Math Prep Notes[1]

Joel Sobel and Joel Watson

2008, under revision

Preface

These notes are the starting point for a math-preparation book, primarily for use by UCSD students enrolled in Econ 205 (potentially for use by folks outside UCSD as well). The first draft consists of a transcription of Joel Watson's handwritten notes, as well as extra material added by Philip Neary, who worked on the transcription in 2006. Joel Sobel and Joel Watson have revised parts of these notes and added material, but the document is still rough and disorganized. Surely there are many mistakes. These material here is incomplete and contain many mistakes. If you find an error, a notational inconsistency, or other deficiency, please let one of the Joels know.

Contents

Notes on notation and items to be corrected:

1. Sets are typically denoted by italic (math type) upper case letters; elements are lower case.

2. PN used script letters to denote sets and there may be remnants of this throughout (such as in some figures, all which must be redrawn anyway).

3. ε (varepsilon) is the preferred epsilon symbol.

4. Standard numerical sets (reals, positive integers, etc.) are written using the mathbb symbols: $\mathbb{R}$, $\mathbb{P}$, and so on.

Chapter 1

Sets, Functions, and the Real Line

This chapter reviews some basic definitions regarding sets and functions, and it contains a brief overview of the construction of the real line.

1.1 Sets

The most basic of mathematical concepts is a *set*, which is simply a collection of objects. For example, the "days of the week" is a set comprising the following objects: Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday. The set of Scandinavian countries consists of: Sweden, Norway, Denmark, Finland and Iceland. Sets can be composed of any objects whatsoever.

We often use capital italic letters to denote sets. For instance, we might let D denote the set of days of the week and S denote the set of Scandinavian countries. We will use lowercase italic letters to denote individual objects (called *elements* or *points*) in sets. Using the symbol "$\in$," which means "is an element of," we thus write $x \in X$ to indicate that x is an element of set X. By the way, the symbol "$\notin$" means "is not an element of," so we would write $x \notin X$ to mean that X does not contain x.

To define a set, it is sometimes convenient to list its elements. Formally, we do this by enclosing the list in curly brackets and separating the elements with commas. For instance, the set of Scandinavian countries is

$$S = \{\text{Sweden, Norway, Denmark, Finland, Iceland}\}$$

and the set of days of the week is

$$D = \{\text{Monday, Tuesday, Wednesday, } \ldots, \text{ Sunday}\}.$$

When a set contains many elements, it is useful to define it by making reference to a generic element x, using the "such as" symbol "|", and including qualifying statements (properties that x is required to have). For instance, the set of numbers between 2 and 5 (including the endpoints) is

$$\{x \mid 2 \leq x \leq 5\}.$$

The set that contains no elements is called the *empty set* and is denoted $\emptyset$. If a set has at least one element, it is called *nonempty*.

Here are other common definitions that we will use frequently:

Definition 1. *A set Y is called a* **subset of** *set X, written $Y \subset X$, if every element of Y is also an element of X (that is, $x \in Y$ implies $x \in X$). If, in addition, $Y \neq X$, then we say that Y is a* **proper subset** *of X.*

Note that the set of "weekend days," {Saturday, Sunday}, is a proper subset of the set of days of the week.

Observe that if sets X and Y are both subsets of each other, then they must be equal. Using the symbol $\Rightarrow$, which means "implies," this conclusion can be expressed as: $X \subset Y$ and $Y \subset X \Rightarrow X = Y$. Sometimes when we want to prove that two sets are equal, we perform the two steps of showing that each is contained in the other.

Definition 2. *The* **union** *of sets X and Y, written $X \cup Y$, is the set formed by combining the distinct elements of the two sets. That is,*

$$X \cup Y \equiv \{x \mid x \in X \text{ or } x \in Y \text{ (or both)}\}.$$

The symbol "$\equiv$", used first in the definition above, means "is defined as" or "is equivalent to."

Definition 3. *The* **intersection** *of sets X and Y, written $X \cap Y$, is the set formed by collecting only the points that are common to both sets. That is,*

$$X \cap Y \equiv \{x \mid x \in X \text{ and } x \in Y\}.$$

Definition 4. *The* **difference** *between sets X and set Y, written $X \setminus Y$, is the set formed by removing from X all points in the intersection of X and Y. That is,*

$$X \setminus Y \equiv \{x \mid x \in X \text{ and } x \notin Y\}.$$

When we are analyzing various sets, we typically have in mind a grand set that contains all objects of interest, such that every set we study is a subset of the grand set. This grand set is called the *space*, or *universe*, and is sometimes denoted U. The next definition makes reference to the space.

Figure 1.1: Proper subset.

$$\mathcal{A} \subset \mathcal{B}$$

Figure 1.2: Union of two sets.

$$\mathcal{A} \cup \mathcal{B}$$

Definition 5. *The* **complement** *of the set X, denoted X' or X^c, is defined as $U \setminus X$.*

Example 1. *Suppose that the space is all of the lowercase letters:*

$$U \equiv \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \ldots, \mathrm{z}\}.$$

Consider $X = \{a, b, c, d\}$ and $Y = \{d, e, f, g\}$. Then we see that $X \cup Y = \{a, b, c, d, e, f, g\}$, $X \cap Y = \{\mathrm{d}\}$, $X \setminus Y = \{a, b, c\}$, and $X^c = \{e, f, g, \ldots, z\}$.

For visual thinkers, Venn diagrams can be useful to represent the relations between sets. Figure 1.1 represents that set A (the points inside the smaller circle) is a proper subset of set B (the points inside the larger circle). The shaded region of Figure 1.2 shows the union of sets A and B, whereas the shaded region in Figure 1.3 is the intersection.

1.2 Functions

Often we are interested in representing ways in which elements of one set might be related to, or associated with, elements of another set. For example, for sets X

Figure 1.3: Intersection of two sets

$$\mathcal{A} \cap \mathcal{B}$$

Figure 1.4: A function α..

and Y, we might say that every $x \in X$ points to, or "maps to," a point $y \in Y$. The concept of a "function" represents such a mapping.

Definition 6. *A* **function** f *from a set* X *to a set* Y *is a specification (a mapping) that assigns to each element of* X *exactly one element of* Y*. Typically we express that* f *is such a mapping by writing* $f : X \to Y$*. The set* X *(the points one "plugs into" the function) is called the* **domain** *of the function, and the set* Y *(the items that one can get out of the function) is called the function's* **codomain**.

Note that the key property of a function is that every element in X is associated with *exactly one* element in Y. If you plug a point from X into the function, it specifies just one point in Y that is associated with it. It is not the case that some point $x \in X$ maps to both y and y' such that $y \neq y'$.

We write $f(x)$ as the point in Y that the function associates with $x \in X$. Also, for any subset of the domain $Z \subset X$, we define

$$f(Z) \equiv \{f(x) \mid x \in Z\}.$$

We refer to this as the **image** of the set Z for the function f. The set $f(X)$, which is clearly a subset of Y, is called the **range** of the function. Finally, we define the **inverse image** of a set $W \subset Y$ as

$$f^{-1}(W) \equiv \{x \in X \mid f(x) \in W\}.$$

This is the set of points in X that map to points in W. Note that the inverse image does not necessarily define a function from Y to X, because it could be that there are two distinct elements of X that map to the same point in Y.

Figures 1.4 and 1.5 depict functions from S to T. Figure 1.6 depicts a map-

Figure 1.5: A function β..

ping that is not a function. The mapping does not associate any element in T with the point $y \in S$. Another violation of the requirements for a function is that the mapping associates *two* elements of T (both 1 and 3) with the single point $x \in S$.

Figure 1.6: A mapping that is not a function.

Example 2. *With α and β shown in Figures 1.4 and 1.5, we have $\alpha(\{x, z\}) = \{2, 3\}$ and $\beta^{-1}(\{1, 2\}) = \{x, z\}$.*

Definition 7. *Consider a function $f : X \to Y$. If $f(X) = Y$ then f is said to be* **onto***. That is, f is onto if for each $y \in Y$ there is at least one $x \in X$ that maps to it.*

Note that the function α in Example 2 is onto while the function β is not, since its image is $\{1, 3\}$ which is a proper subset of $\{1, 2, 3\}$, the codomain of β.

Definition 8. *Consider a function $f : X \to Y$. This function is said to be* **one-to-one** *if $f(x) = f(x')$ implies that $x = x'$. An equivalent condition is that for every two points $x, x' \in X$ such that $x \neq x'$, it is the case that $f(x) \neq f(x')$.*

Note that the function α in Example 2 is one-to-one, wherease the function β is not one-to-one because $\beta(x) = \beta(z)$ and yet $x \neq z$.

Definition 9. *A function $f : X \to Y$ is called* **invertible** *if the inverse image mapping is a function from Y to X. Then, for every $y \in Y$, $f^{-1}(y)$ is defined to be the point $x \in X$ for which $f(x) = y$.*

Theorem 1. *A function is invertible if and only if it is both one-to-one and onto.*

This section concludes with the definition of a composition function, which is basically a function formed by performing the operations of two functions in order.

Definition 10. *Suppose we have functions $g : X \to Y$ and $f : Y \to Z$. Then the* **composition function** *$f \circ g$ is a mapping from X to Z (that is, $f \circ g : X \to Z$. Writing $h = f \circ g$, this function is defined by $h(x) \equiv g(f(x))$ for every $x \in X$.*

1.3 The Real Line

There are several sets of numbers that we will work with often. Perhaps the most important for everyday mathematics is the set of **real numbers**, denoted by $\mathbb{R}$.

The set of reals comprises all of the ordinary numbers you are used to dealing with, such as 0, 1, fractions like $1/2$, decimals such as 4.23, and more exotic numbers like the "natural number" e. The set includes both positive and negative numbers, arbitrarily high numbers, and arbitrarily low (large negative) numbers.

Although many of its elements are quite familiar to most people, the definition of $\mathbb{R}$ is not as straightfoward. In fact, $\mathbb{R}$ is defined in relation to the **real number system**, which includes the set $\mathbb{R}$, operators of addition and multiplication, standard "identiity numbers" 0 and 1, and some axioms (assumptions). Here is a brief description of how the set $\mathbb{R}$ is defined. It is okay to not study a complete treatment, but realize that the real number system has some nuances.

A good place to start is with the set of positive integers:

$$\mathbb{P} \equiv \{1, 2, 3, \ldots\}.$$

One way to think of this set is that it is defined by the special "multiplicative identity" number 1 and the idea of addition. The number 1 is called the multiplicative identify because 1 times any number is the same number. The set of positive integers comprises the number 1, $1 + 1$, $1 + 1 + 1$, and all such numbers formed by adding 1 to itself multiple times.

Example 3. *Define functions α and β from the set of positive integers $\mathbf{P} = \{1, 2, 3, \ldots\}$ to itself by $\alpha(n) = 2n$ for all n, and*

$$\beta = \begin{cases} (n+1)/2, & \text{if } n \text{ is odd}, \\ n/2, & \text{if } n \text{ is even}. \end{cases}$$

These are pictured in Figure 1.7. Note that α is one-to-one but not onto, and β is

Figure 1.7: α and β.

onto but not one-to-one.

Definition 11. *If there exists a 1-1 function of $\mathcal{X}$* onto *$\mathcal{Y}$, we say that $\mathcal{X}$ and $\mathcal{Y}$ can be put in a 1-1* correspondence*, or that $\mathcal{X}$ and $\mathcal{Y}$ have the same cardinal number, or briefly, that $\mathcal{X}$ and $\mathcal{Y}$ are* equivalent*, and we write $\mathcal{X} \sim \mathcal{Y}$.*

This relation has the following properties:

- *It is reflexive: $\mathcal{X} \sim \mathcal{X}$.*

- *It is symmetric: If $\mathcal{X} \sim \mathcal{Y}$, then $\mathcal{Y} \sim \mathcal{X}$.*
- *It is transitive: If $\mathcal{X} \sim \mathcal{Y}$ and $\mathcal{Y} \sim \mathcal{Z}$, then $\mathcal{X} \sim \mathcal{Z}$.*

Any relation with these three properties is called an equivalence relation.

Note: there may be lots of functions from $\mathcal{X}$ onto $\mathcal{Y}$ that are not 1-1, but for this condition we only need to be able to find 1 such function!

Definition 12. *For any positive integer n, let $\mathbb{P}_n$ be the set whose elements are the integers 1, 2, ..., n; Let $\mathbb{P}$ be the set consisting of all positive integers (which is the set of natural numbers we already saw in Example 3).*

For any set $\mathcal{A}$, we say:

- $\mathcal{A}$ is *finite* if $\mathcal{A} \sim \mathbb{P}_n$ for some n.
- $\mathcal{A}$ is *infinite* if $\mathcal{A}$ is not finite.
- $\mathcal{A}$ is *countable* if $\mathcal{A} \sim \mathbb{P}$.
- $\mathcal{A}$ is *uncountable* if $\mathcal{A}$ is neither finite nor countable.
- $\mathcal{A}$ is *at most countable* if $\mathcal{A}$ is finite or countable.

Note: see section **??** for a further discussion of countable versus uncountable.

Definition 13. *The rational numbers, $\mathbb{Q}$[1], are informally defined as the set of all numbers of the form m/n, where m and n are integers and $n \neq 0$.*

Definition 14. *The Real Numbers, $\mathbb{R}$[2], are informally defined as the set of all regular numbers from negative to positive.*

$$\mathbb{R} = \{\, x \mid -\infty < x < \infty \}$$

Formally[3] *$\mathbb{R}$ is defined as a set of objects associated with a set of "operators" (function from $\mathcal{X} \times \mathcal{X} \to \mathcal{X}$) "+" and ".", and special elements 0 and 1, satisfying certain axioms:*

[1] $\mathbb{Q}$ is an example of a countable set. This may seem very confusing since there must be more rational numbers than integers (mustn't there be!), and will be discussed further in section **??**

[2] $\mathbb{R}$ is an example of an uncountable set

[3] For mathsy types in the class; really first of all you have to define a *field*, then an ordered *field*, and then $\mathbb{R}$ can be defined as the ordered field which has the least upper bound property (to be defined soon) and which contains $\mathbb{Q}$ as a subfield. Finally you must show that $\mathbb{R} \setminus \mathbb{Q}$, the irrationals, can be written as infinite decimal expansions and are considered "approximated" by the corresponding finite decimals. If after all this you still care, see Rudin chapter 1.

- *Addition: If* $x, y \in \mathbb{R} \Rightarrow x + y \in \mathbb{R}$.
- *Additive Identity:* $x + 0 = 0 + x, \ \ \forall\, x \in \mathbb{R}$
- *Multiplicative Identity:* $1.x = x.1, \ \ \forall\, x \in \mathbb{R}$
- *Associativity:* $(x + y) + z = x + (y + z), \ \ \forall\, x, y, z \in \mathbb{R}$

Definition 15. *There is a set* $\mathbb{R}_{++} \subset \mathbb{R}$ *such that*

- *If* $x, y \in \mathbb{R}_{++} \Rightarrow x + y \in \mathbb{R}_{++}$.
- *If* $x \in \mathbb{R} \Rightarrow x \in \mathbb{R}_{++}$ *or* $x = 0$ *or* $-x \in \mathbb{R}_{++}$
  where -x $\equiv$ *so that (-x) + x = 0*

It is fairly obvious from this that $\mathbb{R}_{++}$ is just the set of strictly positive real numbers.

Definition 16. *We say*

$$x > y \quad \Longleftrightarrow \quad x - y \in \mathbb{R}_{++}$$

where

$$-y \equiv +(-y)$$

Definition 17. *For* $a, b \in \mathbb{R}$ *where* $a \leq b$

$$\begin{array}{ll} [a, b] \equiv \{\, x \mid a \leq x \leq b\} & \textit{(closed interval)} \\ (a, b] \equiv \{\, x \mid a < x \leq b\} & \textit{(half-open interval)} \\ [a, b) \equiv \{\, x \mid a \leq x < b\} & \textit{(half-open interval} \\ (a, b) \equiv \{\, x \mid a < x < b\} & \textit{(open interval / segment)} \end{array}$$

Definition 18. *Take* $\mathcal{X} \subset \mathbb{R}$
Then we say $a \in \mathbb{R}$ *is an* upper bound *for* $\mathcal{X}$ *if*

$$a \geq x, \ \ \forall\, x \in \mathcal{X}$$

And we say $b \in \mathbb{R}$ *is a* lower bound *for* $\mathcal{X}$ *if*

$$b \leq x, \ \ \forall\, x \in \mathcal{X}$$

Note that there can be lots of upper bounds and lots of lower bounds.

We say "The set $\mathcal{X}$ is bounded" if "$\mathcal{X}$ is bounded from above" and "$\mathcal{X}$ is bounded from below".

Definition 19. *We say $a \in \mathbb{R}$ is a* least upper bound of $\mathcal{X}$*, or the* supremum *of $\mathcal{X}$, if $a \geq x, \forall\, x \in \mathcal{X}$ (a is an upper bound), and if a' is also an upper bound then $a' \geq a$. We write*

$$a = \sup \mathcal{X}$$

Definition 20. *We say $b \in \mathbb{R}$ is a* greatest lower bound of $\mathcal{X}$*, or the* infimum *of $\mathcal{X}$, if $b \leq x, \forall\, x \in \mathcal{X}$ (b is an lower bound), and if b' is also an lower bound then $b' \leq b$. We write*

$$b = \inf \mathcal{X}$$

Example 4. *The following illustrates the important point that the "*$\sup$*" and "*$\inf$*" of a set do not have to be contained in the set.*
For example let

$$\mathcal{X} = [0, 1], \text{ and } \mathcal{Y} = (0, 1)$$

We can see that

$$\textit{sup } \mathcal{X} = \textit{sup } \mathcal{Y} = 1$$

and that

$$\textit{inf } \mathcal{X} = \textit{inf } \mathcal{Y} = 0$$

However, obviously the points $\{0\}$ and $\{1\}$ are not contained in the set $\mathcal{Y}$.

Lemma 1. *Any set $\mathcal{X} \subset \mathbb{R}$, ($\mathcal{X} \neq \phi$), that has an upper bound, has a least upper bound (sup).*

Proof. Left as exercise to the reader.[4] □

[4] Any reader able to complete this exercise without assistance should have a successful career as a mathematician. Although the result seems intuitive, it is a deep property of the real numbers.

Definition 21. *We define the following*

$$\textit{max } \mathcal{X} \equiv \textit{a number } a \textit{ such that } a \in \mathcal{X} \textit{ and } a \geq x,\ \forall\, x \in \mathcal{X}$$

and

$$\textit{min } \mathcal{X} \equiv \textit{a number } b \textit{ such that } b \in \mathcal{X} \textit{ and } b \leq x,\ \forall\, x \in \mathcal{X}$$

Referring back to Example 4 we can see that $\mathcal{X}$ = [0,1] has a max and min

$$\max \mathcal{X} = 1 \qquad \text{and} \qquad \min \mathcal{X} = 0$$

While $\mathcal{Y}$ = (0,1) does not have a max or min.

Note It should now be clear that the max and min do not always exist even if the set is bounded, but the sup and the inf do always exist if the set is bounded.

Theorem 2. *If max $\mathcal{X}$ exists then*

$$\textit{max } \mathcal{X} = \textit{sup } \mathcal{X}$$

Proof. Let $a \equiv \max \mathcal{X}$
We will demonstrate that a is sup $\mathcal{X}$.
To do this we must show that:
(i) a is an upper bound
(ii) Every other upper bound a' satisfies $a' \geq a$
so
(i) a is an upper bound on $\mathcal{X}$ since by the definition of max

$$a \geq x, \quad \forall\, x \in \mathcal{X}$$

(ii) Consider any other upper bound a' of the set $\mathcal{X}$ such that $a' \neq a$.
Since a' is an upper bound, and $a \in \mathcal{X}$ (by the fact that a = max $\mathcal{X}$), we must have $a' > a$. □

Definition 22 (Certain Sets)**.** *Positive Integers:*

$$\mathbb{P} \equiv \{\, 1,\ 1+1,\ 1+1+1,\ \ldots \}$$

Integers:

$$\mathbb{Z} \equiv \mathbb{P} \cup -\mathbb{P} \cup \{\, 0\}$$

Rational Numbers:

$$\mathbb{Q} \equiv \left\{ \frac{m}{n} \mid m, n \in \mathbb{Z},\ n \neq 0 \right\}$$

It can be shown that

$$\mathbb{Q} \subset \mathbb{R}$$

and

$$\mathbb{Q} \neq \mathbb{R}$$

1.4 Methods of Proof

Proof by Direction Implication/Construction

Example 5. *Show that*

$$max\ \mathcal{X} = sup\ \mathcal{X}$$

when max$\mathcal{X}$ exists.

Proof. Refer back to Theorem 2 □

Theorem 3 (Direct Implication)**.** *Distributive Law for Sets*

$$\mathcal{A} \cap (\mathcal{B} \cup \mathcal{C}) = (\mathcal{A} \cap \mathcal{B}) \cup (\mathcal{A} \cap \mathcal{C})$$

To prove this, let the left hand side and right hand side of the above equation be denoted by $\mathcal{D}$ and $\mathcal{E}$ respectively. We will show that

$$\mathcal{D} \subset \mathcal{E} \quad \textit{and that} \quad \mathcal{E} \subset \mathcal{D}$$

Note. *Subsets going in each direction imply equality!*

Proof. Suppose $x \in \mathcal{D}$. Then $x \in \mathcal{A}$ and $x \in \mathcal{B} \cup \mathcal{C}$, that is, $x \in \mathcal{B}$ or $x \in \mathcal{C}$ (possibly both). Hence $x \in \mathcal{A} \cap \mathcal{B}$ or $x \in \mathcal{A} \cap \mathcal{C}$, so that $x \in \mathcal{E}$. Thus $\mathcal{D} \subset \mathcal{E}$.

Next, suppose $x \in \mathcal{E}$. Then $x \in \mathcal{A} \cap \mathcal{B}$ or $x \in \mathcal{A} \cap \mathcal{C}$. That is, $x \in \mathcal{A}$, and $x \in \mathcal{B} \cup \mathcal{C}$. Hence $x \in \mathcal{A} \cap (\mathcal{B} \cup \mathcal{C})$, so that $\mathcal{E} \subset \mathcal{D}$.

It follows that $\mathcal{D} = \mathcal{E}$.

Proof by Contradiction These proofs usually begin with, *let's assume the statement does not hold* and then at the end we find that this cannot be.

Example 6. *Suppose there is a two person world consisting of John and Mary, and in this world everybody wears properly fitting clothes. Suppose people fit into the clothes of people the same size or bigger, but do not fit into the clothes of smaller people. We denote John's height by x, and Mary's height by y. We want to show that John is taller than Mary, i.e. that*

$$x > y$$

So the way we proceed is that we make the initial assumption that Mary is at least as tall as John, i.e. $y \geq x$. But then we note that John's clothes are all bigger than Mary's clothes. So John must not fit into Mary's clothes. Thus we must have that our initial assumption is false. So

$$x > y$$

Now this is a pretty simple and stupid example but highlights the way to proceed.

Example 7. *Suppose we have 2 sets $\mathcal{A}$ and $\mathcal{B}$. And we have that $\mathcal{A}$ is a subset of $\mathcal{B}$. They are both subsets of the universal set $\mathcal{U}$.*

FIGURE 9 GOES HERE

The following statements are all equivalent (you should convince yourself of this)

- If $x \in \mathcal{A} \Rightarrow x \in \mathcal{B}$,
  $\Rightarrow \mathcal{A} \subset \mathcal{B}$
- If $x \in \mathcal{B}^c \Rightarrow x \in \mathcal{A}^c$,
  $\Rightarrow \mathcal{B}^c \subset \mathcal{A}^c$
- $x \notin \mathcal{B} \Rightarrow x \notin \mathcal{A}$

Basically this method of proof hinges on showing that if we have 2 statements p and q. Then

$$p \Rightarrow q \quad \Longleftrightarrow \quad \neg q \Rightarrow \neg p$$

Lemma 2.

$$\begin{aligned} &\mathcal{X} \subset \mathbb{R},\ a = \sup\mathcal{X},\ \epsilon > 0 && \text{(let this be statement p)} \\ \Rightarrow\ &\exists\, x \in \mathcal{X} \text{ such that } a - \epsilon < x \leq a && \text{(let this be statement q)} \end{aligned}$$

Proof. Suppose the conclusion (i.e. statement q) does not hold.
Assume there is no $x \in \mathcal{X}$ such that

$$a - \epsilon < x \leq a$$

By the definition of supremum, we know that a is an upper bound.
Since there are no points in $\mathcal{X}$ between $a - \epsilon$ and a (and obviously nothing above a), we must conclude that $a - \epsilon$ is an upper bound of $\mathcal{X}$.
But this contradicts that a is sup$\mathcal{X}$. □

Proof by Induction This involves making a statement for n,

$$S(n) = \text{“claim } \ldots \text{ has to do with } n\text{”}$$

To prove this claim for all n, it's sufficient to prove

- $S(1)$ is true
- $S(n) \Rightarrow S(n+1)$

Example 8. *Show that*

$$\sum_{k=1}^{n} k^3 = \frac{1}{4}n^2(n+1)^2$$

Proof.

- Observe that

$$\begin{aligned} S(1) &= \frac{1}{4}(1)^2(1+1)^2 \\ &= 1 \end{aligned}$$

- Suppose the claim is true for $n = m$

$$\sum_{k=1}^{m} k^3 = \frac{1}{4}m^2(m+1)^2$$

Now look at $n = m + 1$

$$\begin{aligned}\sum_{k=1}^{m+1} k^3 &= \sum_{k=1}^{m} k^3 + (m+1)^3 \\ &= \tfrac{1}{4}m^2(m+1)^2 + (m+1)^3 \\ &= (m+1)^2(\tfrac{1}{4}m^2 + m + 1) \\ &= \tfrac{1}{4}(m+1)^2(m^2 + 4m + 4) \\ &= \tfrac{1}{4}(m+1)^2(m+2)^2\end{aligned}$$

So you can see that if we assume the original statement is true for m and now everywhere that we had an m we have an $m + 1$. Thus it must be true for $m + 1$. And so we are done. □

1.5 Some helpful notes

If and only if

You will have noticed in preceding sections mentioning of things like "*if and only if*" (which is often abbreviated to *iff*), and phrases like "A is necessary for B", or "x is sufficient for y". You will also have noticed the symbol $\iff$. So what the hell do all these things mean? Thankfully they are all very closely related.

The symbol $\iff$ is just the mathematical symbol for "if and only if"[5]. But that's not particularly helpful if we don't know what "if and only if" means.

The difference between "if" and "iff" is as follows. Compare the 2 sentences below:

1. I will drink beer if the beer is a Guinness. (equivalently: If the beer is a Guinness, then I will drink it)

2. I will drink beer if and only if the beer is a Guinness.

Sentence (1) says only that I will drink Guinness. It does not rule out that I may also drink Budweiser. Maybe I will, maybe I won't - there just is not enough information to determine. All we know for sure is that I will drink Guinness.

[5] in fact many symbols are used for *if and only if*, such as $\iff$, $\equiv$, and $\longleftrightarrow$, so watch out as different authors may use different ones.

Sentence (2) makes it quite clear that I will drink Guinness and Guinness only. I won't drink any other type of beer[6]. Also, I will definitely drink the beer if it is a Guinness.

This may seem confusing so perhaps we should look at how proofs involving "iff" are special. Consider the First Fundamental Theorem of Asset Pricing due to Harrison and Kreps[7].

Theorem 4. *The finite market model is viable* if and only if *there exists an equivalent martingale measure.*

So perhaps it is better to look at the statement of the theorem more carefully (with some shortening so as to fit the page better).

Theorem 5. $\underbrace{\textit{viable}}_{A}$ if and only if $\underbrace{\textit{equivalent martingale measure.}}_{B}$

Or in mathematical symbols

Theorem 6. $A \iff B$.

Now you might have noticedn that the $\iff$ is sort of a fusion of the symbols $\Longleftarrow$, and $\Longrightarrow$. This was not an accident. To prove an *iff* statement, you must prove the implication both ways. To see what is meant by this let's look at a mock proof for the above theorem.

Basically to prove a statement $A \iff B$, you must show

- $A \Longrightarrow B$
- $B \Longrightarrow A$

Of course you can also prove that

- $A \Longrightarrow B$
- $\neg A \Longrightarrow \neg B$

since recall that proving $B \Rightarrow A$ is the same as proving $\neg A \Rightarrow \neg B$[8].

Ok, so now let's run through a mock proof[9] of The First Fundamental Theorem of Asset Pricing

[6] I'm not really like this:)

[7] Don't worry if you don't understand what the Theorem means, the method of proof is what's important

[8] or now to show we really get it, we could write $[B \Rightarrow A] \iff [\neg A \Rightarrow \neg B]$!

[9] There is not even a hint of how to prove the theorem here. I'm just trying to show you how you go about proving such types of statements.

Proof. ($\Rightarrow$) So to prove the implication this direction, we first assume that the finite market model is viable. Or more simply, we assume A. Thus, having made this assumption, we now have to show that armed with this assumption, that there exists an equivalent martingale measure. We have to show B.

($\Leftarrow$) Now we must go the other way. So we start by assuming there is an equivalent martingale measure, i.e. we assume B. Then using only this assumption, we must some how end up with the statement that the finite market model is viable (statement A).

Once we do both of these we are done. Make sense?? □

A word of caution. If you are reading a paper, and it makes a **Definition** statement, then any *if* in this type of statement should be read as *iff*. This is true since it is a definition, so by definition the *if* is an *iff* (what!).

For example, see back to Definition 18 on page 12. For clarity's sake I'll restate the first part.

Definition 23. *Take* $\mathcal{X} \subset \mathbb{R}$
Then we say $a \in \mathbb{R}$ *is an* upper bound *for* $\mathcal{X}$ *if*

$$a \geq x, \quad \forall\, x \in \mathcal{X}$$

So obviously if $a \geq x, \quad \forall\, x \in \mathcal{X}$, then we can say $a \in \mathbb{R}$ is an *upper bound* for $\mathcal{X}$. But by the same token, since it is a definition, if $a \in \mathbb{R}$ is an *upper bound* for $\mathcal{X}$, then we can also say that $a \geq x, \quad \forall\, x \in \mathcal{X}$. Make sense???

Necessary and Sufficient

Supposing we have a statement $P \iff Q$. Then we can also say P is "necessary and sufficient" for Q. Since proving $P \iff Q$, requires implications going 2 different directions, and "necessary" and "sufficient" are a list of 2 words, you can probably guess that there is a connection between them. In fact they mean equivalent things. Or to be quite cryptic

$$[P \iff Q] \iff [P \text{ is necessary and sufficient for } Q]$$

Necessary Condition

As usual, maybe it's best to start with an example.

Example 9. *Suppose that I am the greatest football*[10] *player in the world. I am therefore the best football player in Ireland too. So we could make a statement like:*

"I am the best football player in Ireland if I am the best football player in the world" (equivalently: If I am the best football player in the world, then I am the best football player in Ireland).

Note however that this is not an if and only if *statement. It does not go both ways. If I am the best football player in Ireland, we cannot conclude that I am the best football player in the world (one of those Brazilians might be better than me!).*

Thus we would say that being the best football player in Ireland is a necessary condition for being the best football player in the world, since if I am the best football player in the world, then I am necessarily the best football player in Ireland.

Sufficient Condition

Again it's probably best to start with an example.

Example 10. *Supposing you are studying for the mathcamp final while sitting on a bench on the cliffs overlooking Torrey Pines. Suppose you are becoming increasingly frustrated with the course and are really getting sick of studying. You decide to hurl your notes off the cliff so that they land in the sea where they will never be seen again. Then we would say that the hurling of the notes is a sufficient condition for the notes to land in the sea. But the notes landing in the sea does not imply that they were hurled: you might have slipped and dropped them, or a gust of wind may have come and taken them. Thus the hurl was sufficient but not necessary for the notes to land in the sea.*

[10] football refers to soccer not "American Football"

Relationship between necessity and sufficiency

You guessed it, necessity and sufficiency are dual to one another. Look back to example 9. We have that the following are all equivalent:

- "(best football player in Ireland) if (best football player in the world)"
- (best football player in the world) $\Rightarrow$ (best football player in Ireland)
- (best football player in the world) *is sufficient for* (best football player in Ireland)
- (best football player in Ireland) *is necessary for* (best football player in the world)

Infinity

The notion of infinity is one of the most seemingly simple yet difficult concepts in all of mathematics. The symbol for infinity is ∞. Intuitively, ∞ just represents a really "big" number that "beats" all other numbers. We got an introduction to how to measure the size of sets on pages 10 and 11.

But before we get to comparing sizes of sets, let's just look at how infinity is defined.

Definition 24. *The* extended real number system*,* $\overline{\mathbb{R}}$*, consists of* $\mathbb{R}$[11] *and two symbols,* $+\infty$ *and* $-\infty$*. We define*

$$-\infty < x < +\infty$$

for every $x \in \mathbb{R}$*.*

We make the following assumptions concerning ∞[12].

[11] recall from definition 14 that $\mathbb{R}$ is defined as the set $\{ x \mid -\infty < x < \infty\}$, where the inequalities are strict.

[12] for those who care, the extended real number system does not form a field, but the adding of $+\infty$ and $-\infty$ does form a 2-point compactification of the topological space of real numbers.

- If $x \in \mathbb{R}$ then

$$x + \infty = \infty$$

$$x - \infty = -\infty$$

$$\frac{x}{+\infty} = \frac{x}{-\infty} = 0$$

- If $x > 0$ then

$$x \cdot (+\infty) = +\infty$$

$$x \cdot (-\infty) = -\infty$$

- If $x < 0$ then

$$x \cdot (+\infty) = -\infty$$

$$x \cdot (-\infty) = +\infty$$

So point is infinity wins, always. However the following is a very important assumption:

$$\infty - \infty = \text{undefined}$$

It is not equal to zero![13] This is because it is possible for one number to be "more infinite" than another as we will soon see.

Now recall definitions 11, 12. In fact let's just state them again

Definition 25. *If there exists a 1-1 function of $\mathcal{X}$* onto *$\mathcal{Y}$, we say that $\mathcal{X}$ and $\mathcal{Y}$ can be put in a 1-1* correspondence*, or that $\mathcal{X}$ and $\mathcal{Y}$ have the same cardinal number, or briefly, that $\mathcal{X}$ and $\mathcal{Y}$ are* equivalent*, and we write $\mathcal{X} \sim \mathcal{Y}$.*

Note: there may be lots of functions from $\mathcal{X}$ onto $\mathcal{Y}$ that are not 1-1, but for this condition we only need to be able to find 1 such function!

Definition 26. *For any positive integer n, let $\mathbb{P}_n$ be the set whose elements are the integers 1, 2, ..., n; Let $\mathbb{P}$ be the set consisting of all positive integers (which is the set of natural numbers we already saw in Example 3).*

For any set $\mathcal{A}$, we say:

- $\mathcal{A}$ is *finite* if $\mathcal{A} \sim \mathbb{P}_n$ for some n.
- $\mathcal{A}$ is *infinite* if $\mathcal{A}$ is not finite.

[13] this is why $\overline{\mathbb{R}}$ cannot be a field since every element minus itself must be the "zero" element.

- $\mathcal{A}$ is *countable* if $\mathcal{A} \sim \mathbb{P}$.
- $\mathcal{A}$ is *uncountable* if $\mathcal{A}$ is neither finite nor countable.
- $\mathcal{A}$ is *at most countable* if $\mathcal{A}$ is finite or countable.

The notion of cardinality was always used to compare finite sets, and is very intuitive for such finite sets. But the notion, as described above to describe infinite sets was first used by Georg Cantor. He said that any set which could be put in one-to-one correspondence with the set of natural numbers $\mathbb{Z}$ is countably infinite. Loosely, what this means is that we can label or index all the elements of any countable set with the integers. Now it turns out that the rational numbers $\mathbb{Q}$ can be put in a one-to-one correspondence with $\mathbb{Z}$ and as such are countable. Now this seems very unintuitive since there are obviously more rationals since they also include fractions. The point is you can keep going further out along the set of integers and as such can keep indexing the rationals. Every time you need another index it is there. This is not an easy concept. In fact the natural numbers $\mathbb{Z}$ have the same cardinality as the positive integers, $\mathbb{P}$ even though $\mathbb{P} \subset \mathbb{Z}$. Confusing eh?? To make it worse, a countable union of countable sets is a new set that is also countable! For a proof of the countability of the rationals $\mathbb{Z}$ see Rudin Theorem 2.13.

Now the reals, $\mathbb{R}$, are an example of an uncountable set. Loosely, what this means is that they can not be put in a one-to-one correspondence with $\mathbb{Z}$. There are just too many of them to index by the integers. Below we list some other sets of cardinality $\mathcal{C}$.

Example 11 (Sets with cardinality $\mathcal{C}$)**.**

- the real numbers $\mathbb{R}$
- any closed or open interval in $\mathbb{R}$ (for example the unit interval $[0, 1]$)
- the irrational numbers $(\mathbb{R} \setminus \mathbb{Q})$

Yes that's right, there are more numbers within $[0, 1]$ than there are in $\mathbb{Q}$! And the sets $\mathbb{R}$ and $[0, 1]$ are of the same cardinality even though obviously $[0, 1]$ is a tiny subset of $\mathbb{R}$.

Finally, to really confuse you, $\mathbb{Q}$ is *dense* in $\mathbb{R}$ which means if $x \in \mathbb{R}, y \in \mathbb{R}$, and $x < y$, then there exists a $z \in \mathbb{Q}$ such that $x < z < y$. So what this says is that between any two real numbers there is a rational one.

But then mustn't the cardinality of $\mathbb{R}$ be at most twice that of $\mathbb{Q}$? And from the statement above, twice a countable set would definitely be countable! What's wrong with this?

