IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

NPTEL

Nonlinear Programming with equality constraint

Dr. Debjani Chakraborty
Department of Mathematics
IIT Kharagpur

Constrained optimization

- We solved NLP problems that only had objective functions, with no constraints.

- Now we will look at methods on how to solve problems that include constraints.

Multivariable functions with both equality and inequality constraints.

NLP with equality constraint

Let us consider the nonlinear programming problem with equality constraint as:

$$\text{Minimize } f(X)$$

Subject to,

$$g_j(X) = b_j, \quad j = 1, 2 \cdots m$$

$$\text{where } X = \{x_1, x_2 \cdots x_n\}^T$$

Graphical Illustration

Let us take

$$\text{Minimize } 2x_1 + 3x_2$$

$$\text{Subject to } x_1^2 + x_2^2 = 1.$$

- The feasible region is a circle with a radius of one. The possible objective function curves are lines with a slope of - (2/3).
- The minimum will be the point where the lowest line still touches the circle. Similarly, if we consider the maximization of the same objective function the maximum point will be the point where the upper line touches the circle.

$$\text{Minimize } 2x_1 + 3x_2$$
$$\text{Subject to, } x_1^2 + x_2^2 = 1.$$

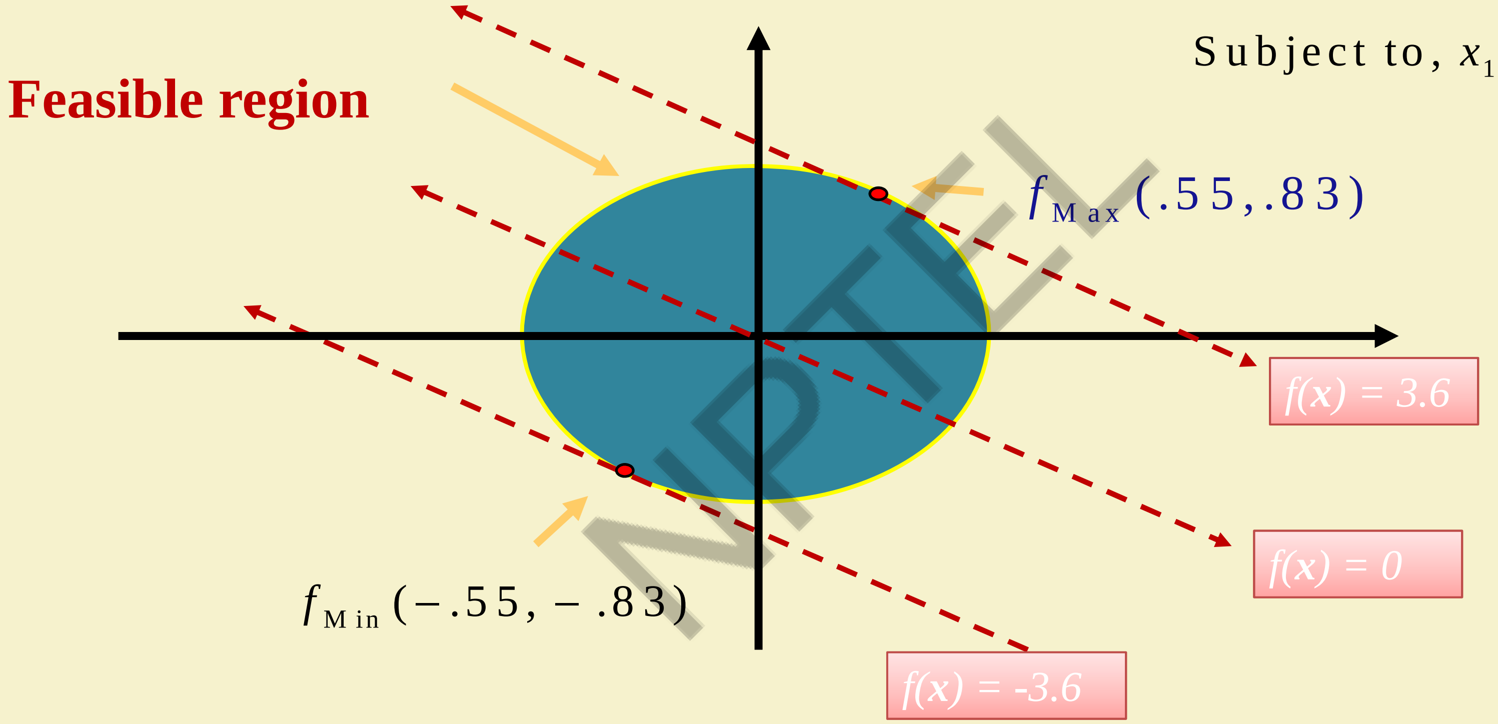
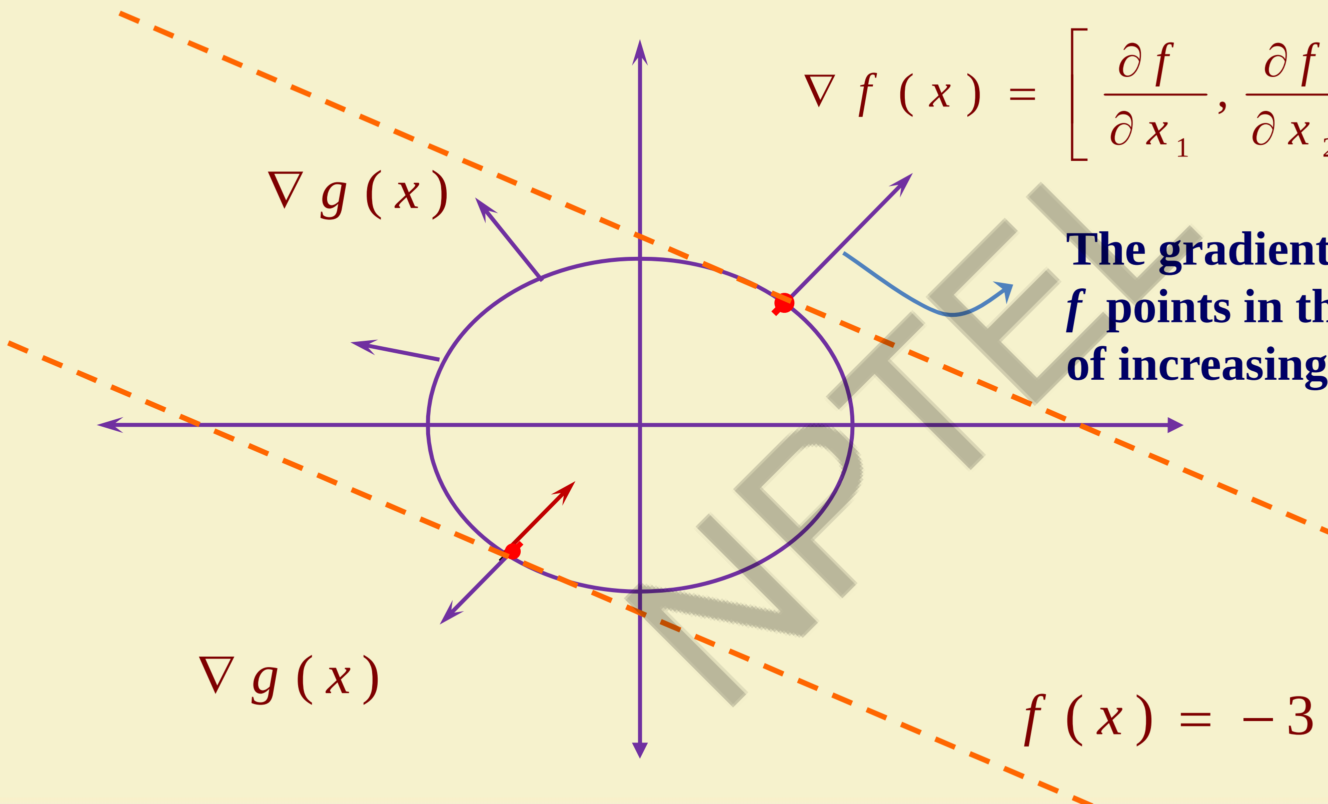


Illustration on gradient

- Objective function lines are straight parallel lines, the gradient of objective function is a straight line pointing toward the direction of increasing objective function, which is to the upper right.

