Logic and Set Theory

These notes correspond to mathematical appendix 1 in the text.

1 Logical statements

A conditional statement is simply an "if, then" statement. Typically we will write, if p, then q. There are many other details and plenty of terminology that I will skip. You can consult Epp, Susanna. *Discrete Mathematics with Applications.* 2^{nd} ed. PWS Publishing Company, Boston, MA, 1995 for additional details beyond those in the textbook.

Related to our conditional statement, which we state as *if p, then q*, (we will write this as $p \to q$, or "p implies q") we have three other statements:

Converse: $q \to p$

Inverse: $\sim p \to \sim q$ (the symbol $\sim$ is defined as "not" in this context; later we will define it as "indifferent to")

Contrapositive: $\sim q \to \sim p$

Of these 3 statements, only the contrapositive is *logically equivalent* to the conditional statement. For two items to be logically equivalent, they need to have the same *truth table.* A truth table is a table that lists the truth-values of a proposition that result from all the possible combinations of the truth-values of its components. Table 1 shows the truth table for the conditional statement, $p \to q$. Note that the conditional statement is only false when p is true and q is false. If both are true then the statement is true. However, the statement says nothing about whether or not q will occur if p does NOT occur, so the statement is still true in those instances when p does not occur regardless of whether or not q occurs. Thus, if the conditional statement is: If you show up to class, then you will receive an A, the only time we can say that the statement is false is when you show up to class and do not receive an A. If you fail to show up to class then all bets are off – you may receive an A, you may not, but the conditional statement does not tell us anything about what happens in those instances.

Table 2 shows a truth table for the statement, the converse, the inverse, and the contrapositive. Note that we will need the truth values for $\sim p$ and $\sim q$ as well. We can see that the truth values for the contrapositive are identical to those of the statement, so that the two are logically equivalent. This will be useful information momentarily.

2 Three Methods of Proof

We will discuss three specific methods of proving theorems that may prove useful to you. Those three methods are direct proof, proof by contradiction, and proof by contraposition. Proofs can be written out in paragraph form, with correct grammatical structure (and should be for journal articles). However, that sometimes obscures the thought process, and I will write proofs in a t-table.

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Table 1: Truth table for conditional statement

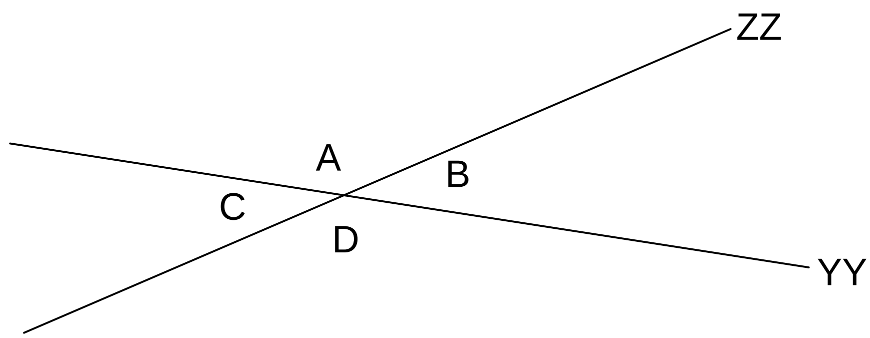
p	q	$\sim p$	$\sim q$	$p \to q$	$q \to p$	$\sim p \to \sim q$	$\sim q \to \sim p$
T	T	F	F	T	T	T	T
T	F	F	T	F	T	T	F
F	T	T	F	T	F	F	T
F	F	T	T	T	T	T	T

Table 2: Truth table for conditional statement, converse, inverse, and contrapositive

2.1 Direct Proof

Direct proof is straightforward – if we have a conditional statement and want to prove that q is true, we assume that p is true and deduce, using "what we know", that q is true. "What we know" typically consists of previously proven theorems and results, definitions, and other common knowledge.

Suppose we have the following diagram as in Figure 2.1:



Two intersecting lines.

