CONSTRAINED AND UNCONSTRAINED OPTIMIZATION

INTRODUCTION

In the previous lecture we have learnt how to determine the convexity and concavity of functions of single and two variables. For functions of single and two variables we have also learnt about determining stationary points and examining higher derivatives to check for convexity and concavity, and tests were recommended to evaluate stationary points as local minima, local maxima or points of inflection.

In this lecture, functions of multiple variables, which are more difficult to be analyzed owing to the difficulty in graphical representation and tedious calculations involved in mathematical analysis, will be studied for unconstrained optimization. This is done with the aid of the gradient vector and the Hessian matrix. Examples are discussed to show the implementation of the technique.

In this lecture we will also learn the optimization of functions of multiple variables subjected to equality constraints using the method of constrained variation.

UNCONSTRAINED OPTIMIZATION

If a convex function is to be minimized, the stationary point is the global minimum and analysis is relatively straightforward as discussed earlier. A similar situation exists for maximizing a concave variable function. The necessary and sufficient conditions for the optimization of unconstrained function of several variables are given below.

Necessary condition

In case of multivariable functions a necessary condition for a stationary point of the function $f(\mathbf{X})$ is that each partial derivative is equal to zero. In other words, each element of the gradient vector defined below must be equal to zero. i.e. the gradient vector of $f(\mathbf{X})$, $\Delta_x f$ at $\mathbf{X}=\mathbf{X}^*$, defined as follows, must be equal to zero:

$$\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(X^*) \\ \frac{\partial f}{\partial x_2}(X^*) \\ \vdots \\ \vdots \\ \frac{\partial f}{\partial dx_n}(X^*) \end{bmatrix} = 0 \tag{1}$$

The proof given for the theorem on necessary condition for single variable optimization can be easily extended to prove the present condition.

Sufficient condition

For a stationary point **X*** to be an extreme point, the matrix of second partial derivatives (Hessian matrix) of *f*(**X**) evaluated at **X*** must be:

(i) positive definite when **X*** is a point of relative minimum, and

(ii) negative definite when **X*** is a relative maximum point.

Proof *(Formulation of the Hessian matrix)*

The Taylor's theorem with reminder after two terms gives us

$$f(\mathrm{X}^*+h) = f(\mathrm{X}^*) + \sum_{i=1}^{n} h_i \frac{df}{dx_i}(\mathrm{X}^*) + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathrm{X}=\mathrm{X}^*+\theta h} \tag{2}$$

$$0<\theta<1$$

Since **X*** is a stationary point, the necessary condition gives

$$\frac{df}{dx_i} = 0, \qquad i = 1,2,\ldots,n \tag{3}$$

Thus

$$f(\mathrm{X}^*+h) - f(\mathrm{X}^*) = \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathrm{X}=\mathrm{X}^*+\theta h} \tag{4}$$

For a minimization problem the left hand side of the above expression must be positive. Since the second partial derivative is continuous in the neighborhood of **X*** the sign of $\partial^2 f/\partial x_i \partial x_j \big| \mathrm{X} = \mathrm{X}^* + \theta h$ is the same as the sign of $\partial^2 f/\partial x_i \partial x_j \big| \mathrm{X} = \mathrm{X}^*$. And hence $f(\mathrm{X}^*+h) - f(\mathrm{X}^*)$ will be a relative minimum, if

$$f(\mathrm{X}^*+h) - f(\mathrm{X}^*) = \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathrm{X}=\mathrm{X}^*} \tag{5}$$

is positive. This can also be written in the matrix form as

$$Q = \mathbf{h}^T \mathbf{H} \mathbf{h}\Big|_{\mathrm{X}=\mathrm{X}^*} \tag{6}$$

where

$$\mathbf{H}\Big|_{\mathrm{X}=\mathrm{X}^*} = \left[\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathrm{X}=\mathrm{X}^*} \right] \tag{7}$$

is the matrix of second partial derivatives and is called the *Hessian matrix* of *f*(**X**).

Q will be positive for all **h** if and only if **H** is positive definite at **X=X*.** i.e., the sufficient condition for **X*** to be a relative minimum is that the Hessian matrix evaluated at the same point is positive definite, which completes the proof for the minimization case. In a similar manner, it can be proved that the Hessian matrix will be negative definite if **X*** is a point of relative maximum.

