Operations Research
Chapter 3: Existence and Optimality Conditions for Minimizers

Dr. Kareem Elgindy

Lecturer
Mathematics Department, Faculty of Science
Assiut University

10 March, 2016

Outline

Outline

Main Questions

In this chapter we shall try to answer the following TWO questions:

1. **How Do We Ensure the Existence of a Minimum?**
2. **What are the Necessary and Sufficient Conditions for Local Minimizers?**

Main Questions

In this chapter we shall try to answer the following TWO questions:

1. **How Do We Ensure the Existence of a Minimum?**
2. **What are the Necessary and Sufficient Conditions for Local Minimizers?**

Outline

Existence of a Minimum

- Generally one attempts to solve a minimization problem without knowing in advance if a minimum even exists.
  - **Even when the DOF analysis indicates that a problem requires optimization, the problem may not admit an optimal solution.**

In certain cases **we can ensure existence of a minimum, even though we may not know how to find it**.

- **Weierstrass Theorem** guarantees the existence of a minimum when certain conditions are satisfied.

Existence of a Minimum

To use the theorem, we must first understand the meaning of a **closed** and **bounded set**.

Definition 1 (**Closed Set**)

A set $\boldsymbol{\Omega}$ is closed if:

- it includes all of its boundary/limit points, or,
- every sequence of points in it has a subsequence that converges to a point in the set.

- For instance, any closed interval $\boldsymbol{S_1} = [a, b]$ is a closed set while any open interval $\boldsymbol{S_2} = (a, b)$ is an open set, since $\boldsymbol{S_2}$ does not include its boundary points a and b. Moreover, for $\boldsymbol{S_2} = (0, 1)$, say, the sequence $\{1/n\}_{n=2}^{\infty} \subset \boldsymbol{S_2}$, but $1/n \to 0 \notin \boldsymbol{S_2}$.

Existence of a Minimum

Corollary 1

We cannot have strict "$<$ type" inequality constraints in the formulation of the optimization problem if the feasible (constraint) set $\mathbf{\Omega}$ is closed.

Definition 2 (Bounded Set)

Let $n \in \mathbb{Z}^+$. A set $\mathbf{\Omega}$ is bounded if for any point $\boldsymbol{x} \in \mathbf{\Omega} \subset \mathbb{R}^n$,

$$\boldsymbol{x}^T \boldsymbol{x} \le c < \infty,$$

where c is a finite number.

For example:

- The sets $\mathbf{S_1} = (0, 2), \mathbf{S_2} = (0, 2]$, and $\mathbf{S_3} = \{(x, y)^T | x^2 + y^2 < 1\}$, are bounded.
- The sets $\mathbb{R}, \mathbb{Q}, \mathbb{N}, \mathbb{Z}$, of all real, rational, natural, and integer numbers are unbounded, because there is no such a number c.

Existence of a Minimum

Theorem 1 (Weierstrass Theorem– Existence of a Global Minimum)

*If $f(x)$ is **continuous** on a non-empty feasible set $\mathbf{\Omega}$ that is **closed** and **bounded** (**compact**), then $f(x)$ has a global minimum in $\mathbf{\Omega}$.*

Example 1

Consider the function $f(x) = -1/x$ defined on the set $\mathbf{\Omega} = \{x|0 < x \leq 1\} = (0, 1]$. Check the existence of a global minimum for the function.

Solution 1

*The feasible set $\mathbf{\Omega}$ **is not closed**, since it does not include the boundary point $x = 0$. Hence the conditions of the Weierstrass Theorem are not satisfied, and **the existence of a global minimum is not guaranteed**. Indeed, there is no point x^* satisfying $f(x^*) \leq f(x) \, \forall x \in \mathbf{\Omega} \backslash \{x^*\}$.*

Existence of a Minimum

Example 2

Consider the function $f(x) = -1/x$ defined on the set $\mathbf{\Omega} = \{x|0 \leq x \leq 1\} = [0, 1]$. Check the existence of a global minimum for the function.

Solution 2

The feasible set $\mathbf{\Omega}$ is closed and bounded. However, f is not defined at $x = 0$ (hence not continuous), so the conditions of the theorem are still not satisfied. Hence there is no guarantee of a global minimum for f in the set $\mathbf{\Omega}$.