And you are told that segments ZZ and YY are straight lines. The following statement is then made: If ZZ and YY are straight lines, then $\angle B$ and $\angle C$ are congruent. You are then asked to prove this:

Proof. If ZZ and YY are straight lines, then angle B and angle C are congruent

Statement	Reason
1. $\angle A + \angle B = 180°$; $\angle A + \angle C = 180°$	1. Definition of a straight line
2. $\angle B = 180° - \angle A$ $\angle C = 180° - \angle A$	2. Subtraction
3. $\angle B = \angle C$	3. Substitution

$\therefore$ $\angle B$ and $\angle C$ are congruent by the definition of congruent. ■

2.2 Proof by Contradiction

A second method of proof is proof by contradiction. Here are the steps.

1. Suppose the statement to be proved is false.
2. Show that this supposition leads naturally to a contradiction.
3. Conclude that the statement to be proved is true.

Why does this work? Essentially, when you contradict the negation of the statement (which is what you are assuming is true in step 1), you are proving it false. If the negation is false, then the statement is true. We will use proof by contradiction to show that there is no greatest integer.

Proof. *There is no greatest integer.*

By contradiction. Assume there is a greatest integer, M.

Statement	Reason
1. $M > m$, $\forall$ integers m	1. Definition of greatest
2. Let $N = M + 1$	2. Defining a new number by addition
3. N is an integer	3. The sum of integers is an integer
4. $M + 1 > M$	4. Definition of greater
5. $N > M$	5. Substitution

Step 5 contradicts the original assumption, that M is the greatest integer, because we have found an integer that is greater than M. $\therefore$ There is no greatest integer. ■

Although there are no set rules as to when to use proof by contradiction, here are two guidelines as to when you might find proof by contradiction useful.

1. When you want to show that there is no object with a certain property.
2. When you want to show that a certain object does not have a certain property.

2.3 Proof by Contraposition

A third method of proof is proof by contraposition. Here are the steps:

1. State the contrapositive of the statement.
2. Prove the contrapositive by direct proof.

Why does this work? Recall that the contrapositive and the conditional statement are logically equivalent. Thus, if the contrapositive is true, then the conditional statement is also true. We will show that for all integers n, if n^2 is even, then n is even.

Proof. $\forall$ integers n, if n^2 is even, then n is even

By contraposition. $\forall$ integers n, if n is odd, then n^2 is odd

Statement	Reason
1. k is an integer	1. assumption
2. $2k + 1$ is an odd integer	2. definition of an odd integer
3. $n = 2k + 1$	3. we have now defined what is given to us
4. $n^2 = (2k+1)(2k+1)$	4. Definition of square
5. $n^2 = 4k^2 + 4k + 1$	5. Multiplication (FOIL)
6. $n^2 = 2\left(2k^2 + 2k\right) + 1$	6. Distribution
7. $x = 2k^2 + 2k$	7. Defining a new number
8. x is an integer	8. Addition and multiplication of integers results in an integer
9. $n^2 = 2x + 1$	9. Substitution

$\therefore n^2$ is odd by definition of odd. Since we have proven the contrapositive true, we know the statement is also true, so that if n^2 is even, then n is even $\forall$ integers n ■

Note that the proof in the book skips some steps because they "know" that the product of two odd integers is odd, so they can invoke what they "know" to conclude that n^2 is odd after step 4. However, we did not know that.

Again, there are no set rules for using proof by contraposition. One benefit of using proof by contraposition is that we know exactly what we need to prove, so there is some guidance. With proof by contradiction, we do not necessarily know what contradiction we are looking for. However, the benefit of proof by contradiction is that once we have shown any contradiction then we are finished. Also, if a theorem is proved using proof by contraposition, then it can also be proved using proof by contradiction. However, the converse of that statement is NOT true (but you already knew that).

2.4 NOT Methods of Proof

Here is a list of things that people might want to use to prove a theorem. However, they are NOT methods of proof.

1. Proof by converse – we know that a statement and its converse are NOT logically equivalent, so proving the converse is true does not necessarily prove the statement is true.

2. Proof by inverse – replace "converse" with "inverse" in the statement above.

   Note: I have been thinking about this, and although I have no evidence, I believe that people like to use "proof by converse" and "proof by inverse" as methods of proving the truth of a conditional statement is because if p and q are both true then the conditional, converse, inverse, and contrapositive are all true (just look at the truth table). However, one needs to consider ALL the possible combinations of the truth of p and q, and the truth table shows us that the conditional and the inverse (and converse) are not logically equivalent.

