Computational Intelligence

Classical Optimization

Introduction

- Optimization is the act of obtaining the best result under given circumstances.

- Optimization can be defined as the process of finding the conditions that give the maximum or minimum of a function.

- The optimum seeking methods are also known as *mathematical programming techniques* and are generally studied as a part of operations research.

- *Operations research* is a branch of mathematics concerned with the application of scientific methods and techniques to decision making problems and with establishing the best or optimal solutions.

Introduction

- **Operations research** (in the UK) or **operational research (OR)** (in the US) or **yöneylem araştırması** (in Turkish) is an interdisciplinary branch of mathematics which uses methods like:

  - mathematical modeling
  - statistics
  - algorithms to arrive at optimal or good decisions in complex problems which are concerned with optimizing the maxima (profit, faster assembly line, greater crop yield, higher bandwidth, etc) or minima (cost loss, lowering of risk, etc) of some objective function.

- The eventual intention behind using operations research is to elicit a best possible solution to a problem mathematically, which improves or optimizes the performance of the system.

Introduction

TABLE 1.1 Methods of Operations Research

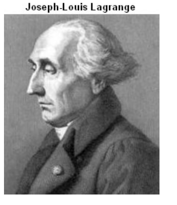
Mathematical Programming Techniques	Stochastic Process Techniques	Statistical Methods
Calculus methods	Statistical decision theory	Regression analysis
Calculus of variations	Markov processes	Cluster analysis, pattern recognition
Nonlinear programming	Queueing theory	Design of experiments
Geometric programming	Renewal theory	Discriminate analysis (factor analysis)
Quadratic programming	Simulation methods	
Linear programming	Reliability theory	
Dynamic programming		
Integer programming		
Stochastic programming		
Separable programming		
Multiobjective programming		
Network methods: CPM and PERT		
Game theory		
Simulated annealing		
Genetic algorithms		
Neural networks		

Introduction

Historical development

- Isaac Newton (1642-1727)
  (The development of differential calculus
  methods of optimization)

- Joseph-Louis Lagrange (1736-1813)
  (Calculus of variations, minimization of functionals,
  method of optimization for constrained problems)



- Augustin-Louis Cauchy (1789-1857)
  (Solution by direct substitution, steepest
  descent method for unconstrained optimization)

Introduction

Historical development

- Leonhard Euler (1707-1783)
  (Calculus of variations, minimization of functionals)

- Gottfried Leibniz (1646-1716)
  (Differential calculus methods of optimization)

İsim: Gottfired Wilhelm von Leibniz

Introduction

Historical development

- George Bernard Dantzig (1914-2005)
  (Linear programming and Simplex method (1947))

- Richard Bellman (1920-1984)
  (Principle of optimality in dynamic
  programming problems)

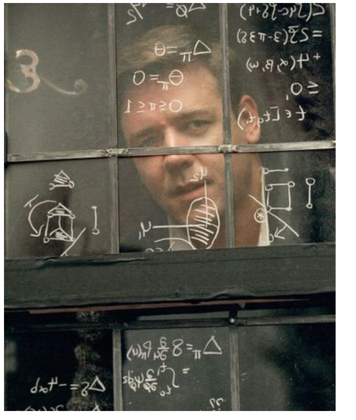
- Harold William Kuhn (1925-)
  (Necessary and sufficient conditions for the optimal solution of
  programming problems, game theory)

Introduction

Historical development

- Albert William Tucker (1905-1995) (Necessary and sufficient conditions for the optimal solution of programming problems, nonlinear programming, game theory: his PhD student was John Nash)

- Von Neumann (1903-1957) (game theory)



Introduction

- **Mathematical optimization problem:**

$$\text{minimize } f_0(x)$$
$$\text{subject to } g_i(x) \le b_i, \quad i = 1, \ldots, m$$

- f_0 : $\boldsymbol{R^n} \longrightarrow \boldsymbol{R}$: objective function
- $x=(x_1,\ldots,x_n)$: design variables (unknowns of the problem, they must be linearly independent)
- g_i : $\boldsymbol{R^n} \longrightarrow \boldsymbol{R}$: $(i=1,\ldots,m)$: inequality constraints

- The problem is a constrained optimization problem

Introduction

- If a point x^* corresponds to the minimum value of the function $f(x)$, the same point also corresponds to the maximum value of the negative of the function, $-f(x)$. Thus optimization can be taken to mean minimization since the maximum of a function can be found by seeking the minimum of the negative of the same function.

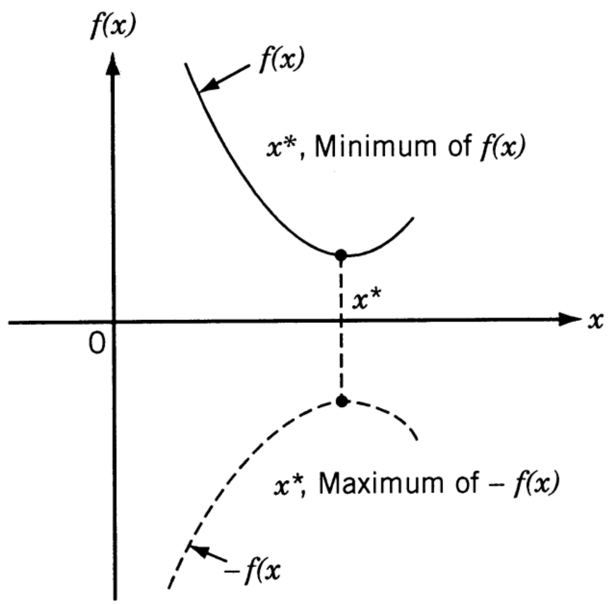
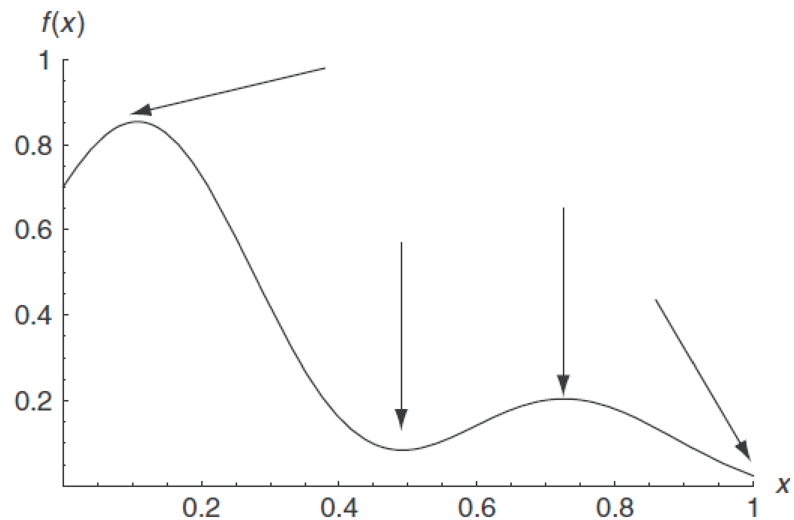


Figure 1.1 Minimum of $f(x)$ is same as maximum of $-f(x)$.

