Department of Mechanical Engineering, IIT Bombay

ME 782 Design Optimization

Final Exam, November 11, 2024

Part 2 (Paper and pen, open notes) Maximum marks: 25 6:30 pm to 8:30 pm

---

1. Minimize the function, $f\left(x_1, x_2\right) = x_1 x_2$, using the principle of **quadratic approximation**, subject to the constraints: $\frac{6x_1}{x_2} + \frac{x_2}{x_1^2} - 5 = 0$ and $-x_1 - x_2 + 1 \le 0$. Beginning with the initial values of $x_1 = 2$ and $x_2 = 1$, carry out one iteration and determine the next best values. **[6 marks]**

The objective function is given by

$$f\left(x_1, x_2\right) = x_1 x_2 \tag{1}$$

The gradient of the objective function is given by

$$\nabla f(\mathbf{x}) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} \tag{2}$$

The Hessian matrix of the objective function is given by

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{3}$$

The equality constraint is given by

$$h(x_1, x_2) = \frac{6x_1}{x_2} + \frac{x_2}{x_1^2} - 5 = 0 \tag{4}$$

The gradient of the equality constraint is given by

$$\nabla h(\mathbf{x}) = \begin{bmatrix} \dfrac{6}{x_2} - \dfrac{2x_2}{x_1^3} \\ \dfrac{1}{x_1^2} - \dfrac{6x_1}{x_2^2} \end{bmatrix} \tag{5}$$

The inequality constraint is given by

$$g\left(x_1, x_2\right) = -x_1 - x_2 + 1 \leq 0 \tag{6}$$

The gradient of the inequality constraint is given by

$$\nabla g(\mathbf{x}) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \tag{7}$$

From equation (2), the gradient of the objective function evaluated at the starting point (2,1) is given by

$$\nabla f(\mathbf{x}^0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \tag{8}$$

From equation (4), the equality constraint evaluated at the starting point (2,1) is given by

$$h(\mathbf{x}^0) = \frac{6 \times 2}{1} + \frac{1}{2^2} - 5 = 7.25 \tag{9}$$

From equation (5), the gradient of the equality constraint evaluated at the starting point (2,1) is given by

$$\nabla h(\mathbf{x}^0) = \begin{bmatrix} \dfrac{6}{1} - \dfrac{2 \times 1}{2^3} \\ \dfrac{1}{2^2} - \dfrac{6 \times 2}{1^2} \end{bmatrix} = \begin{bmatrix} 5.75 \\ -11.75 \end{bmatrix} \tag{10}$$

The gradient of the inequality constraint is a constant and hence its value at the starting point remains the same as equation (7).

$$\nabla g(\mathbf{x}^0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \tag{11}$$

The small step that has to be taken from the starting point can be defined as

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{12}$$

The minimization of the given problem can now be cast as minimizing

$$\nabla f(\mathbf{x}^{(0)})^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^{(0)})\mathbf{d} \tag{13}$$

$$= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = d_1 + 2d_2 + d_1 d_2$$

Subject to

$$h(\mathbf{x}^{(0)}) + \nabla h(\mathbf{x}^{(0)})^T \mathbf{d} = 7.25 + \begin{bmatrix} 5.75 & -11.75 \end{bmatrix} \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} \tag{14}$$

$$= 5.75 d_1 - 11.75 d_2 + 7.25 = 0$$

And

$$g(\mathbf{x}^{(0)}) + \nabla g(\mathbf{x}^{(0)})^T \mathbf{d} \le 0 \tag{15}$$

$$-2 + \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} = -2 - d_1 - d_2 \le 0$$

