Review of Calculus and Optimization Theory

Tapan Mitra

Intermediate Mathematical Economics I
(ECON 6170)
Fall 2011

Contents

Part I
Sets and Functions

1 Sets

A *set* is a primitive concept in mathematics. We think of a set as a collection of objects. The objects belonging to a set are called *elements* of the set. If S is a set, and x is an element belonging to S, we write $x \in S$. A set consisting of no elements is called an *empty set* and denoted by $\emptyset$.

If S and T are two sets, then the *union* of the two sets denoted by $S \cup T$ and defined as:

$$S \cup T = \{x : x \in S \text{ or } x \in T\}$$

The *intersection* of the two sets is denoted by $S \cap T$ and defined as:

$$S \cap T = \{x : x \in S \text{ and } x \in T\}$$

If S and T are two sets, and every element of S is also an element of T, we say that S is a subset of T, and write $S \subset T$. If $S \subset T$ and $T \subset S$, then the sets S and T are the same, and we write $S = T$.

If S and T are two sets such that $S \cap T = \emptyset$, we say that the sets S and T are *disjoint*.

If $S \subset T$, then the *complement* of S relative to T is the set of elements of T not in S. We write this as:

$$\sim S = \{x \in T : x \notin S\}$$

If S and T are two sets, the *difference* $T \sim S$ is the set of elements in T which are not in S :

$$T \sim S = \{x : x \in T \text{ and } x \notin S\}$$

When $S \subset T$, the difference $T \sim S$ coincides with the complement of S relative to T, which we have denoted by $\sim S$.

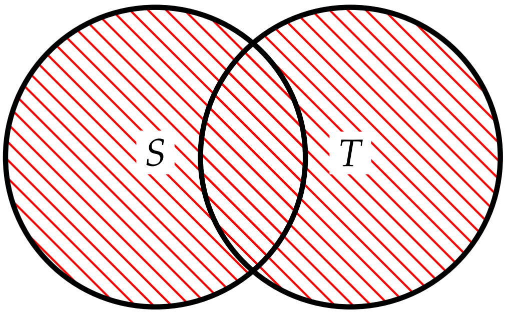
All of the concepts introduced above can be illustrated with *Venn Diagrams*. Such diagrams can be useful visual aids in understanding the various set theoretic operations.

If S and T are two sets, the Cartesian product of the two sets is denoted by $S \times T$ and defined as:

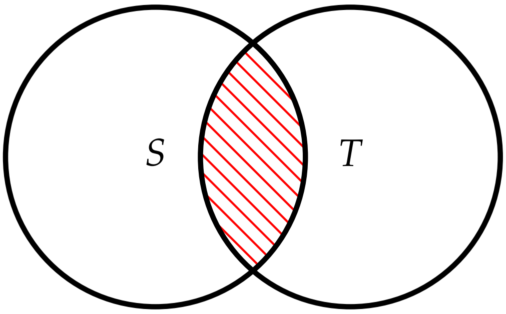
$$S \times T = \{(x, y) : x \in S \text{ and } y \in T\}$$

1.1 Sets of Real Numbers

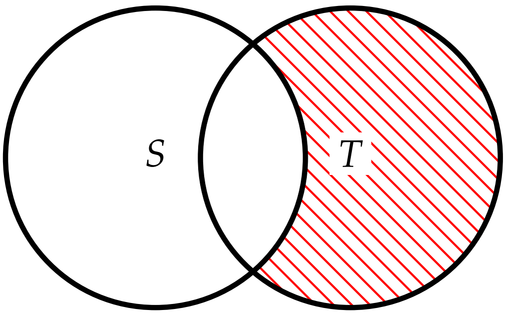
We will be concerned with the set of *real numbers*, denoted by $\mathbb{R}$, and subsets of this set. The set of real numbers is geometrically represented by the *real line.*



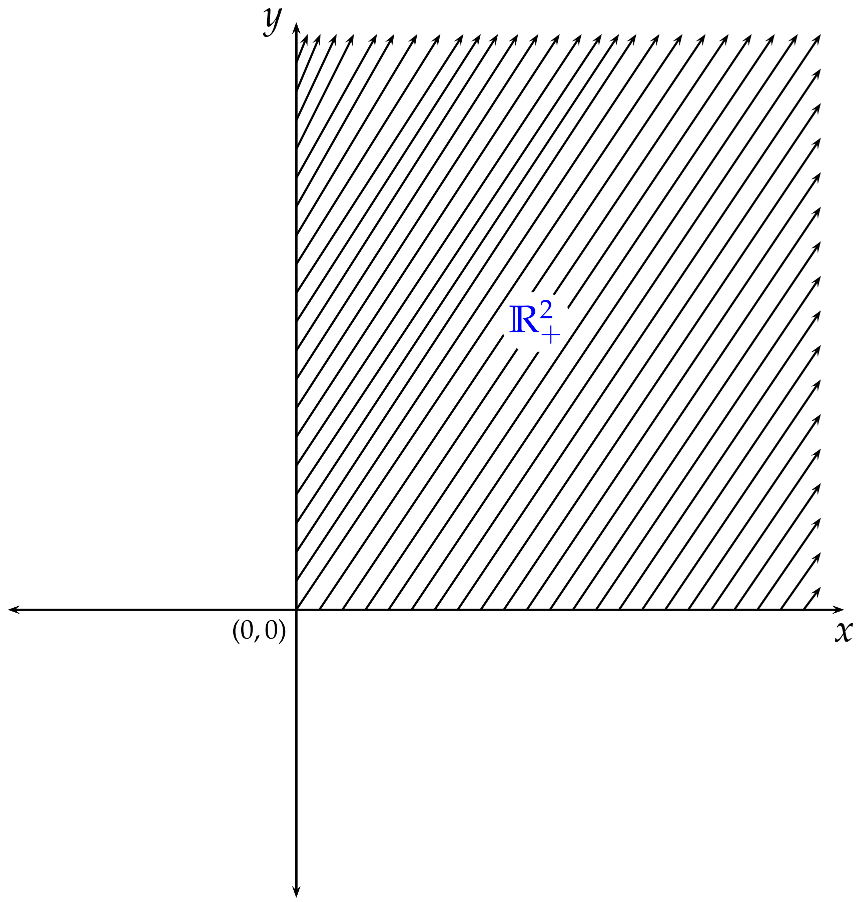
$S \cup T$



$S \cap T$



$T \sim S$


y
$\mathbb{R}^2_+$
$(0,0)$
x

The set of *non-negative* real numbers is denoted by $\mathbb{R}_+$ and defined as:

$$\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$$

The set of *positive* real numbers is denoted by $\mathbb{R}_{++}$ and defined as:

$$\mathbb{R}_{++} = \{x \in \mathbb{R} : x > 0\}$$

If $a, b \in \mathbb{R}$, with $a < b$, then the *open interval* (a, b) is the set defined by:

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}$$

and the *closed interval* $[a, b]$ is the set defined by:

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

The intervals $(a, b]$ and $[a, b)$ are analogously defined.

1.2 Sets of Real Vectors

The Cartesian product $\mathbb{R} \times \mathbb{R}$ is denoted by $\mathbb{R}^2$ and defined as:

$$\mathbb{R}^2 = \{(x_1, x_2) : x_1 \in \mathbb{R} \text{ and } x_2 \in \mathbb{R}\}$$

This set is the set of all real vectors with two co-ordinates. It is geometrically represented by the *real plane*.

The set of *non-negative* vectors in $\mathbb{R}^2$ is denoted by $\mathbb{R}^2_+$ and defined as:

$$\mathbb{R}^2_+ = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0 \text{ and } x_2 \geq 0\}$$

The set of *positive* vectors in $\mathbb{R}^2$ is denoted by $\mathbb{R}^2_{++}$ and defined as:

$$\mathbb{R}^2_{++} = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0 \text{ and } x_2 > 0\}$$

These concepts can be generalized to discuss real vectors with n co-ordinates, where n is any positive integer.

2 Functions

A function $f : S \to T$ is a rule which associates with each $x \in S$ a unique element $y \in T$. In this case, we denote the element $y \in T$ by $f(x)$. The set S is called the *domain* of the function of f. The *range* of f is denoted by $f(S)$ and defined as:

$$f(S) = \{y \in T : \text{there is some } x \in S, \text{ satisfying } f(x) = y\}$$

For example, if $S = (0, 1)$ and $T = \mathbb{R}$, and $f : S \to T$ is defined by:

$$f(x) = x + 2 \quad \text{for all } x \in S$$

then $f(S) = (2, 3)$.

Given a function $f : S \to T$, we define the *graph* of the function as the set:

$$G(f) = \{(x, y) \in S \times T : y = f(x)\}$$

If $f : S \to T$ is a function and $T \subset \mathbb{R}$ then f is called a *real-valued function*. If S is also a subset of $\mathbb{R}$, then f is a real valued function of a *real variable*. In this case, the set defined as the graph of f corresponds exactly to the usual graph of the function that we would draw on graph paper.

3 Limit of a Function

Let S be a subset of $\mathbb{R}$ and let f be a function from S to $\mathbb{R}$. If $c \in \mathbb{R}$, the expression:

$$\lim_{x \to c} f(x) = L$$

means that given any $\varepsilon > 0$, there is $\delta > 0$ such that:

$$x \in S \text{ and } 0 < |x - c| < \delta \ \text{ imply } |f(x) - L| < \varepsilon$$

It is worth noting two things about the definition. First, the point $c \in \mathbb{R}$ need not belong to S, the set on which the function is defined. Second, even if $c \in S$, the value of the function at the point c is of no significance in the concept of the limit because of the condition that $|x - c| > 0$.

Informally, the concept of the limit conveys the information that as x gets close to c, $f(x)$ gets close to L. The formal epsilon-delta definition conveys the same information more rigorously by specifying what we mean by "close to". Thus, it says that whenever x is delta-close to c, $f(x)$ is epsilon-close to L.

A function $f : S \to \mathbb{R}$ might not have a limit when $x \to c$, where $c \in \mathbb{R}$. This can happen because as x approaches c from above (that is, from the right on the real line) the function values approach L_1 and as x approaches c from below (that is, from the left on the real line) the function values approach L_2, and $L_1 \neq L_2$. Here is an example. Let $S = [0, 2]$ and let $f : S \to \mathbb{R}$ be defined as:

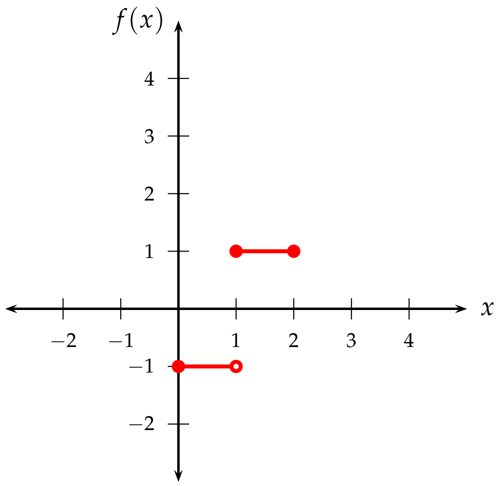
$$f(x) = \begin{cases} -1 & \text{for } x \in [0, 1) \\ 1 & \text{for } x \in [1, 2] \end{cases}$$

Then, $\lim_{x \to 1} f(x)$ does not exist, because $L_1 = 1$ while $L_2 = -1$.

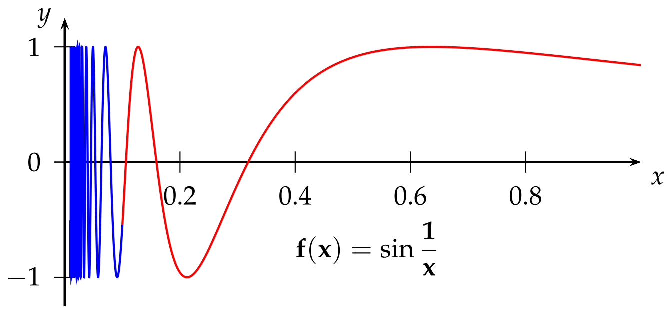
A function $f : S \to \mathbb{R}$ might also not have a limit when $x \to c$ (where $c \in \mathbb{R}$) because the function oscillates and therefore does not approach a definite value L when x approaches c. Here is an example. Let $S = (0, 1)$ and let $f : S \to \mathbb{R}$ be defined as:

$$f(x) = \sin(1/x)$$

In this case $\lim_{x \to 0} f(x)$ does not exist because the function takes on all values between 1 and -1 no matter how close $x \in S$ gets to 0.



$$f(x) = \begin{cases} -1 & \text{for } x \in [0,1) \\ 1 & \text{for } x \in [1,2] \end{cases}$$



The value of the function oscillates between 1 and -1 no matter how close x gets to 0.

4 Problems for Discussion

4.1 Sets

1. Let S be a set. Show that:
(a) $(S \cup \emptyset) = S$
(b) $(S \cap \emptyset) = \emptyset$

2. Let A and B be subsets of the set X. Interpret complement of all sets below to be relative to the set X. Show that:
(a) $\sim (A \cup B) = (\sim A) \cap (\sim B)$
(b) $\sim (A \cap B) = (\sim A) \cup (\sim B)$
(c) Illustrate the results in (a) and (b) with appropriate Venn Diagrams.

3. Let A and B be arbitrary sets. Define the set S by:

$$S = (A \cup B) \sim (A \cap B)$$

and the set T by:

$$T = (A \sim B) \cup (B \sim A)$$

(a) Show that $S = T$.
(b) Illustrate the result in (a) with an appropriate Venn Diagram.

4.2 Functions

4. Let $S = [0, 2]$ and let f be a function from S to $\mathbb{R}$ defined by:

$$f(x) = 2x - x^2 \text{ for all } x \in S$$

(a) Draw a graph of the function f.
(b) What is the range $f(S)$ of the function f? Explain.

5. Let $S = (0, 1)$ and let f be a function from S to $\mathbb{R}$ defined by:

$$f(x) = \ln x \text{ for all } x \in S$$

(a) Draw a graph of the function f.
(b) What is the range $f(S)$ of the function f? Explain.

6. Let $S = \mathbb{R}$ and let f be a function from S to $\mathbb{R}$ defined by:

$$f(x) = e^x \text{ for all } x \in S$$

(a) Draw a graph of the function f.
(b) What is the range $f(S)$ of the function f? Explain.

