Course Notes: Public Economics

Greg Leo

Contents

I Prologue 6

1 Binary Relations and Preferences 6
- 1.1 Definition . . . 6
- 1.2 Properties of Relations . . . 6
- 1.3 Preference Relation . . . 7
- 1.4 Indifference and Strict Preference . . . 7
- 1.5 Why Complete and Transitive? . . . 8
- 1.6 Indifference Curves and Other Sets . . . 10
- 1.7 Exercises . . . 10

2 Utility 11
- 2.1 Definition . . . 12
- 2.2 Ordinal Utility / Cardinal Utility . . . 13
- 2.3 Exercises . . . 14

II Public Decision Making 15

3 Framework 15
- 3.1 Public vs Private . . . 15
- 3.2 Ordinal Models . . . 15
- 3.3 Cardinal Models . . . 16
- 3.4 Exercises . . . 17

4 Pareto 18
- 4.1 Pareto Dominance . . . 18
- 4.2 Pareto as a Relation . . . 18
- 4.3 Strict Pareto Dominance . . . 19
- 4.4 Pareto Efficiency . . . 20
- 4.5 Weak Pareto Efficiency . . . 20
- 4.6 Pareto Efficiency Under Cardinal Preferences . . . 21
- 4.7 Geometry of Pareto Efficiency . . . 22

4.8 Exercises 23

5 Social Preferences **24**

5.1 Definition 25

5.2 Strict Preferences 25

5.3 Preference Aggregation Rule 25

6 Some Preference Aggregation Rules **26**

6.1 Dictatorship 26

6.2 Unanimity Rule 27

6.3 Majority Rule 28

6.3.1 Condorcet Paradox 29

6.4 Copeland's Method 29

6.5 Borda 30

6.6 Exercises 31

7 Preference Aggregation Properties **31**

7.1 Basic Properties 31

7.2 Classifying Preference Aggregation Rules 32

7.2.1 Dictatorship 32

7.2.2 Unanimity 32

7.2.3 Majority Rule 32

7.2.4 Copeland's Method 32

7.2.5 Borda 33

7.2.6 A Chart of Preference Aggregation Rules 33

7.3 Independence of Irrelevant Alternatives 33

7.4 Why IIA Matters 34

7.4.1 Example: Borda Count 34

7.4.2 Example: Copeland's Method 35

7.5 Arrow's Impossibility 35

7.6 Exercises 36

8 Making Choices **36**

8.1 Social Choice from Preference Aggregation 37

8.2 Some Social Choice Functions 37

8.2.1 Dictatorship- Social Choice . . . 37
8.2.2 Unanimity Rule- Social Choice . . . 38
8.2.3 Plurality Vote- Social Choice . . . 38
8.2.4 Borda Count- Social Choice . . . 39

9 Social Choice Function Properties **40**
9.1 Basic Properties . . . 40
9.2 Classifying Social Choice Functions . . . 40
9.2.1 Unanimity and Majority . . . 40
9.2.2 Plurality Vote . . . 41
9.2.3 Borda . . . 41
9.2.4 A Chart of Social Choice Functions . . . 42
9.3 Why IIA Matters for Social Choice . . . 42
9.4 Arrow's Impossibility Again . . . 43
9.5 exercises . . . 43

10 Strategic Voting **43**
10.1 Manipulation . . . 44
10.2 Another Impossibility . . . 44
10.3 Exercises . . . 45

III Appendix **46**

A Optimization **46**
A.1 Unconstrained One-Dimensional Optimization . . . 46
A.2 Unconstrained Multi-Dimensional Optimization . . . 47
A.3 Constrained Multi-Dimensional Optimization . . . 49
A.4 Contours . . . 50
A.5 Monotonicity . . . 52
A.6 Three Possibilities for an Optimal Point . . . 52
A.7 Slopes of Contours . . . 54
A.8 Solution - Two Dimensional Constrained Optimization . . . 55
A.9 Gradients . . . 56
A.10 First Order Condition and Solution . . . 56

A.11 Lagrange 57
A.12 Examples 58
A.13 Exercises 59

A More Preference Aggregation Rules **59**
A.1 Plurality Vote 59
A.2 Veto 60

IV Solutions **61**

B Chapter 1 **62**

B Chapter 2 **64**

Part I

Prologue

1 Binary Relations and Preferences

1.1 Definition

In mathematics, a binary relation is a concept that describes a relationship between things. They allow us to express various kinds of relationships.

Definition 1.1: Relation. Formally, a **binary relation** R on A is a subset of the Cartesian product $A \times A$. That is, $R \subseteq A \times A$. If $(a, b) \in R$, then we say that a is related to b by R, often denoted as aRb.[a]

[a]Technically, a relation can be *between* two different sets A and B, but in this course we are usually representing relationships among elements of a single set.

This is a somewhat intense definition, but relations are very familiar. Here are some examples of mathematical binary relations.

Example 1.1: Siblings. Let R be a relation on the set of all people. If Laura l and Mike m are siblings, then $(l, m) \in R$. We can also write lRm. Similarly $(m, l) \in R$ and we can write mRl.

Example 1.2: Friends. Friend of: Let R be a relation on the set of people where aRb means "person a is a friend of person b". If Michael m and Sarah s are friends, then mRs. Similarly sRm.

Example 1.3: Height. Human height: Let R be the "at least as tall as relation on the set of people" where aRb means "person a is at least as tall as person b". For example, if John j is taller than Alice a, then jRa. Notice that unlike the previous two examples, we would not say aRj since Alice is not at last as tall as John.

As we can see, binary relations can capture a wide range of relationships.

1.2 Properties of Relations

Notice how in the examples Example 1.1 (friends) and Example 1.2 (siblings) there is a symmetry to the relation. If person a is a sibling of person b then person b is also a sibling of a. The same is the case with friends (I think). In either case, if aRb, then bRa. We say that such a relation is *symmetric*. Can you think of some other relations on the set of humans that are symmetric?

There are many properties such as *symmetry* that we should know about. Here is a list of some properties a relation can have.

Definition 1.2: Reflexive. A relation R on a set A is reflexive if every element is related to itself, i.e., **Formally:** $\forall a \in A, (a, a) \in R$.

Definition 1.3: Complete. A relation R on a set A is total if every pair of elements is related in at least one direction.
Formally: $\forall a, b \in A$, aRb or bRa or both.

Definition 1.4: Transitive. A relation R on a set A is transitive if a is related to b and b is related to c, then a is related to c.
Formally: $\forall a, b, c \in A, aRb \& bRc \Rightarrow aRc$.

Definition 1.5: Symmetric. A relation R on a set A is symmetric if any time a is related to b, then b is also related to a.
Formally: $\forall a, b \in A, aRb \Rightarrow bRa$.

Definition 1.6: Asymmetric. A relation R on a set A is asymmetric if any time a is related to b, then b is *not* related to a
Formally: $\forall a, b \in A, aRb \Rightarrow b \not R a$.

1.3 Preference Relation

A preference relation is a set of statements about outcomes, objects, or pairs of bundles. The statement x *is at least as good as* y is shortened to $x \succsim y$.

Definition 1.7: Rational Preference Relation. A rational preference relation is a *complete* and *transitive* preference relation $\succsim$ where we interpret the statement $a \succsim b$ as "a is at least as good as b.

Info 1.1: What it means to be rational.. There is a lot of misunderstanding about the formal meaning of the word rational in economics, even among economists' textbook writers. Rationality has little to do with self-interest, being fully informed, or happiness. Though rationality certainly does not preclude these things.
Rational consumers have preferences. Preferences allow the consumer to rank alternatives (ties are allowed). They can have any ranking they want. Rational consumers choose the highest ranked alternative among the set of alternatives they can afford.
Economists represent these rankings with a utility function that gives higher ranked alternatives a higher score. Representing preferences with a utility function allows economists to use the tools of mathematics to study choices.

1.4 Indifference and Strict Preference

Let (x_1, x_2) represent bowls of ice cream x_1 scoops of vanilla and x_2 scoops of chocolate.

Suppose that someone likes a scoop of vanilla more than a scoop of chocolate. Then the following would be true for them: $(1,0) \succsim (0,1)$. They might also like *any* the number of scoops of vanilla more than the same number of chocolate. Then the following would also be true for their preferences: $(2,0) \succsim (0,2)$ and $(3,0) \succsim (0,3)$ and $(100,0) \succsim (0,100)$.

The following is true for a consumer who does not care about flavor at all just the total amount of ice cream: $(1,0) \succsim (0,1)\,, (0,1) \succsim (1,0)$. Notice that we have both $(1,0) \succsim (0,1)$ and $(0,1) \succsim (1,0)$. That is, a scoop of vanilla is just as good as a scoop of chocolate and a scoop of chocolate is just as good as a scoop of vanilla. When this is the case, we say that the consumer is indifferent and write $(1,0) \sim (0,1)$.

Definition 1.8: Indifference Relation. $a \sim b$ when $a \succsim b$ and $b \succsim a$. "a is indifferent to b".

If a consumer is not indifferent between two things, we say that they have strict preference.

Definition 1.9: Strict Preference Relation. $a \succ b$ when $a \succsim b$ and $b \not\succsim a$. "a is strictly preferred to b"

Note that $\succsim$ is symmetric and $\succ$ is asymmetric.

1.5 Why Complete and Transitive?

You might wonder why completeness and transitivity are the two key assumptions we make. Here's why.

Economics is about choice. We assume people draw on their preferences to make these choices. Suppose we have the following complete and transitive preference relation on the set $\{a, b, c\}$.

$$a \succsim b, a \succsim c, b \succsim c, a \succsim a, b \succsim b, c \succsim c$$

. Below, I have created a plot of these preferences where there is an arrow pointing from one letter to another if the first is preferred to the second. For instance, there is an arrow pointing from a to b since $a \succsim b$. I have left off the arrows from each letter to themselves since they do not add much to this figure. Just know that technically they should also be there.

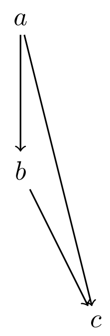


Figure 1.1: A Complete and Transitive Relation on $\{a, b, c\}$

Look how such a complete and transitive relation creates a natural ordering of the objects. Things higher up, like a, are better than everything lower. Even when we get more objects and some indifferences, the same kind of shape appears again. Let's plot the following complete and transitive relation (it is much easier to visualize with the graph).

$$a \succsim b, b \succsim a, a \succsim c, a \succsim d, b \succsim c, b \succsim d, c \succsim d, d \succsim c, a \succsim e,$$
$$b \succsim e, c \succsim e, d \succsim e, a \succsim a, b \succsim b, c \succsim c, d \succsim d, e \succsim e$$

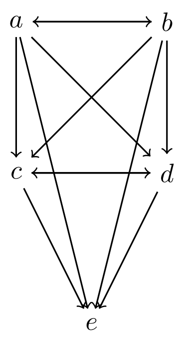


Figure 1.2: A Complete and Transitive Relation on $\{a, b, c, d, e\}$

One nice property of such a shape is that, whatever set of objects a consumer might have to choose from, there is some object in that subset that is at least as good as everything else in that set. In the above example, for instance, a and b are at least as good as everything in $\{a, b, c, d, e\}$. b is at least as good as everything in $\{b, d, e\}$. Whatever set we pick, there is at least one object like that. For whatever set the consumer might be asked to choose from, there is at least one **best** object– something they would be happy to choose.

> **Definition 1.10: Best.** x is **best** from some set B (that includes x) if $x \succsim y$ for every y in B.

We sometimes denote the set of **best** outcomes/options from a set as $C(B)$. This is called the **choice function**. It is the set of best things from a set. Or rather, the things the consumer would be willing to choose. As an example, for the preferences graphed above, $C(a, b, c, d, e) = \{a, b\}$ and $C(b, c, e) = \{b\}$.

What if we make the relation incomplete by removing $a \succsim c$, what is best from the set of $\{a, c\}$? There is nothing as good as everything else, because the consumer has no idea how to compare a and c. That is $C(\{a, c\}) = \emptyset$. So, somewhat trivially, when a relation is not complete, there are menus of objects that the consumer cannot choose form. In a less trivial way, this also happens when we make the relation intransitive.

Suppose we have the following complete but **intransitive** preference relation on the set $\{a, b, c\}$.

$$a \succsim b, b \succsim c, c \succsim a, a \succsim a, b \succsim b, c \succsim c$$

.

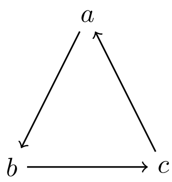


Figure 1.3: A Complete and Intransitive Relation on $\{a, b, c\}$

What is best from $\{a, b, c\}$? That is, what is $C(\{a, b, c, \})$. There is nothing at least as good as everything else. a is not at least as good as c, b is not at least as good as a, c is not at least as good as b. What would the consumer choose?!? We have $C(\{a, b, c\}) = \emptyset$. Intransitive preferences create these kinds of cycles (look at the figure again), and when there are cycles, there are sets that the consumer cannot choose from.

1.6 Indifference Curves and Other Sets

For every object, we can use the preference relation to define a few sets. $\succsim (x)$ is the set of objects that is at least as good as x. $\succ (x)$ is the set of objects that is strictly better than x. $\sim (x)$ is the set of objects indifferent to x.

Definition 1.11: Weakly Preferred Set. The set of points weakly preferred to x is: $\succsim (x) = \{y | y \in X, y \succsim x\}$

Definition 1.12: Strictly Preferred Set. The set of points strictly preferred to x is: $\succ (x) = \{y | y \in X, y \succ x\}$

Definition 1.13: Indifference Set. The set of points indifferent to x is: $\sim (x) = \{y | y \in X, y \sim x\}$

Sets of indifferent bundles are very important in studying preferences. We call such a set of bundles an "indifference curve". We use indifference curves to visualize preferences. Note: There are many indifference curves. We only sketch a few to get an idea of the "shape" of the preferences.

1.7 Exercises

For the relations R below, when a pair is not listed, assume that the relation is not true of that pair.

