Optimization with Equality and Inequality Constraints

Consider the optimization of continuous functions subjected to equality constraints:

Minimize
$$f = f(\mathbf{x})$$

subject to $g_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, m,$ (1)

where $\mathbf{x} = [x_1, x_2, \dots, x_n]^{\top}$. Here $m \leq n$; otherwise (if m > n), the problem becomes over-defined and, in general, there will be no solution.

1 Solution by Direct Substitution

For a problem with n variables and m equality constraints, it is theoretically possible to solve simultaneously the m equality constraints and express any set of m variables in terms of the remaining n-m variables. When these expressions are substituted into the original objective function, there results a new objective function involving only n-m variables. The new objective function is not subjected to any constraint, and hence its optimum can be found.

This method of direct substitution, although it appears to be simple in theory, is not convenient from a practical point of view. The reason for this is that the constraint equations will be nonlinear for most of practical problems, and often it becomes impossible to solve them and express any m variables in terms of the remaining n-m variables. However, the method of direct substitution might prove to be very simple and direct for solving simpler problems.

Example 1. Find the dimensions of a box of largest volume that can be inscribed in a sphere of unit radius.

Solution:

Let the origin of the Cartesian coordinate system x_1 , x_2 , x_3 be at the center of the sphere and the sides of the box be $2x_1$, $2x_2$, and $2x_3$. The volume of the box is given by

$$f(x_1, x_2, x_3) = 8x_1 x_2 x_3. (2)$$

Since the corners of the box lie on the surface of the sphere of unit radius, x_1 , x_2 , and x_3 have to satisfy the constraint

$$x_1^2 + x_2^2 + x_3^2 = 1. (3)$$

This problem has three design variables and one equality constraint. Hence the equality constraint can be used to eliminate any one of the design variables from the objective function. If we choose to eliminate x_3 , Eq. (3) gives

$$x_3 = (1 - x_1^2 - x_2^2)^{1/2}. (4)$$

Thus the objective function becomes

$$f(x_1, x_2) = 8x_1 x_2 (1 - x_1^2 - x_2^2)^{1/2}, (5)$$

which can be maximized as an unconstrained function in two variables.

The necessary conditions for the maximum of f give

$$\frac{\partial f}{\partial x_1} = 8x_2 \left((1 - x_1^2 - x_2^2)^{1/2} - \frac{x_1^2}{(1 - x_1^2 - x_2^2)^{1/2}} \right) = 0, \tag{6}$$

$$\frac{\partial f}{\partial x_2} = 8x_1 \left((1 - x_1^2 - x_2^2)^{1/2} - \frac{x_2^2}{(1 - x_1^2 - x_2^2)^{1/2}} \right) = 0.$$
 (7)

Equations (6) and (7) can be simplified to obtain

$$1 - 2x_1^2 - x_2^2 = 0 (8)$$

$$1 - x_1^2 - 2x_2^2 = 0, (9)$$

from which it follows that $x_1^* = x_2^* = 1/\sqrt{3}$ and hence $x_3^* = 1/\sqrt{3}$. This solution gives the maximum volume of the box as

$$f_{max} = \frac{8}{3\sqrt{3}}. (10)$$

To find whether the solution found corresponds to a maximum or a minimum, we apply the sufficiency conditions to $f(x_1, x_2)$ of Eq. (5). The second-order partial derivatives of f at (x_1^*, x_2^*) are given by

$$\frac{\partial^2 f}{\partial x_1^2} = -\frac{32}{\sqrt{3}}\tag{11}$$

$$\frac{\partial^2 f}{\partial x_2^2} = -\frac{32}{\sqrt{3}}\tag{12}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -\frac{16}{\sqrt{3}}. (13)$$

Since

$$\frac{\partial^2 f}{\partial x_1^2} < 0$$
, and $\frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 > 0$, (14)

the Hessian matrix of f is negative definite at (x_1^*, x_2^*) . Hence the point corresponds to the maximum of f.

2 Method of Constrained Variation

The basic idea used in the method of constrained variation is to find a closed-form expression for the first-order differential of f(df) at all points at which the constraints $g_j(\mathbf{x}) = 0$, $j = 1, 2, \dots, m$, are satisfied. The desired optimum points are then obtained by setting the differential df equal to zero.