4.3 Limit of a Function

7. Let $S = \mathbb{R}$ and let f be a function from S to $\mathbb{R}$ defined by:

$$f(x) = 2x + 1 \text{ for all } x \in S$$

Use the epsilon-delta definition of a limit of a function to show that:

$$\lim_{x \to 1} f(x) = 3$$

[The important part of this exercise is to demonstrate, given the $\varepsilon > 0$ in the epsilon-delta definition, how to choose the $\delta > 0$].

8. Let $S = \mathbb{R}$ and let f be a function from S to $\mathbb{R}$ defined by:

$$f(x) = x^2 + 3 \text{ for all } x \in S$$

Use the epsilon-delta definition of a limit of a function to show that:

$$\lim_{x \to 1} f(x) = 4$$

9. Let $S = (1, 2)$ and let f be a function from S to $\mathbb{R}$ defined by:

$$f(x) = \frac{x^2 - 1}{x - 1} \text{ for all } x \in S$$

Use the epsilon-delta definition of a limit of a function to show that:

$$\lim_{x \to 1} f(x) = 2$$

Part II
Continuity

5 Continuous Functions

Let $S = [a, b]$ with $a < b$, and let f be a function from S to $\mathbb{R}$. If $c \in S$, then f is *continuous at c* if $\lim_{x \to c} f(x)$ exists and is equal to $f(c)$.

Note that the value of the function f at c is important in the concept of continuity, unlike in the concept of limit.

Consider the following example. Suppose $S = [-1, 1]$, and let $f : S \to \mathbb{R}$ be defined by:

$$f(x) = \begin{cases} |x| & \text{for } x \neq 0 \\ k & \text{for } x = 0 \end{cases} \tag{1}$$

Then, it can be checked that:

$$\lim_{x \to 0} f(x) = 0 \tag{2}$$

based simply on the information provided in the first line of (1), and therefore independent of what k is. But, if we want to check the continuity of f at $c = 0$, then it is clear that the information provided in the second line of (1) is crucial; f is continuous at $c = 0$ if and only if $k = 0$.

Based on the definition of continuity of a function f at a point c in its domain, we can define the concept of continuity of a function on a set. Let $S = [a, b]$ with $a < b$, and let f be a function from S to $\mathbb{R}$. Then f is *continuous on S* if f is continuous at every point $x \in S$.

5.1 Some Rules for Continuity

Let $S = [a, b]$ with $a < b$, and let f and g be functions from S to $\mathbb{R}$. Let $c \in S$. Then:

(i) If f and g are continuous at c, and $h : S \to \mathbb{R}$ is defined as $h(x) = f(x) + g(x)$ for $x \in S$, then h is continuous at c.

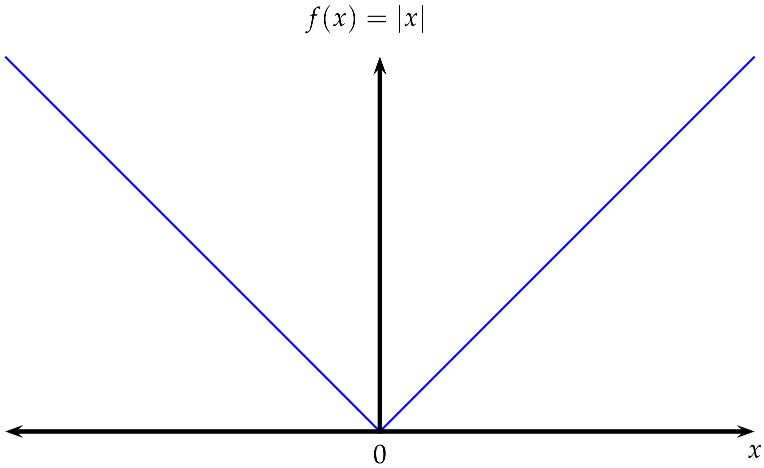
(ii) If f and g are continuous at c, and $h : S \to \mathbb{R}$ is defined as $h(x) = f(x)g(x)$ for $x \in S$, then h is continuous at c.

(iii) If f and g are continuous at c, and $h : S \to \mathbb{R}$ is defined as $h(x) = f(x)/g(x)$ for those $x \in S$ for which $g(x) \neq 0$, and $g(c) \neq 0$, then h is continuous at c.

These rules can be formally derived from the definition of the continuity of a function at a point in its domain. They are useful because they allow us to infer from the continuity of a simple function at a point, the continuity of more "complicated" functions at that point.

For example if $S = [a, b]$ and $f : S \to \mathbb{R}$ is defined by:

$$f(x) = b + ax \quad \text{for all } x \in S$$



Here $f(x)$ is a continuous function

where a, b are arbitrary constants, it is easy to check that f is continuous at every $c \in S$. It now follows by using rules (i) and (ii) above that any *polynomial function* $p : S \to \mathbb{R}$:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

where $a_0, a_1, a_2, ..., a_n$ are arbitrary constants, is continuous at every $c \in S$.

5.2 Intermediate Value Theorem

One of the useful implications of continuity of a function is conveyed by the following result.

Intermediate Value Theorem:

Let $S = [a, b]$ and let f be a continuous function from S to $\mathbb{R}$. Suppose $f(a) > f(b)$, and $z \in [f(b), f(a)]$. Then there is some $c \in S$, such that $f(c) = z$.

This means that *all* the values between $f(b)$ and $f(a)$ are attained by the function f. It does *not* mean that these are the only values attained by the function.

The theorem is especially useful in locating "roots" of polynomial equations. Here is a simple example. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by:

$$f(x) = 1 + 2x - 5x^5 \text{ for } x \in \mathbb{R} \tag{3}$$

Then, $f(0) = 1$ and $f(1) = -2$. Consequently, we can conclude that there is a root of the associated polynomial equation between 0 and 1. That is, there is $c \in (0, 1)$, such that $1 + 2c - 5c^5 = 0$, even if we don't know how to explicitly solve the quintic equation:

$$1 + 2x - 5x^5 = 0 \tag{4}$$

5.3 Weierstrass Theorem on the Maximum and Minimum of a Function

In the theory of optimization, it is sometimes useful to know that a maximum (or a minimum) of a function exists, without actually solving for it. The theorem of Weierstrass provides an easily applicable criterion for doing this.

Weierstrass Theorem:

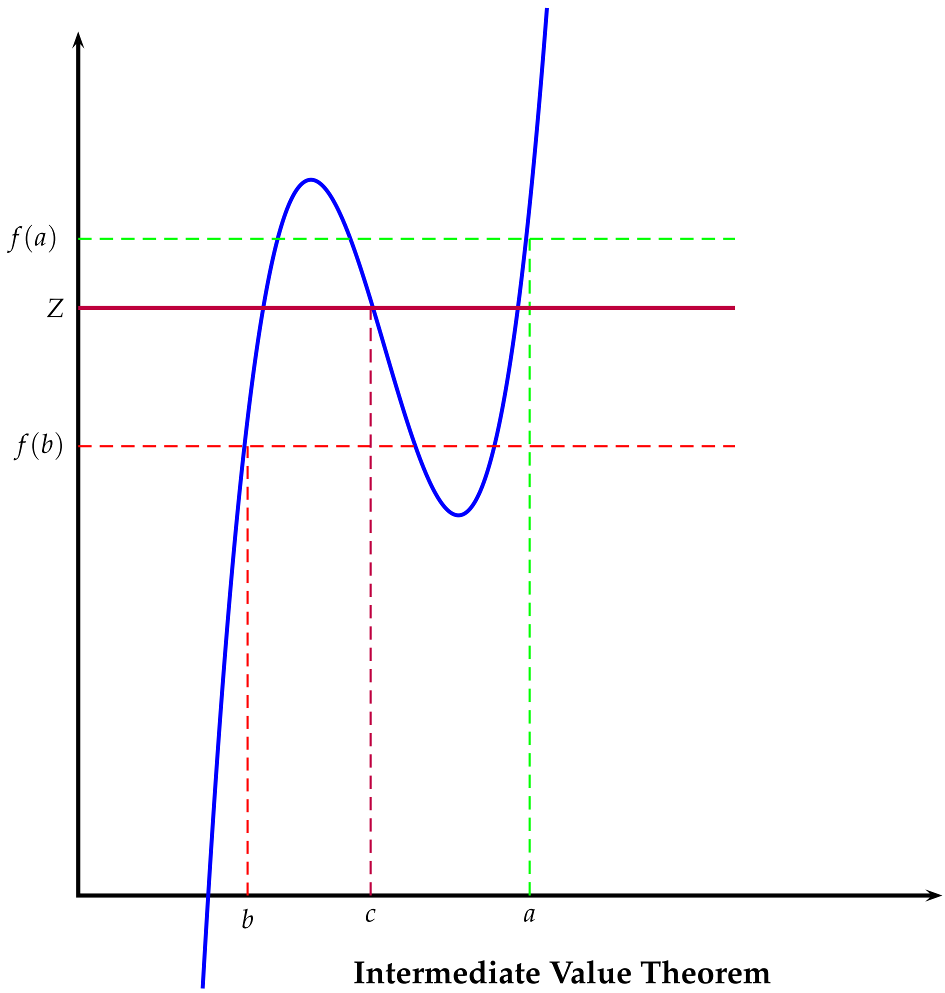
Let $S = [a, b]$ and let f be a continuous function from S to $\mathbb{R}$. Then there is some $c \in S$ such that:

$$f(x) \leq f(c) \text{ for all } x \in S \tag{5}$$

and some $c' \in S$ such that:

$$f(x) \geq f(c') \text{ for all } x \in S \tag{6}$$

Condition (5) says that c is a point in S at which f attains a maximum value on the set S. Similarly, (6) says that c' is a point in S at which f attains a minimum value on the set S.



Intermediate Value Theorem

As an example, consider $S = [0, 1]$, and $f : S \to \mathbb{R}$, where f is defined by the formula in (3). Then, the theorem ensures us that the function attains its maximum on S, without actually calculating it.

There can be points other than c at which the function also attains a maximum value on the set S. Similarly, there can be points other than c' at which the function also attains a minimum value on the set S.

The continuity of f on $S = [a, b]$ is *sufficient* to guarantee the conclusions of the theorem; it is not necessary, since one can easily construct examples of functions which are *not* continuous on $S = [a, b]$, but for which the conclusions of the theorem still hold.

6 Problems for Discussion

6.1 Definition of Continuity

1. (i) Let $S = [a, b]$ with $a < b$, and let f and g be real-valued continuous functions on S. Let $h : S \to \mathbb{R}$ be defined by $h(x) = f(x) + g(x)$ for $x \in S$. Show, by using only the definition of continuity on a set (that is, without using the rules of continuity stated in Section 5.1) that h is continuous on S.

(ii) Let $S = [a, b]$ with $a < b$, and let f and g be real-valued continuous functions on S. Let $h : S \to \mathbb{R}$ be defined by $h(x) = f(x)g(x)$ for $x \in S$. Show, by using only the definition of continuity on a set (that is, without using the rules of continuity stated in Section 5.1) that h is continuous on S.

2. Let $S = [0, 1]$ and let f be a real-valued continuous function on S. Show that there is a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = f(x)$ for all $x \in S$. [In this case, g is called a *continuous extension* of the function f].

6.2 Rules for Continuity

3. Let $S = [a, b]$ with $a < b$, and let p and q be real-valued functions on S defined by:

$$\left.\begin{array}{l} p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \\ q(x) = b_0 + b_2 x^2 \end{array}\right\}$$

where $a_0, a_1, a_2, ..., a_n$ are arbitrary constants, and b_0, b_2 are arbitrary positive constants.

(i) Verify that $q(x) > 0$ for all $x \in S$.

(ii) Let $h : S \to \mathbb{R}$ be defined by $h(x) = p(x)/q(x)$ for $x \in S$. Show (using the rules of continuity stated in Section 5.1) that h is continuous on S.

4. Let $S = (0, 1)$ and let f be a real-valued function on S defined by:

$$f(x) = (1/x) \text{ for all } x \in S$$

(i) Show (using the rules of continuity stated in Section 5.1) that f is a continuous function on S.

(ii) Show that there is no continuous function $g : \mathbb{R} \to \mathbb{R}$ satisfying $g(x) = f(x)$ for all $x \in S$. [Compare this result with the result in problem 2 above].

6.3 Intermediate Value Theorem

5. Let $S = [0, 1]$ and let f be a real-valued continuous function from S to S. Show that there is some $c \in S$, such that $f(c) = c$. [In this case, c is called a *fixed point* of the function f].

6.4 Weierstrass Theorem

6. Let $f : \mathbb{R}_+ \to \mathbb{R}$ be defined by:

$$f(x) = \frac{x}{1+x} \text{ for all } x \in \mathbb{R}_+$$

(i) Show (using the rules of continuity stated in Section 5.1) that f is a continuous function on $\mathbb{R}_+$.

(ii) Let $S = [0, a]$, where a is an arbitrary positive real number. Show (using Weierstrass theorem) that there is some $c \in S$ such that $f(c) \geq f(x)$ for all $x \in S$.

(iii) Show that there is no $z \in \mathbb{R}_+$, satisfying $f(z) \geq f(x)$ for all $x \in \mathbb{R}_+$.

Part III
Differential Calculus

7 Differentiation

Let $S = (a, b)$, and f be a function from S to $\mathbb{R}$. If $c \in S$, then if the limit:

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \tag{1}$$

exists, then we say that f is *differentiable at c*, and we define the *derivative at c* as this limit. We write this as:

$$f'(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \tag{2}$$

If f is differentiable at every point $c \in S$, then we say that f is *differentiable on S*.

Note that since $S = (a, b)$ and $a < c < b$, we can allow h to be negative or positive as it approaches zero; that is, for $h \neq 0$, and $|h|$ sufficiently small, we can ensure that $(c + h) \in S$, and therefore the expression $f(c + h)$ in (1) is well-defined.

Sometimes we are interested in defining the derivative of a function on an interval which is *not* open. For example, we might have $S = [a, b]$, and f a function from S to $\mathbb{R}$. If we are interested in a derivative of the function at a, the above definition will not apply, since we cannot allow $h < 0$ in the expression in (1). In such a case, we can define the limit:

$$\lim_{h \to 0+} \frac{f(a+h) - f(a)}{h} \tag{3}$$

where the expression $h \to 0+$ means that $h \to 0$, with $h > 0$. If the limit in (3) exists, it is called the *right-hand derivative* of f at a. Clearly, we can do this for all $c \in [a, b)$. Similarly, one can define the *left-hand derivative* of f at all $c \in (a, b]$.