3. Proof by example – one million examples may suggest that something is true; however, all it takes is one counterexample to prove that something is NOT true. And if you give one million examples of something and state that it is true when it really is not true, someone with time on his or her hands will find that one counterexample.

4. Proof by appealing to someone else's proof for a case that is related to their theory, but not exactly as it is specified – economists will typically appeal to someone else's theory (which will typically be done in a continuous space for mathematical tractability) and say that it suggests certain results in their example (which will typically be a discrete space if they are writing about something that actually happened). If you have a specific example about which you are writing, take a little time to make sure simple counterexamples to the theory do not exist.

2.5 Elements of set theory

A set is a collection of elements. We can list the elements $S = \{1, 8, 22\}$ or describe the elements $S' = \{x|x < 10\}$. We use $\in$ to denote membership in the set, as in $1.82 \in S'$.

A set S is a subset of the set T if all the elements of S are in T. If $S = \{1, 2\}$ and $T = \{x|x < 10\}$ then S is a subset of T because all of the elements of S are in T. We can write $S \subset T$ (S is contained by T) or $T \supset S$ (T contains S). If $S \subset T$, then this implies that an element of S is also in T.

A set S is empty if it contains no elements. We denote the empty set by $\varnothing$.

The complement of a set $S \subset U$ are all the elements in U which are not in S. If $U = \{1, 2, 3, 4, 5\}$ and $S = \{1, 2\}$ the complement of S is $S^c = \{3, 4, 5\}$.

The union of two sets S and T is the set of all elements which are contained in either set. We denote union by $\cup$ and write $S \cup T \equiv \{x|x \in S \text{ or } x \in T\}$.

The intersection of two sets S and T is the set of all elements which are contained in both sets. We denote intersection by $\cap$ and write $S \cap T \equiv \{x|x \in S \text{ and } x \in T\}$.

The product of two sets S and T is the set of ordered pairs in the form (s, t) where $s \in S$ and $t \in T$. We denote the product as $S \times T \equiv \{(s, t)\,|s \in S, t \in T\}$. The Cartesian plane is a product of two sets. We let $\mathbb{R}$ denote the set of real numbers, so that $\mathbb{R} = \{x| -\infty < x < \infty\}$. The Cartesian plane is $\mathbb{R} \times \mathbb{R} \equiv \{(x_1, x_2)\,|x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$. This set is often called two-dimensional Euclidean space and denoted as $\mathbb{R}^2$. We can also have n-dimensional Euclidean space, which is $\mathbb{R}^n \equiv \mathbb{R} \times \mathbb{R} \times ... \times \mathbb{R}$, where there are n $\mathbb{R}$'s. A vector from a set product consists of one element from each individual set, so that if the space is $\mathbb{R}^2$ then we will have an ordered pair (which you all should be familiar with), and if the space is $\mathbb{R}^n$ then we will have an n-vector. Note that with two vectors x and y, $x \geq y$ if and only if each element of x is greater than or equal to the corresponding element of y, or $x_i \geq y_i$ for all $i = 1, ..., n$. We state that x is strictly greater than y, or $x >> y$, if and only if $x_i > y_i$ for all $i = 1, ..., n$.

3 Convex sets

Convex sets are very important in microeconomic theory. Many results hinge upon some set being convex. Consider a two-dimensional set, or a set in $\mathbb{R}^2$. The basic notion of convexity is that one can take any two points in the set (including those on the boundary), draw a line between them, and all the points on the line are also in the set. Figure 1 provides some examples of convex and nonconvex sets.