Local vs Global



The arrows indicate the location of local and global optima for a function defined on the closed interval [0, 1]. The leftmost arrow points to the global optimum, whereas the rightmost arrow indicates the global minimum. Note that the global minimum is not, strictly speaking, a local minimum, since a neighbourhood cannot be defined at $x = 1$.

In addition to local optima, the concept of *global* optima is essential in optimization: a function $f : \mathbf{D} \to \mathbf{R}$, has a **global minimum** at a point x^* if $f(\mathbf{x}) \geq f(\mathbf{x}^*)\ \forall x \in \mathbf{D}$.

Constraints

- **Mathematical optimization problem:**

$$\text{minimize } f_0(x)$$
$$\text{subject to } g_i(x) \le b_i, \quad i = 1, \ldots, m$$

- $f_0 : \boldsymbol{R}^n \rightarrow \boldsymbol{R}$: objective function
- $x=(x_1, \ldots, x_n)$: design variables (unknowns of the problem, they must be linearly independent)
- $g_i : \boldsymbol{R}^n \rightarrow \boldsymbol{R}$: $(i=1, \ldots, m)$: inequality constraints

- The problem is a constrained optimization problem

Linear Programming

A common special case of continuous optimization is **linear programming**, in which the objective function is linear, and the constraint functions are affine, i.e. of the form $g(\mathbf{x}) = \mathbf{a}^\mathsf{T}\mathbf{x} + b$, where $\mathbf{a}$ and $\mathbf{x}$ are n-dimensional vectors. More formally, linear programming is defined as

$$\left.\begin{array}{lll} \text{minimize} & \mathbf{c}^\mathsf{T}\mathbf{x} & \\ \text{subject to} & \mathbf{a}_i^\mathsf{T}\mathbf{x} + b_i \leq 0, & i = 1, \ldots, m \\ \text{and} & -x_i \leq 0, & i = 1, \ldots, n. \end{array}\right\}$$

Non-Linear Programming

The restriction to non-negative variables is motivated by the fact that integer programming usually concerns problem involving, for example, production, transportation or investment, in which the quantities involved are, of course, non-negative. If the objective function or at least one of the constraint functions are non-linear, the problem is one of **non-linear programming**. An important special case of non-linear programming is **quadratic programming**, which, in vector notation, takes the form

$$\left.\begin{array}{lll} \text{minimize} & \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} + \mathbf{c}^{\mathrm{T}}\mathbf{x}, & \\ \text{subject to} & \mathbf{a}_i^{\mathrm{T}}\mathbf{x} + b_i \leq 0, & i = 1, \ldots, m \\ \text{and} & -x_i \leq 0, & i = 1, \ldots, n. \end{array}\right\}$$

Review of Mathematics…1

Concepts from linear algebra: Positive definiteness

- **Test 1:** A matrix **A** will be positive definite if all its eigenvalues are positive; that is, all the values of λ that satisfy the determinental equation

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

should be positive. Similarly, the matrix **A** will be negative definite if its eigenvalues are negative.

Review of Mathematics…2

Positive definiteness

- **Test 2:** Another test that can be used to find the positive definiteness of a matrix **A** of order *n* involves evaluation of the determinants

$$A = |a_{11}|$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \cdots a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots a_{2n} \\ a_{31} & a_{32} & a_{33} \cdots a_{3n} \\ \vdots & & \\ a_{n1} & a_{n2} & a_{n3} \cdots a_{nn} \end{vmatrix}$$

- The matrix **A** will be **positive definite** if and only if all the values A_1, A_2, A_3, …A_n are positive
- The matrix **A** will be **negative definite** if and only if the sign of A_j is

$(-1)^j$ for $j=1,2,\ldots,n$

- If some of the A_j are positive and the remaining A_j are zero, the matrix **A** will be **positive semidefinite**

Review of mathematics…3

Negative definiteness

- Equivalently, a matrix is **negative-definite** if all its **eigenvalues** are **negative**
- It is **positive-semidefinite** if all its **eigenvalues are all greater than or equal to zero**
- It is **negative-semidefinite** if all its **eigenvalues are all less than or equal to zero**

Review of mathematics…4

Concepts from linear algebra:

Nonsingular matrix: The determinant of the matrix is not zero.

Rank: The rank of a matrix **A** is the order of the largest nonsingular square submatrix of **A,** that is, the largest submatrix with a determinant other than zero.

Example

Example - Is the following matrix positive definite?

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

2

$$\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3$$

$$\begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 8 - 2 - 2 = 4$$

$2, 3, 4 > 0 \Rightarrow$ Positive definite.

Example

Example - For what numbers b is the following matrix positive semidefinite?

$$\begin{pmatrix} 2 & -1 & b \\ -1 & 2 & -1 \\ b & -1 & 2 \end{pmatrix}$$

$$\begin{vmatrix} 2 & -1 & b \\ -1 & 2 & -1 \\ b & -1 & 2 \end{vmatrix} = 8 + b + b - 2 - 2 - 2b^2$$

$$= -2b^2 + 2b + 4 \geq 0$$

$$\Rightarrow b^2 - b - 2 \leq 0$$

$$(b-2)(b+1) \leq 0$$

$$\boxed{-1 \leq b \leq 2}$$

2. Classical optimization techniques
Single variable optimization

- Useful in finding the optimum solutions of continuous and differentiable functions

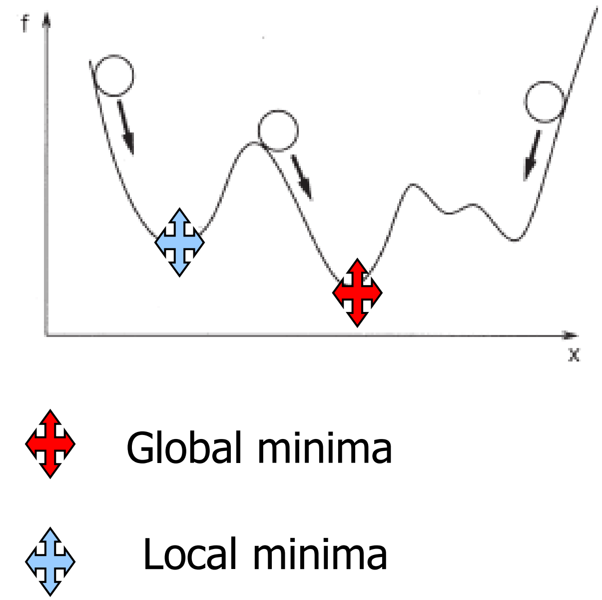
- These methods are analytical and make use of the techniques of differential calculus in locating the optimum points.