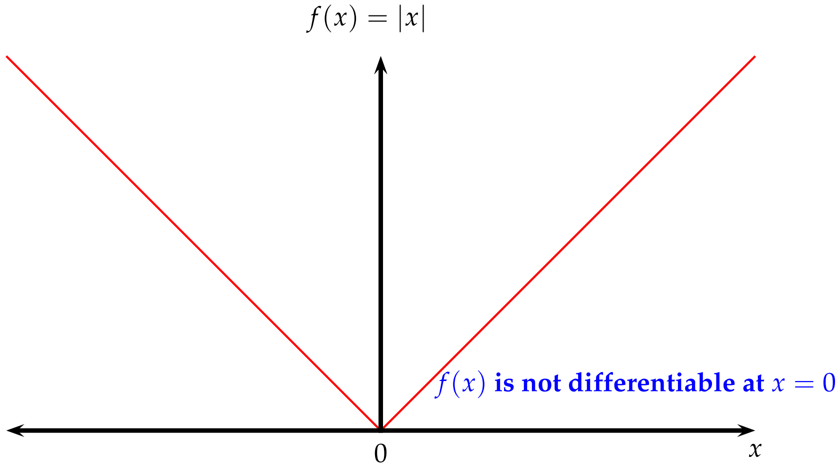
For the function f defined by:

$$f(x) = \begin{cases} |x| & \text{for } x \neq 0 \\ k & \text{for } x = 0 \end{cases}$$

it can be checked that if $k = 0$, then f is continuous at $x = 0$, but f is not differentiable at $x = 0$. The right hand derivative of f exists at $x = 0$ and is equal to 1; the left-hand derivative of f also exists at $x = 0$ and is equal to -1.

7.1 Some Rules for Differentiation

Let $S = (a, b)$ with $a < b$, and let f and g be functions from S to $\mathbb{R}$. Let $c \in S$. Then:



Continuity does not imply differentiability

(i) If f and g are differentiable at c, and $h : S \to \mathbb{R}$ is defined as $h(x) = f(x) + g(x)$ for $x \in S$, then h is differentiable at c, and:

$$h'(c) = f'(c) + g'(c)$$

(ii) If f and g are differentiable at c, and $h : S \to \mathbb{R}$ is defined as $h(x) = f(x)g(x)$ for $x \in S$, then h is differentiable at c, and:

$$h'(c) = f'(c)g(c) + g'(c)f(c)$$

(iii) If f and g are differentiable at c, and $h : S \to \mathbb{R}$ is defined as $h(x) = f(x)/g(x)$ for those $x \in S$ for which $g(x) \neq 0$, and $g(c) \neq 0$,then h is differentiable at c, and:

$$h'(c) = \frac{[f'(c)g(c) - g'(c)f(c)]}{[g(c)]^2}$$

These rules can be derived rigorously from the definition of a derivative.

We collect here a few basic formulae for differentiation of functions commonly encountered in applications of differential calculus. Applying the above rules to these basic formulae one can obtain differentiation formulae for a large class of functions.

Let $S = (a, b)$ with $a < b$, and let f be a function from S to $\mathbb{R}$. Let $c \in S$.

(i) If $f(x) = k$ for all $x \in S$, where k is a constant, then $f'(c) = 0$.

(ii) If $f(x) = x^n$ for all $x \in S$, where n is a positive integer, then $f'(c) = nc^{n-1}$.

(iii) If $f(x) = e^x$ for all $x \in S$, then $f'(c) = e^c$.

(iv) If $a \geq 0$, and $f(x) = \ln x$ for all $x \in S$, then $f'(c) = (1/c)$.

7.2 Chain Rule

Many functions commonly encountered in applications of differential calculus can be conveniently viewed as a *composite* of two functions. The *chain rule* provides a rule of differentiating such a function, based on the knowledge of the differentiability of the two individual functions which make up the composite function.

Chain Rule of Differentiation

Let $S = (a, b)$ with $a < b$, and let g be a function from S to $\mathbb{R}$. Let $T = (c, d)$ with $c < d$, and let f be a function from T to $\mathbb{R}$. Suppose $g(S) \subset T$, then the function $h : S \to \mathbb{R}$ defined by:

$$h(x) = f(g(x)) \text{ for all } x \in S$$

is called the composite function of f and g on S, and is written as $f \circ g$.

If g is differentiable on S and f is differentiable on T, then h is differentiable on S, and:

$$h'(x) = f'(g(x))g'(x) \text{ for all } x \in S \tag{4}$$

As an application of the chain rule, consider the function $h : S \to \mathbb{R}$ defined by:

$$h(x) = \ln(1 + x^4) \text{ for all } x \in S$$

where $S = (0,1)$. Then h can be viewed as a composite of two functions, $g : S \to \mathbb{R}$ defined by:

$$g(x) = (1 + x^4) \text{ for all } x \in S$$

and a function $f : T \to \mathbb{R}$ defined by:

$$f(y) = \ln y \text{ for all } y \in T$$

where $T = (0, 2)$. Then $g(S) = (1,2) \subset T$, and so:

$$h(x) = f(g(x)) \text{ for all } x \in S$$

is well-defined. [Note that it would be legitimate to choose T to be different from $(0,2)$, so long as $g(S)$ was a subset of T].

By our above rules of differentiation, g is differentiable on S and f is differentiable on T. Thus, by the chain rule h is differentiable on S, and using (4), we obtain:

$$h'(x) = \frac{4x^3}{(1 + x^4)} \text{ for all } x \in S$$

Applying the chain rule to the basic formulae described in the previous subsection, one can obtain additional useful formulae for differentiation. For example, if $S = \mathbb{R}_{++}$ and $f : S \to \mathbb{R}$ is defined by:

$$f(x) = x^a \text{ for all } x \in S$$

where a is a non-zero constant, then one can apply the chain rule to obtain:

$$f'(x) = ax^{a-1} \text{ for all } x \in S$$

7.3 Necessary Condition for a Maximum or Minimum

One of the key results in differential calculus relates to optimization theory. It provides a *necessary condition* for the existence of a point of *interior* maximum or minimum. Here is a formal statement.

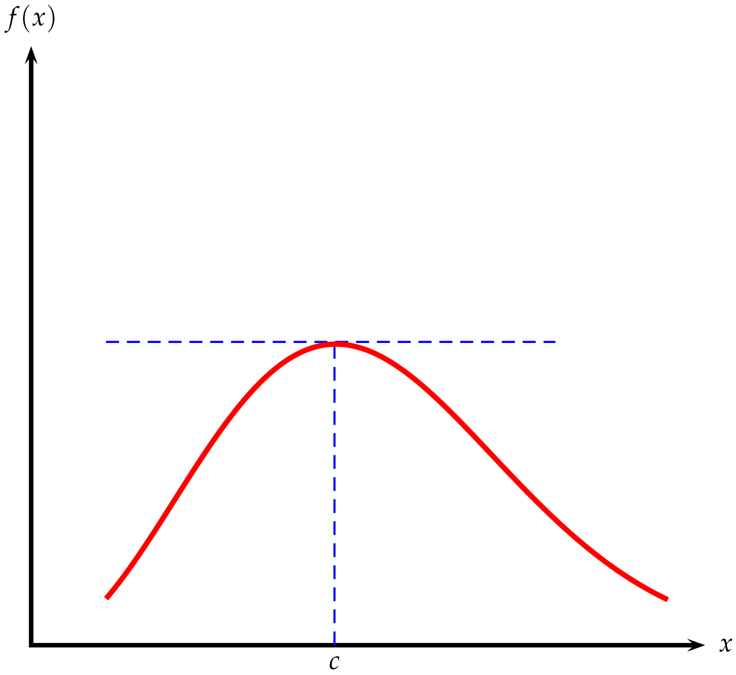
Necessary Condition for a Maximum or Minimum

Let $S = [a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a continuous function on S. Suppose $c \in (a, b)$ is a point of maximum or minimum of f on S. If f is differentiable at c, then $f'(c) = 0$.

We can make a few clarificatory remarks.

(i) By Weierstrass theorem there will always exist a point k in S where f reaches a maximum value among all points in S [A similar statement can be made for a minimum value of f on S].

(ii) The point k could be a or b, in which case the above result is not applicable. If the point k happens to be in (a, b), then the result can be applied provided f has a derivative at k.



Necessary condition for a maximum

(iii) If $k \in (a, b)$ and has a derivative at k, and $f'(k) = 0$, this does not by itself tell us whether k is a point of maximum or minimum of f on S. It might be a point of maximum, it might be a point of minimum, or it might be neither (an inflexion point).

Here is an application. Let $S = [a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a polynomial of degree $n = 3$ on S (that is, a cubic). Then, f is continuous on S, and differentiable on (a, b). The derivative, $f'(x)$, is a polynomial of degree $(n - 1) = 2$ on S (that is, a quadratic). Then, $f'(x) = 0$ will yield a quadratic equation, which we can solve. This can have at most two real roots in (a, b). This means that by evaluating f at four points (at most) in S, we can definitely find out where f attains a maximum and a minimum on S. [It is useful for you to spell out the full argument for this last statement]. Notice that this is possible without any information about higher order derivatives of f (which we will briefly discuss below).

7.4 The Mean Value Theorem

An important theorem which follows from the result on the necessary condition for a maximum or minimum of a function is the Mean Value Theorem.

The Mean Value Theorem

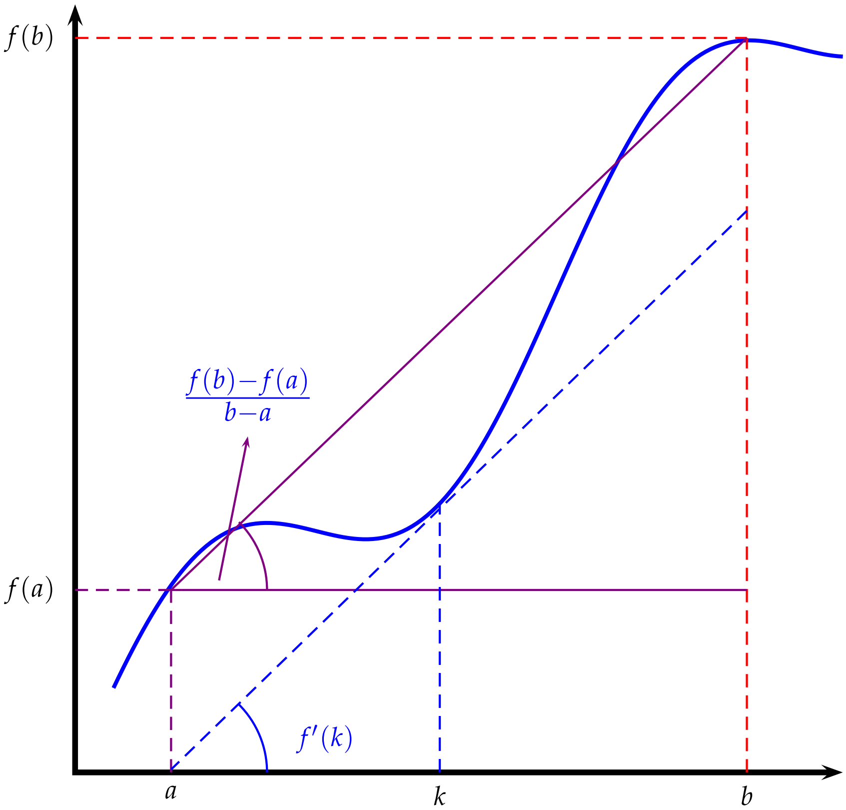
Let $S = [a, b]$, with $a < b$ and let $f : S \to \mathbb{R}$ be a continuous function on S, which is differentiable on (a, b). Then, there is some $k \in (a, b)$ such that:

$$\left[\frac{f(b) - f(a)}{b - a}\right] = f'(k) \tag{5}$$

The left hand side of (5) depends only on the function values at the end-points of the interval, and gives the slope of the line obtained by joining the points $(a, f(a))$ and $(b, f(b))$ with a straight line. The Mean Value Theorem says that there is some $k \in (a, b)$ at which the slope of the tangent to the function is exactly equal to the slope of this straight line. The value of k will clearly depend on the behavior of the function in between a and b.

The name "Mean-Value Theorem" derives from the fact that the left-hand side is the mean-value of $f'(x)$ on (a, b). [You should try to see that this is indeed the case after we have discussed the Second Fundamental Theorem of Integral Calculus]. The theorem says that the mean value of $f'(x)$ on (a, b) is actually attained by the function $f'(x)$ at some point $x = k \in (a, b)$.

As an application, let $S = [a, b]$, with $a < b$ and let $f : S \to \mathbb{R}$ be a continuous function on S, which is differentiable on (a, b), with $f'(x) > 0$ for all $x \in (a, b)$. Then we can use the Mean-Value theorem to infer that f is *increasing on S*; that is, if $x', x'' \in S$ with $x'' > x'$, then $f(x'') > f(x')$.



Mean Value Theorem

$$f'(k) = \frac{f(b)-f(a)}{b-a}$$

7.5 Second Order Derivatives and Taylor's Theorem

Let f be a differentiable function from $S = (a,b)$ to $\mathbb{R}$. Then $f'(x)$ is well-defined for every $x \in S$, and we can define a function $g : S \to \mathbb{R}$ by:

$$g(x) = f'(x) \text{ for all } x \in S$$

If $c \in S$, and g has a derivative at c, denoted by $g'(c)$, then this is called the *second derivative* of f at c, and we write:

$$f''(c) \equiv g'(c)$$

We check whether g has a derivative at c by the usual process of checking whether the limit:

$$\lim_{h \to 0} \frac{g(c+h) - g(c)}{h}$$

exists. When f has a second derivative at every $x \in S$, we say that f is *twice differentiable* on S.

When f is twice differentiable on S, an extension of the Mean Value Theorem can be established. We will refer to this as Taylor's theorem, even though the conventional statement of Taylor's theorem is more general, dealing with derivatives of a function of any finite order.