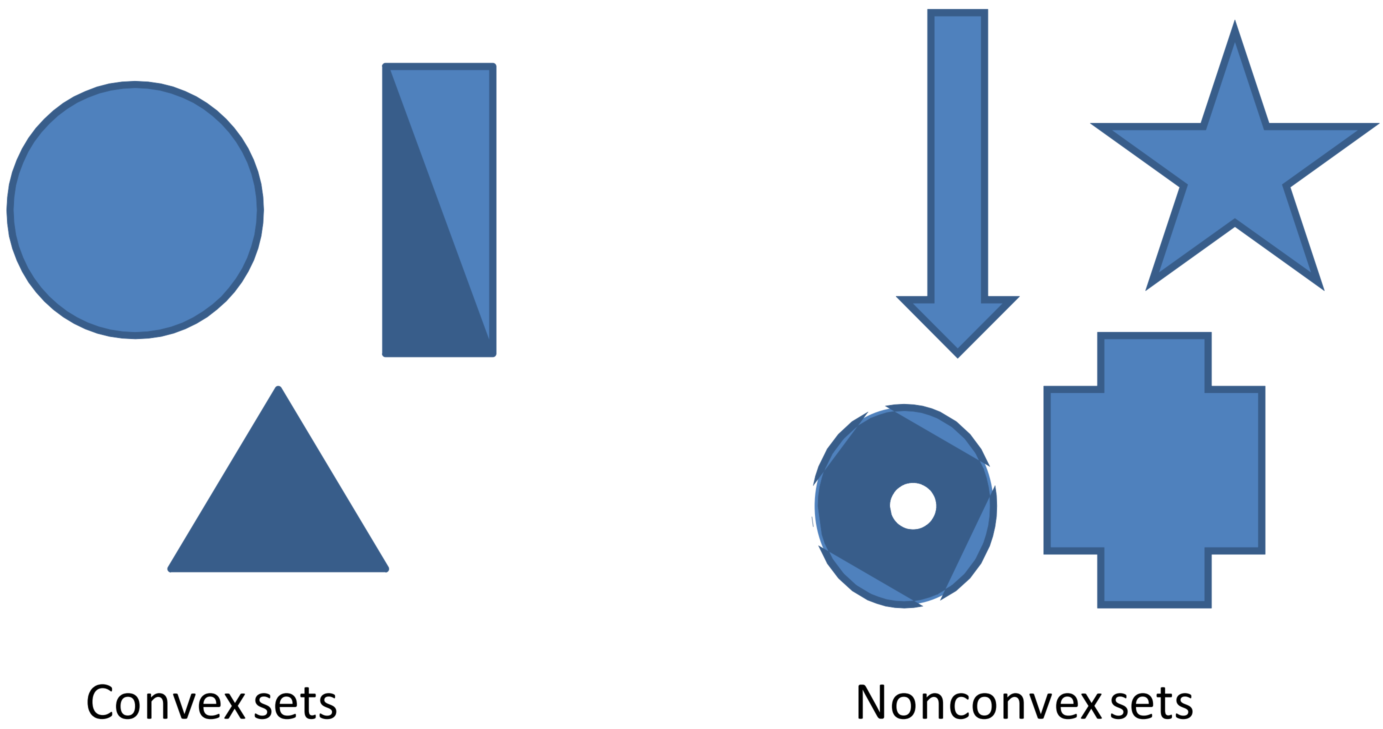


Figure 1: Some convex and nonconvex sets

Extending the definition to $\mathbb{R}^n$, a set $S \subset \mathbb{R}^n$ is a convex set if for all $x^1 \in S$ and all $x^2 \in S$, we have $tx^1 + (1-t)\,x^2 \in S$ for all $t \in [0,1]$. This simply means that we can take any two n-vectors from the set S and find all the weighted averages between those points and that if all those weighted averages are in the set then the set is convex.

Theorem 1 *Let S and T be convex sets in $\mathbb{R}^n$. Then $S \cap T$ is a convex set.*

Proof. Let S and T be convex sets. Let x^1 and x^2 be any two points in $S \cap T$. Because $x^1 \in S \cap T$, $x^1 \in S$ and $x^1 \in T$. The same is true for x^2. Let $z = tx^1 + (1-t)\,x^2$ for $t \in [0,1]$. Because S and T are convex $z \in S$ and $z \in T$. By the definition of the term intersection, $z \in S \cap T$. Because every convex combination of any two points in $S \cap T$ is also in $S \cap T$, the set $S \cap T$ is convex. $\blacksquare$

4 Relations

A binary relation is defined by specifying some meaningful relationship that holds between the elements of the the ordered pair. Familiar binary relations should be $\geq$, $>$, $\leq$, and $<$. Some useful concepts concerning binary relations are as follows.