A matrix **A** will be positive definite if all its eigen values are positive. i.e., all values of λ that satisfy the equation

$$\left|\mathbf{A}-\lambda\mathbf{I}\right|=0 \tag{8}$$

should be positive. Similarly, the matrix **A** will be negative definite if its eigen values are negative. When some eigen values are positive and some are negative then matrix A is neither positive definite or negative definite.

When all eigen values are negative for all possible values of **X**, then **X*** is a global maximum, and when all eigen values are positive for all possible values of **X**, then **X*** is a global minimum.

If some of the eigen values of the Hessian at **X*** are positive and some are negative, or if some are zero, the stationary point, **X***, is neither a local maximum nor a local minimum.

Example

Analyze the function $f(x)=-x_1^2-x_2^2-x_3^2+2x_1x_2+2x_1x_3+4x_1-5x_3+2$ and classify the stationary points as maxima, minima and points of inflection.

Solution

$$\Delta_x f=\begin{bmatrix}\frac{\partial f}{\partial x_1}(\mathrm{X}^*)\\ \frac{\partial f}{\partial x_2}(\mathrm{X}^*)\\ \frac{\partial f}{\partial x_3}(\mathrm{X}^*)\end{bmatrix}=\begin{bmatrix}-2x_1+2x_2+2x_3+4\\ -2x_2+2x_1\\ -2x_3+2x_1-5\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

Solving these simultaneous equations we get $\mathbf{X}^*$=[1/2, 1/2, -2]

$$\frac{\partial^2 f}{\partial x_1{}^2}=-2;\frac{\partial^2 f}{\partial x_2{}^2}=-2;\frac{\partial^2 f}{\partial x_3{}^2}=-2$$

$$\frac{\partial^2 f}{\partial x_1\partial x_2}=\frac{\partial^2 f}{\partial x_2\partial x_1}=2$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_2} = 0$$
$$\frac{\partial^2 f}{\partial x_3 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_3} = 2$$

Hessian of $f(\mathbf{X})$ is

$$\mathbf{H} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]$$

$$\mathbf{H} = \begin{bmatrix} -2 & 2 & 2 \\ 2 & -2 & 0 \\ 2 & 0 & -2 \end{bmatrix}$$

$$|\lambda \mathbf{I} - \mathbf{H}| = \begin{vmatrix} \lambda+2 & -2 & -2 \\ -2 & \lambda+2 & 0 \\ -2 & 0 & \lambda+2 \end{vmatrix} = 0$$

or $\quad (\lambda+2)(\lambda+2)(\lambda+2) - 2(\lambda+2)(2) + 2(2)(\lambda+2) = 0$

$$(\lambda+2)[\lambda^2 + 4\lambda + 4 - 4 + 4] = 0$$

$$(\lambda+2)^3 = 0$$

or $\lambda_1 = \lambda_2 = \lambda_3 = -2$

Since all eigen values are negative the function attains a maximum at the point

$\mathbf{X}^*$=[1/2, 1/2, -2]

CONSTRAINTED OPTIMIZATION: EQUALITY CONSTRAINTS

A function of multiple variables, $f(x)$, is to be optimized subject to one or more equality constraints of many variables. These equality constraints, $g_j(x)$, may or may not be linear. The problem statement is as follows:

Maximize (or minimize) $f(\mathbf{X})$, subject to $g_j(\mathbf{X}) = 0$, j = 1, 2, … , m

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \qquad (9)$$

with the condition that $m \le n$; or else if $m > n$ then the problem becomes an over defined one and there will be no solution. Of the many available methods, the method of constrained variation is discussed here. Another method using Lagrange multipliers will be discussed in the next lecture.

Solution by method of Constrained Variation

For the optimization problem defined above, let us consider a specific case with n = 2 and m=1 before we proceed to find the necessary and sufficient conditions for a general problem using Lagrange multipliers. The problem statement is as follows:

Minimize $f(x_1,x_2)$, subject to $g(x_1,x_2) = 0$

For $f(x_1,x_2)$ to have a minimum at a point $\mathbf{X}^* = [x_1^*, x_2^*]$, a necessary condition is that the total derivative of $f(x_1,x_2)$ must be zero at $[x_1^*, x_2^*]$.

