac{d^2 f}{d\alpha^2} = \frac{12}{9} > 0$.

$\therefore \alpha = \frac{5}{12}$ **represents a minimizer of** $f(X_2)$.

$\therefore X_2 = \left[\frac{4}{9} - \frac{8}{9}\alpha, -\frac{1}{9} + \frac{4}{9}\alpha\right]^T = \left[\frac{4}{9} - \frac{8}{9}\left(\frac{5}{12}\right), -\frac{1}{9} + \frac{4}{9}\left(\frac{5}{12}\right)\right]^T = \left[\frac{2}{27}, \frac{2}{27}\right]^T$

<u>H.W.</u>

Compute the first two iterates of the method of steepest descent applied to the objective function $f(X) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ ***with initial point*** $X_0 = [0, 0]^T$.