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 4

Causes of Failure Newton′s Method

Case 1:

The inverse of Hessian matrix exists and it is positive definite but the search direction s_k is so large that $f(X_{k+1}) > f(X_k)$.

Case 2:

The search direction s_k is orthogonal to the $gradient\ vector\ g_k$.

Case 3:

The inverse of Hessian matrix exists and it is not positive definite.

Case 4:

The inverse of Hessian matrix does not exist.

We now consider some strategies for reducing the probability of failure du to above cases.

Case 1:

$\boldsymbol{Take\ X_{k+1} = X_k + \alpha_k s_k}$, **where the value of** $\boldsymbol{\alpha_k \in (0, 1)}$ $\boldsymbol{must\ be\ found\ such\ that\ f(X_k + \alpha_k s_k) < f(X_k)}$. **A simple strategy for computing a value of** $\boldsymbol{\alpha_k}$ **is given in the following algorithm.**

Algorithm (4):

1: **Set** $\boldsymbol{\alpha = 1}$.

2: **Compute** $\boldsymbol{f(X_k + \alpha s_k)}$.

3: **If** $\boldsymbol{f(X_k + \alpha s_k) < f(X_k)}$ **go to step 5.**

4: **Set** $\boldsymbol{\alpha = \frac{\alpha}{2}}$ **and go to step 2.**

5: **Set** $\boldsymbol{\alpha_k = \alpha}$ **and** $\boldsymbol{f(X_{k+1}) = f(X_k)}$.

6: **Stop.**

Replace s_k $with$ $-g_k$.
Thus if g_k and s_k are orthogonal, we take a steepest descent direction.

<u>Case 3:</u>
Replace s_k $with$ $-s_k$ and take a steepest descent direction.

<u>Case 4:</u>
Replace s_k $with$ $-g_k$ and take a steepest descent direction.

Now, we give two examples show that a failure of Newton′s method.

Example (1):

$\boldsymbol{Let\ f: R^2 \to R\ is\ defined\ by}$

$$\boldsymbol{f(X) = x_1^4 + x_1x_2 + (x_2+1)^2\ with\ X_0 = \begin{bmatrix}0\\0\end{bmatrix}}.$$

First, we find the gradient vector $\boldsymbol{g(X)}$ as:

$$\boldsymbol{g(X) = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}]^T = [4x_1^3 + x_2, x_1 + 2(x_2+1)]^T}.$$

We find the Hessian matrix $\boldsymbol{G(X)}$ as:

$$\boldsymbol{G = G(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial {x_2}^2} \end{bmatrix} = \begin{bmatrix} 12x_1^2 & 1 \\ 1 & 2 \end{bmatrix}}.$$

$$\boldsymbol{\therefore g_0 = g(X_0) = [0, 2]^T}.$$

$$\boldsymbol{G(X_0) = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \to G^{-1}(X_0) = -\begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}}.$$

Now, notice that

$$g_0{}^T s_0 = [0 \quad 2]\begin{bmatrix} -2 \\ 0 \end{bmatrix} = 0\,.$$

Hence the Newton′s method is failure (see case 2).

<u>Example (2):</u>

$\boldsymbol{Let\ f:\ R^2\ \to R\ is\ defined\ by}$

$$f(X) = x_1^4 - 3x_1x_2 + (x_2 + 2)^2 \ with\ X_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

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

$$g(X) = [\frac{\partial f}{\partial x_1}\,, \frac{\partial f}{\partial x_2}]^T = [4x_1^3 - 3x_2\,, -3x_1 + 2(x_2 + 2)]^T.$$

We find the Hessian matrix $G(X)$ as:

$$G = G(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial {x_2}^2} \end{bmatrix} = \begin{bmatrix} 12x_1^2 & -3 \\ -3 & 2 \end{bmatrix}.$$

$$\therefore g_0 = g(X_0) = [0, 4]^T.$$

$$G(X_0) = \begin{bmatrix} 0 & -3 \\ -3 & 2 \end{bmatrix} \rightarrow G^{-1}(X_0) = -\frac{1}{9}\begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{9} & -\frac{3}{9} \\ -\frac{3}{9} & 0 \end{bmatrix}$$

We find the direction s_0 as:

$$s_0 = -G^{-1}(X_0)g_0 = -\begin{bmatrix} -\frac{2}{9} & -\frac{3}{9} \\ -\frac{3}{9} & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{2}{9} & \frac{3}{9} \\ \frac{3}{9} & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{12}{9} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ 0 \end{bmatrix}.$$

$$\therefore X_1 = X_0 + s_0 = \begin{bmatrix} \frac{4}{3} \\ 0 \end{bmatrix}.$$

$$\therefore f(X_1) = (\tfrac{4}{3})^4 + 4 = \frac{256}{81} + 4 = \frac{590}{81} \approx 7.2840 \,.$$

$$f(X_0) = 4 \,.$$

$$\therefore f(X_1) > f(X_0) \,.$$

Hence the Newton′s method is failure (see case 1).