Exercise 1.1: Is the relation 'is a sibling of' on the set of all people complete? Is it transitive? Is it symmetric?

Exercise 1.2: Is the relation 'is at least as tall as' on the set of all people complete? Is it transitive? Is it symmetric?

Exercise 1.3: Is the relation 'has same birthday as' on the set of all people complete? Is it transitive? Is it symmetric?

Exercise 1.4: Explain why a relation that is complete and symmetric is trivial in the sense that it relates all pairs to each other.

Exercise 1.5: For the set $X = \{x, y, z\}$, identify if the following relations are transitive:

1. $R : xRy, yRz, xRz$
2. $R : xRx, yRy, zRz$
3. $R : xRy, yRz, zRx$

Exercise 1.6: For the set $X = \{p, q, r\}$, identify if the following relations are complete and transitive. When a relation is not both of these, say which assumption fails and why.

1. $R : pRp, qRq, rRr, pRq, qRr$
2. $R : pRp, qRq, rRr, pRq, qRr, pRr$
3. $R : pRp, qRq, rRr, pRq, qRp, qRr, rRq, pRr, rRp$
4. $R : pRp, qRq, rRr, pRq, qRp, pRr$

Exercise 1.7: Consider the preference relation that describes someone's preferences over left l and right r shoes, where they only care about the number of usable pairs of shoes they consume. Sketch the indifference curves $\sim (1, 1)$ and $\sim (2, 2)$ on graph that has l on the x-axis and r on the y-axis. Label the set $\succ (2, 2)$.

Exercise 1.8: Consider the preference relation that describes someone's preferences for red apples r and green apples g, where they only care about the total number of apples they have but not the color. Sketch the indifference curves $\sim (1, 1)$ and $\sim (2, 2)$ on graph that has r on the x-axis and g on the y-axis. Label the set $\succ (2, 2)$.

Exercise 1.9: Plot the following rational preference relation using a graph similar to those used in this chapter.

$$a \succsim b, a \succsim c, a \succsim d, b \succsim c, b \succsim d, c \succsim b, c \succsim d, a \succsim a, b \succsim b, c \succsim c, d \succsim d$$

2 Utility

A utility function is a way to assign "scores" to bundles, so that better bundles according to $\succsim$ get a higher score. Utility functions allow us to use familiar tools of mathematics to study preferences.

Let's return to our plot of a complete and transitive preference relation from the last chapter. Recall that here, things higher up are better than anything lower down. It is possible to graph preferences this way as long as preferences are complete and transitive. This time, let's add some numbers to each level of the graph where things higher up get higher numbers.

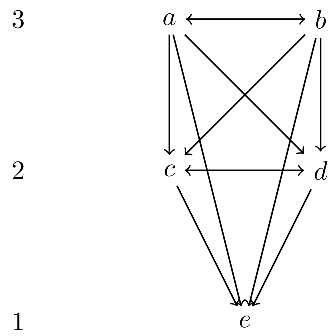


Figure 2.1: A Complete and Transitive Relation on $\{a, b, c, d, e\}$ with Utility

Notice that the number represents "how good" an object is here. a and c get a number 3. They are both indifferent to each other, but strictly better than everything else. c and d get the number 2. They are indifferent to each other but strictly better than e which gets the number 1. We can think of these numbers as "scores" that represent the preferences. In fact, this is precisely what we call a utility function.

2.1 Definition

Definition 2.1: Utility Function. A utility function $U(x)$ represents preferences $\succsim$ when, for every pair of bundles x and y, $U(x) \geq U(y)$ if and only if $x \succsim y$.

Example 2.1: Utility Function Example. Suppose we have the preference relation plotted above.

$$a \succsim b, b \succsim a, a \succsim c, a \succsim d, b \succsim c, b \succsim d, c \succsim d, d \succsim c, a \succsim e,$$
$$b \succsim e, c \succsim e, d \succsim e, a \succsim a, b \succsim b, c \succsim c, d \succsim d, e \succsim e$$

Written more succintly, $a \succ b \succ c \sim d \succ e$.
Some utility functions that represent these preferences are $U(a) = 10$, $U(b) = 10$, $U(c) = 5, U(d) = 5, U(e) = 0$ and also $U(a) = 12$, $U(b) = 12$, $U(c) = 10, U(d) = 10, U(e) = -100$.

Note that utility functions simply represent the underlying preference relation $\succsim$ of a consumer. When there is a large number of alternatives, the preference relation itself can be cumbersome to work with. However, a utility function can effectively characterize a preference relation in a succinct way.

Example 2.2: Perfect Substitutes Preferences. Suppose a consumes red apples r and green apples g. They like green apples twice as much as red apples, so they would always give up two red apples in exchange for one green apple.
For combinations of red apples and green apples (r, g) this consumer has preference where, for example:

$(2, 0) \sim (0, 1)$
$(2, 1) \succ (0, 1)$
$(0, 1) \succ (1, 0)$

We can summarize these preferences with the utility function $u(r, g) = r + 2g$.

Example 2.3: Perfect Complements Preferences. Suppose a consumes only apple pies. An apple pie is made from apples a and crusts c. It takes exactly 2 apples and 1 crust to make a pie.

For combinations of apples and crusts (a, c) this consumer has preference where, for example:

$(2, 0) \sim (0, 1)$ (Since both make zero pies.)
$(2, 1) \sim (2, 2)$ (Since both make one pie.)
$(4, 2) \succ (2, 1)$ (Since the first makes two and the second makes one pie.)

We can summarize these preferences with the utility function $u(r, g) = min\left\{\frac{1}{2}a, c\right\}$.

2.2 Ordinal Utility / Cardinal Utility

Often, the magnitude of utility is meaningless and only the relationships between scores matter. In this case, we say that the utility is **ordinal**. In Example 2.1, the fact that $U(a) = 10$ and $U(c) = 5$ do not imply that a is twice as good as c. In fact, in the second set of utilities, a gets a utility that is only 1.2 times greater than c.

Sometimes, however, there is meaningful information encoded in a particular representation. Suppose that we have a consumer who consumes two things. t tacos and m money, and their preferences can be represented by $u(t, m) = \sqrt{t} + m$. The utility of the combination $u(4, 10) = 12$ is the same as the utility of $u(0, 12) = 12$. In terms of preferences $(4, 10) \sim (0, 12)$. The utility function directly encodes the amount of money (and no tacos) that some combination is worth to the consumer. **Utility is measured in terms of dollars**. In this sense, the bundle $(36, 18)$ which has utility $u(36, 18) = 24$ is worth "twice" as much as the bundle $u(4, 10) = 12$. The utility is measured in terms of some tangible thing, in this case money. When this is the case, we say that the utility function is **cardinal**.

Definition 2.2: Ordinal Utility. Utility function in which utility numbers have no meaning beyond relative comparisons.

Definition 2.3: Cardinal Utility. Utility function in which utility numbers are measured in terms of something with meaningful magnitude (like money).

One common utility function that we will use in this course is the *quasi-linear* utility function that measures everything in terms of money.

Definition 2.4: Dollar-Denominated Quasi-Liner Utility. A utility function of the form $u(x, m) = f(x) + m$ where m is money.

With this utility function, the utility number $u = u(x, m)$ says "the combination (x, m) worth the equivalence of $\$u$ to the decision maker.

2.3 Exercises

Exercise 2.1: Consider bundles A, B, and C with the given utilities $U(A) = 8$, $U(B) = 15$, and $U(C) = 10$. What complete and transitive relation $\succsim$ does this represent?

Exercise 2.2: Provide an alternative utility function that represents the same preferences as those in the previous exercise.

Exercise 2.3: Suppose that a consumer's preferences can be represented by the utility function $u(t, m) = \sqrt{t} + m$. Which is true of this consumer's preferences? $(16, 3) \succ (4, 5)$, $(4, 5) \succ (16, 3)$, or $(4, 5) \sim (16, 3)$

Exercise 2.4: Suppose that a consumer's preferences can be represented by the utility function $u(t, m) = \sqrt{t} + m$. What is the utility of $(9, 4)$. What amount of money m solves the following $(9, 4) \sim (0, m)$?

Exercise 2.5: Suppose that a consumer's preferences can be represented by the utility function $u(t, m) = \sqrt{t} + m$. Sketch the indifference curve $\sim (9, 4)$ on a graph with t on the x-axis and m on the y-axis. Feel free to use a computer to help you with this.

Exercise 2.6: Discuss the following statement: *Economists do not have to believe that utility functions exist in the minds of consumers for the concept to be useful.*

Part II

Public Decision Making

3 Framework

3.1 Public vs Private

We now begin our study of public economics in earnest. The way in which public economics differs from the type of economics you might have studies in a course like intermediate economics is that in this course, we focus on situations, outcomes, or choices that affect more than one person. Here, we will differentiate between *private outcomes* and *public outcomes*.

> **Example 3.1: Private Outcome.** Alice is in her studio apartment on Saturday afternoon. She decides to microwave a leftover fish to enjoy for lunch. Besides us, she is the only one who will ever know this happened.

When studying private outcomes, we assume that an individual assesses the available options according to their preferences and chooses their favorite alternative. This is the end of the story with private choice. Choices are optimal, or they are not.

This course examines the complexities that arise when outcomes are **public** in nature. A **public outcome** impacts multiple individuals. In contrast to private choice, what is optimal for one person may be suboptimal for others.

> **Example 3.2: Public Outcome.** Alice is at work on a Monday afternoon. She decides to microwave a leftover fish to enjoy for lunch. The lingering smell of warm fish reduces office productivity for three days.

3.2 Ordinal Models

The real world is complex. The value of creating theoretical models is that they simplify scenarios down to their core elements. For many scenarios we want to study in this course, I think we can get away with focusing on three elements:

- Who are the people involved?
- What are the potential outcomes?
- What are each person's preferences over those outcomes?

For *ordinal models*, we define only ordinal preference relations over the outcomes.

> **Definition 3.1: Ordinal Model.** An *ordinal public outcome model* is: O: the set of outcomes. P: the set of decision makers. And for every decision maker $i \in P$: $\succsim_i$ their preferences over the set O.

Let's look at an example. In this scenario, Alice and Bob are co-workers. Alice sometimes microwaves fish. Bob hates the smell of microwaved fish. Here we have people Alice a and Bob b and there are two outcomes "Alice microwaves fish" (y) and "Alice does not microwave fish" (n).

Example 3.3: Microwaving Fish.

$$\begin{aligned} P &= \{\text{a}, \text{b}\} \\ O &= \{\text{yes}, \text{no}\} \\ \text{yes} &\succ_a \text{no} \\ \text{no} &\succ_b \text{yes} \end{aligned}$$

Let's look at a slightly more complex example. Alice and Bob share the office kitchen. Sometimes, it needs to be cleaned. It can be cleaned by only one person, or the work can be shared. However, both prefer that the kitchen be clean, even if that means doing the work alone. Let's formalize this model.

Example 3.4: Cleaning the Kitchen.

$$\begin{aligned} P &= \{\text{a}, \text{b}\} \\ O &= \{\text{both,Alice,Bob,neither}\} \\ \text{Bob} \succ_a \text{both} \succ_a \text{Alice} &\succ_a \text{neither} \\ \text{Alice} \succ_b \text{both} \succ_b \text{Bob} &\succ_b \text{neither} \end{aligned}$$

3.3 Cardinal Models

In some situations, we need or have a little bit more information– not only relative preferences over outcomes, but also the strength of those preferences measured through some common cardinal utility measure, like there value of each outcome in terms of dollars. To model these scenarios, we use a *cardinal model*. Instead of defining the preference reltion $\succsim_i$ for each person, we define their utility function u_i over the outcomes.

Definition 3.2: Cardinal Model. A *cardinal public outcome model* is: O: the set of outcomes. P: the set of decision makers. And for every decision maker $i \in P$: $u_i()$ their cardinal utility function over the set O.

Let's update Example 3.3 and Example 3.4 to cardinal models. For both, let u_a be the utility of Alice and u_b be the utility of Bob. Let's assume that these utilities are measured in terms of dollars so that we have a valid means for assessing the magnitudes of utility. That is, these are cardinal utilities.

Example 3.5: Microwaving Fish: Utility Version.

$$P = \{\text{a,b}\}$$
$$O = \{\text{yes,no}\}$$
$$u_a(o) = \begin{cases} 10 & o = \text{yes} \\ 9 & o = \text{no} \end{cases}$$
$$u_b(o) = \begin{cases} 1 & o = \text{yes} \\ 10 & o = \text{no} \end{cases}$$

One way for us to interpret these utilities is that while Alice would pay up to \$1 to be able to microwave fish, Bob would pay up to \$9 to prevent it.

Example 3.6: Cleaning the Kitchen: Utility Version.

$$P = \{a, b\}$$
$$O = \{\text{both, Alice, Bob, Neither}\}$$
$$u_a(o) = \begin{cases} 12 & \text{if } o = \text{both} \\ 10 & \text{if } o = \text{Alice} \\ 25 & \text{if } o = \text{Bob} \\ 5 & \text{if } o = \text{neither} \end{cases}$$
$$u_b(o) = \begin{cases} 12 & \text{if } o = \text{both} \\ 25 & \text{if } o = \text{Alice} \\ 10 & \text{if } o = \text{Bob} \\ 5 & \text{if } o = \text{neither} \end{cases}$$

3.4 Exercises

Exercise 3.1: Add a third person (Camden) to Example 3.3 who likes the smell of warm fish.

Exercise 3.2: Add a third person (Camden) to Example 3.4 who is so inept that if he attempts to clean, even with the help of others, he makes the kitchen worse than if no one had tried to clean at all. For preferences, *there are many right answers*, justify your answers with an explanation of why Alice, Bob, and Camden might have those particular preferences in the context of this "story".

Exercise 3.3: Add a third person (Camden) to Example 3.5 who likes the smell of warm fish. Ensure that the utilities you choose are consistent with your solution to Exercise 3.1.