Completeness

Definition 2 *A relation $\Re$ on S is complete if, for all distinct elements x and y in S, $x \Re y$ or $y \Re x$.*

Let S be the set of all people, then the relation "is older than" is not complete because we can always find two people who are the same age. If the relation "is at least as old as", then that relation is complete.

Reflexivity

Definition 3 *A relation $\Re$ on S is reflexive if, for all elements in S, $x\Re x$.*

Again, the relation "is older than" is not reflexive as people are not older than themselves, but the relation "is at least as old as" is reflexive because people are at least as old as themselves.

Transitivity

Definition 4 *A relation $\Re$ on S is transitive if, for any three elements x, y, and z in S, $x\Re y$ and $y\Re z$ implies $x\Re z$.*

Both "is older than" and "is at least as old as" are transitive.

4.1 Functions

A function is a common kind of relation. A function maps from one element of one set to another single, unique element of another set. We can write the function f maps from set D to set R as: $f : D \rightarrow R$, where D is the domain and R is the range. The image of f is the set of points in the range into which some point in the domain is mapped. If every point in the range is assigned to at most a single point in the domain, the function is called *one-to-one*. If the image is equal to the range – if every point in the range is mapped into by some point in the domain – the function is called *onto*.

5 Topology

Topology is the study of properties of sets and mappings. A *metric* is a measure of distance. A *metric space* is a set with some associated distance measure. Consider the real number line $\mathbb{R}$ with the distance given between two points on the real number line as $d\left(x^1, x^2\right) = \left|x^1 - x^2\right|$.

Now consider the Cartesian plane with the following metric:

$$d\left(x^1, x^2\right) = \sqrt{\left(x_1^2 - x_1^1\right)^2 + \left(x_2^2 - x_2^1\right)^2} \tag{1}$$

Note that this metric is an application of the Pythagorean theorem, which we can generalize to n dimensions. Now some definitions.

Definition 5 *The open ball with center x^0 and radius $\varepsilon > 0$ (a real number) is the subset of points in $\mathbb{R}^n$:*
$B_\varepsilon\left(x^0\right) \equiv \left\{x \in \mathbb{R}^n : d\left(x^0, x\right) < \varepsilon\right\}$, where $d\left(x^0, x\right)$ is the distance from x^0 to x
The closed ball with center x^0 and radius $\varepsilon > 0$ (a real number) is the subset of points in $\mathbb{R}^n$:
$B_\varepsilon^*\left(x^0\right) \equiv \left\{x \in \mathbb{R}^n : d\left(x^0, x\right) \leq \varepsilon\right\}$, where $d\left(x^0, x\right)$ is the distance from x^0 to x

Note that the difference between an open ball and a closed ball (besides the asterisk) is in the inequality. So, an open ball contains all the points within a circle of given radius ε but NOT the boundary of the circle, while a closed ball contains all the points within a a circle of given radius ε including the boundary points (circle is for $\mathbb{R}^2$). So if a set is open then any element of the set can be contained within SOME open ball (not every, but at least one) centered at that point. Think of the open unit interval $(0, 1)$ (the term open kind of gives it away). Even for points very close to 0 and 1 we can draw SOME open ball around them and all the points in the open ball will be in the set. Now, think of the closed unit interval $[0, 1]$. A set is closed if its complement (the complement is the set of all elements in the universal set – in our current example, the real number line – that are not in our set) is open. The complement to the closed unit interval is $(-\infty, 0) \cup (1, \infty)$, which is an open set.

Definition 6 *A set S is bounded if it is entirely contained within some open or closed ball. That is, S is bounded if there exists some $\varepsilon > 0$ such that $S \subset B_\varepsilon(x)$ or $S \subset B_\varepsilon^*(x)$ for some $x \in \mathbb{R}^n$.*

Basically, for $\mathbb{R}^2$, if one can draw a circle around the set, then it is bounded.