- The gradient of constraint will be pointing out from the circle and so its direction will depend on the point at which the gradient is evaluated.



$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right]$$

The gradient of *f* points in the direction of increasing *f*.

$$f(x) = -3.605$$

Conclusions

- At the minimum point, $\nabla f(\mathbf{x})$ and $\nabla g(x)$ are parallel and opposite to each other. But the magnitude of the gradient vectors are generally not equal.

- At the maximum point, $\nabla f(\mathbf{x})$ and $\nabla g(x)$ are parallel in the same direction
  - *Otherwise, we could improve the objective function by changing position*

Classical optimization technique

Nonlinear functions involved must be at least twice differentiable.

- *Method of Direct Substitution* (*transforming the constrained optimization to an unconstrained one using substitution*)
- *Lagrange's Multiplier Technique* (*transforming the constrained optimization to an unconstrained one after introducing Lagrange multiplier*)

Method of Direct Substitution

Let us consider a problem with n decision variables and m equality constraints. Assuming condition $m < n$.

$$\text{Minimize } f(X)$$

$$\text{Subject to}$$

$$g_j(X) = b_j, \quad j = 1, 2 \cdots m$$

$$\text{where } X = \{x_1, x_2 \cdots x_n\}^T$$

Method of Direct Substitution

Step 1: Using these *m* relations, the original objective function involving *n* variables can newly be represented with *(n - m)* variables.

Step 2: Thus the original optimization model takes its equivalent form which is an unconstrained problem with *m* variables

Find the minimum value of $x^2 + y^2 + z^2$;
where $x + y + 2z = 12$.

Here the objective function is $f(X) = x^2 + y^2 + z^2$

and from the given constraint condition, we get $z = \frac{1}{2}(12 - x - y)$

Substituting the value of z in f $f(X) = x^2 + y^2 + \frac{1}{4}(12 - x - y)^2$

Now this is unconstrained optimization problem with two variables…….

The necessary conditions give

$$\frac{\partial f}{\partial x} = 0 \qquad \Rightarrow \quad 2x + \tfrac{2}{4}(12 - x - y).(-1) = 0$$

$$\frac{\partial f}{\partial y} = 0 \qquad \Rightarrow \quad 2y + \tfrac{2}{4}(12 - x - y).(-1) = 0$$

Thus the possible extreme points are:

$$x^* = 2, \quad y^* = 2, \quad z^* = 4.$$

To check the sufficient condition for we need to construct the Hessian matrix H.

Hessian Matrix:

$$H = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} \dfrac{5}{2} & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{5}{2} \end{pmatrix}$$

$$\frac{\partial f}{\partial x} = 2x + \tfrac{2}{4}(12 - x - y).(-1)$$

$$\frac{\partial f}{\partial y} = 2y + \tfrac{2}{4}(12 - x - y).(-1)$$

The matrix H is positive definite as $H_2 = 6$ and $H_1 = \frac{5}{2}$

Hence, (2,2,4) is the point of minima and the minimum value of the objective function is 24.

Method of Lagrangian Multipliers

Let us consider the nonlinear programming problem with two variables and one equality constraint.

$$\text{Minimize } f(x_1, x_2)$$

$$\text{Subject to, } g(x_1, x_2) = b, \quad x_1, x_2 \geq 0.$$

This is done by converting a constrained problem to an equivalent unconstrained problem with the help of certain unspecified parameters known as *Lagrange multipliers.*

Method of Lagrangian Multipliers

The following classical problem

$$Minimize \;\; f(x_1, x_2) \quad \text{Subject to} \;\; g(x_1, x_2) = b$$

can be converted to

$$Minimize \;\; L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda \{ g(x_1, x_2) - b \}$$

Multi-dimensional derivatives

- As in the one-dimensional case, we can recognize convexity and concavity by looking at the objective function's derivatives
- The *gradient vector* is the vector of first-order partial derivatives:

$$\nabla f(x) = \begin{pmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

Recognizing multi-dimensional convex and concave functions

- The *Hessian matrix* is the matrix of second-order partial derivatives:

