Department of Mechanical Engineering, IIT Bombay

ME 782 Design Optimization Quiz 2 September 4,

Part 1

1. $\mathbf{x}_0 = \begin{bmatrix} 2.3449 \\ 1.8164 \end{bmatrix}$

2. $\mathbf{x}_0 = \begin{bmatrix} -0.1875 \\ -0.1250 \end{bmatrix}$

Part 2

---

1. Consider the following function:

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

Beginning from [1.25, 1.5] carry out one iteration (including one-dimensional search to determine the magnitude of search direction) towards minimization of the above function, using the steepest descent method. Assume any other necessary data and clearly state them. **[7 marks]**

$\mathbf{x}_0 = \begin{bmatrix} 1.25 \\ 1.5 \end{bmatrix}$

The given function can be written as

$$f(x_1, x_2) = 100x_1^4 - 200x_1^2 x_2 + 100x_2^2 + x_1^2 - 2x_1 + 1 \qquad (1)$$

The gradient of the above equation is given by

$$\nabla f = \begin{bmatrix} 2x_1 - 400x_1\left(-x_1^2 + x_2\right) - 2 \\ 200\left(x_2 - x_1^2\right) \end{bmatrix} \qquad (2)$$

The search direction based on steepest descent at the given starting point is given by

$$s_1 = -\nabla f = -\begin{bmatrix} 2x_1 - 400x_1\left(-x_1^2 + x_2\right) - 2 \\ 200\left(x_2 - x_1^2\right) \end{bmatrix} \qquad (3)$$

$$= -\begin{bmatrix} 2 \times 1.25 - 400 \times 1.25\left(-1.25^2 + 1.5\right) - 2 \\ 200\left(1.5 - 1.25^2\right) \end{bmatrix} = \begin{bmatrix} -31.75 \\ 12.5 \end{bmatrix}$$

The unit vector along the search direction is given by

$$\hat{s}_1 = \frac{1}{\sqrt{31.75^2 + 12.5^2}} \begin{bmatrix} -31.75 \\ 12.5 \end{bmatrix} = \begin{bmatrix} -0.93 \\ 0.37 \end{bmatrix} \tag{4}$$

The next coordinates, based on steepest descent are given by

$$\mathbf{x}_1 = \mathbf{x}_0 + \gamma \hat{\mathbf{s}}_1 = \begin{bmatrix} 1.25 \\ 1.5 \end{bmatrix} + \gamma \begin{bmatrix} -0.93 \\ 0.37 \end{bmatrix} = \begin{bmatrix} 1.25 - 0.93\gamma \\ 1.5 + 0.37\gamma \end{bmatrix} \tag{5}$$

Method 1

From equations (1) and (5), the objective function becomes

$$\begin{aligned} f(\gamma) = {} & 100\left(1.25 - 0.93\gamma\right)^4 - 200\left(1.25 - 0.93\gamma\right)^2\left(1.5 + 0.37\gamma\right) + 100\left(1.5 + 0.37\gamma\right)^2 \\ & + \left(1.25 - 0.93\gamma\right)^2 - 2\left(1.25 - 0.93\gamma\right) + 1 \end{aligned} \tag{6}$$

Equation (6) can be simplified as

$$f(\gamma) = 74.8052\gamma^4 \text{-} 466.1811\gamma^3 + 737.9787\gamma^2 \text{-} 34.1525\gamma + 0.4531 \tag{7}$$

The derivative of equation (7) is given by

$$f^{'}(\gamma) = 1000\left(0.2992\gamma^3 \text{-} 1.3985\gamma^2 + 1.4760\gamma \text{-} 0.0342\right) = 0 \tag{8}$$

The function within the parenthesis of equation (8) can be written as

$$\phi(\gamma) = \left(0.2992\gamma^3 \text{-} 1.3985\gamma^2 + 1.4760\gamma \text{-} 0.0342\right) = 0 \tag{9}$$

Using Newton-Raphson's method the following roots of equation (9) are obtained for the value of γ : 3.08,1.56,0.0237. All these values of γ should be used in equation (5) to get the next set of coordinates and its value that gives the least value of the function defined in equation (1) will give the corresponding next coordinates as shown in Table 1.