Exercise 3.4: Add a third person (Camden) to Example 3.6 who is so inept that if he attempts to clean, even with the help of others, he makes the kitchen worse than if no one had tried to clean at all. Ensure that the utilities you choose are consistent with your solution to Exercise 3.2.

4 Pareto

4.1 Pareto Dominance

Imagine yourself as a benevolent ruler, making decisions for society. How would you choose outcomes? One thing that I think should be true of every *benevolent* ruler's preferences is that if there were two outcomes and one was at least as good for *everyone* as the other, the ruler should prefer that better outcome.

Think about the cleaning example in Example 3.4. How would you choose what to implement? Having both clean seems fair, but maybe only having one clean is more efficient? I don't think there is a clear best outcome, but I think essentially everyone would agree that having no one clean is not a desirable outcome. **Any** of the other outcomes are better for both Alice and Bob.

This notion of being *at least as good for everyone* is what we call "Pareto domination" . An outcome o' Pareto dominates o if o' is at least as good for everyone. Formally:

Definition 4.1: Pareto Dominates. An outcome o' Pareto dominates o if $o' \succsim_i o$ for all $i \in P$.

4.2 Pareto as a Relation

Let's define Pareto dominance as a relation. We will say aPb if outcome a Pareto dominates b according to the definition above.

Example 4.1: Pareto Dominance in Cleaning the Kitchen. In Example 3.4. For convenience, let's simplify the outcome names from $O = \{both, Alice, Bob, neither\}$ to $O = \{ab, a, b, n\}$. We can now define the Pareto relation P on this set: $abPn$, aPn, bPn, $abPab$, aPa, bPb, nPn.

Is P a complete and transitive relation? It is **always transitive** as long as all of the individuals in the model have transitive preferences. This is because, for example, if *everyone* likes a over b and everyone likes b over c then since everyone has transitive preferences, everyone will like a over c and hence a will Pareto dominate c.

However, Pareto dominance **is not always complete**. In fact, in the example above it is not complete. There is no relationship between ab and a or between ab and b or between a and b. Pareto dominance cannot compare these outcomes.

Info 4.1: Pareto Dominance is Incomplete. The **Pareto dominates** relation is always transitive but not always complete.

Since Pareto dominance is incomplete, for any given set of options, there may not be an outcome that Pareto dominates all the others. If there were, it would be pretty clearly the best outcome.

Let's look at an example of Pareto dominance.

Example 4.2: Example of Pareto Dominance for Two People.

$$a \succ_1 c \succ_1 b \succ_1 d \succ_1 e$$

$$b \succ_2 d \succ_2 a \succ_2 c \succ_2 e$$

Everything Pareto dominates e. b Pareto dominates d. a Pareto dominates c. But otherwise, the outcomes are not comparable in Pareto terms. The Pareto dominance relation P is:

$$aPc, aPe, bPd, bPe, cPe, dPe,$$
$$aPa, bPb, cPc, dPd, ePe$$

Let's plot this relation using the same type of graph from the chapter on preferences.

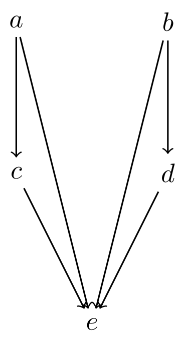


Figure 4.1: A graph Pareto dominance relation from Example 4.2

4.3 Strict Pareto Dominance

Just like with preferences, where we defined the strict preference relation such that $a \succ b$ any time $a \succsim b$ but $b \not\succsim a$, we can do the same with Pareto dominance.

Definition 4.2: Strictly Pareto Dominates. An outcome o' Strictly Pareto dominates o if $o'Po$ but $o \not{P} o'$.

What will this mean in terms of preferences of the individuals? Well, first, $o'Po$ says everyone should like o' at least as well as o. But at the same time, $o \not{P} o'$ says that it is not the case that everyone likes o at least as well as o'. This means at least one person must like o' strictly better than o. (Take a moment to convince yourself of this.) Combining these, o' will strictly Pareto dominate o if everyone likes o' at least as well **and** at least one person likes it strictly more.

Definition 4.3: Strictly Pareto Dominates. An outcome o' Strictly Pareto dominates o if $o' \succsim_i o$ for all $i \in P$ and there is some i such that $o' \succ_i o$.

In the example above, everything strictly dominates e. a strictly dominates c and b strictly dominates d. Notice that in each case, on the graph, an outcome that is strictly dominates by some other outcome has at least one one-way arrow leading into it.

4.4 Pareto Efficiency

As we have seen, Pareto dominance is not complete. In that sense, for any given set of options, there may not be an outcome that Pareto dominates all the others. If there were, it would be pretty clearly the best outcome. However, just like in our case of the cleaning example, we can at least use Pareto dominance to eliminate the clearly undesirable options. What are those clearly undesirable options? **The ones that are strictly Pareto dominated.** This is because for any outcome that is strictly Pareto dominated there must be a way to make everyone at least as well off and at least one person strictly better off. That's a win/win.

One great thing about Pareto efficiency is that, even though there will not always be some outcome that Pareto dominates all others, **there will always be at at least one Pareto efficient outcome.**

> **Definition 4.4: Pareto Efficiency.** A outcome o is Pareto efficient if there is no other outcome o' that strictly Pareto dominates o.

If we graph preferences, such as in Figure 4.1, a Pareto efficient outcome has no one-way arrow leading into it. We can also define Pareto efficiency in terms of the preferences of individuals since we have defined strict Pareto dominance in terms of individual preference above.

> **Definition 4.5: Pareto Efficiency in Terms of Individual Preferences.** A outcome o is Pareto efficient if there is no other outcome o' such that for all people $o' \succsim_i o$ and for at least one person $o' \succ_i o$.

Put another way, we have what might be the most familiar definition of Pareto efficiency.

> **Definition 4.6: Pareto Efficiency Alternative Definition.** A outcome o is Pareto efficient if there is no other outcome o' that makes someone strictly better off without making anyone else strictly worse off.

In Example 4.2, the Pareto efficient outcomes are a and b. From a there is no way to make someone better off without making someone strictly worse off. The same goes for b. Looking at the graph, they are also the only outcomes that have no one-way arrow that points to them.

In out cleaning example, a, b and ab are all Pareto efficient since, starting from any of these, there is no way to make someone strictly better off without also making someone strictly worse off.

4.5 Weak Pareto Efficiency

In some applications, we use a slightly weaker notion of Pareto efficiency.

Definition 4.7: Weak Pareto Efficiency. A outcome o is Pareto efficient if there is no other outcome o' that **makes everyone strictly better off**. That is for all people, $i \in P$, $o'_i o$.

Notice that the difference between Pareto efficiency and Weak Pareto efficient is that we can say some outcome is not Pareto efficient if we can make even just one person strictly better off while others remain just as well off. We can only say an outcome is not weak Pareto efficient if we can find a way to make **everyone** strictly better off. If something is Pareto efficient it will definitely be weakly Pareto efficient since if we can not make even one person strictly better off while others remain indifferent then we can't make everyone strictly better off.

Here is an example:

Exercise 4.1: Alice and Bob have the following preferences over outcomes $\{a, b, c, d\}$. Graph the Pareto dominance relation for this example as demonstrated in this chapter.

- Alice: $a \succ b \succ c$
- Bob: $a \sim b \succ c$
- Camden: $a \sim b \succ c$

a is Pareto efficient. We cannot make anyone strictly better off at all. What about b? It is not Pareto efficient because a makes eveyrone at least as well off and also makes Alice strictly better off. However, b is weakly Pareto efficient since we cannot make everyone strictly better off. c is neither since we can make everyone strictly better off.

4.6 Pareto Efficiency Under Cardinal Preferences

When preferences are Cardinal, that is, the preferences are given a magnitude through the utility function that is measured in terms of some natural "measuring stick" like money, we can still use Pareto efficiency. We just need to rewrite the definition in terms of utilities rather than the preference relation.

Definition 4.8: Pareto Efficiency. A outcome o is Pareto efficient if there is no other outcome o' such that strictly Pareto dominates o. That is, there is no o' such tht $U_i(o') \geq U_i(o)$ for all $i \in P$ and $U_j(o') > U_i(o)$ for at least one $j \in P$.

It is important to note that Pareto efficiency does not imply fairness or equity. An allocation can be Pareto efficient even if it is highly unequal. For example, look at the cardinal version of the "Microwaving Fish" model in 3.5. Alice likes to microwave fish, but Bob *really* hates the smell of warm fish. Like, he throws up. Both outcomes are Pareto efficient, but letting Alice cook the fish seems a little unfair.

Thus, while achieving Pareto efficiency is often a goal of policy interventions. Policymakers will often need consider other criteria, such as equity and fairness, when designing policies to social outcomes. We will return to this later. For now, let's look at a slightly more complex model to solidify understanding about Pareto efficiency.

4.7 Geometry of Pareto Efficiency

When we are using cardinal utilities, we can visualize Pareto dominance and Pareto efficiency. We begin with a plot of the utility pairs from our running example Example 3.6 shown in Figure 4.2.

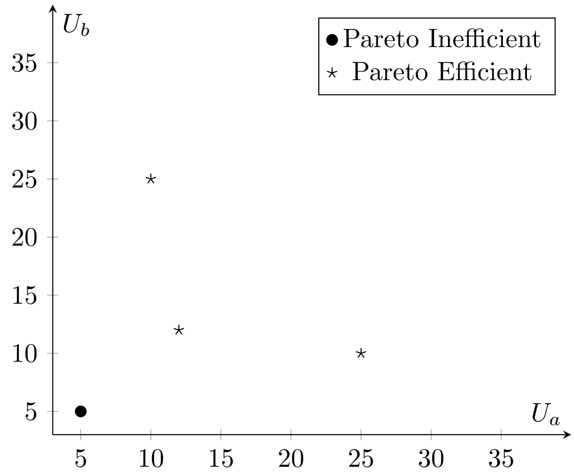


Figure 4.2: Cleaning Utility Pairs

What does it look like when a choice combination is not Pareto efficient? Have a look at Figure 4.3 which shows the possible points that Strictly Pareto dominate $(5, 5)$ from Example 3.6. Notice that $(12, 12$, $(25, 10)$ and $(10, 25)$ are all in the blue region. They all strictly Pareto dominate $(5, 5)$.

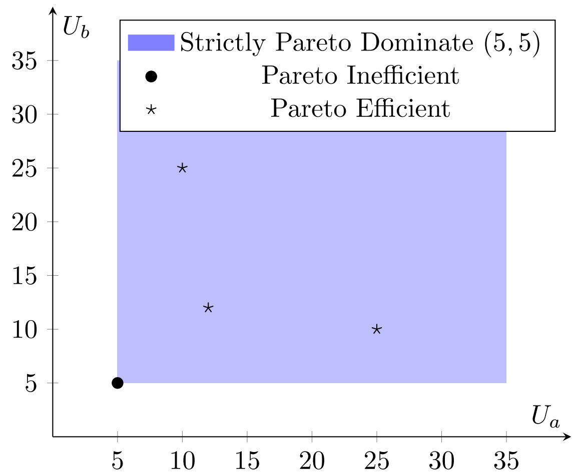


Figure 4.3: Outcomes that Strictly Pareto Dominate $(5, 5)$ from Example 3.6.

Here, $(5, 5)$ is not Pareto efficient because there are outcomes that strictly Pareto dominate it.

On the other hand, if we repeat this exercise with the point $(12, 12)$, we see there are no outcomes in the blue region. It is Pareto efficient! This shown in Figure 4.4.

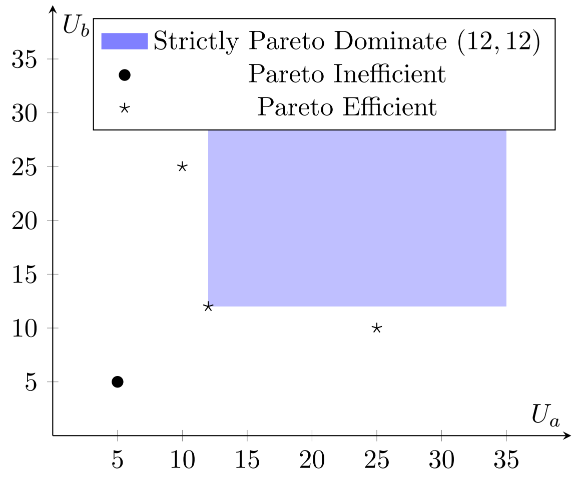


Figure 4.4: Outcomes that Strictly Pareto Dominate $(12, 12)$ **from Example 3.6.**

4.8 Exercises

Exercise 4.2: Alice and Bob have the following preferences over outcomes $O = \{a, b, c\}$. What are the Pareto efficient outcomes?

- Alice: $a \succ b \succ c$
- Bob: $b \succ a \succ c$

Exercise 4.3: Alice, Bob, and Camden have the following preferences over outcomes $O = \{a, b, c, d\}$. What are the Pareto efficient outcomes?

- Alice: $b \succ d \succ c \succ a$
- Bob: $d \succ c \succ a \succ b$
- Camden: $a \succ c \succ d \succ b$

Exercise 4.4: Alice, Bob, Camden, and Dave have the following preferences over outcomes $O = \{a, b, c, d, e, f\}$. What are the Pareto efficient outcomes?

- Alice: $a \succ b \succ c \succ d \succ e \succ f$
- Bob: $a \succ d \succ c \succ b \succ e \succ f$
- Camden: $a \succ c \succ b \succ e \succ f \succ d$
- Dave: $b \succ c \succ a \succ f \succ d \succ e$

Exercise 4.5: Alice and Bob have the following preferences over outcomes $\{a, b, c, d\}$. Graph the Pareto dominance relation for this example as demonstrated in this chapter.

- Alice: $a \sim b \sim c \succ d$
- Bob: $a \succ b \sim c \succ d$

Exercise 4.6: For the example above, which outcomes Pareto dominante others? Which outcomes strictly Pareto dominate others?