Existence of a Minimum

Example 3

Consider the function $f(x) = -1/(x^2+1)$ defined on the set $\mathbf{\Omega} = \{x|0 \leq x \leq 1\} = [0,1]$. Check the existence of a global minimum for the function.

Solution 3

*The function f is **continuous** on a nonempty **closed and bounded (compact)** feasible set $\mathbf{\Omega}$, so the conditions of the theorem are satisfied and a global minimum for f in $\mathbf{\Omega}$ exists. Indeed, the point $x^* = 0$ satisfies $f(x^*) = -1 < f(x) \forall x \in \mathbf{\Omega} \backslash \{0\}$.*

Existence of a Minimum

Remark 1.

Note that when the conditions of the Weierstrass Theorem are satisfied, the existence of a global optimum is guaranteed. However, **it is important to realize that when they are not satisfied, a global solution may still exist**; that is, it is not an "if-and-only-if" theorem. In other words, the conditions of the Weierstrass Theorem are **sufficient** but **not necessary**.

Existence of a Minimum

Remark 2.

Weierstrass Theorem **does not provide a method for finding a global minimum point** even if its conditions are satisfied; it is only an existence theorem.

Outline

Conditions for Local Minimizers

- Given an optimization problem with constraint set Ω, and its boundary is denoted by $\partial\Omega$. In this section we derive conditions for a point x^* to be a local minimizer.

  A minimizer x^* may lie either in $\Omega\backslash\partial\Omega$ (the interior of Ω) or on the boundary $\partial\Omega$. To study these cases, we need to understand the concepts of:

  1. **the gradient vector** and **the Hessian matrix**,
  2. **feasible directions**,
  3. **Taylor's expansion**, and
  4. **directional derivatives**.

Outline

Conditions for Local Minimizers

First-Order Partial Derivatives. For a function $f(\boldsymbol{x})$ of n variables, the first-order partial derivatives are written as,

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i}, \quad i = 1, \ldots, n.$$

The n partial derivatives are usually arranged in a column vector known as the gradient of the function $f(\boldsymbol{x})$, and is denoted by $\nabla f(\boldsymbol{x})$ (or $\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}$); i.e.,

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

The gradient is a **vector-valued function** (also called **vector field**), since $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$, and it represents a generalization of the usual concept of the derivative of a single-variable function to functions of several variables.

Conditions for Local Minimizers

Geometrically, the gradient vector $\nabla f(\boldsymbol{x}^*)$ is **normal to the tangent plane** at the point x^*. Also, it points in the **direction of the greatest rate of increase of the function**.

Second-Order Partial Derivatives. Each component of the gradient vector can be differentiated again with respect to a variable to obtain the second-order partial derivatives for the function $f(\boldsymbol{x})$ written as:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_j}\right) = \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n.$$

Conditions for Local Minimizers

The n^2 partial derivatives can be arranged in a matrix known as the matrix of second partial derivatives of $f(\boldsymbol{x})$ (or simply the **Hessian matrix**[1]), and written as $\mathbf{H}(\mathbf{x})$ ($\nabla^2 f(\boldsymbol{x})$, $\frac{\partial^2 f(\boldsymbol{x})}{\partial \mathbf{x} \partial \mathbf{x}}$, or $[\frac{\partial^2 f}{\partial x_i \partial x_j}]_{n \times n}$):

$$\mathbf{H}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

[1] The Hessian matrix was developed in the 19th century by the German mathematician Ludwig Otto Hesse and later named after him.

Conditions for Local Minimizers

Remark 3.

If $f(\boldsymbol{x})$ is twice continuously differentiable, then the mixed/cross partial derivatives are equal; that is,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}, \quad i, j = 1, \ldots, n. \tag{1}$$

Therefore, the Hessian $\mathbf{H}(\mathbf{x})$ is always a symmetric matrix, i.e. $\mathbf{H}(\boldsymbol{x}) = \mathbf{H}^T(\boldsymbol{x})$.

Conditions for Local Minimizers

Example 4

Let $f(x_1, x_2) = 5x_1 + 8x_2 + x_1x_2 - x_1^2 - 2x_2^2$. Then,

$$\nabla^T f(\boldsymbol{x}) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x})\right]^T = [5 + x_2 - 2x_1, 8 + x_1 - 4x_2]^T,$$

and

$$\mathbf{H}(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}) \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -4 \end{bmatrix}.$$