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

Outlines

- **Optimality Criteria**
- **Direct Search Methods**
  - Nelder and Mead (Simplex Search)
  - Hook and Jeeves (Pattern Search)
  - Powell’s Method (The Conjugated Direction Search)
- **Gradient Based Methods**
  - Steepest Descent (Cauchy’s) Method
  - Newton’s Method
  - Modified Newton’s Method
  - Marquardt’s Method
  - **Conjugate Gradient Method**
  - Quasi-Newton Method
  - Trust Regions
  - Gradient-Based Algorithm
  - Numerical Gradient Approximations

Conjugate Gradient (Fletcher–Reeves) Method

The Fletcher–Reeves method is developed by modifying the steepest descent method to make it quadratically convergent. Starting from an arbitrary point $\mathbf{X}_1$, the quadratic function

$$f(\mathbf{X}) = \frac{1}{2}\mathbf{X}^{\mathrm{T}}[\mathbf{A}]\mathbf{X} + \mathbf{B}^{\mathrm{T}}\mathbf{X} + C$$

can be minimized by searching along the search direction $\mathbf{S}_1 = -\nabla f_1$ (steepest descent direction) using the step length

$$\lambda_1^* = -\frac{\mathbf{S}_1^T \nabla f_1}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1}$$

The second search direction $\mathbf{S}_2$ is found as a linear combination of $\mathbf{S}_1$ and $-\nabla f_2$:

Conjugate Gradient (Fletcher–Reeves) Method

$$\mathbf{S}_2 = -\nabla f_2 + \beta_2 \mathbf{S}_1$$

where the constant β_2 can be determined by making $\mathbf{S}_1$ and $\mathbf{S}_2$ conjugate with respect to $[A]$.

$$\beta_2 = -\frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \mathbf{S}_1} = \frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \nabla f_1}$$

This process can be continued to obtain the general formula for the ith search direction as

$$\mathbf{S}_i = -\nabla f_i + \beta_i \mathbf{S}_{i-1}$$

where $\beta_i = \dfrac{\nabla f_i^T \nabla f_i}{\nabla f_{i-1}^T \nabla f_{i-1}}$

Conjugate Gradient (Fletcher–Reeves) Method

The iterative procedure of Fletcher–Reeves method can be stated as follows:

1. Start with an arbitrary initial point $\mathbf{X}_1$.
2. Set the first search direction $\mathbf{S}_1 = -\nabla f(\mathbf{X}_1) = -\nabla f_1$.
3. Find the point $\mathbf{X}_2$ according to the relation

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^* \mathbf{S}_1$$

   where λ_1^* is the optimal step length in the direction $\mathbf{S}_1$. Set $i = 2$ and go to the next step.
4. Find $\nabla f_i = \nabla f(\mathbf{X}_i)$, and set

$$\mathbf{S}_i = -\nabla f_i + \frac{|\nabla f_i|^2}{|\nabla f_{i-1}|^2}\mathbf{S}_{i-1}$$

Conjugate Gradient (Fletcher–Reeves) Method

5. Compute the optimum step length λ_i^* in the direction $\mathbf{S}_i$, and find the new point

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

6. Test for the optimality of the point $\mathbf{X}_{i+1}$. If $\mathbf{X}_{i+1}$ is optimum, stop the process. Otherwise, set the value of $i = i + 1$ and go to step 4.

EXAMPLE

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

Conjugate Gradient (Fletcher–Reeves) Method – Example

Iteration 1

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}$$

$$\nabla f_1 = \nabla f(\mathbf{X}_1) = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$$

The search direction is taken as $\mathbf{S}_1 = -\nabla f_1 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$. To find the optimal step length λ_1^* along $\mathbf{S}_1$, we minimize $f(\mathbf{X}_1 + \lambda_1 \mathbf{S}_1)$ with respect to λ_1. Here

$$f(\mathbf{X}_1 + \lambda_1 \mathbf{S}_1) = f(-\lambda_1, +\lambda_1) = \lambda_1^2 - 2\lambda_1$$

$$\frac{df}{d\lambda_1} = 0 \quad \text{at} \quad \lambda_1^* = 1$$

Therefore,

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

Conjugate Gradient (Fletcher–Reeves) Method – Example

Iteration 2

Since $\nabla f_2 = \nabla f(\mathbf{X}_2) = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix}$, the next search direction as

$$\mathbf{S}_2 = -\nabla f_2 + \frac{|\nabla f_2|^2}{|\nabla f_1|^2}\mathbf{S}_1$$

where

$$|\nabla f_1|^2 = 2 \quad \text{and} \quad |\nabla f_2|^2 = 2$$

Therefore,

$$\mathbf{S}_2 = -\begin{Bmatrix} -1 \\ -1 \end{Bmatrix} + \left(\frac{2}{2}\right)\begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ +2 \end{Bmatrix}$$

Conjugate Gradient (Fletcher–Reeves) Method – Example

To find λ_2^*, we minimize

$$f(\mathbf{X}_2 + \lambda_2 \mathbf{S}_2) = f(-1, 1 + 2\lambda_2)$$
$$= -1 - (1 + 2\lambda_2) + 2 - 2(1 + 2\lambda_2) + (1 + 2\lambda_2)^2$$
$$= 4\lambda_2^2 - 2\lambda_2 - 1$$

with respect to λ_2. As $df/d\lambda_2 = 8\lambda_2 - 2 = 0$ at $\lambda_2^* = \frac{1}{4}$, we obtain

$$\mathbf{X}_3 = \mathbf{X}_2 + \lambda_2^* \mathbf{S}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{4} \begin{Bmatrix} 0 \\ 2 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1.5 \end{Bmatrix}$$

Thus the optimum point is reached in two iterations. Even if we do not know this point to be optimum, we will not be able to move from this point in the next iteration. This can be verified as follows.

Conjugate Gradient (Fletcher–Reeves) Method – Example

Iteration 3

Now

$$\nabla f_3 = \nabla f(\mathbf{X}_3) = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}, \quad |\nabla f_2|^2 = 2, \quad \text{and} \quad |\nabla f_3|^2 = 0.$$

Thus

$$\mathbf{S}_3 = -\nabla f_3 + (|\nabla f_3|^2/|\nabla f_2|^2)\mathbf{S}_2 = -\begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + \left(\frac{0}{2}\right)\begin{Bmatrix} 0 \\ 2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$$

This shows that there is no search direction to reduce f further, and hence $\mathbf{X}_3$ is optimum.

Thanks