Definition 7 *A set S is open if for all of its elements $x \in S$, there exists some $\varepsilon > 0$ such that the open ball $B_\varepsilon(x)$ is in the set.*

Definition 8 *A set S is closed if its complement S^c is open.*

Definition 9 *A set S is compact if it is closed and bounded.*

Why do we need this information? When we get to consumer theory it will be useful to show that our budget set, $B_{p,w}$ is a compact set.

Definition 10 *A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point x^0 if, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $d\left(x, x^0\right) < \delta$ implies that $d\left(f\left(x\right), f\left(x^0\right)\right) < \varepsilon$. A function is called a continuous function if it is continuous at every point in its domain.*

We will not delve too much into the notion of continuity – it is a fairly simple concept but the simplest concepts tend to be some of the most difficult to explain.

There is a nice result called the Weierstrass Theorem (extreme value theorem) that states that a continuous function attains a maximum (as well as a minimum) on any compact set. While we will not go through the details, it is a useful theorem to have knowledge of as it will guarantee that a solution exists to our consumer's optimization problem.

6 Real-valued functions

A function is real-valued if it maps the elements of the domain D into the real number line $\mathbb{R}$. Formally:

Definition 11 *A function $f : D \to R$ is a real-valued function if D is any set and $R \subset \mathbb{R}$.*

Generally we will deal with functions that either rise or fall as the domain increases. These are increasing and decreasing functions.

Definition 12 *Let $f : D \to \mathbb{R}$, where $D \subset \mathbb{R}^n$. Then f is* ***increasing*** *if $f\left(x^0\right) \geq f\left(x^1\right)$ whenever $x^0 \geq x^1$. If we have $f\left(x^0\right) > f\left(x^1\right)$ when $x^0 >> x^1$, then f is* ***strictly increasing****. If we have $f\left(x^0\right) > f\left(x^1\right)$ whenever $x^0 \neq x^1$ and $x^0 \geq x^1$, then f is* ***strongly increasing****.*

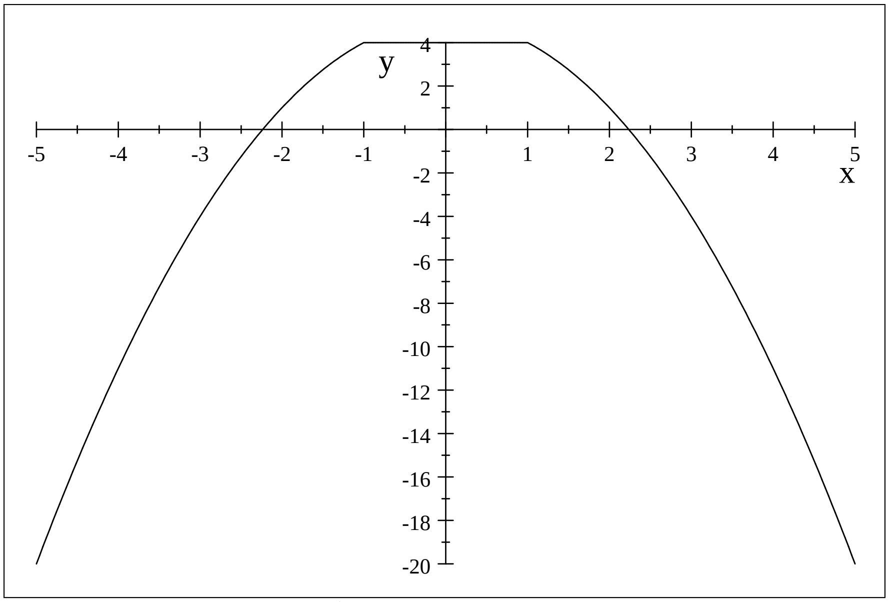
Note that every strongly increasing function is strictly increasing, and every strictly increasing function is increasing.

