08 - Oct - 2021

D - F - P Method

- Approximate the inverse of the Hessian

1. Start with an initial guess point $\underset{\sim}{X}_1$ and an $n \times n$ symmetric positive definite matrix $[B_1]$ to approximate the inverse of the Hessian of f.

   $[B_1] = [I]$ usual choice. $i = 1$

2. $$\underset{\sim}{S}_i = -[B_i]\nabla f_i$$ ← compute ∇f_i at $\underset{\sim}{X}_i$ & then obtain $\underset{\sim}{S}_i$

3. $$\underset{\sim}{X}_{i+1} = \underset{\sim}{X}_i + \lambda_i^* \underset{\sim}{S}_i$$ ← obtain λ_i^* by 1-D minimization

4. Test $\underset{\sim}{X}_{i+1}$ for optimality.

   $$\left|\nabla f_{i+1}\right| \le \varepsilon$$ → small quantity close to zero.

   If X_{i+1} is found to be optimal, terminate the search. Else go to step 5.

5. Update $[B_i]$

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

where $$[M_i] = \lambda_i^* \frac{\underset{\sim}{S}_i \underset{\sim}{S}_i^T}{\underset{\sim}{S}_i^T \underset{\sim}{g}_i} \; ; \; [N_i] = -\frac{([B_i]\underset{\sim}{g}_i)([B_i]\underset{\sim}{g}_i)^T}{\underset{\sim}{g}_i^T [B_i] \underset{\sim}{g}_i}$$

$$\underset{\sim}{g}_i = \nabla f_{i+1} - \nabla f_i$$

6. set $i = i+1$ & go to step 2

<u>Note:</u> $[B_{i+1}]$ remains positive definite only if λ_i^* is computed accurately.

<u>Example:</u> $f(x_1, x_2) = 100(x_1^2 - x_2)^2 + (1-x_1)^2$ with $\underset{\sim}{X}_1 = \begin{Bmatrix} -2 \\ -2 \end{Bmatrix}$.

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 400x_1(x_1^2 - x_2) - 2(1-x_1) \\ -200(x_1^2 - x_2) \end{Bmatrix}$$

1) $i = 1$

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

At $\underset{\sim}{X}_1 = \begin{Bmatrix} -2 \\ -2 \end{Bmatrix}$, $\nabla f_1 = \begin{Bmatrix} -4806 \\ -1200 \end{Bmatrix}$ and $f_1 = 3609$.

$$\underset{\sim}{S}_1 = -[B_1]\nabla f_1$$

$$= \begin{Bmatrix} 4806 \\ 1200 \end{Bmatrix} \cdot \frac{1}{\sqrt{4806^2 + 1200^2}} \begin{Bmatrix} 4806 \\ 1200 \end{Bmatrix} =$$

Normalizing $\underset{\sim}{S}_1 = \begin{Bmatrix} 0.970 \\ 0.244 \end{Bmatrix}$

$$\underset{\sim}{X}_2 = \underset{\sim}{X}_1 + \lambda_1^* \underset{\sim}{S}_1 .$$

To find λ_1^* Minimize $f(\underset{\sim}{X}_1 + \lambda_1 \underset{\sim}{S}_1) = f(-2 + 0.97\lambda_1, -2 + 0.244\lambda_1)$

$$f = 100(6 - 4.124\lambda_1 + 0.938\lambda_1^2)^2 + (3 - 0.97\lambda_1)^2$$

$$\lambda_1^* = 2.201$$

$$\therefore \quad \underset{\sim}{x}_2 = \begin{Bmatrix} -2 \\ -2 \end{Bmatrix} + 2.201 \begin{Bmatrix} 0.970 \\ 0.244 \end{Bmatrix}$$

$$\underset{\sim}{x}_2 = \begin{Bmatrix} 0.135 \\ -1.463 \end{Bmatrix}$$

Test $\underset{\sim}{x}_2$ for optimality $\quad \nabla f_2 = \begin{Bmatrix} 78.29 \\ -296.24 \end{Bmatrix} \neq \underset{\sim}{0}$

$$\therefore \quad [B_2] = [B_i] + [M_i] + [N_i]$$

$$[M_i] = \frac{\lambda_i^{*} \, \underset{\sim}{S}_i \, \underset{\sim}{S}_i^{T}}{\underset{\sim}{S}_i^{T} \, \underset{\sim}{g}_i} =$$