Consider n = 2, and m = 1: Minimize $f(x_1, x_2)$, subject to $g(x_1, x_2) = 0$. A necessary condition for f to have a minimum at some point (x_1^*, x_2^*) is that the total derivative of $f(x_1, x_2)$ with respect to x_1 must be zero at (x_1^*, x_2^*) . By setting the total differential of $f(x_1, x_2)$ equal to zero, we obtain

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0. \tag{15}$$

Since $g(x_1^*, x_2^*) = 0$ at the minimum point, any variations dx_1 and dx_2 taken about the point (x_1^*, x_2^*) are called admissible variations provided that the new point lies on the constraint:

$$g(x_1^* + dx_1, x_2^* + dx_2) = 0. (16)$$

The Taylors series expansion of the function in Eq. (16) about the point (x_1^*, x_2^*) gives

$$g(x_1^* + dx_1, x_2^* + dx_2) \approx g(x_1^*, x_2^*) + \frac{\partial g}{\partial x_1}(x_1^*, x_2^*) dx_1 + \frac{\partial g}{\partial x_2}(x_1^*, x_2^*) dx_2 = 0, (17)$$

where dx_1 and dx_2 are assumed to be small. Since $g(x_1^*, x_2^*) = 0$, Eq. (17) reduces to

$$dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0, \quad \text{at } (x_1^*, x_2^*). \tag{18}$$

Assuming that $\partial g/\partial x_2 \neq 0$, Eq. (18) can be rewritten as

$$dx_2 = \frac{-\partial g/\partial x_1}{\partial g/\partial x_2} dx_1, \quad \text{at } (x_1^*, x_2^*). \tag{19}$$

This relation indicates that once the variation in x_1 (dx_1) is chosen arbitrarily, the variation in x_2 (dx_2) is decided automatically in order to have dx_1 and dx_2 as a set of admissible variations. By substituting Eq. (19) in Eq. (15), we obtain

$$df = \left. \left(\frac{\partial f}{\partial x_1} - \frac{\partial g/\partial x_1}{\partial g/\partial x_2} \frac{\partial f}{\partial x_2} \right) \right|_{(x_1^*, x_2^*)} dx_1 = 0.$$
 (20)

Since dx_1 can be chosen arbitrarily, Eq. (20) leads to

$$\left. \left(\frac{\partial f}{\partial x_1} \frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_1} \right) \right|_{(x_1^*, x_2^*)} = 0.$$
(21)

Equation (21) represents a necessary condition in order to have (x_1^*, x_2^*) as an extreme point (minimum or maximum).

3 Method of Lagrange Multipliers

Consider the problem

Minimize
$$f(x_1, x_2)$$

subject to $g(x_1, x_2) = 0$. (22)

For this problem, the necessary condition for the existence of an extreme point at $\mathbf{x} = \mathbf{x}^*$ is

$$\left. \left(\frac{\partial f}{\partial x_1} - \frac{\partial g/\partial x_1}{\partial g/\partial x_2} \frac{\partial f}{\partial x_2} \right) \right|_{(x_1^*, x_2^*)} = 0.$$
(23)

By defining a quantity λ called the Lagrange multiplier, as

$$\lambda = -\left. \left(\frac{\partial f/\partial x_2}{\partial g/\partial x_2} \right) \right|_{(x_1^*, x_2^*)},\tag{24}$$

Equation (23) can be expressed as

$$\left. \left(\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} \right) \right|_{(x_1^*, x_2^*)} = 0.$$
(25)

and Eq. (24) can be written as

$$\left. \left(\frac{\partial f}{\partial x_2} + \lambda \frac{\partial g}{\partial x_2} \right) \right|_{(x_1^*, x_2^*)} = 0. \tag{26}$$

In addition, the constraint equation has to be satisfied at the extreme point, that is,

$$g(x_1, x_2)|_{(x_1^*, x_2^*)} = 0.$$
 (27)

Thus Eqs. (25) to (27) represent the necessary conditions for the point (x_1^*, x_2^*) to be an extreme point.

The necessary conditions given by Eqs. (25) to (27) are more commonly generated by constructing a function L, known as the Lagrange function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2). \tag{28}$$

By treating L as a function of the three variables x_1 , x_2 , and λ , the necessary conditions for its extremum are given by

$$\begin{split} \frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) \\ &= \frac{\partial f}{\partial x_1}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_1}(x_1, x_2) = 0 \\ \frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) \\ &= \frac{\partial f}{\partial x_2}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_2}(x_1, x_2) = 0 \\ \frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) &= g(x_1, x_2) = 0. \end{split}$$

Theorem 1. Sufficient Condition

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

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^{2} L}{\partial x_{i} \partial x_{j}} dx_{i} dx_{j},$$
(29)

evaluated at $\mathbf{x} = \mathbf{x}^*$ must be positive definite for all values of dx for which the constraints are satisfied.