Chapter 2

Sequences

2.1 Introduction

2.2 Sequences

Definition 27. *By a* sequence*, we mean a function f from $\mathbb{P}$ to $\mathbb{R}$.*

$$f : \mathbb{P} \longmapsto \mathbb{R}$$

If $f(n) = a_n$, for $n \in \mathbb{P}$, it is customary to denote the sequence f by the symbol $\{ a_n\}_{n=1}^{\infty}$, or sometimes by $\{a_1, a_2, a_3, \dots\}$. The values of f, that is, the elements a_n, are called the terms *of the sequence.*

Note that the terms $a_1, a_2, a_3, \dots$ of a sequence need not be distinct.

Example 12. *Here are some sequences:*

$$\begin{array}{ll} a_n = n & \{ a_n\}_{n=1}^{\infty} = \{1, 2, 3, \dots\} \\ a_{n+1} = a_n + 2, \forall\, n \geq 1 & \{ a_n\}_{n=1}^{\infty} = \{1, 3, 5, \dots\} \\ a_n = (-1)^n, n \geq 1 & \{ a_n\}_{n=1}^{\infty} = \{-1, 1, -1, 1, \dots\} \end{array}$$

Definition 28. *[Convergent Sequence] A sequence $\{ a_n\}$ is said to* converge *in $\mathbb{R}$, if there is a point $a \in \mathbb{R}$ with the following property: For every ϵ[1] > 0, $\exists\, N \in \mathbb{P}$ such that $n \geq N$ implies*

$$|a_n - a| < \epsilon$$

[1]This is our first encounter with ϵ (epsilon). It is just the standard symbol in mathematics to denote a really tiny, tiny, tiny, but still positive amount.

In this case we also say that $\{ a_n\}$ converges to a, or that a is the limit of $\{ a_n\}$, and we write

$$a_n \longrightarrow a$$

or

$$\lim_{n\longrightarrow\infty} a_n = a$$

If $\{ a_n\}$ does not converge, it is said to *diverge*.

Note. that our definition of a convergent sequence depends not only on $\{ a_n\}$ but also on the the space in which it is defined, so it is meaningless to say simply that a sequence "converges", you **must** also specify the space! For example the sequence $\left\{\frac{1}{n}\right\}$ converges in $\mathbb{R}$ (to 0), but fails to converge in the set of all positive real numbers $\mathbb{R}_{++}$.

Theorem 7. *Let* $\{a_n\}$ *be a sequence in* $\mathbb{R}$.
If $b \in \mathbb{R}$, $b' \in \mathbb{R}$, *and if* $\{a_n\}$ *converges to* b *and to* b', *then*

$$b = b'$$

Proof. Let's take $b \neq b'$ and suppose without loss of generality (w.l.o.g.) that $b > b'$

Let $\epsilon = \frac{1}{2}(b - b')$.

i.e. take ϵ so that if you are within a distance of ϵ from b, then you are not within a distance ϵ from b'.

FIGURE GOES HERE

If b' is a limit of $\{a_n\}$, then $\exists\, N'$ such that $\forall\, n \geq N'$

$$|b' - a_n| < \epsilon$$

But this means that there is $\underline{no}\ N$ such that $\forall\, n \geq N$

$$|a_n - b| < \epsilon$$

Thus, we have that b is not a limit of $\{a_n\}$.

□

Proof. Let $\epsilon > 0$ be given. There exist integers N and N' such that

$$n \geq N \Longrightarrow d(a_n, b) < \frac{\epsilon}{2}$$

$$n \geq N' \Longrightarrow d(a_n, b') < \frac{\epsilon}{2}$$

Hence, if $n \geq \max(N, N')$, we have

$$d(b, b') \leq d(b, a_n) + d(b', a_n) < \epsilon$$

Since ϵ was arbitrary, we conclude that $d(b, b') = 0$.[2] □

Definition 29. *Formally, a* subsequence *is a composition mapping*[3]*. A subsequence of the function f from $\mathbb{P}$ to $\mathbb{R}$.*

$$f : \mathbb{P} \longrightarrow \mathbb{R}$$

is given by a function

$$f \circ g : \mathbb{P} \longrightarrow \mathbb{R}$$

where

$$g : \mathbb{P} \longrightarrow \mathbb{P}$$

and g is assumed to be strictly increasing[4]*.*

FIGURE GOES HERE

Alternatively, given a sequence $\{a_n\}$, consider a sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3 < \cdots$. Then the sequence $\{a_{n_i}\}$ is called a *subsequence* of $\{a_n\}$. If $\{a_{n_i}\}$ converges, its limit is called a *subsequential limit* of $\{a_n\}$.

It can be shown that $\{a_n\}$ converges to b if and only if every subsequence of $\{a_n\}$ converges to b.

[2] d here is called a *metric*. You can think of it as being function which takes in any two points in a set, and then spits out the distance between the two of them. It will be more formally defined in Definition 72.

[3] see back to Definition 10 for what exactly is a composition mapping

[4] we will see exactly what it means for a function to be strictly increasing soon, but for now just think of it as always going up

Definition 30. *A sequence of real numbers $\{a_n\}$ of real numbers is said to be*

1. monotonically increasing *if*

$$a_n \leq a_{n+1} \qquad (n = 1, 2, 3, \dots)$$

2. monotonically decreasing *if*

$$a_n \leq a_{n+1} \qquad (n = 1, 2, 3, \dots)$$

Note we say strictly increasing/decreasing if the above weak inequalities are replaced with strict inequalities.

Definition 31. *A function f is said to be*

1. monotonically increasing *on* (a, b) *if*

$$a < x < y < b \quad \Longrightarrow \quad f(x) \leq f(y)$$

2. monotonically decreasing *on* (a, b) *if*

$$a < x < y < b \quad \Longrightarrow \quad f(x) \geq f(y)$$

Note we say a function f is strictly increasing/decreasing if the above weak inequalities are replaced with strict inequalities.

The intuition for this is straightforward. A function is monotonically increasing if it never takes a "step" backwards. Each time you go further on in the sequence you are at a value at least as high as the one before. If the function is strictly increasing, then the function not only never takes a "step" backwards, it always takes a step forwards. In other words, each time you go further on in the sequence, you always take a step forwards. Moving to a value the same as the one you are at is no longer allowed.

Theorem 8. *Suppose*

$$\lim_{n \longrightarrow \infty} a_n = b$$

then all subsequences of $\{a_n\}$ converge to b as well.

Note Be careful! What this is saying is that if a sequence has a limit - then so does every subsequence of it. But if a subsequence has a limit - this does not imply that the sequence necessarily has a limit[5].

Example 13. *Consider the sequence*

$$\{a_n\} = \left\{ 1, 1, 1, \frac{1}{2}, 1, \frac{1}{3}, 1, \frac{1}{4}, 1, \frac{1}{5}, 1, \dots \right\}$$

So we can tell from this that the subsequence obtained by taking the even elements of the above sequence has a limit of 0[6]*. But this is clearly not the limit of the original sequence since the number* 1 *keeps appearing! Moreover, the subsequence obtained by taking the odd elements of the original sequence clearly has a limit of 1 (since we are only taking the element 1 over and over again)!! (since recall that the terms of a sequence need not be distinct, and a subsequence is a sequence in its own right).*

Theorem 9. *Suppose* $\{a_n\}$ *and* $\{b_n\}$ *are sequences, and*

$$\lim_{n \longrightarrow \infty} a_n = a$$

$$\lim_{n \longrightarrow \infty} b_n = b$$

Then

1. $$\lim_{n \longrightarrow \infty} (a_n + b_n) = a + b$$

2. $$\lim_{n \longrightarrow \infty} ca_n = ca, \ c \in \mathbb{R}$$

3. $$\lim_{n \longrightarrow \infty} a_n b_n = ab$$

4. $$\lim_{n \longrightarrow \infty} (a_n)^k = a^k$$

[5] As a test of whether or not you've grasped necessary versus sufficiency, which is which in this case?

[6] thought I didn't define the space on which the sequence is defined so technically this is not correct!

5.

$$\lim_{n\longrightarrow\infty}\frac{1}{a_n}=\frac{1}{a}$$

provided $a_n\neq 0(n=1,2,3,\dots)$*, and* $a\neq 0$

6.

$$\lim_{n\longrightarrow\infty}\frac{a_n}{b_n}=\frac{a}{b}$$

provided $b_n\neq 0(n=1,2,3,\dots)$*, and* $b\neq 0$

Proof. 1. [7] Given $\epsilon>0$, $\exists$ integers N_1, N_2 such that

$$n\geq N_1\Longrightarrow |a_n-a|<\frac{\epsilon}{2},$$

$$n\geq N_2\Longrightarrow |b_n-b|<\frac{\epsilon}{2}.$$

If we take $N=\max(N_1,N_2)$, then $n\geq N\Longrightarrow$

$$|(a_n+b_n)-(a+b)|\leq |a_n-a|+|b_n-b|<\epsilon$$

2. This proof is very easy - just factor out the c and then put it in again at the end

3. This uses a trick

$$a_nb_n-ab=(a_n-a)(b_n-b)+a(b_n-b)+b(a_n-a) \tag{2.1}$$

Given $\epsilon>0$, there are integers N_1,N_2 such that

$$n\geq N_1\Longrightarrow |a_n-a|<\sqrt{\epsilon},$$

$$n\geq N_2\Longrightarrow |b_n-b|<\sqrt{\epsilon}.$$

If we take $N=\max(N_1,N_2)$, $n\geq N\Longrightarrow$

$$|(a_n-a)(b_n-b)|\leq\epsilon,$$

so that

$$\lim_{n\longrightarrow\infty}(a_n-a)(b_n-b)=0.$$

We now apply results (1) and (2) to Equation 2.1, and conclude that

$$\lim_{n\longrightarrow\infty}(a_nb_n-ab)=0$$

[7] this proof really uses the triangle inequality which will not show up til Definition 72, but which intuitively just states that the shortest distance between any two points is a straight line between them.

4. This is left as an exercise to the reader

5. Choosing m such that $|a_n - a| < \frac{1}{2}|a|$ if $n \geq m$, we see that

$$|a_n| > \frac{1}{2}|a|, \qquad (n \geq m)$$

Given $\epsilon > 0$, $\exists$ an integer $N > m$ such that $n \geq N \Longrightarrow$

$$|a_n - a| < \frac{1}{2}|a|^2 \epsilon.$$

Hence for $n \geq N$,

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a_n - a}{a_n a}\right| < \frac{2}{|a|^2}|a_n - a| < \epsilon$$

6. This is really just combining previous results

□

Definition 32. *We say the sequence $\{a_n\}$ is bounded above if $\exists\ \overline{m} \in \mathbb{R}$ such that*

$$a_n \leq \overline{m}, \quad \forall\, n \in \mathbb{P}$$

We say the sequence $\{a_n\}$ is bounded below if $\exists\ \underline{m} \in \mathbb{R}$ such that

$$a_n \geq \underline{m}, \quad \forall\, n \in \mathbb{P}$$

Note: that a bounded sequence does not necessarily converge, for example the sequence $\{0, 1, 0, 1, 0, 1, \dots\}$ just oscillates back forth between $\{0\}$ and $\{1\}$ forever.

Theorem 10. *If $\{a_n\}$ converges then $\{a_n\}$ is bounded.*

Proof. Suppose $a_n \longrightarrow a$. There is an integer N such that $n > N$ implies $|a_n - a| < 1$. Put

$$r = max\{1, |a_1 - a|, \dots, |a_N - a|\}$$

Then $|a_n - a| \leq r$ for $n = 1, 2, 3, \dots$. □

Another way to think of this is that after a certain n (i.e. $n = 1000$) we don't have to worry since it must converge - so it's definitely bounded. So then what happens before this? Well before this, it is a finite sequence - and a finite sequence must be bounded since it must have a max and a min. This last result can be proved by induction:

Is a set of 1 element bounded?	Obviously!
Is a set of 2 elements bounded?	Obviously!
$\vdots$	$\vdots$
Is a set of n elements bounded?	Obviously!

Theorem 11. *Let* $\{a_n\}$, $\{b_n\}$, *and* $\{c_n\}$ *be sequences such that*

$$a_n \leq b_n \leq c_n, \qquad \forall\, n \in \mathbb{R}$$

and we have that both

$$a_n \longrightarrow a, \qquad c_n \longrightarrow a$$
$$\Longrightarrow b_n \longrightarrow a$$

Proof. The proof is left as an exercise to the reader[8]. □

Theorem 12. *A monotone sequence is convergent if and only if it is bounded*[9].

Proof. This is our first full "if and only if" proof. Recall from section **??** that to prove this we have to prove 2 things.

- ($\overset{only\,if}{\Longrightarrow}$) We have a monotone sequence that is convergent, and must prove that this implies that it's bounded.
- ($\overset{if}{\Longleftarrow}$) We have a monotone sequence that is bounded and must prove that this implies convergence.

To reiterate *if and only if* (iff) statements must be proved going both ways.

[8] Think about this. If the first sequence converges, and the third sequence converges, and we are told that every element of the second sequence is "sandwiched" between the corresponding element of the first and third sequence then this sequence should also converge to the same point. Makes sense right???!!! So you'll find that often times you know exactly what you have to prove, but it can still be difficult to actually do it.

[9] compare this to Theorem 10. What's different? Why has "if" gone to "iff"?

1. ($\overset{only\,if}{\Longrightarrow}$) We already showed that any convergent sequence is bounded in Theorem 10. So obviously a monotone sequence is bounded![10]

2. ($\overset{if}{\Longleftarrow}$) Suppose $\{\, a_n\}$ is bounded. Thus by Lemma 1, we know that the set $\mathcal{X} \equiv \{\, a_n \mid n \in \mathbb{P}\}$ has a sup which we will denote $\overline{a}$. We must show that $\overline{a}$ is the limit of $\{\, a_n\}$. We need to show that $\forall\, \epsilon > 0, \exists\, N \in \mathbb{P}$ such that

$$|\overline{a} - a_n| < \epsilon, \quad \forall\, n \geq N$$

Take any $\epsilon > 0$. Since $\overline{a} - \epsilon$ is not an upper bound of $\mathcal{X}$, we know from Lemma 2 that $\exists\, N$ such that $a_N > \overline{a} - \epsilon$. Since $\{a_n\}$ is increasing, for every $n \geq N$ we have

$$\overline{a} - \epsilon < a_N \leq a_n \leq \overline{a}$$

since we also have that $a_n \leq \overline{a}$, $\forall\, n$.
Thus,

$$|a_n - \overline{a}| < \epsilon, \qquad \forall\, n \geq N$$

□

Theorem 13. *Every real sequence has a monotone subsequence.*

Proof. Take a sequence $\{x_n\}$. We must find a subsequence $\{x_{n_k}\}$ that is either increasing or decreasing.

So obviously there can be only 2 cases:

1. the set $\{x_n \mid n \geq N\}$ has a maximum for all N.
   And if this is true we should be able to form a decreasing sequence

2. the set $\{x_n \mid n \geq N\}$ doesn't have a maximum for some $N \in \mathbb{P}$

Note: that this theorem is not saying that there exists a unique monotone subsequence. Indeed, if there is one monotone subsequence, there will be infinitely many (generate a new one by deleting the first term of the original one).

[10] Since the class of monotonic sequences is a subset of the class of all sequences, then obviously this applies to monotone sequences.

1. Form a sequence of $\mathbb{P}$ by:

$$x_{n_1} = \max_{n>1} x_n$$

$$x_{n_2} = \max_{n>n_1} x_n$$

$$\vdots$$

$$x_{n_{k+1}} = \max_{n>n_k} x_n$$

So we see by construction that this sequence $\{x_{n_k}\}$ is decreasing. (Convince yourself of this!) Also note that the set $\{x_n \mid n > N\}$ is decreasing in size (number of elements) as N increases.

2. For each $n > N$, we can find $m > n$ such that $x_m > x_n$. If we can't do such a thing then the set has a maximum so we are in case 1 and there is no reason to proceed.

Define,

$$\begin{array}{lcl} x_{n_1} & \equiv & x_N, \\ x_{n_2} & \equiv & \text{the first term following } x_{n_1} \text{ for which } x_{n_2} > x_{n_1} \\ \vdots & & \vdots \\ x_{n_{k+1}} & \equiv & \text{the first term following } x_{n_k} \text{ for which } x_{n_{k+1}} > x_{n_k}. \end{array}$$

And we can see that $\{x_{n_k}\}$ is increasing by construction.

□

Theorem 14 (Bolzano-Weierstrauss)**.** *Every bounded, real sequence has a convergent subsequence.*

Proof. This is got by applying the last 2 theorems.

1. Every real sequence has a monotone subsequence

2. We know from applying Theorem 12 using the ($\overset{if}{\Longleftarrow}$) direction case, that every bounded monotone sequence is bounded.

□

Definition 33. *A sequence* $\{a_n\}$ *is said to be a* Cauchy Sequence *if for every* $\epsilon > 0$ *there is an integer N such that*

$$|a_n - a_m| < \epsilon$$

if $n \geq N$ *and* $m \geq N$.

Refer back to the definition of a convergent sequence (Definition 28) to see the difference between the definition of convergence and the definition of a Cauchy sequence. The difference is that the limit is explicitly involved in the definition of convergence but not in that of a Cauchy sequence[11].

Theorem 15. *A real sequence* $\{a_n\}$ *is convergent if and only if it is a Cauchy sequence*[12].

Definition 34. $\mathcal{X} \subset \mathbb{R}$ *is a* compact space *if every sequence in* $\mathcal{X}$ *has a convergent subsequence in* $\mathcal{X}$.

Example 14. *Consider the sequence defined by* $a_n = \frac{1}{n}$

$$\implies \quad \{\, a_n\}_{n=1}^{\infty} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots \right\}$$

On $\mathbb{R}$ *it is obvious that this converges to* 0. *But on the set* $\mathcal{X} = (0, 1)$ *it does not converge (as will be explained below). Here we say that the set* $\mathcal{X}$ *is not compact.*

Definition 35. *Take any set* $\mathcal{X} \subset \mathbb{R}$.
Call $a \in \mathbb{R}$ *a* limit point *of* $\mathcal{X}$ *if there is a sequence* $\{a_n\}$, *(where* $\{\, a_n \mid n \in \mathbb{P}\} \subset \mathcal{X}$*) such that*

$$a_n \longrightarrow a.$$

Note: that $a \notin \mathcal{X}$ is possible, so a set need not include it's limit points.

Definition 36. $\mathcal{X}$ *is called a* closed *set if it contains all its limit points.*

(That is, every sequence that converges in $\mathcal{X}$ does so to a point in $\mathcal{X}$.)

[11] note above that I did not specify a space in which the sequence is Cauchy. This is sloppy.

[12] This result is usually called the Cauchy criterion. Also note that $\{a_n\}$ above was stated to be real, so the space was stated this time.

Definition 37. *$\mathcal{X}$ is called* open *if for every $x \in \mathcal{X}$, $\exists\, \epsilon > 0$ such that*

$$(x - \epsilon, x + \epsilon) \subset \mathcal{X}$$

A *note* on notation:
As you may have seen before that a closed set from the point x to the point y is denoted with *square* brackets

$$[x, y]$$

while an open set from x to y is denoted with *round* brackets

$$(x, y)$$

So the only difference is that the open set does not contain the points x and y but does include every point right up next to them!

Recall the set $\mathcal{X}$ from example 14. We can see that this set $\mathcal{X}$ has a limit point (namely the point $a = 0$) but that no point of $\mathcal{X}$ is a limit point of $\mathcal{X}$. It's important to understand the difference between having a limit point and containing one!

Theorem 16. *$\mathcal{X} \subset \mathbb{R}$ is compact if and only if it is closed and bounded.*

Chapter 3

Functions and Limits of Functions

Take any function f,

$$f : \mathcal{X} \longrightarrow \mathbb{R}$$

for $\mathcal{X} = (c, d)$ (so note that it is an open set) and take $a \in \mathcal{X}$.

Definition 38. *y is called the* limit of f from the right at a *(right hand limit) if, for every $\epsilon > 0$, there is a δ such that*

$$0 < x - a < \delta \quad \Longrightarrow \quad |f(x) - y| < \epsilon$$

this is also denoted as

$$y = \lim_{x \to a^+} f(x)$$

Note:We are only considering $x > a$, so x is to the right of a. Also $x \neq a$.

Definition 39. *y is called the* limit of f from the left at a *(left hand limit) if, for every $\epsilon > 0$, there is a $\delta > 0$ such that*

$$0 < a - x < \delta \quad \Longrightarrow \quad |f(x) - y| < \epsilon$$

this is also denoted as

$$y = \lim_{x \to a^-} f(x)$$

Note: Now it is $a - x$ since we are coming from the left, so this must be positive.

FIGURE GOES HERE

Definition 40. *y is called the* limit of f at a *if*

$$\begin{aligned} y &= \lim_{x \to a^-} f(x) \\ &= \lim_{x \to a^+} f(x) \end{aligned}$$

and we write

$$y = \lim_{x \to a} f(x)$$

y is the limit of f at a if for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$0 < |x - a| < \delta \qquad \Longrightarrow \qquad |f(x) - y| < \epsilon$$

Note: The definition of the limit of a function at the point a does not require the function to be defined at a.

The use of ϵ and δ in these definitions is a standard part of mathematical analysis, but takes practice to appreciate. The formal definition of limit at a point is a way to capture the idea of the behavior of a function "near" the point. If $\lim_{x \to a} f(x) = L$ these means, roughly, that values of $f(x)$ are near L when x is near to a.

In order to use the definition, you must answer infinitely many questions: for each ϵ you must say how close x must be to a to guarantee that $f(x)$ is within ϵ of L. The way that you do this is to define δ as a function of ϵ and then attempt to confirm the definition.

Example 15. *Consider the function*

$$f(x) = \begin{cases} 0, & \text{if } x \neq 0, \\ 1, & \text{if } x = 0. \end{cases}$$

FIGURE GOES HERE

It is clear from the picture that

$$\lim_{x \to 0^-} f(x) \quad = \quad 0 \quad = \quad \lim_{x \to 0^+} f(x)$$

$$\Rightarrow \quad \lim_{x \to 0} f(x) \quad = \quad 0$$

But the value of the function at the point 0 is not equal to the limit. The example shows that it is possible for $\lim_{x \to a} f(x)$ to exist but for it to be different from $f(a)$. Since you can define the limit without knowing the value of $f(a)$, this observation is mathematically trivial. It highlights a case that we wish to avoid, because we want a function's value to be approximated by nearby values of the function.

Theorem 17. *Limits are unique. That is, if* $\lim_{x \to a} f(x) = L$ *and* $\lim_{x \to a} f(x) = L'$*, then* $L = L'$.

Proof. Assume that $L \neq L'$ and argue to a contradiction. Let $\epsilon = |L - L'|/2$. Given this ϵ let $\delta^* > 0$ have the property that $|f(x) - L|$ and $|f(x) - L'|$ are less than ϵ when $0 < |x - a| < \delta^*$. (This is possible by the definition of limits.) Since

$$|f(x) - L| + |f(x) - L'| \geq |L - L'| \tag{3.1}$$

it follows that $\epsilon \geq |L - L'|$, which is not possible. □

The theorem simply states that the function cannot be close to two different things at the same time. By the way, the critical inequality (inequality (3.1)) is called the triangle inequality. In one dimension you can check that the inequality is true by using the definition of absolute value, by drawing a simple picture, or by thinking about the meaning of the three different terms. Roughly it says that the distance between two numbers is no larger that the distance between the first number and a third number plus the distance between the second number and the third number.

If a limit exists, then both limit from the left and limit from the right exist (and they are equal) and, conversely, if both limit from right and limit from left exist, then the limit exists. These statements require proofs, but the proofs are simple.

You do not want to use $\epsilon - \delta$ proofs whenever you need to find a limit. To avoid tedium, you need to collect a few "obvious" limits (for example, if the function f is constant, then $\lim_{x \to a} f(x)$ exists for all a and is equal to the constant value of f), and then some basic results that permit you to compute limits.

Theorem 18. *If* f *and* g *are functions defined on a set* S*,* $a \in (\alpha, \beta) \subset S$ *and* $lim_{x \to a} f(x) = M$ *and* $lim_{x \to a} g(x) = N$*, then*

1. $lim_{x \to a} (f + g)(x) = M + N$
2. $lim_{x \to a} (fg)(x) = MN$

3. $lim_{x\to a}\left(\frac{f}{g}\right)(x) = M/N$ *provided* $N \neq 0$.

Proof. For the first part, given $\epsilon > 0$, let $\delta_1 > 0$ be such that if $0 < |x-a| < \delta_1$ then $|f(x) - M| < \epsilon/2$ and $\delta_2 > 0$ be such that if $0 < |x-a| < \delta_2$ then $|g(x) - N| < \epsilon/2$. This is possible by the definition of limit. If $\delta = \min\{\delta_1, \delta_2\}$, then $0 < |x-a| < \delta$ implies

$$|f(x) + g(x) - M - N| \leq |f(x) - M| + |g(x) - N| < \epsilon/2 + \epsilon/2 < \epsilon.$$

The first inequality follows from the triangle inequality while the second uses the definition of δ. This proves the first part of the theorem.

For the second part, you can use the same type of argument, setting δ_1 so that if $0 < |x-a| < \delta_1$ then $|f(x) - M| < \sqrt{\epsilon}$ and so on.

For the third part, note that when $g(x)$ and N are not equal to zero

$$\frac{f(x)}{g(x)} - \frac{M}{N} = \frac{g(x)\left(f(x) - M\right) + f(x)\left(N - g(x)\right)}{g(x)N}.$$

So given $\epsilon > 0$, find δ so small that if $0 < |x-a|\,\delta$, then $\frac{|f(x)|}{|g(x)N|} < \frac{2|M|}{N^2}$, $|f(x) - M| < N\epsilon/2$ and $|g(x) - N| < \frac{N^2}{4|M|}\epsilon/2$. This is possible provided that $N \neq 0$. (In this part I constructed an upper bound for $\left|\frac{f(x)}{g(x)N}\right|$ using the fact that $f(x)$ is near M and $g(x)$ is near N when x is a. □

Using this theorem you can generate the limits of many functions by combining more basic information.

Definition 41. *We say* $f : \mathcal{X} \longmapsto \mathbb{R}$ *is* continuous at a *if*

$$f(a) = \lim_{x\to a} f(x)$$

Note: f is said to be continuous at the point a. If f is continuous at every point of $\mathcal{X}$, then f is said to be *continuous* on $\mathcal{X}$.
Note: the limit exists and the limit of the function is equal to the function of the limit. It should also be noted that the function has to be defined at a. *Note*: In order for a function to be continuous at a, it must be defined "in a neighborhood" of a – that is, the function must be defined on an interval (α, β) with $a \in (\alpha, \beta)$. We extend the definition to take into account "boundary points" in a natural way: We say that f defined on $[a, b]$ is continuous at a (resp. b) if $\lim_{x\to a^+} f(x) = f(a)$

(resp. $\lim_{x \to b^-} f(x) = f(b)$). *Note*: Informally one could remember the definition of continuity as:

$$f(\lim_{x \to a} x) = \lim_{x \to a} f(x).$$

Continuity at a point requires two things: a limit exists and the limit is equal to the right thing. It is easy to think of examples in which the limit exists, but is not equal to the value of the function. The limit of a function can fail to exist for two different reasons. It could be that the left and right hand limits exist, but are different. Alternatively, either left or right-hand limit may fail to exist. This could happen either because the function is growing to infinity (for example, $f(x) = 1/x$ for $x > 0$ is continuous at any point on $a > 0$, but cannot be continuous at 0 no matter how $f(0)$ is defined. Alternatively, the function can oscillate wildly (a standard example is $\sin(\frac{1}{x})$ near $x = 0$.

Lemma 3. $f : \mathcal{X} \longmapsto \mathbb{R}$ *is* continuous at a *if and only if for every* $\epsilon > 0$*, there is a* $\delta > 0$ *such that*

$$0 < |a - x| < \delta \qquad \Longrightarrow \qquad |f(x) - f(a)| < \epsilon$$

We say "f is continuous" if f is continuous at every point $a \in \mathcal{X}$.

Note: a very informal way of describing a continuous function is that you can draw the function without lifting your hand off the page, by sweeping the pen in one fluid, *continuous* motion!

As in the case of limits, we can identify whether combinations of functions are continuous without using the definition directly. The sums, products, and ratios of continuous functions are continuous (in the last case the denominator of the ratio must be non-negative in the limit).

Theorem 19. *For*

$$g : \mathcal{X} \longrightarrow \mathcal{Y}$$

and

$$f : \mathcal{Y} \longrightarrow \mathbb{R}$$

where both $\mathcal{X}$ *and* $\mathcal{Y}$ *are open intervals of* $\mathbb{R}$*, if* g *is continuous at* $a \in \mathcal{X}$ *and if* f *is continuous at* $g(a) \in \mathcal{Y}$*, then*

$$f \circ g : \mathcal{X} \longrightarrow \mathbb{R}$$

is continuous at a.[1]

[1] again note that this is continuity at a point.

Proof. Let $\epsilon > 0$ be given. Since f is continuous at $g(a)$, there exists $\gamma > 0$ such that

$$|f(y) - f(g(a))| < \epsilon$$

if

$$|y - g(a)| < \gamma$$

and $y \in g(\mathcal{X})$.
Since g is continuous at a, there exist $\delta > 0$ such that

$$|g(x) - g(a)| < \gamma$$

if

$$|x - a| < \delta$$

and $x \in \mathcal{X}$.
It follows that

$$|f(g(x)) - f(g(a))| < \epsilon$$

if

$$|x - a| < \delta$$

and $x \in \mathcal{X}$.

Thus $f \circ g$ is continuous at a. □

It is easy to show that constant functions and linear functions are continuous. By the combining properties, this allows you to conclude polynomials and ratios of polynomials are continuous.

Theorem 20. *The function*

$$f : (a, b) \longrightarrow \mathbb{R}$$

is continuous at a point x if and only if, for <u>all</u> sequences $\{x_n\} \subset (a, b)$ with $x_n \longrightarrow x$, it is the case that

$$f(x_n) \longrightarrow f(x)$$

Proof. 1. ($\overset{only\,if}{\Longrightarrow}$) Let's use contradiction method of $\neg B \implies \neg A$ (to prove $A \implies B$)
Suppose $\exists$ a sequence $\{x_n\} \subset (a, b)$ with $x_n \to x$ but $f(x_n) \not\to f(x)$. This means that $\exists\, \epsilon > 0$ such that $\forall\, \delta > 0, \exists\, n \in \mathbb{P}$ (i.e. $\exists\ \{x_n\}$) such that

$$|x_n - x| < \delta$$

but

$$|f(x_n) - f(x)| > \epsilon$$

But this clearly violates the continuity of f.

2. ($\overset{if}{\Longleftarrow}$) Suppose f is not continuous at x. Then $\exists\, \epsilon > 0$ such that $\forall\, \delta > 0, \exists\, x' \in (a, b)$ satisfying $|x' - x| < \delta$ and $|f(x') - f(x)| > \epsilon$.
We can construct $\{x_n\}$ by letting x_n be this x' for $\delta = \frac{1}{n}$.
Then we get

$$x_n \to x \quad \text{but} \quad f(x_n) \not\to f(x)$$

□

The next result states has two important consequences.

Theorem 21. *The continuous image of a closed and bounded interval is a closed and bounded interval. That is, if*

$$f : [a, b] \longrightarrow \mathbb{R}$$

is continuous, then there exists $d \geq c$ such that $f([a, b]) = [c, d]$.

The assumptions in the theorem are important. If f is not continuous, then there generally nothing that can be said about the image. If the domain is an open interval, the image could be a closed interval (it is a point of f is constant) or it could be unbounded even if the interval is finite (for example if $f(x) = 1/x$ on $(0, 1)$).

The first consequence of the result is the existence of a maximum (and minimum).

Definition 42. *We say that $x^* \in \mathcal{X}$ maximizes the function f on $\mathcal{X}$ if*

$$f(x^*) \geq f(x),$$

for every $x \in \mathcal{X}$.

Definition 43. *We say that $x^* \in \mathcal{X}$ minimizes the function f on $\mathcal{X}$ if*

$$f(x^*) \leq f(x),$$

for every $x \in \mathcal{X}$.

We write

$$\max_{x \in \mathcal{X}} f(x) = \max f(\mathcal{X})$$

and

$$\max f(\mathcal{X}) = y^*$$

is the max value.

Since the image of f is a closed, bounded interval, f attains both its maximum (the maximum value is d) and its minimum. This means that if a function is continuous and it is defined on a "nice" domain, then it has a maximum. This result is much more general than the result above. It applies to all real-valued continuous functions (defined on arbitrary sets, not just the real numbers) provided that the domain is "compact." Bounded closed intervals contained in the real line are examples of compact sets. More generally, any set that is "closed" (contains its boundary points) and bounded is compact.

Theorem 22. *If the function*

$$f : \mathcal{X} \longrightarrow \mathbb{R}$$

is continuous and $\mathcal{X}$ is compact, then $f(\mathcal{X})$ is compact.

Proof. Take any $\{y_n\} \subset f(\mathcal{X})$, we must show that $\exists$ a subsequence that converges to a point in $f(\mathcal{X})$. Construct the sequence $\{x_n\} \subset \mathcal{X}$ such that $y_n = f(x_n) \, \forall \, n$. Since $\mathcal{X}$ is compact, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$; that is, $x_{n_k} \longrightarrow x \in \mathcal{X}$. By continuity of f and our previous theorem, we have that $y_{n_k} \longrightarrow f(x)$ as well. □

FIGURES GO HERE

Theorem 23. *If the function*

$$f : \mathcal{X} \longrightarrow \mathbb{R}$$

is continuous and $\mathcal{X}$ is compact, then $\max f(\mathcal{X})$ *exists.*

Proof. Note that $f(\mathcal{X})$ is bounded because $f(\mathcal{X})$ is compact due to $\mathcal{X}$ being compact and using Theorem 22.

Thus using this result and Lemma 1, we know that $\sup f(\mathcal{X})$ exists. So we can find a sequence $\{y_n\} \subset f(\mathcal{X})$ such that

$$y_n \longrightarrow \sup f(\mathcal{X}).$$

Since $f(\mathcal{X})$ is compact, it is thus closed by Theorem 16, so it therefore contains its limit points, and so

$$\lim_{n\to\infty} y_n = \sup f(\mathcal{X}) \in f(\mathcal{X})$$

And now by Theorem 2 when the $\max$ exists it equals the $\sup$.

So finally we have that $\max f(\mathcal{X})$ exists and

$$\begin{aligned}\lim_{n\to\infty} y_n &= \sup f(\mathcal{X}) \\ &= \max f(\mathcal{X})\end{aligned}$$

□

FIGURE GOES HERE

The second consequence of Theorem 21 is the Intermediate Value Theorem.

Theorem 24. *If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, and z is between $f(a)$ and $f(b)$, then there exists $z \in [a, b]$ such that $f(z) = [a, b]$*

In the statement of the theorem, "between" means either $z \in [f(a), f(b)]$ (when $f(a) \leq f(b)$ or $z \in [f(b), f(a)]$ if $f(b) < f(a)$.

Proof. From Theorem 21 it follows that the image of f is an interval. Hence if two points are in the image, then all points between these two points are in the image. □

Theorem 24 is a method of solving showing that equations have solutions. It is common to conclude that there must be a c for which the continuous function f is zero ($f(c) = 0$) from the observation that sometimes f is positive and sometimes f is negative. The most important existence theorems in micro (the existence of a market clearing price system, for example) follow from (harder to prove) versions of this result.

Economists often prove existence results using fixed-point theorems. The easiest fixed point theorem is a consequence of the Intermediate Value Theorem.

The point $x^* \in S$ is called a *fixed point* of the function $f : S \longrightarrow S$ if $f(x^*) = x^*$.

Theorem 25. *Let $S = [a, b]$ be a closed bounded interval and $f : S \longrightarrow S$ a continuous function. There exists $x^* \in S$ such that $f(x^*) = x^*$.*

Proof. Consider the function $h(x) = f(x) - x$. Since $f(a) \geq a$, $h(a) \geq 0$. Since $f(b) \leq b$, $h(b) \leq 0$. Since h is a continuous function on a closed bounded interval, there must be an x^* such that $h(x^*) = 0$. It is clear that for this value, $f(x^*) = x^*$. □

This theorem generalizes: One can show that if S satisfies some technical properties (convexity and compactness) then a continuous function from S to itself must have a fixed point. This result does not require that S is one-dimensional and it is instrumental in many existence proofs. The more general fixed-point theorems are easy to state and understand, but are hard to prove.

It is not hard to give examples of functions that fail to have fixed points (if the domain is not right). A simple one is: $f(x) = x + 1$.

Chapter 4

Differentiation

Calculus works because of two insights. The first is that linear functions are easy to understand. The second insight is that although not all interesting functions are linear, there is a large set of functions that can be approximated by a linear function. The derivative is the best linear approximation to a function. You obtain a lot of analytic power by taking a general function and studying it by learning about linear approximations to the function.

How do you approximate a function? The first step is to treat the approximation as local. Given a point in the domain of the function, a, the idea is to come up with a function that is easy to deal with and is close to the given function when x is close to a. A possible attempt to do this is with what I will call a *zero-th order approximation*.

Definition 44. *The* zero-th order approximation *of the function f at a point a in the domain of f is the function $A_0(x) \equiv f(a)$.*

The symbol $\equiv$ indicates an identity. I could also have written $A_0(x) = a$ for all x.

One way to approximate a function is with the constant function that is equal to the value of the function at a point. $A_0(x)$ is certainly a tractable function and if f is continuous at a it is the only function that satisfies $lim_{x \to a} (f(x) - A_0(x)) = 0$. That is the good news. The bad news is that A_0 tells you almost nothing about the behavior of f.