Exercise 4.7: For the example above, what are the Pareto efficient outcomes.

Exercise 4.8: Consider the following model based on Example 3.5, but where Bob can choose to leave. Plot the possibly utility combinations and mark the Pareto efficient outcomes.

$$
\begin{aligned}
P =& \{\text{Alice,Bob}\} \\
O =& \{\text{yes/stays,yes/leaves,no/stays,no/leaves}\}
\end{aligned}
$$

$$
U_a(o) = \begin{cases} 10 & o = \text{yes/stays} \\ 10 & o = \text{yes/leaves} \\ 5 & o = \text{no/stays} \\ 5 & o = \text{no/leaves} \end{cases}
$$

$$
U_b(o) = \begin{cases} 1 & o = \text{yes/stays} \\ 5 & o = \text{yes/leaves} \\ 10 & o = \text{no/stays} \\ 5 & o = \text{no/leaves} \end{cases}
$$

5 Social Preferences

Pareto efficiency is a lovely property. I think it is somewhat indisputably desirable as a property to strive for. Unfortunately, as we have seen, Pareto efficiency is not always enough to make a choice among outcomes. It is a property that is transitive but not complete. It helps us choose outcomes, but it often fails to provide complete guidance on what to choose. In a sense, it is not enough to help us construct a preference relation on how *we* might asses the outcomes.

We might want to try to extend Pareto efficiency in a way that offers a complete preference relation over outcomes that we can use to assess options on behalf of society. We call such preferences **Social Preferences** since they are preferences about the outcomes that affect society.

5.1 Definition

Definition 5.1: Social Preference Relation. A social preference relation $\succsim_*$ is a complete and transitive relation on the set of outcomes O. It is used to evaluate outcomes by an administrator to evaluate outcomes on behalf of society.

Think about your preferences about the outcomes in Example 3.3 and Example 3.4. What social preferences do you have over these outcomes?

Example 5.1: Social Preferences for Example 3.4. In Example 3.4 "Cleaning the Kitchen" a social preference relation might be the following:

$$\text{both} \succ_* \text{Alice} \sim_* \text{Bob} \succ_* \text{neither}$$

Notice that this is a preference relation that does not match the preferences of any of the individuals in the model.

5.2 Strict Preferences

For the next few chapters, I am going to assume that everyone has strict preferences over all outcomes. That is for each individual $i \in P$, there are no two *distinct* outcomes $x \neq y$ such that $x \sim_i y$. This means that everyone's preferences will look like this: $x \succ_i y \succ_i z$... with no indifferences. This will make things a little easier, but know that much of what we will discuss will apply to situations where indifferences are allowed.

When there are no indifferences, then our definition of Pareto efficiency and Pareto dominance can be updated. Recall that x Pareto dominates y if everyone likes it at least as well. But if x and y are distinct and everyone has strict preferences, then the only way this can happen is if everyone likes x **strictly more** than y.

Definition 5.2: Pareto Dominance Under Strict Preferences. If everyone's preferences are strict, then x Pareto dominates y if $x \succ_i y$ for all $i \in P$.

We can update Pareto efficiency accordingly as well.

Definition 5.3: Pareto Efficiency Under Strict Preferences. If everyone's preferences are strict, then x is Pareto efficient if there is no y that everyone likes strictly better. That is, not y such that $y \succ_i y$ for all $i \in P$.

5.3 Preference Aggregation Rule

Technically, a social preference relation can be *any* complete and transitive relation on the outcomes. But really, an administrator should consider how their constituents care about the outcomes, that is, the social preference relation should somehow be constructed by referencing the individual preferences.

Definition 5.4: Preference Aggregation Rule. A **preference aggregation rule** (also known as a social welfare function) is a way to turn individual preferences into the social preference relation. Formally, it is a *mapping* from the set of possible individual preferences over the outcomes into a social preference relation over the outcomes.

These can be somewhat hard to define using formal notation, so I will use intuitive descriptions of various rules when necessary. Let's look at a few preference aggregation rules in the context of the following two models:

6 Some Preference Aggregation Rules

In this section, I present many preference aggregation rules. These are meant to show you *some* of the possibilities and demonstrate how different rules can prioritize different kinds of goals. These should get your mind working on different ways to construct interesting rules. By the end of the section, I hope you will be able to come up with your own rules and think about how they might work in practice.

I will use the following two running examples throughout this section. Before you begin reading through each rule, think about what you would choose as the social preferences for each example.

Example 6.1: Example 1.
There are three people. They have these preferences:
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

Example 6.2: Example 2.
There are five people. They have these preferences:
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

6.1 Dictatorship

A dictatorship where there is one person whose preferences completely determine the social preferences. The preferences of everyone else are ignored.

This may seem like a strange rule, but I think it is surprisingly common. For example, the faculty of the LMU Economics department occasionally meets for lunch. The faculty rotates, being the *one* who gets to choose the restaurant. We could vote every time, or submit our rankings, or do something else. But instead, we rotate being the "dictator". It is simple and decisive, and for scenarios like this that happen over and over, rotating through "dictators" makes what would usually be a very unfair rule fair in the long-run.

Definition 6.1: Dictatorship. Pick a person $i \in P$. The social preferences are that person's preferences. $\succsim^* = \succsim_i$.

Example 6.3: Example 1.
Let's assume person 1 is the dictator.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

$$a \succ^* b \succ^* c$$

Example 6.4: Example 2.
For the examples, let's assume person 1 is the dictator.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

$$a \succ^* c \succ^* b$$

Another nice thing we can say about dictatorships is that at least they always give us a complete and transitive social preference ordering. Let's not undervalue that. There are reasonable rules that do not return a complete and transitive ordering.

6.2 Unanimity Rule

The unanimity rule is the rule that results from using **Pareto dominance** to create social preferences. Here, we rank one option over another if it Pareto dominates: everyone agrees it is better. Since preferences are strict here, that means everyone likes it strictly more.

Definition 6.2: Unanimity Rule. $x \succ^* y$ if $x \succ_i y$ for everyone. That is, one outcome is ordered above another if *everyone* thinks it is better.

Example 6.5: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

$$a \succ^* b$$

This is an incomplete social preference. It says nothing about the preferences between a, c and b, c.

Example 6.6: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$
Unanimity gives us nothing here!

As we can see, and probably as you expected, the unanimity rule is almost never complete! However, it is transitive. If everyone likes x over y and everyone likes y over z they will certainly like x over z (assuming they all have transitive preferences). Now let's look at a rule that is complete, but (perhaps surprisingly) may be intransitive.

6.3 Majority Rule

Aka. Pairwise Voting.

The pairwise majority social welfare function is a rule where the preference of each pair of alternatives is determined by the majority of voters. In the late 18th century, this was considered the sort of definition of social preferences. If most of society likes x over y then we can say that $x \succsim^* y$ truly represents "society's" preferences between x and y. There's a problem though...

Definition 6.3: Majority Rule. $x \succ^* y$ if more than half of the people prefer x to y. In other words, x is better than y in the social preferences if it wins a pairwise vote between those outcomes.

Example 6.7: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

$$a \succ^* c \succ^* b$$

Example 6.8: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

$$a \succ^* c \succ^* b$$

In these two cases, we get a complete and transitive social preference. Since some outcome will always win each pairwise vote, the rule will always produce a complete social preference relation, but now we will see that it will not always be transitive.

6.3.1 Condorcet Paradox

The Condorcet paradox, named after the French mathematician and philosopher Marquis de Condorcet, who discovered in the late 18th century, that social preferences can be cyclic, even if the individual preferences are not.

Example 6.9: Condorcet Cycle.
1: $a \succ b \succ c$
2: b$\succ$ $c \succ a$
3: $c \succ a \succ b$

$$a \succ^* b, b \succ^* c, c \succ^* a$$

Intransitive Social Preference Relation

6.4 Copeland's Method

Copeland's method, is named after Arthur Copeland who popularized it in the 1950s, though it appears to date back to work by Ramon Llull in the early 1300s! It handles Condorcet's paradox by assigning 1 point for each pairwise win and 0.5 points for each pairwise tie (though in all the examples below there will be an odd number of people and so there will be no ties). The candidate with the highest total points is the winner.

This is also how many sports tournaments are conducted. In sports we can easily get intransitivity. Team a beats b, b beats c, but c beats a. You might see how that would be a problem with picking a winner. For instance, a similar method is used to determine which teams advance from the "group" round to the "knockout" round in the World Cup.

Definition 6.4: Copeland's Method. Conduct a pairwise vote for every pair. If an outcome wins a vote, add one to its score. (If there is a tie add $\frac{2}{2}$. The social preferences are ranked by score. So if x gets a higher score than y it is ranked higher.

Example 6.10: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$
a beats b. a beats c. c beats b.
a wins 2 votes. c wins 1 vote. b wins 0 votes.

$$a \succ^* c \succ^* b$$

$$a :, b :, c :$$

Example 6.11: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$
a wins 2 votes. c wins 1 vote. b wins 0 votes.

$$a \succ^* c \succ^* b$$

Example 6.12: Condorcet Cycle.
1: $a \succ b \succ c$
2: b$\succ c \succ a$
3: $c \succ a \succ b$
All outcomes win one of their pairwise votes.

$$a \sim^* b \sim^* c$$

6.5 Borda

Plurality vote focuses on maximizing the number of people who get their favorite outcome and veto attempts to minimize the number of people who get their least favorite outcome. What if we want to create a rule that balances both of these goals? We can think of the Borda count as a rule that attempts to balance both goals.

Definition 6.5: Borda Count. Each rank is assigned a certain number of points, with higher ranks receiving more points. The option outcome with the highest total points wins. If there are 3 outcomes, we might assign 3 for a first-rank, 2 for a second, and 1 for a third.

With Borda count, it is sort of "traditional" to assign a score of 1 for the last rank and work up from there. For example, using scores of $3, 2, 1$ for three outcomes. However, the outcome will be the same regardless of what numbers we use as long as they are all one-apart. For example, we could use $1, 0, -1$.

Example 6.13: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$
$a : 3 + 3 + 2 = 8$
$b : 2 + 1 + 1 = 4$
$c : 1 + 2 + 3 = 6$

$$a \succ^* c \succ^* b$$

Example 6.14: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$
$a : 11, b : 9, c : 10$

$$a \succ^* c \succ^* b$$

6.6 Exercises

The following exercises make use of this example.

Example 6.15: Example 3.. There are five people and three outcomes. Their preferences are:
1: $a \succ b \succ c$
2: $a \succ b \succ c$
3: $b \succ c \succ a$
4: $b \succ c \succ a$
5: $c \succ a \succ b$

Exercise 6.1: Informally, what social preferences would you assign for this example and why?

Exercise 6.2: What social preferences result from applying *Majority Rule* to this example? Is this an example of the Condorcet Paradox? How do you know?

Exercise 6.3: What social preferences result from applying *Copeland's Method* to this example?

Exercise 6.4: What social preferences result from applying *Borda Count* to this example where a top-ranked outcome gets a score of 1, a second-ranked outcome gets a score of 0 and a last-ranked outcome gets a score of -1.

Exercise 6.5: Come up with your own preference aggregation rule and apply it to examples 1 and 2 from this section and example 3 above.

7 Preference Aggregation Properties

7.1 Basic Properties

There are several properties that we might want in a rule. Let's begin with two properties that ensure that the resulting social preference relation is indeed a rational preference relation:

Definition 7.1: Complete. A preference aggregation rule is **complete** if $\succsim_*$ is complete for all profiles of individual preferences: $(\succsim_1, \ldots \succsim_n)$.

Definition 7.2: Transitive. A preference aggregation rule is **transitive** if $\succsim_*$ is transitive for all profiles of individual preferences: $(\succsim_1, \ldots \succsim_n)$.

I hope I also argued that Pareto efficiency is a very desirable property. Let's add that to the mix. Here, we will write a property that says that the rule needs to respect Pareto dominance. This definition will be given in the context of models where *everyone has strict preferences*.

Definition 7.3: Pareto Efficient. A preference aggregation rule is **Pareto Efficient** if for every x and y such that for every person $i \in P$, $x \succ_i y$ then $x \succ_* y$. That is, if everyone likes x strictly better than y, then the social preference also strictly prefers x to y.

7.2 Classifying Preference Aggregation Rules

7.2.1 Dictatorship

Since a **Dictatorship** just uses an individual's preferences as the social preference, it is always complete and transitive. It is also Pareto efficient. If everyone likes x over y, so will the dictator, and so $x \succ^* y$. Thus, it is also Pareto efficient.

7.2.2 Unanimity

We have seen that **Unanimity Rule** is transitive, Pareto efficient (by definition), but it is **not complete**.

7.2.3 Majority Rule

We have seen that **Majority Rule** is **complete**. It is also **Pareto efficient**. If everyone prefers x to y then a more than a majority will vote for x, thus $x \succ^* y$. However, due to the Condorcet paradox it is **not transitive**.

7.2.4 Copeland's Method

Copeland's method is **non-dictatorial** since it uses the preferences of everyone to determine the social preferences.

Any rule that uses a score is complete and transitive. For these rules, if the score of an outcome is at least as high then $x \succsim^* y$. Since every pair of outcomes gets a score we can compare every outcome. Similarly, if the score of x is higher than y and the score of y is higher than z then the score of x is higher than z. Thus these Since Copeland's method assigns scores based on how many pair-wise votes an outcome wins, is **complete** and **transitive**.

It is also **Pareto efficient**. To see this, suppose everyone prefers a to b. We need $a \succ^* b$ which is true if a beats strictly more other outcomes in a pair-wise vote than b. But since preferences are transitive, anyone who would vote for b over some other outcome, would also vote for a over that outcome. Thus, a beats everything b does, plus it also beats b. Thus the score of a is larger than b.

7.2.5 Borda

Borda count is **non-dictatorial** since it uses the preferences of everyone to determine the social preferences.

Since the Borda count also assigns scores, it is **complete** and **transitive**.