- Since some of the practical problems involve objective functions that are not continuous and/or differentiable, the classical optimization techniques have limited scope in practical applications.

2. Classicial optimization techniques
Single variable optimization

- A function of one variable $f(x)$ has a relative or local minimum at $x = x^*$ if $f(x^*) \leq f(x^*+h)$ for all sufficiently small positive and negative values of h

- A point x^* is called a relative or local maximum if $f(x^*) \geq f(x^*+h)$ for all values of h sufficiently close to zero.



2. Classicial optimization techniques
Single variable optimization

- A function $f(x)$ is said to have a global or absolute minimum at x^* if $f(x^*) \le f(x)$ for all x, and not just for all x close to x^*, in the domain over which $f(x)$ is defined.

- Similarly, a point x^* will be a global maximum of $f(x)$ *if* $f(x^*) \ge f(x)$ for all x in the domain.

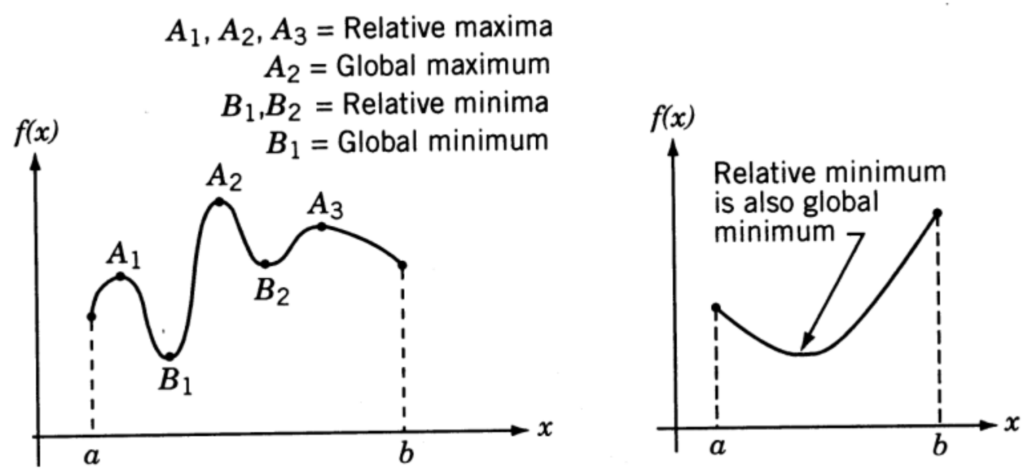
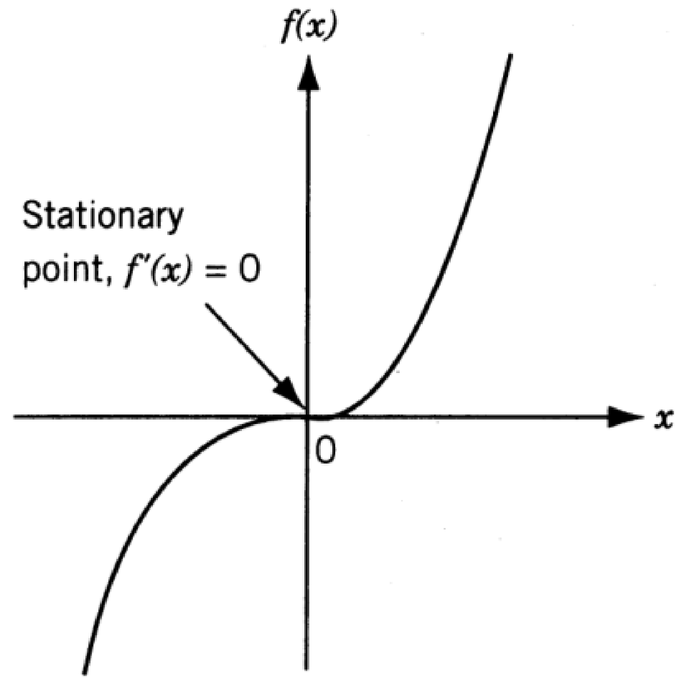


Figure 2.1 Relative and global minima.

Necessary condition

- If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $df(x) / dx = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*)=0$

- The theorem does not say that the function necessarily will have a minimum or maximum at every point where the derivative is zero. *e.g.* $f'(x)=0$ at $x = 0$ for the function shown in figure. However, this point is neither a minimum nor a maximum. In general, a point x^* at which $f'(x^*)=0$ is called a *stationary point.*



Necessary condition

- The theorem does not say what happens if a minimum or a maximum occurs at a point x^* where the derivative fails to exist. For example, in the figure

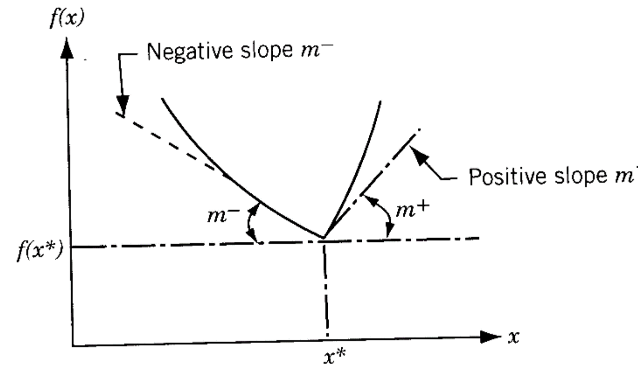


Figure 2.2 Derivative undefined at x^*.

$$\lim_{h \to 0} \frac{f(x^*+h) - f(x^*)}{h} = m^+ \text{ (positive) or } m^- \text{ (negative)}$$

depending on whether h approaches zero through positive or negative values, respectively. Unless the numbers m^+ or m^- are equal, the derivative $f'(x^*)$ does not exist. If $f'(x^*)$ does not exist, the theorem is not applicable.