The next step is to try to replace A_0 with the best linear approximation to f. In order to do this, imagine a line that goes intersects the graph of f at the points $(x, f(x))$ and the point $(x + \delta, f(x + \delta)$. This line has slope given by:

$$\frac{f(x+\delta) - f(x)}{(x+\delta) - x} = \frac{f(x+\delta) - f(x)}{\delta}$$

FIGURE GOES HERE

When f is linear, the line with this slope is f (just like when f is constant, $A_0 \equiv f$). Otherwise, it is not. If we are interesting only in the local behavior of f, then it makes sense to consider slopes defined when δ is small. The notion of limit tells us how to do this:

$$\lim_{\delta \to 0} \frac{f(x+\delta) - f(x)}{(x+\delta) - x} \qquad = \qquad \lim_{\delta \to 0} \frac{f(x+\delta) - f(x)}{\delta}$$

The demoninator of the expression does not make sense when $\delta = 0$, but we do not need to this value to evaluate the limit of the ratio as δ approaches zero. On the other hand, for the limit to make sense the ratio must be defined for all sufficiently small non-zero δ – that is x must be an interior point of the domain of f. If the limit exists, we say that f is differentiable at f and call the limit the derivative of f at x.

As with the two earlier definitions (limit of function and continuity), there are two "one-sided" versions of this definition.

Take

$$f : (a, b) \longrightarrow \mathbb{R}$$

and $x \in (a, b)$.

Define if they exist

$$f'_+ \equiv \lim_{y \to x^+} \frac{f(y) - f(x)}{y - x}$$

$$(\lim_{\delta \to 0^+} \frac{f(x+\delta) - f(x)}{\delta})$$

and

$$f'_- \equiv \lim_{y \to x^-} \frac{f(y) - f(x)}{y - x}$$

$$(\lim_{\delta \to 0^-} \frac{f(x+\delta) - f(x)}{\delta})$$

Note we do not let $y = x$!!!!!!!!!

Definition 45. *The derivative of a function f is defined at a point x when the left hand derivative equals the right hand derivative. There are many ways to denote the derivative of the function f.*

$$\begin{aligned} f'(x) &= Df \\ &= \frac{df}{dx} \\ &\equiv \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \end{aligned}$$

Example 16. $f(x) = x^2$

$$\begin{aligned} f'(x) &= \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \\ &= \lim_{y \to x} \frac{y^2 - x^2}{y - x} \\ &= \lim_{y \to x} \frac{(y - x)(y + x)}{y - x} \\ &= \lim_{y \to x} (y + x) \\ &= 2x \end{aligned}$$

FIGURE GOES HERE

Example 17. *Consider the function*

$$f(x) = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0. \end{cases}$$

So we get from this that

$$\begin{aligned} f'_+(0) &= 1 \quad && \textit{right hand derivative} \\ f'_-(0) &= -1 \quad && \textit{left hand derivative} \end{aligned}$$

So obviously no derivative exists since

$$f'_+(0) \quad \neq \quad f'_-(0)$$

Now we can see how the derivative creates a first-order approximation. Assume the function f is defined on an open interval containing x. Let $A_1(y) = f(x) + f'(x)(y - x)$. This is the equation of the line with slope $f'(x)$ that passes through $(x, f(x))$. It follows from the definition of the derivative that

$$lim_{y \to x} \frac{f(y) - A_1(y)}{y - x} = 0. \tag{4.1}$$

Equation (4.1) explains why A_1 is a better approximation to f than A_0. Not only is the linear approximation A_1 close to f when y is close to x – this would be true if the limit of the numerator in (4.1) converged to zero as y converged to x. It is also the case that the linear approximation is close to f if you derive the difference by something really close to zero $(y - x)$. From this interpretation and the examples, it is not surprising that it is harder to be differentiable than it is to be continuous.

Theorem 26. *Consider the function*

$$f : \mathcal{X} \longrightarrow \mathbb{R}$$

where $\mathcal{X}$ is an open interval. If f is differentiable at a point x (i.e. $f'(x)$ exists), then f is continuous at x.

Proof.

$$\begin{aligned} \lim_{y \to x} f(y) - f(x) &= \lim_{y \to x} \frac{f(y) - f(x)}{y - x}(y - x) \\ &= f'(x) \cdot 0 \\ &= 0 \end{aligned}$$

where the second last equality is got from the fact that if you have 2 functions g and h and $lim_{y \to x} h(y)$ exists and $lim_{y \to x} g(y)$ exists, then

$$\lim_{y \to x} g(y)h(y) = \left(\lim_{y \to x} g(y) \right) \left(\lim_{y \to x} h(y) \right)$$

which is the corresponding version of Theorem 9 (3) for functions. □

As with continuity, it is tedious to use the definition to confirm that functions are differentiable. Instead, you note that certain functions are differentiable (constant functions and linear functions) and then apply a higher-level theorem. The next result should look familiar (compare the related statement about continuity).

Theorem 27. *Suppose f and g are defined on an open interval containing point x and both functions are differentiable at x. Then, $(f+g)$, $f \cdot g$, and f/g are differentiable at x (the last of these provided $g'(x) \neq 0$).*[1]

1.
$$(f+g)'(x) = f'(x) + g'(x)$$

2.
$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$$

3.
$$(\frac{f}{g})'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$$

Proof. 1. This should be clear by Theorem 9 (1), though again we are using the "functional version".

2. Let $h = fg$. Then

$$h(y) - h(x) = f(y)[g(y) - g(x)] + g(x)[f(y) - f(x)]$$

If we divide this by $y - x$ and note that $f(y) \longrightarrow f(x)$, and $g(y) \longrightarrow g(x)$ as $y \longrightarrow x$ by Theorem 26, then the result follows.
Then the result follows.

3. Now let $h = f/g$. Then

$$\frac{h(y) - h(x)}{y - x} = \frac{1}{g(y)g(x)} \left[g(x)\frac{f(y) - f(x)}{y - x} - f(x)\frac{g(y) - g(x)}{y - x} \right]$$

Now let $y \longrightarrow x$ and you are done.

□

Exercise 1. *A good exercise at this stage would be to prove that if*

$$f(x) = x^k$$

$$\Rightarrow f'(x) = kx^{k-1}$$

Hint *Prove it by induction and using property (2) of Theorem 27. We already have it for the case where $k = 2$ from Example 16*

[1] $f \cdot g$ denotes multiplication of functions. Composition, $f \circ g$, is different.

Theorem 28 (Chain Rule)**.** *Suppose that*

$$g : \mathcal{X} \longrightarrow \mathcal{Y}$$

and

$$f : \mathcal{Y} \longrightarrow \mathbb{R}$$

that g is differentiable at x and that f is differentiable at $y = g(x) \in \mathcal{Y}$.[2] *Then*

$$\begin{aligned}(f \circ g)'(x) &= f'(y)g'(x) \\ &= f'(g(x))g'(x)\end{aligned}$$

Proof. Define the function

$$h(z) = \begin{cases} \frac{f(g(z))-f(g(x))}{g(z)-g(x)} & \text{if } g(z) - g(x) \neq 0 \\ f'(g(x)) & \text{if } g(z) - g(x) = 0 \end{cases}$$

$$\begin{aligned}\lim_{z\to x} \frac{f(g(z)) - f(g(x))}{z - x} &= \lim_{z\to x} \left(h(z) \cdot \frac{g(z) - g(x)}{z - x} \right) \\ &= \left[\lim_{z\to x} h(z)\right] \left[\lim_{z\to x} \frac{g(z) - g(x)}{z - x}\right]\end{aligned}$$

To complete the proof it suffices to show that $\lim_{z\to x} h(z) = f'(g(x))$. We do this using the definition of the limits. Let $\epsilon > 0$. Because f' is differentiable at $g(x)$ I can find $\delta_1 > 0$ such that if $0 < |y - g(x)| < \delta_1$, then $\left|\frac{f(y)-f(g(x))}{y-g(x)} - f'(g(x))\right| < \epsilon$. By continuity of g at x, I can find $\delta > 0$ such that if $0 < absz - x < 0$, then $|g(z) - g(x)| < \delta_1$. It follows that if $0 < absz - x < \delta$ and $g(z) \neq g(x)$, then $\left|\frac{f(g(z))-f(g(x))}{g(z)-g(x)} - f'(g(x))\right| < \epsilon$. This means that if $0 < absz - x < \delta$ and $g(z) \neq g(x)$, then $|h(z) - f(g(x))| < \epsilon$. To complete the proof just note that if $g(z) = g(y)$ then $h(z) = f'(g(x))$. □

[2] In this theorem we are talking about $f \circ g$.

Definition 46. *Recall Definition 42. This is technically the definition of a* global max. *We say* globally *since it maximizes the function over the whole domain. However, we say x^* is a* local maximizer *of the function f if $\exists$ a segment (a, b) such that*

$$f(x^*) \geq f(x), \qquad \forall\, x \in (a, b)$$

Definition 47. *We say x^* is a* local minimizer *of the function f if $\exists$ a segment (a, b) such that*

$$f(x^*) \leq f(x), \qquad \forall\, x \in (a, b)$$

Note that a maximum that occurs at the boundary of the domain is not a local maximum, but that a maximum that occurs in the interior of the domain is a local maximum.

Theorem 29. *Suppose f is defined on $[a, b]$. If f has a local max at $c \in (a, b)$ and if $f'(c)$ exists, then*

$$f'(c) = 0$$

Proof. For $|x - c| < \delta$, we have

$$f(x) - f(c) \leq 0$$

Therefore, if $x \in (c, c + \delta)$, then

$$\frac{f(x) - f(c)}{x - c} \geq 0 \tag{4.2}$$

while for $x \in (c - \delta, c)$

$$\frac{f(x) - f(c)}{x - c} \leq 0. \tag{4.3}$$

Since f is differentiable at c, both left and right derivatives exist and they are equal. Inequality (4.2) states that the derivative from above must be nonpositive (if it exists). Inequality (4.3) states that the derivative from below must be nonnegative (if it exists). Since the derivative exists, these observations imply that the derivative must be both nonpositive and nonnegative. The only possibility is that $f'(c) = 0$. □

You can use essentially the same argument to conclude that if c is a local minimum and $f'(c)$ exists, then $f'(c) = 0$. The equation $f'(c) = 0$ is called a first-order condition. The theorem states that satisfying a first-order condition is

a necessary condition for c to be a local maximum or minimum. This observation may be the most important result in calculus for economics. Economists are interested in optimizing functions. The result says that you can replace solving an optimization problem (which seems complicated) with solving an equation (which perhaps is easier to do). The approach is powerful and generalizes. It suffers from several limitations. One limitation is that it does not distinguish local maxima from local minima. This is a major limitation of the statement of the theorem. If you examine the proof carefully, you will see that the calculus does distinguish maxima from minima. If you attempt to carry out the proof when c is a local minimum, the inequalities in (4.2) and (4.3) will be reversed. This means, loosely speaking, for a local maximum $f'(x) \geq 0$ for $x < c$ and $f'(x) \geq 0$ for $x > c$. That means, it appears that f' is decreasing. While for a local minimum the derivative is negative to the left of c and positive to the right.

A second limitation is that the theorem only applies to local extrema. It is possible that the maximum occurs at the boundary or there are many local maxima. Calculus still has something to say about boundary optima: If f is defined on $[a, b]$ and a is a maximum, then $f(a) \geq f(x)$ for all $x \in [a, b]$ so that the derivative from above must be nonpositive. Analogous statements are available for minima or for the right-hand endpoint. As for identifying which local maximum is a true maximum, calculus typically does not help. You must compare the values of the various candidates.

Finally, it is possible that $f'(c)$ equals zero, but c is neither a local maximum nor a local minimum. The standard example of this is $f(x) = x^3$ at the point $x = 0$.

In spite of the drawbacks, Theorem 29 gives a procedure for solving an optimization problem for a differentiable function on the interval $[a, b]$: Solve the equation $f'(c) = 0$. Call the set of solutions Z. Evaluate $f(x)$ for all $x \in Z \cup \{a, b\}$. The highest value of f over this set is the maximum. The lowest value is the minimum. This means that instead of checking the values of f over the entire domain, you need only check over a much smaller set.

An implication of the algorithm is that if a differentiable function's derivative is never zero, then it only has maxima and minima at the boundaries of its domain.

The previous result tells you that there is something special about places where the derivative is zero. It is also possible to interpret places where the derivative is positive or negative.

Definition 48. *The function f is increasing if $x > y$ implies $f(x) \geq f(y)$. The function f is strictly increasing if $x > y$ implies $f(x) > f(y)$. The function is increasing in a neighborhood of x if there exists $\delta > 0$ such that if $y \in (x - \delta, x + \delta)$ then $x > y$ implies $f(x) \geq f(y)$. The function is strictly increasing in a*

neighborhood of x if there exists $\delta > 0$ such that if $y \in (x - \delta, x + \delta)$ then $x > y$ implies $f(x) > f(y)$.

There are analogous definitions for deceasing and strictly decreasing. A function that is either (strictly) increasing everyone or decreasing everywhere is called (strictly) monotonic. There is a little bit of ambiguity about the term "increasing." Some people use "non-decreasing" to describe a function that we call increasing.

Theorem 30. *If f is a real valued function differentiable on (a, b), then*

1. *If $f'(x) > 0$, then f is strictly increasing in the neighborhood of x.*
2. *If $f'(x) > 0$ for all $x \in (a, b)$ then f is strictly increasing.*
3. *If f is increasing in the neighborhood of x, then $f'(x) \geq 0$.*
4. *If f is increasing, then $f'(x) \geq 0$ for all $x \in (a, b)$.*

The theorem almost says that differentiable functions are (strictly) increasing if and only if the derivative is nonnegative (positive). Alas, you can have functions that are strictly increasing but whose derivative is sometimes zero: $f(x) = x^3$ is again the standard example.

The proof of Theorem 30 is a straightforward exercise in the definition of the derivative. It requires only writing down inequalities similar to (4.2) and (4.3) and a little bit of care.

The previous two results given you excellent tools for graphing functions. You can use derivatives to identify when the function is increasing, decreasing, or has local maxima and minima. If you can figure out when the function crosses zero and its behavior at infinity, then you have a nice insight into into behavior.

Theorem 31 (Mean Value Theorem)**.** *If f is real valued and continuous on $[a, b]$, and differentiable on (a, b), then $\exists$ a point $c \in (a, b)$ such that*

$$f(b) - f(a) = (b - a)f'(c)$$

FIGURE GOES HERE

We can see from the picture above what this is saying

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

That is, there must be a point between a and b such that the derivative at that point (i.e. slope of the tangent to the curve) is the same as that of the line connecting our two points.

The mean value theorem is another way to think about how derivatives generate first-order approximations. Instead of saying that $f(x)$ is approximately $f(x_0) + f'(x_0)(x - x_0)$ it says that $f(x)$ is exactly $f(x_0) + f'(c)(x - x_0)$ for some value of c between x_0 and x. So one approach gives you an approximation, but to get the approximation you know exactly where to evaluate f'. The other approach gives you an exact expression, but you do not know where to evaluate the derivative.

Proof. Define

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

We know that $g(x)$ is continuous on compact $[a, b]$. Thus $g(x)$ has a max at some point $c \in (a, b)$ and $g'(c) = 0$. Thus

$$\begin{aligned} g'(c) &= f'(c) - \frac{f(b) - f(a)}{b - a} \\ &= 0. \end{aligned}$$

□

The Mean Value Theorem has several applications.

Theorem 32. *Suppose f is real valued and continuous on $[a, b]$, and differentiable on (a, b). If $f'(x) \equiv 0$ for $x \in (a, b)$, then f is constant.*

The result is intuitively obvious and perhaps something that you would implicitly assume. It does require proof. (Notice that the converse is true too: If f is constant, then it is differentiable and its derivative is always zero.)

Proof. By the Mean Value Theorem, for all $x \in [a, b]$,

$$f(x) - f(a) = f'(c)(x - a)$$

for some $c \in [a, x]$. Since the right-hand side is zero by assumption, we have $f(x) = f(a)$ for all x, which is the desired result. □

This is your first differential equation! That is, you found all of the functions that satisfy an equation involving a derivative.

Here is another application.

Theorem 33. *Suppose that f is a continuous real-valued function defined on the interval $(-1, 1)$. Further suppose that $f'(x)$ exists for all $x \neq 0$ and that $lim_{x\to 0} f'(x)$ exists. Prove that f is differentiable at $x = 0$ and that f' is continuous at $x = 0$.*

Proof. By the Mean Value Theorem,

$$\frac{f(x) - f(0)}{x} = f'(c)$$

for c between 0 and x. Since f' is continuous at 0, $\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^-} f'(x)$. It follows that

$$\lim_{x\to 0^+} \frac{f(x) - f(0)}{x} = \lim_{c\to 0^+} f'(c) = \lim_{c\to 0^-} f'(c) = \lim_{x\to 0^-} \frac{f(x) - f(0)}{x},$$

which establishes that $f'(0)$ exists and is equal to $\lim_{c\to 0} f'(c)$. □

Theorem 34. *If*

$$g : \mathbb{R} \longmapsto \mathbb{R}$$

is the inverse of

$$f : \mathbb{R} \longmapsto \mathbb{R}$$

and if f is strictly increasing and differentiable, then

$$g'(f(x)) = \frac{1}{f'(x)}$$

Note: f absolutely must be strictly increasing since otherwise it's not invertible and hence g is not well defined.

Theorem 35. ***(L'Hopital's Rule)*** *Consider functions f and g differentiable on $[a, b)$, where $g'(x) \neq 0$ on $[a, b)$.*
If either

1. $\lim_{x\to b^-} f(x) = 0$, *and* $\lim_{x\to b^-} g(x) = 0$

   OR

2. $\lim_{x\to b^-} f(x) = \infty$, *and* $\lim_{x\to b^-} g(x) = \infty$

and further

$$\lim_{x \to b^-} \frac{f'(x)}{g'(x)} = L \in \mathbb{R}$$

Then

$$\lim_{x \to b^-} \frac{f(x)}{g(x)} = L$$

L'Hopital's Rule is a useful way to evaluate indeterminate forms $(0/0)$. It looks a bit magical: How can the ratio of the functions be equal to the ratio of derivatives? You prove that the rule works by using a variation of the mean value theorem. In Case 1, what is going on is that $f(x) = f(b) + f'(c)(x - b)$ for some $c \in (x, b)$ and similarly $g(x) = g(b) + g'(d)(x - b)$. Since $f(b) = g(b) = 0$ (loosely), the ratio of f to g is the ratio of derivatives of f and g. The trouble is that these derivatives are evaluated at different points. The good news is that you can prove a version of the Mean-Value Theorem in which allows you to take $c = d$. This enables you to prove the theorem.

Warning: The Rule needs the two conditions to hold. If you try to evaluate

$$lim_{x \to 0} \frac{x + 1}{x^2 + 1}$$

by setting the ratio equal to $1/(2x)$ and taking the limit you would be making a big mistake. (The ratio is a rational function and the denominator is positive, so it is continuous. Therefore the limit is just the function's value at 0: 1.)

Chapter 5

Taylor's Theorem

Using the first derivative, we were able to come up with a way to find the best linear approximation to a function. It is natural to ask whether it is possible to find higher order approximations. What does this mean? By analogy with zeroth and first order approximations, we first decide what an appropriate approximating function is and then what the appropriate definition of approximation is.

First-order approximations were affine functions. In general, an nth-order approximation is a polynomial of degree n, that is a function of the form

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n.$$

Technically, the degree of a polynomial is the largest power of n that appears with non-zero coefficient. So this polynomial has degree n if and only if $a_n \neq 0$.

Plainly a zeroth degree polynomial is a constant, a first degree polynomial is an affine function, a second degree polynomial is a quadratic, and so on. An n^{th} order approximation of the function f at x is a polynomial of degree at most n, A_n that satisfies

$$\lim_{y \to x} \frac{f(y) - A_n(y)}{(y-x)^n} = 0.$$

This definition generalizes the earlier definition. Notice that the denominator is a power of $y - x$. When y approaches x the denominator is really small. If the ratio has limit zero it must be that the numerator is really, really small. We know that zeroth order approximations exist for continuous functions and first-order approximations exist for differentiable functions. It is natural to guess that higher-order approximations exist under stricter assumptions. This guess is correct.

Definition 49. *The nth derivative of a function f, denoted f^n, is defined inductively to be the derivative of $f^{(n-1)}$.*

We say f is of class C^n *on (a,b) ($f \in C^N$) if $f^{(n)}(x)$ exists and is continuous $\forall\, x$.*

One can check that, just as differentiability of f implies continuity of f, if f^n exists, then f^{n-1} is continuous.

Theorem 36 (Taylor's Theorem)**.** *Let $f \in C^n$ and assume that f^n exists on (a,b) and let c and d be any points in (a,b). Then there exists a point t between c and d and a polynomial A_n of degree at most n such that*

$$f(d) = A_n(d) + \frac{f^{(n+1)}(t)}{(n+1)!}(d-c)^{n+1}, \tag{5.1}$$

where A_n is the Taylor Polynomial for f centered at c:

$$A_n(d) = \sum_{j=0}^{n} \frac{f^{(k)}(c)}{k!}(d-c)^k.$$

The theorem decomposes f into a polynomial and an error term $E_n = \frac{f^{(n+1)}(t)}{(n+1)!}(d-c)^{n+1}$. Notice that

$$\lim_{d\to c} \frac{E_n}{(d-c)^n} = 0$$

so the theorem states that the Taylor polynomial is, in fact, the nth order approximation of f at c.

The form of the Taylor approximation may seem mysterious at first, but the coefficients can be seen to be the only choices with the property that $f^{(k)}(c) = A_n^{(k)}(c)$ for $k \le n$. As impressive as the theorem appears, it is just a disguised version of the mean-value theorem.

FIGURE GOES HERE

Proof. Define

$$F(x) \equiv f(d) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(d-x)^k$$

and

$$G(x) \equiv F(x) - \left(\frac{d-x}{d-c}\right)^{n+1} F(c).$$

It follows that $F(d) = 0$ and (lots of terms cancel) $F'(x) = -\frac{f^{(n+1)}(x)}{n!}(d-x)^n$. Also, $G(c) = G(d) = 0$. It follows from the mean value theorem that there exists a t between c and d such that $G'(t) = 0$. That is, there exists a t such that

$$0 = -\frac{f^{(n+1)}(x)}{n!}(d-x)^n + \left(\frac{d-t}{d-c}\right)^n F(c)$$

or

$$F(c) = \frac{f^{(n+1)}(t)}{(n+1)!}(d-c)^{n+1}.$$

An examination of the definition of F confirms that this completes the proof. □

Taylor's Theorem has several uses. As a conceptual tool it makes precise the notion that well behaved functions have polynomial approximations. This permits you to understand "complicated" functions like the logarithm or exponential by using their Taylor's expansion. As a computational tool, it permits you to compute approximate values of functions. Of course, doing this is not practical (because calculators and computers are available). As a practical tool, first- and second-order approximations permit you to conduct analyses in terms of linear or quadratic approximations. This insight is especially important for solving optimization problems, as we will see soon.

Next we provide examples of the first two uses.

Consider the logarithm function: $f(x) = \log x$.[1] This function is defined for $x > 0$. It is not hard to show that $f^{(k)}(x) = x^{-k}(-1)^{k-1}(k-1)!$. So $f^{(k)}(1) = (-1)^{k-1}(k-1)!$ Hence:

$$f(x) = \sum_{k=1}^{N}(-1)^{k-1}\frac{(x-1)^k}{k} + E_N$$

where $E_N = (-1)^N\frac{(y-1)^N}{N+1}$ for y between 1 and x. Notice that this expansion is done around $x_0 = 1$. This is a point at which the function is nicely behaved. Next notice that the function $f^{(k)}$ is differentiable at 1 for all k. This suggests that you can extend the polynomial for an infinite number of terms. It is the case that

$$\log(x) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{(x-1)^k}{k}.$$

Sometimes this formula is written in the equivalent form:

[1] Unless otherwise mentioned, logarithms are always with respect to the base e.

$$log(y+1) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{y^k}{k}.$$

The second way to use Taylor's Theorem is to find approximations. The formula above can let you compute logarithms. How about square roots? The Taylor's expansion of the square root of x around 1 is:

$$\sqrt{2} = 1 + .5 - .125 + E_2.$$

$E_2 = x^{-2.5}/16$ for some $x \in [1, 2]$. Check to make sure you know where the terms come from. The approximation says that $\sqrt{2} = 1.375$ up to an error. The error term is largest when $x = 1$. Hence the error is no more than $.0625$. The error term is smallest when $x = 2$. I'm not sure what the error is then, but it is certainly positive. Hence I know that the square root of 2 is at least 1.375 and no greater than 1.4275. Perhaps this technique will come in handy the next time you need to compute a square root without the aid of modern electronics.

Chapter 6

Univariate Optimization

Economics is all about optimization subject to constraints. Consumers do it. They maximize utility subject to a budget constraint. Firms do it. They maximize profits (or minimize costs) subject to technological constraints. You will study optimization problems in many forms. The one-variable theory is particularly easy, but it teaches many lessons that extend to general settings.

There are two features of an optimization problem: the *objective function*, which is the real-valued function[1] that you are trying to maximize or minimize, and the constraint set.

You already know a lot about one-variable optimization problems. You know that continuous functions defined on closed and bounded intervals attain their maximum and minimum values. You know that local optima of differentiable functions satisfy the first-order condition.

The next step is to distinguish between local maxima and local minima.

Suppose that f differentiable on an open interval. The point x^* is a critical point if $f'(x^*) = 0$. We know that local minima and local maxima must be critical points. We know from examples that critical points need not be local optima. It turns out that properties of the second derivative of f classify critical points.

Theorem 37 (Second-Order Conditions)**.** *Let f be twice continuously differentiable on an open interval (a, b) and let $x^* \in (a, b)$ be a critical point of f. Then*

1. *If x^* is a local maximum, then $f''(x^*) \leq 0$.*

2. *If x^* is a local minimum, then $f''(x^*) \geq 0$.*

[1] Although the domain of the function you are trying to optimize varies in applications, the range is typically the real numbers. In general, you need the range to have an ordering that enables you to compare any pair of points.

3. *If $f''(x^*) < 0$, then x^* is a local maximum.*

4. *If $f''(x^*) > 0$, then x^* is a local minimum.*

Conditions (1) and (3) are almost converses (as are (2) and (4)), but not quite. Knowing that x^* is a local maximum is enough to guarantee that $f''(x^*) \leq 0$. Knowing that $f''x^*)$ is not enough to guarantee that you have a local minimum (you may have a local maximum or you may have neither a minimum nor a maximum). (All the intuition you need comes from thinking about the behavior of $f(x) = x^n$ at $x = 0$ for different values of $x = 0$.) You might think that you could improve the statements by trying to characterize strict local maxima. It is true that if $f''(x^*) < 0$, then x^* is a strict local maximum, but it is possible to have a strict local maximum and $f''(x^*) = 0$. The conditions in Theorem 37 parallel the results about first derivatives and monotonicity stated earlier.

Proof. By Taylor's Theorem we can write:

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2}f''(t)(x - x^*)^2 \tag{6.1}$$

for t between x and x^*. If $f''(x^*) > 0$, then by continuity of f'', $f''(t) > 0$ for t sufficiently close to x^* and so, by (6.1), $f(x) > f(x^*)$ for all x sufficiently close to x^*. Consequently, if x^* is a local maximum, $f''(x^*) \leq 0$, proving (1). (2) is similar.

If $f''(x^*) < 0$, then by continuity of f'', there exists $\delta > 0$ such that if $0 < |x - t| < \delta$, then $f''(t) < 0$. By (6.1), it follows that if $0 < |x - x^*| < \delta$, then $f(x) < f(x^*)$, which establishes (3). (4) is similar. □

The theorem allows us to refine our method for looking for maxima. If f is defined on an interval (and is twice continuously differentiable), the maximum (if it exists) must occur either at a boundary point or at a critical point x^* that satisfies $f(x^*) \leq 0$. So you can search for maxima by evaluating f at the boundaries and at the appropriate critical points.

This method still does not permit you to say when a local maximum is really a global maximum. You can do this only if f satisfies the appropriate global conditions.

Definition 50. *We say a function f is* concave *over an interval $\mathcal{X} \subset \mathbb{R}$ if $\forall x, y \in \mathcal{X}$ and $\delta \in (0, 1)$, we have*

$$f(\delta x + (1 - \delta)y) \geq \delta f(x) + (1 - \delta)f(y) \tag{6.2}$$

If f is only a function of one argument you can think of this graph as having an inverted "u" shape.

Geometrically the definition says that the graph of the function always lies above segments connecting two points on the graph. Another way to say this is that the graph of the function always lies below its tangents (when the tangents exist). If the inequality in (6.2) is strict, then we say that the function is strictly concave. A linear function is concave, but not strictly concave.

Concave functions have nicely behaved sets of local maximizers. It is an immediate consequence of the definition that if x and y are local maxima, then so are all of the points on the line segment connecting x to y. A fancy way to get at this result is to note that concavity implies

$$f(\delta x + (1-\delta)y) \geq \delta f(x) + (1-\delta)f(y) \geq \min\{f(x), f(y)\}. \tag{6.3}$$

Moreover, the inequality in (6.3) is strict if $\delta \in (0,1)$ and either (a) $f(x) \neq f(y)$ or (b) f is strictly concave. Suppose that x is a local maximum of f. It follows that $f(x) \geq f(y)$ for all y. Otherwise $f(\lambda x + (1-\lambda)y) > f(x)$, for all $\lambda \in (0,1)$, contradicting the hypothesis that x is a local maximum. This means that any local maximum of f must be a global maximum. It follows that if x and y are both local maxima, then they both must be global maximal and so $f(x) = f(y)$. In this case it follows from (6.3) that all of the points on the segment connecting x and y must also be maxima. It further implies that $x = y$ if f is strictly concave.

These results are useful. They guarantee that local extrema are global maxima (so you know that a critical point must be a maximum without worrying about boundary points or local minima) and they provide a tractable sufficient condition for uniqueness. Notice that these nice properties follow from (6.3), which is a weaker condition than (6.2).[2] This suggests that the following definition might be useful.

Definition 51. *We say a function f is* quasi concave *over an interval $\mathcal{X} \subset \mathbb{R}$ if $\forall x, y \in \mathcal{X}$ and $\delta \in (0,1)$, we have $f(\delta x + (1-\delta)y) \geq \min\{f(x), f(y)\}$.*

If quasi-concavity is so great, why bother with concavity? It turns out that concavity has a characterization in terms of second derivatives.

We can repeat the same analysis with signs reversed.

Definition 52. *We say a function f is* convex *over an interval $\mathcal{X} \subset \mathbb{R}$ if $\forall x, y \in \mathcal{X}$ and $\delta \in (0,1)$, we have*

$$f(\delta x + (1-\delta)y) \leq \delta f(x) + (1-\delta)f(y)$$

If f is only a function of one argument you can think of this graph as having an "u" shape.

[2] As an exercise, try to find a function that satisfies (6.3) but not (6.2).

Convex functions have interior minima and differentiable convex functions have positive second derivatives.

Theorem 38. *Let $f\colon \mathcal{X} \longrightarrow \mathbb{R}$, $\mathcal{X}$ an open interval and $f \in C^2$ on $\mathcal{X}$.*
$f''(x) \leq 0, \quad \forall x \in \mathcal{X}$*, if and only if f is concave on $\mathcal{X}$.*

Proof. If f is concave, then for all $\lambda \in (0,1)$

$$\frac{f(\lambda x + (1-\lambda)y) - f(x)}{1-\lambda} \geq \frac{f(y) - f(\lambda x + (1-\lambda)y)}{\lambda}.$$

Routine manipulation demonstrates that the limit of the left-hand side (if it exists) as $\lambda \to 1$ is equal to $(y-x)f'(x)$ (note that $f(\lambda x + (1-\lambda)y) - f(x) = f(x + (1-\lambda)(y-x)) - f(x)$) and the limit of the right-hand side $\lambda \to 1$ is equal to $(y-x)f'(y)$. It follows that if f is concave and differentiable, then

$$(y-x)f'(x) \geq (y-x)f'(y),$$

which in turn implies that f' is decreasing. Consequently, if f'' exists, then it is non-positive.

Conversely, if f is differentiable, then by the Mean Value Theorem

$$f(\lambda x + (1-\lambda)y) - f(x) = (1-\lambda)f'(c)(y-x)$$

for some c between x and $\lambda x + (1-\lambda)y$. You can check that this means that if f' is decreasing, then

$$f(\lambda x + (1-\lambda)y) - f(x) \geq (1-\lambda)f'(\lambda x + (1-\lambda)y)(y-x). \tag{6.4}$$

Similarly,

$$f(\lambda x + (1-\lambda)y) - f(y) = -\lambda f'(c)(y-x)$$

for some c between $\lambda x + (1-\lambda)y$ and y and if f' is decreasing, then

$$f(\lambda x + (1-\lambda)y) - f(y) \geq -\lambda f'(\lambda x + (1-\lambda)y)(y-x). \tag{6.5}$$

The result follows from adding λ times inequality (6.4) to $1-\lambda$ times inequality (6.5). □

Note

- The previous theorem used the fact that f' was decreasing (rather than $f'' \leq 0$).

- The Mean-Value Theorem says that if f is differentiable, then $f(y)-f(x) = f'(c)(y-x)$ for some c between x and y. You can check that this means that if f' is decreasing, then $f(y) - f(x) \leq f'(x)(y-x)$. This inequality is the algebraic way of expressing the fact that the tangent line to the graph of f at $(x, f(x))$ lies above the graph.

- If $f'' < 0$ so that f is strictly concave, then f' is strictly decreasing. That means that f can have at most one critical point. Since $f'' < 0$, this critical point must be a local maximum. These statements simply reiterate earlier observations: Local maxima of strictly concave functions are global maxima and these must be global maxima.

Chapter 7

Integration

FIGURES GO HERE

7.1 Introduction

Integration is a technique that does three superficially different things. First, it acts an the inverse to differentiation. That is, if you have the derivative of a function and want to know the function, then the process of "antidifferentiation" is closely related to the integration. Finding antiderivatives is essential if you want to solve differential equations.

Second, integration is a way to take averages of general functions. This interpretation is the most natural and insightful one for the study of integration in probability and statistics.

Third, for non-negative functions, the integral is a way to compute areas.

The connection between the second and third interpretations is fairly straightforward. The connection between the second and first interpretation is important and surprising and is a consequence of "The Fundamental Theorems of Calculus."

We will motivate the definition of integral as a generalized average. If you are given a set of N numbers, you average them by adding the numbers up and then dividing by N. How do you generalize this to a situation in which you are given an infinite set of numbers to average?

Assume that $f : [a, b] \longrightarrow \mathbb{R}$. A naive way to guess the average value of f would be to pick some $x \in [a, b]$ and say that the average is equal to $f(x)$. This would work great if f happened to be constant. If you could find the smallest and largest values of f on the interval (this would be possible if f were continuous),

then you could use these are lower and upper bounds for the average. If f did not vary too much, perhaps you could use these values to get a good estimate of the value of f's average. You could get an every better estimate if you subdivided the interval $[a, b]$ and repeated the process: The lower bound for the average value of f would be the average of the minimum value of f on $[a, (a+b)/2]$ and $[(a+b)/2, b]$. When you do this, the estimate for the lower bound will be higher (at least it won't be lower) and the estimate for the upper bound will be no higher than the original estimate. Maybe if you keep repeating the process the upper and lower bounds converge to something that would be a good candidate for the average. The theory of integration is primarily about finding conditions under which this kind of argument works.

A partition P of $[a, b]$ is a finite set of points $x_0, x_1, \ldots, x_n$, where

$$a = x_0 \leq x_1 \leq \cdots \leq x_{n-1} \leq x_n \leq b$$

Given such a partition, define

$$L_f(P) = \sum_{k=1}^{n} m_k \Delta_k$$

where

$$\begin{aligned} m_k &\equiv \inf_{x \in [x_{k-1}, x_k]} f(x) \\ &= \inf \{ f(x) \mid x \in [x_{k-1}, x_k] \} \end{aligned}$$

and

$$\Delta_k \equiv x_k - x_{k-1}$$

$$U_f(P) = \sum_{k=1}^{n} M_k \Delta_k$$

where

$$\begin{aligned} M_k &\equiv sup_{x \in [x_{k-1}, x_k]} f(x) \\ &= \sup \{ f(x) \mid x \in [x_{k-1}, x_k] \} \end{aligned}$$

Since these definitions are defined in terms of "sup" and "inf" they are well defined even if f is not continuous. (If f is continuous, then the maxima and

minima are attained on each subinterval, so you can replace $\sup$ by $\max$ and $\inf$ by $\min$ in the definitions.) It is clear that for each partitition P, $L_f(P) \leq U_f(P)$. Also, $L_f(P)$ is an underestimate of the value that we want while $U(P)$ is an overestimate.[1]

Now imagine subdividing the partition P by dividing each subset of P into two non-empty pieces. This leads to a new partition P' and new values $L_f(P') \geq L_f(P)$ and $U_f(P') \leq U_f(P)$. The reason that subdividing increasing the lower sums is that

$$\inf_{x\in[x_{k-1},x_k]} f(x) \leq \inf_{x\in[x_{k-1},z]} f(x) + \inf_{x\in[z,x_k]} f(x) \text{ for all } z \in [x_{k-1}, x_k]$$

because the value of x that makes f smallest in the expression on the left may be in only one of the two subintervals.

So far we have a process that generates an increasing sequence of lower estimates of the average value of f and a process that generates a decreasing sequence of upper estimates of the average value of f. We know that bounded monotone sequences converge. This motivates the following definition.

