1. Graphical Overview of Calculus and Optimization Theory

This section will take about three-four lectures. The idea is partly to provide a purely graphical roadmap of much of the territory we are going to cover later, partly to introduce a number of hard conceptual issues (e.g., necessary and sufficient conditions) in a low-tech environment. Doesn’t mean that the material is necessarily easy. Just that I’m going to concentrate entirely on the pictures to the exclusion of (almost) all symbols. Later on, we’ll come back and redo all these concepts. Some duplication, but it will be useful.

1.1. Necessary and Sufficient Conditions

Everybody takes a while to get the hang of these relationships. They will pop up throughout our calculus overview, so we’ll try to deal with them in an abstract context.

Consider the following groups of people

(1) Economists
(2) World renowned experts in the economics of spinach production
(3) Really clever people who know about agriculture.
(4) Agricultural economists

What can we say about these groups of people in terms of necessary and sufficient conditions? In particular, we’re interested in necessary and sufficient conditions for Callista Flockhart to be an ag economist. Mathematics is a bit more precise than this example, but it’s good to think about imprecise things.

Think: does belonging to one group imply anything about belonging to another?

(1) If you are an ag economist, then you pretty much have to be an economist
(2) If you are an expert in the economics of spinach production, you pretty much have to be an agricultural economists
(3) What about in the other direction? Obviously, there are a lot of economists who aren’t agricultural economists. And a lot of ag economists who know nothing about spinach.
(4) What about cleverness? Do you have to be really clever to be an ag economist?

So the answers are:

(1) A necessary condition for Callista Flockhart to be an ag economist is that she’s an economist
(2) A sufficient condition for Callista Flockhart to be an ag economist is that she’s a world renowned expert on the economics of spinach production.
(3) Cleverness and knowledge of spinach production are neither necessary nor sufficient.

The best way to think about necessary and sufficient conditions is in terms of set containment. Think about drawing the above four conditions as sets. Three of the four sets can be ordered by containment.

(1) the set of all world renowned experts in the economics of spinach production is contained in the set of all ag economists
(2) the set of all ag economists is contained in the set of all economists
(3) the set of really clever people who know about agriculture is neither contained in, or contains, any of the above sets.

So we have the following relationships

(1) $Y$ is necessary for $X$ if the set of all elements satisfying $Y$ contains the set of elements satisfying $X$. 
Figure 1. Necessary and Sufficient Conditions for an Ag Economist.

(2) $Z$ is sufficient for $X$ if the set of all elements satisfying $Z$ is contained in the set of elements satisfying $X$.

(3) $W$ is neither necessary nor sufficient for $X$ if there are elements in $X$ that are not in $W$ and vice versa.

Think:

(1) Necessary ... big set (economists $Y$)

(2) Sufficient ... small set (spinach guys $Z$)

NASC and “$$\rightarrow$$”:
Writing $X \rightarrow Y$ is the same thing as saying $Y$ is necessary for $X$. Thinking about it set-theoretically this is clear: the necessary property $Y$ is the containing set: being in $X$ (the contained set) implies being in $Y$ the containing set. Alternatively, $X \rightarrow Y$ is the same thing as saying $X$ is sufficient for $Y$.

NASC and “$$\neg$$”: saying $Y$ is necessary for $X$ is the same thing as writing $\neg Y \rightarrow \neg X$. (If you are not in the bigger set, you can’t be in a subset of the bigger set.

Typically, in math/economics, we try to find assumptions that enable us to make necessity/sufficiency statements.

(1) If $f$ is concave, then...

(2) If the Hessian of the Hessian of $f$ is globally negative definite, then
In set-theoretic terms, if you want to modify a property like the yellow set \( W \) so that it becomes sufficient for something \( (X) \), then you need to

1. either tighten \( W \) so that it fits inside of the thing you want it to be sufficient for, i.e., \( X \).
2. or expand \( X \)

Similarly, if you want to make \( W \) necessary for \( X \), you have to

1. either expand \( W \) so that \( X \) fits inside of it
2. or shrink \( X \) so that it fits inside of \( W \) (this is actually what we usually do)

E.g., you could take the ‘clever’ set \( W \) and qualify it to: \( W' = \) “the set of really clever people who’ve published at least seven articles related to spinach production in the AJAE”. We can now say a sufficient condition for belonging to \( Z \) is membership in \( W' \).