$$[N_i] = -\frac{\left([B_i] \, \underset{\sim}{g}_i\right) \left([B_i] \, \underset{\sim}{g}_i\right)^{T}}{\underset{\sim}{g}_i^{T} \, \underset{\sim}{S}_i \, \underset{\sim}{g}_i}$$

$$\underset{\sim}{g}_i = \nabla f_2 - \nabla f_1 = \begin{Bmatrix} 78.29 \\ -296.24 \end{Bmatrix} - \begin{Bmatrix} -4806 \\ -1200 \end{Bmatrix}$$

Algorithm for BFGS method.

1. Start with $\underset{\sim}{X}_1$ & $[B_1] = I_{n\times n}$. Compute ∇f_1 set $i=1$

2. set

$$[S_i] = -[B_i]\nabla f_i$$

3. Find optimal step length λ_i^* along $\underset{\sim}{S}_i$ to get

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

4. Test $\underset{\sim}{X}_{i+1}$ for optimality. If $\|\nabla f_{i+1}\| \leq \varepsilon$. $\underset{\sim}{X}^* \approx \underset{\sim}{X}_{i+1}$ & stop
Else go to step 5.

5. update the Hessian.

$$[B_{i+1}] = [B_i] + \left(1 + \frac{\underset{\sim}{g}_i^T [B_i] \underset{\sim}{g}_i}{\underset{\sim}{d}_i^T \underset{\sim}{g}_i}\right) \frac{\underset{\sim}{d}_i \underset{\sim}{d}_i^T}{\underset{\sim}{d}_i^T \underset{\sim}{g}_i}$$

$$- \frac{\underset{\sim}{d}_i \underset{\sim}{g}_i^T [B_i]}{\underset{\sim}{d}_i^T \underset{\sim}{g}_i} - \frac{[B_i] \underset{\sim}{g}_i \underset{\sim}{d}_i^T}{\underset{\sim}{d}_i^T \underset{\sim}{g}_i}.$$

$$\underset{\sim}{d}_i = \underset{\sim}{X}_{i+1} - \underset{\sim}{X}_i = \lambda_i^* \underset{\sim}{S}_i$$

$$\underset{\sim}{g}_i = \nabla f_{i+1} - \nabla f_i$$

6. Set $i = i+1$ & go to step 2

124

Test functions:

1. Rosenbrock's parabolic valley.

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1-x_1)^2$$

$$\underset{\sim}{X}_1 = \begin{Bmatrix} -1.2 \\ 1.0 \end{Bmatrix}, \quad \underset{\sim}{X}^* = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

$$f_1 = 24 \qquad f^* = 0.$$

2. Quadratic function:

$$f(x_1, x_2) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$$

$$\underset{\sim}{X}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \qquad \underset{\sim}{X}^* = \begin{Bmatrix} 1 \\ 3 \end{Bmatrix}$$

$$f_1 = 74 \qquad f^* = 0$$

3. powell's quartic function:

$$f(x_1, x_2, x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

$$\underline{X}_1 = \begin{Bmatrix} 3 \\ -1 \\ 0 \\ 1 \end{Bmatrix} \qquad \underline{X}^* = \begin{Bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{Bmatrix}$$

$$f_1 = 215 \qquad f^* = 0$$

4. Fletcher & powell's helical valley

$$f(x_1, x_2, x_3) = 100\left\{\left[x_3 - 10\,\underbrace{\theta(x_1, x_2)}\right]^2 + \left[\sqrt{x_1^2 + x_2^2} - 1\right]^2\right\} + x_3^2$$

where

$$2\pi\,\theta(x_1, x_2) = \begin{cases} \tan^{-1}\left(\dfrac{x_2}{x_1}\right) & \text{if } x_1 > 0 \\ \pi + \tan^{-1}\left(\dfrac{x_2}{x_1}\right) & \text{if } x_1 < 0. \end{cases}$$

$$\underset{\sim}{x}_1 = \begin{Bmatrix} -1 \\ 0 \\ 0 \end{Bmatrix} \qquad \underset{\sim}{x}^* = \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix}$$

$$f_1 = 25{,}000 \qquad f^* = 0$$

12