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

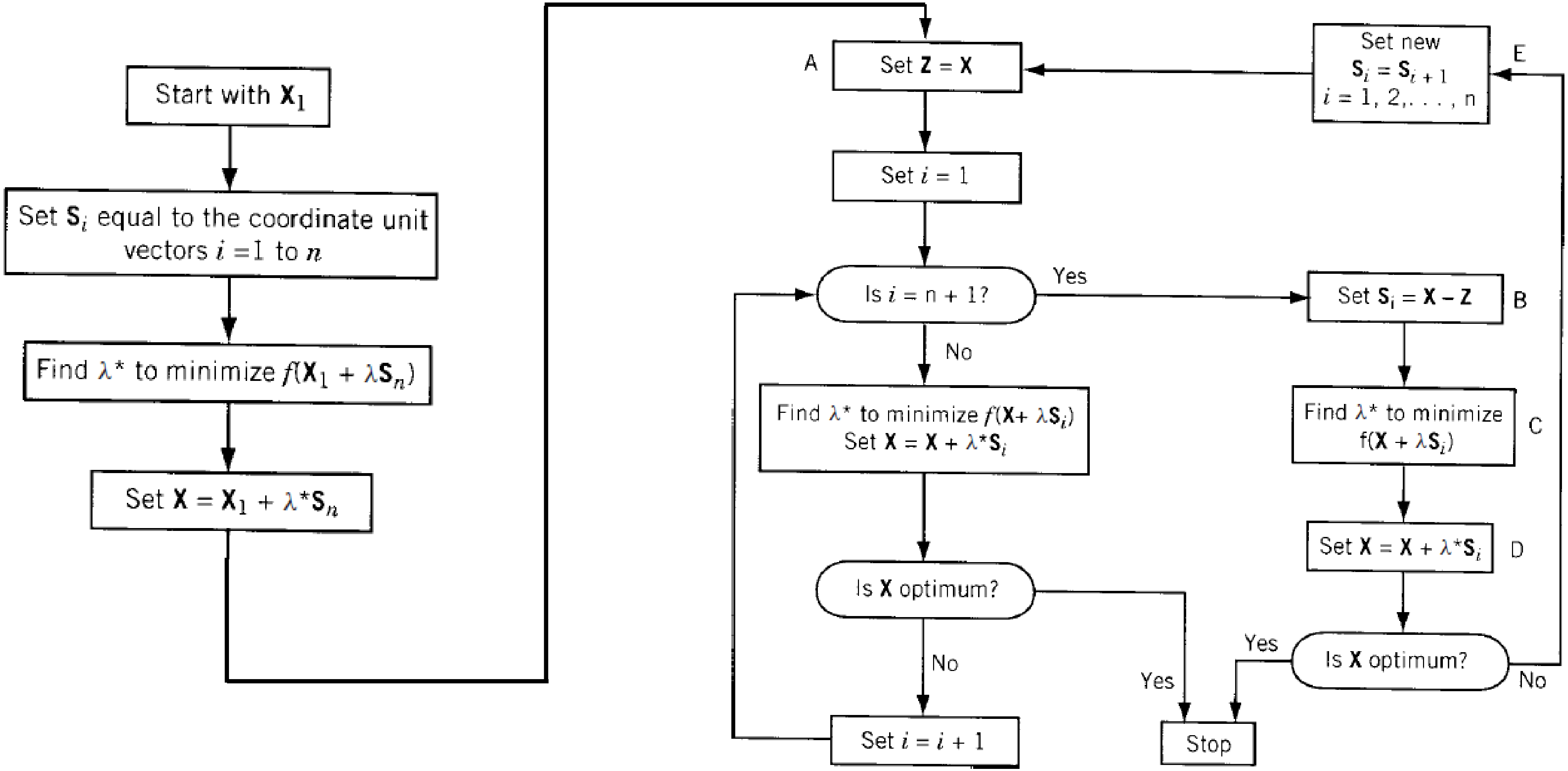
Module 4

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**
  - 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

Powell’s Method



Powell's Method

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 Powell's method.

Solution

Cycle 1: Univariate Search

We minimize f along $\mathbf{S}_2 = \mathbf{S}_n = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ from $\mathbf{X}_1$. To find the correct direction ($+\mathbf{S}_2$ or $-\mathbf{S}_2$) for decreasing the value of f, we take the probe length as $\varepsilon = 0.01$. As $f_1 = f(\mathbf{X}_1) = 0.0$, and

$$f^+ = f(\mathbf{X}_1 + \varepsilon \mathbf{S}_2) = f(0.0, 0.01) = -0.0099 < f_1$$

f decreases along the direction $+\mathbf{S}_2$. To find the minimizing step length λ^* along $\mathbf{S}_2$, we minimize

$$f(\mathbf{X}_1 + \lambda \mathbf{S}_2) = f(0.0, \lambda) = \lambda^2 - \lambda$$

Powell's Method

As $df/d\lambda = 0$ at $\lambda^* = \frac{1}{2}$, we have $\mathbf{X}_2 = \mathbf{X}_1 + \lambda^*\mathbf{S}_2 = \begin{Bmatrix} 0 \\ 0.5 \end{Bmatrix}$.