Taylor's Theorem

Let $S = (a,b)$ with $a < b$, and let f be a function from S to $\mathbb{R}$. Suppose $c, d \in S$, with $c < d$ and let $T = [c,d]$. Suppose f has a second derivative at every $x \in T$, then there is $k \in T$, such that:

$$f(d) - f(c) = f'(c)(d-c) + (1/2)f''(k)(d-c)^2 \tag{6}$$

In contrast to the Mean-Value Theorem, the first derivative of f appearing on the right hand side of (6) is evaluated at c, the lower end-point of the interval T. The second derivative of f is evaluated at a point k in between c and d.

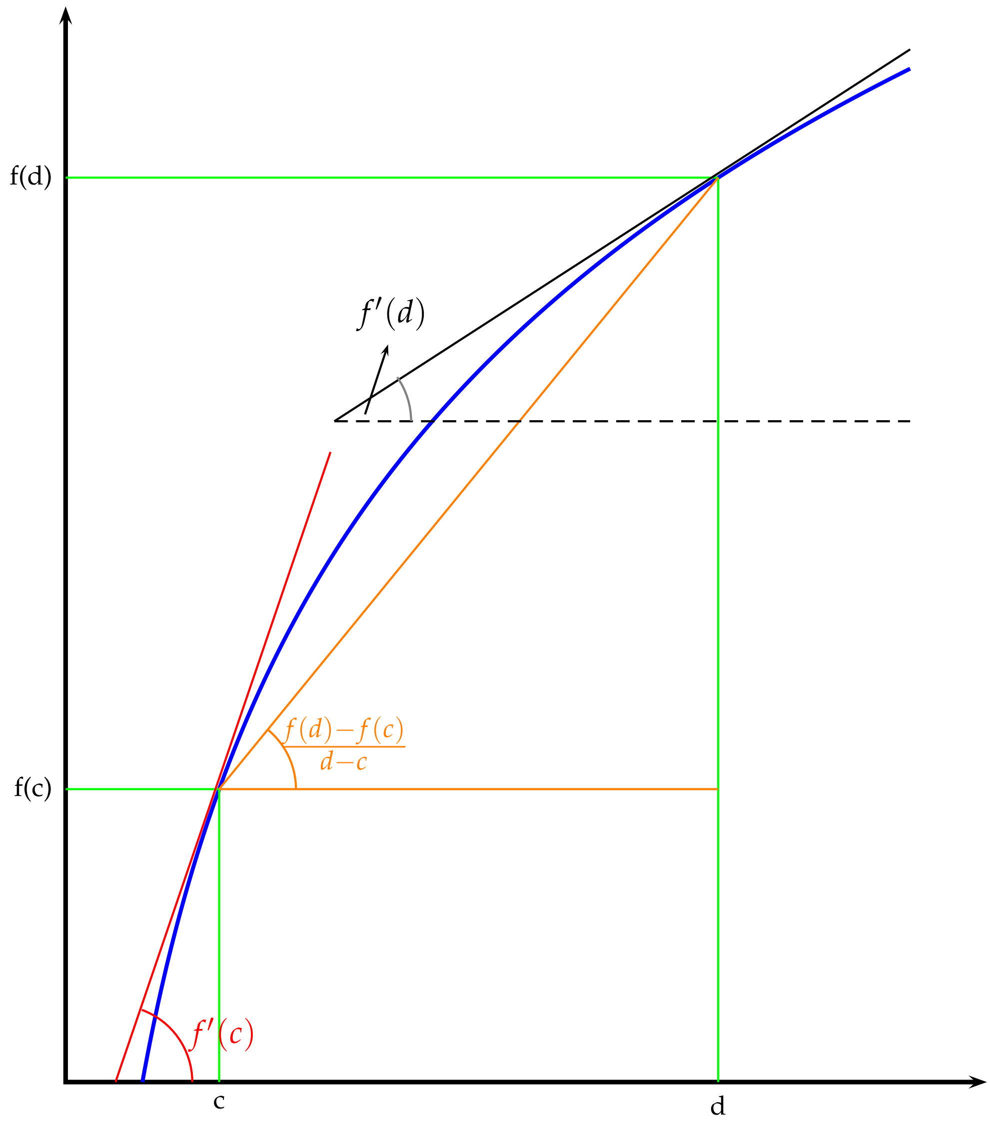
Using the above theorem, one can write an expression similar to (6), but where the first derivative of f appearing on the right hand side is evaluated at d, the upper end-point of the interval T. The statement would be written as follows. [It is a useful exercise for you to derive this from the above-stated result of Taylor's Theorem].

Let $S = (a,b)$ with $a < b$, and let f be a function from S to $\mathbb{R}$. Suppose $c, d \in S$, with $c < d$ and let $T = [c,d]$. Suppose f has a second derivative at every $x \in T$, then there is $k' \in T$, such that:

$$f(d) - f(c) = f'(d)(d-c) - (1/2)f''(k')(d-c)^2 \tag{7}$$

As an application, consider $S = (a,b)$ with $a < b$, and let f be a function from S to $\mathbb{R}$. Suppose $c, d \in S$, with $c < d$ and let $T = [c,d]$. Suppose f has a second derivative at every $x \in T$, and $f''(x) < 0$ for every $x \in T$, then:

$$f'(d) < \frac{f(d) - f(c)}{d - c} < f'(c) \tag{8}$$



A Concave Function where

$$f'(d) < \frac{f(d)-f(c)}{d-c} < f'(c)$$

This is an important inequality in the theory of *concave functions* which will be discussed in detail in the second part of this course. Draw a graph of f (for the specific case in which f is increasing on T) to geometrically depict the relationship given in (8).

8 Problems for Discussion

8.1 Continuity and Differentiability of Functions

1. (a) Let $S = [0, 2]$, and let $f : S \to \mathbb{R}$ be defined by:

$$f(x) = \begin{cases} 2x & \text{for } x \in [0,1] \\ x+1 & \text{for } x \in (1,2] \end{cases}$$

Show (using the definitions of continuity and differentiability) that f is continuous on S, and f is not differentiable at $x = 1$.

(b) Let $S = [a, b]$, and f a function on S. Suppose f is differentiable at a point $c \in (a, b)$. Show that f must be continuous at c.

8.2 Rules for Differentiation

2. Let $S = (a, b)$ with $a < b$, and let f and g be real-valued differentiable functions on S. Let $h : S \to \mathbb{R}$ be defined by $h(x) = f(x)g(x)$ for $x \in S$. Show, by using only the definition of differentiability on a set (that is, without using the rules of differentiability stated in Section 3.2) that h is differentiable on S.

8.3 Chain Rule of Differentiation

3. Differentiate the following functions (on the specified domains) using the chain rule of differentiation.

(a) $f : \mathbb{R}_{++} \to \mathbb{R}$ defined by $f(x) = x^a$, where a is a constant, with $a \neq 0$.

(b) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = a^x$, where a is a constant, with $a > 0$.

8.4 Necessary Condition for a Maximum or Minimum

4. Let $S = [a, b]$, and let $f : S \to \mathbb{R}$ be a continuous function on S, which is differentiable on (a, b).

(a) Suppose $f(a) = f(b) = 0$. Use the necessary condition of a maximum or minimum of a function to show that there is some $c \in (a, b)$ such that $f'(c) = 0$.

(b) Define $g : S \to \mathbb{R}$ by:

$$g(x) = [f(x) - f(a)] - \left[\frac{f(b) - f(a)}{b - a}\right](x - a) \quad \text{for all } x \in S$$

Note that $g(a) = g(b) = 0$. Now apply the result in (a) above to the function $g : S \to \mathbb{R}$ to infer that there is some $k \in (a, b)$ such that:

$$\left[\frac{f(b) - f(a)}{b - a}\right] = f'(k)$$

[The result in (a) is known as Rolle's theorem; the result in (b) is the Mean Value Theorem].

8.5 Mean Value Theorem

5. We say that a function $f : \mathbb{R} \to \mathbb{R}$ is *increasing* on $\mathbb{R}$ if whenever $x, x' \in \mathbb{R}$, and $x' > x$, we have $f(x') > f(x)$. Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable on $\mathbb{R}$.

(a) If $f'(x) > 0$ for all $x \in \mathbb{R}$, use the Mean Value Theorem to show that f is increasing on $\mathbb{R}$.

(b) If f is increasing on $\mathbb{R}$, does it follow that $f'(x) > 0$ for all $x \in \mathbb{R}$? Explain.

8.6 Maximum and Minimum at the Boundary

6. Let S denote the closed interval $[0, 3]$. Let $f : S \to \mathbb{R}$ by:

$$f(x) = 20 - 24x + 18x^2 - 4x^3 \quad \text{for all } x \in S$$

(a) Show that $f'(1) = 0$ and $f''(1) > 0$, but f does not attain a minimum at $x = 1$ among all $x \in S$.

(b) Show that $f'(2) = 0$ and $f''(2) < 0$, but f does not attain a maximum at $x = 2$ among all $x \in S$.

(c) Show that f attains a maximum at $x = 0$ and a minimum at $x = 3$ among all $x \in S$.

7. Let T be the open interval (a, b), and let S be the closed interval $[c, d]$, with $a < c < d < b$. Let f be a differentiable function on T.

(i) Suppose that $f(x) \leq f(c)$ for all $x \in S$. Show that $f'(c) \leq 0$.

(ii) Suppose that $f(x) \leq f(d)$ for all $x \in S$. Show that $f'(d) \geq 0$.

Part IV

Integral Calculus

9 Integration

We will confine our review to the theory of *Riemann Integration*. The basic subject matter of integration is the evaluation of the area "under a curve". Our exposition will be simplified by presenting the theory of Riemann Integration when the curve is generated by a *continuous* function on a closed interval. For most applications, this will suffice.

Let S be the set $[a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a continuous function on S. We can divide the interval $[a, b]$ into smaller subintervals by picking points in $[a, b]$, which we will label as x_i (with $i = 1, ..., n$), such that:

$$a < x_1 < x_2 \cdots < x_{n-1} < x_n = b \tag{1}$$

Then, $P \equiv \{a, x_1, x_2, ..., x_{n-1}, b\}$ is called a *partition* of the interval $[a, b]$. This partition produces n sub-intervals.

For the first closed sub-interval $[a, x_1]$, we can find a maximum value M_1 and a minimum value m_1 attained by the function f on the interval $[a, x_1]$, by using Weierstrass theorem. It follows that the area (A_1) under the curve generated by f on this sub-interval $[a, x_1]$ must satisfy:

$$m_1(x_1 - a) \leq A_1 \leq M_1(x_1 - a) \tag{2}$$

The idea here is to approximate the area A_1 (above and below) with the areas of rectangles, which we know how to calculate. We can repeat this step for the other closed sub-intervals $[x_1, x_2], ..., [x_{n-1}, b]$ to obtain:

$$m_{i+1}(x_{i+1} - x_i) \leq A_{i+1} \leq M_{i+1}(x_{i+1} - x_i) \text{ for all } i = 1, ..., n - 1 \tag{3}$$

Adding up the inequalities in (2) and (3), we obtain:

$$\begin{aligned} m_1(x_1 - a) + \textstyle\sum_{i=1}^{n-1} m_{i+1}(x_{i+1} - x_i) &\leq \textstyle\sum_{i=1}^{n} A_i \\ &\leq M_1(x_1 - a) + \textstyle\sum_{i=1}^{n-1} M_{i+1}(x_{i+1} - x_i) \end{aligned} \tag{4}$$

The middle expression in (4) is the area under the curve generated by f on the entire interval $[a, b]$. The left-hand expression in (4) is the *lower sum* $L(P)$ corresponding to the partition P, and the right-hand expression in (4) is the *upper sum* $U(P)$ corresponding to the partition P.

Now, we let the partition become *finer* by increasing the number of subintervals in the partition, and simultaneously decreasing the length of the largest subinterval. If this process of making the partition more fine is taken to the limit (that is, the length of the largest subinterval approaches zero), the upper sum $U(P)$ and the lower sum $L(P)$ approach the same number. [When f is continuous on $[a, b]$, this can be shown rigorously, by using the definition of a continuous function]. This number is the area under the curve generated by the function f. We define it as the *Riemann integral* of f on the interval $[a, b]$, and we denote it by the symbol:

$$\int_a^b f(x)dx$$

Note that if $a = b$, then the process described above stops at the very first step, yielding the result:

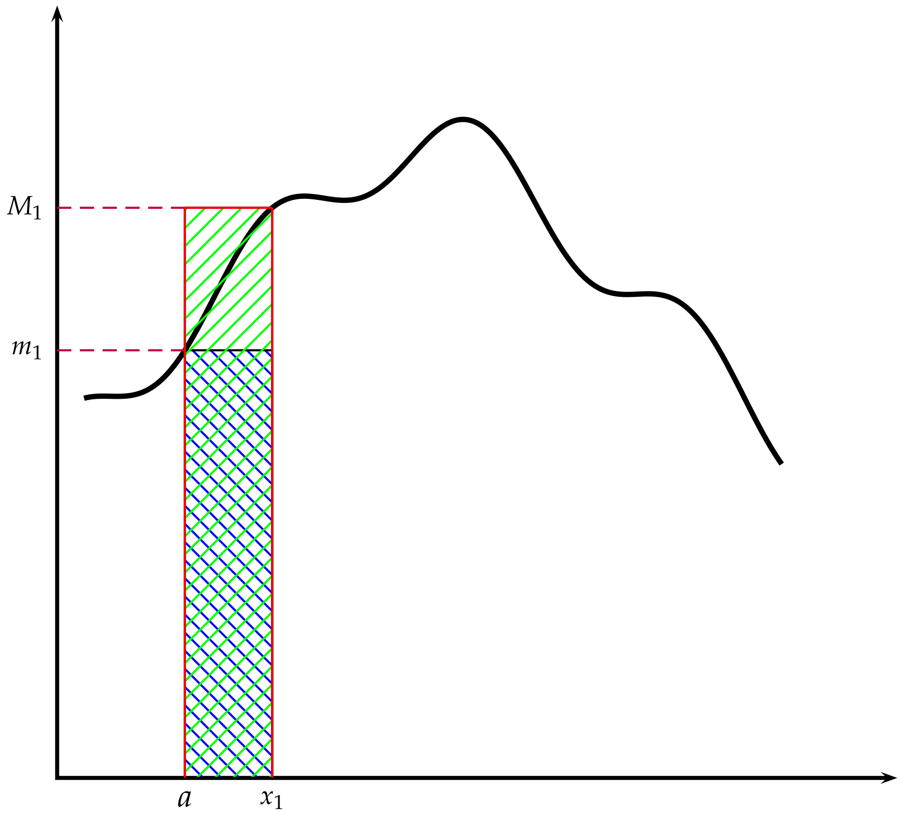
$$\int_a^a f(x)dx = 0$$

In the theory of integration, it is sometimes useful to also write the symbol:

$$\int_b^a f(x)dx \tag{5}$$

even when $a < b$. The symbol in (5) is *defined* as:

$$\int_b^a f(x)dx \equiv - \int_a^b f(x)dx$$



$$m_1(x_1 - a) \le A_1 \le M_1(x_1 - a)$$

(where A_1 is the area under the curve)

9.1 Fundamental Theorems of Integral Calculus

The first fundamental theorem of integral calculus states that integration is a process opposite to differentiation in the following sense. If we integrate a function f over the sub-interval $[a, y]$ where $a < y < b$, then we will get a function F of the variable y. If we now differentiate the function F, we will recover the original function, f.