Sufficient condition

- Let $f'(x^*)=f''(x^*)=\ldots=f^{(n-1)}(x^*)=0$, but $f^{(n)}(x^*) \neq 0$. Then $f(x^*)$ is
  - A **minimum** value of $f(x)$ if $f^{(n)}(x^*) > 0$ and n is **even**
  - A **maximum** value of $f(x)$ if $f^{(n)}(x^*) < 0$ and n is **even**
  - Neither a minimum nor a maximum if n is **odd**

Example

Determine the maximum and minimum values of the function:

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$$

Solution: Since $f'(x)=60(x^4-3x^3+2x^2)=60x^2(x-1)(x-2)$,

$f'(x)=0$ at $x=0, x=1$, and $x=2$.

The second derivative is:

$$f''(x) = 60(4x^3 - 9x^2 + 4x)$$

At $x=1$, $f''(x)=-60$ and hence $x=1$ is a relative maximum. Therefore,

$$f_{max} = f(x=1) = 12$$

At $x=2$, $f''(x)=240$ and hence $x=2$ is a relative minimum. Therefore,

$$f_{min} = f(x=2) = -11$$

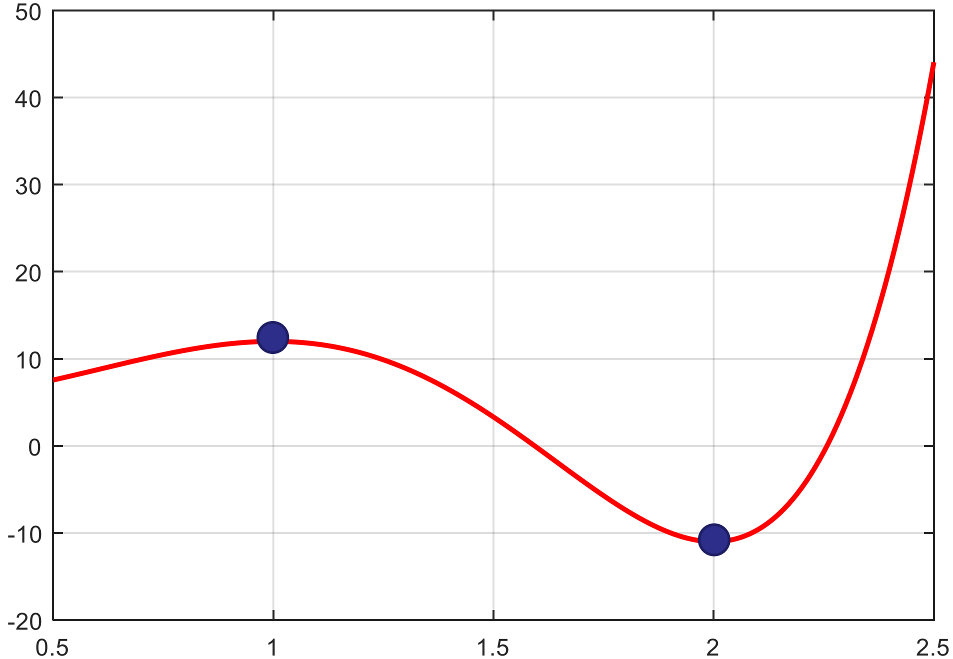
Example

Solution cont'd:

At *x=0, f"(x)=0* and hence we must investigate the next derivative.

$$f'''(x) = 60(12x^2 - 18x + 4) = 240 \text{ at } x = 0$$

Since $f'''(x) \neq 0$ at *x=0, x=0* is neither a maximum nor a minimum, and it is an inflection point (n is **<u>ODD</u>**).

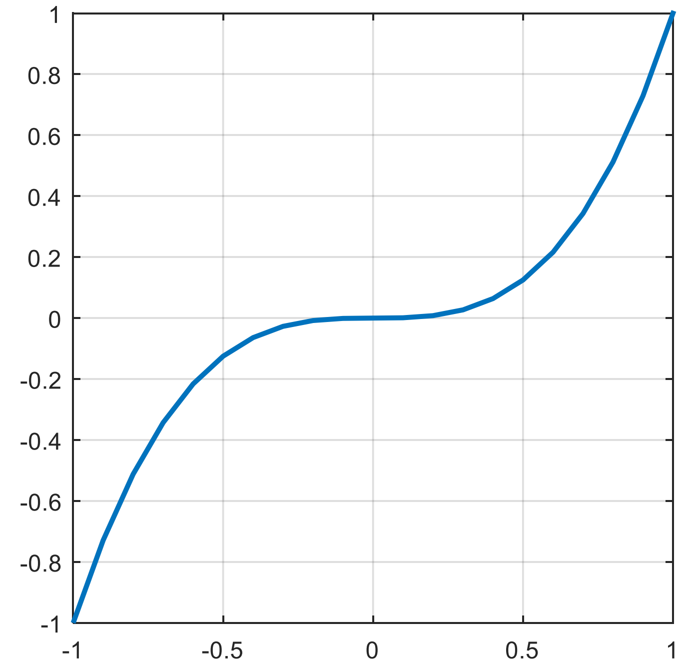


Example

Determine the maximum and minimum values of the function:

$$f(x)=x^3$$

Solution: Since $f'(x)=3x^2$, $f'(x)=0$ at $x=0$.



The second derivative is:

At $x=0$, $f''(x)=6x=0$.

The Third derivative is:

At $x=0$, $f'''(x)=6$.

Since $f'''(x) \neq 0$ at $x=0$, $x=0$ is neither a maximum nor a minimum, and it is an inflection point (n is **ODD**).

Multivariable optimization with no constraints

- **Necessary condition**

  If $f(\mathbf{X})$ has an extreme point (maximum or minimum) at $\mathbf{X}=\mathbf{X}^*$ and if the first partial derivatives of $f(\mathbf{X})$ exist at $\mathbf{X}^*$, then

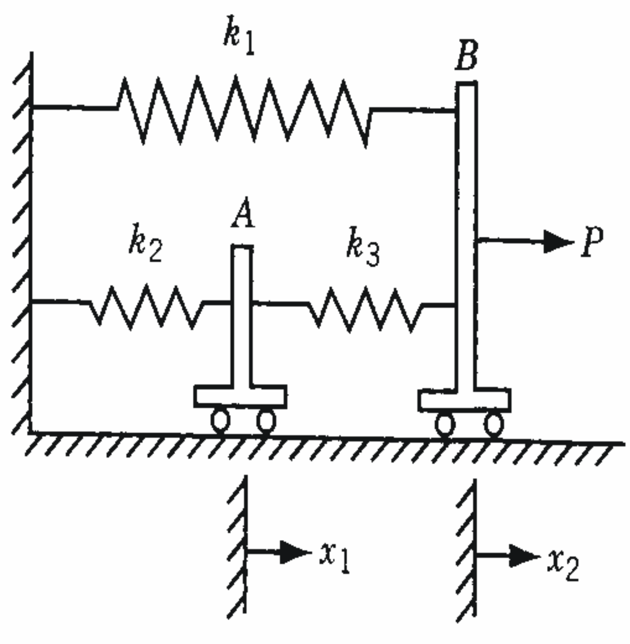
$$\frac{\partial f}{\partial x_1}(\mathbf{X}^*) = \frac{\partial f}{\partial x_2}(\mathbf{X}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{X}^*) = 0$$