From equation (14)

$$d_1 = \frac{11.75 d_2 - 7.25}{5.75} = 2.043 d_2 - 1.26 \tag{16}$$

From equations (13) and (16), the objective function becomes

$$\left(2.043 d_2 - 1.26\right) + 2d_2 + \left(2.043 d_2 - 1.26\right) d_2 \tag{17}$$

$$= \left(2.043\right) d_2^2 + 2.783 d_2$$

Hence, the next best variable is given by

Now equation (17) becomes an objective function of a single variable d_2. Equating the derivative of equation (17) to zero

$$d_2 = -\frac{2.783}{2 \times 2.043} = -0.68 \tag{18}$$

From equations (16) and (18)

$$d_1 = 2.043\left(-0.68\right) \; -1.26 = -2.65 \tag{19}$$

The next best values are given by

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{d} = \begin{Bmatrix} 2 \\ 1 \end{Bmatrix} + \begin{Bmatrix} -2.65 \\ -0.68 \end{Bmatrix} = \begin{Bmatrix} -0.65 \\ 0.32 \end{Bmatrix} \tag{20}$$

2. Consider the unconstrained minimization of the following objective function:

   $f(x_1, x_2) = -(2x_1 + 3x_2 - x_1^3 - 2x_2^2)$

   From the starting point (1,1), carry out **one** iteration using the following methods. **Wherever it is necessary to determine γ (optimum step length), assume it as 0.4. [7 marks]**

   The objective function is given by

$$f(x_1, x_2) = -(2x_1 + 3x_2 - x_1^3 - 2x_2^2) \tag{1}$$

   The starting coordinates are given by

$$\mathbf{x}_0 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \tag{2}$$

   The objective function evaluated at the starting point is given by

$$f(1,1) = -2 \tag{3}$$

   The gradient of the objective function is given by

$$\nabla f(x_1, x_2) = \begin{Bmatrix} 3x_1^2 - 2 \\ 4x_2 - 3 \end{Bmatrix} \tag{4}$$

   The gradient of the objective function evaluated at the starting point is given by

$$\nabla f(1,1) = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \tag{5}$$

   The Hessian matrix of the objective function is given by

$$\mathbf{H}_0 = \begin{bmatrix} 6x_1 & 0 \\ 0 & 4 \end{bmatrix} \tag{6}$$

   The Hessian matrix evaluated at the starting point is given by

$$\mathbf{H}_0 = \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix} \tag{7}$$

(i) Univariate method

For horizontal movement

The objective function evaluated at $\begin{Bmatrix} 1+0.001 \\ 1 \end{Bmatrix}$ is given by

$$f^{+}(1.001,1) = -1.999 \tag{8}$$

The objective function evaluated at $\begin{Bmatrix} 1-0.001 \\ 1 \end{Bmatrix}$ is given by

$$f^{-}(0.999,1) = -2.0010 \tag{9}$$

Therefore, minimization occurs in the negative horizontal direction. Since γ (optimum step length), is given as 0.4, and hence, the next point is given by

$$\mathbf{x}_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} + 0.4 \begin{Bmatrix} -1 \\ 0 \end{Bmatrix} = \begin{Bmatrix} 0.6 \\ 1 \end{Bmatrix} \tag{10}$$

For vertical movement

The objective function evaluated at $\begin{Bmatrix} 1 \\ 1+0.001 \end{Bmatrix}$ is given by

$$f^{+}(1,1.001) = -1.999 \tag{11}$$

The objective function evaluated at $\begin{Bmatrix} 1 \\ 1-0.001 \end{Bmatrix}$ is given by

$$f^{-}(1,0.999) = -2.0010 \tag{12}$$

Therefore, minimization occurs in the negative vertical direction. Since γ (optimum step length), is given as 0.4, the next point is given by

$$\mathbf{x}_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} + 0.4 \begin{Bmatrix} 0 \\ -1 \end{Bmatrix} = \begin{Bmatrix} 1 \\ 0.6 \end{Bmatrix} \tag{13}$$

(Movement in either direction will be considered as one iteration.)