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0 \qquad (10)$$

Since $g(x_1^*, x_2^*) = 0$ at the minimum point, variations dx_1 and dx_2 about the point $[x_1^*, x_2^*]$ must be *admissible variations*, i.e. the point lies on the constraint:

$$g(x_1^* + dx_1 , x_2^* + dx_2) = 0 \qquad (11)$$

assuming dx_1 and dx_2 are small the Taylor's series expansion of this gives us

$$g(x_1^* + dx_1 \ , \ x_2^* + dx_2) = g(x_1^*, \ x_2^*) + \frac{\partial g}{\partial x_1}(x_1^*, x_2^*) \ dx_1 + \frac{\partial g}{\partial x_2}(x_1^*, x_2^*) \ dx_2 = 0 \qquad (12)$$

or

$$dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0 \text{ at } [x_1^*, x_2^*] \qquad (13)$$

which is the condition that must be satisfied for all *admissible variations*.

Assuming $\partial g / \partial x_2 \neq 0$ (13) can be rewritten as

$$dx_2 = -\frac{\partial g / \partial x_1}{\partial g / \partial x_2}(x_1^*, x_2^*) dx_1 \qquad (14)$$

which indicates that once variation along x_1 ($d\,x_1$) is chosen arbitrarily, the variation along x_2 ($d\,x_2$) is decided automatically to satisfy the condition for the *admissible variation.*

Substituting equation (14) in (10) we have:

$$df = \left(\frac{\partial f}{\partial x_1} - \frac{\partial g / \partial x_1}{\partial g / \partial x_2} \frac{\partial f}{\partial x_2} \right) \Bigg|_{(x_1^*, \, x_2^*)} dx_1 = 0 \qquad (15)$$

The equation on the left hand side is called the constrained variation of f. Equation (14) has to be satisfied for all dx_1, hence we have

$$\left(\frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2}-\frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1}\right)\Bigg|_{(x_1^*,\, x_2^*)}=0 \tag{16}$$

This gives us the necessary condition to have [x_1*, *x2**] as an extreme point (maximum or minimum)

WATER DISTRIBUTION SYSTEMS

INTRODUCTION

The main purpose of water distribution network is to supply water to the users according to their demand with adequate pressure. Water distribution systems are composed of three major components: pumping stations, storage tanks and distribution piping. These systems are designed according to the loading conditions i.e., pressure and demand at nodal points. The loading conditions may include fire demands, peak daily demands or critical demands when the pipes are broken. A reliable design should consider all the loading conditions including the critical conditions.

COMPONENTS OF WATER DISTRIBUTION SYSTEMS

Various components of water distribution systems are:

(i) Pipes: These are the principal elements in the system. The flow or velocity is usually described using Hazen – Williams equation

$$V = 1.318 C_{HW} R^{0.63} S_f^{0.54} \tag{1}$$

where V is the average flow velocity. C_{HW} is the Hazen – Williams roughness coefficient, R is the hydraulic radius and S_f is the slope.

In terms of headloss h_L, the above equation can be expressed as,

$$h_L = \frac{KLQ^{1.852}}{C_{HW}^{1.852} D^{4.87}} = K_p Q^{1.852} \tag{2}$$

where L is the length of the pipe, D is the diameter and Q is the flow rate.

Headloss can also be determined using Darcy – Weisbach equation as

$$h_L = f\frac{L}{D}\frac{V^2}{2g} = \frac{8fL}{\pi^2 g D^2} Q^2 = K_p Q^2 \tag{3}$$

where f is the friction factor (determined from Moody's diagram) and g is the acceleration due to gravity.

(ii) Node: Junction nodes are connections of pipes to transfer the water. The diameter of pipe is changed at these nodes. Fixed grade nodes is where pressure and elevation are fixed i.e., reservoirs, tanks etc.

(iii) Valves: These are used to vary the head loss or to control the flow.

(iv) Tanks: It stores water and acts as a buffer by storing water at low demands and releasing at high demands.

(v) Pumps: Used to increase the energy

SIMULATION OF NETWORK

The flow distribution through a network should satisfy the conservation of mass and conservation of energy. Consider the network structure in Figure 1 with 6 pipes and 5 nodes.