- **Sufficient condition**

  A sufficient condition for a stationary point $\mathbf{X}^*$ to be an extreme point is that the matrix of second partial derivatives (Hessian matrix) of $f(\mathbf{X}^*)$ evaluated at $\mathbf{X}^*$ is
  - **Positive definite** when $\mathbf{X}^*$ is a **relative minimum point**
  - **Negative definite** when $\mathbf{X}^*$ is a **relative maximum point**

Example

Figure shows two frictionless rigid bodies (carts) A and B connected by three linear elastic springs having spring constants k_1, k_2, and k_3. The springs are at their natural positions when the applied force P is zero. Find the displacements x_1 and x_2 under the force P by using the principle of minimum potential energy.



Example

Solution: According to the principle of minimum potential energy, the system will be in equilibrium under the load *P* if the potential energy is a minimum. The potential energy of the system is given by:

Potential energy (U)

= Strain energy of springs-work done by external forces

$$=[\frac{1}{2}k_2x_1^2+\frac{1}{2}k_3(x_2-x_1)^2+\frac{1}{2}k_1x_2^2]-Px_2$$

The necessary condition for the minimum of U are

$$\left.\begin{aligned}\frac{\partial U}{\partial x_1}&=k_2x_1-k_3(x_2-x_1)=0\\ \frac{\partial U}{\partial x_2}&=k_3(x_2-x_1)+k_1x_2-P=0\end{aligned}\right\}\qquad \begin{aligned}x_1^*&=\frac{Pk_3}{k_1k_2+k_1k_3+k_2k_3}\\ x_2^*&=\frac{P(k_2+k_3)}{k_1k_2+k_1k_3+k_2k_3}\end{aligned}$$

Example

Solution cont'd: The sufficiency conditions for the minimum at (x_1*,x_2*) can also be verified by testing the positive definiteness of the Hessian matrix of U. The Hessian matrix of U evaluated at (x_1*,x_2*) is:

$$\mathbf{J}\Big|_{(x_1^*,x_2^*)} = \begin{bmatrix} \dfrac{\partial^2 U}{\partial x_1^2} & \dfrac{\partial^2 U}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 U}{\partial x_1 \partial x_2} & \dfrac{\partial^2 U}{\partial x_2^2} \end{bmatrix}_{(x_1^*,x_2^*)} = \begin{bmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{bmatrix}$$

The determinants of the square submatrices of **J** are

$$J_1 = \left| k_2 + k_3 \right| = k_2 + k_3 > 0$$

$$J_2 = \begin{vmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{vmatrix} = k_1 k_2 + k_1 k_3 + k_2 k_3 > 0$$

Since the spring constants are always positive. Thus the matrix **J** is positive definite and hence (x_1*,x_2*) corresponds to the minimum of potential energy.

Semi-definite case

The sufficient conditions for the case when the Hessian matrix of the given function is semidefinite:

- In case of a function of a single variable, the higher order derivatives in the Taylor's series expansion are investigated

Semi-definite case

The sufficient conditions for a function of several variables for the case when the Hessian matrix of the given function is semidefinite:

- Let the partial derivatives of f of all orders up to the order $k \geq 2$ be continuous in the neighborhood of a stationary point **X***, and

$$d^r f|_{\mathbf{X}=\mathbf{X}^*} = 0 \quad 1 \leq r \leq k-1$$

$$d^k f|_{\mathbf{X}=\mathbf{X}^*} \neq 0$$

so that $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is the first nonvanishing higher-order differential of f at **X*.**

- **If *k* is even:**
  - **X*** is a relative minimum if $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is positive definite
  - **X*** is a relative maximum if $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is negative definite
  - If $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is semidefinite, no general conclusions can be drawn
- **If *k* is odd**, **X*** is not an extreme point of f**(X*)**

Saddle point

- In the case of a function of two variables $f(x,y)$, the Hessian matrix may be neither positive nor negative definite at a point (x^*,y^*) at which

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$$

In such a case, the point (x^*,y^*) is called a saddle point.

- The characteristic of a saddle point is that it corresponds to a relative minimum or maximum of $f(x,y)$ wrt one variable, say, x (the other variable being fixed at $y=y^*$) and a relative maximum or minimum of $f(x,y)$ wrt the second variable y (the other variable being fixed at x*).

Saddle point

Example: Consider the function $f(x,y)=x^2-y^2$. For this function:

$$\frac{\partial f}{\partial x} = 2x \text{ and } \frac{\partial f}{\partial y} = -2y$$

These first derivatives are zero at $x^* = 0$ and $y^* = 0$. The Hessian matrix of f at (x^*,y^*) is given by:

$$\mathbf{J} = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix}$$

Since this matrix is neither positive definite nor negative definite, the point ($x^*=0$, $y^*=0$) is a saddle point.

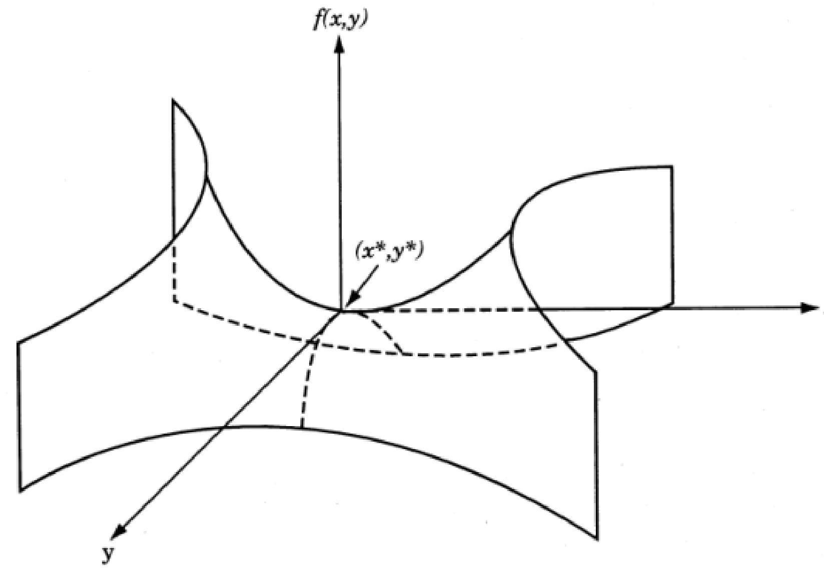


Figure 2.5 Saddle point of the function $f(x,y) = x^2 - y^2$.