Finally, it is **Pareto efficient** since if everyone agrees that $a \succ_i b$ then a gets a higher score in everyones preference than b. Thus the sum of the scores for a must be strictly higher than b and so $a \succ^* b$.

7.2.6 A Chart of Preference Aggregation Rules

Rule	Complete	Transitive	Pareto
Dictatorship	✓	✓	✓
Unanimity Rule	×	✓	✓
Majority Rule	✓	×	✓
Copeland's Method	✓	✓	✓
Borda Count	✓	✓	✓

Table 1: Comparison of Preference Aggregation Rules

7.3 Independence of Irrelevant Alternatives

There's something a little weird about Copeland's Method and Borda count...

Let's stat with some preferences:

- Person 1: $a \succ b \succ c$
- Person 2: $b \succ a \succ c$
- Person 3: $c \succ a \succ b$

In both Borda and Copeland's method, the social preferences are $a \succ^* b \succ^* c$. Let's focus on the fact that $a \succ^* b$ here. Let's change person 2's preference over a and c to be $c \succ a$ instead of $a \succ c$. We get

- Person 1: $a \succ b \succ c$
- Person 2: $b \succ c \succ a$

- Person 3: $c \succ a \succ b$

In both Borda and Copeland's method, the social preferences are now $a \sim^* b \sim^* c$ and $a \sim^* b$. But we did not change anything about anyone's preferences over a and b, and yet the social preference changed. When this is possible for a preference aggregation rule, we say it fails **Independence of Irrelevant Alternatives**.

> **Definition 7.4: Independence of Irrelevant Alternatives.** A preference aggregation rule obeys **Independence of Irrelevant Alternatives** [IIA] if for any two sets of preferences where the preference for x and y is the same between the two sets, they should have the same social preference between x and y.

7.4 Why IIA Matters

7.4.1 Example: Borda Count

Suppose there are 100 in a society and there are three types of preferences:

- Type 1 (25 People): $a \succ b \succ c$
- Type 2 (40 People): $b \succ c \succ a$
- Type 3 (35 People): $c \succ a \succ b$

In Borda Count, the social preferences are $c \succ^* b \succ^* a$ since the scores are:

- a: $25 * 3 + 40 * 1 + 35 * 2 = 185$
- b: $25 * 2 + 40 * 3 + 35 * 1 = 205$
- c: $25 * 1 + 40 * 2 + 35 * 3 = 210$

But if we remove a. Preferences are:

- Type 1 (25 People): $b \succ c$
- Type 2 (40 People): $b \succ c$
- Type 3 (35 People): $c \succ b$

In Borda Count, the social preferences are $b \succ^* c$ since the scores are:

- b: $25 * 2 + 40 * 2 + 35 * 1 = 165$
- c: $25 * 1 + 40 * 1 + 35 * 2 = 135$

7.4.2 Example: Copeland's Method

In Copeland's method, we can get similar oddities.

- Type 1 (45 People): $a \succ b \succ c$
- Type 2 (15 People): $b \succ c \succ a$
- Type 3 (40 People): $c \succ b \succ a$

In Copeland's Method, the social preferences are $c \succ^* b \succ^* a$ since b beats a, b beats c, c beats a.

Let's remove a.

- Type 1 (45 People): $b \succ c$
- Type 2 (15 People): $b \succ c$
- Type 3 (40 People): $c \succ b$

In Copeland's Method, the social preferences are $b \succ^* c$. Again, we get a reversal!

7.5 Arrow's Impossibility

Rule	Complete	Transitive	Pareto	IIA
Dictatorship	✓	✓	✓	✓
Unanimity Rule	×	✓	✓	✓
Majority Rule	✓	×	✓	✓
Copeland's Method	✓	✓	✓	×
Borda Count	✓	✓	✓	×

Table 2: Comparison of Preference Aggregation Rules

Our goal in this whole process was to look for rule that built on Pareto efficiency but also filled in the gaps to create a complete and transitive social preference relation. We found two options in Borda and Copeland's method. Unfortunately, we ran into a new problem. IIA.

So far, the only thing we have seen that meets all of our assumptions is a dictatorship! As it turns out, that is the only preference aggregation rule that is complete and transitive, Pareto efficient, and does is not susceptible to violations of IIA.

> **Info 7.1: Arrow's Impossibility Theorem.** If there are at least three outcomes, the only preference aggregation rule that is **complete**, **transitive**, **Pareto efficient**, and does not violate **IIA** is a **dictatorship**!

7.6 Exercises

In **Plurality Vote** the preference aggregation rule, the **score** of an outcome is the number of people who rank that outcome highest. Social preferences are determined by score. See subsection A.1 for more info.

In **Veto**, the score of an outcome is the negative of the number of people who rank it last. Social preferences are determined by the score as in the other scoring methods above with a higher score being ranked higher. See subsection A.2 for more info.

Rule	Complete	Transitive	Pareto	IIA
Dictatorship	✓	✓	✓	✓
Unanimity Rule	×	✓	✓	✓
Majority Rule	✓	×	✓	✓
Copeland's Method	✓	✓	✓	×
Borda Count	✓	✓	✓	×
Plurality Vote				
Veto				

Table 3: Comparison of Preference Aggregation Rules

Exercise 7.1: Ask ChatGPT or another *AI* of your choice to provide an example of a preference aggregation rule that meets completeness, transitivity, Pareto efficiency, and IIA. Can you convince it to lie to you? If so, provide the prompt you used.

Exercise 7.2: Fill out the table above for **Plurality Vote**. For each entry, explain why you place a ✓or an ×.

Exercise 7.3: Fill out the table above for **Veto**. For each entry, explain why you place a ✓or an ×.

Exercise 7.4: Come up with a preference aggregation rule that is complete, transitive, IIA, but **not Pareto efficient**.

8 Making Choices

Preference aggregation rules produce a preference relation. In most real-world situations, we do not need a preference relation, we just need to make a choice. There is a concept related to preference aggregation rules that has the focus not of making a preference relation, but rather just making a choice. We call these social chocie functions.

Definition 8.1: Social Choice Function. A **social choice function** (also known as a social welfare function) is a way to turn individual preferences into a choice. Formally, it is a *mapping* from the set of possible individual preferences over the outcomes into an outcome or subset of the outcomes- the choice/choices. That is, from the set of possibilities, it picks a winner or winners (if there are ties).

8.1 Social Choice from Preference Aggregation

We learned in section 1 that when a preference relation is complete and transitive, we can use it to make a choice from any subset of the outcomes through the notion of best outcomes. Recall from Definition 1.10 that an outcome x is **best** from some set B according $\succsim$ if for every other y in the set B, $x \succsim y$. That is, x is best from a set if it is preferred to all other outcomes in the set.

In this sense, if a preference aggregation rule results in a complete and transitive social preference relation $\succsim^*$, we can use that relation to make a choice from any subset of the outcomes. In this sense, every preference aggregation rule is a social choice function.

Info 8.1: Social Choice Functions from Preference Aggregation Rule. Any preference aggregation rule can be converted into a social choice function by taking the **best** outcomes from the resulting social preference relation.

However, some social choice functions are more suited towards picking a winner than creating an entire preference ordering, and the properties we look for in a social choice function are not the same as those of a preference aggregation rule.

8.2 Some Social Choice Functions

8.2.1 Dictatorship- Social Choice

Definition 8.2: Dictatorship. Pick a person $i \in P$. The social choice is that person's favorite outcome.

Example 8.1: Example 1. Let's assume person 1 is the dictator. 1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$
The social choice is a.

Example 8.2: Example 2. Assume person 1 is the dictator. 1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$
The social choice is a.

8.2.2 Unanimity Rule- Social Choice

The unanimity rule is a social choice function that selects an outcome if it is unanimously preferred over another by all individuals.

Definition 8.3: Unanimity Rule- Social Choice. Choose x if for every person and for all other outcomes y $x \succ_i y$ for everyone. That is, an outcome is chosen if *everyone* thinks it is better than every other outcome.

Example 8.3: Example 1. 1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$
The social choice is a.

Example 8.4: Example 2. 1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$
Unanimity gives no choice here.

8.2.3 Plurality Vote- Social Choice

Plurality vote focuses on the goal of giving as many people as possible their top-ranked outcome.

Definition 8.4: Plurality Vote- Social Choice. The social choice is the outcome which the most number of people rank first.

Example 8.5: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

a is the social choice.

Example 8.6: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

a is the social chocie.

8.2.4 Borda Count- Social Choice

The Borda count is a social choice function that balances the preferences of individuals by assigning points to ranks. Higher ranks receive more points, and the outcome with the highest total points is chosen.

Definition 8.5: Borda Count- Social Choice. Each rank is assigned a certain number of points, with higher ranks receiving more points. The outcome with the highest total points is the social choice. If there are 3 outcomes, for example, we might assign 3 points for a first-rank, 2 points for a second, and 1 point for a third.

Example 8.7: Example 1. 1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

Scores:
$a : 3 + 3 + 2 = 8$
$b : 2 + 1 + 1 = 4$
$c : 1 + 2 + 3 = 6$
a is the social choice.

Example 8.8: Example 2. 1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

Scores:
$a : 3 + 3 + 1 + 2 = 9$
$b : 2 + 2 + 3 + 1 = 8$
$c : 1 + 1 + 2 + 3 = 7$

a is the social choice.

The following exercises make use of this example.

Example 8.9: Example 3.. There are five people and three outcomes. Their preferences are:
1: $a \succ b \succ c$
2: $a \succ b \succ c$
3: $b \succ c \succ a$
4: $b \succ c \succ a$
5: $c \succ a \succ b$

Exercise 8.1: What choice/choices result from applying *Plurality Vote- Social Choice* to this example?

Exercise 8.2: What choice/choices result from applying *Veto- Social Choice* to this example?

Exercise 8.3: What choice/choices result from applying *Borda Count- Social Choice* to this example?

9 Social Choice Function Properties

9.1 Basic Properties

A preference aggregation rule results in a preference relation. For such a relation to be capable of making choices from *any* subset of outcomes, it needs to be complete and transitive. However, a social choice function only needs to pick a winner or winners from the entire set of alternatives.

We can replace *completeness* and *transitivity* with the following property:

> **Definition 9.1: Nonempty.** A social choice function is **nonempty** if the set of choices is nonempty for all profiles of individual preferences: $(\succsim_1, \dots \succsim_n)$.

In preference aggregation rules, if there is an outcome y such that there is another x that everyone strictly prefers to y, then y could never be ranked highest since *Pareto efficiency* of the social welfare function will require $x \succ^* y$. In that sense, y should also never *win*– it should never be the choice. That leads to the following extension of Pareto efficiency to social choice functions. This says that if y is strictly Pareto dominated, it cannot be chosen.

> **Definition 9.2: Pareto Efficient.** A social choice function is **Pareto Efficient** if for every y where there is another outcome x such that every person $i \in P$, $x \succ_i y$ then y cannot be in the set of choies. That is, if everyone likes x strictly better than y, then the social preference also strictly prefers x to y.

Lastly, we can extend independence of irrelevant alternatives to social choice functions:

> **Definition 9.3: Independence of Irrelevant Alternatives.** A social choice function obeys **Independence of Irrelevant Alternatives** [IIA] if for any two sets of preferences where the preference for x and y is the same between the two sets, if x is chosen in in the first set and y is not, then y cannot be chosen in the second set.

9.2 Classifying Social Choice Functions

9.2.1 Unanimity and Majority

Among the social choice functions we looked at above, **Unanimity** and **Majority Vote** are **not nonempty**. For instance, neither pick a winner under the Condorcet paradox preferences below:

1: $a \succ b \succ c$

2: $b \succ c \succ a$

3: $c \succ a \succ b$

There is no outcome that is unanimously better than all others, so unanimity does not pick a choice. Similarly, majority vote results in an intransitive cycle where $a \succ^* b$, $b \succ^* c$, $c \succ^* a$, thus there is no outcome at least as good as all others.

They are, however, **Pareto efficient** and **IIA**.

9.2.2 Plurality Vote

Since some outcome always wins the plurality vote, it is **nonempty**.

Plurality vote is Pareto efficient because if some outcome x is preferred by everyone to y, then y cannot be anyone's favorite. Thus, it cannot win.

Plurality vote is **not IIA**. Consider the following sets of preferences:

Set 1:

1: $a \succ b \succ c$

2: $a \succ b \succ c$

3: $b \succ a \succ c$

Set 2:

1: $a \succ b \succ c$

2: $c \succ a \succ b$

3: $b \succ a \succ c$

a is the choice in the first set, and a and b are the chocie in the second set even though everyon has the same preferences over a and b. This violates IIA.

9.2.3 Borda

Since some outcome always wins the Borda count, it is **nonempty**.

It is Pareto efficient since if x is strictly prefered by everyone over y, x must get a strictly higher score, thus, y cannot get the highest score and be a chocie.

It is **not IIA**. Consdier the following sets of preferences:

Set 1:

- 1: $a \succ b \succ c$
- 2: $b \succ a \succ c$
- 3: $c \succ a \succ b$

Set 2:

- 1: $a \succ b \succ c$
- 2: $b \succ c \succ a$
- 3: $c \succ a \succ b$

a is the choice in set 1, but a, b and c are all choices in set 2. This violates IIA.

9.2.4 A Chart of Social Choice Functions

Rule	Nonempty	Pareto	IIA
Dictatorship	✓	✓	✓
Unanimity Rule	×	✓	✓
Majority Rule	×	✓	✓
Plurality Vote	✓	✓	×
Borda	✓	✓	×

Table 4: Comparison of Social Choice Functions

9.3 Why IIA Matters for Social Choice

This example is identical to the example used to show why IIA matters for preference aggregation. I have duplicated it here for convenience.