First Fundamental Theorem of Integral Calculus

Let S be the set $[a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a continuous function on S. For each $y \in (a, b)$, define:

$$F(y) = \int_a^y f(x)dx$$

Then F is a differentiable function on (a, b), and:

$$F'(y) = f(y) \text{ for all } y \in (a, b) \tag{6}$$

To see intuitively why the result is true, note that for $y \in (a, b)$, and $y + h < b$, we have:

$$F(y + h) - F(y) \approx f(y)h \tag{7}$$

You can draw a diagram to see that $f(y)h$ is a good approximation to the left-hand side of (6), and that the approximation gets better as h becomes smaller. Thus,

$$\frac{F(y + h) - F(y)}{h} \approx f(y) \tag{8}$$

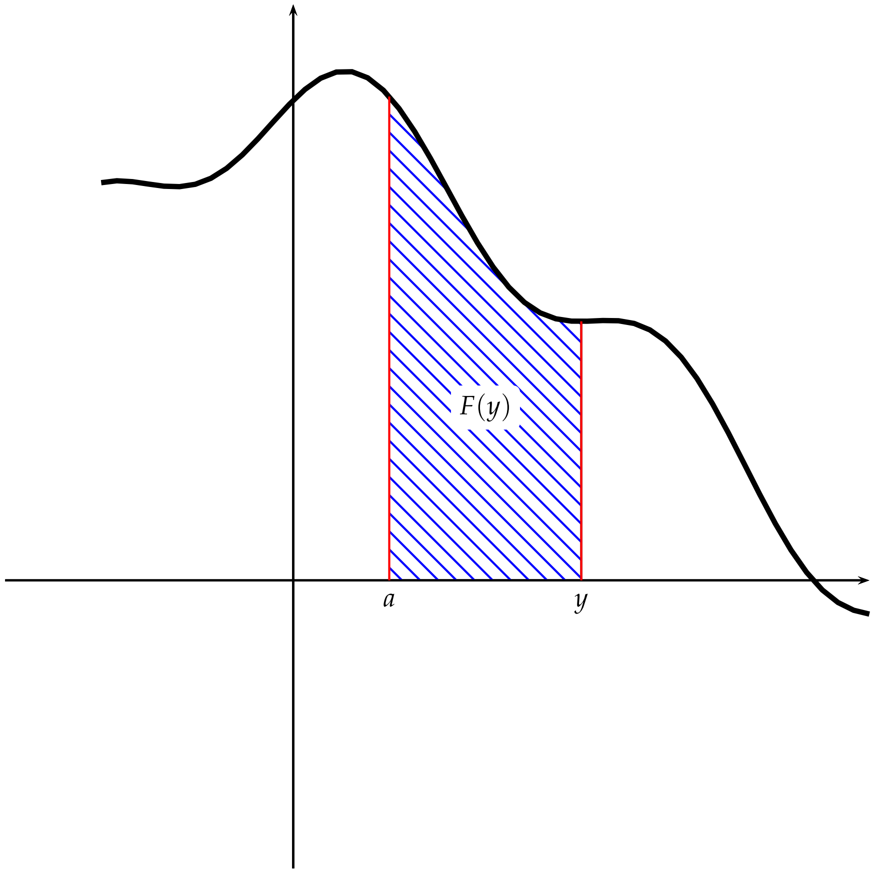
and letting $h \to 0$, and noting that the left-hand side of (8) converges to $F'(y)$, we get the equality claimed in (6). While the first fundamental theorem links the operations of integration and differentiation and is therefore useful in an understanding of both operations, it is the second fundamental theorem of integral calculus that really provides us with the tool we need in actually evaluating an integral.

Second Fundamental Theorem of Integral Calculus

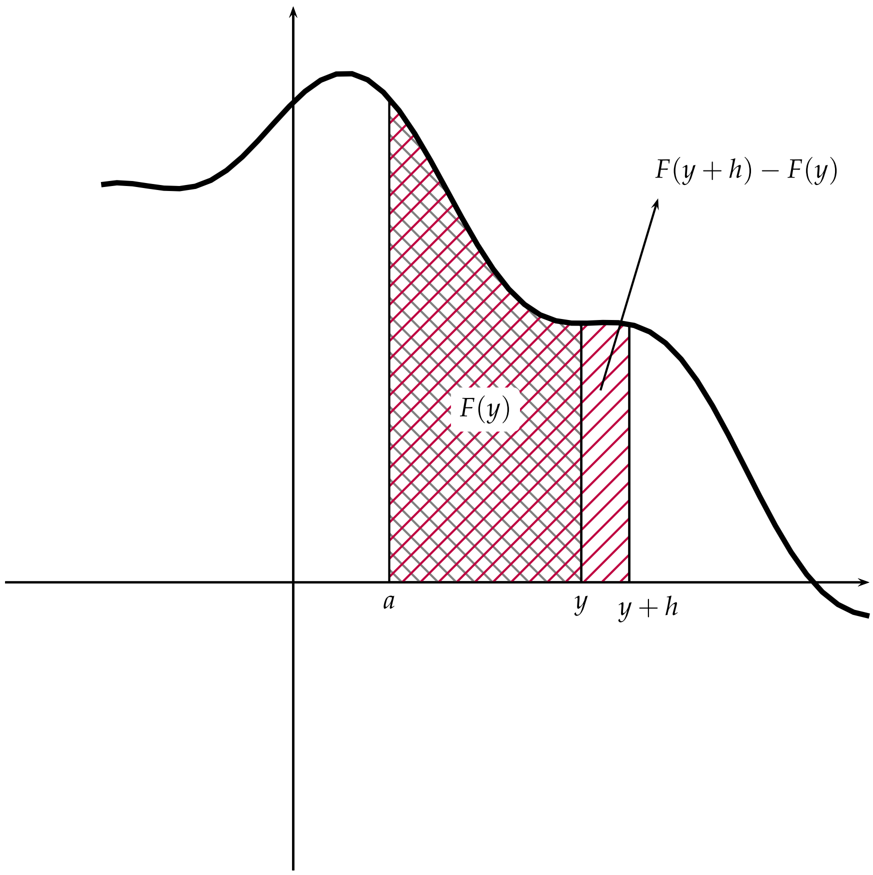
Let S be the set $[a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a continuous function on S. If F is a differentiable function on (c, d), such that $[a, b] \subset (c, d)$ and $F'(x) = f(x)$ for all $x \in [a, b]$, then:

$$F(b) - F(a) = \int_a^b f(x)dx \tag{9}$$

This means, given a continuous function $f : [a, b] \to \mathbb{R}$, which we want to integrate on $[a, b]$, we first find a function, F, whose *derivative* is the function f. Then, the integral on the right hand side of (9) can simply be evaluated by finding the values of F at a and at b, and thereby evaluating the left-hand side of (9). Clearly, familiarity with rules of differentiation will allow us to "guess" the function F in many cases, given the function f.



$$F(y) = \int_a^y f(x)dx$$



$$\frac{F(y+h)-F(y)}{h} \approx f(y)$$

9.2 Rules for Integration

The second fundamental theorem can be used to obtain several useful rules for integration. We list a few below.

(i) Let S be the set $[a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a continuous function on S. If k is a constant, then:

$$\int_a^b kf(x)dx = k \int_a^b f(x)dx$$

(ii) Let S be the set $[a, b]$ with $a < b$, and let $f : S \to \mathbb{R}$ be a continuous function on S. If $c \in (a, b)$, then:

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

(iii) Let S be the set $[a, b]$ with $a < b$, and let f and g be continuous functions from S to $\mathbb{R}$. Then,

$$\int_a^b f(x)dx + \int_a^b g(x)dx = \int_a^b [f(x) + g(x)]dx$$

The rules for differentiating functions commonly encountered in applications help us (using the second fundamental theorem of integral calculus) to obtain rules for integrating functions commonly encountered in applications.

(a) If $f : [a, b] \to \mathbb{R}$ is given by $f(x) = e^x$, then f is continuous on $[a, b]$. Since $F : \mathbb{R} \to \mathbb{R}$ defined by $F(x) = e^x$ is a function satisfying $F'(x) = e^x = f(x)$ for all $x \in [a, b]$, we obtain:

$$\int_a^b e^x dx = e^b - e^a$$

(b) If $f : [a, b] \to \mathbb{R}$ (with $a > 0$) is given by $f(x) = (1/x)$, then f is continuous on $[a, b]$. Since $F : \mathbb{R}_{++} \to \mathbb{R}$ defined by $F(x) = \ln x$ is a function satisfying $F'(x) = (1/x) = f(x)$ for all $x \in [a, b]$, we obtain:

$$\int_a^b (1/x)dx = \ln b - \ln a$$

(c) If $f : [a, b] \to \mathbb{R}$ is given by $f(x) = x^n$, where n is a positive integer, then f is continuous on $[a, b]$. Since $F : \mathbb{R} \to \mathbb{R}$ defined by $F(x) = (x^{n+1}/(n + 1))$ is a function satisfying $F'(x) = x^n = f(x)$ for all $x \in [a, b]$, we obtain:

$$\int_a^b x^n dx = \frac{b^{n+1}}{(n+1)} - \frac{a^{n+1}}{(n+1)}$$

9.3 Integration by Parts

The rule of differentiation for the sum of two functions leads to the sum rule of integration noted as rule (iii) in the previous subsection. The rule of differentiation for the product of two functions leads to the rule of integration known as *integration by parts.*

Integration by Parts:

Let (c,d) and $[a,b]$ be intervals with $[a,b] \subset (c,d)$. Let f and g be differentiable functions on (c,d), whose derivatives are continuous on $[a,b]$. Define:

$$h(x) = f(x)g(x) \text{ for all } x \in (c,d)$$

Then:

$$f(b)g(b) - f(a)g(a) = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx \tag{IP}$$

To see how this formula is derived note that the product rule of differentiation yields:

$$h'(x) = f'(x)g(x) + f(x)g'(x) \text{ for all } x \in (c,d)$$

Applying rule (iii) of integration (noted in the previous subsection), we obtain:

$$\int_a^b h'(x)dx = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx \tag{10}$$

Using the second fundamental theorem of integral calculus on the left-hand side of (10), we get:

$$f(b)g(b) - f(a)g(a) = h(b) - h(a) = \int_a^b h'(x)dx \tag{11}$$

Combining (10) and (11) yields the integration by parts formula:

$$f(b)g(b) - f(a)g(a) = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx \tag{12}$$

As an application of this formula, one can show that if $0 < a < b$, then:

$$\int_a^b \ln x \, dx = (b \ln b - a \ln a) - (b - a) \tag{13}$$

You should check that (13) does indeed follow from (IP); the idea is to define f and g appropriately before applying (IP).

9.4 Integration by Change of Variable

The chain rule of differentiation leads to a rule of integration, known as integration *by substitution* or *change of variable.*

Integration by Change of Variable:

Let (a',b') and $[a,b]$ be intervals with $S = [a,b] \subset (a',b') = T$. Let g be a differentiable function on T, whose derivative is continuous on $[a,b]$. If f is a continuous function from $g(S)$ to $\mathbb{R}$, then:

$$\int_{g(a)}^{g(b)} f(y)dy = \int_a^b f(g(x))g'(x)dx \tag{ICV}$$

Note that the statement of the result is made convenient because of the symbol introduced in (5).

To see how the formula in the above result is obtained, note first that $g(S)$ is a closed interval in $\mathbb{R}$ [it is useful for you to spell out the argument for this step]. If the closed interval is degenerate, then (ICV) holds trivially. Thus, assume that the closed interval $g(S)$ is not degenerate and denote it by $[c, d]$, where $c < d$.

Since f is a continuous function from $[c, d]$ to $\mathbb{R}$, we can apply the first fundamental theorem of integral calculus to define a function F from (c, d) to $\mathbb{R}$ such that:

$$F(y) = \int_a^y f(x)dx \text{ and } F'(y) = f(y) \text{ for all } y \in (c, d)$$

Define:

$$F(y) = \begin{cases} f(c)(y - c) & \text{for all } y \leq c \\ \int_a^b f(y)dy + f(d)(y - d) & \text{for all } y \geq d \end{cases}$$

Then, F is differentiable on $\mathbb{R}$, and its derivative is continuous on $\mathbb{R}$, with:

$$F'(y) = f(y) \text{ for all } y \in [c, d] \tag{14}$$

[It is useful for you to spell out the argument for this step]. Now, define the composite function:

$$G(x) = F(g(x)) \text{ for all } x \in T \tag{CF}$$

Since g is differentiable on T and F is differentiable on $\mathbb{R}$, and $g(T) \subset \mathbb{R}$, the composite function $G : T \to \mathbb{R}$ is well-defined. Since $S \subset T$, we can apply the Chain Rule of differentiation to get:

$$G'(x) = F'(g(x))g'(x) \text{ for all } x \in S \tag{15}$$

Note that the right hand side of (15) is a continuous function on $S = [a, b]$. By (15), so is the left-hand side. We can therefore obtain:

$$\int_a^b G'(x)dx = \int_a^b F'(g(x))g'(x)dx \tag{16}$$

Using (14) in (16), one gets:

$$\int_a^b G'(x)dx = \int_a^b f(g(x))g'(x)dx \tag{17}$$

Using the second fundamental theorem of integral calculus, and using the definition of G in (CF), we can write (17) as:

$$F(g(b)) - F(g(a)) = G(b) - G(a) = \int_a^b f(g(x))g'(x)dx \tag{18}$$

Employing the second fundamental theorem of integral calculus once again, and using (14), we can rewrite (18) as:

$$\int_{g(a)}^{g(b)} f(y)dy = \int_{g(a)}^{g(b)} F'(y)dy = F(g(b)) - F(g(a)) = \int_a^b f(g(x))g'(x)dx$$

which yields (ICV).