Saddle point

Example cont'd:

It can be seen from the figure that $f(x, y^*) = f(x, 0)$ has a relative minimum and $f(x^*, y) = f(0, y)$ has a relative maximum at the saddle point (x^*, y^*).

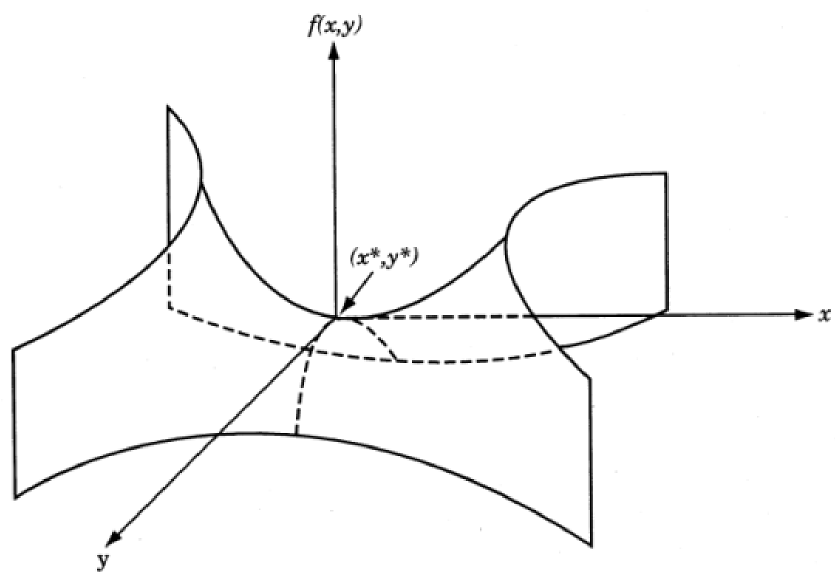


Figure 2.5 Saddle point of the function $f(x,y) = x^2 - y^2$.

Example

Find the extreme points of the function

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

Solution: The necessary conditions for the existence of an extreme point are:

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) = 0$$

$$\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8) = 0$$

These equations are satisfied at the points: (0,0), (0,-8/3), (-4/3,0), and (-4/3,-8/3)

Example

Solution cont'd: To find the nature of these extreme points, we have to use the sufficiency conditions. The second order partial derivatives of *f* are given by:

$$\frac{\partial^2 f}{\partial {x_1}^2} = 6x_1 + 4 \qquad \frac{\partial^2 f}{\partial {x_2}^2} = 6x_2 + 8 \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$$

The Hessian matrix of *f* is given by:

$$\mathbf{J} = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$$

Example

Solution cont'd:

$$\mathbf{J} = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$$

If $J_1 = |6x_1 + 4|$ and $\mathbf{J_2} = \begin{vmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{vmatrix}$, the values of J_1 and J_2 and

the nature of the extreme point are as given in the next slide:

Example

Point **X**	Value of $\mathbf{J}_1$	Value of $\mathbf{J}_2$	Nature of **J**	Nature of **X**	$f(\mathbf{X})$
(0,0)	+4	+32	Positive definite	Relative minimum	6
(0,-8/3)	+4	-32	Indefinite	Saddle point	418/27
(-4/3,0)	-4	-32	Indefinite	Saddle point	194/27
(-4/3,-8/3)	-4	+32	Negative definite	Relative maximum	50/3

Multivariable optimization with equality constraints

- **Problem statement:**

  Minimize $f = f(\mathbf{X})$ subject to $g_j(\mathbf{X})=0$, $j=1,2,.....,m$ where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$$

  Here m is less than or equal to n, otherwise the problem becomes overdefined and, in general, there will be no solution.

- **Solution:**
  - **Solution by direct substitution**
  - **Solution by the method of constrained variation** **(not covered)**
  - **Solution by the method of Lagrange multipliers**

Solution by direct substitution

For a problem with n variables and m equality constraints:

- Solve the m equality constraints and express any set of m variables in terms of the remaining n-m variables

- Substitute these expressions into the original objective function, the result is 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 by using the unconstrained optimization techniques.

Solution by direct substitution

- Simple in theory
- Not convenient from a practical point of view as the constraint equations will be nonlinear for most of the problems
- Suitable only for simple problems

Example

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:

$$f(x_1, x_2, x_3) = 8x_1x_2x_3$$

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

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

Example

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:

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

Thus, the objective function becomes:

$$f(x_1,x_2)=8x_1x_2(1-x_1^2-x_2^2)^{1/2}$$

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

Example

The necessary conditions for the maximum of f give:

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

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

which can be simplified as:

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

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

From which it follows that $x_1^* = x_2^* = 1/\sqrt{3}$ and hence $x_3^* = 1/\sqrt{3}$

Example

This solution gives the maximum volume of the box as:

$$f_{\max} = \frac{8}{3\sqrt{3}}$$

To find whether the solution found corresponds to a maximum or minimum, we apply the sufficiency conditions to $f(x_1,x_2)$ of the equation $f(x_1,x_2)=8x_1x_2(1-x_1^2-x_2^2)^{1/2}$. The second order partial derivatives of f at (x_1^*,x_2^*) are given by:

$$\frac{\partial^2 f}{\partial x_1^2} = -\frac{8x_1x_2}{(1-x_1^2-x_2^2)^{1/2}} - \frac{8x_2}{1-x_1^2-x_2^2}\left[\frac{x_1^3}{(1-x_1^2-x_2^2)^{1/2}} + 2x_1(1-x_1^2-x_2^2)^{1/2}\right]$$

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

Example

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

$$\frac{\partial^2 f}{\partial x_2^2} = -\frac{8x_1x_2}{(1-x_1^2-x_2^2)^{1/2}} - \frac{8x_1}{1-x_1^2-x_2^2}\left[\frac{x_2^3}{(1-x_1^2-x_2^2)^{1/2}} + 2x_2(1-x_1^2-x_2^2)^{1/2}\right]$$

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

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 8(1-x_1^2-x_2^2)^{1/2} - \frac{8x_2^2}{(1-x_1^2-x_2^2)^{1/2}}$$

$$-\frac{8x_1^2}{1-x_1^2-x_2^2}[(1-x_1^2-x_2^2)^{1/2} + \frac{x_2^2}{(1-x_1^2-x_2^2)^{1/2}}]$$

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

Example

Since

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

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.