Suppose there are 100 in a society and there are three types of preferences:

- Type 1 (25 People): $a \succ b \succ c$
- Type 2 (40 People): $b \succ c \succ a$
- Type 3 (35 People): $c \succ a \succ b$

In Borda Count, the choice is c since the scores are:

- a: $25 * 3 + 40 * 1 + 35 * 2 = 185$
- b: $25 * 2 + 40 * 3 + 35 * 1 = 205$
- c: $25 * 1 + 40 * 2 + 35 * 3 = 210$

But if we remove a. Preferences are:

- Type 1 (25 People): $b \succ c$
- Type 2 (40 People): $b \succ c$
- Type 3 (35 People): $c \succ b$

In Borda Count, the choice is b since the scores are:

- b: $25 * 2 + 40 * 2 + 35 * 1 = 165$
- c: $25 * 1 + 40 * 1 + 35 * 2 = 135$

9.4 Arrow's Impossibility Again

Arrow's impossibility theorem says only a dictatorship can aggregating preferences in a way that is complete, transitive, Pareto efficient and respects IIA. A preference aggregation rule creates a social preference that lets a decision-maker make a choice from *any* subset of the outcomes. That's what a preference relation is good for. It might seem like if we were not worried about being able to make a choice from every subset, but just wanted to make a choice from the whole set of outcomes, it might be easier to find a suitable rule.

Notice in the chart above that even though all a social choice function has to do is pick a winner, we cannot seem to get one that has the three properties we might want. In fact, it is still impossible.

> **Info 9.1: Arrow's Impossibility Theorem for Social Choice.** If there are at least three outcomes, the only social choice function that is **nonempty**, **Pareto efficient**, and does not violate **IIA** is a **dictatorship**!

As far as I can tell, this result was formalized by Denicolò Vincenzo in [1].

9.5 exercises

Exercise 9.1: Describe Arrow's impossibility theorem as it applies to *social choice functions* in three paragraphs to someone who has never taken economics or mathematics. Be sure to describe the properties that cannot simultaneously be achieved in as simple terms as possible.

Exercise 9.2: Consider the **veto** rule. This social choice function chooses the outcome or outcomes that are the least favorite of as few people as possible. Is this social choice function **nonempty**? Is it **Pareto efficient**? Is it **IIA**?

10 Strategic Voting

So far we have we have looked at instances where preferences are *known*. But what happens if preferences are not known? We have to collect them from constituents.

Let's add one more property to the mix. Arrow's impossbility theorem above says that even if we allow ties for the "choice" then there is no social choice function that always makes choice that is IIA and Pareto efficient.

However, in the real world, *some* outcome actually needs to get chosen. A real-world social choice function needs to have a way of breaking ties. Any social choice function that breaks ties and just picks one outcome is called **decisive**:

> **Definition 10.1: Decisive.** A social choice function is **decisive** if there is always a single choice for all profiles of individual preferences: $(\succsim_1, \ldots \succsim_n)$.

Note that decisiveness is **stronger** than **nonempty**. For a rule to be decisive, it has to be nonempty **and** there always has to be just one choice. Most real-world social choice functions have some built-in tie-breaking rule. However, we can also easily convert any non-empty social choice function into a decisive one by adding a simple tie-breaking rule. Below, I will use the rule that breaks ties by choosing the outcome lowest in the alphabet.

For example suppose we have the following preferences:

1: $a \succ b \succ c$

2: $b \succ c \succ a$

3: $c \succ a \succ b$

Plurality Vote is **not decisive** since there is a tie for number of first place votes between a and b and c. However, if we add a tie-breaking rule that the outcome lowest in the alphabet wins any tie then the rule is decisive. The choice will be a.

> **Definition 10.2: Plurality Vote with Alphabetical Tie-Breaker.** The social choice is the outcome which the most number of people rank first. If there is a tie, the outcome lowest in the alphabet wins.

10.1 Manipulation

In the example above, notice that if person 2 changed their vote to c instead of b by pretending to have preferences $c \succ b \succ a$ then c would win instead, and person 3 likes this better!

This is an example of **manipulation**.

> **Definition 10.3: Non-manipulable.** A social choice function is **Non-manipulable** if no individual can achieve a more preferred outcome by misrepresenting their preferences.

Manipulation can undermine the fairness and accuracy of voting outcomes. It can lead to scenarios where the chosen outcome does not reflect the true preferences of the voters.

10.2 Another Impossibility

Arrows impossibility says that we have to be ok with either relaxing Pareto efficiency or IIA. Suppose we were ok with relaxing IIA– I think it is the more natural option to give up. There are many social choice rules that are **decisive** and **Pareto efficient**.

For instance, **Plurality Vote with Alphabetical Tie-Breaker** or **Borda with Alphabetical Tie-Breaker** are both Pareto efficient and decisive. However, they are also both manipulable. We have seen that in the example above for Plurality vote with Alphabetical Tie-Breaker, but the same example works for Borda count. Person 2 would still like to pretend to have preferences $c \succ b \succ a$ giving the win to outcome c instead of outcome a.

As it turns out, as long as a social choice rule is **Pareto efficient** and **decisive** we will always be able to find an example where the rule can be manipulated.

Info 10.1: Gibbard-Satterthwaite. If there are at least three outcomes, the only social choice function that is **decisive**, **Pareto efficient**, and **non-Manipulable** is a **dictatorship**!

10.3 Exercises

Exercise 10.1: Describe the Gibbard-Satterthwaite theorem in simple terms, detailing why any non-dictatorial voting system with at least three choices is manipulable.

Exercise 10.2: Provide an example with 5 people and 3 outcomes where someone can manipulate their preferences and make the chosen outcome better for themselves in **Plurality Vote**

Exercise 10.3: Provide an example with 5 people and 3 outcomes where someone can manipulate their preferences and make the chosen outcome better for themselves in **Borda Count**

Part III

Appendix

A Optimization

A.1 Unconstrained One-Dimensional Optimization

Optimization involves finding the minimum or maximum of a function $f(x)$. Here, we focus on instances where f is one-dimensional. The goal is to determine the value of x that maximizes (or minimizes) $f(x)$. *Unconstrained* means that we will not place any restrictions on what x can be.

Imagine that you are hiking on a mountain trail. If the slope of the trail is positive, then moving forward will bring you to a higher point. If the slope of the trail is negative, then moving *backward* will bring you to a higher point. Thus, **the slope must be zero at the peak**. This is demonstrated in Figure Figure A.1.

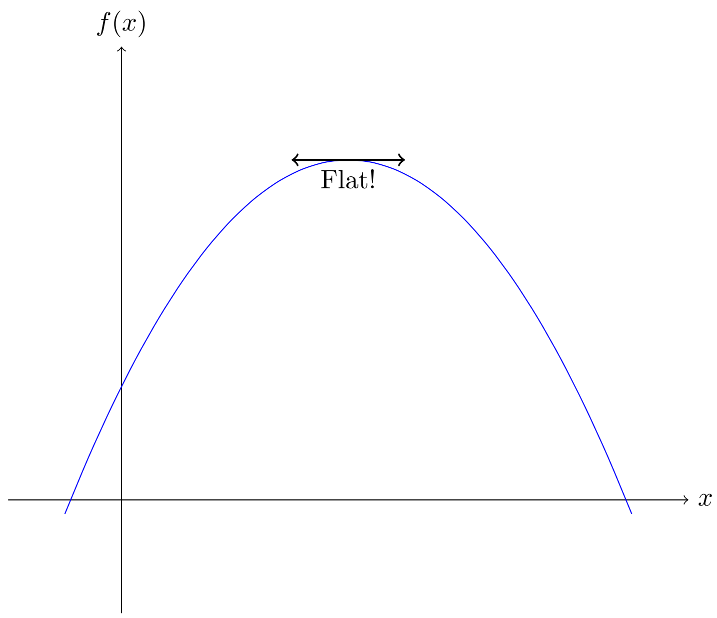


Figure A.1: Slope is Zero at the Peak

One issue with using this fact to find a maximum is that the slope can also be zero at a minimum and also at places that are "local" maxima.

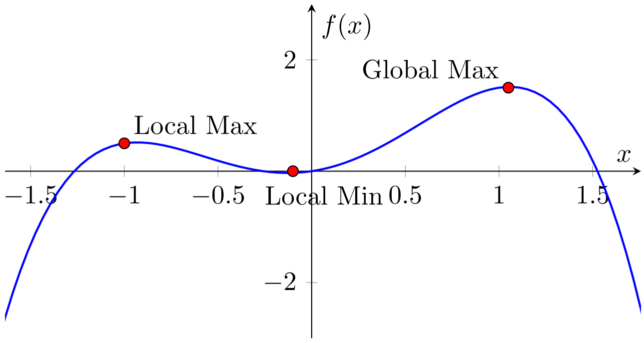


Figure A.2: Not every point of zero slope is a global maximum!

To account for this, we should remember that when we find a point of zero slope, it is only a *candidate* for a global maximum.

> **Info A.1: Unconstrained Optimization.** How to find the unconstrained maximum of a one-dimensional function: For a function $f(x)$:
>
> 1. Find the first derivative $f'(x)$.
> 2. Set the first derivative to zero: $f'(x) = 0$.
> 3. Solve for x. These are your candidates.
> 4. Which, if any is a global max?

A.2 Unconstrained Multi-Dimensional Optimization

The intuition of the slope being zero at the maximum holds even when there are multiple directions in which you can move. Imagine trying to find the peak of a mountain when you are not on a trail. You can move east/west or north/south. In fact, you can also move in combinations of these directions, like the northwest. But at the peak, you better not be able to move east-/west and get to a higher altitude. **The slope has to be zero in the east/west direction**. Similarly, **the slope has to be zero in the north/south direction.** One of the nice things about *smooth* functions is that if the slope is zero in these two cardinal directions, it will be zero even if you try to move northwest, or southeast, or any other direction. Figure A.3 demonstrates this. Notice that at the peak, the slope is zero in both the x direction and the y direction, and also in all other directions.

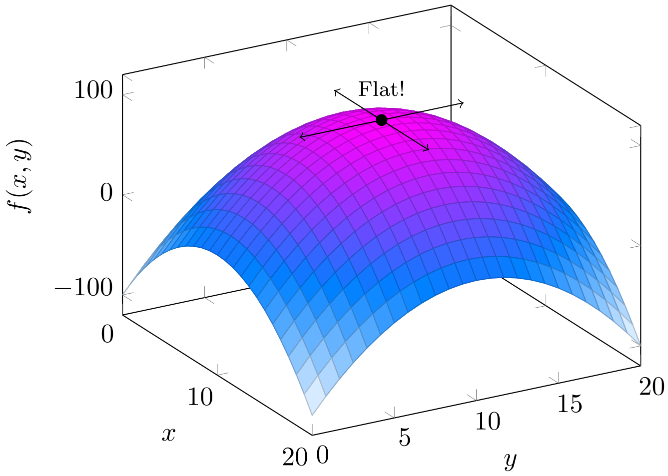
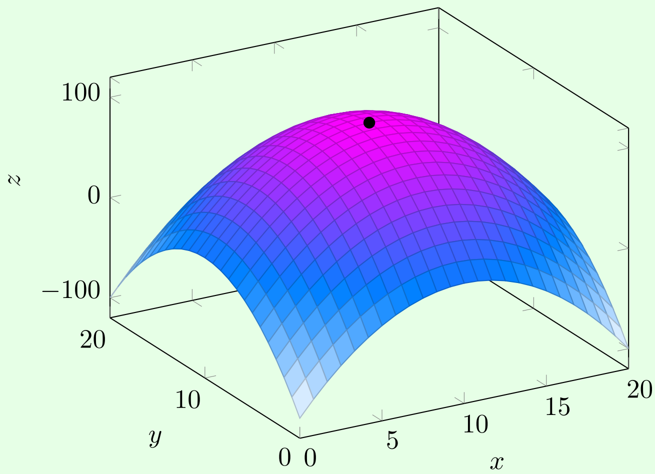


Figure A.3: Slope is Zero in All Directions!

Info A.2: Unconstrained Multi-Dimensional Optimization. To maximize a function $f(\mathbf{x})$ where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$:

1. Find all partial derivatives $\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$.
2. Set all partial derivatives to zero: $\frac{\partial f}{\partial x_i} = 0$.
3. Solve the resulting system of equations for $(x_1, ..., x_n)$. These solutions are your candidates.
4. Determine which, if any, of these solutions is a global maximum.

Example A.1: Unconstrained Maximum. Maximize $100 - (x-10)^2 - (y-10)^2$. Let's look at this function first. The global maximum (black dot) occurs where $x = 10$ and $y = 10$.



Now, we confirm that this is the maximum formally using the procedure in A.2. The partial derivatives are $\frac{f(x,y)}{x} = -2(x-10)$ and $\frac{f(x,y)}{y} = -2(y-10)$.
Setting these to zero, we get the equations:

$$\frac{f(x,y)}{x} = -2(x-10) = 0$$

$$\frac{f(x,y)}{y} = -2(y-10) = 0$$

Solving these gives us the (x, y) where the function has zero slope. The only solution is $x = 10$, $y = 10$.
We can see by inspecting the function that this must be the global maximum.

A.3 Constrained Multi-Dimensional Optimization

Suppose that we want to maximize a function $f(x, y)$ (**the objective**) but where the set of x and y we can choose from is constrained in some way (**the constraint**).

Let's have a look at how adding a constrained complicates Example A.1.

Example A.2: Constrained Maximum. Maximize $100-(x-10)^2-(y-10)^2$ subject to $x+y \leq 10$.
Let's look at this function first. As we found in Example A.1, the global maximum (black dot) occurs where $x = 10$ and $y = 10$, but that violates the constraint since $10+10 > 10$. We are not allowed to go past the red line. The maximum within that area occurs at $x = 5$ and $y = 5$ (green dot).

