Optimization

Fourth Class

2020 - 2021

By

Dr. Jawad Mahmoud Jassim

Dept. of Math.

Education College for Pure Sciences

University of Basrah

Iraq

Chapter Three

The Steepest Descent and Newton Methods

Lecture 3

Theorem (4): (Convergence Theorem of Newton′s Method)

1: Let f is twice continuously differentiable in R^n.

2: Let X_k be close enough to the solution X^* of the minimization problem with $g(X^*) = 0$.

3: If the Hessian $G(X^*)$ is positive definite matrix.

4: If $G(X)$ satisfies Lipschitz condition

$$\left|G_{ij}(X) - G_{ij}(Y)\right| \le \beta\|X - Y\|, \text{ for some } \beta \text{ , for all } i, j$$

$\text{where } G_{ij}(X) \text{ is the } (i, j) - \text{element of } G(X)$.

Then

1: For all k Newton′s iteration $X_{k+1} = X_k - {G_k}^{-1} g_k$ is well defined.

2: The generated sequence $\{X_k\}$ converges to X^* with a quadratic rate.

Algorithm (3): (Newton′s Method with Line Search)

Step 1: Given $X_0 \epsilon R^n, \varepsilon > 0, k = 0$.

Step 2: Compute g_k. $If \; \|g_k\| \leq \varepsilon$ stop.

Step 3: Solve $G_k s_k = -g_k$ for s_k.

Step 4: Find α_k such that $f(X_k + \alpha_k s_k) = \min_{\alpha \geq 0} f(X_k + \alpha s_k)$.

Step 5: $Set \; X_{k+1} = X_k + \alpha_k s_k$

$, k = k + 1, and \; go \; to \; step \; 2$.

<u>Theorem (5):</u>

1: Let $f: R^n \to R$ be twice continuously differentiable on an open $convex\ set\ D \subset R^n$.

2: Assume that for any $X_0 \epsilon D$ there exists a constant $m > 0$ such $that\ f(x)\ satisfies$

$$u^T \nabla^2 f(X) u \geq m\|u\|^2, for\ all\ u\ \epsilon\ R^n, X\ \epsilon\ L(X_0), where$$

$$L(X_0) = \{X: f(X) \leq f(X_0)\ is\ the\ corresponding\ level\ set.$$

Then the sequence $\{X_k\}$ generated by Algorithm (3) satisfies:

1: when $\{X_k\}$ $is\ a\ finite\ sequence,\ g_k = 0\ for\ some\ k$.

2: when $\{X_k\}$ is an infinite sequence, $\{X_k\}$ converges to the unique $minimizer\ X^*\ of\ f$.

Example:

Use the Newton′s Method with Line Search to find the minimizer of the objective function $f(X) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$.

Take $X_0 = [0, 0]^T$ and $\varepsilon = 0.0001$.

Solution:

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

$$g(X) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right]^T = [1 + 4x_1 + 2x_2, -1 + 2x_1 + 2x_2]^T .$$

Now, find the Hessian matrix $G(X)$ as:

$$G = G(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial {x_2}^2} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}.$$

$Now, compute\ g_0 = g(X_0) = [1, -1]^T.$

Compute $G^{-1} = \frac{1}{4}\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$.

$$\textit{Compute } s_0 = -G^{-1}g_0 = -\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} =$$

$$\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}.$$

Compute $X_1 = X_0 + \alpha s_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \alpha\begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} -\alpha \\ \frac{3}{2}\alpha \end{bmatrix}$.

Evaluate $f(X_1) = -\alpha - \frac{3}{2}\alpha + 2\alpha^2 - 3\alpha^2 + \frac{9}{4}\alpha^2 = -\frac{5}{2}\alpha + \frac{5}{4}\alpha^2$.

$df \quad 5 + 5\,\alpha$

Set $\frac{df}{d\alpha} = 0 \rightarrow -\frac{5}{2} + \frac{5}{2}\alpha = 0 \rightarrow \alpha = 1$.

$\therefore \frac{d^2 f}{d\alpha^2} = \frac{5}{2} > 0 \rightarrow \alpha \; is \; a \; minimizer \; of \; f(X_1)$.

$\therefore X_1 = \begin{bmatrix} -\alpha \\ \frac{3}{2}\alpha \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}$.

$Now, we \; compute \; g_1 = g(X_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

$Since \; \|g_1\| = 0 < \varepsilon$.

$\therefore X_1 = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}$ is the minimizer of $f(X)$.

H.W.

Use the Newton′s Method with Line Search to find the minimizer of the objective $function\ f(X) = x_1^2 + 2x_2^2$ **.**

Take $X_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ **and** $\varepsilon = 0.0001$.