Solution by Lagrange multipliers

Problem with two variables and one constraint:

Minimize $f\,(x_1,x_2)$

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

For this problem, the necessary condition was found to be:

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

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

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

Solution by Lagrange multipliers

Problem with two variables and one constraint:

Necessary conditions for the point (x_1*,x_2*) to be an extreme point
The problem can be rewritten as:

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

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

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

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

Solution by Lagrange multipliers

Problem with two variables and one constraint:

- The derivation of the necessary conditions by the method of Lagrange multipliers requires that at least one of the partial derivatives of $g(x_1,x_2)$ be nonzero at an extreme point.

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

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) = 0$$

Solution by Lagrange multipliers

Problem with two variables and one constraint:

- 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:

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

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

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

Example

Example: Find the solution using the Lagrange multiplier method.

Minimize $f(x,y) = kx^{-1}y^{-2}$

subject to

$g(x,y) = x^2 + y^2 - a^2 = 0$

Solution

The Lagrange function is

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y) = kx^{-1}y^{-2} + \lambda(x^2 + y^2 - a^2)$$

The necessary conditions for the minimum of $f(x,y)$

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

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

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

Example

Solution cont'd

which yield:

$$2\lambda = \frac{k}{x^3 y^2} = \frac{2k}{xy^4}$$

from which the relation $x^* = (1/\sqrt{2})y^*$ can be obtained. This relation along with

$\frac{\partial L}{\partial \lambda} = x^2 + y^2 + a^2 = 0$ gives the optimum solution as :

$$\mathrm{x}^* = \frac{\mathrm{a}}{\sqrt{3}} \text{ and } y^* = \sqrt{2}\,\frac{\mathrm{a}}{\sqrt{3}}$$

Solution by Lagrange multipliers

Necessary conditions for a general problem:

Minimize $f(\mathbf{X})$

subject to

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

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

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

Solution by Lagrange multipliers

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 corresponds to the solution of the original problem are given by:

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

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

The above equations represent $n+m$ equations in terms of the $n+m$ unknowns, x_i and λ_j

Solution by Lagrange multipliers

The solution:

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

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

Solution by Lagrange multipliers

Sufficient Condition

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

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

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

If

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \partial x_j} (\mathbf{X^*}, \lambda^*) dx_i dx_j$$

is negative for all choices of the admissable variations dx_i, $\mathbf{X^*}$ will be a constrained maximum of $f(\mathbf{X})$

Solution by Lagrange multipliers

A necessary condition for the quadratic form Q to be positive (negative) definite for all admissable variations $d\mathbf{X}$ is that each root of the polynomial z_i, defined by the following determinantal equation, be positive (negative):

$$\begin{vmatrix} L_{11}-z & L_{12} & L_{13} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22}-z & L_{23} & \cdots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn}-z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0 \tag{2.44}$$

where

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \, \partial x_j}(\mathbf{X}^*, \lambda^*) \tag{2.45}$$

$$g_{ij} = \frac{\partial g_i}{\partial x_j}(\mathbf{X}^*) \tag{2.46}$$

- The determinantal equation, on expansion, leads to an $(n\text{-}m)$th-order polynomial in z. If some of the roots of this polynomial are positive while the others are negative, the point $\mathbf{X}^*$ is not an extreme point.

Example 1

Find the dimensions of a cylindirical tin (with top and bottom) made up of sheet metal to maximize its volume such that the total surface area is equal to $A_0=24\pi$.

Solution

If x_1 and x_2 denote the radius of the base and length of the tin, respectively, the problem can be stated as:

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

subject to

$$2\pi x_1^2 + 2\pi x_1 x_2 = A_0 = 24\pi$$

Example 1

Solution

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

subject to $2\pi x_1^2 + 2\pi x_1 x_2 = A_0 = 24\pi$

The Lagrange function is:

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

and the necessary conditions for the maximum of f give:

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

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

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0$$

Example 1

Solution

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

that is,

$$x_1 = \frac{1}{2} x_2$$

The above equations give the desired solution as:

$$x_1^* = \left(\frac{A_0}{6\pi}\right)^{1/2}, \quad x_2^* = \left(\frac{2A_0}{3\pi}\right)^{1/2}, \text{and } \lambda^* = -\left(\frac{A_0}{24\pi}\right)^{1/2}$$

Example 1

Solution

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

If $A_0 = 24\pi$, the optimum solution becomes

$$x_1^* = 2, \quad x_2^* = 4, \quad \lambda^* = -1, \text{ and } f^* = 16\pi$$

To see that this solution really corresponds to the maximum of f, we apply the sufficiency condition of equation

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22} - z & L_{23} & \cdots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn} - z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0 \tag{2.44}$$

where

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \, \partial x_j} (\mathbf{X}^*, \boldsymbol{\lambda}^*) \tag{2.45}$$

$$g_{ij} = \frac{\partial g_i}{\partial x_j} (\mathbf{X}^*) \tag{2.46}$$

Example 1

Solution

In this case:

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

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

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

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

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

Example 1

Solution

Thus, equation

$$\begin{vmatrix} L_{11}-z & L_{12} & L_{13} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22}-z & L_{23} & \cdots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn}-z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0 \tag{2.44}$$

where

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \, \partial x_j}(\mathbf{X}^*, \boldsymbol{\lambda}^*) \tag{2.45}$$

$$g_{ij} = \frac{\partial g_i}{\partial x_j}(\mathbf{X}^*) \tag{2.46}$$

becomes

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

Example 1

Solution

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.

Example 2

Find the maximum of the function $f(\mathbf{X}) = 2x_1 + x_2 + 10$ subject to $g(\mathbf{X}) = x_1 + 2x_2^2 = 3$ using the Lagrange multiplier method. Also find the effect of changing the right-hand side of the constraint on the optimum value of f.