Definition 53. *Let $f : [a, b] \longrightarrow \mathbb{R}$ and suppose that P_r is a sequence of partitions*

$$a = x_0(r) \leq x_1(r) \leq \cdots \leq x_{n-1}(r) \leq x_{n_r}(r) \leq b$$

such that $\Delta_k(r) = x_k(r) - x_{k-1}(r)$ goes to zero as r approaches infinity for all k. f is integrable *if*

$$\lim_{r\to\infty} L_f(P_r) \text{ exists and is equal to } \lim_{r\to\infty} U_f(P_r). \tag{7.1}$$

If f is integrable then we denote the common limit in (7.1) by $\int_a^b f(x)dx$.

In the definition, there are lots of ways to take "finer and finer" partitions. It turns out that if the upper and lower limits exist and are equal, then it does not matter which partition you take. They will all converge to the same limit as the length of each element in the partition converges to zero. It also turns out that if a function is integrable, then it does not matter whether you evaluate f inside each partition element using the $\sup$, the $\inf$ or any value in between. All choices will lead to the same value.

[1]The formulas for U and L are the standard ones, but they are not quite right for the "average" interpretation. If $f(x) \equiv c$, then we would have $U_f(P) = L_f(P) = c(b - a)$ – the length of the interval appears in the formula. In order to maintain the interpretation as average, you must divide the formulas by the length of the interval, $b - a$.

If a function is integrable, then the integral (divided by $b - a$) is a perfect generalization of average value. The next question is: Which functions are integrable? It is not hard to come up with examples of non-integrable functions: The function that is 1 on the rationals and 0 on the irrationals has the property that the lower sums is equal to 0 and the upper sum is equal to $b - a$. One can show that continuous functions must be integrable (because you can make the difference between $U_f(P)$ and $L_f(P)$ arbitrarily small by making the size of the pieces of the partition small). It is useful in some situations to look at families of functions that are discontinuous and still integrable. For example, monotonic functions may be discontinuous, but bounded monotonic functions are always integrable.

The definition above is the definition of the *Riemann Integral*. You may hear about other kinds of integral. A *Riemann-Stieltjes Integral* is a generalization of the Riemann Integral in which the "length" of the interval is not uniform. Measure theory is a mathematical area that allows you to define integrals of functions defined on abstract sets equipped with a "measure" that gives you the "size" of various sets. One example of a measure is a probability distribution. The integrals used in measure theory are called expectations by probability theorists. Lebesgue Measure is the most common measure. The Lebesgue measure defines the size of the interval $[a, b]$ to be $b - a$ and, more generally, the size of rectangles to be their (standard) area. The basic theory is similar to the theory of Riemann integration and you should not be intimidated when people talk about Lebesgue integrals.[2] The theory of stochastic calculus, which is used in continuous-time finance, uses a different kind of integral. For this theory it is important to compute integrals of functions that are not integrable in the sense of Riemann. (That is, the class of functions that includes all continuous things and some other objects just isn't big enough.) In this theory it does matter where you evaluate f in the definition of the approximating sums. This leads to a theory that is different from classic integration theory in several respects.

Theorem 39. *If*

$$f(x) = c, \quad \forall x \in [a, b]$$

Then

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

[2]To define the Lebesgue integral you approximate a function by a sequence of "simple" functions that take on only finitely many values and hence are easy to average.

Definition 54. *Take* $f\colon (a,b) \longrightarrow \mathbb{R}$.
Suppose $\exists\, F\colon (a,b) \longrightarrow \mathbb{R}$*, continuous and differentiable on* (a,b).
If

$$F'(x) = f(x),$$

$\forall x \in (a,b)$.
Then F *is called the* antiderivative *of* f.

Example 18.

$$f(x) = x^2$$

$$\implies F(x) = \frac{1}{3}x^3$$

also could have

$$F(x) = \frac{1}{3}x^3 + 6$$

Theorem 40. *If* F *and* G *are both antiderivatives of* f *on* $[a,b]$*, then*

$$G(x) = F(x) + c$$

Note what this is saying is that any 2 antiderivatives differ only by a constant.

7.2 Fundamental Theorems of Calculus

Theorem 41 (Fundamental Theorem of Calculus)**.** *If*

$$f\colon \mathcal{X} \longrightarrow \mathbb{R}$$

is continuous and $\mathcal{X}$ *is an open interval then, for* $a \in \mathcal{X}$ *and* F *defined by*

$$F(x) \equiv \int_a^x f(t)dt, \quad \forall\, x \in \mathcal{X}$$

it is the case that F *is differentiable and*

$$F'(x) = f(x), \quad \forall\, x \in \mathcal{X}$$

The theorem states that differentiation and integration are inverse operations in the sense that if you start with a function and integrate it, then you get a differentiable function and the derivative of that function is the function you started with.

Proof. By the definition of the integral,

$$h \sup_{y\in[x,x+h]} f(y) \geq F(x+h) - F(x) = \int_x^{x+h} f(x)dx \geq h \inf_{y\in[x,x+h]} f(y)$$

the result follows from dividing through by h and taking limits. □

Theorem 42. *If F is an antiderivative of*

$$f\colon [a,b] \longrightarrow \mathbb{R}$$

then,

$$\begin{aligned}\int_a^b f(x)dx &= F(b) - F(a) \\ &= [F(x)]_a^b\end{aligned}$$

Proof. By the mean value theorem $F(y)-F(x) = f(t)(y-x)$ for some t between x and y. It follows that for any partition P

$$U_f(P) \geq \sum_{k=1}^{n} (F(x_k) - F(x_{k-1}) \geq L_f(P) \tag{7.2}$$

because for each element of the partition $F(x_k) - F(x_{k-1}) = f(t_k)(x_k - x_{k-1})$ and $f(t_k) \in [\inf_{x\in[x_{k-1},x_k]} f(x), \sup_{x\in[x_{k-1},x_k]} f(x)]$. Since

$$\sum_{k=1}^{n} (F(x_k) - F(x_{k-1}) = F(b) - F(a)$$

the result follows by taking limits of finer and finer partitions in (7.2). □

This theorem is a converse of the first result.

A bit of terminology: An antiderivative of f is a function whose derivative is f. This is sometimes called a primitive of f or the indefinite integral of f and is denoted $\int f$ (where the limits of integration are not specified). When the limits of integration are specified, $\int_a^b f$ is a number, called the definite integral of f on the interval $[a,b]$.

Up until now, we have only talked about integration over closed and bounded intervals. It is sometimes convenient to talk about "improper" integrals in which the limits of the integrals may be infinity. These integrals are defined to be limits of integrals over bounded intervals (provided that the limits exist).

7.3 Properties of Integrals

Theorem 43. *1.*

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

where $\lambda, \mu \in \mathbb{R}$.

2. *If f and g are continuous on $[a, b]$ with antiderivatives F and G, then*

$$\int_a^b f(x)G(x) \quad = \quad [F(x)G(x)]_a^b \quad - \quad \int_a^b F(x)g(x)dx$$

3. *Suppose π is strictly monotonic on $[a, b]$ and f is continuous on an open interval containing $\pi([a, b])$. Then*

$$\int_{\pi(a)}^{\pi(b)} f(x)dx \quad = \quad \int_a^b f(\pi(t))\pi'(t)dt$$

4. *If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, then there exists $t \in [a, b]$ such that*

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

Proof. 1. A direct consequence of the definitions.

2. There are 2 ways to see this

   (a) We know from Theorem 27 (2) that

   $$\frac{d}{dx}(FG) = fG + Fg$$

   $$\implies fG = \frac{d}{dx}(FG) - Fg$$

   and we can just integrate to get the desired result

   (b) Alternatively we can set $H(x) = F(X)G(X)$ an apply Theorem 42.

3. Exercise

4. This result is a simple consequence of the intermediate value theorem. By the definition of the integral it cannot be that

$$\int_a^b f(x)dx > f(t)(b-a)$$

for all $t \in [a,b]$ nor can it be that

$$\int_a^b f(x)dx < f(t)(b-a)$$

for all $t \in [a,b]$. Hence

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

is a continuous function of t that changes sign on $[a,b]$ so it must equal zero for some value of t.

□

The second formula is called integration by parts and it comes up a lot. The third formula is the change of variables formula. The fourth result is called the mean value theorem for integrals and states that at some point in an interval a function must take on the average value of the function on the interval.

7.4 Computing Integrals

Since every continuous function is integrable, we know a lot of functions that are integrable. Coming up with formulas for the integrals is not easy in practice. Polynomials are easy. One can integrate the exponential function ($\int e^x = e^x$) and the identity

$$\int_1^x \frac{1}{t}dt = \log x$$

is often taken to be the definition of the $\log$ function. Other than that, the change of variables formula and integration by parts are the primary methods we have to find integrals.

Chapter 8

Basic Linear Algebra

Many of the elements of one-variable differential calculus extend naturally to higher dimensions. The definition of continuity is (stated appropriately) identical. Maxima exist for real-valued continuous functions defined on the appropriate domain. The derivative still plays the role of linear approximation. Critical points play the same role in optimization theory. Taylor's Theorem and Second-Order Conditions reappear. Before we return to calculus, however, we need to record a few facts about domains that are rich enough to permit many variables.

8.1 Preliminaries

Definition 55. *n-dimensional Euclidean Space:*

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} \times \mathbb{R}$$

Note if $\mathcal{X}$ and $\mathcal{Y}$ are sets, then

$$\mathcal{X} \times \mathcal{Y} \equiv \{(x, y) \mid x \in \mathcal{X}, y \in \mathcal{Y}\}$$

so

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}, \forall\, i = 1, 2, \ldots, n\}$$

There are different ways to keep track of elements of $\mathbb{R}^n$. When doing calculus, it is standard to think of a point in $\mathbb{R}^n$ as a list of n real numbers, and write $x = (x_1, \ldots, x_n)$. When doing linear algebra, it is common to think of the

elements of $\mathbb{R}^n$ as column vectors. When we do this we write

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Given $\mathbf{x} \in \mathbb{R}^n$, we understand x_i to be the i^{th} coordinate. In these notes, we will try (and almost certainly fail) to denote vectors using bold face ($\mathbf{x}$) and elements more plainly (x). Although there may be exceptions, for the most part you'll see $\mathbf{x}$ in discussions of linear algebra and x in discussions about calculus.

Definition 56. *The zero element:*

$$\boldsymbol{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Definition 57 (Vector Addition)**.** *For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we have*

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Vector addition is commutative

$$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$$

Definition 58 (Scalar Multiplication)**.** *For $\boldsymbol{x} \in \mathbb{R}^n$, and $a \in \mathbb{R}$ we have*

$$a\boldsymbol{x} = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}$$

In other words every element of $\boldsymbol{x}$ gets multiplied by a.

You may hear people talk about vector spaces. Maybe they are showing off. Maybe they really need a more general structure. In any event a vector space is a general set of $\mathcal{V}$ in which the operation of addition and multiplication by a scalar make sense, where addition is commutative and associative (as above), there is a special zero vector that is the additive identity ($\mathbf{0} + \mathbf{v} = \mathbf{v}$), additive inverses exist (for each $\mathbf{v}$ there is a $-\mathbf{v}$, and where scalar multiplication is defined as above. Euclidean Spaces are the leading example of vector spaces. We will need to talk about subsets of Euclidean Spaces that have a linear structure (they contain $\mathbf{0}$, and if $\mathbf{x}$ and $\mathbf{y}$ are in the set, then so is $\mathbf{x} + \mathbf{y}$ and all scalar multiples of $\mathbf{x}$ and $\mathbf{y}$). We will call these subspaces (this is a correct use of the technical term), but we have no reason to talk about more general kinds of vector spaces.

8.2 Matrices

Definition 59. *An $m \times n$ matrix is an element of $\mathcal{M}^{m\times n}$ written as in the form*

$$\boldsymbol{A} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} = [\alpha_{ij}]$$

where m denotes the number of rows and n denotes the number of columns.

Note An $m \times n$ matrix is just of a collection of nm numbers organized in a particular way. Hence we can think of a matrix as an element of $\mathbb{R}^{m\times n}$. The extra notation $\mathcal{M}^{m\times n}$ makes it possible to distinguish the way that the numbers are organized. *Note* We denote vectors in **boldface** lower-case letters. Matrices are represented in capital **boldface**.

Note Vectors are just a special case of matrices. e.g.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}^{n\times 1}$$

In particular real numbers are just another special case of matrices. e.g. the number 6

$$6 \in \mathbb{R} = \mathbb{R}^{1\times 1}$$

This notation emphasizes that we think of a vector with n components as a matrix with n rows and 1 columns.

Example 19.

$$\underset{2\times 3}{\boldsymbol{A}} = \begin{pmatrix} 0 & 1 & 5 \\ 6 & 0 & 2 \end{pmatrix}$$

Definition 60. *The* transpose *of a matrix* $\boldsymbol{A}$*, is denoted* $\boldsymbol{A}^t$*. To get the transpose of a matrix, we let the first row of the original matrix become the first column of the new (transposed) matrix. Using Definition 59 we would get*

$$\boldsymbol{A}^t = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{1n} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{1m} & \alpha_{2m} & \cdots & \alpha_{nm} \end{pmatrix} = [\alpha_{ji}]$$

Definition 61. *A matrix* $\boldsymbol{A}$ *is* symmetric *if* $\boldsymbol{A} = \boldsymbol{A}^t$.

So we can see that if $\mathbf{A} \in \mathcal{M}^{m\times n}$, then $\mathbf{A}^t \in \mathcal{M}^{n\times m}$.

Example 20. *Using Example 19 we see that*

$$\underset{3\times 2}{\boldsymbol{A}^t} = \begin{pmatrix} 0 & 6 \\ 1 & 0 \\ 5 & 2 \end{pmatrix}$$

8.2.1 Matrix Algebra

You have probably guessed then how to add 2 $(m \times n)$ matrices - term by term! *Note* as we saw beforehand with vectors, it is totally meaningless to add matrices that are of different dimensions.

Definition 62 (Addition of Matrices)**.** *If*

$$\underset{m\times n}{\boldsymbol{A}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} = [\alpha_{ij}]$$

and

$$\underset{m\times n}{\boldsymbol{B}} = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & & \vdots \\ \beta_{m1} & \beta_{n2} & \cdots & \beta_{mn} \end{pmatrix} = [\beta_{ij}]$$

then

$$\underset{m\times n}{\boldsymbol{A}} + \underset{m\times n}{\boldsymbol{B}} = \underset{m\times n}{\boldsymbol{D}} = \begin{pmatrix} \alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \cdots & \alpha_{1n}+\beta_{1n} \\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \cdots & \alpha_{2n}+\beta_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1}+\beta_{m1} & \alpha_{m2}+\beta_{m2} & \cdots & \alpha_{mn}+\beta_{mn} \end{pmatrix} = [\delta_{ij}] = [\alpha_{ij}+\beta_{ij}]$$

$$\underbrace{\underset{m\times n}{\boldsymbol{A}} + \underset{m\times n}{\boldsymbol{B}}}_{m\times n} = \begin{pmatrix} \alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \cdots & \alpha_{1n}+\beta_{1n} \\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \cdots & \alpha_{2n}+\beta_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1}+\beta_{m1} & \alpha_{m2}+\beta_{m2} & \cdots & \alpha_{mn}+\beta_{mn} \end{pmatrix} = [\alpha_{ij}+\beta_{ij}]$$

Definition 63 (Multiplication of Matrices)**.** *If* $\underset{m\times k}{\boldsymbol{A}}$ *and* $\underset{k\times n}{\boldsymbol{B}}$ *are given, then we define*

$$\underset{m\times k}{\boldsymbol{A}} \cdot \underset{k\times n}{\boldsymbol{B}} = \underset{m\times n}{\boldsymbol{C}} = [c_{ij}]$$

such that

$$c_{ij} \equiv \sum_{l=1}^{k} a_{il} b_{lj}$$

so note above that the only index being summed over is l.

The above expression may look quite daunting if you have never seen summation signs before so a simple example should help to clarify.

Example 21. *Let*

$$\underset{2\times 3}{\boldsymbol{A}} = \begin{pmatrix} 0 & 1 & 5 \\ 6 & 0 & 2 \end{pmatrix}$$

and

$$\underset{3\times 2}{\boldsymbol{B}} = \begin{pmatrix} 0 & 3 \\ 1 & 0 \\ 2 & 3 \end{pmatrix}$$

Then

$$\underbrace{\underset{2\times3}{\boldsymbol{A}} \cdot \underset{3\times2}{\boldsymbol{B}}}_{2\times2} = \begin{pmatrix} 0 & 1 & 5 \\ 6 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 3 \\ 1 & 0 \\ 2 & 3 \end{pmatrix}$$
$$= \begin{pmatrix} (0\times0)+(1\times1)+(5\times2), & (0\times3)+(1\times0)+(5\times3) \\ (6\times0)+(0\times1)+(2\times2), & (6\times3)+(0\times0)+(2\times3) \end{pmatrix}$$
$$= \begin{pmatrix} 11 & 15 \\ 4 & 24 \end{pmatrix}$$

Note that $\underset{2\times3}{\mathbf{A}} \cdot \underset{4\times5}{\mathbf{B}}$ is meaningless. The second dimension of **A** must be equal to the first dimension of **B**.

Note further that this brings up the very important point that matrices do not multiply like regular numbers. They are **NOT** commutative i.e.

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$$

For example

$$\underset{2\times3}{\mathbf{A}} \cdot \underset{3\times4}{\mathbf{B}} \neq \underset{3\times4}{\mathbf{B}} \cdot \underset{2\times3}{\mathbf{A}}$$

in fact, not only does the LHS not equal the RHS, the RHS does not even exist. We will see later that one interpretation of a matrix is as a representation of a linear function. With that interpretation, matrix multiplication takes on a specific meaning and there will be another way to think about why you can only multiply certain "conformable" pairs of matrices.
The point of all this is to ensure that you are very careful with the dimensions when multiplying matrices!

Definition 64. *Any matrix which has the same number of rows as columns is known as a square matrix, and is denoted* $\underset{n\times n}{\boldsymbol{A}}$.

For example the matrix

$$\underset{3\times3}{\mathbf{A}} = \begin{pmatrix} 0 & 1 & 5 \\ 6 & 0 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

is a square (3×3) matrix.

Definition 65. *There is a special square matrix known as the* identity matrix *(which is likened to the number 1 (the multiplicative identity) from Definition 14), in that any matrix multiplied by this* identity matrix *gives back the original matrix. The Identity matrix is denoted* $\boldsymbol{I}_n$ *and is equal to*

$$\underset{n\times n}{\boldsymbol{I}_n} == \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

Definition 66. *A square matrix is called a* diagonal matrix *if* $a_{ij} = 0$ *whenever* $i \neq j$.[1]

Definition 67. *A square matrix is called an* upper triangular matrix *(resp.* lower triangular *if* $a_{ij} = 0$ *whenever* $i > j$ *(resp.* $i < j$*).*

Diagonal matrices are easy to deal with. Triangular matrixes are also somewhat tractable. You'll see that for many applications you can replace an arbitrary square matrix with a related diagonal matrix.

For any matrix $\underset{m\times n}{\mathbf{A}}$ we have the results that

$$\underset{m\times n}{\mathbf{A}} \cdot \mathbf{I}_n = \underset{m\times n}{\mathbf{A}}$$

and

$$\mathbf{I}_m \cdot \underset{m\times n}{\mathbf{A}} = \underset{m\times n}{\mathbf{A}}$$

Note that unlike normal algebra it is not the same matrix which multiplies $\underset{m\times n}{\mathbf{A}}$ on both sides to give back $\underset{m\times n}{\mathbf{A}}$ (unless $n = m$).

Definition 68. *We say a matrix* $\underset{n\times n}{\boldsymbol{A}}$ *is* invertible *or* non-singular *if* $\exists\ \underset{n\times n}{\boldsymbol{B}}$ *such that*

$$\underbrace{\underset{n\times n}{\boldsymbol{A}} \cdot \underset{n\times n}{\boldsymbol{B}}}_{n\times n} = \underbrace{\underset{n\times n}{\boldsymbol{B}} \cdot \underset{n\times n}{\boldsymbol{A}}}_{n\times n} = \boldsymbol{I}_n$$

If $\boldsymbol{A}$ *is invertible, we denote it's inverse as* $\boldsymbol{A}^{-1}$.
So we get

$$\underbrace{\underset{(n\times n)}{\boldsymbol{A}} \cdot \underset{(n\times n)}{\boldsymbol{A}^{-1}}}_{n\times n} = \underbrace{\underset{n\times n}{\boldsymbol{A}^{-1}} \cdot \underset{n\times n}{\boldsymbol{A}}}_{n\times n} = \boldsymbol{I}_n$$

[1] The main diagonal is always defined as the diagonal going from top left corner to bottom right corner i.e. $\searrow$

A square matrix that is not invertible is called *singular*. *Note* that this only applies to square matrices[2].
Note We will see how to calculate inverses soon.

Definition 69. *The* determinant *of a matrix* $\boldsymbol{A}$ *(written* $det\boldsymbol{A} = |\boldsymbol{A}|$*) is defined inductively.*

$n = 1$ $\underset{(1\times 1)}{\boldsymbol{A}}$

$$det\boldsymbol{A} = |\boldsymbol{A}| \equiv a_{11}$$

$n \geq 2$ $\underset{(n\times n)}{\boldsymbol{A}}$

$$det\boldsymbol{A} = |\boldsymbol{A}| \equiv a_{11}\,|\boldsymbol{A}_{-11}| - a_{12}\,|\boldsymbol{A}_{-12}| + a_{13}\,|\boldsymbol{A}_{-13}| - \cdots \pm a_{1n}\,|\boldsymbol{A}_{-1n}|$$

where $\boldsymbol{A}_{-1j}$ *is the matrix formed by deleting the first row and jth column of* $\boldsymbol{A}$*.*

Note $\boldsymbol{A}_{-1j}$ *is an* $(n-1)\times(n-1)$ *dimensional matrix.*

Example 22. *If*

$$\underset{2\times 2}{\boldsymbol{A}} = [a_{ij}] = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$
$$\Longrightarrow |\boldsymbol{A}| = a_{11}a_{22} - a_{12}a_{21}$$

Example 23. *If*

$$\underset{3\times 3}{\boldsymbol{A}} = [a_{ij}] = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$\Longrightarrow |\boldsymbol{A}| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

The determinant is useful primarily because of the following result:

[2] You can find one-sided "pseudo inverses" for all matrices, even those that are not square.

Theorem 44. *A matrix is invertible if and only if its determinant* $\neq 0$.

Definition 70. *The Inverse of a matrix* $\underset{n\times n}{\boldsymbol{A}}$ *is defined as*

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|} \cdot adj\boldsymbol{A}$$

*where adj**A** is the* adjoint *of **A** and we will not show how to calculate it here.*

Example 24. *If **A** is a* (2×2) *matrix and invertible then*

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|} \cdot \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Note It is worthwhile memorizing the formula for the inverse of a 2×2 matrix. Leave inverting higher-order matrices to computers.

8.2.2 Inner Product and Distance

Definition 71 (Inner Product)**.** *If* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{M}^{n\times 1}$*, then the* inner product *(or dot product or scalar product) is given by*

$$\begin{aligned} \boldsymbol{x}^t\boldsymbol{y} &= x_1y_1 + x_2y_2 + \cdots + x_ny_x \\ &= \sum_{i=1}^{n} x_iy_i \end{aligned}$$

Note that $\mathbf{x}^t\mathbf{y} = \mathbf{y}^t\mathbf{x}$. We will have reason to use this concept when we do calculus, and will write $x \cdot y = \sum_{i=1}^{n} x_iy_i$.

Definition 72 (Distance)**.** *We generalize the notion of distance.*
For $\mathbb{R}^n$*, a function*

$$d\colon \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

*is called a metric if for any two points **x** and* $\boldsymbol{y} \in \mathbb{R}^n$

1. $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$
2. $d(\boldsymbol{x}, \boldsymbol{y}) = 0 \Longleftrightarrow \boldsymbol{x} = \boldsymbol{y}$

3. $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x})$

4. $d(\boldsymbol{x},\boldsymbol{y}) \leq d(\boldsymbol{x},\boldsymbol{z}) + d(\boldsymbol{z},\boldsymbol{y})$, *for any* $\boldsymbol{z} \in \mathbb{R}^n$

The last of these properties is called the triangle inequality.
If you think about an example on a map (e.g. in $\mathbb{R}^2$*), all this is saying is that it is a shorter distance to walk in a straight line from* $\boldsymbol{x}$ *to* $\boldsymbol{y}$*, than it is to walk from* $\boldsymbol{x}$ *to* $\boldsymbol{z}$ *and then from* $\boldsymbol{z}$ *to* $\boldsymbol{y}$*.*

A metric is a generalized distance function. It is possible to do calculus on abstract spaces that have a metric (they are called metric spaces). Usually there are many possible ways to define a metric. Even in $\mathbb{R}^n$ there are many possibilities:

Example 25.

$$d(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 1, & \text{if } \boldsymbol{x} \neq \boldsymbol{y}, \\ 0, & \text{if } \boldsymbol{x} = \boldsymbol{y}. \end{cases}$$

This satisfies the definition of a metric. It basically tells you whether $\mathbf{x} = \mathbf{y}$. The metric defined by

Example 26.

$$d(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1,\ldots,n} |x_i - y_i|$$

states that the distance between two points is the length of the path connecting the two points using segments parallel to the coordinate axes.

We will be satisfied with the standard *Euclidean metric*:

Example 27.

$$d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$$

where

$$\begin{aligned} \|\boldsymbol{z}\| &= \sqrt{z_1^2 + z_2^2 + \cdots + z_n^2} \\ &= \sqrt{\sum_{i=1}^{n} z_i^2} \end{aligned}$$

Under the Euclidean metric, the distance between two points is the length of the line segment connecting the points. We call $\|\mathbf{z}\|$, which is the distance between $\mathbf{0}$ and $\mathbf{z}$ the *norm* of $\mathbf{z}$.

Notice that $\|\mathbf{z}\|^2 = \mathbf{z} \cdot \mathbf{z}$.

When $\mathbf{x}\cdot\mathbf{y} = 0$ we say that $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal/at right angles/perpendicular*.

It is a surprising geometric property that two vectors are perpendicular if and only if their inner product is zero. This fact follows rather easily from "The Law of Cosines." The law of cosines states that if a triangle has sides A,B, and C and the angle θ opposite the side c, then

$$c^2 = a^2 + b^2 - 2ab\cos(\theta),$$

where a, b, and c are the lengths of A, B, and C respectively. This means that:

$$(\mathbf{x}-\mathbf{y})\cdot(\mathbf{x}-\mathbf{y}) = \mathbf{x}\cdot\mathbf{x} + \mathbf{y}\cdot\mathbf{y} - 2\left\|\mathbf{x}\right\|\left\|\mathbf{x}\right\|\cos(\theta),$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. If you multiply everything out you get the identity:

$$\left\|\mathbf{x}\right\|\left\|\mathbf{x}\right\|\cos(\theta) = \mathbf{x}^t\mathbf{y}. \tag{8.1}$$

Equation (8.1) has two nice consequences. First, it justifies the use of the term orthogonal: The inner product of two non-zero vectors is zero if and only if the cosine of the angle between them is zero. Second, it gives you an upper bound of the inner product (because the absolute value of the cosine is less than or equal to one):

$$\left\|\mathbf{x}\right\|\left\|\mathbf{x}\right\| \geq \left|\mathbf{x}^t\mathbf{y}\right|.$$

8.3 Systems of Linear Equations

Consider the system of n equations in m variables:

$$\begin{aligned}
y_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n \\
y_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n \\
&\vdots \\
y_i &= \alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n \\
&\vdots \\
y_m &= \alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n
\end{aligned}$$

Here the variables are the x_j. This can be written as

$$\underset{(m\times 1)}{\mathbf{y}} = \underbrace{\underset{(m\times n)}{\mathbf{A}} \cdot \underset{(n\times 1)}{\mathbf{x}}}_{(m\times 1)}$$

where

$$\underset{(m\times 1)}{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \quad , \quad \underset{(n\times 1)}{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\underset{m\times n}{\mathbf{A}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} = [\alpha_{ij}]$$

or, putting it all together

$$\underset{(m\times 1)}{\mathbf{y}} = \underbrace{\underset{(m\times n)}{\mathbf{A}} \cdot \underset{(n\times 1)}{\mathbf{x}}}_{(m\times 1)}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Note you should convince yourself that if you multiply out the RHS of the above equation and then compare corresponding entries of the new $(m\times 1)$ vectors that the result is equivalent to the original system of equations. The value of matrices is that they permit you to write the complicated system of equations in a simple form. Once you have written them a system of equations in this way, you can use matrix operations to solve some systems.

Example 28. *In high school you probably solved equations of the form:*

$$\begin{aligned} 3x_1 - 2x_2 &= 7 \\ 8x_1 + x_2 &= 25 \end{aligned}$$

Well matrix algebra is just a clever way to solve these in one go.
So here we have that

$$\underset{(2\times 2)}{\boldsymbol{A}} = \begin{pmatrix} 3 & -2 \\ 8 & 1 \end{pmatrix} \quad , \quad \underset{(2\times 1)}{\boldsymbol{x}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad , \quad \underset{(2\times 1)}{\boldsymbol{y}} = \begin{pmatrix} 7 \\ 25 \end{pmatrix}$$

And can write this as

$$\underset{(2\times 2)}{\boldsymbol{A}} \cdot \underset{(2\times 1)}{\boldsymbol{x}} = \underset{(2\times 1)}{\boldsymbol{y}}$$

And obviously we want to solve for $\boldsymbol{x}$.
So we multiply both sides of the equation on the left (recall that it does matter what side of the equation you multiply) by $\boldsymbol{A}^{-1}$,

$$\underbrace{\boldsymbol{A}^{-1}\boldsymbol{A}}_{\boldsymbol{I}_2}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y}$$

$$\Longleftrightarrow \boldsymbol{I}_2\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y}$$

$$\Longleftrightarrow \boldsymbol{x} = \underbrace{\underset{(2\times 2)}{\boldsymbol{A}^{-1}} \cdot \underset{(2\times 1)}{\boldsymbol{y}}}_{(2\times 1)}$$

and we have that

$$\begin{aligned}
\boldsymbol{A}^{-1} &= \frac{1}{|\boldsymbol{A}|} \cdot \begin{pmatrix} 1 & 2 \\ -8 & 3 \end{pmatrix} \\
&= \frac{1}{|(3\times 1) - (-2\times 8)|} \cdot \begin{pmatrix} 1 & 2 \\ -8 & 3 \end{pmatrix} \\
&= \frac{1}{|19|} \cdot \begin{pmatrix} 1 & 2 \\ -8 & 3 \end{pmatrix}
\end{aligned}$$

So

$$\begin{aligned}
\boldsymbol{x} &= \boldsymbol{A}^{-1}\boldsymbol{y} \\
&= \frac{1}{|19|} \cdot \begin{pmatrix} 1 & 2 \\ -8 & 3 \end{pmatrix} \cdot \begin{pmatrix} 7 \\ 25 \end{pmatrix} \\
&= \frac{1}{19} \cdot \begin{pmatrix} 57 \\ 19 \end{pmatrix} \\
&= \begin{pmatrix} 3 \\ 1 \end{pmatrix}
\end{aligned}$$

The example illustrates some general properties. If you have exactly as many equations as unknowns (so that the matrix **A** is square, then the system has a

unique solution if and only if $\mathbf{A}$ is invertible. If $\mathbf{A}$ is invertible, it is obvious that the solution is unique (and given by the formula: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$). If $\mathbf{A}$ is not invertible, it is the case that there is a nonzero $\mathbf{z}$ such that $\mathbf{Az} = \mathbf{0}$. This means that if you can find one solution to $\mathbf{Ax} = \mathbf{y}$,[3] then you can find infinitely many solutions (by adding arbitrary multiples of $\mathbf{z}$ to the original solution. Intuitively, it is hard for a matrix to be singular, so "most" of the time systems of n equations and n unknowns have unique solutions.

These comments do not tell you about situations where the number of equations is not equal to the number of unknowns. In most cases, when there are extra equations, a system of linear equations will have no solutions. If there are more unknowns than equations, typically the system will have infinitely many solutions (it is possible for the system to have no solutions, but it is not possible for the system to have a unique solution).

A system of equations of the form $\mathbf{Ax} = \mathbf{0}$ is called a *homogeneous* system of equations. Such a system always has a solution ($\mathbf{x} = \mathbf{0}$). The solution will not be unique if there are more unknowns than equations.

The standard way to establish these results is by applying "elementary row operations" to $\mathbf{A}$ to transform a system of equations into an equivalent system that is easier to analyze.

8.4 Linear Algebra: Main Theory

Euclidean Spaces (and Vector Spaces more generally) have a nice structure that permits you to write all elements in terms of a fixed set of elements. In order to describe this theory, we need a few definitions.

A *linear combination* of a collection of vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ is a vector of the form $\sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ for some scalars $\lambda_1, \ldots, \lambda_k$.

$\mathbb{R}^n$ has the property that sums and scalar multiples of elements of $\mathbb{R}^n$ remain in the set. Hence if we are given some elements of the set, all linear combinations will also be in the set. Some subsets are special because they contain no redundancies:

Definition 73. *A collection of vectors* $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ *is* linearly independent *if* $\sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i = \boldsymbol{0}$ *if and only if* $\lambda_i = 0$ *for all* i.

Here is the way in which linear independence captures the idea of no redundancies:

[3] When $\mathbf{A}$ is singular, there is no guarantee that a solution to this equation exists.

Theorem 45. *If $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ is a linearly independent collection of vectors and $\boldsymbol{z} \in S(X)$, then there are unique $\lambda_1, \ldots, \lambda_k$ such that $\boldsymbol{z} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$.*

Proof. Existence follows from the definition of span. Suppose that there are two linear combinations that of the elements of X that yield $\mathbf{z}$ so that

$$\mathbf{z} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i$$

and

$$\mathbf{z} = \sum_{i=1}^{k} \lambda_i' \mathbf{x}_i.$$

Subtract the equations to obtain:

$$\mathbf{0} = \sum_{i=1}^{k} (\lambda_i' - \lambda_i)\, \mathbf{x}_i.$$

By linear independence, $\lambda_i = \lambda_i'$ for all i, the desired result. □

Next, let us investigate the set of things that can be described by a collection of vectors.

Definition 74. *The* span *of a collection of vectors $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ is $S(X) = \{\boldsymbol{y} : \boldsymbol{y} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$ for some scalars $\lambda_1, \ldots, \lambda_k\}$.*

$S(X)$ is the smallest vector space containing all of the vectors in X.

Definition 75. *The* dimension *of a vector space is N, where N is the smallest number of vectors needed to span the space.*

We deal only with finite dimensional vector spaces. We'll see this definition agrees with the intuitive notion of dimension. In particular, $\mathbb{R}^n$ has dimension n.

Definition 76. *A* basis *for a vector span V is any collection of linearly independent vectors than span V.*

Theorem 46. *If $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ is a set of linearly independent vectors that does not span V, then there exists $\boldsymbol{v} \in V$ such that $X \cup \{\boldsymbol{v}\}$ is linearly independent.*

Proof. Take $\mathbf{v} \in V$ such that $\mathbf{v} \neq \mathbf{0}$ and $\mathbf{v} \notin S(X)$. $X \cup \{\mathbf{v}\}$ is a linearly independent set. To see this, suppose that there exists λ_i $i = 0, \ldots, k$ such that at least one $\lambda_i \neq 0$ and

$$\lambda_0 \mathbf{v} + \sum_{i=1}^{k} \lambda_i \mathbf{x}_i. \tag{8.2}$$

If $\lambda_0 = 0$, then X are not linearly independent. If $\lambda_0 \neq 0$, then equation (8.2) can be rewritten

$$\mathbf{v} = \sum_{i=1}^{k} \frac{\lambda_i}{\lambda_0} \mathbf{x}_i.$$

In either case, we have a contradiction. □

Definition 77. *The* standard basis *for $\mathbb{R}^n$ consists of the set of N vectors e_i, $i = 1, \ldots, N$, where e_i is the vector with component 1 in the ith position and zero in all other positions.*

You should check that the standard basis really is a linearly independent set that spans $\mathbb{R}^n$. Also notice that the elements of the standard basis are mutually orthogonal. When this happens, we say that the basis is *orthogonal*. It is also the case that each basis element has unit length. When this also happens, we say that the basis is *orthonormal*. It is always possible to find an orthonormal basis.[4] They are particularly useful because it is easy to figure out how to express an arbitrary element of X in terms of the basis.

It follows from these observations that each vector $\mathbf{v}$ has a unique representation in terms of the basis, where the representation consists of the λ_i used in the linear combination that expresses $\mathbf{v}$ in terms of the basis. For the standard basis, this representation is just the components of the vector.

It is not hard (but a bit tedious) to prove that all bases have the same number of elements. (This follows from the observation that any system of n homogeneous equations and $m > n$ unknowns has a non-trivial solution, which in turn follows from "row-reduction" arguments.)

8.5 Eigenvectors and Eigenvalues

An *eigenvalue* of the square matrix $\mathbf{A}$ is a number λ with the property $\mathbf{A} - \lambda\mathbf{I}$ is singular. If λ is an eigenvalue of $\mathbf{A}$, then any $\mathbf{x} \neq \mathbf{0}$ such that $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ is called an *eigenvector* of $\mathbf{A}$ associated with the eigenvalue λ.

[4] We have exhibited an orthonormal basis for $\mathbb{R}^n$. It is possible to construct an orthonormal basis for any vector space.

Eigenvalues are those values for which the equation

$$\mathbf{Ax} = \lambda\mathbf{x}$$

has a non-zero solution. You can compute eigenvalues (in theory) by solving the equation $\det(\mathbf{A}-\lambda\mathbf{I}) = 0$. If $\mathbf{A}$ is an $n\times n$ matrix, then this *characteristic equation* is a polynomial equation of degree n. By the Fundamental Theorem of Algebra, it will have n (not necessarily distinct and not necessarily real) roots. That is, the characteristic polynomial can be written

$$P(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k},$$

where $r_1, r_2 \ldots, r_k$ are the distinct roots ($r_i \neq r_j$ when $i \neq j$) and m_i are positive integers summing to n. We call m_i the *multiplicity* of root r_i. Eigenvalues and their corresponding eigenvectors are important because they enable one to relate complicated matrices to simple ones.

