Lecture Assignment 8 a)

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

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

$$f(x1, x2, x3) = 8x_1x_2x_3$$

Since the corners of the box lie on the surface of the sphere of unit radius, x1, x2, and x3 have to satisfy the constraint

$$x_1^2 + x_2^2 + x_3^2 = 1$$

Lecture Assignment 8 a)

This problem has three design variables and one equality constraint.

The equality constraint can be used to eliminate any one of the design variables from the objective function

$$x_3 = \sqrt{1 - x_1^2 - x_2^2}$$

or

$$f(x1, x2, x3) = 8x_1x_2\sqrt{1 - x_1^2 - x_2^2}$$

Lecture Assignment 8 a)

The necessary conditions for the maximum of f give

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

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

equations can be simplified to obtain,

$$1 - 2x_1^2 - x_2^2 = 0$$

$$1 - x_1^2 - 2x_2^2 = 0$$

Lecture Assignment 8 a)

Solving this gives

$$x_1^* = x_2^* = 1/\sqrt{3} \text{ and hence } x_3^* = 1/\sqrt{3}.$$

This solution gives the maximum volume of the box as

$$fmax = \frac{8}{3\sqrt{3}}$$

To find whether the solution found corresponds to a maximum or a minimum, we apply the sufficiency conditions to $f(x1, x2)$.

Lecture Assignment 8 a)

The second-order partial derivatives of f at (x_1^*, x_2^*) are given by

$$\frac{\partial^2 f}{\partial x_1^2} = -\frac{32}{\sqrt{3}} \text{ at } (x_1^*,\ x_2^*)$$

$$\frac{\partial^2 f}{\partial x_2^2} = -\frac{32}{\sqrt{3}} \text{ at } (x_1^*,\ x_2^*)$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -\frac{16}{\sqrt{3}} \text{ at } (x_1^*,\ x_2^*)$$

and

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

Lecture Assignment 8 a)

Since $\dfrac{\partial^2 f}{\partial x_1^2} < 0$ and $\dfrac{\partial^2 f}{\partial x_1^2}\dfrac{\partial^2 f}{\partial x_2^2} - \left(\dfrac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 > 0$

the Hessian matrix of f is negative definite at

$$(x_1^*, x_2^*)$$

Hence the point (x_1^*, x_2^*) corresponds to the maximum of f.

Multivariable optimization with equality constraints

Method of Lagrange Multipliers:

Problem with two variables and one constraint.

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

subject to

$$g(x_1, x_2) = 0$$

For this problem, the necessary condition for the existence of an extreme point at

$$\boldsymbol{x} = \boldsymbol{x}^*$$

is

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

Method of Lagrange Multipliers:

By defining a quantity λ, called the *Lagrange multiplier*, as

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

Equations can be expressed as

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

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

Method of Lagrange Multipliers:

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

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

Notice that the partial derivative

has to be nonzero to be able to define λ

Method of Lagrange Function:

The necessary conditions are more commonly generated by constructing a function L, known as the Lagrange function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

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

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

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

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

Method of Lagrange Function:

The necessary conditions are more commonly generated by constructing a function L, known as the Lagrange function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

By treating L as a function of the three variables x1, x2, and ,
the necessary conditions for its extremum are given by

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

Method of Lagrange Functions

Method of Lagrange Function:

The necessary conditions for general problem

$$\text{Minimize } f(\boldsymbol{x})$$

subject to

$$g_j(\boldsymbol{x}) = 0, \qquad j = 1, 2, \ldots, m$$

The Lagrange function, L, in this case is defined by introducing one Lagrange multiplier λ_j for each constraint $g_j(\boldsymbol{x})$ as

$$L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$$
$$= f(\boldsymbol{x}) + \lambda_1 g_1(\boldsymbol{x}) + \lambda_2 g_2(\boldsymbol{x}) + \cdots + \lambda_m g_m(\boldsymbol{x})$$

Method of Lagrange Functions

Method of Lagrange Function:

By treating L as a function of the $n + m$ unknowns,

$$x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m,$$

the necessary conditions for the extremum of L, which also correspond to the solution of the original problem are given by

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1, 2, \ldots, n$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{X}) = 0, \qquad j = 1, 2, \ldots, m$$

Equations represent $n + m$ equations in terms of the $n + m$ unknowns, x_i and λ_j .