A necessary condition for the quadratic form Q, defined by Eq. (29), to be positive (negative) definite for all admissible variations dx is that each root of the polynomial z_i , defined by the following determinantal equation, be positive (negative):

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & \cdots & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22} - z & L_{23} & \cdots & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn} - z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0, \quad (30)$$

where

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j} (\mathbf{x}^*, \lambda^*)$$
 (31)

$$g_{ij} = \frac{\partial g_i}{\partial x_j}(\mathbf{x}^*). \tag{32}$$

Example 2. Find the dimensions of a cylindrical tin (with top and bottom) made up of sheet metal to maximize its volume such that the total surface area is equal to $A = 24\pi$.

Solution:

Let x_1 and x_2 denote the radius of the base and height of the tin, respectively, the problem can be stated as

Maximize
$$f(x_1, x_2) = \pi x_1^2 x_2$$

subject to $2\pi x_1^2 + 2\pi x_1 x_2 = A = 24\pi$. (33)

The Lagrange function is

$$L(x_1, x_2, \lambda) = \pi x_1^2 x_2 + \lambda (2\pi x_1^2 + 2\pi x_1 x_2 - A), \tag{34}$$

and the necessary conditions for the maximum of f give

$$\frac{\partial L}{\partial x_1} = 2\pi x_1 x_2 + 4\pi \lambda x_1 + 2\pi \lambda x_2 = 0 \tag{35}$$

$$\frac{\partial L}{\partial x_2} = \pi x_1^2 + 2\pi \lambda x_1 = 0 \tag{36}$$

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A = 0. \tag{37}$$

Equations (35) and (36) lead to

$$\lambda = \frac{-x_1 x_2}{2x_1 + x_2} = -\frac{1}{2} x_1,\tag{38}$$

that is

$$x_1 = \frac{1}{2}x_2. (39)$$

Eqs. (37) and (39) give the desired solution as

$$x_1^* = \left(\frac{A}{6\pi}\right)^{1/2}, \quad x_2^* = \left(\frac{2A}{3\pi}\right)^{1/2}, \quad \lambda^* = -\left(\frac{A}{24\pi}\right)^{1/2}.$$
 (40)

This gives the maximum value of f as

$$f^* = \left(\frac{A^3}{54\pi}\right)^{1/2}. (41)$$

If $A = 24\pi$, the optimum solution becomes

$$x_1^* = 2, \quad x_2^* = 4, \quad \lambda^* = -1, \quad f^* = 16\pi.$$
 (42)

To see that this solution really corresponds to the maximum of f, we apply the sufficiency condition, Eq. (30). In this case

$$L_{11} = \frac{\partial^2 L}{\partial x_1^2} \Big|_{(\mathbf{x}^*, \lambda^*)} = 2\pi x_2^* + 4\pi \lambda^* = 4\pi$$

$$L_{12} = \frac{\partial^2 L}{\partial x_1 \partial x_2} \Big|_{(\mathbf{x}^*, \lambda^*)} = L_{21} = 2\pi x_1^* + 2\pi \lambda^* = 2\pi$$

$$L_{22} = \frac{\partial^2 L}{\partial x_2^2} \Big|_{(\mathbf{x}^*, \lambda^*)} = 0$$

$$g_{11} = \frac{\partial g_1}{\partial x_1} \Big|_{(\mathbf{x}^*, \lambda^*)} = 4\pi x_1^* + 2\pi x_2^* = 16\pi$$

$$g_{12} = \frac{\partial g_1}{\partial x_2} \Big|_{(\mathbf{x}^*, \lambda^*)} = 4\pi.$$
(43)

This gives

$$\begin{vmatrix} L_{11} - z & L_{12} & g_{11} \\ L_{12} & L_{22} - z & g_{12} \\ g_{11} & g_{12} & 0 \end{vmatrix} = \begin{vmatrix} 4\pi - z & 2\pi & 16\pi \\ 2\pi & 0 - z & 4\pi \\ 16\pi & 4\pi & 0 \end{vmatrix} = 0, \tag{44}$$

that is

$$272\pi^2 z + 192\pi^3 = 0, (45)$$

gives

$$z = -\frac{12\pi}{17}. (46)$$

Since the value of z is negative, the point (x_1^*, x_2^*) corresponds to the maximum of f.

