Chapter 4: Basics of Unconstrained Optimization Problem.

OL Sela

Optimization

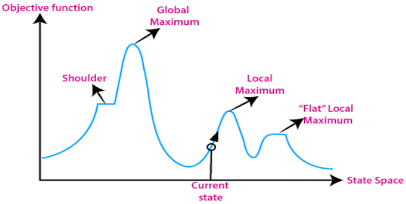
July 8, 2023

Outline

Outline

Brief Historical Reference

⋆ 17th century: Leonhard Euler (1707–1783) ,the problem of finding extreme value serve as one of motivation in the invention of differential calculus.



Outline

Introduction of Problem

In this chapter we consider the optimization problem ,
minimize

$$f(x)$$

subject to

$$x \in S$$

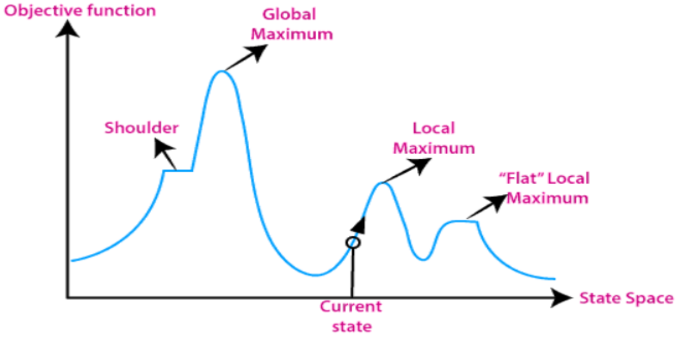
- The function $f : \mathbb{R}^n \to \mathbb{R}$ that we wish to minimize is a real-valued function called the objective function or cost function.

Introduction of Problem

- The vector x is an n-vector of independent variables: $x = [x_1, x_2, ..., x_n]^T \in \mathbb{R}^n$.
- The variable $x_1, x_2, ..., x_n$ are often referred to as decision variables.
- The set $S \subseteq \mathbb{R}^n$ called the constraint set or feasible set.
- If $S = \mathbb{R}^n$,the problem above is called Unconstrained optimization.

Outline

What is Global and Local Optimization?



Global and Local Optimization

Definition : Global minimum and maximum

Let $f : S \to \mathbb{R}$ be defined on a set $S \subseteq \mathbb{R}^n$.Then,

- $x^\star$ is called a global minimum point of f over S if $f(x) \geq f(x^\star)$ for any $x \in S$,
- $x^\star$ is called a strict global minimum point of f over S if $f(x) > f(x^\star)$ for any $x^\star \neq x \in S$,

Global and Local Optimization

Definition : Global minimum and maximum

- $x^\star$ is called a global maximum point off over S if $f(x) \leq f(x^\star)$ for any $x \in S$,
- $x^\star$ is called a strict global maximum point off over S if $f(x) < f(x^\star)$ for any $x^\star \neq x \in S$,

- The set S on which the optimization of f is performed is also called the feasible set,
- Any point $x \in S$ is called a feasible solution.

Global and Local Optimization

- A vector $x^\star \in S$ is called a global optimum of f over S if it is either a global minimum or a global maximum.
- The maximal value of f over S is defined as the supremum of $f$over S:

$$\max\{f(x) : x \in S\} = \sup\{f(x) : x \in S\}.$$

- If $x^\star \in S$ is a global maximum of f over S, then the maximum value of f over S is $f(x^\star)$. Similarly the minimal value of f over S is the infimum off over S,

$$\min\{f(x) : x \in S\} = inf\{f(x) : x \in S\}.$$

1. The set of all global minimizers of f over S is denoted by

$$argmin\{f(x) : x \in S\}$$

2. The set of all global maximizers of f over S is denoted by

$$argmax\{f(x) : x \in S\}$$

Example 1

Consider the function $f : \mathbb{R} \to \mathbb{R}$ define by

$$f(x) = (x+1)^2 + 3$$

What is $argmin\{f(x) : x \in S\}$,$argmax\{f(x) : x \in S\}$?

Example

Example 2

Consider the two-dimensional linear function $f(x, y) = x + y$ defined over the unit ball

$$\mathrm{S} = \mathrm{B}[0, 1] = \{(\mathrm{x}, \mathrm{y})^T : x^2 + y^2 \leq 1\}$$

Can we find maximal or minimal vale of f ??