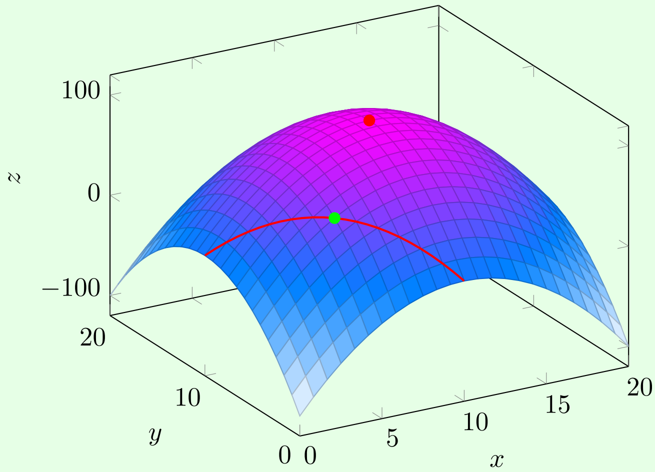
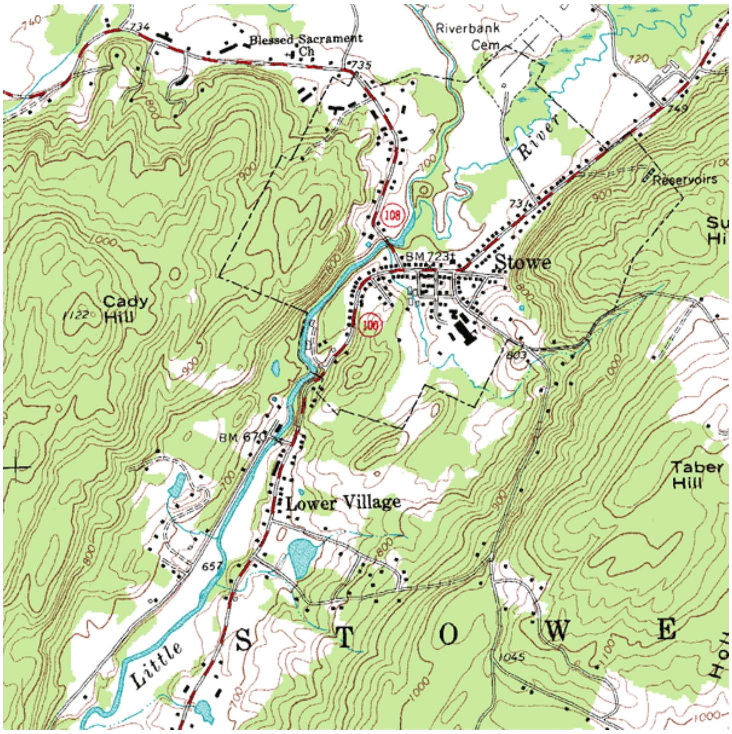


Figure A.4: 3d Plot of $100-(x-10)^2-(y-10)^2$

How should we formalize the process of finding the constrained optimum? Let's work through a few concepts and return to this example later in the chapter.

A.4 Contours

It can be very useful to think of three-dimensional plots in terms of their contours. Figure A.5 shows a real-world example of how contours are used on a topographic map, which is a 2d map that includes information about elevation through contour lines. Look at the line labeled "1000" near Cady Hill. This is a line connecting places that all have an elevation of 1000 feet.



Taken from the public domain USGS Digital Raster Graphic file o44072d6.tif for the Stowe, VT quadrangle.

Figure A.5: A topographic map of Stowe Mountain.

Let's add some contours to our function at an "elevation" of 25, 50, 75, and 99 (right near the peak). In the context of mathematics, this is known as a "contour" plot.

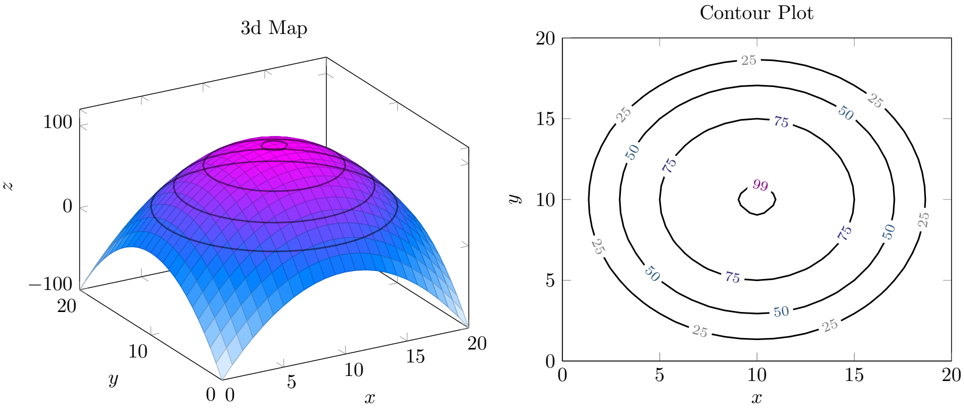


Figure A.6: A function and its contours.

A.5 Monotonicity

Imagine standing at the point $(0, 0)$ on the function plotted in Figure A.6. If you walk in the northwest direction (increasing x and y) the function increases. That is, you are increasing in elevation. In fact, this is true whenever $0 \leq x \leq 10$ and $0 \leq y \leq 10$. Figure A.7 shows a plot of the function in Example A.2 limited to this region. Notice how the function always slopes up when moving in the northwest direction regardless of where you are.

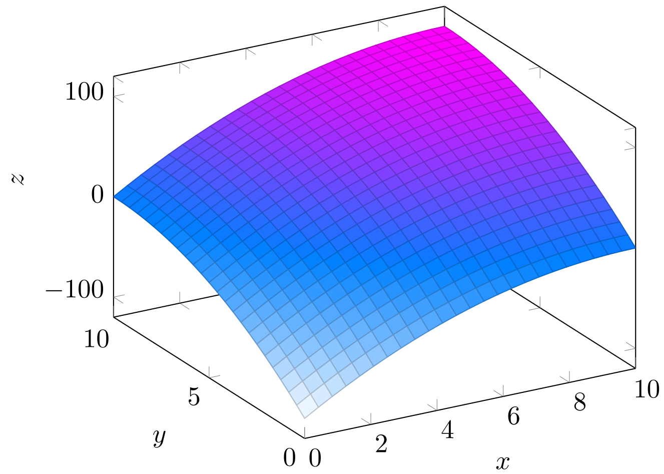


Figure A.7: 3d Plot of $100 - (x - 10)^2 - (y - 10)^2$ **where** $x \leq 10$ **and** $y \leq 10$**.**

When a function increases when you increase all of its variables, we say that it is monotonic. As we will see, this property comes in handy.

Definition A.1: Monotone. $f(x, y)$ is said to be **monotone** when:

1. $x' \geq x$ and $y' \geq y$ implies $f(x', y') \geq f(x, y)$.
2. $x' > x$ and $y' > y$ implies $f(x', y') > f(x, y)$.

A.6 Three Possibilities for an Optimal Point

Let's continue looking at Example A.2. Let's look at the contour plot where $x \leq 10$ and $y \leq 10$ and add the line $x + y = 10$. This is shown in Figure A.8.

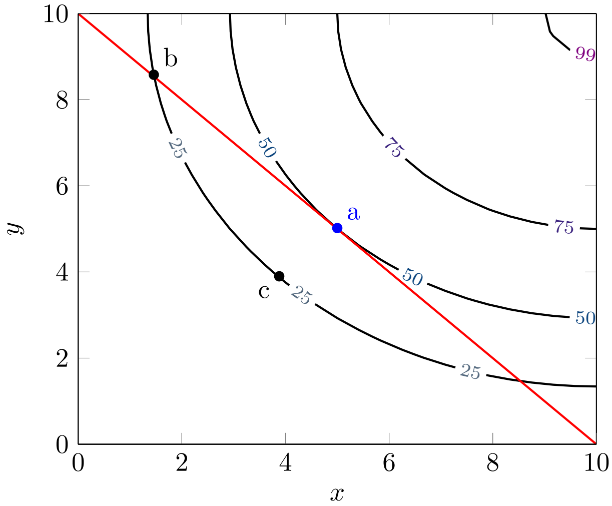


Figure A.8: Contour plot with constraint.

Here, the constraint is the area southwest of the red line. The red line is the "boundary" of the constraint, and the area to the south-west of that line is called the interior of the constraint. For example, the points b and c are on the boundary of the constraint and c is on the interior.

First, notice that point c could never be optimal. Why? If we are on the interior, we could always move up and to the right a little (increasing both x and y) and still meet the constraint. *Since the function is monotonic*, the result **must be better!** Here, for example, we could move from c to a. Thus, we can see that if a function is monotonic, the optimal point cannot be on the interior of the constraint. But it has to meet the constraint. Thus, **when a function is monotonic, the optimal point must be on the boundary of the constraint.**

The point b is on the boundary. Can it be optimal? No, it is on the same contour as c. Because c cannot be optimal, neither can b. They have the same value. This shows us that a point like b, which is on a contour that passes through the interior of the constraint, can never be optimal.

What we have seen so far is that whatever point is optimal must be on the boundary of the constraint and not on a contour that passes into the interior of the constraint. The only way for this to happen is if the contour at the optimal point *just touches* the constraint. See point c for instance. When the contours are smooth, the only way for this to happen is if the contour and the constraint have the same slope.

There are only three possibilities for an optimal point. These are enumerated below.

Info A.3: Three Possibilities for a Constrained Optimum. When the objective is monotonic, the optimum must meet one of the following three conditions.

1. (**Tangent**) It is at a point where the contour of the objective at that point had the same slope constraint.
2. (**Touching but not Smooth**) The point is a "non-smooth" point on the contour of the objective, but the that point just touches the constraint.
3. (**Boundary**) The point is at one of the boundaries of the constraint.

A.7 Slopes of Contours

Many of the optimization problems we will encounter in this course will be "smooth". In that case, the first possibility "tangent" from box A.3 is relevant. To find such a tangency point, we need to know how to find the slopes of functions.

In most cases, we will be dealing with functions of just two variables. For instance, lines like $ax_1 + bx_2 = 10$ or non-linear functions like $x_1^2 x_2^2 = 10$. How do we find the slope of functions like these at particular points?

For the linear case $ax_1 + bx_2 = 10$, we can put it in the conical form of a line $x_2 = -\frac{a}{b}x_1 + \frac{10}{b}$ then read the slope right off. Here it is $\frac{a}{b}$ and since this is a line, that is the slope at every point on the function. But what about a nonlinear function like $x_1^2 x_2^2 = 10$? That slope depends on the point you are considering. Here is a plot.

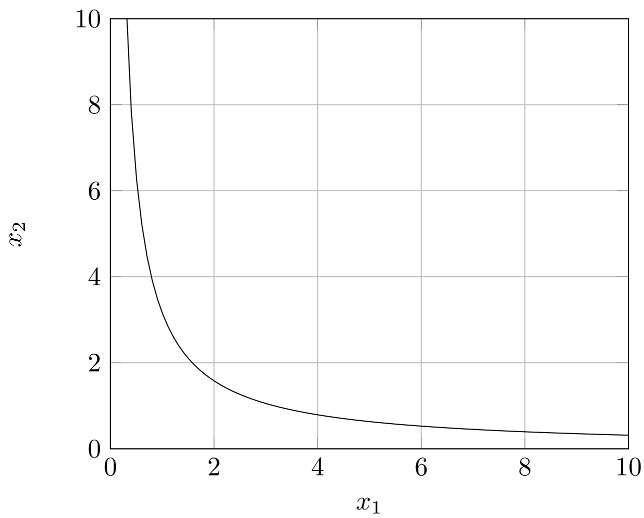


Figure A.9: Plot of $x_1^2 x_2^2 = 10$

We can use something called the implicit function theorem to find the slope, but let's work through it intuitively before I give you the general result.

The slope of a function is really measuring something like this: When you move horizontally a little, how much do you have to move vertically to get back to the function? For a line with a slope of -1 if you move to the right by one unit, you have to move down by one unit to get back on the function. Look at the plot above. When x_1 is small and x_2 is large, if you move to the right a little, you have to move down a lot to get back to the function. If x_1 is large but x_2 is small, if you move to the right a little, you don't have to move down much at all to get back to the function.

A function like $x_1^2 x_2^2 = 10$ defines a set of points (x_1, x_2) that all meet some condition. Here, the value of the function $f(x_1, x_2) = x_1^2 x_2^2$ in the set of points is equal to 10.

The partial derivative of f with respect to x_1 tells us how much f changes when we increase x_1 a little. This is denoted by $\frac{\partial f}{\partial x_1}$. The partial derivative with respect to x_2, denoted $\frac{\partial f}{\partial x_2}$ tells us how much f changes when we increase x_2 by a little.

Suppose $\frac{\partial f}{\partial x_1} = 1$ and $\frac{\partial f}{\partial x_2} = 1$. Roughly, if we increase x_1 by a little, $f()$ increases by 1 unit. What do we have to do to get back to the function? We cannot increase x_2. That will only make $f()$ even bigger. We have to **decrease** x_2 by one unit. The slope is -1.

Suppose $\frac{\partial f}{\partial x_1} = 1$ and $\frac{\partial f}{\partial x_2} = 2$. Roughly, if we increase x_1 by a little, $f()$ increases by 1 unit. What do we have to do to get back to the function? If we decrease x_2 by one unit, $f()$ will decrease by 2. That's too much! Instead, we decrease it by $\frac{1}{2}$. Then $f()$ will decrease by 1. The slope is $-\frac{1}{2}$.

Finally, suppose $\frac{\partial f}{\partial x_1} = 1$ and $\frac{\partial f}{\partial x_2} = \frac{1}{2}$. Roughly, if we increase x_1 by a little, $f()$ increases by 1 unit. What do we have to do to get back to the function? If we decrease x_2 by one unit, $f()$ will be decreased by $\frac{1}{2}$. That is not enough! Instead, we decrease it by 2. Then $f()$ will decrease by 1. The slope is -2.

Note that in each case, we find that the slope is the **negative** of the **ratio** of the **partial derivatives**. That is a general result.

Info A.4: Slope of an Implicit Function. The slope of a function $f(x_1, x_2) = y$ at the point (x_1, x_2) is $-\frac{\frac{\partial f(x_1,x_2)}{\partial x_1}}{\frac{\partial f(x_1,x_2)}{\partial x_2}}$.