Solution

The Lagrange function is given by:

$$L(\mathbf{X}, \lambda) = 2x_1 + x_2 + 10 + \lambda(3 - x_1 - 2x_2^2)$$

The necessary conditions for the solution of the problem are:

$$\frac{\partial L}{\partial x_1} = 2 - \lambda = 0$$

$$\frac{\partial L}{\partial x_2} = 1 - 4\lambda x_2 = 0$$

$$\frac{\partial L}{\partial \lambda} = 3 - x_1 - 2x_2^2$$

Example 2

Solution

The solution of the equation is:

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \end{Bmatrix} = \begin{Bmatrix} 2.97 \\ 0.13 \end{Bmatrix}$$

$$\lambda^* = 2$$

The application of the sufficiency condition yields:

$$\begin{vmatrix} L_{11} - z & L_{12} & g_{11} \\ L_{21} & L_{22} - z & g_{12} \\ g_{11} & g_{12} & 0 \end{vmatrix} = 0$$

$$\begin{vmatrix} -z & 0 & -1 \\ 0 & -4\lambda - z & -4x_2 \\ -1 & -4x_2 & 0 \end{vmatrix} = \begin{vmatrix} -z & 0 & -1 \\ 0 & -8 - z & -0.52 \\ -1 & -0.52 & 0 \end{vmatrix} = 0$$

Example 2

Solution

$$0.2704z + 8 + z = 0$$

$$z = -6.2972$$

Hence **X*** will be a maximum of f with $f^* = f(\mathbf{X}^*) = 16.07$

Example (Text Book)

Use the Lagrange multiplier method to find the minimum of the function

$$f(x_1, x_2) = x_1 x_2^2$$

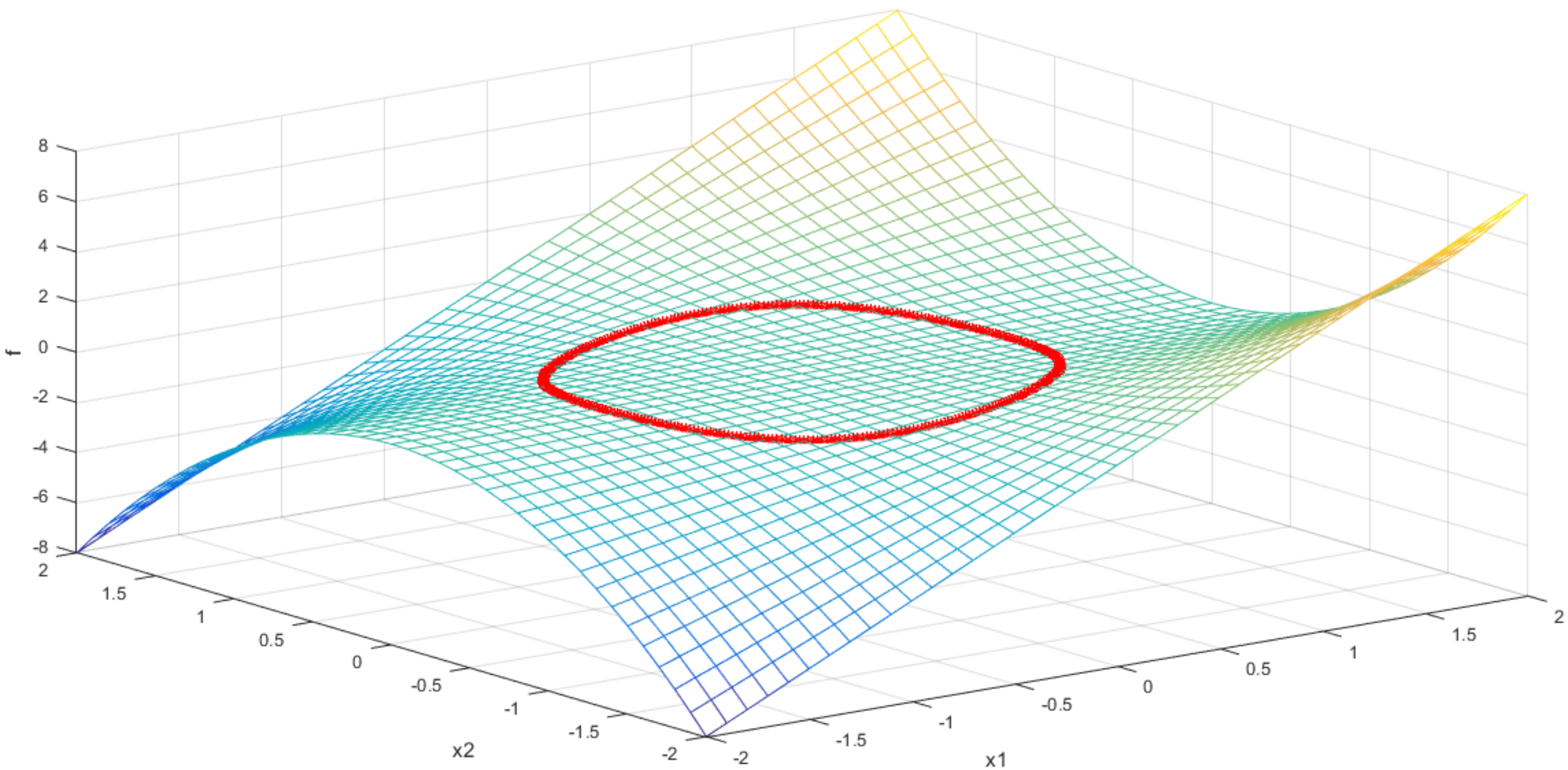
(see Fig. 2.9), subject to the constraint

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

Solution Introducing the Lagrange multiplier λ, we can form the function L as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2) = x_1 x_2^2 + \lambda \left(x_1^2 + x_2^2 - 1 \right).$$

Example (Text Book) Visualization



Example (Text Book) Conditions

Solution Introducing the Lagrange multiplier λ, we can form the function L as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2) = x_1 x_2^2 + \lambda \left(x_1^2 + x_2^2 - 1 \right).$$

Thus, stationary points of L occur where the following three equations hold:

$$\frac{\partial L}{\partial x_1} = x_2^2 + 2\lambda x_1 = 0,$$

Example (Text Book) Conditions

$$\frac{\partial L}{\partial x_2} = 2x_1x_2 + 2\lambda x_2 = 0, \tag{2.59}$$

$$\frac{\partial L}{\partial \lambda} = x_1^2 + x_2^2 - 1 = 0. \tag{2.60}$$

From eqn (2.59) it follows that either $x_2 = 0$ or $\lambda = -x_1$. If $x_2 = 0$, the constraint equation gives $x_1 = \pm 1$ and, therefore, $\lambda = 0$. On the other hand, if $\lambda = -x_1$ then eqn (2.58) becomes

$$x_2^2 - 2x_1^2 = 0, \tag{2.61}$$

so that, in combination with the constraint equation, we obtain $x_1 = \pm 1/\sqrt{3}$, $x_2 = \pm\sqrt{2/3}$. Hence, there are six critical points, for which the function f takes the value $f(\pm 1, 0) = 0$, $f(1/\sqrt{3}, \pm\sqrt{2/3}) = 2/(3\sqrt{3})$, $f(-1/\sqrt{3}, \pm\sqrt{2/3}) = -2/(3\sqrt{3})$. Thus, the minima occur at the points $(-1/\sqrt{3}, \pm\sqrt{2/3})^T$. ■

Example (Text Book) Visualization