As an application, consider a function $g : \mathbb{R}_{++} \to \mathbb{R}_{++}$, which is differentiable on $\mathbb{R}_{++}$, and assume that its derivative is continuous on $\mathbb{R}_{++}$. Let $S = [a, b]$, with $0 < a < b$. Then (ICV) can be applied to obtain the integral:

$$\int_a^b \frac{g'(x)}{g(x)} dx = \ln g(b) - \ln g(a) \tag{19}$$

You should check that (19) does indeed follow from (ICV); the idea is to define f (and therefore F) appropriately before applying (ICV).

10 Problems for Discussion

10.1 Fundamental Theorems of Integral Calculus

1. If f is a continuous function on $[a, b]$, show that there is some $c \in (a, b)$ such that:

$$\int_a^b f(x)dx = f(c)(b - a)$$

This is called the First Mean Value Theorem of Integral Calculus.

2. If f and g are continuous functions on $[a, b]$, and $g(x) \geq 0$ for all $x \in [a, b]$, show that there is some $c \in (a, b)$ such that:

$$\int_a^b f(x)g(x)dx = f(c) \int_a^b g(x)dx$$

This is called the Second Mean Value Theorem of Integral Calculus.

3. Let $g : \mathbb{R}_{++} \to \mathbb{R}_{++}$ be a function, which is differentiable on $\mathbb{R}_{++}$ and satisfies:

$$\frac{xg'(x)}{g(x)} = \alpha \quad \text{for all } x \in \mathbb{R}_{++}$$

where α is a constant, satisfying $\alpha \in (0, 1)$. Show that there is a positive constant A such that:

$$g(x) = Ax^{\alpha} \quad \text{for all } x \in \mathbb{R}_{++}$$

10.2 Integration by Parts

4. Using the Integration by Parts formula, show that if $0 < a < b$, then:

$$\int_a^b \ln x \, dx = (b \ln b - a \ln a) - (b - a)$$

10.3 Integration by Change of Variable

5. Suppose $g : \mathbb{R}_{++} \to \mathbb{R}_{++}$, is differentiable on $\mathbb{R}_{++}$, and assume that its derivative is continuous on $\mathbb{R}_{++}$. Let $S = [a, b]$, with $0 < a < b$. Using the Integration by Change of Variable formula, show that:

$$\int_a^b \frac{g'(x)}{g(x)} dx = \ln g(b) - \ln g(a)$$

Part V
Partial Differentiation

11 Notation

For this part and the next two parts of the review, we will denote $\mathbb{R}^2_+$ by X, and $\mathbb{R}^2_{++}$ by Y. The set $X \sim Y$ will be denoted by $B(Y)$; this is the set of points in X which are *not* in Y.

12 Functions of Two Real Variables

For this part and the next two parts of the review, we will be exclusively concerned with real valued functions of two real variables, whose domain is a subset of X. The definitions of limit and continuity of a function $f : X \to \mathbb{R}$ can then be given as follows.

12.1 Limit of a Function

Let f be a function from X to $\mathbb{R}$. If $c = (c_1, c_2) \in X$, the expression:

$$\lim_{x \to c} f(x) = L$$

means that given any $\varepsilon > 0$, there is $\delta > 0$ such that:

$$x = (x_1, x_2) \in X \text{ and } 0 < |x_1 - c_1| + |x_2 - c_2| < \delta \text{ imply } |f(x) - L| < \varepsilon$$

12.2 Continuity of a Function

Let f be a function from X to $\mathbb{R}$. If $c = (c_1, c_2) \in X$, then f is *continuous at c* if $\lim_{x \to c} f(x)$ exists and is equal to $f(c)$.

Using the definition of continuity, it is possible to check that the following functions, commonly encountered in applications, are continuous functions on X.

(i) $f : X \to \mathbb{R}$ is given by $f(x_1, x_2) = ax_1 + bx_2$, where a and b are constants.

(ii) $f : X \to \mathbb{R}$ is given by $f(x_1, x_2) = x_1^{\alpha} x_2^{\beta}$, where α and β are positive constants.

(iii) $f : X \to \mathbb{R}$ is given by $f(x_1, x_2) = x_1 x_2 / (x_1 + x_2)$ if $(x_1 + x_2) > 0$, and $f(x_1, x_2) = 0$ if $(x_1 + x_2) = 0$.

(iv) $f : X \to \mathbb{R}$ is given by $f(x_1, x_2) = g(x_1) + h(x_2)$, where g is a continuous function on $\mathbb{R}_+$ and h is a continuous function on $\mathbb{R}_+$.

(v) $f : X \to \mathbb{R}$ is given by $f(x_1, x_2) = g(x_1)h(x_2)$, where g is a continuous function on $\mathbb{R}_+$ and h is a continuous function on $\mathbb{R}_+$.

13 Partial Derivatives

13.1 First-Order Partial Derivatives

Let f be a function from X to $\mathbb{R}$. If $x = (x_1, x_2) \in Y$, the limit

$$\lim_{h\to 0} \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h}$$

if it exists, is called the (first-order) *partial derivative* of f at x with respect to the first variable, and is denoted by $D_1 f(x)$. Similarly, the limit

$$\lim_{h\to 0} \frac{f(x_1, x_2 + h) - f(x_1, x_2)}{h}$$

if it exists, is called the (first-order) *partial derivative* of f at x with respect to the second variable, and is denoted by $D_2 f(x)$.[1]

This means that we can compute partial derivatives just like ordinary derivatives of a function of *one* variable. That is, if $f(x_1, x_2)$ is given by some formula involving (x_1, x_2), then we find $D_1 f(x)$ by differentiating the function whose value at x_1 is given by the formula when x_2 is thought of as a constant; similarly, we find $D_2 f(x)$ by differentiating the function whose value at x_2 is given by the formula when x_1 is thought of as a constant.

For example, if $f : X \to \mathbb{R}$ is given by $f(x_1,\ x_2) = 2x_1^2 + 4x_2^3 + 3x_1x_2$, then for every $x \in Y$, we have $D_1 f(x) = 4x_1 + 3x_2$, $D_2 f(x) = 12x_2^2 + 3x_1$.

When $f : X \to \mathbb{R}$ has partial derivatives at $x \in Y$, we write the vector $[D_1 f(x), D_2 f(x)]$ as $\nabla f(x)$; this is called the *gradient vector* of f at x.

When $f : X \to \mathbb{R}$ has partial derivatives at each $x \in Y$, we say that f has *partial derivatives on* Y.

Sometimes, we are interested in extending the above definition of (first-order) partial derivatives to the boundary of Y. If $x \in B(Y)$, and $x_1 = 0$, we will say that f has a partial derivative at x with respect to the first variable if the limit:

$$\lim_{h\to 0+} \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h}$$

exists; that is if its right-hand partial derivative $D_1^+ f(x_1, x_2)$ exists. Similarly, if $x \in B(Y)$, and $x_2 = 0$, we will say that f has a partial derivative at x with respect to the second variable if the limit:

$$\lim_{h\to 0+} \frac{f(x_1, x_2 + h) - f(x_1, x_2)}{h}$$

exists; that is if its right-hand partial derivative $D_2^+ f(x_1, x_2)$ exists.

[1] These partial derivatives of f at x are also, alternatively, denoted by $f_1(x)$ and $f_2(x)$ respectively.

13.2 Second-Order Partial Derivatives

When $f : X \to \mathbb{R}$ has (first-order) partial derivatives on Y, these first-order partial derivatives are themselves functions from Y to $\mathbb{R}$. If these functions have (first-order) partial derivatives on Y, *these* partial derivatives are called the *second-order partial derivatives* of f on Y.

To elaborate, if $D_1 f(x)$ exists for all $x \in Y$, we can define the function $D_1 f : Y \to \mathbb{R}$. If this function has (first-order) partial derivatives on Y, then the (first-order) partial derivative of $D_1 f$ at x with respect to the first variable is denoted by $D_{11} f(x)$ and the (first-order) partial derivative of $D_1 f$ at x with respect to the second variable is denoted by $D_{12} f(x)$. Similarly, if $D_2 f(x)$ exists for all $x \in Y$, we can define the function $D_2 f : Y \to \mathbb{R}$. If this function has (first-order) partial derivatives on Y, then the (first-order) partial derivative of $D_2 f$ at x with respect to the first variable is denoted by $D_{21} f(x)$ and the (first-order) partial derivative of $D_2 f$ at x with respect to the second variable is denoted by $D_{22} f(x)$. These four partial derivatives are the *second-order partial derivatives* of f at x.

In the example described above, $D_{11} f(x) = 4$, $D_{12} f(x) = 3 = D_{21} f(x)$; $D_{22} f(x) = 24x_2$. We note in this example that the "cross partials" $D_{12} f(x)$ and $D_{21} f(x)$ are equal. This is not a coincidence; it is a more general phenomenon as noted in the following result.

Young's Theorem:

Suppose $f : X \to \mathbb{R}$ has first and second-order partial derivatives on Y. If $D_{12} f$ and $D_{21} f$ are continuous on Y, then $D_{12} f(x) = D_{21} f(x)$ for all $x \in Y$.

When $f : X \to \mathbb{R}$ has first and second-order partial derivatives on Y, we can arrange them in the form of a 2×2 matrix:

$$H_f(x) = \begin{bmatrix} D_{11} f(x) & D_{12} f(x) \\ D_{21} f(x) & D_{22} f(x) \end{bmatrix}$$

This matrix is called the *Hessian matrix* of f at $x \in Y$ and is denoted by $H_f(x)$.

In the example described above:

$$H_f(x) = \begin{bmatrix} 4 & 3 \\ 3 & 24x_2 \end{bmatrix}$$

is the Hessian matrix of f for all $(x_1, x_2) \in Y$.

14 Problems for Discussion

14.1 Cobb-Douglas Functions

1. Let $f : X \to \mathbb{R}_+$ be given by:

$$f(x_1, x_2) = x_1^{\alpha} x_2^{\beta} \text{ for all } (x_1, x_2) \in X$$

where $\alpha > 0$ and $\beta > 0$ are given parameters.

(i) Obtain the gradient vector of f for $(x_1, x_2) \in Y$.

(ii) Obtain the Hessian matrix of f for $(x_1, x_2) \in Y$.

14.2 CES Functions

2. Let $f : X \to \mathbb{R}_+$ be defined by:

$$f(x_1, x_2) = \begin{cases} [\theta x_1{}^{-\alpha} + (1-\theta)x_2{}^{-\alpha}]^{-(1/\alpha)} & \text{for } (x_1, x_2) \in \mathbb{R}^2_{++} \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 0$ and $\theta \in (0, 1)$ are given parameters.

(i) Obtain the gradient vector of f for $(x_1, x_2) \in Y$.

(ii) Obtain the Hessian matrix of f for $(x_1, x_2) \in Y$.

Part VI
Convex Analysis

15 Convex Sets and Concave Functions

If $x, y \in X$, the *line segment* joining x and y is given by the set of points $\{z \in X : z = \theta x + (1-\theta)y\}$ for some $0 \leq \theta \leq 1\}$. A set $S \subset X$ is a *convex set* if for every $x, y \in S$, the line segment joining x and y is contained in S.

Applying this definition, you can check that $X \equiv \mathbb{R}^2_+$ and $Y \equiv \mathbb{R}^2_{++}$ are convex sets.

Let A be a convex set in X. Then $f : A \to \mathbb{R}$ is a *concave function* (on A) if for all $x^1, x^2 \in A$, and for all $0 \leq \theta \leq 1$,

$$f[\theta x^1 + (1-\theta)x^2] \geq \theta f(x^1) + (1-\theta)f(x^2)$$

The function f is *strictly concave* on A if $f[\theta x^1 + (1-\theta)x^2] > \theta f(x^1) + (1-\theta)f(x^2)$ whenever $x^1, x^2 \in A$, $x^1 \neq x^2$ and $0 < \theta < 1$.

16 Properties of Concave Functions

The following results summarize some of the most important and useful properties of concave functions.

(i) If $f : X \to \mathbb{R}$ is a concave function on X, then f is a continuous function on Y.

(ii) If $f : X \to \mathbb{R}$ is a concave function on X, with continuous first order partial derivatives on Y, then:

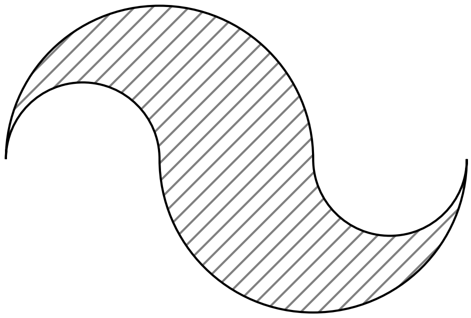
$$f(x_1', x_2') - f(\bar{x}_1, \bar{x}_2) \leq D_1 f(\bar{x}_1, \bar{x}_2)(x_1' - \bar{x}_1) + D_2 f(\bar{x}_1, \bar{x}_2)(x_2' - \bar{x}_2)$$

whenever $x' \equiv (x_1', x_2')$ and $\bar{x} \equiv (\bar{x}_1, \bar{x}_2)$ are in Y.

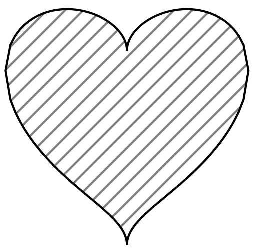
(iii) If $f : X \to \mathbb{R}$ is continuous on X, with continuous first and second order partial derivatives on Y, then f is concave on X if and only if:

$$\left.\begin{array}{l} D_{11}f(x_1, x_2) \leq 0, D_{22}f(x_1, x_2) \leq 0, \text{ and} \\ D_{11}f(x_1, x_2) D_{22}f(x_1, x_2) \geq [D_{12}f(x_1, x_2)]^2 \end{array}\right\}$$

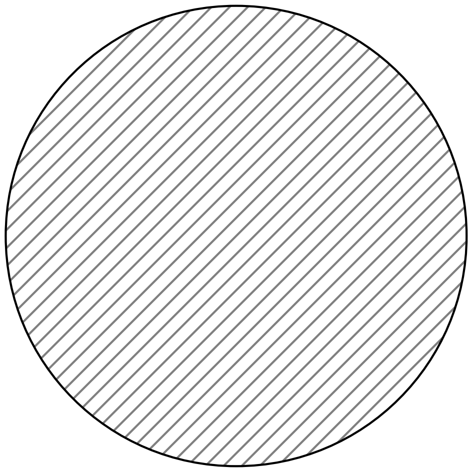
for all $(x_1, x_2) \in Y$.