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

Unconstrained optimization

- A point x where $\nabla f(x) = 0$ is called a *stationary point* of f
  - These are called *first-order conditions* for optimality
- Let x^* be a stationary point, i.e., $\nabla f(x^*) = 0$
  - If $H(x^*)$ is positive definite then x^* is a *local minimum*
  - If $H(x^*)$ is negative definite then x^* is a *local maximum*
  - If H(x*) is neither negative definite nor positive definite,
    - If $\det H(x^*) = 0$ then x^* is a *local minimum*, *local maximum*, or *saddle point*
    - If $\det H(x^*) \neq 0$ then x^* is not a optimum

Method of Lagrangian Multipliers

The following classical problem

$$Minimize\ \ f(x_1, x_2) \quad \text{Subject to} \ \ g(x_1, x_2) = b$$

can be converted to

$$Minimize\ \ L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda \{g(x_1, x_2) - b\}$$

Method of Lagrangian Multipliers

$$\frac{\partial L}{\partial x_1} = 0 \Rightarrow \frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} = 0$$

$$\frac{\partial L}{\partial x_2} = 0 \Rightarrow \frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} = 0$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow g(x_1, x_2) = 0.$$

Let us consider an example………………

An Example

$$\text{Minimize } f(x,y) = 5/xy^2$$

$$\text{Subject to } g(x,y) = x^2 + y^2 - 1 = 0.$$

The Lagrange function is

$$L(x,y,\lambda) = 5/xy^2 + \lambda(x^2 + y^2 - 1)$$

Necessary conditions will be ……….

Necessary conditions are:

$$\frac{\partial L}{\partial x} = -5x^{-2}y^{-2} + 2x\lambda = 0$$

$$\frac{\partial L}{\partial y} = -10x^{-1}y^{-3} + 2y\lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - 1 = 0$$

Thus we get from above $2\lambda = \frac{5}{x^3 y^2} = \frac{10}{x y^4}$

$$\frac{5}{x^3 y^2} = \frac{10}{x y^4} \quad \text{gives } x^* = \frac{1}{\sqrt{2}} y^* \quad \text{for optimal } (x^*, y^*)$$

Here the extreme point is

$$x^* = \frac{1}{\sqrt{3}} \quad \text{and} \quad y^* = \frac{\sqrt{2}}{\sqrt{3}}$$

Sufficient conditions need to be verified …………

Lagrange method for multidimensional cases

Let us now develop the Lagrangian function for problem with n independent variables and m constraints ($m < n$) which is defined as follows:

$$\text{Minimize} \quad f(x_1, x_2 \cdots x_n)$$

$$\text{Subject to} \quad g_j(x_1, x_2 \cdots x_n) = b_j \qquad j = 1, 2 \cdots m$$

Same reasoning may be applied and can be converted to

$$\text{Minimize} \quad L(x_1, x_2 \cdots x_n, \lambda_1, \lambda_2 \cdots \lambda_m)$$

$$= f(x_1, x_2 \cdots x_n) + \sum_j \lambda_j \left\{ g_j(x_1, x_2 \cdots x_n) - b_j \right\}$$

Lagrange method for multidimensional cases

Necessary condition for optimum The same reasoning may be applied. Take derivatives of $L(x_1, x_2 \cdots x_n, \lambda_1, \lambda_2 \cdots \lambda_m)$ with respect to x_i and λ_j set them equal to zero.

So, we can treat the Lagrangian as an unconstrained optimization problem with variables $x_1, x_2 \cdots x_n$ and $\lambda_1, \lambda_2 \cdots \lambda_m$

we can solve it by solving the equations

Thank You!!