(ii) Steepest descent method

The normalized search direction based on steepest descent is given by

$$\mathbf{s} = -\frac{1}{\sqrt{2}} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \tag{14}$$

Since γ (optimum step length), is given as 0.4, the next point is given by

$$\mathbf{x}_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} - \frac{0.4}{\sqrt{2}} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 0.7172 \\ 0.7172 \end{Bmatrix} \tag{15}$$

(iii) Newton's method

The next point as per Newton's method is given by

$$\mathbf{x}_1 = \mathbf{x}_0 - \mathbf{H}_0^{-1} \nabla f(\mathbf{x}_0) \tag{16}$$

$$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.8333 \\ 0.7500 \end{bmatrix}$$

(iv) Modified Newton's method

The search direction is given by

$$\mathbf{s} = -\mathbf{H}_0^{-1} \nabla f(\mathbf{x}_0) = -\begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}^{-1} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = -\begin{Bmatrix} 0.1667 \\ 0.2510 \end{Bmatrix} \tag{17}$$

Since γ (optimum step length), is given as 0.4, the next point using the normalized equation (17) is given by

$$\mathbf{x}_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} - \frac{0.4}{\sqrt{0.1667^2 + 0.25^2}} \begin{Bmatrix} 0.1667 \\ 0.2510 \end{Bmatrix} = \begin{Bmatrix} 0.7787 \\ 0.6681 \end{Bmatrix} \tag{18}$$

(v) Marquardt's method (assume λ=100)

The search direction is given by

$$\mathbf{s} = -\left[\mathbf{H}_0 + \lambda \mathbf{I}\right]^{-1} \nabla f(\mathbf{x}_0) = -\left\{ \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix} + 100 \begin{bmatrix} 1 & 0 \\ 9 & 1 \end{bmatrix} \right\}^{-1} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = -\begin{Bmatrix} 0.0094 \\ 0.0096 \end{Bmatrix} \tag{19}$$

Since γ (optimum step length), is given as 0.4, the next point is given by

$$\mathbf{x}_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} - \begin{Bmatrix} 0.0094 \\ 0.0096 \end{Bmatrix} = \begin{Bmatrix} 0.9906 \\ 0.9904 \end{Bmatrix} \tag{20}$$

3. Minimize the function, $f\left(x_1, x_2\right) = x_1^2 - x_2$, using **KT conditions**, subject to the constraints:

$x_1 + x_2 - 6 = 0$ and $1 - x_1 \le 0$. **[6 marks]**

Using the objective functions and the constraints, the Lagrangian can be written as

$$L(x_1, x_2, \alpha, \lambda) = x_1^2 - x_2 + \alpha\left(1 - x_1 + s^2\right) + \lambda\left(x_1 + x_2 - 6\right) \tag{1}$$

The gradient of the Lagrangian defined in equation (1) is given by

$$\begin{Bmatrix} \dfrac{\partial L}{\partial x_1} \\ \dfrac{\partial L}{\partial x_2} \\ \dfrac{\partial L}{\partial \alpha} \\ \dfrac{\partial L}{\partial \lambda} \\ \dfrac{\partial L}{\partial s} \end{Bmatrix} = \begin{bmatrix} 2x_1 - \alpha + \lambda \\ \lambda - 1 \\ s^2 - x_1 + 1 \\ x_1 + x_2 - 6 \\ 2\alpha s \end{bmatrix} = 0 \tag{2}$$

Case (1) α=0

Equation (2) now becomes

$$\begin{Bmatrix} \dfrac{\partial L}{\partial x_1} \\ \dfrac{\partial L}{\partial x_2} \\ \dfrac{\partial L}{\partial \alpha} \\ \dfrac{\partial L}{\partial \lambda} \\ \dfrac{\partial L}{\partial s} \end{Bmatrix} = \begin{bmatrix} 2x_1 + \lambda \\ \lambda - 1 \\ s^2 - x_1 + 1 \\ x_1 + x_2 - 6 \\ 0 \end{bmatrix} = 0 \tag{3}$$

Collecting equations containing only x_1, x_2 and λ

$$\begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ \lambda \end{Bmatrix} = \begin{Bmatrix} 0 \\ 1 \\ 6 \end{Bmatrix} \tag{4}$$

Solving equation (4), $x_1 = -0.5;\ x_2 = 6.5;\ \lambda = 1$

Substituting x_1 and x_2 in one of the equations of (2) containing the slack variable

$$s = \pm\sqrt{x_1 - 1} = \pm j\sqrt{1.5} \tag{5}$$

Since the slack variable is a complex number, the above points are not feasible.