Theorem 47. *If $\mathbf{A}$ is an $n \times n$ matrix that has n distinct eigen-values or is symmetric, then there exists an invertible $n \times n$ matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$. Moveover, the diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$ and the columns of $\mathbf{P}$ are the corresponding eigenvectors.*

The theorem says that if $\mathbf{A}$ satisfies certain conditions, then it is "related" to a diagonal matrix. It also tells you how to find the diagonal matrix. The relationship $\mathbf{A} = \mathbf{PDP}^{-1}$ is quite useful. For example, it follows from the relationship that $\mathbf{A}^k = \mathbf{PD}^k\mathbf{P}^{-1}$. It is much easier to raise a diagonal matrix to a power than find a power of a general matrix.

Proof. Suppose that λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ is an eigenvector. This means that $\mathbf{Ax} = \lambda\mathbf{x}$. If $\mathbf{P}$ is a matrix with column j equal to an eigenvector associated with λ_i, it follows that $\mathbf{AP} = \mathbf{PD}$. The theorem would follow if we could guarantee that $\mathbf{P}$ is invertible.

When A is symmetric, one can prove that A has only real eigenvalues and that one can find n linearly independent eigenvectors even if the eigenvalues are not distinct. This result is elementary (but uses some basic facts about complex numbers), but we omit the proof.

In general, one can prove that the eigenvectors of distinct eigenvalues are distinct. To see this, suppose that $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues and $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are associated eigenvectors. In order to reach a contradiction, suppose that the vectors are linearly dependent. Without loss of generality, we may assume that

$\{\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}\}$ are linearly independent, but that $\mathbf{x}_k$ can be written as a linear combination of the first $k-1$ vectors. This means that there exists α_i $i = 1, \ldots, k-1$ not all zero such that:

$$\sum_{i=1}^{k-1} \alpha_i \mathbf{x}_i = \mathbf{x}_k. \tag{8.3}$$

Multiply both sides of equation (8.3) by $\mathbf{A}$ and use the eigenvalue property to obtain:

$$\sum_{i=1}^{k-1} \alpha_i \lambda_i \mathbf{x}_i = \lambda_k \mathbf{x}_k. \tag{8.4}$$

Multiply equation (8.3) by λ_k and subtract it from equation (8.4) to obtain:

$$\sum_{i=1}^{k-1} \alpha_i (\lambda_i - \lambda_k) \mathbf{x}_i = \mathbf{0}. \tag{8.5}$$

Since the eigenvalues are distinct, equation (8.5) gives a non-trivial linear combination of the first $k-1$ $\mathbf{x}_i$ that is equal to $\mathbf{0}$, which contradicts linear independence. □

Here are some useful facts about determinants and eigenvalues. (The proofs range from obvious to tedious.)

1. det $\mathbf{AB}$ = det $\mathbf{BA}$

2. If $\mathbf{D}$ is a diagonal matrix, then det $\mathbf{D}$ is equal to the product of its diagonal elements.

3. det $\mathbf{A}$ is equal to the product of the eigenvalues of $\mathbf{A}$.

4. The *trace* of a square matrix $\mathbf{A}$ is equal to the sum of the diagonal elements of $\mathbf{A}$. That is, tr $(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$. Fact: tr $(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$, where λ_i is the ith eigenvalue of $\mathbf{A}$ (eigenvalues counted with multiplicity).

One variation on the symmetric case is particularly useful in the next section. When $\mathbf{A}$ is symmetric, then we take take the eigenvectors of $\mathbf{A}$ to be orthonormal. In this case, the $\mathbf{P}$ in the previous theorem has the property that $\mathbf{P}^{-1} = \mathbf{P}^t$. Eigenvalues turn out to be important in many different places. They play a role in the study of stability of difference and differential equations. They make certain computations easy. They make it possible to define a sense in which matrices can be positive and negative that allows us to generalize the one-variable second-order conditions. The next topic will do this.

8.6 Quadratic Forms

Definition 78. *A* quadratic form *in n variables is any function $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ that can be written $Q(\boldsymbol{x}) = \boldsymbol{x}^t\boldsymbol{A}\boldsymbol{x}$ where $\boldsymbol{A}$ is a symmetric $n \times n$ matrix.*

When $n = 1$ a quadratic form is a function of the form ax^2. When $n = 2$ it is a function of the form $a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2$ (remember $a_{12} = a_{21}$). When $n = 3$, it is a function of the form $a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2+2a_{13}x_1x_3+2a_{23}x_2x_3$. A quadratic form is second-degree polynomial that has no constant term.

We will see soon that the second derivative of a real-valued function on $\mathbb{R}^n$ is not a single function, but a collection of n^2 functions. Quadratic forms will be the way in which we study second derivatives. In particular, they are important for checking second-order conditions and concavity of functions.

Definition 79. *A quadratic form $Q(\boldsymbol{x})$ is*

1. positive definite *if $Q(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq 0$.*
2. positive semi definite *if $Q(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$.*
3. negative definite *if $Q(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \neq 0$.*
4. negative semi definite *if $Q(\boldsymbol{x}) \leq 0$ for all $\boldsymbol{x}$.*
5. indefinite *if there exists $\boldsymbol{x}$ and $\boldsymbol{y}$ such that $Q(\boldsymbol{x}) > 0 > Q(\boldsymbol{y})$.*

The definition provides a notion of positivity and negativity for matrices. You should check to confirm that the definitions coincide with ordinary notions of positive and negative when Q is a quadratic form of one variable (that is $Q(x) = Ax^2$). In this case the quadratic form is positive definite in $A > 0$, negative (and positive) semi-definite when $A = 0$ and negative definite when $a < 0$. When $n > 1$ it is not hard to find indefinite matrices. We will see this soon.

If **A** happened to be a diagonal matrix, then it is easy to classify the associated quadratic form according to the definition. $Q(\mathbf{x}) = \mathbf{x}^t\mathbf{A}\mathbf{x} = \sum_{i=1}^n a_{ii}x_i^2$. This quadratic form is positive definite if and only if all of the $a_{ii} > 0$, negative definite if and only if all of the $a_{ii} < 0$, positive semi definite if and only if $a_{ii} \geq 0$, for all i negative semi definite if and only if $a_{ii} \leq 0$ for all i, and indefinite if **A** has both negative and positive diagonal entries.

The theory of diagonalization gives us a way to translate use these results for all matrices. We know that if **A** is a symmetric matrix, then it can be written **A** = $\mathbf{R}^t\mathbf{D}\mathbf{R}$, where **D** is a diagonal matrix with (real) eigenvalues down the diagonal and **R** is an orthogonal matrix. This means that the quadratic form: $Q(\mathbf{x}) = \mathbf{x}^t\mathbf{A}\mathbf{x} = \mathbf{x}^t\mathbf{R}^t\mathbf{D}\mathbf{R}\mathbf{x} = (\mathbf{R}\mathbf{x})^t\,\mathbf{D}\,(\mathbf{R}\mathbf{x})$. This expression is useful because it means

that the definiteness of **A** is equivalent to the definiteness of its diagonal matrix of eigenvectors, **D**. (Notice that if I can find an **x** such that $\mathbf{x}^t\mathbf{A}\mathbf{x} > 0$, then I can find an **x** such that $\mathbf{y}^t\mathbf{D}\mathbf{y}$ ($\mathbf{y} = \mathbf{R}\mathbf{x}$) and conversely.)

Theorem 48. *The quadratic form $Q(\boldsymbol{x}) = \boldsymbol{x}^t\boldsymbol{A}\boldsymbol{x}$ is*

1. positive definite *if $\lambda_i > 0$ for all i.*
2. positive semi definite *if $\lambda_i \geq 0$ for all i.*
3. negative definite *if $\lambda_i < 0$ for all i.*
4. negative semi definite *if $\lambda_i \leq 0$ for all i.*
5. indefinite *if there exists j and k such that $\lambda_j > 0 > \lambda_k$.*

There is a computational trick that often allows you to identify definiteness without computing eigenvalue.

Definition 80. *A* principal submatrix *of a square matrix **A** is the matrix obtained by deleting any k rows and the corresponding k columns. The determinant of a principal submatrix is called the* principal minor *of **A**. The* leading principal submatrix of order k *of an $n \times n$ matrix is obtained by deleting the last $n - k$ rows and column of the matrix. The determinant of a leading principal submatrix is called the* leading principal minor *of **A**.*

Theorem 49. *A matrix is*

1. *positive definite if and only if all of its leading principal minors are positive.*
2. *negative definite if and only if its odd principal minors and negative and its even principal minors are positive.*
3. *indefinite if one of its kth order leading principal minors is negative for an even k or if there are two odd leading principal minors that have different signs.*

The theorem permits you to classify the definiteness of matrices without finding eigenvalues. It may seem strange at first, but you can remember it by thinking about diagonal matrices.

Chapter 9

Multivariable Calculus

The goal is to extend the calculus from real-valued functions of a real variable to general functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. Some ideas generalize easily, but going from one dimensional domains to many dimensional domains raises new issues that need discussion. Raising the dimension of the range space, on the other hand, raises no new conceptual issues. Consequently, we begin our discussion to real-valued functions. We will explicitly consider higher dimensional ranges only when convenient or necessary. (It will be convenient to talk about linear functions in general terms. It is necessary to talk about the most interesting generalization of the Chain Rule and for the discussions of inverse and implicit functions.)

9.1 Linear Structures

To do calculus, you need to understand linear "objects" in $\mathbb{R}^n$. In the plane there are three kinds of subspace (recall that a subspace is a set L with the property that if x and y are in the set, then so are $x + y$ and λx for all λ). These sets are: the entire plane, any line through the origin, and the origin. In higher dimensional spaces, there are more linear subsets.

It turns out that the most useful to us are one dimensional linear subsets, lines, and $n - 1$ dimensional subsets, hyperplanes.

Definition 81. *A* line *is described by a point x and a direction v. It can be represented as $\{z :$ there exists $t \in \mathbb{R}$ such that $z = x + tv\}$.*

If we constrain $t \in [0, 1]$ in the definition, then the set is the line segment connecting x to $x + v$. Two points still determine a line: The line connecting x to y can be viewed as the line containing x in the direction v. You should check that this is the same as the line through y in the direction v.

Definition 82. *A* hyperplane *is described by a point x_0 and a* normal *direction $p \in \mathbb{R}^n$, $p \neq 0$. It can be represented as $\{z : p \cdot (z - x_0) = 0\}$. p is called the normal direction of the plane.*

The interpretation of the definition is that a hyperplane consists of all of the z with the property that the direction $z - x_0$ is normal to p.

In $\mathbb{R}^2$ lines are hyperplanes. In $\mathbb{R}^3$ hyperplanes are "ordinary" planes.

Lines and hyperplanes are two kinds of "flat" subset of $\mathbb{R}^n$. Lines are subsets of dimension one. Hyperplanes are subsets of dimension $n - 1$ or *co-dimension* one. You can have a flat subsets of any dimension less than n. Although, in general, lines and hyperplanes are not subspaces (because they do not contain the origin) you obtain these sets by "translating" a subspace that is, by adding the same constant to all of its elements.

Definition 83. *A* linear manifold *of $\mathbb{R}^n$ is a set S such that there is a subspace V on $\mathbb{R}^n$ and $x_0 \in \mathbb{R}^n$ with $S = V + \{x_0\}$.*

In the above definition, $V + \{x_0\} \equiv \{y : y = v + x_0 \text{ for some } v \in V\}$.

The definition is a bit pedantic. Officially lines and hyperplanes are linear manifolds and not linear subspaces.

It is worthwhile reviewing the concepts of line and hyperplane. Given two points x and y, you can construct a line that passes through the points. This line is

$$\{z : z = x + t(y - x) \text{ for some } t.\}$$

This formulation is somewhat different from the one normally sees, but it is equivalent. Writing out the two-dimensional version yields:

$$z_1 = x_1 + t(y_1 - x_1) \text{ and } z_2 = x_2 + t(y_2 - x_2).$$

If you use the equation for z_1 to solve for t and substitute out you get:

$$z_2 = x_2 + \frac{(y_2 - x_2)\,(z_1 - x_1)}{y_1 - x_1}$$

or

$$z_2 - x_2 = \frac{y_2 - x_2}{y_1 - x_1}\,(z_1 - x_1)\,,$$

which is the standard way to represent the equation of a line (in the plane) through the point (x_1, x_2) with slope $(y_2 - x_2)(y_1 - x_1)$. This means that the "parametric" representation is essentially equivalent to the standard representation in $\mathbb{R}^2$.[1] The

[1]The parametric representation is actually a bit more general, since it allows you to describe lines that are parallel to the vertical axis. Because these lines have infinite slope, they cannot be represented in standard form.

familiar ways to represent lines do not work in higher dimensions. The reason is that one linear equation in $\mathbb{R}^n$ typically has an $n-1$ dimensional solution set, so it is a good way to describe a one dimensional set only if $n = 2$.

You need two pieces of information to describe a line. If the information consists of a point and a direction, then the parametric version of the line is immediately available. If the information consists of two points, then you form a direction by subtracting one point from the other (the order is not important).

You can describe a hyperplane easily given a point and a (normal) direction. Note that the direction of a line is the direction you follow to stay on the line. The direction for a hyperplane is the direction you follow to go away from the hyperplane. If you are given a point and a normal direction, then you can immediately write the equation for the hyperplane. What other pieces of information determine a hyperplane? In $\mathbb{R}^3$, a hyperplane is just a standard plane. Typically, three points determine a plane (if the three points are all on the same line, then infinitely many planes pass through the points). How can you determine the equation of a plane in $\mathbb{R}^3$ that passes through three given points? A mechanical procedure is to note that the equation for the plane can always be written $Ax_1 + Bx_2 + Cx_3 = D$ and, use the three points to find values for the coefficients. For example, if the points are $(1, 2-3), (0, 1, 1), (2, 1, 1)$, then we can solve:

$$\begin{array}{rcrcrcl} A & + & 2B & - & 3C & = & D \\ & & B & + & C & = & D \\ 2A & + & B & + & C & = & D \end{array}$$

Doing so yields $(A, B, C, D) = (0, .8D, .2D, D)$. (If you find one set of coefficients that work, any non-zero multiple will also work.) Hence an equation for the plane is: $4x_2 + x_3 = 5$ you can check that the three points actually satisfy this equation.

An alternate computation technique is to look for a normal direction. A normal direction is a direction that is orthogonal to **all** directions in the plane. A direction in the plane is a direction of a line in the plane. You can get such a direction by subtracting any two points in the plane. A two dimensional hyperplane will have two independent directions. For this example, one direction can come from the difference between the first two points: $(1, 1, -4)$ and the other can come from the difference between the second and third points $(-2, 0, 0)$ (a third direction will be redundant, but you can do the computation using the direction of the line connecting $(1, 2, -3)$ and $(2, 1, 1)$ instead of either of directions computed above). Once you have two directions, you want to find a normal to both of them. That is, a p such that $p \neq 0$ and $p \cdot (1, 1, -4) = p \cdot (-2, 0, 0) = 0$. This is a system of

two equations and three variables. All multiples of $(0, 4, 1)$ solve the equations.[2] Hence the equation for the hyperplane is $(0, 4, 1) \cdot (x_1 - 1, x_2 - 2, x_3 + 3) = 0$. You can check that this agrees with the equation will found earlier. It also would be equivalent to the equation you would obtain if you used either of the other two given points as "the point on the plane."

9.2 Linear Functions

Definition 84. *A function*

$$L\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

is linear *if and only if*

1. *If for all x and y, $f(x + y) = x + y$ and*
2. *for all scalars λ, $f(\lambda x) = \lambda f(x)$.*

The first condition says that a linear function must be additive. The second condition says that it must have constant returns to scale. The conditions generate several obvious consequences. If L is a linear function, then $L(0) = 0$ and, more generally, $L(x) = -L(-x)$. It is an important observation that any linear function can be "represented" by matrix multiplication. Given a linear function, compute $L(e_i)$, where e_i is the ith standard basis element. Can this a_i and let $\mathbf{A}$ be the square matrix with ith column equal to a_i. Note that $\mathbf{A}$ must have n columns and m rows. Note also that (by the properties of matrix multiplication and linear functions) $L(x) = \mathbf{A}x$ for all x. This means that any linear function can be thought of as matrix multiplication (and the matrix has a column for each dimension in the domain and a row for each dimension in the range). Conversely, every matrix gives rise to a linear function. Hence the identification between linear functions and matrices is perfect. When the linear function is real valued, linear functions can be thought of as taking an inner product.

The identification of linear functions with matrix multiplication supplies another interpretation of matrix multiplication. If $L\colon \mathbb{R}^m \longrightarrow \mathbb{R}^k$ and $M\colon \mathbb{R}^k \longrightarrow \mathbb{R}^n$ are functions, then we can form the composite function $M \circ L\colon \mathbb{R}^m \longrightarrow \mathbb{R}^n$ by $M \circ L(x) = M(L(x))$. If M and L are linear functions, and $\mathbf{A}$ is the matrix that represents M and $\mathbf{B}$ is the matrix that represents L, then $\mathbf{AB}$ is the matrix that represents $M \circ L$. (Officially, you must verify this by checking how the composite function transforms the standard basis elements.)

[2] The "cross product" is a computational tool that allows you to mechanically compute a direction perpendicular to two given directions.

9.3 Representing Functions

Functions of one variable are easy to visualize because we could draw their graphs in the plane. In general, the graph of a function f is $Gr(f) = \{(x, y) : y = f(x)\}$. This means that if $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, then the graph is a subset of $\mathbb{R}^{n+1}$. If $n = 2$, then the graph is a subset of $\mathbb{R}^3$, so someone with a good imagination of a three-dimensional drawing surface could visualize it. If $n > 2$ there is no hope. You can get some intuition by looking at "slices" of the graph obtained by holding the function's value constant.

Definition 85. *A* level set *is the set of points such that the functions achieves the same value. Formally it is defined as the set*

$$\{\boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) = c\} \quad \textit{for any } c \in \mathbb{R}$$

While the graph of the function is a subset of $\mathbb{R}^{n+1}$, the level set (actually, level sets) are subsets of $\mathbb{R}^n$.

Example 29.

$$f(x) = x_1^2 + x_2^2$$

So the function is $\mathbb{R}^2 \longrightarrow \mathbb{R}$ and therefore the graph is in $\mathbb{R}^3$.

The level sets of this function are circles in the plane. (The graph is a cone.) FIGURE GOES HERE

Example 30. *A good example to help understand this is a utility function (of which you will see lots!). A utility function is a function which "measures" a person's happiness. It is usually denoted U. In 200A you will see conditions necessary for the existence of the utility function but for now we will just assume that it exists, and is strictly increasing in each argument. Suppose we have a guy Joel whose utility function is just a function of the number of apples and the number of bananas he eats. So his happiness is determined solely by the number of apples and bananas he eats, and nothing else. Thus we lose no information when we think about utility as a function of two variables:*

$$U\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$$

$U(x_A, x_B)$ where x_A is the number of apples he eats, and x_B is the number of bananas he eats.

A level set is all the different possible combinations of apples and bananas that give him the same utility level *i.e. that leave him equally happy!¡For example Joel might really like apples and only slightly like bananas. So 3 apples and 2 bananas might make him as happy as 1 apple and 10 bananas. In other words he needs lots of bananas to compensate for the loss of the 2 apples.*

If the only functions we dealt with were utility functions, we would call level sets "indifference curves." In economics, typically curves that are "iso-SOMETHING" are level sets of some function.

In fact maybe another example is in order.

Example 31. *Suppose we have a guy Joel who only derives joy from teaching mathematics. Nothing else in the world gives him any pleasure, and as such his utility function U_J is only a function of the number of hours of he spends teaching mathematics H_M. Now we also make the assumption that Joel's utility is strictly increasing in hours spend teaching mathematics, the more he teaches the happier he is*[3]*. So the question is does Joel's utility function U_J have any level sets? Since utility is a function of one variable, Joel's level "sets" are zero dimensional objects – points. Since if his utility function is defined as $U_J(H_M) = H_M^{1/2}$ and Joel teaches for 4 hours (i.e. $H_M = 4$), then $U_J(4) = 4^{1/2} = 2$. So is there any other combination of hours teaching that could leave Joel equally happy? Obviously not, since his utility is only a function of one argument, and it is strictly increasing, no two distinct values can leave him equally happy.*[4]

Example 32. *Recall Example 29*

So plotting level sets of a function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is really just a clever way of representing the information from a 3d graph, but only drawing a 2d graph (much easier to draw!).

A good example of this is a map of a country. When you spread a map out in front of you we know that left-right ($\longleftrightarrow$) represents East-West, and that up-down ($\updownarrow$) represents North-South. So how do we represent mountains and valleys

[3] you may doubt that Joel is really like this but I assure you he is!

[4] So you can see that level sets reference the arguments of a function. And functions with two or more arguments are much more likely to have level sets than functions of one argument, since you can have many different combinations of the arguments.

i.e. points on the earth of different altitude? The answer is with level sets! There are many contour lines drawn all over a map and at each point along these lines the earth is at the same altitude. Of course to be completely rigorous it should also be obvious which direction altitude is increasing in.
So in this case our function would take in coordinates (values for east and west) and spit out the value (altitude) at those coordinates (values).

Definition 86. *We define the* upper contour set *of f as the set*

$$\{\mathbf{x} \in \mathbb{R}^{\mathbf{n}} \mid \mathbf{f}(\boldsymbol{x}) \geq \mathbf{c}\} \quad c \in \mathbb{R}$$

And we define the lower contour set *of f as the set*

$$\{\mathbf{x} \in \mathbb{R}^{\mathbf{n}} \mid \mathbf{f}(\boldsymbol{x}) \leq \mathbf{c}\} \quad c \in \mathbb{R}$$

So referring back to our map example. The upper contour set of a point **x** would be a set of all the coordinates such that if we plugged those coordinates into our altitude function, it would give out a value greater than or equal to the value at the point **x**, i.e. all points that are at a higher altitude than **x**.

9.4 Limits and Continuity

Definition 87 (Limit of a Function)**.**

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$\lim_{\boldsymbol{x} \longrightarrow \boldsymbol{a}} f(\boldsymbol{x}) = c \in \mathbb{R}$$

$\iff \forall\, \epsilon > 0, \exists\, \delta > 0$ *such that*

$$0 < d(\boldsymbol{x}, \boldsymbol{a}) < \delta$$

$$\implies |f(\boldsymbol{x}) - c| < \epsilon$$

This definition agrees with the earlier definition, although there are two twists. First, a general "distance function" replaces absolute values in the condition that says that x is close to a. For our purposes, the distance function will always be the standard Euclidean distance. Second, there we do not define one-sided continuity.

Definition 88 (Continuity of a Function)**.**

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

is called continuous at a point $\boldsymbol{a} \in \mathbb{R}^n$ *if*

$$\lim_{\boldsymbol{x} \longrightarrow \boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a})$$

Again, this definition is a simple generalization of the one-variable definition.

Exercise 2. *Prove that* $f(\boldsymbol{x}) = x_1 x_2$ *is continuous at the point* $(1, 1)$.

We must show that $\forall\, \epsilon > 0, \exists\, \delta > 0$ such that

$$\|(x_1, x_2) - (1, 1)\| < \delta$$

$$\implies |f(x_1, x_2) - 1| < \epsilon$$

Note that

$$\|(x_1, x_2) - (1, 1)\| = \sqrt{(x_1 - 1)^2 + (x_2 - 1)^2}$$

Also

$$\begin{aligned} |f(x_1, x_2) - 1| &= |x_1 x_2 - 1| \\ &= |x_2 x_1 - x_1 + x_1 - 1| \\ &= |x_1(x_2 - 1) + x_1 - 1| \\ &\overset{\triangle}{\leq} \underbrace{|x_1(x_2 - 1)|}_{<\frac{1}{2}\epsilon} + \underbrace{|x_1 - 1|}_{<\frac{1}{2}\epsilon} \\ &< \epsilon \end{aligned}$$

where the second last inequality is got using the triangle inequality hence the $\triangle$ superscipt.

For any given $\epsilon > 0$ let $\delta = \min\left\{\frac{1}{4}\epsilon, 1\right\}$. Then

$$\|(x_1, x_2) - (1, 1)\| < \frac{1}{4}\epsilon$$

$$\implies |x_1 - 1| < \frac{1}{4}\epsilon \qquad \text{and} \qquad |x_2 - 1| < \frac{1}{4}\epsilon$$

Also we have that $x_1 < 2$. Thus

$$\begin{aligned} |x_1(x_2 - 1)| + |x_1 - 1| &< 2 \cdot \frac{1}{4}\epsilon + \frac{1}{4}\epsilon \\ &= \frac{3}{4}\epsilon \end{aligned}$$

implying that

$$\begin{aligned} |f(x_1, x_2) - 1| &< \frac{3}{4}\epsilon \\ &< \epsilon \end{aligned}$$

9.5 Sequences

(we now use superscript for elements of sequence)
$\left\{ \mathbf{x}^k \right\}_{k=1}^{\infty}$ sequences of vectors in $\mathbb{R}^n$.

$$\begin{gathered} \mathbf{x}^k \in \mathbb{R}^n, \quad \forall\, k \\ \mathbf{x}^k = (x_1^k, x_2^k, \ldots, x_n^k) \end{gathered}$$

Definition 89. *A sequence $\left\{ \boldsymbol{x}^k \right\}_{k=1}^{\infty}$ converges to a point $\boldsymbol{x} \in \mathbb{R}^n$, that is $\boldsymbol{x}^k \longrightarrow \boldsymbol{x}$, if and only if $\forall\, \epsilon > 0$, $\exists\, K \in \mathbb{P}$ such that $k \geq K \Longrightarrow$*

$$d(\boldsymbol{x}^k, \boldsymbol{x}) < \epsilon$$

Definition 90. *For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$, we say*

$$\begin{gathered} \boldsymbol{a} \geq \boldsymbol{b} \\ \Longleftrightarrow a_i \geq b_i \quad \forall\, i = 1, 2, \ldots, n \end{gathered}$$

and

$$\begin{gathered} \boldsymbol{a} > \boldsymbol{b} \\ \Longleftrightarrow a_i \geq b_i \quad \forall\, i = 1, 2, \ldots, n, \\ \text{and } a_j > b_j \text{ for some } j \end{gathered}$$

Definition 91. *Let* $\mathcal{X} \subset \mathbb{R}$.
$\mathcal{X}$ *is bounded from above if* $\exists\ \overline{\boldsymbol{M}} \in \mathbb{R}^n$ *such that*

$$\overline{\boldsymbol{M}} \geq \boldsymbol{x}, \quad \forall \boldsymbol{x} \in \mathcal{X}$$

$\mathcal{X}$ *is bounded from below if* $\exists\ \underline{\boldsymbol{m}} \in \mathbb{R}^n$ *such that*

$$\underline{\boldsymbol{m}} \leq \boldsymbol{x}, \quad \forall \boldsymbol{x} \in \mathcal{X}$$

Now we define what mean by vectors being "greater" or "less" than each other.

Definition 92. $\mathcal{X}$ *is said to be* closed *if, for every sequence* $\left\{\boldsymbol{x}^k\right\}$ *from* $\mathcal{X}$*, if*

$$\boldsymbol{x}^k \longrightarrow \boldsymbol{x} \in \mathbb{R}^n$$

(so $\boldsymbol{x}$ *is a limit point of* $\mathcal{X}$*), then* $\boldsymbol{x} \in \mathcal{X}$.

Definition 93. $\mathcal{X}$ *is said to be a compact space if every sequence from* $\mathcal{X}$ *has a subsequence that converges to a point in* $\mathcal{X}$.

Definition 94. *For a metric* d*, a Ball of Radius* ϵ *around a point* $\boldsymbol{x} \in \mathbb{R}^n$ *is defined as*

$$B_\epsilon(x) \quad \equiv \quad \{\boldsymbol{y} \in \mathbb{R}^n \mid d(\boldsymbol{x}, \boldsymbol{y}) < \epsilon\}$$

FIGURE GOES HERE

Definition 95. $\mathcal{X}$ *is called* open *if* $\forall x \in \mathcal{X}, \exists \epsilon > 0$ *such that*

$$B_\epsilon(x) \subset \mathcal{X}$$

Note the following 2 common misconceptions

$$\mathcal{X} \text{ not open, does not imply } \mathcal{X} \text{ closed}$$

$$\mathcal{X} \text{ not closed, does not imply } \mathcal{X} \text{ open}$$

For example, $\mathbb{R}^n$ is both open and closed.
$[0, 1)$ is neither open nor closed.

9.6 Partial Derivatives and Directional Derivatives

Definition 96. *Take* $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$. *The* ith partial derivative *of* f *at* $\boldsymbol{x}$ *is defined as*

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}) \equiv \lim_{h \longrightarrow 0} \frac{f(\boldsymbol{x} + h\boldsymbol{e}_i) - f(\boldsymbol{x})}{h}.$$

Treat every other x_j as a constant and take the derivative as though f we a function of just x_i.

As in the one variable case, partial derivatives need not exist. If the ith partial derivative exists, then the function (when viewed as a function of x_i alone) must be continuous.

The definition illustrates the way that we will think about functions of many variables. We we think of them as many functions of one variable. If it is too hard to figure out how f is behaving when you move several variables at once, hold all but one of the variables fixed and analyze the one-variable function that results. A partial derivative is just an ordinary derivative of a function of one variable. The one variable is one of the components of the function. Roughly speaking, if a function is well behaved, knowing how it changes in every direction allows you to know how it behaves in general.

These considerations suggest a more general concept.

Definition 97. *Take* $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ *and let* $\boldsymbol{v}$ *be a unit vector in* $\mathbb{R}^n$. *The* directional derivative of f in the direction $\mathbf{v}$ at $\mathbf{x}$ *is defined as*

$$D_{\boldsymbol{v}}(\boldsymbol{x}) \equiv \lim_{h \longrightarrow 0} \frac{f(\boldsymbol{x} + h\boldsymbol{v}) - f(\boldsymbol{x})}{h}.$$

It follows from the definition, that

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) \equiv D_{\mathbf{e}_i}(\mathbf{x}).$$

That is, the ith partial derivative is just a directional derivative in the direction $\mathbf{e}_i$.

Notice that to compute directional derivatives you are just computing the derivative of a function of one variable. The one-variable function is the function of h of the form $f(\mathbf{x} + h\mathbf{v})$ with $\mathbf{x}$ and $\mathbf{v}$ fixed.

9.7 Differentiability

Definition 98. *We say a function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is* differentiable *at $\boldsymbol{a} \in \mathbb{R}^n$ if and only if there is a linear function $L\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that*

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{a}} \frac{\left\| \overset{m\times 1}{f(\boldsymbol{x})} - \overset{m\times 1}{f(\boldsymbol{a})} - \overset{n\times 1}{L(\boldsymbol{x}-\boldsymbol{a})} \right\|}{\left\| \underset{n\times 1}{\boldsymbol{x}-\boldsymbol{a}} \right\|} = 0.$$

If L exists we call it the derivative of f at $\boldsymbol{a}$ and denote it by $Df(\boldsymbol{a})$. In the case $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ this is equivalent to

$$\lim_{\boldsymbol{x}\longrightarrow\boldsymbol{a}} \left[\frac{f(\boldsymbol{x}) - f(\boldsymbol{a})}{\|\boldsymbol{x}-\boldsymbol{a}\|} - \frac{L(\boldsymbol{x}-\boldsymbol{a})}{\|\boldsymbol{x}-\boldsymbol{a}\|} \right] = 0$$

This implies that if f is differentiable, then for any directions defined by $\mathbf{y} \in \mathbb{R}^n$ and a magnitude given by $\alpha \in \mathbb{R}$

$$\begin{aligned} \lim_{\alpha\longrightarrow 0} \frac{f(\mathbf{a}+\alpha\mathbf{y}) - f(\mathbf{a})}{|\alpha|\,\|\mathbf{y}\|} &= \lim_{\alpha\longrightarrow 0} \frac{Df(\mathbf{a})\alpha\mathbf{y}}{|\alpha|\,\|\mathbf{y}\|} \\ &= \frac{Df(\mathbf{a})\mathbf{y}}{\|\mathbf{y}\|} \end{aligned}$$

That is, if a function in differentiable at a point, then all directional derivatives exist at the point.

Stated more officially:

Theorem 50. *If $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable then all of its directional derivatives exist. The directional derivative of f at $\boldsymbol{a} \in \mathbb{R}^n$ in direction $\boldsymbol{v} \in \mathbb{R}^n$ is given by*

$$Df(\boldsymbol{a})\boldsymbol{v}$$

We assume that the direction $\mathbf{v}$ in the definition of directional derivative is of unit length.

The theorem says two things. First, it says the if a function is differentiable, then the matrix representation of the derivative is just the matrix of partial derivatives. (This follows because the way that you get the matrix representation is to evaluate the linear function on the standard basis elements. Here, this would give you partial derivatives.) Second, it gives a formula for computing derivational

derivatives is any direction. A directional derivative is just a weighted average of partial derivatives.

The definition of differentiability should look like the one-variable definition. In both cases, the derivative is the best linear approximation to the function. In the multivariable setting a few things change. First, Euclidean distance replaces absolute values. Notice that the objects inside the $\|\cdot\|$ are vectors and not scalars. Taking their norms replaces a difference with a non-negative scalar. If we did not do this (at least for the denominator), then the ratios would not make sense. Second, because the linear function has the same domain and range as f, it is more complicated than in the one variable case. In the one variable case, the derivative of f evaluated at a point is a single function of a single variable. This allows us to think about f' as a real-valued function. When $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$, the derivative is a linear function from $\mathbb{R}^n$ into $\mathbb{R}^m$. This means that it can be represented by matrix multiplication of a matrix with m rows and n columns. That is, the derivative is described by mn numbers. What are these numbers? The computation after the definition demonstrates that the entries in the matrix that represents the derivative are the partial derivatives of (the component functions of) f. This is why we typically think of the derivatives of multivariable functions as "matrices of partial derivatives."

Sometimes people represent the derivative of a function from $\mathbb{R}^n$ to $\mathbb{R}$ as a vector rather than a linear function.

Definition 99. *Given*

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$$

the gradient *of f at $\boldsymbol{x} \in \mathbb{R}^n$ is*

$$\begin{aligned} Df(\boldsymbol{x}) &= \nabla f(\boldsymbol{x}) \\ &= \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}), \frac{\partial f}{\partial x_2}(\boldsymbol{x}), \ldots, \frac{\partial f}{\partial x_n}(\boldsymbol{x}) \right) \end{aligned}$$

The definition just introduces notation. In terms of the notation, Theorem 50 states that

$$D_{\mathbf{v}} f(a) = \nabla f(a) \cdot \mathbf{v}. \tag{9.1}$$

We have a definition of differentiability. We also have definitions of partial derivatives. Theorem 50 says that if you know that a function is differentiable, then you know that is has partial derivatives and that these partial derivatives tell

you everything you need to know about the derivative. What about the converse? That is, if a function has partial derivatives, then do we know it is differentiable? The answer is "almost".

Theorem 51. *Take $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, if f's partial derivatives exist and are continuous ($f \in C^1$), then f is differentiable.*

Theorem 51 states that if a function has partial derivatives in all directions, then the function is differentiable **provided that these partial derivatives are continuous.** There are standard examples of functions that have partial derivatives in all directions but still are not differentiable. In these examples the partial derivatives are discontinuous. These examples also provide situations in which equation (9.1) fails to be true (because the gradient can exist although the function is not differentiable).

Example 33.

$$f(\boldsymbol{x}) = x_1 x_2$$

$$\begin{aligned} \Longrightarrow \frac{\partial f}{\partial x_1}(0,0) &= x_2 \mid_{(x_1=0,x_2=0)} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \Longrightarrow \frac{\partial f}{\partial x_2}(0,0) &= x_1 \mid_{(x_1=0,x_2=0)} \\ &= 0 \end{aligned}$$

$$\lim_{h \longrightarrow 0} \frac{f((0,0) + h(1,1)) - f(0,0)}{h\sqrt{2}}$$

from $(0,0)$ the distance traveled to the point $(1,1)$ is $\sqrt{2}$, so that is why we normalize and divide in the denominator by $\sqrt{2}$.

$$\begin{aligned} &= \lim_{h \longrightarrow 0} \frac{f((h,h)) - f(0,0)}{h\sqrt{2}} \\ &= \lim_{h \longrightarrow 0} \frac{h^2}{h\sqrt{2}} \\ &= 0 \end{aligned}$$

So for this function the derivative in every direction is zero.

Example 34.

$$f(\boldsymbol{x}) = |x_1|^{\frac{1}{2}} |x_2|^{\frac{1}{2}}$$

So

$$\begin{aligned}\lim_{h\longrightarrow 0} \frac{f((0,0)+h(1,1))-f(0,0)}{h\sqrt{2}} &= \lim_{h\longrightarrow 0} \frac{f((h,h))-f(0,0)}{h\sqrt{2}} \\ &= \lim_{h\longrightarrow 0} \frac{|h|^{\frac{1}{2}} |h|^{\frac{1}{2}}}{h\sqrt{2}} \\ &= \lim_{h\longrightarrow 0} \frac{h}{h\sqrt{2}} \\ &= \frac{1}{\sqrt{2}}\end{aligned}$$

Why include this example? The computation above tells you that the directional derivative of $f$at $(0,0)$ in the direction $(1/\sqrt{2}, 1/\sqrt{2})$ exists and is equal to $1/\sqrt{2}$. On the other hand, you can easily check that both partial derivatives of the function at $(0,0)$ exist and are equal to zero. Hence the formula for computing the direction directive as the average of partials fails. Why? Because f is not differentiable at $(0,0)$.