Example 3

Consider the two-dimensional function

$$f(x, y) = \frac{x + y}{x^2 + y^2 + 1}$$

defined over the entire space $\mathbb{R}^2$

. The contour and surface plots of the function are given . This function has two optima points:

- a global maximizer $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$
- a global minimizer $(x, y) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$

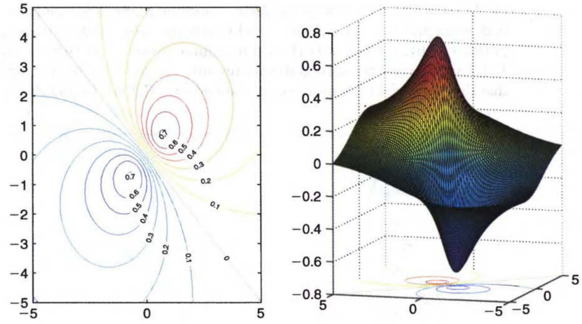


Figure 2.1. *Contour and surface plots of* $f(x,y) = \frac{x+y}{x^2+y^2+1}$.

- The maximal value of the function is $\frac{1}{\sqrt{2}}$.
- and the minimal value is $-\frac{1}{\sqrt{2}}$.

Local minima and maxima

Definition: Local minima and maxima

Let $f : S \to \mathbb{R}$ be defined on a set $S \subseteq \mathbb{R}^n$. Then

- $x^\star \in S$ is called a local minimum point of f over S if there exists $r > 0$ for which $f(x^\star) \leq f(x)$ for any $x \in S \cap B(x^\star, r)$,
- $x^\star \in S$ is called a strict local minimum point of f over S if there exists $r > 0$ for which $f(x^\star) < f(x)$ for any $x \in S \cap B(x^\star, r)$,
- $x^\star \in S$ is called a local maximum point of f over S if there exists $r > 0$ for which $f(x^\star) \geq f(x)$ for any $x \in S \cap B(x^\star, r)$,
- $x^\star \in S$ is called a strict local maximum point of f over S if there exists $r > 0$ for which $f(x^\star) > f(x)$ for any $x \in S \cap B(x^\star, r)$,

Of course, a global minimum (maximum) point is also a local minimum (maximum) point. As with global minimum and maximum points, we will also use the terminology local minimizer and local maximizer for local minimum and maximum points, respectively.

Example 4

Consider the one-dimensional function

$$f(x) = \begin{cases} (x-1)^2+2, & -1 \le x \le 1, \\ 2, & 1 \le x \le 2, \\ -(x-20^2, & 2 \le x \le 2.5, \\ (x-3)^2+1.5, & 2.5 \le x \le 4, \\ -(x-5)^2+3.5, & 4 \le x \le 6, \\ -2x+14.5, & 6 \le x \le 6.5, \\ 2x-11.5, & 6.5 \le x \le 8 \end{cases}$$

described in Figure 2.2 and defined over the interval $[-1, 8]$.

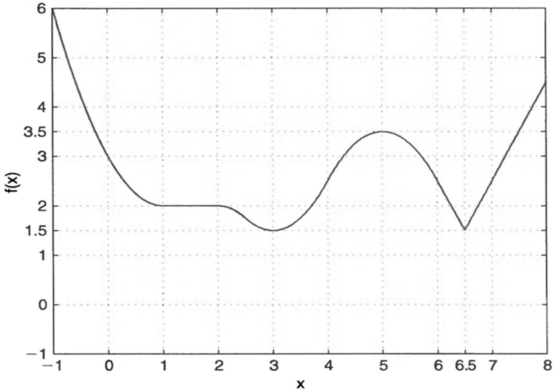


Figure 2.2. *Local and global optimum points of a one-dimensional function.*

- The point $x = -1$ is a strict global maximum point.
- The point $x = 1$ is a nonstrict local minimum point.
- All the points in the interval $(1, 2)$ are nonstrict local minimum points as well as nonstrict local maximum points.
- The point $x = 2$ is a local maximum point.
- The point $x = 3$ is a strict local minimum, and a non-strict global minimum point.
- The point $x = 5$ is a strict local maximum .
- $x = 6.5$ is a strict local minimum, which is a nonstrict global minimum point.
- $x = 8$ is a strict local maximum point.
  Note that, as already mentioned, $x = 3$ and $x = 6.5$ are both global minimum points of the function,

First Order Optimality Condition

Theorem:
first order optimality condition for local optima points

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$ .Suppose that $x^\star \in int(U)$ is a local optimum point and that all the partial derivatives of f exist at $x^\star$ .Then $\nabla f(x^\star) = 0$.