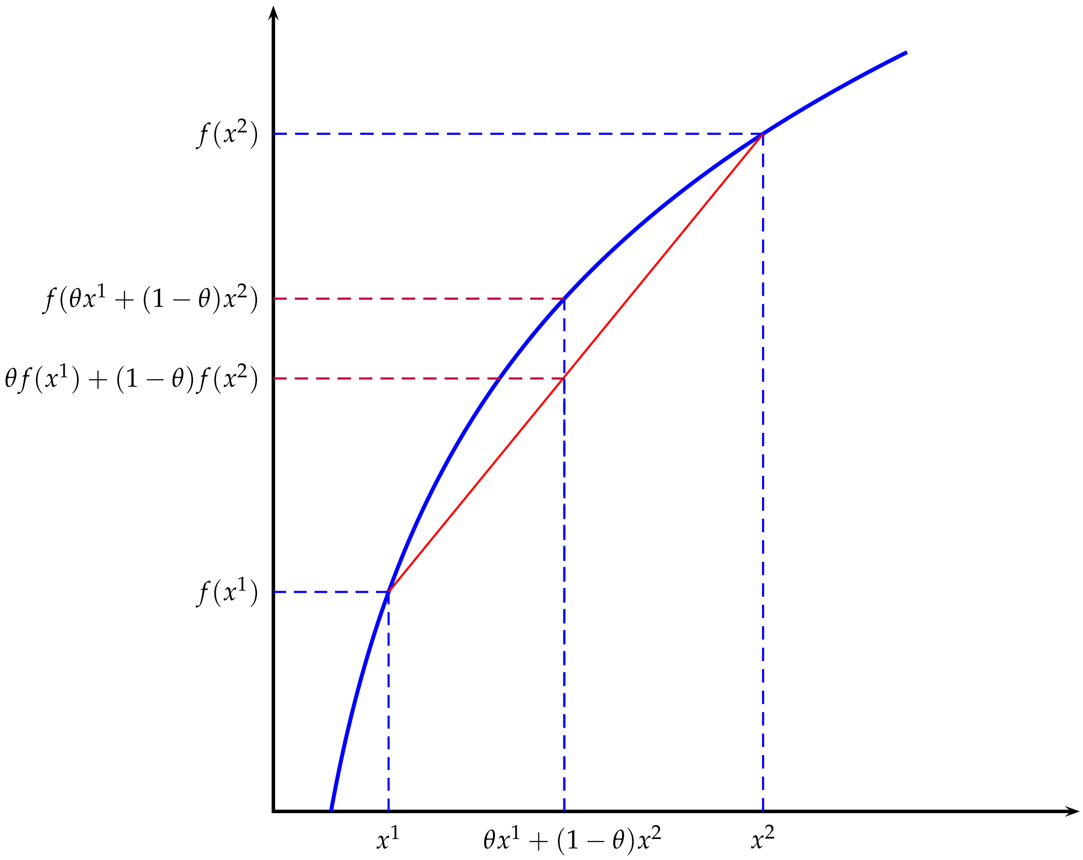
Not a convex Set



Not a convex set



A Convex Set



$$f(\theta x^1 + (1-\theta)x^2) \geq \theta f(x^1) + (1-\theta) f(x^2)$$

17 Problems for Discussion

17.1 Discontinuity of a Concave Function

1. Give an example of a concave function $f : X \to \mathbb{R}$ which is not continuous on X. Note that since we know that f must be continuous on Y (by property (i) of concave functions in Section 16), the discontinuity must be at the boundary of X.

17.2 Addition of Concave Functions

2. Let f and g be concave functions from X to $\mathbb{R}$. Let $h : X \to \mathbb{R}$ be defined by:

$$h(x_1, x_2) = f(x_1, x_2) + g(x_1, x_2) \text{ for all } (x_1, x_2) \in X$$

Show that h is a concave function from X to $\mathbb{R}$.

17.3 Cobb-Douglas Functions

3. Let $f : X \to \mathbb{R}_+$ be given by:

$$f(x_1, x_2) = x_1^{\alpha} x_2^{\beta} \text{ for all } (x_1, x_2) \in X$$

where $\alpha > 0$ and $\beta > 0$ are given parameters. Show that f is concave on X if and only if $(\alpha + \beta) \leq 1$.

17.4 CES Functions

4. Let $f : X \to \mathbb{R}_+$ be defined by:

$$f(x_1, x_2) = \begin{cases} [\theta x_1{}^{-\alpha} + (1-\theta) x_2{}^{-\alpha}]^{-(1/\alpha)} & \text{for } (x_1, x_2) \in \mathbb{R}^2_{++} \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 0$ and $\theta \in (0, 1)$ are given parameters. Apply property (iii) of concave functions (listed in Section 16) to determine whether f is a concave function on X.

Part VII
Introduction to Optimization Theory

18 A Constrained Maximization Problem

For the purpose of this review, we will be concerned exclusively with the following constrained maximization problem:

$$\left.\begin{array}{ll} \textit{Max} & f(x_1,x_2) \\ \textit{subject to} & g(x_1,x_2) \geq 0 \\ \textit{and} & (x_1,x_2) \geq 0 \end{array}\right\} (P)$$

where f and g are functions from X to $\mathbb{R}$. The set:

$$C = \{(x_1,x_2) \geq 0 : g(x_1,x_2) \geq 0\} \tag{1}$$

is called the *constraint set* for the problem (P). Thus, (P) can also be written as:

$$\left.\begin{array}{ll} \textit{Max} & f(x_1,x_2) \\ \textit{subject to} & (x_1,x_2) \in C \end{array}\right\} (P')$$

We will discuss two results, which can be applied to solve a class of problems, which can be written in the form of (P), and which are commonly encountered in microeconomic theory.

To this end, we will maintain the following assumptions on the functions f and g :

(A1) f and g are continuous functions on X, with continuous first and second order partial derivatives on Y.

(A2) f and g are concave functions on X.

Thus, before applying either of the results to be discussed below, one should check that the maintained assumptions are satisfied. We observe here that it is assumption (A2) that is crucial for the results, but we maintain assumption (A1) because it gives us a convenient method of checking that assumption (A2) holds, by applying property (iii) of concave functions listed in Section 16.

18.1 The Lagrangean

Given problem (P), one can denote by λ an auxiliary variable (known as the Lagrange multiplier), and define the Lagrangean, L, as follows:

$$L(x_1,x_2,\lambda) = f(x_1,x_2) + \lambda g(x_1,x_2) \text{ for all } (x_1,x_2) \in X \text{ and } \lambda \in \mathbb{R} \tag{2}$$

At any $(x_1, x_2) \in X$, where the first-order partial derivatives of f and g exist, and any $\lambda \in \mathbb{R}$, we can obtain the partial derivatives of L as follows:

$$\left.\begin{array}{l} D_1 L(x_1, x_2, \lambda) = D_1 f(x_1, x_2) + \lambda D_1 g(x_1, x_2) \\ D_2 L(x_1, x_2, \lambda) = D_2 f(x_1, x_2) + \lambda D_2 g(x_1, x_2) \\ D_3 L(x_1, x_2, \lambda) = g(x_1, x_2) \end{array}\right\} \quad (3)$$

These derivatives are used to write the first-order conditions in the results below.

18.2 First Result

The first result provides sufficient conditions for a point in the constraint set to be a solution to problem (P); these conditions involve solving a system of simultaneous equations.

Theorem 1 *Suppose assumptions (A1) and (A2) hold. Suppose we can find* $(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \geq 0$, *such that the following two conditions hold:*

(i) the first-order partial derivatives of f and g with respect to x_1 and x_2 exist and are continuous at $(\bar{x}_1, \bar{x}_2)$*;*

(ii) $(\bar{x}_1, \bar{x}_2, \bar{\lambda})$ *satisfies:*

$$\left.\begin{array}{l} D_1 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \\ D_2 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \\ D_3 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \end{array}\right\} (FOC)$$

Then, $(\bar{x}_1, \bar{x}_2)$ *solves problem* (P).

The result can be applied to a variety of problems, which can be written in the form of (P). A standard application is to a consumer's utility maximizing problem, with a Cobb-Douglas utility function, subject to a budget constraint.

Consider the problem:

$$\left.\begin{array}{ll} Max & x_1^{\alpha} x_2^{\beta} \\ subject\ \ to & p_1 x_1 + p_2 x_2 \leq M \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (UM)$$

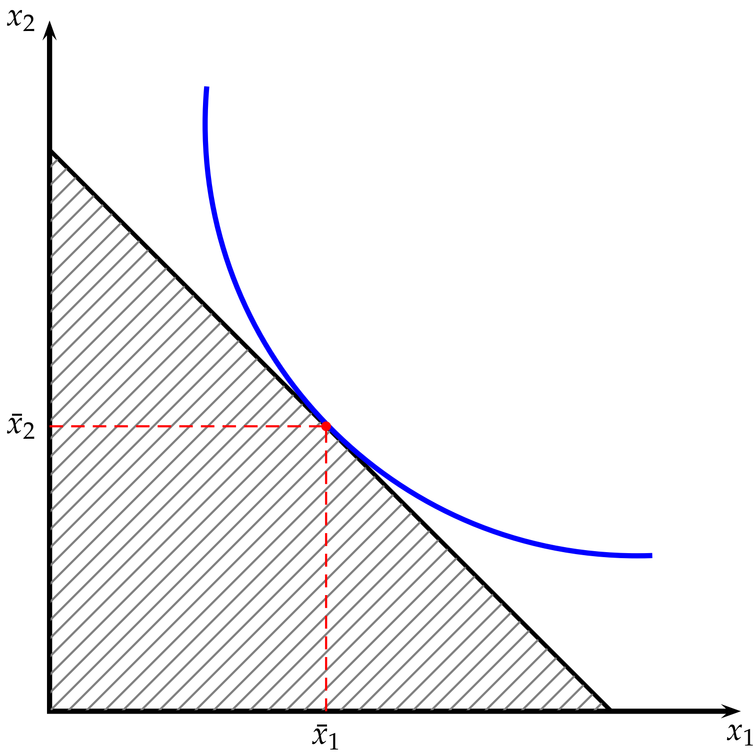
where α, β, p_1, p_2, and M are positive parameters, with $\alpha + \beta \leq 1$. Using Theorem 1, one can show that:

$$(\bar{x}_1, \bar{x}_2) = (\alpha M/(\alpha+\beta) p_1, \beta M/(\alpha+\beta) p_2)$$

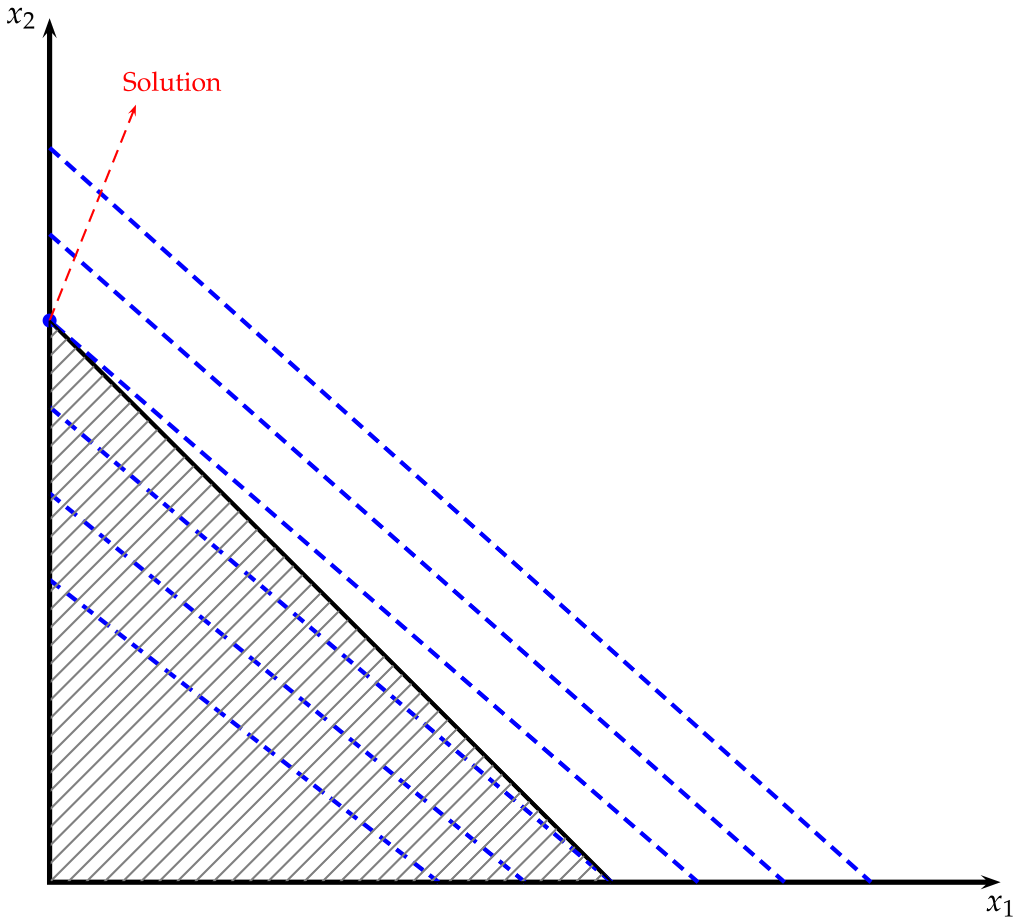
solves problem (UM).

18.3 Second Result

In some cases, Theorem 1 will not be applicable, either because there is no solution to the equations in condition (ii) of the theorem, or because the solution to the equations in condition (ii) is not non-negative.



