# ARE211, Fall 2004

LECTURE #14: TUE, OCT 26, 2004 PRINT DATE: DECEMBER 17, 2004 (GRAPHICAL3)

CONTENTS



## 3. GRAPHICAL OVERVIEW OF OPTIMIZATION THEORY (CONT)

### 3.6. Level Sets, upper and lower contour sets and Gradient vectors (cont)

Vectors: Recall that a vector in $\mathbb{R}^n$ is a collection of $n$ scalars. A vector in $\mathbb{R}^2$ is often depicted as an arrow. Properly the base of the arrow should be at the origin, but often you see vectors that have been "picked up" and placed elsewhere. Example below.

Gradient vectors: When economists draw level sets through a point, they frequently attach arrows to the level sets. These arrows are pictorial representation of the *gradient vector*, i.e., the slope of $f$ at $\mathbf{x}$, $\mathbf{f}'(\mathbf{x})$. Its components are the partial derivatives of the function $f$, evaluated at $\mathbf{x}$, i.e., $(f_1(\mathbf{x}), \cdots, f_n(\mathbf{x}))$

Example: $f(\mathbf{x}) = 2x_1x_2$, evaluated at $(2,1)$, i.e., $\mathbf{f}'(2,1) = (2x_2, 2x_1) = (2,4)$. Draw the level set through $(2,1)$, draw the gradient through the origin, lift it up and place its base at $(2,1)$. Generally, the gradient of a function with $n$ arguments is a point in $\mathbb{R}^n$, and for this reason, you often see the gradient vector drawn in the domain of the function, e.g., for functions in $\mathbb{R}^2$, you often draw the gradient vector in the horizontal plane.

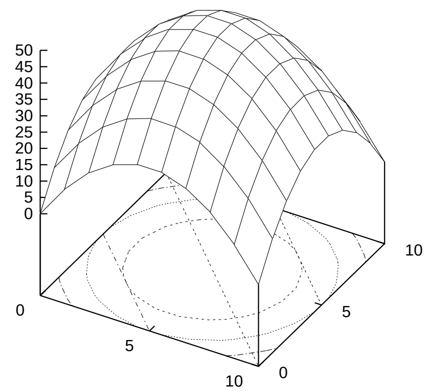


FIGURE 11. Level set and gradient vector through a point

The gradient vector points in the direction of steepest ascent: Consider Fig. 11. Let **x** denote the point in the domain where the first straight line touches the circle. The graph represents a nice symmetric mountain which you are currently about to scale. You are currently at the point **x**. You're a macho kind of person and you want to go up the mountain in the steepest way possible. Ask yourself the question, looking at the figure. What direction from **x** is the steepest way up? Answer is: the direction perpendicular to the straight line. Draw an arrow pointing in this direction. Now the *gradient vector* of $f$ at **x** is an arrow pointing in precisely the direction you've drawn.

The following things about the gradient vector are useful to know:

- its length is a measure of the steepness of the function at that point (i.e., the steeper the function, the longer is the arrow.)
- as we've seen it is perpendicular to the level set at the point **x**
- it points inside the upper contour set. **Note Well: It could point into the upper contour set, but then go out the other side!**
- as we've seen, it points in the direction of steepest ascent of the function.

When we get to constrained optimization, we'll talk a lot more about this vector.

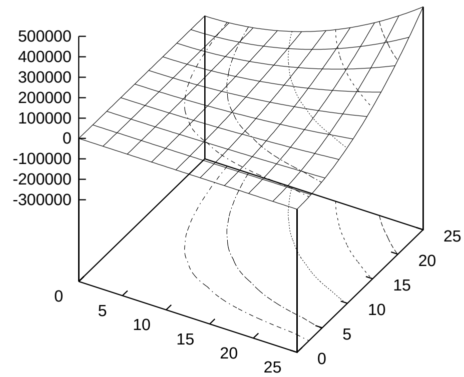


FIGURE 12. Level and contour sets of a quasiconcave function

### 3.7. **Quasiconcavity, quasiconvexity**

Recall that convex and concave functions were characterized by whether or not the area above or below their graphs were convex. There's another class of functions that are characterized by whether or not their *upper or lower contour sets* are convex.

It turns out that for economists, the critical issue about a function is not whether it is convex or concave, but simply whether its *upper or lower contour sets* are convex.

This should be very familiar to you: recall that what matters about a person's utility function is not the *amount* of utility that a person receives from different bundles but the shape of the person's indifference curves. Well, indifference curves are just the level sets. Notice that you always draw functions that have convex upper contour sets: called the law of diminishing marginal rate of substitution.