Proof: See in references, author:Amir Beck.
Noted: $\nabla f(x)$ is called gradient of f defined by

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right]^T$$

Stationary point

Definition 5

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$. Suppose that $x^\star \in int(U)$ and that f is differentiable over some neighborhood of $x^\star$ .Then $x^\star$ is called a stationary point of f if $\nabla f(x^\star) = 0$

Example

Example 6

Consider the one-dimensional quartic function

$$f(x) = 3x^4 - 20x^3 + 42x^2 - 36x$$

- Find all its stationary points.
- Find its local and global optima points over $\mathbb{R}$.

Classification of Matrices

In order to be able to characterize the second order optimality conditions, which are expressed via the Hessian matrix, the notion of "positive definiteness" must be defined.

Definition 7

1. A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called positive semidefinite, denoted by $A \geq 0$,if $x^T Ax \geq 0$ for every $x \in \mathbb{R}^n$.
2. A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called positive definite, denoted by $A > 0$,if $x^T Ax > 0$ for every $0 \neq x \in \mathbb{R}^n$.

Example 8

Consider the matrix

1. $$A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

2. $$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

3. $$B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Question: Positive semidefinite or positive definite?

Lemma 9

1. *Let* $A \in \mathbb{R}^{n\times n}$ *a positive definite matrix. Then the diagonal elements of* A *are positive.*
2. *Let* $A \in \mathbb{R}^{n\times n}$ *a positive semidefinite matrix. Then the diagonal elements of* A *are non-negative.*

Definition 10

1. A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called negative semidefinite, denoted by $A \preceq 0$,if $x^T A x \leq 0$ for every $x \in \mathbb{R}^n$.
2. A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called negative definite, denoted by $A \prec 0$,if $x^T A x < 0$ for every $0 \neq x \in \mathbb{R}^n$.
3. A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called indefinite,if there exist $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ such that $x^T A x > 0$ and $y^T A y < 0$

Lemma

1. Let A be a negative definite matrix. Then the diagonal elements of A are negative.
2. Let A be a negative semidefinite matrix. Then the diagonal elements of A are nonpositive.

When the diagonal of a matrix contains both positive and negative elements, then the matrix is indefinite. The reverse claim is not correct

lemma

Let A be a symmetric $n \times n$ matrix. If there exist positive and negative elements in the diagonal of A, then A is indefinite.

Second Order Optimality Conditions

We begin by stating the necessary second order optimality condition.

Theorem: necessary second order optimality conditions

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$. Suppose that f is twice continuously differentiable over U and that $x^\star$ is a stationary point. Then the following hold.

1. If $x^\star$ is a local minimum point of f over U, then $\nabla^2 f(x^\star) \geq 0$.
2. If $x^\star$ is a local maximum point of f over U, then $\nabla^2 f(x^\star) \leq 0$.

Second Order Optimality Conditions

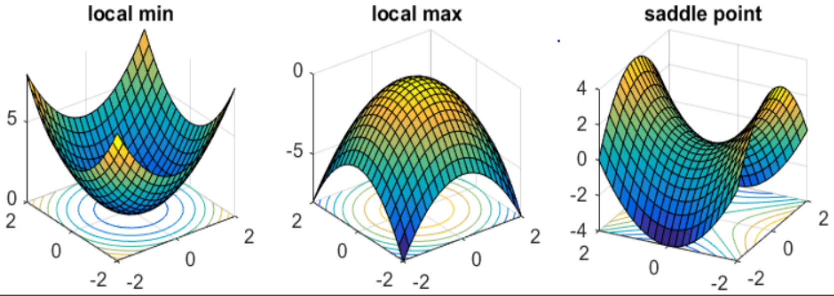
The latter result is a necessary condition for local optimality. The next theorem states a sufficient condition for strict local optimality.

Theorem:(sufficient second order optimality condition

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$. Suppose that f is twice continuously differentiable over U and that $x^\star$ is a stationary point. Then the following hold.

1. If $\nabla^2 f(x^\star) \geq 0$.Then $x^\star$ is a local minimum point of f over U, .
2. If $\nabla^2 f(x^\star) \leq 0$.Then$x^\star$ is a local maximum point of f over U.

Saddle point



Saddle point

Definition 11

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$. Suppose that f is continuously differentiable over U. A stationary point $x^\star$ is called a saddle point of f over U if it is neither a local minimum point nor a local maximum point off over U.