Case (2), s=0

Equation (2) now becomes

$$\begin{Bmatrix} \dfrac{\partial L}{\partial x_1} \\ \dfrac{\partial L}{\partial x_2} \\ \dfrac{\partial L}{\partial \alpha} \\ \dfrac{\partial L}{\partial \lambda} \\ \dfrac{\partial L}{\partial s} \end{Bmatrix} = \begin{bmatrix} 2x_1 - \alpha + \lambda \\ \lambda - 1 \\ -x_1 + 1 \\ x_1 + x_2 - 6 \\ 0 \end{bmatrix} = 0 \tag{6}$$

Collecting equations containing only x_1, x_2, α and λ

$$\begin{bmatrix} 2 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ \alpha \\ \lambda \end{Bmatrix} = \begin{Bmatrix} 0 \\ 1 \\ -1 \\ 6 \end{Bmatrix} \tag{7}$$

Solving equation (7) $x_1 = 1;\ x_2 = 5, \alpha = 3, \lambda = 1$. **Therefore, these points are feasible**.

4. Minimize the function, $f\left(x_1,x_2\right)=2x_1^2+3x_2^2+x_1+x_2$, using appropriate penalty functions, subject to the constraints: $5-x_1\le 0$ and $x_1+x_2=5$. **[6 marks]**

$$f\left(x_1,x_2\right)=2x_1^2+3x_2^2+x_1+x_2 \tag{1}$$

The penalty function for the above problem can be written as

$$\phi(x_1,x_2,R)=f(x_1,x_2)=f\left(x_1,x_2\right)=2x_1^2+3x_2^2+x_1+x_2+\frac{1}{R}\left(x_1+x_2-5\right)^2-R\left(\frac{1}{5-x_1}\right) \tag{2}$$

Equating the gradient of equation (2) to zero gives the stationary value of equation (1) and are given by

$$4x_1+1+\frac{2x_1+2x_2-10}{R}-\frac{R}{\left(x_1-5\right)^2}=0 \tag{3a}$$

$$6x_2+1+\frac{2x_1+2x_2-10}{R}=0 \tag{3b}$$

From equation (3) it is clear that the equation cannot be solved through hand calculations. Since the penalty function, with a common controlling parameter for both constraints, is written such that the limiting value of R tends to zero in the final solutions, the third term of equation (3a) can assumed to be much smaller than the other terms; effectively, we are neglecting the inequality constraint for the moment. Therefore, the approximation of equation (3) becomes

$$\begin{aligned}4x_1+1+\frac{2x_1+2x_2-10}{R}&=0\\ 6x_2+1+\frac{2x_1+2x_2-10}{R}&=0\end{aligned} \tag{4}$$

Equation (4) can be simplified as

$$\begin{bmatrix}4R+2 & 2\\ 2 & 6R+2\end{bmatrix}\begin{Bmatrix}x_1\\ x_2\end{Bmatrix}=\begin{Bmatrix}10-R\\ 10-R\end{Bmatrix} \tag{5}$$

Solving the linear non-homogenous equations of (5)

$$x_1 = -\frac{3(R-10)}{2(6R+5)} \qquad (6)$$
$$x_2 = -\frac{(R-10)}{(6R+5)}$$

The limiting values of equation (6) are obtained as the controlling parameter R tends to zero. Thus, the stationary values can be obtained as

$$x_1 = 3 \qquad (7)$$
$$x_2 = 2$$

Since we had neglected the inequality constraint while determining the stationary values, the above stationary values can be checked for the inequality constraint: $5 - x_1 = 5 - 3 = 2 \geq 0$. Therefore, the inequality constraint is not satisfied away from its boundary and hence, the stationary points can be expected to be only on the boundary of the inequality constraint. Hence, the stationary points can as well be solved using the equations $5 - x_1 = 0$ and $x_1 + x_2 = 5$, which result in $x_1^* = 5$ and $x_2^* = 0$.

Another approach could be to use the given equality constraint to replace $x_1 + x_2$ with 5 in the objective function and then solve the resulting problem only with the inequality constraint that would result in the same result as above. However, converting the inequality constraint directly into an equality constraint without any basis and solving the resulting problem will not be acceptable.