Method of Lagrange Function:

The solution gives

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{Bmatrix} \quad \text{and} \quad \lambda^* = \begin{Bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_m^* \end{Bmatrix}$$

The vector $\boldsymbol{X}^*$ corresponds to the relative constrained minimum of $f(\boldsymbol{X})$ (sufficient conditions are to be verified) while the vector λ^* provides the sensitivity information.

Method of Lagrange Function

Sufficient Condition

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

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

evaluated at $\boldsymbol{X} = \boldsymbol{X}^*$ must be positive definite for all values of $d\boldsymbol{X}$ for which the constraints are satisfied..

Method of Lagrange Function

Sufficient Condition

It has been shown that a necessary condition for the quadratic form Q, to be positive (negative) definite for all admissible variations $d\boldsymbol{X}$ is that each root of the polynomial z_i , defined by the following determinantal equation, be positive (negative), see next page, where,

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

$$g_{ij} = \frac{\partial g_i}{\partial x_j}(\mathbf{X}^*)$$

Method of Lagrange Function

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

Lecture Assignment #8 b

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 $A_0 = 24\pi$.

x_1 : the radius of the base

x_2 : the length

$$\begin{array}{ll} \text{Maximize} & f(x_1, x_2) = \pi x_1^2 x_2 \\ \text{subject to} & 2\pi x_1^2 + 2\pi x_1 x_2 = A_0 = 24\pi \end{array}$$

The Lagrange function is:

$$L(x_1, x_2, \lambda) = \pi x_1^2 x_2 + \lambda(2\pi x_1^2 + 2\pi x_1 x_2 - A_0)$$

Lecture Assignment #8 b

Necessary conditions:

$$\frac{\delta L}{\delta x_1} = 2\pi x_1 x_2 + 4\pi x_1 + 2\pi\lambda x_2 = 0 \;\Rightarrow\; \lambda = -\frac{x_1 x_2}{2x_1 + x_2}$$

$$\frac{\delta L}{\delta x_2} = \pi x_1^2 + 2\pi\lambda x_1 \qquad \Rightarrow\; \lambda = -\frac{x_1}{2} \;\text{ or }\; x_1 = \frac{1}{2}x_2$$

$$\frac{\delta L}{\delta \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 \quad \Rightarrow x_1^* = \left(\frac{A_0}{6\pi}\right)^{\frac{1}{2}}$$

$$x_2^* = \left(\frac{2A_0}{3\pi}\right)^{\frac{1}{2}} \text{ and } \lambda^* = \left(\frac{A_0}{24\pi}\right)^{\frac{1}{2}}$$

This gives the maximum value of $f^* = \left(\frac{A_0^3}{54\pi}\right)^{\frac{1}{2}}$

Lecture Assignment #8 b

Sufficiency conditions:

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

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

$$L_{22} = \left.\frac{\partial^2 L}{\partial x_2^2}\right|_{(\mathbf{X}^*,\lambda^*)} = 0$$

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

$$g_{12} = \left.\frac{\partial g_1}{\partial x_2}\right|_{(\mathbf{X}^*,\lambda^*)} = 2\pi x_1^* = 4\pi$$

Lecture Assignment #8 b

Sufficiency conditions:

$$\begin{vmatrix} 4\pi - z & 2\pi & 16\pi \\ 2\pi & 0 - z & 4\pi \\ 16\pi & 4\pi & 0 \end{vmatrix} = 0$$

that is, $272\pi^2 z + 192\pi^3 = 0$

This gives $z = -\frac{12}{17}\pi$

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

Kuhn-Tucker Conditions

Multivariable optimization with inequality constraints

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

subject to

$$g_j(\boldsymbol{X}) \leq 0, \quad j = 1, 2, \ldots, m$$

The inequality constraints can be transformed to equality constraints by adding nonnegative slack variables, y_j^2

as
$$g_j(\boldsymbol{X}) + y_j^2 = 0, \quad j = 1, 2, \ldots, m$$

where the values of the slack variables are yet unknown.

Kuhn-Tucker Conditions

Minimize $f(X)$

subject to

$$G_j(\boldsymbol{X}, \boldsymbol{Y}) = g_j(\boldsymbol{X}) + y_j^2 = 0 \quad \text{j = 1, 2, . . . ,m}$$

where $\boldsymbol{Y} = \{y_1, y_2, \ldots, y_m\}^T$ is the vector of slack variables.