Next we minimize f along $\mathbf{S}_1 = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$ from $\mathbf{X}_2 = \begin{Bmatrix} 0.5 \\ 0.0 \end{Bmatrix}$. Since

$$f_2 = f(\mathbf{X}_2) = f(0.0, 0.5) = -0.25$$

$$f^+ = f(\mathbf{X}_2 + \varepsilon\mathbf{S}_1) = f(0.01, 0.50) = -0.2298 > f_2$$

$$f^- = f(\mathbf{X}_2 - \varepsilon\mathbf{S}_1) = f(-0.01, 0.50) = -0.2698$$

f decreases along $-\mathbf{S}_1$. As $f(\mathbf{X}_2 - \lambda\mathbf{S}_1) = f(-\lambda, 0.50) = 2\lambda^2 - 2\lambda - 0.25$, $df/d\lambda =$ 0 at $\lambda^* = \frac{1}{2}$. Hence $\mathbf{X}_3 = \mathbf{X}_2 - \lambda^*\mathbf{S}_1 = \begin{Bmatrix} -0.5 \\ 0.5 \end{Bmatrix}$.

Powell's Method

Now we minimize f along $\mathbf{S}_2 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ from $\mathbf{X}_3 = \begin{Bmatrix} -0.5 \\ 0.5 \end{Bmatrix}$. As $f_3 = f(\mathbf{X}_3) = -0.75$, $f^+ = f(\mathbf{X}_3 + \varepsilon\mathbf{S}_2) = f(-0.5, 0.51) = -0.7599 < f_3$, f decreases along $+\mathbf{S}_2$ direction. Since

$$f(\mathbf{X}_3 + \lambda\mathbf{S}_2) = f(-0.5, 0.5 + \lambda) = \lambda^2 - \lambda - 0.75, \quad \frac{df}{d\lambda} = 0 \text{ at } \lambda^* = \frac{1}{2}$$

This gives

$$\mathbf{X}_4 = \mathbf{X}_3 + \lambda^*\mathbf{S}_2 = \begin{Bmatrix} -0.5 \\ 1.0 \end{Bmatrix}$$

Powell's Method

Cycle 2: Pattern Search

Now we generate the first pattern direction as

$$\mathbf{S}_p^{(1)} = \mathbf{X}_4 - \mathbf{X}_2 = \begin{Bmatrix} -\frac{1}{2} \\ 1 \end{Bmatrix} - \begin{Bmatrix} 0 \\ \frac{1}{2} \end{Bmatrix} = \begin{Bmatrix} -0.5 \\ 0.5 \end{Bmatrix}$$

and minimize f along $\mathbf{S}_p^{(1)}$ from $\mathbf{X}_4$. Since

$$f_4 = f(\mathbf{X}_4) = -1.0$$

$$f^+ = f(\mathbf{X}_4 + \varepsilon \mathbf{S}_p^{(1)}) = f(-0.5 - 0.005, 1 + 0.005)$$

$$= f(-0.505, 1.005) = -1.004975$$

Powell's Method

f decreases in the positive direction of $\mathbf{S}_p^{(1)}$. As

$$f(\mathbf{X}_4 + \lambda \mathbf{S}_p^{(1)}) = f(-0.5 - 0.5\lambda, 1.0 + 0.5\lambda)$$
$$= 0.25\lambda^2 - 0.50\lambda - 1.00,$$

$\dfrac{df}{d\lambda} = 0$ at $\lambda^* = 1.0$ and hence

$$\mathbf{X}_5 = \mathbf{X}_4 + \lambda^* \mathbf{S}_p^{(1)} = \begin{Bmatrix} -\frac{1}{2} \\ 1 \end{Bmatrix} + 1.0 \begin{Bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{Bmatrix} = \begin{Bmatrix} -1.0 \\ 1.5 \end{Bmatrix}$$

The point $\mathbf{X}_5$ can be identified to be the optimum point.

Powell's Method

If we do not recognize $\mathbf{X}_5$ as the optimum point at this stage, we proceed to minimize f along the direction $S_2 = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ from $\mathbf{X}_5$. Then we would obtain

$$f_5 = f(\mathbf{X}_5) = -1.25, \quad f^+ = f(\mathbf{X}_5 + \varepsilon \mathbf{S}_2) > f_5,$$
$$\text{and} \quad f^- = f(\mathbf{X}_5 - \varepsilon \mathbf{S}_2) > f_5$$

This shows that f cannot be minimized along $\mathbf{S}_2$, and hence $\mathbf{X}_5$ will be the optimum point. In this example the convergence has been achieved in the second cycle itself. This is to be expected in this case, as f is a quadratic function, and the method is a quadratically convergent method.