Question: What is the direction from x that most increases the value of f? Answer: It's the direction given by the gradient.

Theorem 52. *Suppose $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable at x, then the direction that maximizes the directional derivative at x is given by*

$$\boldsymbol{v} = \nabla f(\boldsymbol{x})$$

9.8 Properties of the Derivative

Theorem 53. *$g, f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ both differentiable at $\boldsymbol{a} \in \mathbb{R}^n$*
Then

1. $$D[cf](\boldsymbol{a}) = cDf(\boldsymbol{a}) \forall c \in \mathbb{R}$$

2.

$$D[f+g](\boldsymbol{a}) = Df(\boldsymbol{a}) + Dg(\boldsymbol{a})$$

For the case m=1

3.

$$\underset{1\times n}{D[g\cdot f](\boldsymbol{a})} = \underset{1\times 1}{g(\boldsymbol{a})}\cdot\underset{1\times n}{Df(\boldsymbol{a})} + \underset{1\times 1}{f(\boldsymbol{a})}\cdot\underset{1\times n}{Dg(\boldsymbol{a})}$$

4.

$$D\left[\frac{f}{g}\right](\boldsymbol{a}) = \frac{g(\boldsymbol{a})\cdot Df(\boldsymbol{a}) - f(\boldsymbol{a})\cdot Dg(\boldsymbol{a})}{[g(\boldsymbol{a})]^2}$$

Theorem 54 (Chain Rule Theorem)**.** *Suppose $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{V} \subset \mathbb{R}^m$ are open sets and suppose $g\colon \mathcal{U} \longrightarrow \mathbb{R}^m$ is differentiable at $\boldsymbol{x} \in \mathcal{U}$ and $f\colon \mathcal{V} \longrightarrow \mathbb{R}^l$ is differentiable at $\boldsymbol{y} \equiv g(\boldsymbol{x}) \in \mathcal{V}$. Then $f \circ g$ is differentiable at $\boldsymbol{x}$ and*

$$\underbrace{D[f\circ g](\boldsymbol{x})}_{l\times n} = \underset{l\times m}{Df(\boldsymbol{y})}\underset{m\times n}{Dg(\boldsymbol{x})}$$

Proof. Though not a complete proof it is much more intuitive to consider the case

$$f\colon \mathbb{R}^m \longrightarrow \mathbb{R}, \qquad \text{and} \qquad g\colon \mathbb{R} \longrightarrow \mathbb{R}^m$$

So we can see that $l = n = 1$ and

$$f\circ g\colon \mathbb{R} \longrightarrow \mathbb{R}$$

Then

$$D(f\circ g)'(t) = Df(\mathbf{x})Dg(t) \quad \text{for } \mathbf{x} = g(t)$$

That is

$$D[f\circ g](t) = \frac{\partial f_1}{\partial x_1}\cdot\frac{\partial x_1}{\partial t} + \frac{\partial f_2}{\partial x_2}\cdot\frac{\partial x_2}{\partial t} + \cdots + \frac{\partial f_m}{\partial x_m}\cdot\frac{\partial x_m}{\partial t}$$

□

Example 35. *Let the variable t denote the price of oil. This one variable induces an array of population responses (thus becomes a vector valued function) like*

1. *What type of car to buy?*

2. *How fast to expand your business?*
3. *How far to drive on holiday?*
4. *etc.*

and then these responses in turn have their own effect like determining GNP, the variable y (which was got by the function f using these population responses).

$$\begin{array}{ccccccc} t & \xrightarrow{g} & \boldsymbol{x} & = g(t) \xrightarrow{f} & y & = f(g(t)) \in \mathbb{R} \\ \downarrow & & \downarrow & & \downarrow & \\ \textit{price of oil} & & \textit{actions taken by individuals} & & \textit{GNP} & \end{array}$$

$$\begin{aligned} D[f \circ g](t) &= \frac{\partial y}{\partial t} \\ &= D(f(g(t))Dg(t) \\ &= \left(\frac{\partial f}{\partial x_1}(g(t)), \ldots, \frac{\partial f}{\partial x_m}(g(t)) \right) \cdot \begin{pmatrix} \frac{dg_1}{dt} \\ \vdots \\ \frac{dg_m}{dt} \end{pmatrix} \\ &= \sum_{i=1}^{m} \frac{\partial y}{\partial x_1} \cdot \frac{dg_1}{dt} \end{aligned}$$

Example 36.

$$g(x) = x - 1$$
$$f(y) = \begin{pmatrix} 2y \\ y^2 \end{pmatrix}$$

So note that

$$g \colon \mathbb{R} \longrightarrow \mathbb{R}, \qquad \textit{and} \qquad f \colon \mathbb{R} \longrightarrow \mathbb{R}^2$$
$$[f \circ g](x) = \begin{pmatrix} 2(x-1) \\ (x-1)^2 \end{pmatrix}$$
$$D[f \circ g](x) = \begin{pmatrix} 2 \\ 2(x-1) \end{pmatrix}$$

Now let's see if we get the same answer doing it the chain rule way:

$$Dg(x) = 1$$

$$Df(y) = \begin{pmatrix} 2 \\ 2y \end{pmatrix}$$

$$Df(g(x))Dg(x) = \begin{pmatrix} 2 \\ 2(x-1) \end{pmatrix}$$

Example 37.

$$\begin{aligned} f(\boldsymbol{y}) &= f(y_1, y_2) \\ &= \begin{pmatrix} y_1^2 + y_2 \\ y_1 - y_1 y_2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} g(\boldsymbol{x}) &= g(x_1, x_2) \\ &= \begin{pmatrix} x_1^2 - x_2 \\ x_1 x_2 \end{pmatrix} \\ &= \boldsymbol{y} \end{aligned}$$

So we note here that both g and f take in two arguments and spit out a (2×1) vector, so we must have

$$g\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad \textit{and} \qquad f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

Again, it is very very important to keep track of the dimensions!

$$\begin{aligned} D[f \circ g](\boldsymbol{x}) &= Df(g(\boldsymbol{x})Dg(\boldsymbol{x}) \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{pmatrix} \\ &= \begin{pmatrix} 2y_1 & 1 \\ 1 - y_2 & -y_1 \end{pmatrix} \cdot \begin{pmatrix} 2x_1 & -1 \\ x_2 & x_1 \end{pmatrix} \end{aligned}$$

and we know that

$$y_1 = x_1^2 - x_2$$

$$y_2 = x_1 x_2$$

So

$$= \begin{pmatrix} 2x_1^2 - 2x_2 & 1 \\ 1 - x_1x_2 & x_2 - x_1^2 \end{pmatrix} \cdot \begin{pmatrix} 2x_1 & -1 \\ x_2 & x_1 \end{pmatrix}$$
$$= \begin{pmatrix} 4x_1(x_1^2 - x_2) + x_2 & x_1 - 2(x_1^2 - x_2) \\ 2x_1(1 - x_1x_2) + x_2(x_2 - x_1^2) & x_1(x_2 - x_1^2) + x_1x_2 \end{pmatrix}$$

9.9 Gradients and Level Sets

EXAMPLES AND FIGURES GO HERE

Example 38.

$$f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}$$

so the graph is in $\mathbb{R}^4$ (pretty difficult to draw!), but the graph of the level set is in $\mathbb{R}^3$.

$$f(\boldsymbol{x}) = x_1^2 + x_2^2 + x_3^2 = 1$$

Note *that this is just a sphere of radius 1.*
So at any point along the sphere there is not just one tangent line to the point - there are lots of them.

In general, a *surface* in $\mathbb{R}^{n+1}$ can be viewed as the solution to a system of equations. For convenience, represent a point in $\mathbb{R}^{n+1}$ as a pair (x, y), with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$. If $F\colon \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$, then the set $\{(x, y) : F(x, y) = 0\}$ is typically an n dimensional set. We can talk about what it means to be a tangent to this surface. Certainly the tangent at (x_0, y_0) should be an n dimensional linear manifold in $\mathbb{R}^{n+1}$ that contains (x_0, y_0). It should also satisfy the approximation property: if (x, y) is a point on the surface that is close to (x_0, y_0), then it should be approximated up to first order by a point on the tangent. One way to think about this property is to think about directions on the surface. Consider a function $G\colon \mathbb{R} \longrightarrow \mathbb{R}^{n+1}$ such that $G(0) = (x_0, y_0)$ and $F \circ G(t) \equiv 0$ for t in a neighborhood of 0. G defines a curve on the surface through (x_0, y_0). A direction on the

surface at (x_0, y_0) is just a direction of a curve through (x_0, y_0) or $DG(0)$. By the chain rule it follows that

$$\nabla F(x_0, y_0) \cdot DG(0) = 0,$$

which implies that $\nabla F(x_0, y_0)$ is orthogonal to all of the directions on the surface. This generates a non-trivial hyperplane provided that $DF(x_0, y_0) \neq \mathbf{0}$. We summarize these observations below. (You can decide whether this is a definition or a theorem.)

Definition 100. *Let $F\colon \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$ be differentiable at the point (x_0, y_0). Assume that $F(x_0, y_0) = 0$ and that $DF(x_0, y_0) \neq \mathbf{0}$. The equation of the hyperplane tangent to the surface $F(x, y) = 0$ at the point (x_0, y_0) is*

$$\nabla F(x_0, y_0) \cdot ((x, y) - (x_0, y_0)) = 0. \tag{9.2}$$

Now suppose that $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable at $x \in \mathbb{R}^n$. Consider the function $F(x, y) = f(x) - y$. The surface $F(x, y) = 0$ is exactly the graph of f. Hence the tangent to the surface is the tangent to the graph of f. This means that the formula for the equation of the tangent hyperplane given above can be used to find the formula for the equation of the tangent to the graph of a function.

Theorem 55. *If $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable at $x_0 \in \mathbb{R}^n$ then the vector $\nabla f(x_0)$ is normal (perpendicular) to the tangent vector of the level set of f at value $f(\boldsymbol{x})$ at point $x \in \mathbb{R}^n$ and the equation of the hyperplane tangent to the graph of f at the point $(x_0, f(x_0))$ is*

$$\nabla f(x_0) \cdot (x - x_0)) = y - y_0.$$

Proof. Substitute $\nabla F(x_0, y_0) = (\nabla f(x_0, -1)$ into equation (9.2) and re-arrange terms. □

Example 39. *Find the tangent plane to $\{\boldsymbol{x} \mid x_1 x_2 - x_3^2 = 6\} \subset \mathbb{R}^3$ at $\widehat{\boldsymbol{x}} = (2, 5, 2)$.* Notice *if you let $f(\boldsymbol{x}) = x_1 x_2 - x_3^2$, then this is a level set of f for value 6.*

$$\begin{aligned} \nabla f(\boldsymbol{x}) &= (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}) \\ &= (x_2, x_1, -2x_3) \end{aligned}$$

$$\begin{aligned}\nabla f(\widehat{\boldsymbol{x}}) &= \nabla f(\boldsymbol{x}) \mid_{\boldsymbol{x}=(2,5,2)} \\ &= (5, 2, -4)\end{aligned}$$

Tangent Plane:

$$\begin{aligned}\{\widehat{\boldsymbol{x}} + \boldsymbol{y} \mid \boldsymbol{y} \cdot \nabla f(\widehat{\boldsymbol{x}}) = 0\} &= \{(2,5,2) + (y_1, y_2, y_3) \mid 5y_1 + 2y_2 - 4y_3 = 0\} \\ &= \{\boldsymbol{x} \mid 5x_1 - 10 + 2x_2 - 10 - 4x_3 + 8 = 0\} \\ &= \{5x_1 + 2x_2 - 4x_3 = 12\}\end{aligned}$$

Example 40. *Suppose* $f(x, y, z) = 3x^2 + 2xy - z^2$. $\nabla f(x, y, z) = (6x + 2y, 2x, -2z)$. *Notice that* $f(2, 1, 3) = 7$. *The level set of* f *when* $f(x, y, z) = 7$ *is* $\{(x, y, z) : f(x, y, z) = 7\}$. *This set is a (two dimensional) surface in* $\mathbb{R}^3$*: It can be written* $F(x, y, z) = 0$ *(for* $F(x, y, z) = f(x, y, z) - 7$*). Consequently the equation of the tangent to the level set of* f *is a (two-dimensional) hyperplane in* $\mathbb{R}^3$. *At the point* $(2, 1, 3)$*, the hyperplane has normal equal to* $\nabla f(2, 1, 3) = (12, 4, -6)$. *Hence the equation of the hyperplane to the level set at* $(2, 1, 3)$ *is equal to:*

$$(12, 4, -6) \cdot (x - 2, y - 1, z - 3) = 0$$

or

$$12x + 4y - 6z = 10.$$

On the other hand, the graph of f *is a three-dimensional subset of* $\mathbb{R}^4$*:* $\{(x, y, z, w) : w = f(x, y, z)\}$. *A point on this surface is* $(2, 1, 3, 7) = (x, y, z, w)$. *The tangent hyperplane at this point can be written:*

$$w - 7 = \nabla f(2, 1, 3) \cdot (x - 2, y - 1, z - 3) = 12x + 4y - 6z - 10$$

or

$$12x + 4y - 6z - w = 3.$$

9.10 Homogeneous Functions

Definition 101. *The function* $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ *is homogeneous of degree* k *if* $F(\lambda x) = \lambda^k F(x)$ *for all* λ.

Homogeneity of degree one is weaker than linearity: All linear functions are homogeneous of degree one, but not conversely. For example, $f(x, y) = \sqrt{xy}$ is homogeneous of degree one but not linear.

Theorem 56 (Euler's Theorem)**.** *If $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a differential at x and homogeneous of degree k, then $\nabla F(x) \cdot x = kF(x)$.*

Proof. Fix x. Consider the function $H(\lambda) = F(\lambda x)$. This is a composite function, $H(\lambda) = F \circ G(\lambda)$, where $G : \mathbb{R} \longrightarrow \mathbb{R}^n$, such that $G(\lambda) = \lambda x$. By the chain rule, $DH(\lambda) = DF(G(\lambda))DG(\lambda)$. If we evaluate this when $\lambda = 1$ we have

$$DH(1) = \nabla F(x) \cdot x. \tag{9.3}$$

On the other hand, we know from homogeneity that $H(\lambda) = \lambda^k x$. Differentiating the right hand side of this equation yields $DH(\lambda) = k\lambda^{k-1}F(\lambda x)$ and evaluating when $\lambda = 1$ yields

$$DH(1) = kF(x). \tag{9.4}$$

Combining equations (9.3) and (9.4) yields the theorem. □

In economics functions that are homogeneous of degree zero and one arise naturally in consumer theory. A cost function depends on the wages you pay to workers. If all of the wages double, then the cost doubles. This is homogeneity of degree one. On the other hand, a consumer's demand behavior is typically homogeneous of degree zero. Demand is a function $\phi(p, w)$ that gives the consumer's utility maximizing feasible demand given prices p and wealth w. The demand is the best affordable consumption for the consumer. The consumptions x that are affordable satisfy $p \cdot x \leq w$ (and possibly another constraint like non-negativity). If p and w are multiplied by the same factor, λ, then the budget constraint remains unchanged. Hence the demand function is homogeneous of degree zero.

Euler's Theorem provides a nice decomposition of a function F. Suppose that F describes the profit produced by a team of n agents, when agent i contributes effort x_i. How such the team divide the profit it generates? If F is linear, the answer is easy: If $F(x) = p \cdot x$, then just give agent i $p_i x_i$. Here you give each agent a constant "per unit" payment equal to the marginal contribution of her effort. When you do so, you distribute the entire surplus (and nothing else). When F is non-linear, it is harder to figure out the contribution of each agent. The theorem states that if you pay each agent her marginal contribution ($D_{e_i} f(x)$) per unit, then you distribute the surplus fully if F is homogeneous of degree one. Otherwise it identifies alternative ways to distribute the surplus.

9.11 Higher-Order Derivatives

If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable, then its partial derivatives, $D_{e_i} f(x) = D_i f(x)$ can also be viewed as functions from $\mathbb{R}^n$ to $\mathbb{R}$. So these functions maybe differentiable. This leads to the idea of higher-order partial derivatives, which are the

natural generalization of higher-order derivatives of a function of one variable. You can imagine taking a derivative with respect to one variable, then other then the first again, and so on, creating sequences of the form $D_{i_k} \cdots D_{i_2} D_{i_1} f(x)$. Fortunately, provided that all of the derivatives are continuous in a neighborhood of x, the order in which you take partial derivatives does not matter. Hence we will denote an kth derivative $D^k_{i_1,\ldots,i_n} f(x)$, where k is the total number of derivatives and i_j is the number of partial derivatives with respect to the jth argument (so each i_j is a non-negative integer and $\sum_{j=1}^n i_j = x$.

Except for Taylor's Formula, we will have little interest in third or higher derivatives. Second derivatives are important. A real-valued function of n variables will have n^2 second derivatives, which we sometimes think of as terms in a square matrix:

$$\underset{n\times n}{D^2 f(\mathbf{x})} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}) \end{pmatrix}$$

We say $f \in C^k$ if the kth order partials exist and are continuous.

9.12 Taylor Approximations

The multivariable version of Taylor's Theorem looks intimidating, but it really is a one-variable theorem. Suppose that you know about a function at the point $a \in \mathbb{R}^n$ and you want to approximate the function at x. Consider the function $F(t) = f(xt + a(1-t))$. $F : \mathbb{R} \longrightarrow \mathbb{R}$ and so you can use one-variable calculus. Also $F(1) = x$ and $F(0) = a$. So if you want to know about the function f at x using information about the function f at a, you really just want to know about the one-variable function F at $t = 1$ knowing something about F at $t = 0$. Hence to get the multivariable version of Taylor's Theorem, you apply the one variable version of the theorem to F. You need the chain rule to compute the derivatives of F (in terms of f) and there are a lot of these derivatives. This means that the formula is a bit messy, but it is conceptually no different from the one-dimensional theorem.

1st Order Approx Consider $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ such that f is differentiable. At $\mathbf{a} \in \mathbb{R}^n$ the 1st degree Taylor Polynomial of f is

$$P_1(\mathbf{x}) \equiv f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})$$

The first-order approximation should be familiar. Notice that you have n "derivatives" (the partials).

If we write $f(x) = P_1(x,a) + E_2(x,a)$ for the first-order approximation with error of f at x around the point a, then we have

$$\begin{aligned}\lim_{\mathbf{x}\to\mathbf{a}} \frac{|E_2(\mathbf{x},\mathbf{a})|}{\|\mathbf{x}-\mathbf{a}\|} &= \lim_{\mathbf{x}\to\mathbf{a}} \frac{|f(\mathbf{x}) - f(\mathbf{a}) - Df(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a})|}{\|\mathbf{x}-\mathbf{a}\|} \\ &= 0\end{aligned}$$

Thus as before, as $\mathbf{x}\to\mathbf{a}$, E_2 converges to 0 faster than $\mathbf{x}$ to $\mathbf{a}$.

2nd Order Approx If $f \in C^2$, the 2nd degree Taylor approximation is

$$f(\mathbf{x}) = \underbrace{f(\mathbf{a}) + \underset{1\times n}{\nabla f(\mathbf{a})}\underset{n\times 1}{(\mathbf{x}-\mathbf{a})} + \frac{1}{2}\underset{1\times n}{(\mathbf{x}-\mathbf{a})'}\underset{n\times n}{D^2 f(\mathbf{a})}\underset{n\times 1}{(\mathbf{x}-\mathbf{a})}}_{P_2(\mathbf{x},a)} + E_3(\mathbf{x},\mathbf{a})$$

where

$$\begin{aligned}\frac{1}{2}\underbrace{\underset{1\times n}{(\mathbf{x}-\mathbf{a})'}\underset{n\times n}{D^2 f(\mathbf{a})}\underset{n\times 1}{(\mathbf{x}-\mathbf{a})}}_{1\times 1} &= \frac{1}{2}(x_1-a_1,\ldots,x_n-a_n)\cdot\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{pmatrix}\cdot\begin{pmatrix} x_1 - a \\ \vdots \\ x_n - a \end{pmatrix} \\ &= \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n (x_i-a_i)\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})(x_j-a_j)\end{aligned}$$

Notice that the error term is a quadratic form.

The general form is quite messy. The way to make the mess more pleasant is to let the notation do all of the work. We define $D_h^k f$ to be a kth derivative:

$$D_h^k f = \sum_{j_1+\cdots+j_n=k} \binom{k}{j_1\cdots j_n} h_1^{j_1}\cdots h_n^{j_n} D_1^{j_1}\cdots D_n^{j_n} f,$$

where the summation is taken over all n-tuples of $j_1,\ldots,j_n$ of non-negative integers that sum to k and the symbol

$$\binom{k}{j_1\cdots j_n} = \frac{k!}{j_1!\cdots j_n!}.$$

Here is the general form.

Theorem 57 (Taylor's Theorem)**.** *If f is a real-valued function in C^{k+1} defined on an open set containing the line segment connecting a to x, then there exists a point η on the segment such that*

$$f(x) = P_k(x, a) + E_k(x, a)$$

where $P_k(x, a)$ is a kth order Taylor's Approximation:

$$P_k(x, a) = \sum_{r=0}^{k} \frac{D^r_{x-a}(a)}{r!}$$

and $E_k(x, a)$ is the error term:

$$E_k(x, a) = \frac{D^{k+1}_{x-a}(\eta)}{(k+1)!}.$$

Moveover, the error term satisfies:

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{E_k(\boldsymbol{a}+\boldsymbol{h}, \boldsymbol{a})}{\|\boldsymbol{h}\|^k} = 0$$

Chapter 10

Convexity

This section contains some basic information about convex sets and functions in $\mathbb{R}^n$.

10.1 Preliminary: Topological Concepts

Let $\mathcal{X} \subset \mathbb{R}^n$.

Definition 102. $x \in \mathbb{R}^n$ *is a* limit point *of* $\mathcal{X}$ *if* $\forall\, \epsilon > 0$

$$B_\epsilon(x) \cap \mathcal{X} \neq \phi$$

Definition 103. x *is an* interior point *of* $\mathcal{X}$ *if for some* $\epsilon > 0$

$$B_\epsilon \subset \mathcal{X}$$

Definition 104. x *is a* boundary point *of* $\mathcal{X}$ *if* $\forall\, \epsilon > 0$ *both*

$$B_\epsilon(x) \cap \mathcal{X} \neq \phi$$

and

$$B_\epsilon(x) \cap [\mathbb{R}^n \backslash \mathcal{X}] \neq \phi$$

Definition 105. *$\mathcal{X}$ is called* open *if every $x \in \mathcal{X}$ is an interior point.*

Definition 106. *$\mathcal{X}$ is called* closed *if it contains all of its limit points*

Reminder: A hyperplane in $\mathbb{R}^n$ is of dimension $\mathbb{R}^{n-1}$ and can be expressed as

$$\mathbf{x} \cdot \mathbf{p} = c$$

where
$\mathbf{x} \in \mathbb{R}^n$
$\mathbf{p} \neq \mathbf{0}$ is in $\mathbb{R}^n$
$c \in \mathbb{R}$. $\mathbf{p}$ is the normal to the hyperplane.

FIGURE GOES HERE

10.2 Convex Sets

Definition 107. *$\mathcal{X} \subset \mathbb{R}^n$ is called* convex *if $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ and $\forall\, \alpha \in [0, 1]$ we have*

$$\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in \mathcal{X}$$

FIGURE GOES HERE

If you draw a line (or in higher dimensions a hyperplane) between any 2 points in $\mathcal{X}$, then every point on the line is also in $\mathcal{X}$.
Note that a set is either "convex" or "not convex" - there is no such thing as a concave set!

Theorem 58. *Given a nonempty, closed, convex set $\mathcal{X} \subset \mathbb{R}^n$ and $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x} \notin \mathcal{X}$. There exists $\boldsymbol{p} \in \mathbb{R}^n, \boldsymbol{p} \neq \boldsymbol{0}$, and $c \in \mathbb{R}$ such that*

$$\mathcal{X} \subset \{\boldsymbol{y} \mid \boldsymbol{y} \cdot \boldsymbol{p} \geq c\}$$

and

$$\boldsymbol{x} \cdot \boldsymbol{p} < c$$

That is, $\mathbf{y} \cdot \mathbf{p} = c$ defines a separating hyperplane for $\mathcal{X}$.
FIGURE GOES HERE

Proof. Consider the problem of minimizing the distance between $\mathbf{x}$ and the set $\mathcal{X}$. That is, find $\mathbf{y}^*$ to solve:

$$\min \|\mathbf{x} - \mathbf{y}\| \text{ subject to } y \in \mathcal{X}. \tag{10.1}$$

The norm is a continuous function. While $\mathcal{X}$ is not necessarily bounded, since it is nonempty, we can find some element $\mathbf{z} \in \mathcal{X}$. Without loss of generality we can replace $\mathcal{X}$ in (10.1) by $\{y \in \mathcal{X} : \|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{z} - \mathbf{x}\|\}$. This set is compact because $\mathcal{X}$ is closed. Hence there is a solution $\mathbf{y}^*$ to (10.1). Let $\mathbf{p} = \mathbf{y}^* - \mathbf{x}$. Since $\mathbf{x} \notin \mathcal{X}$, $\mathbf{p} \neq \mathbf{0}$. Let $c = \mathbf{y}^* \cdot \mathbf{p}$. Since $c - \mathbf{p} \cdot \mathbf{x} = \|\mathrm{p}\|^2$, $c > \mathbf{p} \cdot \mathbf{x}\$. To complete the proof we must show that if

$$\mathbf{y} \in \mathcal{X}, \text{ then } \mathbf{y} \cdot p \geq c.$$

This inequality is equivalent to

$$(\mathbf{y} - \mathbf{y}^*) \cdot (\mathbf{y}^* - \mathbf{x}) \geq 0. \tag{10.2}$$

Since $\mathcal{X}$ is convex and $\mathbf{y}^*$ is defined to solve (10.1), it must be that $\|t\mathbf{y} + (1-t)\mathbf{y}^* - x\|^2$ is minimized when $t = 0$ so that the derivative of $\|t\mathbf{y} + (1-t)\mathbf{y}^* - x\|^2$ is non-negative at $t = 0$. Differentiation and simplifying yields inequality (10.2). □

Notice that without loss of generality you can normalize the normal to the separating hyperplane. That is, you can assume that $\|\mathbf{p}\| = 1$.

You can refine the separating hyperplane theorem in two ways. First, if $\mathbf{x}$ is an element of the boundary of $\mathcal{X}$, then you can approximate $\mathbf{x}$ by a sequence $\mathbf{x}_k$ such that each $\mathbf{x}_k \notin \mathcal{X}$. This yields a sequence of $\mathbf{p}_k$, which can be taken to be unit vectors, that satisfy the conclusion of the theorem. A subsequence of the $\mathbf{p}_k$ must converge. The limit point $\mathbf{p}^*$ will satisfy the conclusion of the theorem (except we can only guarantee that $c \geq \mathbf{p}^* \cdot \mathbf{x}$ rather than the strict equality above). Second, one can check that the closure of any convex set is convex. Therefore, given a convex set $\mathcal{X}$ and a point $\mathbf{x}$ not in the interior of the set, we can separate $\mathbf{x}$ from the closure of $\mathcal{X}$. Taking these considerations into account we have the following version of the separating hyperplane theorem.

Theorem 59. *Given a convex set $\mathcal{X} \subset \mathbb{R}^n$ and $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}$. If $\boldsymbol{x}$ is not in the interior of $\mathcal{X}$, then there exists $\boldsymbol{p} \in \mathbb{R}^n$, $\boldsymbol{p} \neq \boldsymbol{0}$, and $c \in \mathbb{R}$ such that*

$$\mathcal{X} \subset \{\boldsymbol{y} \mid \boldsymbol{y} \cdot \boldsymbol{p} \geq c\}$$

and

$$\boldsymbol{x} \cdot \boldsymbol{p} \leq c$$

In the typical economic application, the separating hyperplane's normal has the interpretation of prices and the separation property states that a particular vector costs more than vectors in a consumption set.

10.3 Quasi-Concave and Quasi-Convex Functions

Definition 108. *A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is called* quasiconcave *if $\{\boldsymbol{x} \mid f(\boldsymbol{x}) \geq a\}$ is convex for all $a \in \mathbb{R}$. In other words the upper contour set, as defined in Definition 86, is a convex set.*

Example 41. $f(\boldsymbol{x}) = -x_1^2 - x_2^2$
the upper contour set is the inner shaded part of the sphere.

FIGURE GOES HERE

Definition 109. *A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is called* quasiconvex *if $\{\boldsymbol{x} \mid f(\boldsymbol{x}) \leq a\}$ is convex for all $a \in \mathbb{R}$. In other words the lower contour set, as defined in Definition 86, is a convex set.*

Note: In the one-variable setting, we said that a function was quasiconcave if

$$\text{for all } x, y, \text{ and } \lambda \in (0,1), f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}. \qquad (10.3)$$

To see that this definition is equivalent to Definition 108 note first that if $a = \min\{f(x), f(y)\}$, then Definition 108 implies (10.3). Conversely, if the condition in the Definition 108 fails, then there exists a, x, y such that $f(x), f(y) \geq a$ but

$f(\lambda x + (1-\lambda)y) < a$. Plainly condition (10.3) fails for these values of λ, x, and y.

Quasiconcavity and Quasiconvexity are global properties of a function. Unlike continuity, differentiability, concavity and convexity (of functions), they cannot be defined at a point.

Theorem 60. *If f is concave and $\boldsymbol{x}^*$ is a local maximizer of f, then $\boldsymbol{x}^*$ is a global maximizer.*

Proof. (by contradiction)
Suppose $\mathbf{x}^*$ is not a global maximizer of f. Then there is a point $\widehat{\mathbf{x}}$ such that

$$f(\widehat{\mathbf{x}}) > f(\mathbf{x}^*)$$

but then we would have that

$$f(\alpha\mathbf{x}^* + (1-\alpha)\widehat{\mathbf{x}}) \geq \alpha f(\mathbf{x}^*) + (1-\alpha)f(\widehat{\mathbf{x}})$$

And this contradicts that $\mathbf{x}^*$ is a local max. □

10.3.1 How to check if a function f is quasiconcave or not

There are a few steps to follow.

1. First solve for the level sets and graph a selection of them

2. Decide by inspection which side of the plotted level set is the upper contour set and which side is the lower contour set

3. Are the inspected upper and lower contour sets convex?

Note we are looking for nice shaped graphs. If the graph bends and twists alot then it is not quasiconvex or quasiconcave. We are looking for nice graphs like circles or parabolas.

10.3.2 Relationship between Concavity and Quasiconcavity

Suppose $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ is concave $\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then $\forall\, \alpha \in [0,1]$

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y})$$

FIGURE GOES HERE

Points on the boundary of this level set (convex) give you values of the function equal to some constant c. Points inside the shaded region give you values for the function that are greater than or equal to c. We want to examine the set $\{\, \mathbf{z} \mid f(\mathbf{z}) \geq c\}$. Consider the points $\mathbf{x}, \mathbf{y} \in \{\, \mathbf{z} \mid f(\mathbf{z}) \geq c\}$,

$$\implies f(\mathbf{x}) \geq c \quad \text{and} \quad f(\mathbf{y}) \geq c$$

and this means $\forall\, \alpha \in [0,1]$ that

$$\begin{aligned} f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) &\geq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) \\ &\geq \alpha c + (1-\alpha) c \\ &= c \end{aligned}$$

$$\implies \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in \{\, \mathbf{z} \mid f(\mathbf{z}) \geq c\}$$

This is really providing the intuition for the following Theorem.

Theorem 61. *If $f: \mathbb{R}^n \longrightarrow \mathbb{R}$ is concave, then f is quasiconcave.*
Similarly if f is convex, then f is quasiconvex.

Note that concavity implies quasiconcavity but that the opposite is not true. A quasiconcave function may be concave, but certainly does not have to be.

Example 42. *e^x is quasiconcave but not concave. In fact it is also convex and quasiconvex!*

The easiest way to see this is just to note that $f(x) = e^x$ is just a function of one variable and thus is easy to plot. But we plot the function not the level sets (*Note* level sets are points). Since the exponential function is increasing, upper contour sets are of the form $[a, \infty)$. So the function is quasiconcave. The function is quasiconvex too!

More generally, we have:

Theorem 62. *Any increasing or decreasing function is both quasiconvex and quasiconcave.*

10.3.3 Ordinal "vs" Cardinal

Definition 110. *A concept will be thought of as* ordinal *if "order" is all that matters. For example the statement*

$$f(\boldsymbol{x}) > f(\boldsymbol{y})$$

In other words it doesn't matter how much greater the value $f(\mathbf{x})$ is than $f(\mathbf{y})$, it just matters that it is greater.

Definition 111. *A concept will be thought of as* cardinal *if things like shape, or distances matter. So it is more than just which is greater - the actual values matter.*

Example 43. *To help understand ordinal concept, suppose we have 3 kids: Anne, Bert, and Chris. Suppose Anne is 5 foot tall, Bert is 5'1" tall, while Chris is a giant at 9 foot tall! We can see that Anne is the smallest, Bert is next, and then Chris. The fact that Chris is 3'11" taller than Bert while Bert is only a measly 1' taller than Anne is irrelevant to the ordinal ranking. All that matters is the order. The point is Chris is taller than Bert so gets a higher ranking than him, while Bert is taller than Anne so gets a higher ranking than her. If Bert now grew to be 8'11" tall, this would in fact change nothing as regards the ordinal ranking.*

Example 44. *A good example of an ordinal concept is utility. A utility function is a function which measures the level of happiness (or utility) that a person gets out of the argument. (In 200A you will see all the conditions necessary for the existence of such a function - but for now we will just assume that it exists).*
So suppose we have a guy Joel, and the only thing in the world that gives Joel any joy are apples. Then we can write Joel's utility function as

$$U_J : \mathbb{R}_{++} \longrightarrow \mathbb{R}$$

So the input into Joel's utility function is just a positive real number (i.e. the number of apples he ate) and the output is also just a number (the number of "utils"

that this number of apples gave him). But the important point is that we do not attach any psychological significance to the value attained by the function, because Joel only cares about the ranking *of how much joy (how many utils) he gets from different quantities of apples.*

Suppose

$$U_J(x) = x^{\frac{1}{2}}$$

Then we have that

$$U_J(16) = 16^{\frac{1}{2}} = 4 \quad \textit{and} \quad U_J(9) = 9^{\frac{1}{2}} = 3$$

So we have that

$$U_J(16) > U_J(9)$$

The fact that it is 1 "util" greater is irrelevant - the important thing is that 16 apples gives Joel more utils than 9 apples. So as you would expect he prefers 16 apples to 9!! He ranks the bundles of goods (in this case apples) ordinally.

While it seems obvious that Joel prefers 16 apples to 9 apples, it does get more tricky when Joel's utility function depends on two commodities: say apples and bananas. But you will see all this in 200A.

Definition 112. *We say that property $(*)$ is ordinal if $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ has property $(*)$ and $g\colon \mathbb{R} \longrightarrow \mathbb{R}$ is strictly increasing*

$$\Longrightarrow g \circ f \textit{ has property}(*)$$

Note this is saying that $g \circ f$ preserves order from f. In other words, however we rank the order with f, $g \circ f$ will preserve this order.
For some $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ if

$$f(\mathbf{x}) > f(\mathbf{y})$$
$$\Longrightarrow g \circ f(\mathbf{x}) > g \circ f(\mathbf{y})$$

Example 45.

$$f(x) = e^x$$
$$g(y) = 3y$$
$$g \circ f(x) = 3e^x$$

Example 46.

$$f\colon (0,\infty) \longrightarrow \mathbb{R}$$
$$f(x) = x^{\frac{1}{2}}$$

So if

$$x > y \Longrightarrow f(x) > f(y)$$

This is clearly a concave function. Now suppose

$$g(y) = y^4$$
$$\Longrightarrow g \circ f(x) = x^2$$

And again we have that

$$x > y \Longrightarrow g \circ f(x) > g \circ f(y)$$

But if you graph this is it now a convex function!!!

conclusion: concavity and convexity are not ordinal concepts.

Example 47. *Suppose $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is quasiconcave. That is, the set $\{ x \mid f(x) \geq c\}$ is convex $\forall c \in \mathbb{R}$.*
Now take a strictly increasing function

$$g\colon \mathbb{R} \longrightarrow \mathbb{R}$$

Note *that since g is strictly increasing, g is invertible.*
So the following 2 sets must be equivalent

$$\{ x \mid g \circ f(x) \geq c\} \Longleftrightarrow \left\{ x \mid f(x) \geq g^{-1}(c)\right\}$$

and so the upper contour set remains convex $\Longrightarrow$ quasiconcavity is preserved!!!

conclusion: quasiconcavity and quasiconvexity are ordinal concepts.

Another more technical way to distinguish the two is to say that an ordinal property is preserved under any monotonic transformation while a cardinal property is preserved only under a positive affine transformation. Ordinal properties rank points in the domain while cardinal properties rank the cartesian product of points in the domain.

Chapter 11

Unconstrained Extrema of Real-Valued Functions

11.1 Definitions

The following are natural generalizations of the one-variable concepts.

Definition 113. *Take* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$.