This problem can be solved conveniently by the method of Lagrange multipliers.

$$L(\mathbf{X}, \mathbf{Y}, \boldsymbol{\lambda}) = f(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j G_j(\mathbf{X}, \mathbf{Y})$$

where $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}T$, is the vector of Lagrange multipliers.

Kuhn-Tucker Conditions

The stationary points of the Lagrange function can be found by solving the following equations,

$$\frac{\partial L}{\partial x_i}(\mathbf{X}, \mathbf{Y}, \boldsymbol{\lambda}) = \frac{\partial f}{\partial x_i}(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(\mathbf{X}) = 0, \qquad i = 1, 2, \ldots, n$$

$$\frac{\partial L}{\partial \lambda_j}(\mathbf{X}, \mathbf{Y}, \boldsymbol{\lambda}) = G_j(\mathbf{X}, \mathbf{Y}) = g_j(\mathbf{X}) + y_j^2 = 0, \qquad j = 1, 2, \ldots, m$$

$$\frac{\partial L}{\partial y_j}(\mathbf{X}, \mathbf{Y}, \boldsymbol{\lambda}) = 2\lambda_j y_j = 0, \qquad j = 1, 2, \ldots, m$$

Kuhn-Tucker Conditions

This represent $(n + 2m)$ equations in the $(n + 2m)$ unknowns, $\boldsymbol{X}$, $\boldsymbol{\lambda}$ and $\boldsymbol{Y}$. The solution gives the optimum solution vector, $\boldsymbol{X}^*$; the Lagrange multiplier vector, $\boldsymbol{\lambda}^*$; and the slack variable vector, $\boldsymbol{Y}^*$.

Kuhn-Tucker Conditions

the conditions to be satisfied at a constrained minimum point, $\boldsymbol{X}^*$, can be expressed as

$$\frac{\partial f}{\partial x_i} + \sum_{j \in J_1} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1, 2, \ldots, n$$

$$\lambda_j > 0, \qquad j \in J_1$$

These are called *Kuhn–Tucker conditions,* necessary conditions to be satisfied at a relative minimum of $f\ (\boldsymbol{X})$. These conditions are, in general, not sufficient to ensure a relative minimum.

Kuhn-Tucker Conditions

However, there is a class of problems, called *convex programming problems*, for which the Kuhn–Tucker conditions are necessary and sufficient for a global minimum.
If the set of active constraints is not known, the Kuhn–Tucker conditions can be stated as follows:

$$\frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1, 2, \ldots, n$$

$$\lambda_j g_j = 0,^{\ddagger} \qquad j = 1, 2, \ldots, m$$

$$g_j \leq 0, \qquad j = 1, 2, \ldots, m$$

$$\lambda_j \geq 0, \qquad j = 1, 2, \ldots, m$$

Kuhn-Tucker Conditions

Note that if the problem is one of maximization or if the constraints are of the type $g_j \geq 0$, the λ_j have to be nonpositive.

On the other hand, if the problem is one of maximization with constraints in the form type $g_j \geq 0$, the λ_j have to be nonnegative.

Kuhn-Tucker Conditions

When the optimization problem is stated as

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

subject to

$$g_j(\mathbf{X}) \leq 0, \qquad j = 1, 2, \ldots, m$$

$$h_k(\mathbf{X}) = 0 \qquad k = 1, 2, \ldots, p$$

the Kuhn–Tucker conditions become

$$\nabla f + \sum_{j=1}^{m} \lambda_j \nabla g_j - \sum_{k=1}^{p} \beta_k \nabla h_k = \mathbf{0}$$

$$\lambda_j g_j = 0, \qquad j = 1, 2, \ldots, m$$

Kuhn-Tucker Conditions

$$g_j \leq 0, \qquad j = 1, 2, \ldots, m$$
$$h_k = 0, \qquad k = 1, 2, \ldots, p$$
$$\lambda_j \geq 0, \qquad j = 1, 2, \ldots, m$$

where λ_j and β_k denote the Lagrange multipliers associated with the constraints $g_j \leq 0$ and $h_k = 0$, respectively. Although we found qualitatively that the Kuhn–Tucker conditions represent the necessary conditions of optimality, the following theorem gives the precise conditions of optimality.