Generally, a condition is a “better” condition, the less restrictive it is. So your goal should be to find minimally shrinking the yellow set \( W' \) so that it fits inside of \( Z \).

On the other hand, of course, we could expand the set \( W \) to make it a necessary condition. Clever is too narrow a condition. So we could soften it some: e.g., a necessary condition for being a world renowned expert in spinach production is that you can speak English. (Without being too languagist about it, it would be hard to be world renowned if you couldn’t attend Ag Economists conferences.

Of course, all of the above exercises are rather silly, or at least wouldn’t help you publish in the AJAE. The exercises stop being silly when you start finding conditions that are easy to check.

1. A sufficient condition for lots of things (a local max being a global max, for example) is that a function is concave. This is pretty useful, but
2. A sufficient condition for a function to be concave is that the determinants of the minors of the Hessian alternate in sign. This is more useful (though hideous in practice) because you can actually do the computation (sometimes) and check whether it’s satisfied or not.

So the conditions for a condition to be a “good sufficient condition” are

1. Checkability
(2) relaxedness (bigness): 5 AJAE articles is a better condition than 17. (provided it really ensures world renown)

and the conditions for a condition to be a “good necessary condition” are

(1) Checkability

(2) tightness (smallness): “A necessary condition for Z is English fluency” is a pretty useless condition.

The examples above are not entirely silly. If Jeff Perloff were to endow a prestigious Jeff Perloff Memorial Spinach Production chair in Berkeley, I would certainly want to add as a condition of the endowment that only world renowned experts in Spinach production could be awarded the chair for life. Now think of the search committee. What’s a sufficient condition for this property, so that Perloff doesn’t dismantle the chair in disgust: Well, really clever with several articles in major journals is a natural one. Trouble is, really clever is pretty hard to check. So the tendency is impose conditions that can be checked easily: multiple articles in the AJAE. Trouble is that checkability enlarges the set, so you might end up with a dumb Spinach professor who has many articles on spinach.

1.2. Unconstrained Maximization: One Variable.

Fig. 2 above is the graph of a function f which is defined from $-\infty$ to $\infty$. Think of the graph going down and down forever at either end. We are interested in the relationship between the slope of the function f' at a point x (denoted $f'(x)$) and the extrema (maxima and minima) of the function. First note that Fig. 2 has
three local maxima and two local minima, one global maximum, one inflexion point but no global minimum (because the function keeps getting smaller).

**Informal definition:** A function \( f \) attains a local maximum (minimum) at \( x^* \) if there is a neighborhood of \( x^* \) such that \( f \) is as at least as large (small) at \( x^* \) as it is anywhere on this neighborhood.

What’s a neighborhood of \( x^* \)? For now, it is some small interval that \( x^* \) is “strictly inside,” i.e., not on the edge of.

**Question:** How big does the neighborhood of \( x^* \) have to be for \( f \) to obtain a local maximum at \( x^* \)? What’s the minimum size?

**Answer:** Obviously, there isn’t a minimum size. An arbitrarily small neighborhood will do, so long as there is one.

**Question:** Suppose I know that \( f \) attains a local maximum at a point \( x^* \) (in the interior of the domain of \( f \)). What can I say then about the slope of \( f \) at \( x^* \), i.e., the number \( f'(x^*) \)?

**Answer:** The slope of \( f \) has to be zero at \( x^* \), i.e., \( f'(x^*) = 0 \).

**Question:** What if \( x^* \) is not in the interior of the domain of \( f \)?

**Answer:** \( x^* \) not being in the interior of the domain of \( f \) implies that \( x^* \) is on one of the edges of the restricted (or constrained, we’ll come back to constrained maximization) domain.