$\boldsymbol{x}^*$ *is a* local maximizer $\iff \exists\, \delta > 0$ *such that* $\forall\, \boldsymbol{x} \in B_\delta(\boldsymbol{x}^*)$*, we have*

$$f(\boldsymbol{x}) \leq f(\boldsymbol{x}^*)$$

$\boldsymbol{x}^*$ *is a* local minimizer $\iff \exists\, \delta > 0$ *such that* $\forall\, \boldsymbol{x} \in B_\delta(\boldsymbol{x}^*)$*, we have*

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$$

$\boldsymbol{x}^*$ *is a* local maximizer $\iff \exists\, \delta > 0$ *such that* $\forall\, \boldsymbol{x} \in \mathbb{R}^n$*, we have*

$$f(\boldsymbol{x}) \leq f(\boldsymbol{x}^*)$$

$\boldsymbol{x}^*$ *is a* local minimizer $\iff \exists\, \delta > 0$ *such that* $\forall\, \boldsymbol{x} \in \mathbb{R}^n$*, we have*

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$$

11.2 First-Order Conditions

Theorem 63 (First Order Conditions)**.** *If f is differentiable at $\boldsymbol{x}^*$, and $\boldsymbol{x}^*$ is a local maximizer or minimizer then*

$$\begin{aligned} Df(\boldsymbol{x}) &= \nabla f(\boldsymbol{x}) \\ &= \boldsymbol{0} \end{aligned}$$

That is

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}^*) = 0,$$

$\forall\, i = 1, 2, \ldots, n$.

FIGURE GOES HERE

Proof. We define $h\colon \mathbb{R} \longrightarrow \mathbb{R}$ by

$$h(t) \equiv f(\mathbf{x}^* + t\mathbf{v})$$

for any $\mathbf{v} \in \mathbb{R}^n$, $t \in \mathbb{R}$.
Take the case of maximizer:
Fix a direction $\mathbf{v}$ ($\|\mathbf{v}\| \neq 0$). We have

$$f(\mathbf{x}^*) \geq f(\mathbf{x}),$$

$\forall \mathbf{x} \in B_\delta(\mathbf{x}^*)$, for some $\delta > 0$.

In particular for t small ($t < \delta \, \|\mathbf{v}\|$) we have

$$\begin{aligned} f(\mathbf{x}^* + t\mathbf{v}) &= h(t) \\ &\leq f(\mathbf{x}^*) \end{aligned}$$

Thus, h is maximized locally by $t^* = 0$.
Our F.O.C. from the $\mathbb{R} \longrightarrow \mathbb{R}$ case

$$\implies h'(0) = 0$$

So by the chain rule

$$\implies \nabla f(\mathbf{x}^*) \cdot \mathbf{v} = 0$$

And since this must hold for every $\mathbf{v} \in \mathbb{R}^n$, this implies that

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

□

Note We call $\mathbf{x}^*$ satisfying $Df(\mathbf{x}^*) = \mathbf{0}$ a *critical point*.

The theorem mimics the one variable theorem. The proof works like this. If x^* is a local maximum, then the one variable function you obtain by restricting x to move along a fixed line through x^* (in the direction v) also must have a local maximum. Hence that function of one variable ($f(x^* + hv)$ treated as a function of h) must have derivative zero when $h = 0$. The derivative of this function is simply the directional derivative of f in the direction v. If all of the directional derivatives are zero and the function is differentiable, then the derivative must be zero.

Just like the one-variable case, the first-derivative test cannot distinguish between local minima and and local maxima, but an examination of the proof tells you that at local maxima derivatives decrease in the neighborhood of a critical point. Just like the one-variable case, critical points may fail to be minima or maxima. In the one variable case, this problem arises essentially for one reason: a function decrease if you reduce x (suggesting a local maximum) and increase if you increase x (suggesting a local minimum). In the many variable case, this behavior could happen in any direction. Moreover, it could be that the function restricted to direction has a local maximum, but it has a local minimum with respect to another direction. At this point, the central insight is that it is much less likely than a critical point of a multivariable function is a local extremum than in the one variable case.

We want *necessary* and *sufficient* conditions.

11.3 Second Order Conditions

For a critical point $\mathbf{x}^*$ we have by Taylor's Theorem that

$$\frac{f(\mathbf{x}^* + \mathbf{h}) - f(\mathbf{x}^*)}{\|\mathbf{h}\|^2} = \frac{\frac{1}{2}\mathbf{h}^t D^2 f(\mathbf{z})\mathbf{h}}{\|\mathbf{h}\|^2}$$

We could write $\mathbf{h} = t\mathbf{v}$, where $\|\mathbf{v}\| = 1$. So

$$\begin{aligned}\frac{f(\mathbf{x}^* + t\mathbf{v}) - f(\mathbf{x}^*)}{t^2} &= \frac{\frac{1}{2}t^2\mathbf{v}^t D^2 f(\mathbf{z})\mathbf{v}}{t^2} \\ &= \frac{1}{2}\mathbf{v}^t D^2 f(\mathbf{z})\mathbf{v}\end{aligned}$$

for $\mathbf{z} \in B_{\|\mathbf{h}\|}(x^*)$

The generalization of second-order conditions follows from this expression. If $\mathbf{x}^*$ is a critical point and the quadratic form $\mathbf{v}^t D^2 f(\mathbf{z})\mathbf{v}$ is negative definite, then continuity implies that $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}$ is negative definite and hence $\mathbf{x}^*$ is a local maximum. Conversely (almost conversely), if $\mathbf{x}^*$ is a local maximum, then it must be that $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}$ is negative semi definite. Finally note that if the quadratic form $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}$ is indefinite, then the critical point $\mathbf{x}^*$ cannot be either a local maximum or a local minimum because there are directions in which the function decreases near $\mathbf{x}^*$ and directions in which the function increases near $\mathbf{x}^*$.

FIGURE GOES HERE

What follows is an informal summary.

11.3.1 S.O. Sufficient Conditions

If f is twice continuously differentiable and $\mathbf{x}^*$ is a critical point of f, then:

1. $\mathbf{x}^*$is a local minimizer whenever the quadratic form $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}$ is a positive definite.

2. $\mathbf{x}^*$is a local maximizer whenever the quadratic form $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}$ is negative definite.

11.3.2 S.O. Necessary Conditions

If $D^2 f(\mathbf{x}^*)$ exists, then:

1. if $\mathbf{x}^*$ is a local maximizer, then $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}$ is a negative semi-definite quadratic form.

2. if $\mathbf{x}^*$ is a local minimizer, then
   $\mathbf{v}^t D^2 f(\mathbf{x}^*)\mathbf{v}^t$ is a positive semi-definite quadratic form.

A consequence of the necessary conditions is that when the second-derivative matrix gives rise to an indefinite quadratic form a critical point cannot be either a maximum or a minimum.

Also recall our definitions of concave up and concave down.
We also have that

- $D^2 f(\mathbf{x}^*)$ positive definite everywhere $\Leftrightarrow$ f concave up everywhere
- $D^2 f(\mathbf{x}^*)$ negative definite everywhere $\Leftrightarrow$ f concave down everywhere

Example 48. *Let* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *be*

$$f(x, y) = 4x^3 + y^2 - 6xy + 6x$$

Now (x^*, y^*) *is our max. We have then as usual that*

$$\nabla f(x^*, y^*) = \boldsymbol{0}$$

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 12x^2 - 6y + 6 \\ &= 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial y}(x, y) &= 2y - 6x \\ &= 0 \end{aligned}$$

from the second equation we get

$$2y = 6x$$

And combining this with the first equation we get

$$\begin{aligned} & 12x^2 - 18x + 6 = 0 \\ \Rightarrow \quad & 2x^2 - 6x + 1 = 0 \\ \Rightarrow \quad & (2x - 1)(x - 1) = 0 \end{aligned}$$

gives us

$$\begin{aligned} x = \tfrac{1}{2} \quad & \textit{or} \quad x = 1 \\ x = \tfrac{1}{2} \quad & \Rightarrow \quad y = \tfrac{3}{2} \\ x = 1 \quad & \Rightarrow \quad y = 3 \end{aligned}$$

So we have 2 critical points

$$(\tfrac{1}{2}, \tfrac{3}{2}), (1, 3)$$

$$\begin{aligned} Df(\boldsymbol{x}) &= \nabla f(\boldsymbol{x}) \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \end{aligned}$$

$$(Df)' = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$$

$$D^2 f = D(Df)'$$

$$\begin{aligned} D^2 f &= \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial x^2} \end{pmatrix} \\ &= \begin{pmatrix} 24x & -6 \\ -6 & 2 \end{pmatrix} \end{aligned}$$

So

$$D^2 f(\tfrac{1}{2}, \tfrac{3}{2}) = \begin{pmatrix} 12 & -6 \\ -6 & 2 \end{pmatrix}$$

And we note that the first entry in this matrix (i.e. the a entry) is positive! It can also be seen that

$$\left| D^2 f(\tfrac{1}{2}, \tfrac{3}{2}) \right| = -12 < 0$$

And from our rule for quadratic forms this gives us an indefinite quadratic form. So the point $(\frac{1}{2}, \frac{3}{2})$ is neither a maximizer nor a minimizer.
So

$$D^2 f(1, 3) = \begin{pmatrix} 24 & -6 \\ -6 & 2 \end{pmatrix}$$

And we note that the first entry in this matrix (i.e. the a entry) is positive! It can also be seen that

$$\left| D^2 f(\tfrac{1}{2}, \tfrac{3}{2}) \right| = 12 > 0$$

And from our rule for quadratic forms this gives us an positive definite quadratic form. So the point $(1, 3)$ is a local maximizer.

Chapter 12

Invertibility and Implicit Function Theorem

The standard motto is: Whatever you know about linear functions is true *locally* about differentiable functions. This section discusses two useful properties that can be understood in these terms.

12.1 Inverse Functions

First, let's review invertibility for functions of one variable. In the linear case, $f(x) = ax$, so the function is invertible if $a \neq 0$. Notice that if a linear function is invertible at a point, it is invertible everywhere. More generally, we have this definition and result:

Definition 114. *We say the function $f\colon \mathcal{X} \longrightarrow \mathcal{Y}$ is one-to-one if*

$$f(x) = f(x') \quad \Longrightarrow \quad x = x'$$

We have already defined $f^{-1}(\mathcal{S}) = \{x \in \mathcal{X} \mid f(x) \in \mathcal{S}\}$ for $\mathcal{S} \subset \mathcal{Y}$. Now let's consider $f^{-1}(y)$ for $y \in \mathcal{Y}$. Is f^{-1} a function?

If f is one-to-one, then $f^{-1}\colon f(\mathcal{X}) \longrightarrow \mathcal{X}$ is a function.

FIGURE GOES HERE

f is generally not invertible as the inverse is not one-to-one. But in the neighborhood (circle around a point x_0), it may be strictly increasing so it is one-to-one and therefore invertible.

Theorem 64. *If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is C^1 and*

$$f'(x_0) \neq 0,$$

then $\exists\, \epsilon > 0$ such that f is strictly monotone on the open interval $(x_0 - \epsilon, x_0 + \epsilon)$.

Note this is just saying that if $f'(x_0)$ is positive and f' is continuous, then it is remains positive over an interval.

So f is locally invertible at x_0, then we can define g on $(x_0 - \epsilon, x_0 + \epsilon)$ such that

$$g(f(x)) = x$$

That is, in the one-variable case, linear functions with (constant) derivative not equal to zero are invertible. Differentiable functions with derivative not equal to zero at a point are invertible locally. For one variable functions, if the derivative is always non zero, then the inverse can be defined on the entire range of the function. When you move from functions from $\mathbb{R}^n$ to itself, you can ask whether inverse functions exist. Linear functions can be represented as multiplication by a square matrix. Invertibility of the function is equivalent to inverting the matrix. So a linear function is invertible (globally) if its matrix representation is invertible.

Definition 115. *The function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is locally invertible at x_0 if there is a $\delta > 0$ and a function $g : B_\delta(f(x_0)) \longrightarrow \mathbb{R}^n$ such that $f \circ g(y) \equiv y$ for $y \in B_\delta(f(x_0))$ and $g \circ f(x) \equiv x$ for $x \in B_\delta(x_0)$.*

The definition asserts the existence of an inverse defined in a (possibly small) set containing $f(x_0)$.

Theorem 65. *If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is differentiable at x_0 and $Df(x_0)$ is invertible, then f is locally invertible at x_0. Moreover, the inverse function, g is differentiable at $f(x_0)$ and $Dg(f(x_0)) = (Df(x_0))^{-1}$.*

The theorem asserts that if the linear approximation of a function is invertible, then the function is invertible locally. In contrast to the one variable case, the

assumption that Df is globally invertible does not imply the existence of a global inverse.

Recall Theorem 34 that says

$$g'(y^0) = \frac{1}{f'(x^0)}$$

so the formula for the derivative of the inverse generalizes the one-variable formula.

The proof of the inverse function theorem is hard. One standard technique involves methods that are sometimes used in economics, but the details are fairly intricate and not worth our time.

The problem of finding an inverse suggests a more general problem. Suppose that you have a function $G : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}^n$. Let $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. We might be interested in whether we can solve the system of equations: $G(x, y) = 0$. This is a system of n equations in $n + m$ variables. You might hope therefore that for every choice of y you could solve the equation. That is, you might search for a solution to the equation that gives x as a function of y. The problem of finding an inverse is really a special case where $n = m$ and $G(x, y) = f(x) - y$.

The general case is an important problem in economics. In the typical application, the system of equations characterize an economic equilibrium. Maybe they are the equations that determine market clearing price. Maybe they are the first-order conditions that characterize the solution to an optimization problem. The y variables are parameters. You "solve" a model for a fixed value of the parameters and you want to know what the solution to the problem is when the parameters change by a little bit. The implicit function theorem is a tool for analyzing the problem. This theorem says that (under a certain condition), if you can solve the system at a give y_0, then you can solve the system in a neighborhood of y_0. Furthermore, it gives you expressions for the derivatives of the solution function.

Why call this the implicit function theorem? Life would be great if you could write down the system of equations and solve them to get an explicit representation of the solution function. If you can do this, then you can exhibit the solution (so existence is not problematic) and you can differentiate it (so you do not need a separate formula for derivatives). In practice, you may not be able to find an explicit form for the solution. The theorem is the next best thing.

We will try to illustrate the ideas with some simple examples.

12.2 Implicit Functions

$$f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$$

Suppose $f(x, z) = 0$ is an *identity* relating x and z.

So we are looking for how z depends on x at this value of f which is attained with a certain value of x. This is really the point!

Example 49.

$$f(x, z) = x^3 - z = 0$$

FIGURE GOES HERE

So there is a pair x and z which fits this identity (actually in this example there are an infinite number of them)

In this example it is easy to pull out the z to one side and solve for it in terms of x but supposing you had a much more complicated example like

$$x^2 z - z^2 + \sin x \ln z + \cos x = 0,$$

then this is not as easy to just find z in terms of x.

The important point is that even this messy function is still expressing z in terms of x, although it may be difficult solve for this explicitly - but there are ways around this problem.

Not only might x_0 correspond to a few z-values, but also z_0 might correspond to a few x-values.
But if we examine (x_0, z_0) in a small neighborhood where the function is increasing...

FIGURE GOES HERE

Using the picture, we can see that increasing the x value slightly, the function

no longer takes the value of zero, so you can see that we are no longer on the level set!

$$\triangle x \frac{\partial f}{\partial x} = -\triangle z \frac{\partial f}{\partial z}$$

$$\frac{\triangle z}{\triangle x} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}$$

$$\frac{\partial z}{\partial x} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}$$

So if f gives us a function

$$g\colon (x_0 - \epsilon, x_0 + \epsilon) \longrightarrow (z_0 - \epsilon, z_0 + \epsilon)$$

such that

$$f(x, g(x)) = 0,$$

$\forall\, x \in (x_0 - \epsilon, x_0 + \epsilon)$
then we can define

$$h\colon (x_0 - \epsilon, x_0 + \epsilon) \longrightarrow \mathbb{R}$$

by

$$h(x) = f(x, g(x))$$

So the point of this is that if $h(x) = 0$ for any x, then $h'(x) = 0$ for any x. But the clever part is that $h'(x)$ can be written as a function of $f'(x)$ and $g'(x)$ which are not necessarily zero.

If f and g are differentiable, we calculate

$$y = \begin{pmatrix} x \\ z \end{pmatrix}, \quad G(x) = \begin{pmatrix} x \\ g(x) \end{pmatrix}$$

$$h(x) = [f \circ G](x)$$

$$\begin{aligned} h'(x) &= Df(x_0, z_0)DG(x^0) \\ &= \left(\frac{\partial f}{\partial x}(x_0, z_0),\ \frac{\partial f}{\partial z}(x_0, z_0)\right) \cdot \begin{pmatrix} 1 \\ g'(x^0) \end{pmatrix} \\ &= \frac{\partial f}{\partial x}(x_0, z_0) + \frac{\partial f}{\partial z}(x_0, z_0)g'(x_0) \\ &= 0 \end{aligned}$$

This gives us

$$g'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, z_0)}{\frac{\partial f}{\partial z}(x_0, z_0)}$$

Now we will look at the more general case:

$$f\colon \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

$$\begin{aligned} f(x, \mathbf{z}) &= \begin{pmatrix} f_1(x, \mathbf{z}) \\ f_2(x, \mathbf{z}) \\ \vdots \\ f_m(x, \mathbf{z}) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\ &= \mathbf{0} \end{aligned}$$

where $x \in \mathbb{R}$, and $\mathbf{z} \in \mathbb{R}^m$.
And we have

$$g\colon \mathbb{R} \longrightarrow \mathbb{R}^m$$

$$\mathbf{z} = g(x)$$

Theorem 66. *Suppose*

$$f\colon \mathbb{R} \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

is C^1 and write $F(x, \boldsymbol{z})$ where $x \in \mathbb{R}$ and $\boldsymbol{z} \in \mathbb{R}^m$.
Assume

$$|D_z f(x_0, \boldsymbol{z}_0)| = \begin{vmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_m} \end{vmatrix} \neq 0$$

and

$$f(x_0, \boldsymbol{z}_0) = \boldsymbol{0}.$$

There exists a neighborhood of (x_0, z_0) and a function $g\colon \mathbb{R} \longrightarrow \mathbb{R}^m$ defined on the neighborhood of x_0, such that $\boldsymbol{z} = g(x)$ uniquely solves $f(x, \boldsymbol{z}) = \boldsymbol{0}$ on this neighborhood.
Furthermore the derivatives of g are given by

$$\underset{m\times 1}{Dg(x_0)} = -\underset{m\times m}{[D_z f(x_0, \boldsymbol{z}_0)]^{-1}} \underset{m\times 1}{D_x f(x_0, \boldsymbol{z}_0)}$$

We said that the inverse function theorem was too hard to prove here. Since the implicit function theorem is a generalization, it too must be too hard to prove. It turns out that the techniques one develops to prove the inverse function theorem can be used to prove the implicit function theorem, so the proof is not much harder. It also is that case that the hard thing to prove is the existence of the function g that gives z in terms of x. If you assume that this function exists, then computing the derivatives of g is a simple application of the chain rule. The following proof describes this argument.

Proof. So we have

$$f(x, g(x)) = 0$$

And we define

$$H(x) \equiv f(x, g(x))$$

And thus

$$\begin{aligned} D_x H(x) &= D_x f(x_0, \mathbf{z}_0) + D_z f(x_0, \mathbf{z}_0) D_x g(x_0) \\ &= \mathbf{0} \end{aligned}$$

$$\implies D_z f(x_0, \mathbf{z}_0) D_x g(x_0) = -D_x f(x_0, \mathbf{z}_0)$$

Multiply both sides by the inverse:

$$\underbrace{[D_z f(x_0, \mathbf{z}_0)]^{-1} \cdot [D_z f(x_0, \mathbf{z}_0)]}_{=\mathbf{I}_m} \cdot D_x g(x_0) = -[D_z f(x_0, \mathbf{z}_0)]^{-1} D_x F(x_0, \mathbf{z}_0)$$

$$\implies D_x g(x_0) = -[D_z f(x_0, \mathbf{z}_0)]^{-1} D_x F(x_0, \mathbf{z}_0)$$

□

The implicit function theorem thus gives you a guarantee that you can (locally) solve a system of equations in terms of parameters. As before, the theorem really is a local version of a result about linear systems. The theorem tells you what happens if you have one parameter. The same theorem holds when you have many parameters $x \in \mathbb{R}^n$ rather than $x \in \mathbb{R}$.

Theorem 67. *Suppose*

$$f\colon \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$$

is C^1 and write $F(\boldsymbol{x}, \boldsymbol{z})$ where $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{z} \in \mathbb{R}^m$.
Assume

$$|D_z f(\boldsymbol{x}_0, \boldsymbol{z}_0)| = \begin{vmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial z_1} & \cdots & \frac{\partial f_m}{\partial z_m} \end{vmatrix} \neq 0$$

and

$$f(\boldsymbol{x}_0, \boldsymbol{z}_0) = \boldsymbol{0}.$$

There exists a neighborhood of $(\boldsymbol{x}_0, \boldsymbol{z}_0)$ and a function $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined on the neighborhood of $\boldsymbol{x}_0$, such that $\boldsymbol{z} = g(\boldsymbol{x})$ uniquely solves $f(\boldsymbol{x}, \boldsymbol{z}) = \boldsymbol{0}$ on this neighborhood.
Furthermore the derivatives of g are given by implicit differentiation (use chain rule)

$$\underset{m\times n}{Dg(\boldsymbol{x}_0)} = -[\underset{m\times m}{D_z f(\boldsymbol{x}_0, \boldsymbol{z}_0)}]^{-1} \underset{m\times n}{D_x f(\boldsymbol{x}_0, \boldsymbol{z}_0)}$$

Verifying that this is the correct formula for the derivative is just the chain rule.

Comments:

1. It is useful to keep track of the dimensions of the various matrices. Keep in mind the intuitive idea that you usually need exactly m variables to solve m equations is helpful. This means that if the domain has n extra dimensions, typically you will have n parameters – the solution function will go from $\mathbb{R}^n$ into $\mathbb{R}^m$.

2. The implicit function theorem proves that a system of equations has a solution **if** you already know that a solution exists at a point. That is, the theorem states that if you can solve the system once, then you can solve it locally. The theorem does not guarantee existence of a solution. (This is where the general case is much less informative than the linear case, where the invertibility assumption in the theorem would be sufficient to show existence.)

3. The theorem provides an explicit formula for the derivatives of the implicit function. One could memorize the formula and apply it when needed. That would be wrong. It turns out that you can compute the derivatives of the implicit function by "implicitly differentiating" the system of equations and solving the resulting equations by explicitly. An example follows.

12.3 Examples

Here is a simple economic example of a comparative-statics computation. A monopolist produces a single output to be sold in a single market. The cost to produce q units is $C(q) = q + .5q^2$ dollars and the monopolist can sell q units for the price of $P(q) = 4 - \frac{q^5}{6}$ dollars per unit. The monopolist must pay a tax of one dollar per unit sold.

1. Show that the output $q^* = 1$ that maximizes profit (revenue minus tax payments minus production cost).

2. How does the monopolist's output change when the tax rate changes by a small amount?

The monopolist picks q to maximize:

$$q(4 - \frac{q^5}{6}) - tq - q - .5q^2.$$

The first-order condition is

$$q^5 + q - 3 + t = 0$$

and the second derivative of the objective function is $-5q^4 - 1 < 0$, so there is at most one solution to this equation, and the solution must be a (global) maximum. Plug in $q = 1$ to see that this value does satisfy the first-order condition when $t = 1$. Next the question asks you to see how the solution $q(t)$ to:

$$q^5 + q - 3 + t = 0$$

varies as a function of t when t is close to one. We know that $q(1) = 1$ satisfies the equation. We also know that the left-hand side of the equation is increasing in q, so the condition of the implicit function theorem holds. Differentiation yields:

$$q'(t) = -\frac{1}{5q^4 + 1}. \tag{12.1}$$

In particular, $q'(1) = -\frac{1}{6}$.

In order to obtain the equation for $q'(t)$ (12.1), you could use the general formula or you could differentiate the identity:

$$q(t)^5 + q(t) - 3 + t \equiv 0$$

with respect to one variable (t) to obtain

$$5q(t)q'(t) + q'(t) + 1 = 0,$$

and solve for $q'(t)$. Notice that the equation is linear in q'. This technique of "implicit differentiation" is fully general. In the example you have $n = m = 1$ so there is just one equation and one derivative to find. In general, you will have an identity in n variables and m equations. If you differentiate each of the equations with respect to a fixed parameter, you will get m linear equations for the derivatives of the m implicit functions with respect to that variable. The system will have a solution if the invertibility condition in the theorem is true.

12.4 Envelope Theorem for Unconstrained Optimization

We are interested in

$$V(a) = \max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}, a)$$

for $a \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathcal{X} \subset \mathbb{R}^n$ open.

We will not concern ourselves with conditions under which the maximum exists (we cannot apply the standard existence result when $\mathcal{X}$ is open), but just focus on examples where the max value of f does exist i.e. $V(a)$ exists, and we will ask the question what happens to $V(a)$ as a changes. i.e. What is $V'(a)$? Suppose that we can find a function $g : \mathbb{R} \longrightarrow \mathbb{R}^n$ such that $g(a)$ maximizes $f(x, a)$. We have $V(a) = f(g(a), a)$. That is, the value function V is a real-valued function of a real variable. We can compute its derivative using the chain rule (assuming that f and

g are differentiable). The implicit function theorem gives us a sufficient condition for g to be differentiable and a formula for the derivative provided that solutions to the optimization problem are characterized as solution to the first order condition. That is, if x^* solves $\max_{x \in \mathcal{X}} f(x, a^*)$ if and only if $D_{\mathbf{x}} f(x^*, a^*) = 0$, then x^* is implicitly defined as a solution to a system of equations. Applying the chain rule mechanically we have

$$\underset{1\times 1}{V'(a^*)} = \underset{1\times n}{D_{\mathbf{x}} f(g(a^*), a^*)} \underset{n\times 1}{Dg(a^*)} + \underset{1\times 1}{D_a f(g(a^*), a^*)}.$$

The implicit function theorem tells us when $Dg(a^*)$ exists and gives us a formula for the derivative. In order to evaluate V', however, we only need to know that the derivative exists. This is because at an interior solution to the optimization problem $D_{\mathbf{x}} f(g(a^*), a^*) = \mathbf{0}$. It follows that the change in the value function is given by $V'(a^*) = D_a f(x^*, a^*)$, where $x^* = g(a^*)$ is the solution to the optimization problem at a^*.

The one condition that needs to be checked is the non-singularity condition in the statement of the implicit function theorem. Here the condition is that the matrix of second derivatives of f with respect to the components of $\mathbf{x}$ is non-singular. We often make this assumption (in fact, we assume that the matrix is negative definite), to guarantee that solutions to first-order conditions characterize a maximum.

The result that the derivative of the value function depends only on the partial derivative of the objective function with respect to the parameter maybe somewhat surprising. It is called the envelope theorem (you will see other flavors of envelope theorem).

Suppose that f is a profit function and a is a technological parameter. Currently a^* describes the state of the technology. The technology shifts. You shift your production accordingly (optimizing by selecting $x = g(a)$). Your profits shift. How can you tell whether the technological change makes you better off? Changing the technology has two effects: a direct one that is captured by $D_a f$. For example, if $D_a f(x^*, a^*) < 0$ this means that increasing a leads to a decrease in profits if you hold x^* fixed. The other effect is indirect: you adjust x^* to be optimal with respect to the new technology. The envelope theorem says that locally the first effect is the only thing that matters. If the change in a makes the technology less profitable for fixed x^*, then you will not be able to change x^* to reverse the negative effect. The reason is, loosely, that when you optimize with respect to x and change the technology, then x^* is still nearly optimal for the new problem. Up to "first order" you do not gain from changing it. When $D_a f(x^*, a^*) \neq 0$, changing a leads to a first-order change in profits that dominates the effect of reoptimizing.

The result is called the envelope theorem because (geometrically) the result says that $V(a)$, the value function, is tangent to a family of functions $f(x, a)$ when $x = g(a)$. On the other hand, $V(a) \geq f(x, a)$ for all a, so the V curve looks like an "upper envelope" to the f curves.

For a presentation of value functions that is both more simple and coherent than this, you can read chapters 26 and 27 of Varian. It is a nice treatment since it does lots of simpler examples of functions of only 1 and 2 variables, as opposed to the general n variable case above. Really only chapter 27 deals with value functions but chapter 26 is good background reading too.

Chapter 13

Constrained Optimization

Economic theory is all about solving constrained optimization problems. Consumers maximize utility subject to a budget constraint. Firms minimize cost or maximize profit subject to their outputs being technologically feasible. Governments may maximize social welfare subject to constraints on consumer's independence. The theory of unconstrained optimization is easy because at an interior optima it must be that directional derivatives in all directions are zero. This follows from a simple argument. If you are at a maximum, then moving "forward" in any direction must not increase the objective function. So a directional derivative must be less than or equal to zero. However, moving "backward" in that direction must also lower the function's value, so the directional derivative must be equal to zero. The same insight applies in the constrained case, but since movement in some directions might be ruled out by constraints, first-order conditions come in the form a set of inequalities that hold for all directions that are not ruled out by constraints.

13.1 Equality Constraints

The basic problem is:

$$\max f(x) \tag{13.1}$$

$$x \in S \tag{13.2}$$

Here $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. The function f in (13.1) is called the objective function. We study maximization problems, but since minimizing a function h is the same as maximizing $-h$, the theory covers minimization problems as well. (13.2) is the constraint set.

Recall that a solution to the problem is x^* such that (a) $x^* \in S$ and (b) if $x \in S$, then $f(x^*) \geq f(x)$.

Calculus does not provide conditions for the existence of an optima (real analysis does, because continuous real-valued functions on compact sets attain their maxima). We do not discuss existence in this section. Instead we concentrate on characterizing optima in terms of equations. When S is an open set, any maxima must be interior. Hence $Df = 0$. In general, we must modify this condition.

We begin with the study of equality constrained maximization. That is, assume that there are a set of functions $g_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $S = \{x : g_i(x) = b_i\}$ for all i. Sometimes we restrict g_i to be linear functions (so that they can be written $g_i(x) = a_i \cdot x$). In this case, it is important to have arbitrary constants on the right-hand side of the constraint. When g_i is an arbitrary function, then we can assume without loss of generality that $b_i = 0$ because otherwise we can replace g_i by the function $g_i - b$.

If your problem had just one constraint, then there is an intuitive way to solve it. Use the constraint to write one variable in terms of the others. Substitute out for that variable and solve the unconstrained problem. This method really works as long as it is possible to use the constraint to write one variable in terms of the others. It is also completely general.

Let $G = (g_1, \ldots, g_m)$ so that there are m constraints. $G : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Typically, the problem is uninteresting if $m \geq n$ (because then the constraint set is likely to contain only one point or be empty).[1]

We say that the constraints are non-degenerate (or satisfy a regularity condition or a constraint qualification) at a point x^* if $DG(x^*)$ has full rank ($\{\nabla g_i(x^*)\}$ is a collection of m linearly independent vectors).

To spell this out, divide the variable $x^* = (u^*, w^*)$ where $u^* \in \mathbb{R}^{n-m}$ and $w^* \in \mathbb{R}^m$ and $D_w G(x^*) \neq \mathbf{0}$. The nondegeneracy condition guarantees that there are m variables you can select so that $D_w G(x^*)$ is nonsingular. This means that by the implicit function theorem it is possible to solve the equations $G(x^*) = 0$ for a function $W : \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^m$ such that $W(u^*) = w^*$ and $G(u^*, W(u^*)) = \mathbf{0}$ in a neighborhood of (u^*, w^*). With this observation, solving the constrained optimization problem is equivalent to solving the unconstrained optimization problem:

$$\max f(u, W(u)).$$

In particular, if x^* solves the original problem, then u^* must be a local maximum of the unconstrained problem. Hence the composite function $h(u) = f(u, W(u))$ has a critical point at u^*. Using the chain rule we see that $h'(u^*) = D_u f(x^*) +$

[1] $m \geq n$ is perfectly fine for the case of inequality constraints.

$D_w f(x^*)DW(u^*)$. Furthermore, the implicit function theorem gives you a formula for $DW(u^*)$. It follows that

$$0 = h'(u^*) = D_u f(x^*) - D_w f(x^*) \left(D_w G(x^*)\right)^{-1} D_u G(x^*) \tag{13.3}$$

or

$$D_u f(x^*) = D_w f(x^*) \left(D_w G(x^*)\right)^{-1} D_u G(x^*). \tag{13.4}$$

Let

$$\lambda = D_w f(x^*) \left(D_w G(x^*)\right)^{-1}$$

or

$$D_w f(x^*) = \lambda D_w G(x^*). \tag{13.5}$$

It follows that from equation (13.4) that

$$D_u f(x^*) = \lambda D_u G(x^*),$$

which combined with (13.5) we can write

$$Df(x^*) = \lambda DG(x^*).$$

We summarize the result in the following result.

Theorem 68. *If x^* solves $\max f(x)$ subject to $G(x) = 0$ then there exists λ such that $Df(x^*) = \lambda DG(x^*)$.*

It is standard to write the equation

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*).$$

The λ_i are often called Lagrange Multipliers.

The beauty of first-order conditions is that they replace solving a maximization problem with finding solutions to a system of equations involving first derivatives. This result demonstrates that the result extends to equality constrained optimization problems. An interesting aspect of the construction is that it provides extra information: the multipliers λ have economic interpretation.

Imagine that the constraints of the original problem are $G(x) = b$ and we can solve the problem at x^* when $G(x^*) = \mathbf{0}$. What happens to the solution of the problem as b changes? In particular, is it possible to compute the derivatives of the value function: $V(b) \equiv f(x^*(b))$ where $x^*(b)$ solves: $\max f(x)$ subject to $G(x) = b$. In the same regularity condition we needed for the theorem holds (there is a subset of m variables – denoted by w – such that $D_w G$ is invertible at

x^*), then the implicit function theorem guarantees that we can solve $G(x) = b$ for W as a function of u and b in a neighborhood of $(u, b) = (u^*, 0)$. Call the function that satisfies the first-order condition $x^*(b) = (u^*(b), w^*(b))$ If we carry out the substitution and computation as before we can differentiate $V(b)$ to obtain:

$$DV(0) = \left(\left(D_u f(x^*) + D_w f(x^*)\right) \left(D_u W(x^*, 0)\right) Du^*(0) + D_w f(x^*) D_b W(x^*, 0).\right.$$

The first two terms on the right-hand side are precise those from the first-order condition (13.3) multiplied by $Du^*(0)$. There terms are zero (as in the case of the unconstrained envelope theorem). The final term on the right is the contribution that comes from the fact that W may change with b. We can use the implicit function theorem to get:

$$D_w f(x^*) D_b W(x^*, 0) = D_w f(x^*) \left(D_w G(x^*)\right)^{-1} \mathbf{I} = \lambda,$$

where the last equation is just the definition of λ. Hence we have the following equality constrained envelope theorem:

$$DV(\mathbf{0}) = \lambda. \tag{13.6}$$

This equation provides a nice interpretation of the Lagrange multipliers. The question is: How does the value of the problem change if the right-hand side of the ith constraint increases from 0? The answer is λ_i. That is, the Lagrange multipliers give the incremental value of a change in the available resources on the value of the objective function. Loosely speaking, the decision maker is willing to pay λ_i to have one unit more of the ith resources. When the constraints are equations you cannot state in advance whether an increase in b_i will raise or lower the value of the problem. That is, you cannot state in advance the sign of λ_i. There is a variation of this result that holds in inequality constrained problems. In that case there is a definite sign to λ_i (whether the sign is positive or negative depends on how you write the constraints).

13.2 The Kuhn-Tucker Theorem

This result generalizes the equality constrained case to the general case of inequality constraints. That is, we assume that the constraint set can be written $S = \{x : g_i(x) \geq b_i, i \in I\}$.

The intuition is fairly simple. We know how to solve problems without constraints. We know how to solve problems with equality constraints. If you have inequality constraints and you think that x^* is an optimum, then (locally) the constraints will divide into the group that are satisfied as equations (binding or active

constraints) and the others. The first group we treat as we would in an equality constrained problem. The second group we treat as in an inequality constrained problem. Hence we would expect a first-order condition like the one in the previous section:

$$\nabla f(x^*) = \sum_{i \in J} \lambda_i \nabla g_i(x^*) \tag{13.7}$$

where J is the set of binding constraints. Under appropriate conditions, this turns out to be the appropriate first-order necessary condition for optimality. The condition is often written as a pair of conditions:

$$\nabla f(x^*) = \sum_{i \in I} \lambda_i \nabla g_i(x^*) \tag{13.8}$$

and

$$\lambda_i g(x_i^*) = 0 \text{ for all } i \in I. \tag{13.9}$$

Equation (13.8) differs from equation (13.7) because it includes all constraints (not just the binding ones). Equation (13.8) and (13.9) are equivalent to (13.7). Equation (13.9) can only be satisfied (for a given i) if either $\lambda_i = 0$ or $g_i(x^*) = 0$. That is, it is a clever way of writing that either a constraints is binding or its associated multiplier is zero. Once we know that the multipliers associated with non-binding constraints are zero, it is clear that equations (13.7) and (13.8) are equivalent.

Equation (13.9) is called a complementary slackness condition. (Constraints that do not bind are said to have slack.) It provides a weak interpretation of Langrange multipliers are "prices" or values of resources represented in each constraint. It is useful to think of multiplier λ_i as the amount that value would increase if the ith constraint were relaxed by a unit or, alternatively, how much the person solving the problem would pay to have an "extra" unit to allocate on the ith constraint. If a constraint does not bind, it should be the case that this value is zero (why would you pay to relax a constraint that is already relaxed?). Also, if you are willing to pay to gain more resource ($\lambda_i = 0$) it must be the case that the constraint is binding. It is possible (but rare) to have both $\lambda_i = 0$ and $g_i(x^*) = 0$. This happens when you have "just enough" of the ith resource for your purposes, but cannot use anymore.

This treatment follows the exposition of Harold Kuhn "Lectures on Mathematical Economics" (Danzig and Veinot (ed.), Lectures on Applied Mathematics, Volume 2).

To begin the study, we need a notion of a direction that is not ruled out by constraints.