4 Optimization with Inequality Constraints

Consider the following problem:

Minimize
$$f(\mathbf{x})$$

subject to $g_j(\mathbf{x}) \le 0, \quad j = 1, 2, \dots, m.$ (47)

4.1 Karush-Kuhn-Tucker Conditions

The conditions to be satisfied at a constrained minimum point, \mathbf{x}^* , of the problem stated can be expressed as

$$\frac{\partial f}{\partial x_i} + \sum_{j} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \quad i = 1, 2, \dots, n$$
(48)

$$\lambda_j > 0, \quad j = 1, 2, \cdots, m. \tag{49}$$

The Karush-Kuhn-Tucker conditions can be stated as follows:

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial g_i}{\partial x_i} = 0, \quad i = 1, 2, \dots, n$$

$$\lambda_j g_j = 0, \quad j = 1, 2, \dots, m, \quad g_j \le 0, \quad \lambda_j \ge 0.$$
(50)

When the optimization problem is stated as

Minimize
$$f(\mathbf{x})$$
 subject to (51)

$$g_j(\mathbf{x}) \le 0, \quad j = 1, 2, \dots, m$$

 $h_k(\mathbf{x}) = 0, \quad k = 1, 2, \dots, p.$ (52)

The Karush-Kuhn-Tucker conditions become

$$\nabla f + \sum_{j=1}^{m} \lambda_j \nabla g_j - \sum_{k=1}^{p} \beta_k \nabla h_k = 0$$
$$\lambda_j g_j = 0, \quad g_j \le 0, \quad h_k = 0, \quad \lambda_j \ge 0,$$
 (53)

where λ_j and β_k denote the Lagrange multipliers associated with the constraints $g_i \leq 0$ and $h_k = 0$, respectively.

Theorem 2. Let \mathbf{x}^* be a feasible solution to the problem of Eqs. (51). If $\nabla g_j(\mathbf{x}^*)$ and $\nabla h_k(\mathbf{x}^*)$, are linearly independent, there exist λ^* and β^* such that $(\mathbf{x}^*, \lambda^*, \beta^*)$ satisfy Eqs. (53).

Example 3. A manufacturing firm producing small refrigerators has entered into a contract to supply 50 refrigerators at the end of the first month, 50 at the end of the second month, and 50 at the end of the third. The cost of producing x refrigerators in any month is given by $(x^2 + 1000)$. The firm can produce more refrigerators in any month and carry them to a subsequent month. However, it costs 20 per unit for any refrigerator carried over from one month to the next. Assuming that there is no initial inventory, determine the number of refrigerators to be produced in each month to minimize the total cost.

Solution:

Let x_1 , x_2 , and x_3 represent the number of refrigerators produced in the first, second, and third month, respectively. The total cost (production cost plus holding cost) to be minimized is given by

$$f(x_1, x_2, x_3) = (x_1^2 + 1000) + (x_2^2 + 1000) + (x_3^2 + 1000) + 20(x_1 - 50) + 20(x_1 + x_2 - 100)$$
$$= x_1^2 + x_2^2 + x_3^3 + 40x_1 + 20x_2.$$
 (54)

The constraints can be stated as

$$g_1(x_1, x_2, x_3) = x_1 - 50 \ge 0,$$

$$g_2(x_1, x_2, x_3) = x_1 + x_2 - 100 \ge 0,$$

$$g_3(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 150 \ge 0.$$
(55)

The Karush-Kuhn-Tucker conditions are given by

$$\frac{\partial f}{\partial x_i} + \sum_j \lambda_j \frac{\partial g_j}{\partial x_i}, \quad i = 1, 2, 3; \quad j = 1, 2, 3.$$
 (56)

That is,

$$2x_1 + 40 + \lambda_1 + \lambda_2 + \lambda_3 = 0$$

$$2x_2 + 20 + \lambda_2 + \lambda_3 = 0$$

$$2x_3 + \lambda_3 = 0,$$

(57)

and for $\lambda_j g_j = 0, j = 1, 2, 3,$

$$\lambda_1(x_1 - 50) = 0$$

$$\lambda_2(x_1 + x_2 - 100) = 0$$

$$\lambda_3(x_1 + x_2 + x_3 - 150) = 0.$$
(58)

For $g_j \geq 0$, we have

$$x_1 - 50 \ge 0$$
, $x_1 + x_2 - 100 \ge 0$, $x_1 + x_2 + x_3 - 150 \ge 0$, (59)

and for $\lambda_i \leq 0, = 1, 2, 3$, then

$$\lambda_1 \le 0, \quad \lambda_2 \le 0, \quad \lambda_3 \le 0. \tag{60}$$

We proceed to solve these equations by first nothing that either $\lambda_1 = 0$ or $x_1 = 50$ according to Eq. (58). Using this information, we investigate the following cases to identify the optimum solution of the problem.

