Introduction to Engineering Optimization (ME6806)

Dr. Yogesh Kumar
Assistant Professor
Mechanical Engineering Department
National Institute of Technology Patna
Bihar - 800 005, India
yogesh.me@nitp.ac.in

Module 4

Multi-variable Optimization Algorithms

Davidon–Fletcher–Powell Method

The iterative procedure of the Davidon–Fletcher–Powell (DFP) method can be described as follows:

1. Start with an initial point $\mathbf{X}_1$ and a $n \times n$ positive definite symmetric matrix $[B_1]$ to approximate the inverse of the Hessian matrix of f. Usually, $[B_1]$ is taken as the identity matrix $[I]$. Set the iteration number as $i = 1$.
2. Compute the gradient of the function, ∇f_i, at point $\mathbf{X}_i$, and set

$$\mathbf{S}_i = -[B_i]\nabla f_i$$

3. Find the optimal step length λ_i^* in the direction $\mathbf{S}_i$ and set

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \lambda_i^* \mathbf{S}_i$$

4. Test the new point $\mathbf{X}_{i+1}$ for optimality. If $\mathbf{X}_{i+1}$ is optimal, terminate the iterative process. Otherwise, go to step 5.

Davidon–Fletcher–Powell Method

5. Update the matrix $[B_i]$, as

$$[B_{i+1}] = [B_i] + [M_i] + [N_i]$$

where

$$[M_i] = \lambda_i^* \frac{\mathbf{S}_i \mathbf{S}_i^{\mathrm{T}}}{\mathbf{S}_i^T \mathbf{g}_i}$$

$$[N_i] = -\frac{([B_i]\mathbf{g}_i)([B_i]\mathbf{g}_i)^{\mathrm{T}}}{\mathbf{g}_i^T [B_i]\mathbf{g}_i}$$

$$\mathbf{g}_i = \nabla f(\mathbf{X}_{i+1}) - \nabla f(\mathbf{X}_i) = \nabla f_{i+1} - \nabla f_i$$

6. Set the new iteration number as $i = i + 1$, and go to step 2.

Davidon–Fletcher–Powell Method - Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ from the starting point $\mathbf{X}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$ using the DFP method with

$$[B_1] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \varepsilon = 0.01$$

Solution

Iteration 1 (i = 1)

Here

$$\nabla f_1 = \nabla f(\mathbf{X}_1) = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}\Bigg|_{(0,0)} = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$$

Davidon–Fletcher–Powell Method - Example

hence

$$\mathbf{S}_1 = -[B_1]\nabla f_1 = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

To find the minimizing step length λ_1^* along $\mathbf{S}_1$, we minimize

$$f(\mathbf{X}_1 + \lambda_1 \mathbf{S}_1) = f\left(\begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + \lambda_1 \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} \right) = f(-\lambda_1, \lambda_1) = \lambda_1^2 - 2\lambda_1$$

with respect to λ_1. Since $df/d\lambda_1 = 0$ at $\lambda_1^* = 1$, we obtain

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^* \mathbf{S}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + 1 \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

Davidon–Fletcher–Powell Method - Example

Since $\nabla f_2 = \nabla f(\mathbf{X}_2) = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix}$ and $\|\nabla f_2\| = 1.4142 > \varepsilon$, we proceed to update the matrix $[B_i]$ by computing

$$\mathbf{g}_1 = \nabla f_2 - \nabla f_1 = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix} - \begin{Bmatrix} 1 \\ -1 \end{Bmatrix} = \begin{Bmatrix} -2 \\ 0 \end{Bmatrix}$$

$$\mathbf{S}_1^{\mathrm{T}}\mathbf{g}_1 = \{-1 \quad 1\} \begin{Bmatrix} -2 \\ 0 \end{Bmatrix} = 2$$

$$\mathbf{S}_1\mathbf{S}_1^{\mathrm{T}} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} \{-1 \quad 1\} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$[B_1]\mathbf{g}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} -2 \\ 0 \end{Bmatrix} = \begin{Bmatrix} -2 \\ 0 \end{Bmatrix}$$

$$([B_1]\mathbf{g}_1)^{\mathrm{T}} = \begin{Bmatrix} -2 \\ 0 \end{Bmatrix}^{\mathrm{T}} = \{-2 \quad 0\}$$

Davidon–Fletcher–Powell Method - Example

$$\mathbf{g}_1^{\mathrm{T}}[B_1]\mathbf{g}_1 = \{-2 \;\; 0\}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{Bmatrix}-2\\ 0\end{Bmatrix} = \{-2 \;\; 0\}\begin{Bmatrix}-2\\ 0\end{Bmatrix} = 4$$

$$[M_1] = \lambda_1^* \frac{\mathbf{S}_1\mathbf{S}_1^{\mathrm{T}}}{\mathbf{S}_1^{\mathrm{T}}\mathbf{g}_1} = 1\left(\frac{1}{2}\right)\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix} = \begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix}$$

$$[N_1] = -\frac{([B_1]\mathbf{g}_1)([B_1]\mathbf{g}_1)^{\mathrm{T}}}{\mathbf{g}_1^T[B_1]\mathbf{g}_1} = -\frac{\begin{Bmatrix}-2\\ 0\end{Bmatrix}\{-2 \;\; 0\}}{4} = -\frac{1}{4}\begin{bmatrix}4 & 0\\ 0 & 0\end{bmatrix} = -\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$$

$$[B_2] = [B_1] + [M_1] + [N_1] = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2}\end{bmatrix} + \begin{bmatrix}-1 & 0\\ 0 & 0\end{bmatrix} = \begin{bmatrix}0.5 & -0.5\\ -0.5 & 1.5\end{bmatrix}$$

Davidon–Fletcher–Powell Method - Example

Iteration 2 (i = 2)

The next search direction is determined as

$$\mathbf{S}_2 = -[B_2]\nabla f_2 = -\begin{bmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{bmatrix}\begin{Bmatrix} -1 \\ -1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$$

To find the minimizing step length λ_2^* along $\mathbf{S}_2$, we minimize

$$f(\mathbf{X}_2 + \lambda_2 \mathbf{S}_2) = f\left(\begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \lambda_2 \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}\right) = f\left(\begin{Bmatrix} -1 \\ 1+\lambda_2 \end{Bmatrix}\right)$$

$$= -1 - (1+\lambda_2) + 2(-1)^2 + 2(-1)(1+\lambda_2) + (1+\lambda_2)^2$$

$$= \lambda_2^2 - \lambda_2 - 1$$

with respect to λ_2. Since $df/d\lambda_2 = 0$ at $\lambda_2^* = \frac{1}{2}$,

Davidon–Fletcher–Powell Method - Example

$$f(\mathbf{X}_2 + \lambda_2 \mathbf{S}_2) = \lambda_2^2 - \lambda_2 - 1$$

with respect to λ_2. Since $df/d\lambda_2 = 0$ at $\lambda_2^* = \frac{1}{2}$, we obtain

$$\mathbf{X}_3 = \mathbf{X}_2 + \lambda_2 \mathbf{S}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{2} \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1.5 \end{Bmatrix}$$

This point can be identified to be optimum since

$$\nabla f_3 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad \text{and} \quad ||\nabla f_3|| = 0 < \varepsilon$$

Thanks