Definition 13 *Let $f : D \to \mathbb{R}$, where $D \subset \mathbb{R}^n$. Then f is* ***decreasing*** *if $f\left(x^0\right) \leq f\left(x^1\right)$ whenever $x^0 \geq x^1$. If we have $f\left(x^0\right) < f\left(x^1\right)$ when $x^0 >> x^1$, then f is* ***strictly decreasing****. If we have $f\left(x^0\right) < f\left(x^1\right)$ whenever $x^0 \neq x^1$ and $x^0 \geq x^1$, then f is* ***strongly decreasing****.*

6.1 Functions over convex sets

Note that a general assumption that we will make is that when we have $f : D \to R$ is a real-valued function we assume $D \subset \mathbb{R}^n$ is a convex set. Also note that we define $x^t \equiv tx^1 + (1-t)\,x^2$ for $t \in [0,1]$. That being stated, we now define various classes of functions.

Definition 14 *A function $f : D \to R$ is a concave function if for all x^1, $x^2 \in D$, $f\left(x^t\right) \geq tf\left(x^1\right) + (1-t)\,f\left(x^2\right)\ \forall t \in [0,1]$.*

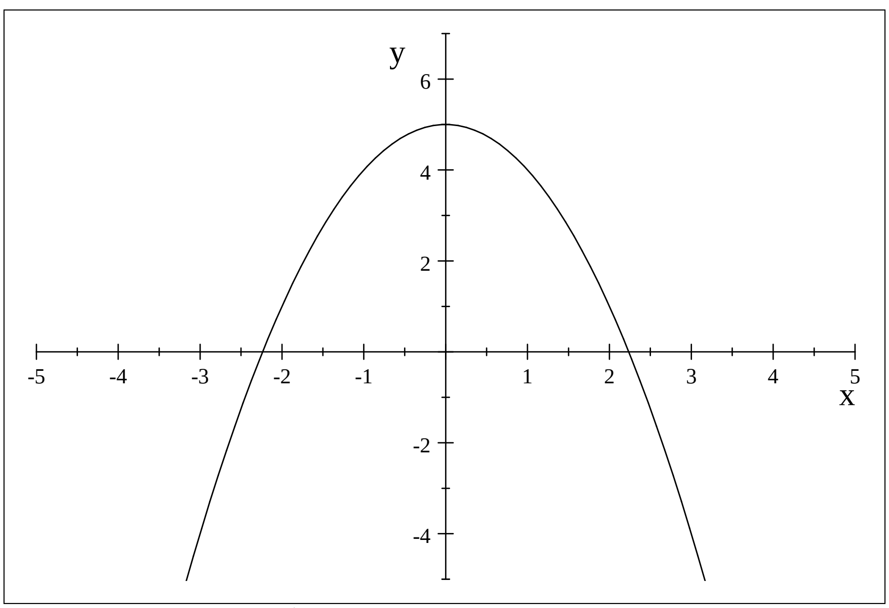


A concave function.

With a concave function we may have flat spots in the function, as in Figure 6.1. Note that all points below a concave function form a convex set.

Definition 15 *A function $f : D \to R$ is a strictly concave function if for all x^1, $x^2 \in D$, $f(x^t) > tf(x^1) + (1-t) f(x^2)$ $\forall t \in (0,1)$.*

Note that this definition of strictly concave function is the same as the one for a concave function except (1) the inequality is now strict and (2) the interval $(0,1)$ is now open.



A strictly concave function.

Figure 6.1 plots a strictly concave function. Again, all the points below a strictly concave function form a convex set.

Definition 16 *A function $f : D \to R$ is quasiconcave if and only if, for all x^1 and x^2 in D, $f(x^t) \geq \min[f(x^1), f(x^2)]$ for all $t \in [0,1]$.*