Theorem

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$. Suppose that f is twice continuously differentiable over U and that $x^\star$ is a stationary point. If $\nabla^2 f(x^\star)$ is an indefinite matrix,then $x^\star$ is a saddle point of f over U.

Example

Example 12

Consider the function

$$f(x_1, x_2) = 2x_1^3 + 3x_2^2 + 3x_1^2 x_2 - 24x_2$$

over $\mathbb{R}^2$ Find all the stationary points off over $\mathbb{R}^2$ and classify them.

Example

Example 13

Consider the function

$$f(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + (x_2^2 - 1)^2$$

over $\mathbb{R}^2$ Find all the stationary points of f over $\mathbb{R}^2$ and classify them.

Example 14

Consider the two-dimensional function

$$f(x, y) = \frac{x + y}{x^2 + y^2 + 1}$$

defined over the entire space $\mathbb{R}^2$

. The contour and surface plots of the function are given . This function has two optima points:

- a global maximizer $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$
- a global minimizer $(x, y) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$

Example

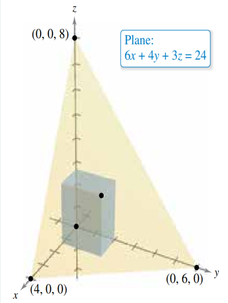
Example 15

Consider the function $f(x, y) = 2x^2 + y^2 + 8x - 6y + 20$

1. Find of all stationary point of f.
2. What is local min or Local max?

Example 16

A rectangular box is resting on the xy-plane with one vertex at the origin. The opposite vertex lies in the plane $6x + 4y + 3z = 24$, . Find the maximum volume of the box.



Applied Optimization Problems

Example 17

A manufacturer determines that the profit P (in dollars) obtained by producing and selling x units of Product 1 and y units of Product 2 is approximated by the model,

$$P(x,y) = 8x + 10y - (0.001)(x^2 + xy + y^2) - 10000$$

1. Find the production level that produces a maximum profit.
2. What is the maximum profit?

Maximum Revenue

Example 18

A company manufactures running shoes and basketball shoes. The total revenue (in thousands of dollars) from x_1 units of running shoes and x_2 units of basketball shoes is

$$R = -5x_1^2 - 8x_2^2 - 2x_1x_2 + 42x_1 + 102x_2$$

where x_1 and x_2 are in thousands of units. Find x_1 and x_2 so as to maximize the revenue.

Example

Example 19

Consider the problem
minimize

$$x_2^2 + 0.5x_2^2 + 3x_2 + 4.5$$

subject to

$$x_1 \geq 0, x_2 \geq 0$$

1. Is the first order optimality condition for a local minimizer satisfied at $x = [1, 3]^T$?
2. Is the first order optimality condition for a local minimizer satisfied at $x = [0, 3]^T$?
3. Is the first order optimality condition for a local minimizer satisfied at $x = [1, 0]$?

Outline

Quadratic functions

Quadratic functions are an important class of functions that are useful in the modeling of many optimization problems. We will now define and derive some of the basic results related to this important class of functions.

Definition 20

A quadratic function over $\mathbb{R}^n$ is a function of the form

$$f(x) = x^T Ax + 2b^T x + c, \tag{1}$$

where $A \in \mathbb{R}^{n \times n}$ is symmetric, $b \in \mathbb{R}^n$, $x \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Quadratic functions

We will frequently refer to the matrix A in (1) as the matrix associated with the quadratic function f. The gradient and Hessian of a quadratic function have simple analytic formulas:

$$\nabla f(x) = 2Ax + 2b \tag{2}$$

$$\nabla^2 f(x) = 2A \tag{3}$$

By the above formulas we can deduce several important properties of quadratic functions, which are associated with their stationary points.

Lemma 21

Let $f(x) = x^T Ax + 2b^T x + c$ *where* $A \in \mathbb{R}^{n\times n}$ *is symmetric,* $b \in \mathbb{R}^n$ *,* $x \in \mathbb{R}^n$ *and* $c \in \mathbb{R}$. *Then*

1. x *is a stationary point off if and only if* $Ax = -b$,
2. *If* $A \geq 0$ *then* x *is a global minimum point of* f *if and only if* $Ax = -b$,
3. *If* $A > 0$ *,then* $x = -A^{-1}b$ *is a strict global minimum point of* f.

Proof.

Question??