Table 1

No	γ	$\mathbf{x}_1 = \mathbf{x}_0 + \gamma \hat{\mathbf{s}}_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ [from equation (5)]	$f(x_1, x_2)$ [from equation (1)]
1	3.08	$\begin{bmatrix} -1.61 \\ 2.64 \end{bmatrix}$	6.9461
2	1.56	$\begin{bmatrix} -0.20 \\ 2.08 \end{bmatrix}$	416.33
3	0.0237	$\begin{bmatrix} 1.23 \\ 1.51 \end{bmatrix}$	0.0520

Hence the value of γ =0.0237 is chosen.

The next coordinates are given by

$$\mathbf{x}_1 = \mathbf{x}_0 + \gamma \hat{\mathbf{s}}_1 = \begin{bmatrix} 1.25 \\ 1.5 \end{bmatrix} + 0.0237 \begin{bmatrix} -0.93 \\ 0.37 \end{bmatrix} = \begin{bmatrix} 1.23 \\ 1.51 \end{bmatrix}$$

Method 2

The Hessian matrix of equation (1) is given by

$$\mathbf{H} = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix} \quad (10)$$

The Hessian matrix evaluated at $\mathbf{x}_0 = \begin{bmatrix} 1.25 \\ 1.5 \end{bmatrix}$ is given by

$$\mathbf{H} = \begin{bmatrix} 1277 & -500 \\ -500 & 200 \end{bmatrix} \quad (11)$$

From equations (3), (4) and (11)

$$\gamma = -\frac{\nabla f(\mathbf{x}_0)^T \mathbf{s}_1}{\mathbf{s}_1^T \mathbf{H}\ \mathbf{s}_1} = -\frac{\begin{bmatrix} 31.75 & 12.5 \end{bmatrix} \begin{bmatrix} -0.93 \\ 0.37 \end{bmatrix}}{\begin{bmatrix} -0.93 & 0.37 \end{bmatrix} \begin{bmatrix} 1277 & -500 \\ -500 & 200 \end{bmatrix} \begin{bmatrix} -0.93 \\ 0.37 \end{bmatrix}} = 0.0231$$

This value of γ=0.0231, corresponds to the least value of the function that can be obtained by the next coordinate and it is given by

$$\mathbf{x}_1 = \mathbf{x}_0 + \gamma \hat{\mathbf{s}}_1 = \begin{bmatrix} 1.25 \\ 1.5 \end{bmatrix} + 0.0231 \begin{bmatrix} -0.93 \\ 0.37 \end{bmatrix} = \begin{bmatrix} 1.23 \\ 1.51 \end{bmatrix}$$

There is a small difference in the value of γ obtained from the two methods that should be acceptable.

---

2. Consider the following function

$$f(x_1, x_2) = 20 + x_1 + 2x_2 + 2.5x_1^2 + 2x_2^2 + 2x_1x_2$$ **[5 marks]**

(a) Obtain the quadratic form of the above equation

The quadratic form of the given equation is given by

$$f(x_1, x_2) = 20 + \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1}$$

The second term of the equation can be recognized as

$$\mathbf{d} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{2}$$

The Hessian matrix can be recognized from the third term of equation (1)

$$\mathbf{H} = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \tag{3}$$

(b) Obtain the stationary values of the function obtained in part(a)

From equations (2) and (3), the stationary values of equation (1) can be obtained as

$$\mathbf{x}^* = -\mathbf{H}^{-1}\mathbf{d} = -\begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \begin{Bmatrix} 1 \\ 2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ -0.5 \end{Bmatrix} \tag{4}$$

(c) Are the stationary values determined in part(b) correspond to the maximum or minimum of the given function and on what basis?

By determining whether the Hessian matrix is positive definite or not, we can determine whether the stationary values represent the maxima or minima of the given function.

By assuming a vector $\mathbf{v} = \begin{Bmatrix} 1 \\ 2 \end{Bmatrix}$, the following equation is checked to determine it has a positive value with respect to the Hessian matrix as follows:

$$\mathbf{v}^T \mathbf{H} \mathbf{v} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 29 \qquad (5)$$

Hence, the Hessian matrix is positive definite and the stationary value corresponds to the minimum of the given function. Similarly, positive definiteness can be proved from the positive eigenvalues of the Hessian matrix.