In figure 4 of the graphical overview, the domain is restricted to the interval \([x, \bar{x}]\) and you have local maxes at the edges of the domain of the function. In figure 1, my function \( f \) is defined from \(-\infty \) to \( \infty \) so there are no edges. \( x^* \) is necessarily in the interior of the function and therefore it has slope \( f'(x^*) = 0 \).

**Question:** Under what conditions is it true to say that \( f \) attains a maximum at \( x^* \) implies that \( f'(x^*) = 0 \).

**Answer:** Two conditions
• $f$ is differentiable, i.e., “smooth”.
• $f$ is defined on a set that has “no edges” (formally, the domain of $f$ is an open set).

Another way of saying this: “for a differentiable function $f$ to attain a local maximum at an interior point $x^*$, it is necessary that $f'(x^*)$ is zero”

We’ll assume from now on that our functions are as differentiable as they need to be, and not bother to add the formally required adjectives.

**Mathematical language:** If $f$ is differentiable at $x^*$, a necessary condition for a maximum of $f$ at an interior point $x^*$ is that the slope of $f$ is zero at $x^*$. The condition is called the first order necessary condition for a maximum. (First order condition because we are talking about the first derivative).

Another way of saying this is: if the slope isn’t zero at $x^*$ then I know I don’t have a maximum at $x^*$.

**Set theoretic conception of necessity:** We started out with an example of $Y$ not being necessary for $X$, where

- $X$ is the set $\{(f,x) : f$ attains a local max at $x}\$.
- $Y$ is the set $\{(f,x) : f'(x) = 0\}$.

$Y$ isn’t big enough to contain $X$; after all, as specified, $f'$ mightn’t even exist. Now we modify things so as to get a necessity relationship. We could expand $Y$ so it contains $X$ or shrink $X$ so it’s contained in $Y$. In this case, the right thing to do is shrink $X$ to

$$X' = \{(f,x) : f$ is differentiable, has open domain and $f$ attains a local max at $x}\$$

We can now say $X' \subset Y$, or $Y$ is necessary for $X'$.

We’ll now move on to sufficiency.

**Question:** Suppose I know that the slope of $f$ is zero at an interior point $x^*$. Can I conclude that $f$ attains a local maximum at $x^*$?
Answer: Obviously not. It could be a local minimum. Or it could be an inflexion point.

Question: Suppose I know that the slope of $f$ is zero at $x^*$. What additional information do I need to know that $f$ attains a local maximum at $x^*$? I want the answer in terms of the slope of $f$ at $x^*$.

Answer: Observe that if $f$ attains a local maximum at $x^*$, then the slope of $f$ is positive to the left of $x^*$ and negative to the right of $x^*$. Thus the condition is that the slope is decreasing at $x^*$. In other words, the slope of the slope is negative. The slope of the slope of $f$ at $x^*$ is written as $f''(x^*)$.

Mathematical language: If $f'(x) = 0$ then a sufficient condition for a (strict) local maximum of $f$ is that the slope of $f'(\cdot)$ is negative at $x$. The condition is called the second order sufficiency condition for a maximum. (Second order condition because we are talking about the second derivative).

The following is almost but not quite true (which means that it is false: a necessary and sufficient condition for a differentiable function $f$ to attain a (strict) local maximum at an interior point $x^*$ is that $f'(x^*) = 0$ and $f''(x^*) < 0$. (It’s not quite true: think about $f(x) = -x^4$. This attains a global maximum at zero, but $f'(0) = f''(0) = f'''(0) = 0$.)

Question: Suppose I have a local maximum for $f$ at a point $x^*$ in the interior of the domain of $f$. What information about the slope of $f$ at $x^*$ would be sufficient to ensure that $f$ attains a global maximum at $x^*$?

Answer: For general differentiable functions, there is no calculus condition on $f$ at $x^*$ that will tell me the answer to this question. Knowing about the slope of $f$ at a point provides only local information about $f$ at $x^*$. Big problem: not much use to me to only know about local maxima. Can’t sell a local maximum to the CEO. However, for a certain (large) class of functions (including, but not restricted to, polynomials), you can answer the question based only on information about $f$ at $x^*$: take an infinite Taylor expansion. More on this later.