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

NPTEL

Nonlinear Programming with equality constraints II

Dr. Debjani Chakraborty
Department of Mathematics
IIT Kharagpur

NLP with equality constraint

Let us consider the nonlinear programming problem with equality constraint as:

$$\text{Minimize } f(X)$$

Subject to,

$$g_j(X) = b_j, \quad j = 1, 2 \cdots m$$

$$\text{where } X = \{x_1, x_2 \cdots x_n\}^T$$

Lagrange method for multidimensional cases

Let us now develop the Lagrangian function for problem with *n* independent variables and *m* constraints ($m < n$) which is defined as follows:

$$\text{Minimize} \quad f(x_1, x_2 \cdots x_n)$$

$$\text{Subject to} \quad g_j(x_1, x_2 \cdots x_n) = b_j \qquad j = 1, 2 \cdots m$$

Same reasoning may be applied and can be converted to

$$\text{Minimize} \quad L(x_1, x_2 \cdots x_n, \lambda_1, \lambda_2 \cdots \lambda_m)$$

$$= f(x_1, x_2 \cdots x_n) + \sum_j \lambda_j \left\{ g_j(x_1, x_2 \cdots x_n) - b_j \right\}$$

Lagrange method for multidimensional cases

Necessary condition for optimum The same reasoning may be applied. Take derivatives of $L(x_1, x_2 \cdots x_n, \lambda_1, \lambda_2 \cdots \lambda_m)$ with respect to x_i and λ_j set them equal to zero.

So, we can treat the Lagrangian as an unconstrained optimization problem with variables $x_1, x_2 \cdots x_n$ and $\lambda_1, \lambda_2 \cdots \lambda_m$

we can solve it by solving the equations

$$\frac{\partial L}{\partial x_1} = 0, \frac{\partial L}{\partial x_2} = 0, \cdots \frac{\partial L}{\partial x_n} = 0$$

&

$$\frac{\partial L}{\partial \lambda_1} = 0, \frac{\partial L}{\partial \lambda_2} = 0, \cdots \frac{\partial L}{\partial \lambda_m} = 0$$

Note: If there are n variables (i.e., x_1, x_2 ... x_n) and m constraints then you will get $m+n$ equations with $m+n$ unknowns (i.e., n variables x_i and m Lagrangian multiplier λ_j)

Sufficient conditions with multiple equality constraints

A sufficient condition for $f(X)$ to have a relative minimum at X^* is that the quadratic Q, defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\partial^2 L}{\partial x_i \partial x_j} dx_i dx_j$$

must be positive for all admissible choices of $dx_i dx_j$ and if Q negative for all admissible choices of $dx_i dx_j$ then X^* will be relative maximum.

Sufficient conditions with multiple equality constraints

In other words, it can be said, the *Hessian matrix* $\nabla^2 L$ must be positive definite for relative minimum and negative definite for relative maximum.

$$\nabla^2 L = \begin{pmatrix} M & V \\ V^T & 0 \end{pmatrix}_{(m+n)\times(m+n)}$$

$$\text{where, } V = \left(\frac{\partial g_j}{\partial x_i}\right)_{n\times m}$$

$$\text{and} \quad M = \left(\frac{\partial^2 L}{\partial x_i \partial x_j}\right)_{n\times m}$$

Sufficient conditions with multiple equality constraints

For checking definiteness of $\nabla^2 L$

$\underline{Positive\ definite} \Rightarrow$ principal minor determinants of A are greater than zero.

$\underline{Negtive\ definite} \Rightarrow$ principal minor determinants of A are in alternate sign starting with negative.

OR we need to check the eigen values of *Hessian matrix*

Sufficient conditions with multiple equality constraints

we need to check the roots of the following polynomial

$$\begin{pmatrix} M - z & V \\ V^T & 0 \end{pmatrix} = 0$$

For the above equations, the roots must be positive for relative minimum and the roots are negative for relative maximum. And if some of the roots are positive, while the others are negative, then is not an extreme point

Interpretation of Lagrange Multiplier

To find the physical meaning of Lagrange multiplier let us consider the following optimization problem involving only a single equality constraint:

Minimize $f(X)$ subject to $g(X) = b$ where b is a constant

The Lagrange function is $L = f(X) + \lambda(g(X) - b)$

The necessary conditions are:

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} = 0$$

$$g(X) = b$$

We are trying to find the effect of small relaxation or tightening the constraint on optimal objective functional values