So whether or not a function is concave or not turns out to be of relatively minor importance to economists. Consider Fig. 12. Though it's not entirely clear from the picture, the function graphed here has a striking resemblance to the concave function in the preceding graph: *the two functions have exactly the same level sets*. The second function is clearly not concave, but from an economic standpoint it works just as well as the concave function.

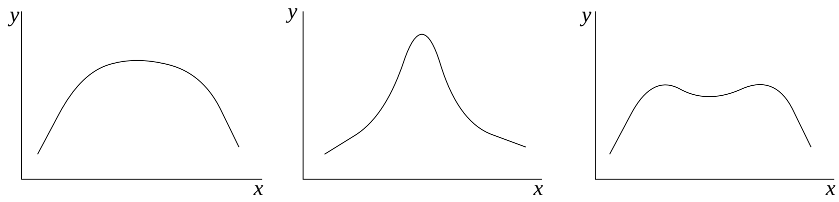


FIGURE 13. A concave, quasi concave and a neither funtion

**Definition:** A function is quasiconcave if *all* of its upper contour sets are convex.

**Definition:** A function is quasiconvex if *all* of its lower contour sets are convex.

So in most of the economics you do, the assumption you will see is that utility functions are quasi-concave.

Most people find this concept rather difficult even though they are quite used to assuming diminishing mrs. A good test of whether you understand quasi concavity or not is to look at the concept in one dimension. Fig. 13 illustrates a concave, quasiconcave and a not quasiconcave function of one variable.

### 3.8. **Strict Quasiconcavity**

Just as there are strictly concave vs weakly concave functions or just concave functions, there are also strictly quasiconcave vs quasiconcave functions. The definition of strict quasi-concavity is less clean than the definition of quasi-concavity, but the properties of strictly quasi-concave functions are MUCH cleaner.

**Definition:** A function $f$ is **strictly quasi-concave** if for any two points $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x} \neq \mathbf{y}$, in the domain of $f$, whenever $f(\mathbf{x}) \leq f(\mathbf{y})$, then $f$ assigns a value strictly higher than $f(\mathbf{x})$ to every point on the line segment joining $\mathbf{x}$ and $\mathbf{y}$ except the points $\mathbf{x}$ and $\mathbf{y}$ themselves. Thus, s.q.c., rules out quasi-concave functions that have:

- straight line level sets (pyramids).
- flat spots (pyramids with helipads on the top).

Note that the above definition would not work if we replaced "$f(\mathbf{x}) \leq f(\mathbf{y})$" with "$f(\mathbf{x}) = f(\mathbf{y})$." More specifically, consider the following, similar but weaker condition: "for any two points $\mathbf{x}$ and $\mathbf{y}$ in the domain of $f$, whenever $f(\mathbf{x}) = f(\mathbf{y})$, then $f$ assigns a value strictly higher than $f(\mathbf{x})$ to every point on the open line segment stricly between $\mathbf{x}$ and $\mathbf{y}$." This condition is satisfied by *any* function which has the property that no point in the range is reached from more than one point in the domain. E.g., consider the function defined on $\mathbb{R}$ by $f(x) = 1/x$, for $x \neq 0$; $f(0) = 0$. This function is not even quasi-concave, and so certainly not strictly so: To see that it's not quasi-concave, note that the levelset corresponding to $-1$ is $[\infty, -1] \cup \mathbb{R}_+$, which is not a convex set. But since there are no points $\mathbf{x}$ and $\mathbf{y}$ s.t. $f(\mathbf{x}) = f(\mathbf{y})$, the function satisfies the latter condition trivially. (Thanks to Rob Letzler (2001) for this example.)

**Definition:** A function is strictly quasiconcave if all of its *upper* contour sets are strictly convex sets and *none* of its *level* sets have any width (i.e., no interior).

**Definition:** A function is strictly quasiconvex if all of its *lower* contour sets are strictly convex sets and *none* of its *level* sets have any width (i.e., no interior).

The first condition rules out straight-line level sets while the second rules out flat spots.

Two questions: Why do economists care so much about quasi-concavity? What is this long discussion doing in an overview of optimization theory?

The answer to both questions is that quasi-concavity is almost, but not quite, as good as concavity in terms of providing second order conditions for a maximum. Recall that if $f$ is concave, then a necessary and sufficient condition for $f$ to attain a global maximum at $\mathbf{x}^*$ is that the *first order conditions* for a max are satisfied at $\mathbf{x}^*$.