A.8 Solution - Two Dimensional Constrained Optimization

Now that we know how to find the slope of implicit functions, we have an easy way to find places where the slope of the contour is the same as the slope of the constraint.

Definition A.2: First-Order Condition. For maximizing $f()$ subject to $g() \leq c$, the **first-order condition** is:

$$-\frac{\frac{\partial f(x_1,x_2)}{\partial x_1}}{\frac{\partial f(x_1,x_2)}{\partial x_2}} = -\frac{\frac{\partial g(x_1,x_2)}{\partial x_1}}{\frac{\partial g(x_1,x_2)}{\partial x_2}}$$

.

Notice that the first-order condition gives us just one equation (the number of variables). But the optimal point has two unknowns. This is not enough for the optimal point. Fortunately, we already know something else about the optimal point. It must occur *on the constraint.*

Info A.5: Solving a Constrained Maximum. To solve a constrained maximum problem:

1. Find $-\frac{\frac{\partial f(x_1,x_2)}{\partial x_1}}{\frac{\partial f(x_1,x_2)}{\partial x_2}}$ and $-\frac{\frac{\partial g(x_1,x_2)}{\partial x_1}}{\frac{\partial g(x_1,x_2)}{\partial x_2}}$
2. Simplify $-\frac{\frac{\partial f(x_1,x_2)}{\partial x_1}}{\frac{\partial f(x_1,x_2)}{\partial x_2}} = -\frac{\frac{\partial g(x_1,x_2)}{\partial x_1}}{\frac{\partial g(x_1,x_2)}{\partial x_2}}$.
3. Plug the result into the constraint to get the solution.

A.9 Gradients

All of the above assumed we are maximizing a two-dimensional function $f(x, y)$. There, we can find a place where the slope of a contour is equal to the slope of the constraint using the formula in Info Box A.4. But for functions with more variables like $f(x, y, z)$ it is more convenient to instead calculate the **gradient** of the objective and constraint. What is that?

Definition A.3: Gradient. The gradient of a function is a vector that points in the direction of the *fastest rate of increase of the function.* It is denoted by ∇f and is the vector of partial derivatives of the function:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)$$

where f is a function of n variables $x_1, x_2, \dots, x_n$.

Info A.6: Gradients are Perpendicular to Contour. The gradient of the function is always perpendicular to the contour lines. This is because the gradient points in the direction of the steepest ascent, while the contour lines represent points of zero ascent.

In summary, the gradient ∇f is perpendicular to the contour lines of the function f, and its magnitude indicates the rate of change of the function in the direction of the gradient.

A.10 First Order Condition and Solution

Now that we know that the gradient of a function at some point is always perpendicular to the contour of a function at some point, we have an easy way to find places where the slope of the contour is the same as the slope of the constraint.

Definition A.4: First-Order Condition. For maximizing $f()$ subject to $g() \leq c$, the **first order condition** is:

$$\nabla f(x) = \lambda \nabla g(x) \tag{1}$$

where $\nabla f(x)$ is the gradient of the objective function, $\nabla g(x)$ is the gradient of the constraint function, and λ is just some number.
This is equivalent to the following n equations:

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= \lambda \frac{\partial g}{\partial x_1} \\ \frac{\partial f}{\partial x_2} &= \lambda \frac{\partial g}{\partial x_2} \\ &... \\ \frac{\partial f}{\partial x_n} &= \lambda \frac{\partial g}{\partial x_n} \end{aligned} \tag{2}$$

The λ comes from the fact that the two gradients need not have the same *magnitude*, only the same direction! The λ allows their magnitude to differ.

Notice that the first-order condition gives us n equations (the number of variables). But the optimal point has n unknowns. This is not quite enough for the optimal point. Fortunately, we already know something else about the optimal point. It must occur *on the constraint*.

Info A.7: Solving a Constrained Maximum. To solve a constrained maximum problem with n variables.

1. For each variable $x_1, x_2, ..., x_n$. Find $\frac{\partial f}{\partial x_i}$ and $\frac{\partial g}{\partial x_i}$.
2. Solve the resulting first-order condition, together with the constraint.

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= \lambda \frac{\partial g}{\partial x_1} \\ \frac{\partial f}{\partial x_2} &= \lambda \frac{\partial g}{\partial x_2} \\ &... \\ \frac{\partial f}{\partial x_n} &= \lambda \frac{\partial g}{\partial x_n} \\ g(x_1, x_2, ..., x_n) &= c \end{aligned} \tag{3}$$

A.11 Lagrange

Notice that first-order condition for *constrained optimization* involves the n equations of the form:

$$\frac{\partial f}{\partial x_i} = \lambda \frac{\partial g}{\partial x_i} \tag{4}$$

... together with the constraint $g(x_1, x_2, ..., x_n) = c$.

As it turns out, there is always an *unconstrained optimization* problem that has the exact same first-order conditions. That is, the optimization of the following Lagrange function for variables $x_1, x_2, ..., x_n, \lambda$.

$L = f(x_1, x_2, ..., x_n) - \lambda(g(x_1, x_2, ..., x_n) - c)$. Recall above that the first-order condition for an unconstrained problem is just that all the first derivatives have to be zero.

Taking those first , we get something familiar:

$$\begin{aligned} \frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} &= 0 \\ \frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} &= 0 \\ &... \\ \frac{\partial f}{\partial x_n} - \lambda \frac{\partial g}{\partial x_n} &= 0 \\ g(x_1, x_2, ..., x_n) - c &= 0 \end{aligned} \tag{5}$$

These can be rearranged exactly to the first-order conditions of the constrained optimization problem.

A.12 Examples

Let's work through Example A.2 using the Lagrange method.

Example A.3: Solution to Example A.2.
We have our objective: $100 - (x - 10)^2 - (y - 10)^2$ and our constraint $x + y \leq 10$.
We begin by turning this constrained optimization into the unconstrained optimization problem of the Lagrange function:
$L(x, y) = 100 - (x - 10)^2 - (y - 10)^2 - \lambda(x + y - 10)$
Now we find the first-order conditions of this unconstrained problem with respect to the three variables x, y, λ. These are:

$$-2(x - 10) - \lambda = 0$$

$$-2(y - 10) - \lambda = 0$$

$$x + y - c = 0$$

We can solve these three equations by first eliminating λ from the first two equations:

$$-2(y - 10) = -2(x - 10)$$

$$x = y$$

Now we plug this into the third (constraint) equation $x + y = 10$ to get:

$$x = 5$$

$$y = 5$$

A.13 Exercises

Assume for each of the following problems that $x \geq 0$ and $y \geq 0$.

Exercise A.1: Maximize the function $f(x) = -x^2 + 4x + 4$.

Exercise A.2: Maximize the function $f(x) = ln(x) - \frac{1}{4}x + 4$.

Exercise A.3: Maximize the function $f(x, y) = -x^2 - y^2 + 2x + 2y$.

Exercise A.4: Maximize the function $f(x, y) = x + y$ subject to the constraint $x + 2y \leq 60$.

Exercise A.5: Maximize $f(x, y) = xy$ subject to the constraint $x + 2y \leq 60$.

Exercise A.6: Maximize the function $f(x, y) = x^{\frac{1}{2}} + y^{\frac{1}{2}}$ subject to the constraint $x + 2y \leq 60$.

Exercise A.7: Maximize the function $f(x, y) = min\{x, y\}$ subject to the constraint $x + 2y \leq 60$.

Exercise A.8: Solve the constrained maximization problem in Example A.2 but change the constraint to $x + 2y \leq 10$

A More Preference Aggregation Rules

A.1 Plurality Vote

Plurality vote focuses on the goal of giving as many people as possible their top-ranked outcome. Because of this, it throws away most of the information about preferences and just focuses on the top of each individuals' ranking.

Definition A.1: Plurality Vote. The **score** of an outcome is the number of people who rank that outcome highest. Social preferences are determined by score.

Example A.1: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

$$\text{Scores. } a : 2, b : 0, c : 1$$

$$a \succ^* c \succ^* b$$

Example A.2: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

$$\text{Scores. } a : 2, b : 2, c : 1$$

$$a \sim^* b \succ^* c$$

A.2 Veto

While plurality vote attempts to maximize the number of people who get their favorite outcome, this method attempts to do the opposite: minimizing the number of people who get their least favorite outcome.

Definition A.2: Veto. The score of an outcome is the negative of the number of people who rank it last. Social preferences are determined by the score as in the other scoring methods above with a higher score being ranked higher.

Example A.3: Example 1.
1: $a \succ b \succ c$
2: $a \succ c \succ b$
3: $c \succ a \succ b$

$$\text{Scores. } a : 0, b : -2, c : -1$$

$$a \succ^* c \succ^* b$$

Example A.4: Example 2.
1: $a \succ c \succ b$
2: $a \succ c \succ b$
3: $b \succ c \succ a$
4: $b \succ a \succ c$
5: $c \succ a \succ b$

$$\text{Scores. } a : -1, b : -3, c : -1$$

$$a \sim^* c \succ^* b$$

Part IV

Solutions

B Chapter 1

Solution 1.1: *No, No, Yes*

It is not complete since there are people who are not siblings of eachother at all. It is not transitive if we think of sibling as being something broader than "having the same biological parents". For instance, person a's step-sibling b might have a step-sibling c who is not a step sibling of a. Sibling is symmetric though since if a is a sibling of b then b is a sibling of a.

Solution 1.2: *Yes, Yes, No*

Solution 1.3: *No, Yes, Yes*

Solution 1.4: If a relation is symmetric then if xRy is true, so is yRx. If yRx is true then so is xRy. Thus, if a relation is symmetric then either both of xRy and yRx are true or neither are true. But by completeness, for all x and y, at least one must be true. Thus, combining these, we have: for all x and y both xRy and yRx are true. Thus, every possible pair of things are related in both directions. This is the **universal** relation.

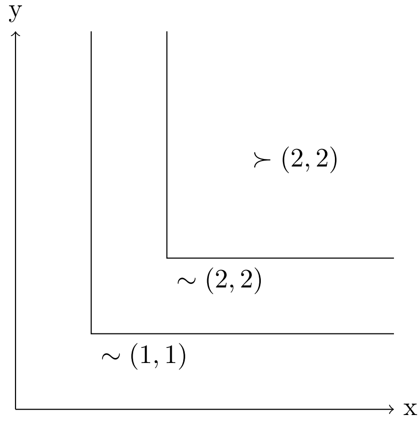
Solution 1.5:

1. *Transitive.*
2. *Transitive.*
3. *Not Transitive. Missing xRz)*

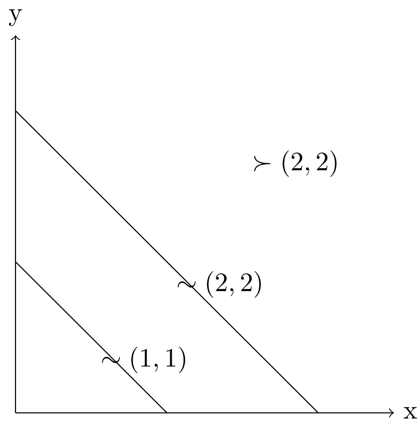
Solution 1.6:

1. *Neither.*
2. *Both.*
3. *Both.*
4. *Neither.*

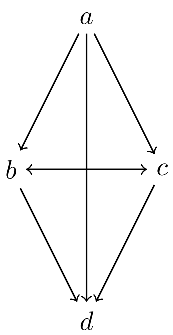
Solution 1.7:



Solution 1.8:



Solution 1.9:



B Chapter 2

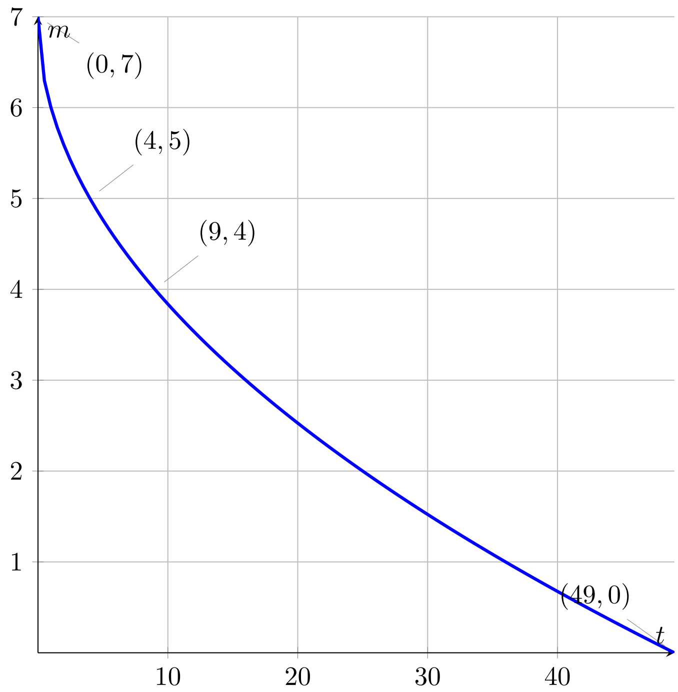
Solution 2.1: $A \succsim A, B \succsim B, C \succsim C, B \succsim A, B \succsim C, C \succsim A$

Solution 2.2: For example, $U(A) = 9, U(B) = 16, U(C) = 12$. But there are **many** possible solutions.

Solution 2.3: $(4,5) \sim (16,3)$

Solution 2.4: $U(9,4) = 7$ and $(9,4) \sim (0,7)$.

Solution 2.5:



Solution 2.6:

Solution D: iscuss the following statement: *Economists do not have to believe that utility functions exist in the minds of consumers for the concept to be useful.*

References

[1] Vincenzo Denicolò. Fixed agenda social choice theory: Correspondence and impossibility theorems for social choice correspondences and social decision functions. *Journal of Economic Theory*, 59(2):324–332, 1993.