A standard UMP with interior solution



A standard UMP with corner solution

We now provide a more general theorem, which is applicable in a larger number of cases than Theorem 1. [If Theorem 1 is applicable, then Theorem 2 is also applicable, but the converse is not true]. This result also provides sufficient conditions for a point in the constraint set to be a solution to problem (P); however, these conditions involve solving a system of simultaneous equations and inequalities.

Theorem 2 *Suppose assumptions (A1) and (A2) hold. Suppose we can find* $(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \geq 0$, *such that the following two conditions hold:*

(i) the first-order partial derivatives of f *and* g *with respect to* x_1 *and* x_2 *exist and are continuous at* $(\bar{x}_1, \bar{x}_2)$;

(ii) $(\bar{x}_1, \bar{x}_2, \bar{\lambda})$ *satisfies:*

$$\left.\begin{array}{l} D_1 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \leq 0 \\ \bar{x}_1 D_1 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \\ D_2 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \leq 0 \\ \bar{x}_2 D_2 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \\ D_3 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \geq 0 \\ \bar{\lambda} D_3 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \end{array}\right\} (KT)$$

Then, $(\bar{x}_1, \bar{x}_2)$ *solves problem* (P).

The conditions (KT) are called the *Kuhn-Tucker conditions*. Note that if $(\bar{x}_1, \bar{x}_2, \bar{\lambda})$ satisfies the equations stated in condition (ii) of Theorem 1, then $(\bar{x}_1, \bar{x}_2, \bar{\lambda})$ automatically satisfies (KT) stated in Theorem 2.

Consider the following utility maximization problem:

$$\left.\begin{array}{ll} Max & ax_1 + bx_2 \\ subject\ to & p_1 x_1 + p_2 x_2 \leq M \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (UM')$$

where a, b, p_1, p_2, and M are positive parameters. You can check that Theorem 1 cannot be applied to find a solution to problem (UM'). However, Theorem 2 can be used to find a solution to problem (UM'), for each of the following parameter configurations:

(i) $(a/b) > (p_1/p_2)$
(ii) $(a/b) < (p_1/p_2)$
(iii) $(a/b) = (p_1/p_2)$.

19 Proof of Theorem 2

We provide in this section the proof of Theorem 2. [Theorem 1 is a special case of Theorem 2 and therefore does not require a separate proof]. While knowing Theorems 1 and 2 is useful, understanding the proof gives one a better sense of the results, and enables one to develop variations to fit different contexts.

The proof is also instructive because it brings together the concepts of continuity, partial dervatives and concavity to establish one of the basic results of optimization theory.

The key idea of the proof is to use the basic inequality for concave functions stated as property (ii) in Section 16, on the Lagrangean L, and combine this with (KT) to get Theorem 2.

There is a technical difficulty though. Property (ii) for concave functions stated in Section 16 applies to the case where $\bar{x}$ and x' are in Y. For Theorems 1 and 2, we would like to allow for optimal solutions which occur at the boundary (as well as the interior) of X. To take account of this, we need to go through some additional steps, so the proof is longer than it would be if the $\bar{x}$, satisfying (KT), was in Y. This should not divert one from the key idea, summarized in the previous paragraph.

We first write down a Lemma, which allows us to conclude that the Lagrangean $L(x_1, x_2, \bar{\lambda})$ reaches a maximum at $(\bar{x}_1, \bar{x}_2)$ on the entire domain X (not just on the constraint set C). It is important to note that only part of the information contained in (KT) is used for this result.

Lemma 1 *Suppose assumptions (A1) and (A2) hold. Suppose we can find* $(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \geq 0$, *such that the following two conditions hold:*

(i) the first-order partial derivatives of f and g with respect to x_1 and x_2 exist and are continuous at $(\bar{x}_1, \bar{x}_2)$;

(ii) $(\bar{x}_1, \bar{x}_2, \bar{\lambda})$ *satisfies:*

$$\left.\begin{array}{l} D_1 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \leq 0 \\ \bar{x}_1 D_1 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \\ D_2 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \leq 0 \\ \bar{x}_2 D_2 L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) = 0 \end{array}\right\} (KT')$$

Then,

$$L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \geq L(x_1, x_2, \bar{\lambda}) \text{ for all } (x_1, x_2) \in X \tag{4}$$

Proof. Suppose, contrary to (4), that there is some $(x_1, x_2) \in X$ such that $L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) < L(x_1, x_2, \bar{\lambda})$. Denote:

$$\delta = L(x_1, x_2, \bar{\lambda}) - L(\bar{x}_1, \bar{x}_2, \bar{\lambda}) \tag{5}$$

Then $\delta > 0$.

Define:

$$M = \max\{|x_1 - \bar{x}_1|, |x_2 - \bar{x}_2|\} \tag{6}$$

Further, define for each $n \in \mathbb{N} = \{1, 2, 3, ...\}$,

$$\bar{x}(n) = \bar{x} + (e/n); \; x(n) = x + (e/n) \tag{7}$$

where $e = (1, 1)$. Then, $\bar{x}(n) \in Y$ and $x(n) \in Y$ for each $n \in \mathbb{N}$.

Pick $N \in \mathbb{N}$ such that for all $n \geq N$, we have:

$$\left.\begin{array}{l} (i)\ |D_1L(\bar{x}(n),\bar{\lambda}) - D_1L(\bar{x},\bar{\lambda})| \leq \delta/8M \\ (ii)\ \ |D_2L(\bar{x}(n),\bar{\lambda}) - D_2L(\bar{x},\bar{\lambda})| \leq \delta/8M \\ (iii)\ |L(\bar{x}(n),\bar{\lambda}) - L(\bar{x},\bar{\lambda})| \leq \delta/8 \\ (iv)\ |L(x(n),\bar{\lambda}) - L(x,\bar{\lambda})| \leq \delta/8 \end{array}\right\} \tag{8}$$

Conditions (iii) and (iv) can be met since f and g are continuous on X. Conditions (i) and (ii) can be met since f and g have partial derivatives at $\bar{x}$ and on Y, and these are continuous at $\bar{x}$.

Since f and g are concave on X, and $\bar{\lambda} \geq 0$, we have $L(x_1, x_2, \bar{\lambda})$ concave in (x_1, x_2) on X. Further $L(x_1, x_2, \bar{\lambda})$ has partial derivatives with respect to the first two arguments, which are continuous on Y.

We can now write:

$$\begin{aligned}
\delta &= L(x_1,x_2,\bar{\lambda}) - L(\bar{x}_1,\bar{x}_2,\bar{\lambda}) = [L(x_1,x_2,\bar{\lambda}) - L(x_1(n),x_2(n),\bar{\lambda})] + [L(x_1(n),x_2(n),\bar{\lambda}) \\
&\quad -L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda})] + [L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda}) - L(\bar{x}_1,\bar{x}_2,\bar{\lambda})] \\
&\leq (\delta/8) + [L(x_1(n),x_2(n),\bar{\lambda}) - L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda})] + (\delta/8) \\
&\leq (\delta/4) + D_1L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda})(x_1(n) - \bar{x}_1(n)) + D_2L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda})(x_2(n) - \bar{x}_2(n)) \\
&= (\delta/4) + [D_1L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda}) - D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})](x_1(n) - \bar{x}_1(n)) + D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_1(n) - \bar{x}_1(n)) \\
&\quad +[D_2L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda}) - D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})](x_2(n) - \bar{x}_2(n)) \\
&\quad +D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_2(n) - \bar{x}_2(n)) \\
&= (\delta/4) + [D_1L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda}) - D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})](x_1 - \bar{x}_1) + D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_1 - \bar{x}_1) \\
&\quad +[D_2L(\bar{x}_1(n),\bar{x}_2(n),\bar{\lambda}) - D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})](x_2 - \bar{x}_2) + D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_2 - \bar{x}_2) \\
&\leq (\delta/4) + (\delta/8M)M + D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_1 - \bar{x}_1) + (\delta/8M)M + D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_2 - \bar{x}_2) \\
&= (\delta/2) + D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_1 - \bar{x}_1) + D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})(x_2 - \bar{x}_2) \\
&= (\delta/2) + D_1L(\bar{x}_1,\bar{x}_2,\bar{\lambda})x_1 + D_2L(\bar{x}_1,\bar{x}_2,\bar{\lambda})x_2 \\
&\leq (\delta/2)
\end{aligned} \tag{9}$$

a contradiction, which establishes the Lemma.

We explain the various steps of display (9) as follows. The inequality in the third line of (9) follows from (8)(iii) and (8)(iv). The inequality in the fourth line is the crucial one, following from the concavity of $L(x_1, x_2, \bar{\lambda})$ in (x_1, x_2) on X, and the fact that $L(x_1, x_2, \bar{\lambda})$ has partial derivatives with respect to the first two arguments, which are continuous on Y. That is, we are using here property (ii) of concave functions, stated in Section 16.

The equality in the eighth line of (9) follows from the definitions of $x(n)$ and $\bar{x}(n)$ in (7). The inequality in the tenth line of (9) follows from (6), (8)(i) and (8)(ii). The equality in the twelfth line of (9) follows from the second and fourth conditions of (KT'). The inequality of the last line of (9) follows from the first and third conditions of (KT') and the fact that $(x_1, x_2) \in X = \mathbb{R}^2_+$. ■

We now provide the proof of Theorem 2.

Proof of Theorem 2. Applying Lemma 1, we have:

$$f(\bar{x}_1, \bar{x}_2) + \bar{\lambda}g(\bar{x}_1, \bar{x}_2) \geq f(x_1, x_2) + \bar{\lambda}g(x_1, x_2) \ \text{ for all } (x_1, x_2) \in X \tag{10}$$

Using the sixth condition of (KT), we get:

$$f(\bar{x}_1, \bar{x}_2) \geq f(x_1, x_2) + \bar{\lambda} g(x_1, x_2) \text{ for all } (x_1, x_2) \in X \tag{11}$$

Now, using the fifth condition of (KT), we know that $(\bar{x}_1, \bar{x}_2) \in C$. Let (x'_1, x'_2) be an arbitrary point in the constraint set C. Then, $g(x'_1, x'_2) \geq 0$. Since $\bar{\lambda} \geq 0$, we must therefore have $\bar{\lambda} g(x'_1, x'_2) \geq 0$. Thus, (11) implies:

$$f(\bar{x}_1, \bar{x}_2) \geq f(x'_1, x'_2) \text{ for all } (x'_1, x'_2) \in C$$

This establishes the Theorem. ■

20 Problems for Discussion

20.1 Application of Theorem 1

1. Consider the optimization problem:

$$\left.\begin{array}{ll} Max & x_1^{\alpha} x_2^{\beta} \\ subject\ to & p_1 x_1 + p_2 x_2 \leq M \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (P)$$

where α, β, p_1, p_2, and M are positive parameters, with $\alpha + \beta \leq 1$. Using Theorem 1, show that:

$$(\bar{x}_1, \bar{x}_2) = (\alpha M/(\alpha+\beta)p_1, \beta M/(\alpha+\beta)p_2)$$

solves problem (P).

2. Consider the optimization problem:

$$\left.\begin{array}{ll} Max & x_1 x_2 \\ subject\ to & p_1 x_1 + p_2 x_2 \leq M \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (Q)$$

where p_1, p_2, and M are positive parameters. Show that, in this case, the objective function:

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

is *not* a concave function on $\mathbb{R}^2_+$. However, Theorem 1 can still be used to solve problem (Q). How would you proceed to do this? Explain.

3. Define a production function, $F : X \to \mathbb{R}_+$ by:

$$F(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{x_1 + x_2} & if\ (x_1 + x_2) > 0 \\ 0 & if\ (x_1 + x_2) = 0 \end{cases}$$

where x_1 and x_2 are the amounts of two inputs used to produce a single output. Given input prices $w_1 > 0$ and $w_2 > 0$, consider the problem of finding the cost minimizing input combination of producing at least an output level of one unit:

$$\left.\begin{array}{ll} Min & w_1x_1 + w_2x_2 \\ subject\ to & F(x_1, x_2) \geq 1 \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (CM)$$

Use Theorem 1 to show that:

$$\bar{x}_1 = [w_1^{\frac{1}{2}} + w_2^{\frac{1}{2}}]/w_1^{\frac{1}{2}}$$
$$\bar{x}_2 = [w_1^{\frac{1}{2}} + w_2^{\frac{1}{2}}]/w_2^{\frac{1}{2}}$$

solves problem (CM).

20.2 Application of Theorem 2

4. Consider the following utility maximization problem:

$$\left.\begin{array}{ll} Max & ax_1 + bx_2 \\ subject\ to & p_1x_1 + p_2x_2 \leq M \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (UM)$$

where a, b, p_1, p_2, and M are positive parameters. Use Theorem 2 to find a solution to problem (UM), for each of the following parameter configurations:
(i) $(a/b) > (p_1/p_2)$
(ii) $(a/b) < (p_1/p_2)$
Draw appropriate diagrams to illustrate your solutions.

5. Consider the following utility maximization problem:

$$\left.\begin{array}{ll} Max & ax_1 + b[x_2/(1 + x_2)] \\ subject\ to & p_1x_1 + p_2x_2 \leq M \\ and & (x_1, x_2) \geq 0 \end{array}\right\} (R)$$

where a, b, p_1, p_2, and M are positive parameters. Use Theorem 2 to find a solution to problem (R), for each of the following parameter configurations:
(i) $(a/b) > (p_1/p_2)$
(ii) $(a/b)[1 + (M/p_2)]^2 < (p_1/p_2)$
(iii) $(a/b) < (p_1/p_2) < (a/b)[1 + (M/p_2)]^2$
Draw appropriate diagrams to illustrate your solutions.

Bibliography for Review Notes

The Review Notes are based on the following books:

1. Apostol, T.M., Mathematical Analysis, Second Edition, Addison-Wesley 1974.
2. Goldberg, R.R., Methods of Real Analysis, Blaisdell 1964.
3. Roberts, A.W. and D.E. Varberg, Convex Functions, Academic Press 1973.
4. Rudin, W., Principles of Mathematical Analysis, Third Edition, McGraw-Hill 1976.
5. Stricharz, R.S., The Way of Analysis, Jones and Bartlett 2000.

The material in the first four parts can be found in any standard text on real analysis. References 2, 4 and 5 are recommended for further reading on this material.

References 1 and 4 are recommended for further reading on the material in the fifth part.

Material in the last two parts are typically not covered in a course on real analysis, but in specialized courses on optimization. Reference 3 is recommended for further reading on this material.