We can almost, but not quite, replace the word concavity by quasiconcavity in the above sentence. "Almost," however, is a very large word in mathematics: it's *not* true that if $f$ is quasiconcave, then a necessary and sufficient condition for $f$ to attain a global maximum at $\mathbf{x}^*$ is that the first order conditions for a max are satisfied at $\mathbf{x}^*$.

For example, consider the problem: max $f(x) = x^3$ on $[-1,1]$, the function is strictly quasi-concave, and the first order conditions for a max are satisfied at $x = 0$, but it doesn't attain a max/min or anything at 0.

The following is true however: if $f$ is quasiconcave *and the gradient of $f$ never vanishes*, then a necessary and sufficient condition for $f$ to attain a global maximum at $\mathbf{x}^*$ is that the first order conditions for a max are satisfied at $\mathbf{x}^*$.

Observe how the caveat takes care of the nasty example.

Now note that while quasi-concavity is a very useful second order condition for a *constrained* maximum, it's a true but pretty useless one for an *unconstrained* max.

The following statement is certainly true, but not particularly helpful. In fact it's particularly useless. If $f$ is quasiconcave and the gradient of $f$ never vanishes, then a necessary and sufficient condition for $f$ to attain an *unconstrained* global maximum at $\mathbf{x}^*$ is that the first order conditions for a max are satisfied at $\mathbf{x}^*$.

Why isn't this particularly helpful?

Quasi-concavity can, however, provide some help in an unconstrained maximization problem, because it guarantees that a strict local maximum is a global maximum. That is, if you know your function is quasi-concave, then if your first order conditions and *local* second order conditions for a *strict* local maximum are satisfied, you know much more than you would know if you didn't know that your function were quasi-concave.

Fact: : If $f$ is quasi-concave, a strict local maximum is a strict global maximum.

Proof: : Consider a function with a strict local maximum that isn't a strict global maximum. We'll show that the function can't be quasiconcave. (Common way to prove things in math: showing "not B implies not A" is equivalent to (and often much easier than) showing that "A implies B". In this case "B" is the existence of a unique global maximum. "A" is the q.c.ness of the function. We're showing not B implies not A.

Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ which has a strict local max at $\mathbf{x}$, but there exists a vector $\mathbf{y} \in \mathbb{R}^2$ such that $f$ is at least as high at $\mathbf{y}$ as at $\mathbf{x}$.

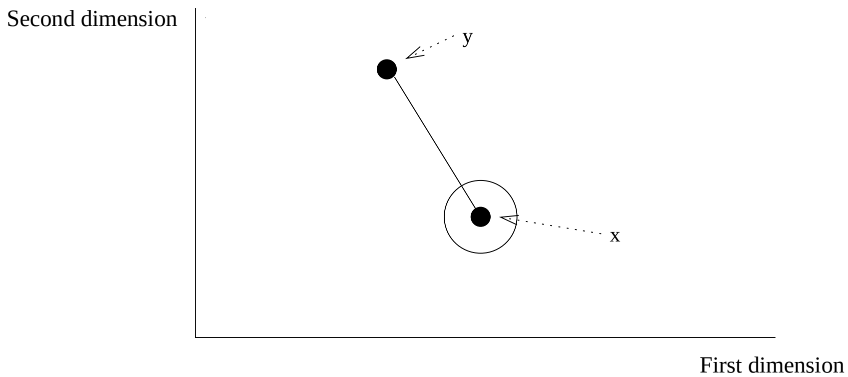


FIGURE 14. A function with a strict local maximum that is not a strict global maximum

- Consider the level set that goes through **x**.
- As we've seen before, the strict local max **x** is an isolated point on this level set (though there may be other points further away that belong to the level set through **x**).
- The point **y** must belong to the upper contour set corresponding to $f(\mathbf{x})$ (because $f$ is at least as high at **y** as it is at **x**.)
- Join up the line between **x** and **y**.
- But we know that at least a part of this line doesn't belong to the upper contour set corresponding to $f(\mathbf{x})$, because $f$ is larger at **x** than everywhere in a nbd of **x**. Remember the upper contour set lives in the domain of the function i.e. in the horizontal plane. Conclude that the upper contour set corresponding to $f(\mathbf{x})$ is *not* a convex set. Therefore, $f(\cdot)$ is not a quasiconcave function.