Case 1: $\lambda_1 = 0$

Equations (57) give

$$x_3 = -\frac{\lambda_3}{2}, \quad x_2 = -10 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2}, \quad x_1 = -20 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2}.$$
 (61)

Substituting Eqs. (61) in Eqs. (58), we obtain

$$\lambda_2(-130 - \lambda_2 - \lambda_3) = 0, \quad \lambda_3(-180 - \lambda_2 - 3\lambda_3/2) = 0.$$
 (62)

The four possible solutions of Eqs. (62) are

1. $\lambda_2=0, -180-\lambda_2-3\lambda_3/2=0$. These equations, along with Eqs. (61), yield the solution

$$\lambda_2 = 0, \quad \lambda_3 = -120, \quad x_1 = 40, \quad x_2 = 50, \quad x_3 = 60.$$
 (63)

This solution violates g_1 , g_2 , hence cannot be optimum.

2. $\lambda_3 = 0, -130 - \lambda_2 - \lambda_3 = 0$. The solution of these equations leads to

$$\lambda_2 = -130, \quad \lambda_3 = 0, \quad x_1 = 45, \quad x_2 = 55, \quad x_3 = 0.$$
 (64)

This solution violates g_1 , g_3 , hence cannot be optimum.

3. $\lambda_2 = 0$, $\lambda_3 = 0$. Equations (61) give

$$x_1 = -20, \quad x_2 = -10, \quad x_3 = 0.$$
 (65)

This solution violates g_1 , g_2 , and g_3 hence cannot be optimum.

4. $-130 - \lambda_2 - \lambda_3 = 0$, $-180 - \lambda_2 - 3\lambda_3/2 = 0$. solution of these equations and Eqs. (61) yields

$$\lambda_2 = -30, \quad \lambda_3 = -100, \quad x_1 = 45, \quad x_2 = 55, \quad x_3 = 50.$$
 (66)

This solution violates g_1 , hence cannot be optimum.

Case 2: $x_1 = 50$

In this case, Eqs. (57) give

$$\lambda_3 = -2x_3, \quad \lambda_2 = -20 - 2x_2 + 2x_3, \quad \lambda_1 = -120 + 2x_2.$$
 (67)

Then

$$(-20-2x_2+2x_3)(x_1+x_2-100)=0$$
, $(-2x_3)(x_1+x_2+x_3-150)=0$. (68)

Once again, it can be seen that there are four possible solutions to Eqs. (68), as indicated below:

1. $-20 - 2x_2 + 2x_3 = 0$, $x_1 + x_2 + x_3 - 150 = 0$. The solution of these equations yields

$$x_1 = 50, \quad x_2 = 45, \quad x_3 = 55.$$
 (69)

This solution can be seen to violate g_2 .

2. $-20-2x_2+2x_3=0$, $-2x_3=0$. These equations lead to the solution

$$x_1 = 50, \quad x_2 = -10, \quad x_3 = 0.$$
 (70)

This solution can be seen to violate g_2 , g_3 .

3. $x_1 + x_2 - 100 = 0$, $-2x_3 = 0$. These equations lead to the solution

$$x_1 = 50, \quad x_2 = 50, \quad x_3 = 0.$$
 (71)

This solution can be seen to violate g_3 .

4. $x_1 + x_2 - 100 = 0$, $x_1 + x_2 + x_3 - 150 = 0$. The solution of these equations yields

$$x_1 = 50, \quad x_2 = 50, \quad x_3 = 50.$$
 (72)

This solution can be seen to satisfy all the constraint g_i . The values of λ_i corresponding to this solution can be obtained from Eqs. (67) as

$$\lambda_1 = -20, \quad \lambda_2 = -20, \quad \lambda_3 = -100.$$
 (73)

The identified optimal solutions are then

$$x_1^* = 50, \quad x_2^* = 50, \quad x_3^* = 50.$$
 (74)