Question: When the function $f$ satisfies a certain property then knowing $f'(x^*) = 0$ is sufficient to conclude that $f$ attains a global maximum at $x^*$. What is that property? Clue. Look at Fig. 2 and try to figure out what needs to happen in order to get rid of the local extrema that aren’t global extrema.
Answer: If the slope of $f$ is everywhere decreasing, i.e., the slope of the slope of $f$ is always negative. A function satisfying this condition is called a concave function. Similarly, if the slope of $f$ is everywhere increasing, i.e., the slope of the slope is always positive, then knowing $f'(x^*) = 0$ is SUFFICIENT to conclude that $f$ attains a GLOBAL minimum at $x^*$. A function satisfying this condition is called a convex function.

1.3. **Convex sets, Concave and Convex Functions.**

Informal definition: A set $S$ is convex if the line joining any two points in $S$ is contained in the set. (See Fig. 3.)

Even more informal definition: A set $S$ is strictly convex if the line joining any two points in $S$ is contained in the interior of the set, i.e., not on the edge of it. This notion of a strictly convex set never appears in math books. Never mention that you heard about them from me, at least to a real mathematician. But it is a useful concept nonetheless.
Informal definition: A function \( f \) is concave if the set of points that lie below the graph of \( f \) (technically, the hypograph of \( f \)) is a convex set. Alternatively, the slope of \( f \) is always nonincreasing. Alternatively, the slope of the slope of \( f \) is always nonpositive. Alternatively, the graph of \( f \) lies everywhere below any tangent line to the graph.

Informal definition: A function \( f \) is convex if the set of points that lie above the graph of \( f \) (technically, the epigraph of \( f \)) is a convex set. Alternatively, the slope of \( f \) is always nondecreasing. Alternatively, the graph of \( f \) lies everywhere above any tangent line to the graph.

Informal definition: A set is concave ....? No such thing as a concave set.

Fact: If \( f \) is a differentiable, concave function, whose domain is an open set, then a necessary and sufficient condition for \( f \) to attain a global maximum at a point \( x^* \) in the interior of its domain is that \( f'(x^*) = 0 \).

Fact: If \( f \) is a differentiable, convex function, whose domain is an open set, then a necessary and sufficient condition for \( f \) to attain a global minimum at a point \( x^* \) in the interior of its domain is that \( f'(x^*) = 0 \).

Note that neither fact would be true without the caveats of differentiability and interiority.

To see that these facts are true, go back to Fig. 2. Observe that the area under the graph isn’t a convex set. To see the relationship of convexity of the set below the graph and the globlaness of local maxima, take any function \( f \) that has a local max that isn’t a global max. The \( f \) in Fig. 2 is an example. I’ll show that \( f \) can’t be concave. Join up the local maxima on the graph: for \( f \) to be concave, the line joining up these maxima HAS to be under the graph. But it can’t be, since both endpoints are local maxima.

That is, local max of \( f \) is not a global max \( \implies \) \( f \) is not concave which is an equivalent statement to the following: \( f \) concave \( \implies \) local max of \( f \) is a global max. Alternatively, a sufficient condition for a local max of \( f \) to be a global max is that \( f \) is concave.

Much of the work in economics involves choosing assumptions so that the functions you are interested in are either concave or convex, depending on whether you want a minimum or a maximum. Much of the art (hype) in economics involves telling economic stories about why the functions you are interested in are
really concave or convex, and pretending that you didn’t choose these functional forms just to make things work.

1.4. **Strictly Concave and Convex Functions and Strict (Locally Unique) Maxima**

A function $f$ is strictly concave if the set of points that lie below the graph of $f$ is a strictly convex set. For one-dimensional functions, this means that the slope of $f$ is always strictly decreasing.

**Informal definition:** A function $f$ is strictly convex if the set of points that lie above the graph of $f$ is a strictly convex set. For one-dimensional functions, this means that the slope of $f$ is always strictly increasing.

In subsection 1.3, dealing with concave functions and maxima, whenever we talk about concave functions and local or global maxima, if we have a strictly concave function, we infer that we have a strict global maximum, which is worth a great deal more.