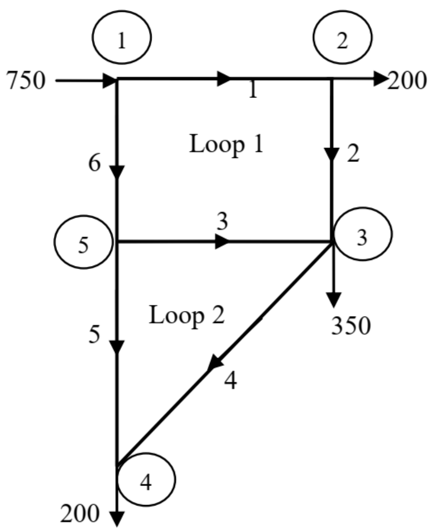


Fig. 1

Conservation of mass: Flow at each junction nodes must be conserved

$$\sum Q_{in} - \sum Q_{out} = Q_{ext} \tag{4}$$

where Q_{in} and Q_{out} are the flows in and out of the node respectively and Q_{ext} is the external supply or demand.

Conservation of energy: For each loop, energy must be conserved i.e., sum of head losses should be zero.

$$\sum h_{L_{i,j}} - \sum H_{pump} = 0 \tag{5}$$

where $h_{Li,j}$ is the head loss in the pipe connecting nodes i and j and H_{pump} is the energy added by the pump (if any). $h_{Li,j}$ can be determined using either eqn. 2 or 3.

Energy must be conserved between the fixed grade nodes which are points of known head (elevation plus pressure head).

$$\Delta E_F = \sum h_{L_{i,j}} - \sum H_{pump} \tag{6}$$

If the number of pipes in the network is N_L, number of junction nodes is N_J and number of fixed grade nodes is N_F, then total number of equations will be $N_L + N_J + (N_F - 1)$.

The set of equations obtained can be solved by any iterative techniques like Hardy-Cross method, linear theory method and Newton - Raphson method.

Hardy-Cross method

In this method, the loop equation (eqn. 5) in terms of flow is used. The loop equations are transformed into so called ΔQ equations in the form

$$\sum_{i,j} K_{p,i,j}\left(Q_{i,j} + \Delta Q_{i,j}\right)^n = 0 \tag{7}$$

Here head loss is determined from equation 2 or 3.

Eqn. 7 is rewritten to account the direction of flow as

$$\sum h_{L,i,j} = \sum_{i,j} K_{p,i,j}\left(Q_{i,j} + \Delta Q_{i,j}\right)^n sign\left(Q_{i,j} + \Delta Q_{i,j}\right) = 0 \tag{8}$$

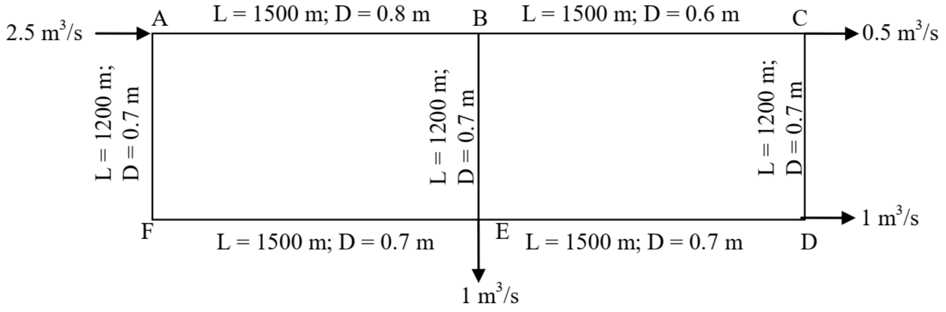
This equation can be finally expressed as

$$\Delta Q_p = -\frac{\sum_{i,j} K_{p,i,j} Q_{i,j}{}^n}{\sum_{i,j} \left| n K_{p,i,j} Q_{i,j}{}^{n-1} \right|} \tag{9}$$

First a flow distribution is assumed across the network. Then the correction ΔQ as given in eqn. 9 is applied in a particular loop p. The numerator in eqn. 9 is the algebraic sum of headloss in loop p taking care of the sign of the flow. If clockwise flows are taken positive, then the corresponding headlosses are positive. The same is applicable while applying correction also i.e., ΔQ_p is added to flows in the clockwise direction and subtracted from flows in counterclockwise direction.

Example:

Consider the pipe network shown below. The friction factor = 0.2. Determine the flow rate in each pipe.