Definition 116. *The vector v enters S from x^* if there exists $\epsilon > 0$ such that $x^* + tv \in S$ for all $t \in [0, \epsilon)$.*

If $x^* \in S$ and v enters S from x^*, then it is possible to start at x^*, go in the direction v, and stay in the feasible set. It follows that if x is a local maximum, then $D_v f(x^*) \geq 0$. To fully describe the set of first-order conditions for a constrained optimization problem, we need to figure out a nice way to describe which vectors enter S at x and try to get a clean way of summarizing the resulting inequalities.

Definition 117. *S is described by linear inequalities if there is a matrix $\mathbf{A}$ such that $S = \{x : Ax \geq b\}$.*

A general constraint set will be written in the form $S = \{x : g_i(x) \geq b_i, i \in I\}$. If the constraint set is described by linear inequalities, then we can take all of the g_i to be linear functions (that is, $g_i(x) = a_i \cdot x$, where a_i is the ith row of the matrix $\mathbf{A}$). Note that a set described by inequalities includes an equality constrained set as a special case (to impose the constraint $g(x) = b$ substitute two inequalities: $g(x) \geq b$ and $-g(x) \geq -b$). The converse is not true: Equality constrained problems are typically easier than inequality constrained problems.

Definition 118. *An inward pointing normal to the boundary of the set $S = \{x : g_i(x) \geq b_i, i = 1, \ldots, I\}$ at x^* is a direction of the form $\nabla g_i(x^*)$ where $g_i(x^*) = b_i$.*

When $g_i(x^*) = b_i$ the ith constraint is *binding*. If you move from x^* in the "wrong" direction, then the constraint will no longer be satisfied. However, if you move "into" S, then the constraint will be satisfied. When S is described by linear inequalities, then the inward directions are simply the a_i associated with binding constraints.

Theorem 69. *Suppose that S is described by linear inequalities. v enters S from x^* if and only if v makes a nonnegative inner product with all of the inward pointing normals to the boundary of the feasible set at x^*.*

Proof. Let $x^* \in S$ and let $a_k \cdot x^* = 0$ for $k \in J$ (J may be empty). Let v enter S from x^*. It follows that there exists $\epsilon > 0$ such that $a_k \cdot (x^* + tv) \geq 0$ for $k \in I$ and $t \in [0, \epsilon)$. In particular, if $i \in J$, then $a_k \cdot x^* = b_k$ and hence

$$a_k \cdot v \geq 0 \text{ for all } k \in J.$$

Conversely, if $a_k \cdot v \geq 0$ for all $k \in J$. For $k \notin K$, it must be that $a_k \cdot x^* > b_k$. Therefore $a_k (x^* + tv) > b_k$ for sufficiently small t. For $k \in J$ we have $a_k \cdot x^* = b_k$ so that $a_k \cdot (x^* + tv) \geq b_k$. This means that v enters S at x^*. □

Now that we have a characterization of what it means for v to enter S, we can get a condition for optimality.

Theorem 70. *If x^* maximizes $f(x)$ subject to $x \in S$, then $\nabla f(x^*) \cdot v \leq 0$ for all v entering S from v.*

Proof. This is the standard one-variable argument. We know that there exists $\epsilon > 0$ such that $x^* + tv \in S$ for $t \in [0, \epsilon)$ and therefore $f(x^*) \geq f(x^* + tv)$ for $t \in [0, \epsilon)$. Hence $D_v f(x^*) \geq 0$, which is equivalent to $\nabla f(x^*) \cdot v \leq 0$. □

Now we can apply a variation of the separating hyperplane theorem. The argument that we used to prove the separating hyperplane theorem proves the following.

Theorem 71. *Exactly one of the following two things holds:*

$$\textit{There exists } \lambda \geq 0 \textit{ such that } \boldsymbol{A}^t\lambda = w.$$

or

$$\textit{There exists } x \textit{ such that } \boldsymbol{A}x \geq 0 \textit{ and } w \cdot x < 0.$$

The first condition in the theorem says that w is in the convex set generated by the rows of the matrix **A**. Hence the theorem says that if w fails to be in a certain convex set, then it can be separated from the set. (In the application of the result, $w = -\nabla f(x^*)$.) The vector λ in the second condition plays the role of the normal of the separating hyperplane. Geometrically, the second condition states that it is possible to find a direction that that makes an angle of less than ninety degrees with all of the rows of **A** and greater than ninety degrees with w. This condition is ruled out by Theorem 71, hence the first condition must hold.

Theorem 72. *Suppose x^* solves: $\max f(x)$ subject to $x \in S$ when S is defined by linear inequalities and f is differentiable. Then there exists λ_k^* such that*

$$\nabla f(x^*) + \sum_{i \in I} \lambda_i \nabla \left(g_i(x^*) - b_i\right) = 0 \tag{13.10}$$

and

$$\sum_{i \in I} \lambda_i \left(g_i(x^*) - b_i\right) = 0. \tag{13.11}$$

This theorem gives us multipliers for an inequality constrained optimization problem provided that the constraint set is linear. The only place where we used linearity of the constraint set was in the characterization of inward pointing normals. It would be nice if this condition held in the non-linear case as well. It does not.

Consider the problem: $\max f(x_1, x_2) = x_1$ subject to $x_1, x_2 \geq 0$ and $(1 - x_1)^3 - x_2 \geq 0$. It is apparent that the solution to this problem is $(x_1, x_2) = (1, 0)$. [The second and third constraints force $x_1 \leq 1$. Since $(1, 0)$ is feasible, it must be optimal.] On the other hand, if you just try to apply the formula in Theorem 72 (in particular, the first component of equation (13.10)), then

$$1 + \lambda_1 + \lambda_3 \left(-3(1 - x^*)^2\right) = 0, \tag{13.12}$$

which plainly has no non-negative solution when $x^* = 1$. On the other hand, equation (13.11) implies that $\lambda_3 = 0$ when $x^* < 1$, which again means that equation (13.12 cannot have a solution. We must conclude that the conditions in Theorem 72 need not hold if S is not described by linear inequalities.

The problem comes from the characterization of vectors entering S. The constraints that define S give inward pointing normals at $(1, 0)$ as $(0, 1)$ (from the binding constraint that $x_2 \geq 0$) and $(0, -1)$ (from the binding constraint that $(1 - x_1)^3 - x_2 \geq 0$). $v = (1, 0)$ makes a zero inner product with both of these normals, but does not satisfy the definition of an entering direction. The characterization in Theorem 72 therefore places "too many" constraints on the derivatives of f so is not a useful characterization. tThe example means that we need to do a bit of work to make sure that the theorem applies to general problems. The work involves a slight modification in the definition of inward normal and then an assumption that rules out awkward cases.

Definition 119. *The vector v enters S from x^* if there is a $\epsilon > 0$ and a curve $\sigma : [0, \epsilon) \longrightarrow S$ such that $\sigma(0) = x^*$, and $\sigma'(0) = v$.*

In the definition, $\sigma'(0)$ denotes the right-hand derivative of σ at 0 (because the function need not be defined for $t < 0$). This definition is more general than Definition 118 because it allows inward directions by paths that are not necessarily straight lines. If S is defined by linear inequalities, the distinction does not matter, but in general directions may enter according to Definition 119 but not Definition 118.[2] It is immediate that $\nabla f(x^*) \cdot v \geq 0$ for all v entering S from x^* (according to Definition 119 as in Theorem 69).

[2] A direction that enters according to Definition 118 must satisfy Definition 119.

Definition 120. *If $S = \{x : g_i(x) \geq 0 \text{ for all } i \in I\}$, g_i are differentiable and $x^* \in S$, then the* inward normals at x^* *are the vectors $\nabla g_i(x^*)$ for i such that $g_i(x^*) = 0$.*

Now we simply rule out troublesome cases:

Definition 121. *Let $S = \{x : g_i(x) \geq 0 \text{ for all } i \in I\}$, g_i are differentiable, and $x^* \in S$. S satisfies the* constraint qualification *at x^* if $n_i \cdot v \geq 0$ for all inward normals n_i at x^* implies that v enters S from x^*.*

Plainly the example does not satisfy the constraint qualification. The constraint qualification will hold if the set of inward normals is linearly independent (check that this condition fails in the example). When linear independence holds, it is possible to find a normal direction that strictly enters: $n_i \cdot w = 1$ for all i and with this you can construct a curve with the property that $\nabla g_i(x^*) \cdot \sigma'(0) > 0$ for all i such that $g_i(x^*)$.

Here is another technical point. The linear independence condition will not hold if some of the constraints were derived from equality constraints (that is, one constraint is of the form $g(x) \geq 0$ and another constraint is of the form $-g(x) \geq 0$). This is why frequently equality constraints are written separately and a linear independence condition is imposed on directions formed by all binding constraints.

We now modify Theorem 72.

Theorem 73. *Suppose x^* solves: $\max f(x)$ subject to $x \in S$ when $S = \{x : g_i(x) \geq 0 \text{ for all } i \in I\}$ and f and g are differentiable. If x^* satisfies the constraint qualification at x^*, then there exists λ_k^* such that*

$$\nabla f(x^*) + \sum_{i \in I} \lambda_i \nabla g_i(x^*) - b_i) = 0 \tag{13.13}$$

and

$$\sum_{i \in I} \lambda_i \left(g_i(x^*) - b_i\right) = 0. \tag{13.14}$$

It is no accident Theorem 73 looks like Theorem 72. All that we did was assume away the pathologies caused by non-linear constraints.

13.3 Saddle Point Theorems

We have shown that if a "regularity" condition holds then

$$x^* \text{ solves } \max f(x) \tag{13.15}$$

subject to

$$g_i(x) \geq \tilde{b}_i, i \in I \tag{13.16}$$

then equations (13.10) and (13.11) must hold. Note that

$$\nabla g_i(x^*) = \nabla \left(g_i(x^*) - b_i \right)$$

so equation (13.16) can be written without the b_i.

In particular, this result holds when the constraints are linear (which guarantees that the regularity condition will hold).

As before, let us concentrate first on the linear case:

$$\max c \cdot x \text{ subject to } Ax \leq b, x \geq 0. \tag{13.17}$$

This is the standard linear programming problem. Here $f(x) = c \cdot x$ and for each $i = 1, \ldots, m, g_i(x) = -a_i \cdot x$, where a_i is a row of A and for $i = m+1, \ldots, m+n, g_i(x) = x_i$. To put (13.17) in the form of the original problem, set the $\tilde{b}$ in the theorem equal (constraint (13.16)) equal to $(-b_1, \ldots, -b_m, 0, \ldots, 0)$. $\tilde{b}$ has $n+m$ components.

Notice that $\nabla f(x) \equiv c$ when $f(x) \equiv c \cdot x$ and similarly for the constraints. Hence, if x^* solves (13.17) there exists $y^*, z^* \geq 0$ such that

$$c - y^* A + z^* = 0 \tag{13.18}$$

and

$$y^* \cdot (b - Ax^*) + z^* \cdot x^* = 0. \tag{13.19}$$

Equations (13.18) and (13.19) divide the multiplier vector λ into two parts. y^* contains the components of λ corresponding to $i \leq m$ – the constraints summarized in the matrix A. z^* are the components of λ corresponding to $i > m$; these are the non-negativity constraints.

It is useful to rewrite the constraints. Let

$$s(x : y, z) = c \cdot x + y \cdot (b - Ax) + v \cdot x. \tag{13.20}$$

We have the following theorem.

Theorem 74. *If x^* solves (13.17), then there exists y^* and $z^* \geq 0$ such that*

1. $s(x^*; y^*, z^*) \geq s(x; y^*, z^*)$ *for all* x.
2. $s(x^*; y^*, z^*) \leq s(x^*; y, z)$ *for all* $y, z \geq 0$.
3. y^* *solves* $\min b \cdot y$ *subject to* $A^t y \geq c$ *and* $y \geq 0$.

4. $b \cdot y^* = c \cdot x^*$.

.

Proof. It follows from equation (13.18) that

$$s(x : y^*, z^*) = \left(c - A^t y * + z^*\right) x + b \cdot y^* = b \cdot y^* \tag{13.21}$$

for all x. Therefore the first claim of the theorem holds as an equation. Furthermore, equation (13.18) implies that

$$y^* \cdot (b - Ax^*) + x^* \cdot z^*. \tag{13.22}$$

Hence $s(x^*; y^*, z^*)$ $b \cdot^*$. Equation 13.22) proves that

$$c \cdot x^* = -(x^*; y^*, z^*) = b \cdot y^*. \tag{13.23}$$

This implies the fourth claim of the theorem.

To prove the second line, we must show that

$$s(x^*; y, z) = c \cdot x^* + y^* \cdot (b - Ax^*) + x \cdot z \geq s(x^*; y^*, z^*) = c \cdot x^* \text{ for all } y, z. \tag{13.24}$$

However, x^* and $b \geq Ax^*$ by the constraints in (13.17). Therefore, if $y, z \geq 0$, then $y\,(b - Ax^*) + z \cdot x$ and so the second claim in the theorem follows.

To prove the third claim, first note that

$$c - A^t y^* + z^* = 0 \text{ and } z^* \geq 0$$

implies that $c \leq A^t y^*$. Therefore y^* satisfies $A^t y^* \geq c$ and $y^* \geq 0$. If y also satisfies $A^t y \geq c$ and $y \geq 0$, then by (2) and

$$\begin{aligned} s(x^*; y, z^*) &= c \cdot x^* + y\,(b - Ax^*) + z^* \cdot x^* && (13.25)\\ &= \left(c - A^t y\right) x^* + b \cdot y + z^* \cdot x^* && (13.26)\\ &= \left(c - A^t y\right) x^* + b \cdot y && (13.27)\\ &= b \cdot y, && (13.28) \end{aligned}$$

where the first equation is the definition of $s(\cdot)$, the second equation is just algebraic manipulation, the third equation follows because equations (13.19), $y^* \geq 0$, and $b \geq Ax^*$ imply that $z^* \cdot x^* = 0$, and the inequality follows because $x^* \geq 0$ and $c \leq A^t y$. It follows that

$$b \cdot y \geq s(x^*; y, z^*) \geq s(x^*; y^*, z^*) = b \cdot y^*$$

so y^* solves the minimization problem (3). □

We have show that if x^* solves the *Primal*:

$$\max c \cdot x \text{ subject to } \; Ax \leq b, x \geq 0,$$

then y^* solves the *Dual*:

$$\min b \cdot y \text{ subject to } A^t y \geq c, y \geq 0.$$

Notice that the Dual is equivalent to

$$\max -b \cdot y \text{ subject to } \; -A^t y \leq -c, y \geq 0$$

and that this problem has the same general form as the Primal (with c replaced by $-b$ and A replaced by $-A^t$ and b replaced by $-c$. Hence the Dual can be written in the form of the Primal and (applying the transformation one more time), the Dual of the Dual is the Primal.

Summarizing the result we have:

Theorem 75. *Associated with any Primal Linear Programming Problem*

$$\max c \cdot x \textit{ subject to } \; Ax \leq b, x \geq 0,$$

there is a Dual Linear Programming Problem:

$$\min b \cdot y \textit{ subject to } A^t y \geq c, y \geq 0.$$

The Dual of the Dual is the Primal. If the Primal has a solution x^*, then the Dual has a solution y^* and $b \cdot y^* = c \cdot x^*$. Moreover, if the Dual has a solution, then the Primal has a solution and the problems have the same value.

This result is called "The Fundamental Duality Theorem" (of Linear Programming). It frequently makes it possible to interpret the "multipliers" economically. Mathematically, the result makes it clear that when you solve a constrained optimization problem you are simultaneously solving another optimization problem for a vector of multipliers.

There is one y_i for every constraint in the primal (in particular, there need not be the same number of variables in the primal as variables in the dual). You therefore cannot compare x and y directly. You can compare the *values* of the two problems. The theorem shows that when you solve these problems, the values are equal: $b \cdot y^* = c \cdot x^*$. It is straightforward to show that if x satisfies the constraints of the Primal then $c \cdot x \leq c \cdot x^*$ and if y satisfies the constraints of the dual then $b \cdot y \geq b \cdot y^*$. Consequently any feasible value of the Primal is less than or equal to

any feasible value for the Dual. The minimization problem provides upper bounds for the maximization problem (and conversely).

It is a useful exercise to look up some standard linear programming problems and try to come up with interpretations of dual variables.

The arguments above demonstrate that $y^* (b - Ax^*) = 0$ and $(c - A^t y^*) x^* = 0$. These equations are informative. Consider $y^* (b - Ax^*) = 0$. This equation is a sum of terms of the form y_i^* (the value of the ith dual variable) times $(b - Ax^*)_i$ (the difference between the larger and smaller side of the ith constraint in the primal. Both of these quantities are positive or zero. When you add all of the terms together, you must get zero. Therefore (since each term is non-negative), each term must be zero. If you multiply two things together and get zero, one of them must be zero. What does this mean? If you know that $y_i > 0$, then you know that the associated constraint in binding (holds as an equation). If you know that a constraint is not binding, then the associated dual variable must be zero. This provides a weak sense in which the dual variables give an economic value to the given resources. Think about the Primal as a production problem and the constraints resource constraints. That is, you have access to a technology that transforms the inputs $b_1, \ldots, b_m$ into outputs $x_1, \ldots, x_n$. Suppose that you solve the problem and there is left over b_1. That is, the first constraint is not binding. How much would you pay for an extra unit of the first input? Nothing. Why? Because you cannot profitably turn that resource into profitable output (if you could do this, you would have used up your original supply). Hence the "economic" value of the first input is 0. Similarly, if you are willing to pay a positive amount for the second good $y_2 > 0$, then it must be that you use up your entire supply (second constraint binding). It turn out that y_i operates as a resource price more generally: If you change the amount of b_i by a small amount δ, then your profits change by δy_i. This story provides another interpretation of the equation $b \cdot y^* = c \cdot x^*$. The left-hand quantity is the value of your inputs. The right-hand side is the value of your output.

The duality theorem of linear programming tells you a lot about the dual if you know that the primal has a solution. What happens if the Primal fails to have a solution? First, since the duality theorem applies when the Dual plays the role of the Primal, it cannot be that the Dual has a solution. That is, the Dual has a solution if and only if the Primal has a solution. There are only two ways in which a linear programming problem can fail to have a solution. Either $\{x : Ax \leq b, x \geq 0\}$ is empty. In this case the problem is *infeasible* – nothing satisfies the constraints. In this case you certainly do not have a maximum. Alternatively, the problem is feasible ($\{x : Ax \leq b, x \geq 0\}$ is not empty), but you cannot achieve a maximum. Of course, since linear functions are continuous, this means that the feasible set cannot be bounded. It turns out that in the second case you can make the objective

function of the Primal arbitrarily large (so the Primal's value is "infinite"). It is not hard to show this. It is even easier to show that it is impossible for the Primal to be unbounded if the Dual is feasible (because any feasible value of the Dual is an upper bound for the Primal's value).

To summarize:

1. If the primal has a solution, then so does the dual, and the solutions have the same value.

2. If the primal has no solution because it is unbounded, then the dual is not feasible.

3. If the primal is not feasible, then either the dual is not feasible or the dual is unbounded.

4. The three statements about remain true if you interchange the words "primal" and "dual."

We becan by showing that if x^* solves the Primal, then there exist y^* and z^* such that

$$s(x^*; y, z) \geq s(x^*; y^*, z^*) \geq s(x; y^*, z^*). \tag{13.29}$$

Expression (13.29) states that the function s is maximized with respect to x at $x = x^*$, $(y, z) = (y^*, z^*)$ and minimized with respect to (y, z) at $(x^*, y^*, z^*), (y, z) \geq 0$. This makes the point (x^*, y^*, z^*) a *saddle point* of the function s.

What about the converse? Let us state the problem more generally.

Given the problem

$$\max f(x) \text{ subject to } g_i(x) \geq b_i, i \in I$$

consider the function

$$s(x, \lambda) = f(x) + \lambda \cdot (g(x) - b) \tag{13.30}$$

where $g(x) = (g_1(x), \ldots, g_I(x)), b = (b_1, \ldots, b_I)$.$(x^*, \lambda^*)$ is a *saddle point* if $\lambda^* \geq 0$ and s is maximized with respect to x and minimized with respect to λ at (x^*, λ^*). That is,

$$s(x, \lambda^*) \leq s(x^*, \lambda^*) \leq s(x^*, \lambda) \tag{13.31}$$

for all $\lambda \geq 0$.

Theorem 76. *If (x^*, λ^*) is a saddle point of s, then x^* solves $\max f(x)$ subject to $g_i(x) \geq b_i$.*

Proof. Since $s(x^*, \lambda^*) \leq s(x^*, \lambda)$ for all $\lambda \geq 0$,

$$\lambda^* \cdot (g(x^*) - b) \leq \lambda \cdot (g(x^*) - b) \text{ for all } \lambda \geq 0. \tag{13.32}$$

Therefore (setting $\lambda = 0$ in inequality (13.32)) it must be that $\lambda^* \cdot (g(x^*) - b) \leq 0$.

We also claim that $g_i(x^*) - b_i \geq 0$ for all i. To prove this, suppose that there exists a k such that $g_j(x^*) - b_j < 0$ and let $\lambda_i = 0$ for $i \neq k$ and $\lambda_k = M$. It follows that

$$\lambda^* \cdot (g(x^*) - b) \leq \lambda \cdot (g(x^*) - b) = M \left(g_k(x^*) - b_k\right).$$

Since the right-hand side gets arbitrarily small as M grows large, which is impossible. This establishes the claim.

The only way to have $\lambda^* \geq 0, g_i(x^*) - b_i \geq 0$ for all i, and $\lambda^* \cdot (g(x^*) - b) \leq 0$ is for $\lambda^* \cdot (g(x^*) - b) = 0$. It follows that $s(x^*, \lambda^*) \leq s(x^*, \lambda)$ implies that x^* is feasible ($g(x^*) \geq b$) and that λ^*, x^* satisfy the complementary slackness condition

$$\lambda^* \cdot (g(x^*) - b) = 0. \tag{13.33}$$

To complete the proof, we must show that $g(x) \geq b$ implies that $f(x^*) \geq f(x)$. However, we know that $s(x^*, \lambda^*) \geq s(x, \lambda^*)$ and we just showed that

$$s(x^{,}\lambda^*) = f(x^*) + \lambda^* \cdot (g(x^*) - b) = f(x^*).$$

On the other hand,

$$s(x, \lambda^*) = f(x) + \lambda^* \cdot (g(x) - b)$$

so if $g(x) - b \geq 0$, $\lambda^* \geq 0$ implies that $s(x, \lambda^*) \geq f(x)$. Summarizing we have

$$f(x^*) = s(x^*, \lambda^*) \geq s(x, \lambda^*) \leq f(x) \text{ whenever } g(x) \geq b.$$

□

Hence we have shown an equivalence between saddle points and optimal in linear programming problems and that, in general, the saddle point property guarantees optimality. Under what conditions is the converse true? That is, when can we say that if x^* solves $\max f(x)$ subject to $g(x) \geq b$, that there exists $\lambda^* \geq 0$ such that (x^*, λ^*) is a saddle point?

Theorem 77. *If f and g_i are concave and x^* solves $\max f(x)$ subject to $g(x) \geq b$ and a "constraint qualification" holds, then there exists $\lambda^* \geq 0$ such that (x^*, λ^*) is a saddle point of $s(x, \lambda) = f(x) + \lambda \cdot (g(x) - b)$.*

Comments:

1. We will discuss the "constraint qualification" later. It plays the role that "regularity" played in the "inward normal" discussion.

2. This result is often called the Kuhn-Tucker Theorem. λ are called Kuhn-Tucker multipliers. (Largrange mulitpliers typically refer only to multipliers from **equality** constrained problems. Remember: equality constraints are a special case of inequality constraints and are generally easier to deal with.)

3. It is possible to prove essentially the same theorem with weaker assumptions (quasi concavity of g). The proof that we propose does not generalize directly. See Sydsaeter or the original paper of Arrow and Enthoven, *Econometrica* 1961.

4. The beauty of this result is that it transforms a constrained problem into an unconstrained one. Finding x^* to maximize $s(x, \lambda^*)$ for λ^* fixed is relatively easy. Also, note that if f and g_i are differentiable, then the first-order conditions for maximizing s are the familiar ones:

$$\nabla f(x^*) + \lambda \cdot \nabla g(x^*) = 0.$$

   The first order conditions that characterization the minimization problem are precisely the complementary slackness conditions. Since $s(x, \lambda)$ is concave in x when f and g are concave, it will turn out that

$$\nabla f(x^*) + \lambda \cdot \nabla g(x^*) = 0, \lambda^* \cdot (g(x^*) - b) = 0, g(x^*) \geq 0$$

   are necessary **and sufficient** for optimization. This will be an immediate consequence of the (ultimate) discussion of the sufficient conditions for optimality.

Proof. We assumed concavity and we need to find "multipliers." It looks like we need a separation theorem. The trick is to find the right set.

Let

$$K = \{(y, z) : \text{ there exists } x \text{ such that } \ g(x) \geq y \text{ and } f(x) \geq z\}$$

K is convex since g_i and f are concave. In this notation, $y \in \mathbb{R}^m$ and $z \in \mathbb{R}$. If x^* solves the maximization problem, then $(b, f(x^*)) \in K$. In fact, it must be a boundary point since otherwise there exists $\epsilon > 0$ and x such that $g_i(x) \geq b_i + \epsilon > b_i$ for all i and $f(x) \geq f(x^*) + \epsilon$. This contradicts the optimality of x^*.

Hence there is a separating hyperplane. That is, there exists $(\lambda, \mu) \neq 0, \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}$ such that

$$\lambda \cdot b + \mu f(x^*) \geq \lambda \cdot y \geq \mu z$$

for all $(y, z) \in K$. This is almost what we need. First, observe that $(\lambda, \mu) \geq 0$. If some component of λ, say $\lambda_j < 0$, you get a contradiction in the usual way by taking points $(y, z) \in K$ where, for example, $z = f(x^*), y_i = b_i, i \neq j, y_j = -M$. If M is large enough, this choice contradicts

$$\lambda \cdot b + \mu f(x^*) \geq \lambda \cdot y + \mu z.$$

Similarly, if $\mu < 0$ you obtain a contradiction because $(b, z) \in K$ for all arbitrarily small (large negative) z. Hence we have $(\lambda, \mu) \geq 0, (\lambda, \mu) \neq 0$ and

$$\lambda \cdot b + \mu f(x^*) \geq \lambda \cdot y + \mu z$$

for all $(y, z) \in K$. It is also true that

$$\lambda \cdot (g(x^*) - b) = 0.$$

This fact follows because $(g(x^*), f(x^*)) \in K$. Hence

$$\lambda \cdot b + \mu f(x^*) \geq \lambda \cdot g(x^*) + \mu f(x^*).$$

Therefore $0 \geq \lambda \left(g(x^*) - b\right)$. However, $\lambda \geq 0$ and $g(x^*) \geq b$ (by feasibility). Hence, $0 = \lambda \left(g(x^*) - b\right)$.

Now **suppose** that $\mu > 0$. (This is where the constraint qualification comes in.) Put $\lambda^* = \lambda/\mu$. We have

$$f(x^*) + \lambda^* \cdot b \geq f(x) + \lambda^* \cdot g(x) \text{ for all } x \tag{13.34}$$

(since $(g(x), f(x)) \in K$) and $\lambda^* \cdot (g(x^*) - b) = 0$.

Now we can confirm the saddle-point property:

$$\begin{aligned} s(x^*, \lambda) \geq s(x^*, \lambda^*) \text{ for } \lambda \geq 0 \equiv \\ f(x^*) + \lambda \cdot (g(x^*) - b) \geq f(x^*) + \lambda^* \cdot (g(x^*) - b) \equiv \\ \lambda \cdot (g(x^*) - b) \geq 0 \text{ for } \lambda \geq 0. \end{aligned}$$

where we use $\lambda^* \cdot (g(x^*) - b) = 0$ and the last line holds because $g(x^*) \geq b$.

$$\begin{aligned} s(x^*, \lambda^*) \geq s(x, \lambda^*) \equiv \\ f(x^*) + \lambda^* \cdot (g(x^*) - b) \geq f(x) + \lambda^* \cdot (g(x) - b) \equiv \\ f(x^*) + \lambda^* \cdot b \ \geq f(x) + \lambda^* \cdot g(x) \end{aligned}$$

since $\lambda^* \cdot (g(x^*) - b) = 0$.

All that remains is to clarify the constraint qualification. We need an assumption that guarantees that μ, the weight on f, is strictly positive. What happens if $\mu = 0$? We have

$$\lambda \neq 0 \text{ such that } \lambda \cdot b \geq \lambda \cdot y \text{ for all } y \text{ such that } (y, z) \in K \text{ for some } z. \quad (13.35)$$

One condition that guarantees this cannot happen is:

$$\text{there exist } x^* \text{ such that } g_i(x^*) > b_i \text{ for all } i.$$

This condition is not consistent with (13.35) because we cannot have $\lambda \cdot b \geq \lambda \cdot g(x^*)$ if $\lambda \geq 0$, $\lambda \neq 0$, and $g_i(x^*) > b_i$ for all i.

The constraint qualification is fairly weak. It just requires that the constraint set has a non-empty interior. □

13.4 Second-Order Conditions

We know from the theory of unconstrained optimization that the first-order conditions for optimality do not distinguish local maxima from local minima from points that are neither. We also know that second-order condition, derived from thinking about how a first derivative changes in the neighborhood of a critical point do provide useful information. When you have an equality constrained maximization problem you derive first-order conditions by requiring that the objective function be maximized in all directions that will satisfy the constraints of the problem. The second-order conditions must hold for exactly these directions. That is, if x^* is a maximum and $x^* + tv$ satisfies the constraints of the problem for t near zero, then $h(t) = f(x^* + tv)$ has a critical point at $t = 0$ and $h''(0) \leq 0$. Two applications of the chain rule yield that $h''(0) = v^t D^2 f(x^*) v$, so $h''(0) < 0$ if and only if a quadratic form is negative. If the problem is unconstrained (that is, if $x^* + tv$ satisfies the constraints of the problem for all v and sufficiently small t, then the second-order conditions require that the matrix of second derivatives be negative semi-definite. In general, this condition need only apply in the directions consistent with the constraints. There is a theory of "Boardered Hessians" that allows you to use some insights from the theory of quadratic forms to classify when quadratic form restricted to a set of directions will be positive definite. This theory is ugly and we will say no more about it.

13.5 Examples

Example 50.

$$\max_{x_1, x_2} x_1^2 - x_2$$

$$\textit{subject to} \quad x_1 - x_2 = 0$$

It is pretty obvious then that $x_1 = x_2$ and thus the problem is really

$$\max_{x_1} x_1^2 - x_1$$

So this example is so simple that it can be done by direct substitution, but with harder examples we must do this implicitly.

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

implicitly define $x_2 = h(x_1)$ around x_1^* (the solution).
Recall that h is well defined if

$$\frac{\partial g}{\partial x_2}(\mathbf{x}^*) \neq 0$$

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

$$\frac{\partial g}{\partial x_1}(\mathbf{x}^*) + \frac{\partial g}{\partial x_2}(\mathbf{x}^*)h'(x_1^*) = 0$$

$$h'(x_1) = -\frac{\frac{\partial g}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial g}{\partial x_2}(\mathbf{x}^*)}$$

This gives us an unconstrained maximization problem, for $\mathbf{x}$ near $\mathbf{x}^*$.

$$\max_{x_1} f(x_1, h(x_1))$$

Because $\mathbf{x}^*$ solves the original maximization problem, it solves this one.

Example 51. *The classic example of constrained optimization is the utility maximization problem subject to a budget constraint. You can think of this as me walking into Costco (the world's greatest store) with I dollars in my pocket, and having surveyed the vast array of wonderful products they have, I make the selection that will leave me happiest on leaving the store.*

$$\max_{\boldsymbol{x}} U(\boldsymbol{x})$$

$$\textit{subject to} \quad \boldsymbol{p}^t \cdot \boldsymbol{x} \leq I$$

where,
***p** is vector of prices*

$\boldsymbol{x}$ is vector of commodities
I is income.
So all the inequality is saying is that a person cannot spend more than the money in their pocket!
For these problems we just assume that the constraint binds (we will see why) since people obviously spend all the money in their pockets. You may question this as people do save, but if people do invest in savings then we just define saving as an extra commodity and this gets rid of the problem.

Example 52. *The other classic example of constrained optimization is the expenditure minimization problem. Suppose I really on Tuesday afternoon my girlfriend is very happy. Her utility level is* 100 *utils. But then on Tuesday night I do something really stupid, and she gets really annoyed. As a result her utility level drops. Now I really want to get her back to her Tuesday afternoon level of happiness again so I decide to buy her some gifts. But I am a cheapskate and as such want to spend as little as possible getting her back to the same level of happiness that she was at on Tuesday afternoon. So now my problem is different: the goal is now to spend as little as possible subject to a utility constraint. I want to spend as little as possible and get her back to a level of utility that is at least as great as it was on Tuesday afternoon.*

Example 53. $\max \log(x_1 + x_2)$ *subject to:* $x_1 + x_2 \leq 5$ *and* $x_1, x_2 \geq 0$.

Since $\log$ *is a strictly increasing function, the problem is equivalent to solving* $\max(x_1 + x_2)$ *subject to the same constraints. Since the objective function is increasing in both variables, the constraint* $x_1 + x_2 \leq 5$ *must bind. Hence the problem is equivalent to maximizing* $x_1 + x_2$ *subject to* $x_1 + x_2 = 5$ *and* $x_1, x_2 \geq 0$. *At this point it should be clear that as long as* $x_1 \in [0, 5]$ *any pair* $(x_1, 5 - x_1)$ *solves the problem.*

Of course, you could solve the problem using the Kuhn-Tucker Theorem. Doing so would give you the additional insight of finding multipliers, but this would not compensate you for the pain.

Example 54. $\max \log(1 + x_1) - x_2^2$ *subject to:* $x_1 + 2x_2 \leq 3$ *and* $x_1, x_2 \geq 0$.

The first-order conditions are

$$\begin{aligned}
\frac{1}{1 + x_1} &= \lambda_1 - \lambda_2 && (13.36)\\
-2x_2^2 &= 2\lambda_1 - \lambda_3 && (13.37)\\
\lambda_1 (3 - x_1 - 2x_2) &= 0 && (13.38)\\
\lambda_2 x_1 &= 0 && (13.39)\\
\lambda_3 x_2 &= 0 && (13.40)
\end{aligned}$$

It should be clear that $3 = x_1 + 2x_2$ *and* $x_1 > 0$ *in the solution. Hence the third constraint is satisfied,* $\lambda_2 = 0$ *by the fourth constraint, and the first constraint gives:* $\lambda_1 = 1/(1+x_1)$ *and therefore the second constraint implies that* $\lambda_3 > 0$*. It follows from the final constraint that* $x_2 = 0$ *and therefore* $x_1 = 3$ *and* $\lambda_1 = 1/4$*.*

Again, one could solve this problem by observation.

Example 55. $\max x_1 \cdots x_n$ *subject to* $\sum_{i=1}^k x_i = K$*.*

This problem has an equality constraint. If $v = x_1 \cdots x_n$*, then we can write the first-order conditions as* $v = \lambda x_i$ *for all* i*. The constraint then implies that* $kv = \lambda K$ *and therefore each* $x_i = K/k$*.*

Example 56. *We can show that the maximum value of the function* $f(x) = x_1^2 x_2^2 \cdots x_n^2$ *on the unit sphere:* $\{x : \sum_{i=1}^n x_i^2 = 1\}$ *is to set* $x_n = (1/n)^n$ *for all* n*. This is because the solution must satisfy*

$$2\frac{f(x)}{x_i} = 2\lambda x_i,$$

which implies that the x_i *are independent of* i*.*[3]

We can use this solution to deduce an interesting property. Given any n *positive numbers* $a_1, \ldots, a_n$*, let*

$$x_i = \frac{a_i^{1/2}}{(a_1 + \cdots + a_n)^{1/2}}, \text{ for } i = 1, \ldots, n.$$

It follows that $\sum_{i=1}^n x_i^2 = 1$ *and so*

$$\left(\frac{a_1 \cdots a_n}{(a_1 + \cdots + a_n)^n}\right)^{1/n} \leq \frac{1}{n}$$

and so

$$(a_1 + \cdots + a_n)^{1/n} \leq \frac{(a_1 + \cdots + a_n)^{1/2}}{n}.$$

The interpretation of the last inequality is that the geometric mean of n *positive numbers is no greater than their arithmetic mean.*

It is possible to use techniques from constrained optimization to deduce other important results (the triangle inequality, the optimality of least squares, . . .).

[3]Careful: We know that this problem has both a maximum and a minimum. How did we know that this was the maximum? And where did the minimum go? The answer is that it is plain that you do not want to set any $x_i = 0$ to get a maximum and that we have the only critical point that does not involve $x_i = 0$. It is also clear that setting one or more $x_i = 0$ does solve the minimization problem.

Example 57. *A standard example in consumer theory is the Cobb-Douglas Utility function,* $f(x) = x_1^{a_1} \cdots x_n^{a_n}$*, where the coefficients are nonnegative and sum to one. A consumer will try to maximize* f *subject to a budget constraint:* $p \cdot x \leq w$*, where* $p \geq 0$*,* $p \neq 0$ *is the vector of prices and* $w > 0$ *is a positive wealth level. Since the function* f *is increasing in its arguments, the budget constraint must hold as an equation. The first-order conditions can be written* $a_i f(x)/x_i = \lambda p_i$ *or*

$$a_i f(x) = \lambda p_i x_i.$$

Summing both sides of this equation and using the fact that the a_i *sum to one yields:*

$$f(x) = \lambda p \cdot x = \lambda w.$$

It follows that

$$x_i = \frac{w a_i}{p_i} \text{ and } \lambda = \left(\frac{a_1}{p_1}\right)^{a_1} \cdots (a_n p_n)^{a_n}.$$

The previous two examples show illustrate that a bit of cleverness allows neat solutions to the first-order conditions.