Note that it is *not true* that a weak local maximum of a quasi-concave function is necessarily a global max. Fig. 15 provides an example of a quasi-concave function $f$ with lots of local maxima that are not global maxima. For example, $f$ attains a local max at $x_0$.

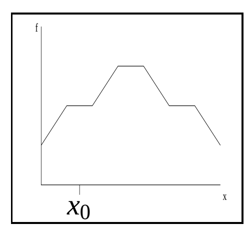


FIGURE 15.

Recapitulate: economists focus on quasi-concave functions because they have precisely the property that they care about: (a) first order necessary conditions for a *constrained* maximum of a quasi-concave function are also sufficient for a constrained max, provided you add a little caveat (b) for any quasi-concave function, a strict local maximum is a strict global maximum.

In economics, the caveat is imposed as *local non-satiation*.

### 3.9. **Constrained Optimization: Several Variables**

As I've said before, economists almost never solve unconstrained maximization problems. The whole mission of economics is to optimize subject to limited resources, and the economic limitations correspond to mathematical constraints.

Throughout this entire subsection, we're going to assume that all the functions we have to deal with are as differentiable as we need them to be. We're going to be interested in "exceptions to the rule," but unlike in the first graphical lecture, we're not interested in exceptions that arise because of non-differentiability.

There are two kinds of constraints: equality constraints and inequality constraints. We'll consider a few examples.

(1) Maximizing subject to being on a line; Economics 101 solution to consumer's utility maximization subject to a budget constraint.
(2) Maximizing subject to being on a line plus nonnegativity constraints. Consumer's problem has the additional condition that you can't consume negative quantities.
(3) Maximizing subject to being inside a convex set (partly defined by nonnegativity constraints.)

Theme of this section. All nonlinear programming problems (NPP) are essentially the same. Focus on the necessary condition for a constrained max: relates the arrows associated with the level sets of the objective function (gradient vector of objective function) with the arrows associated with the level sets of the constraint functions (gradient vectors) that are satisfied with equality. So there's really only one thing you have to know in order to understand NPP. (You need to know quite a lot more to actually go out and *compute* a solution to a real NPP, but if all you want to do is understand NPP theory, this is all you need to know.)

Talk about nonnegative and positive cones:

- math cones are like icecream cones only longer: infinitely long in fact.
- Draw a couple of vectors: the cone consists of all the vectors that are inside the boundaries defined by these vectors.

- Indicate the nonnegative cone defined by them: the vectors that are inside the cone.
- Draw a vector inside the cone
- Ask what is the positive cone defined by a single vector.
- Note that the positive and the nonnegative cones defined by $\{x, -x\}$, $x \neq 0$, is the entire line through $x$ (including zero). That is, in this case, the positive and nonnegative cones are identical.
- Draw three cones in the plane, observe that the inside one will be redundant, i.e., unnecessary.

Mantra: (Except for a couple of bizarre exceptions) a necessary condition for $\mathbf{x}^*$ to solve a constrained maximization problem is that the gradient vector of the objective function at $\mathbf{x}^*$ belongs to the nonnegative cone defined by the gradient vector(s) of the constraint(s) that are satisfied with equality at $\mathbf{x}^*$. (*Parenthetical remark*: The gradient of the objective function also belongs to the *positive* cone defined by the gradient vector(s) of the *binding* constraint(s).)

By a "binding constraint" we mean a constraint with the following property: if you relax the constraint a little bit, then the maximized value of the objective function will be increased. (By a little bit, I mean by an arbitrarily small amount.) We'll see that this is different from saying: "the constraint is satisfied with equality." **(Please note, however, that S&B (p. 430) does *not* make any distinction between binding constraints and those that are satisfied with equality. What they call binding is what we merely call satisfied with equality.)**

NB: you never, never, never worry about the second sentence of the mantra until *after* you've found a point that satisfies the first, i.e., until you've found a point $\mathbf{x}^*$ that belongs to the nonnegative cone defined by the constraints that are satisfied with equality at $\mathbf{x}^*$. Computer algorithms use the *first* sentence primairily: they go around the constraint set looking at points, checking which constraints are satisfied with equality, looking at the nonnegative cones that these constraints generate, and checking the first sentence. *Only when they've found a point where the first sentence is satisfied do they bother with the second sentence*.

Coming back to the mantra, there are two exceptions to the nonnegative cone rule. One arises only because words aren't as precise as mathematical symbols. The other is substantive.