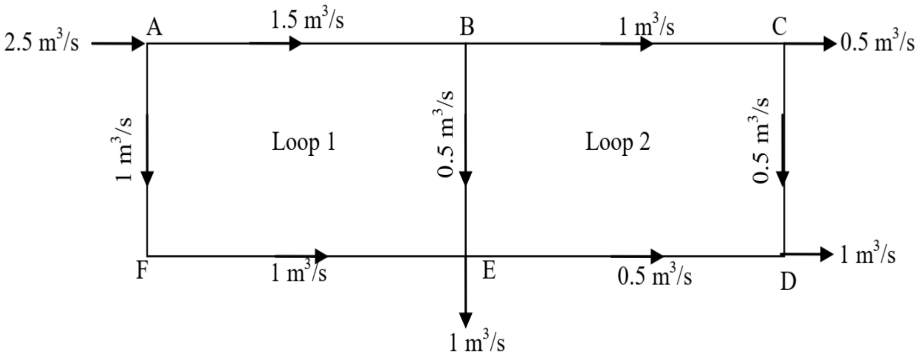
Solution:

Step 1: Determine the K values in eqn. 3

$$h_L = \frac{8fL}{\pi^2 gD^2}Q^2 = KQ^2 \text{ where } K = \frac{8fL}{\pi^2 gD^2}$$

$K_{AB} = 3.88$; $K_{BC} = 6.89$; $K_{FE} = K_{ED} = 5.06$; $K_{AF} = K_{BE} = K_{CD} = 4.05$

Step 2; Assume initial flows in each pipe as shown below



Step 3: Consider loop 1 and calculate ΔQ according to eqn. 9; n = 2. Consider anticlockwise flows positive.

$$\Delta Q_1 = -\frac{1}{2}\frac{K_{AF}Q_{AF}^2 + K_{FE}Q_{FE}^2 - K_{BE}Q_{BE}^2 - K_{AB}Q_{AB}^2}{K_{AF}Q_{AF} + K_{FE}Q_{FE} - K_{BE}Q_{BE} - K_{AB}Q_{AB}}$$
$$= -\frac{1}{2}\frac{4.05\times 1^2 + 5.06\times 1^2 - 4.05\times 0.5^2 - 3.88\times 1.5^2}{4.05\times 1 + 5.06\times 1 + 4.05\times 0.5 + 3.88\times 1.5}$$
$$= 0.0187$$

Step 4: Consider loop 2 and calculate ΔQ

$$\Delta Q_2 = \frac{1}{2}\frac{4.05\times 0.5^2 + 5.06\times 0.5^2 - 4.05\times 0.5^2 - 6.89\times 1^2}{4.05\times 0.5 + 5.06\times 0.5 + 4.05\times 0.5 + 6.89\times 1}$$
$$= 0.2088$$

Step 5: Flows for next iteration

Q_{AF} = 1 + 0.0187 = 1.0187

Q_{FE} = 1 + 0.0187 = 1.0187

Q_{BE} = 0.5 – 0.0187 + 0.2088 = 0.6901

Q_{AB} = 1.5 - 0.0187 = 1.4813

Q_{ED} = 0.5 + 0.2088 = 0.7088

Q_{CD} = 0.5 – 0.2088 = 0.2912

Q_{BC} = 1 – 0.2088 = 0.7912

Repeat steps 2 to 5 with new flows till ΔQ is insignificant.

BIBLIOGRAPHY / FURTHER READING:

1. Bellman, R., *Dynamic Programming*, Princeton University Press, Princeton, N.J, 1957.
2. Hillier F.S. and G.J. Lieberman, *Operations Research*, CBS Publishers & Distributors, New Delhi, 1987.
3. Loucks, D.P., J.R. Stedinger, and D.A. Haith, *Water Resources Systems Planning and Analysis*, Prentice-Hall, N.J., 1981.
4. Rao S.S., *Engineering Optimization – Theory and Practice*, Fourth Edition, John Wiley and Sons, 2009.
5. Taha H.A., *Operations Research – An Introduction*, 8^{th} edition, Pearson Education India, 2008.
6. Vedula S., and P.P. Mujumdar, *Water Resources Systems: Modelling Techniques and Analysis*, Tata McGraw Hill, New Delhi, 2005.