Which implies we need to find the effect of a small change of *b* in optimal *f*

$$\left.\begin{aligned} &\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} = 0 \\ &g(X) = b \end{aligned}\right\} \Rightarrow \left.\begin{aligned} &\frac{\partial f}{\partial x_i} dx_i = -\lambda \frac{\partial g}{\partial x_i} dx_i \\ &dg = db \end{aligned}\right\} \Rightarrow \left.\begin{aligned} &\sum \frac{\partial f}{\partial x_i} dx_i = -\lambda \sum \frac{\partial g}{\partial x_i} dx_i \\ &dg = db \end{aligned}\right\}$$

$$\Rightarrow \left.\begin{aligned} df &= -\lambda\, dg \\ dg &= db \end{aligned}\right\} \qquad \boxed{\Rightarrow\ df = -\lambda\, db}$$

$$\boxed{\Rightarrow df^{*} = -\lambda^{*} db}$$

Depending on the value of λ^* (positive, negative or zero) the following physical meaning can be attributed to λ^*

Case 1: If λ^* > 0, then with one unit increase in positive *b* value, the objective function value decreases. Thus for minimization problem, optimal value of objective function is improving.

$$\boxed{\Rightarrow df^* = -\lambda^* db}$$

Depending on the value of λ^* (positive, negative or zero) the following physical meaning can be attributed to λ^*

Case 2: If λ^* < 0, then with one unit increase in positive *b* value, the objective function value increases. Thus for maximization problem, optimal value of objective function is improving.

$$\boxed{\Rightarrow df^* = -\lambda^* db}$$

Depending on the value of λ^* (positive, negative or zero) the following physical meaning can be attributed to λ^*

Case 3: If λ^* = 0, then incremental change in *b* value does not affect the optimal value of the objective function.

Solve the following nonlinear programming problem :

$$\text{Minimize } 2x_1^2 - 24x_1 + 2x_2^2 - 8x_2 + 2x_3^2 - 12x_3 + 200$$
$$\text{Subject to, } x_1 + x_2 + x_3 = 1.$$

The Lagrangian function can be formulated as follows:

$$L = 2x_1^2 - 24x_1 + 2x_2^2 - 8x_2 + 2x_3^2 - 12x_3 + 200 + \lambda\left(x_1 + x_2 + x_3 - 1\right)$$

The necessary conditions are

$$\frac{\partial L}{\partial x_1} = 4x_1 - 24 + \lambda = 0 \qquad \frac{\partial L}{\partial x_3} = 4x_3 - 12 + \lambda = 0$$
$$\frac{\partial L}{\partial x_2} = 4x_2 - 8 + \lambda = 0 \qquad \frac{\partial L}{\partial \lambda} = x_1 + x_2 + x_3 - 1 = 0$$

By solving the above simultaneous equations we get the stationary point

$$(x_1, x_2, x_3) = \left(\tfrac{8}{3}, -\tfrac{1}{3}, -\tfrac{4}{3}\right), \quad \lambda = \tfrac{40}{3}.$$

Due to sufficient conditions the stationary point $\left(\frac{8}{3}, -\frac{1}{3}, -\frac{4}{3}\right)$ is minimum as the following *Hessian matrix* is positive definite

$$\left(\begin{array}{ccc|c}
\frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial x_3} & \frac{\partial g_1}{\partial x_1} \\
\frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \frac{\partial^2 L}{\partial x_2 \partial x_3} & \frac{\partial g_1}{\partial x_2} \\
\frac{\partial^2 L}{\partial x_3 \partial x_1} & \frac{\partial^2 L}{\partial x_3 \partial x_2} & \frac{\partial^2 L}{\partial x_3^2} & \frac{\partial g_1}{\partial x_3} \\
\hline
\frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \frac{\partial g_1}{\partial x_3} & 0
\end{array}\right) = \left(\begin{array}{ccc|c}
4 & 0 & 0 & 1 \\
0 & 4 & 0 & 1 \\
0 & 0 & 4 & 1 \\
\hline
1 & 1 & 1 & 0
\end{array}\right)$$

We need to check the necessary and sufficient conditions for NLP with inequality constraints…………….

Reference

1. *H.A. Taha*, Operations Research: An Introduction, Prentice hall.

2. S.S. Rao, Engineering Optimization: theory and Practice, New Age Int.

3. M. Bazaraa, H. Sherali,C. Shetty, Nonlinear Programming: Theory and algorithms

Thank You!!