- (purely semantic) if the constrained max problem happens to have an unconstrained solution the mantra as stated above will typically not work. Here the problem is a matter of terminology: when

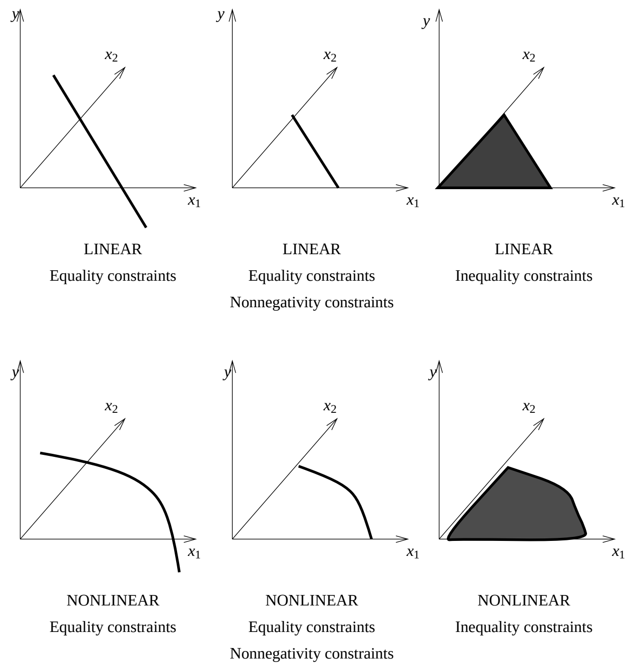


FIGURE 16. Six different (not really) versions of constrained max problems

we write the math version of the mantra, then it certainly works. the words of the mantra don't exactly capture the mathematics.

- (substantive) the other exception is the socalled Constraint Qualification. More on it later.

There are several different versions of constrained optimization (Fig. 16): Each of the three below has a *linear* and a *nonlinear* version. Assume throughout the lecture that $f$ is quasi-concave and $g$ is quasi-convex.

(1) Maximize a function subject to an *equality* constraint. (i.e., solution has to lie on a line.)

$$\max\ f(\mathbf{x}) \text{ subject to } g(\mathbf{x}) = b$$

In other words, maximize $f$ subject to the condition that $x$ lies on a *level set* of the function $g$. We'll call $f$ the objective function and $g$ the constraint function.

(2) Maximize a function subject to an equality constraint and nonnegativity constraints

$$\max\ f(\mathbf{x}) \text{ subject to } g(\mathbf{x}) = b; \mathbf{x} \geq \mathbf{0}.$$

In other words, maximize $f$ subject to the condition that $\mathbf{x}$ lies on a *level set* of the function $g$ and $\mathbf{x}$ is a nonnegative vector.

(3) Maximize a function subject to one or more *inequality* constraints.

a single constraint: $\max\ f(\mathbf{x})$ subject to $g(\mathbf{x}) \leq b$.

multiple constraints: $\max\ f(\mathbf{x})$ subject to $\mathbf{g}(\mathbf{x}) \leq \mathbf{b}$.

In other words, maximize $f$ subject to the condition that $\mathbf{x}$ lies in a *lower contour set* or the intersection of several.

The difference between linear and nonlinear is whether the constraints are linear functions or not: i.e., if the level sets are straight lines, or not.

**Question:** Notice that we assume "$\leq$". What about greater than or equal to...?

**Answer:** If the original constraint is $g(\mathbf{x}) \geq b$, write it as $-g(\mathbf{x}) \leq -b$.

**Fact:** The first five cases (found below) are all special cases of the sixth. This is important: you really want to be able to fit *everything* into the same general framework.

- A linear function is just a special case of a nonlinear function.
- Write $g(\mathbf{x}) = b$ as two constraints: $-g(\mathbf{x}) \leq -b$; $g(\mathbf{x}) \leq b$.
- What about nonnegativity constraints: the set $\left\{\mathbf{x} \in \mathbb{R}^2 : x_i \geq 0\right\}$ is a *lower contour set* of a certain function, i.e., the function $g_i(\mathbf{x}) = -x_i$; Plot $g_i(\mathbf{x})$ in 3D. So we can draw the conventional budget set as the intersection of three lower contour sets.

  Aside: It's useful here to observe the relationship between the *budget set*, which people are used to, and the lower contour set of the function $\mathbf{p} \cdot \mathbf{x}$. People find it hard to see where the gradient vector attached to the budget set comes from. That's because they don't think about the budget set as the lower contour set of a function.