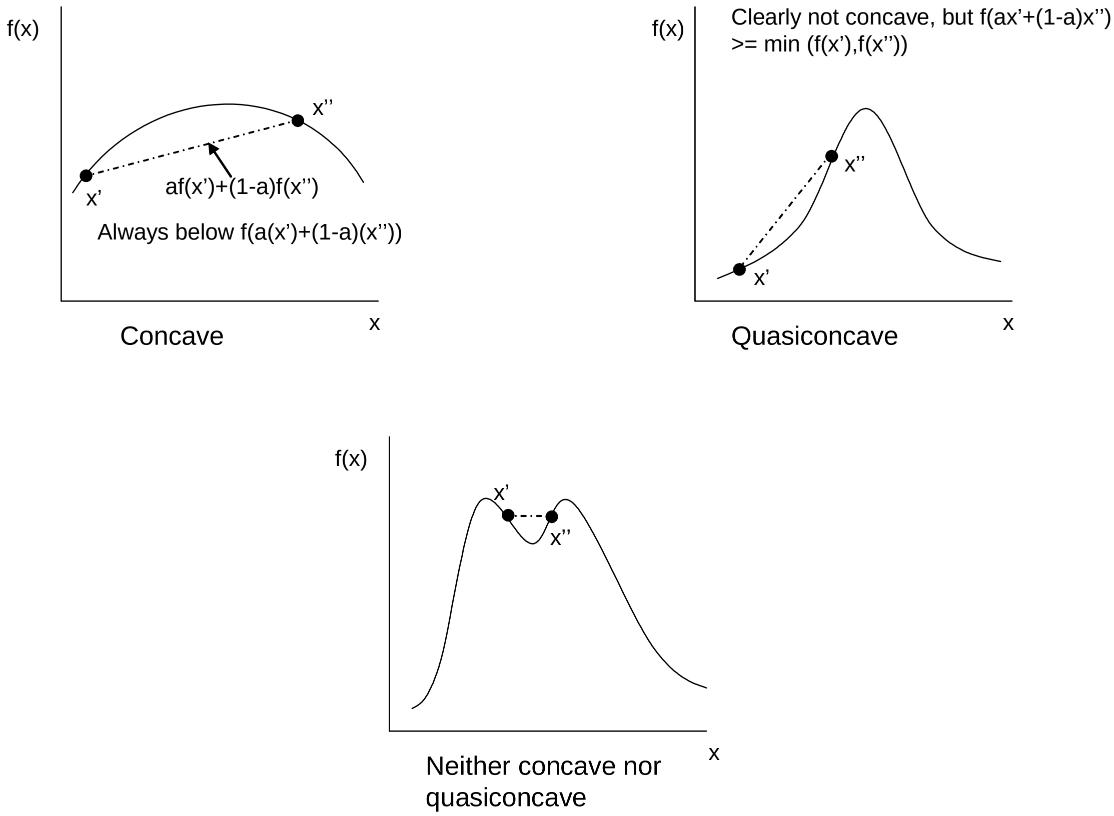


Figure 2: A concave and quasiconcave function as well as one that is neither.

Think about what quasiconcavity means – if we take a weighted average of two points, then the value of that weighted average is greater than or equal to the minimum of the two original points. For strict quasiconcavity replace $\geq$ with $>$ and $\alpha \in [0,1]$ with $\alpha \in (0,1)$. Comparing quasiconcavity with concavity, for a function to be concave, we need $f(tx' + (1-t)x'') > tf(x') + (1-t)f(x'')$ $\forall t \in (0,1)$. Concave simply says that for any two points in the function, any weighted average of the two points evaluated by the function is greater than weighted average of the evaluated values of the two points. Thus, if we pick any two points on the function and draw a line between them the function will lie above that line. Note that concavity is a stronger assumption than quasiconcavity, and the goal is to provide the weakest possible assumptions to obtain various results. Figure 2 shows the difference between a concave, quasiconcave, and a function which is neither concave nor quasiconcave.

Theorem 17 *A function $f(x)$ is (strictly) concave if and only if $-f(x)$ is (strictly) convex.*

Essentially this just involves rearranging the inequality in the definition of a concave function. Note that all points above a convex function will form a convex set. At this point in time it is useful to note that there is a difference between *convex functions* and *convex sets*.

Definition 18 *A function $f : D \to R$ is quasiconvex if and only if, for all $x^1, x^2 \in D$, $f(x^t) \leq \max\left[f\left(x^1\right), f\left(x^2\right)\right]$ $\forall t \in [0,1]$.*

Definition 19 *A function $f : D \to R$ is strictly quasiconvex if and only if, for all $x^1 \neq x^2 \in D$, $f(x^t) < \max\left[f\left(x^1\right), f\left(x^2\right)\right]$ $\forall t \in (0,1)$.*