Unconstrained Optimization

Joshua Wilde, revised by Isabel Tecu, Takeshi Suzuki and María José Boccardi

August 13, 2013

1 Definitions

Economics is a science of optima. We maximize utility functions, minimize cost functions, and find optimal allocations. In order to study optimization, we must first define what maxima and minima are.

Let $f : X \to Y$ be a function. Then

1. We say $x \in X$ is a **local maximum** of f on X if there is $r > 0$ such that $f(x) \geq f(y)$ for all $y \in X \cap B(x, r)$. If the inequality is strict, then we have a **strict local maximum.**

2. We say $x \in X$ is a **local minimum** of f on X if there is $r > 0$ such that $f(x) \leq f(y)$ for all $y \in X \cap B(x, r)$. If the inequality is strict, then we have a **strict local minimum.**

3. We say $x \in X$ is a **global maximum** of f on X if $f(x) \geq f(y)$ for all $y \in X$. If the inequality is strict, then we have a **strict global maximum.**

4. We say $x \in X$ is a **global minimum** of f on X if $f(x) \leq f(y)$ for all $y \in X$. If the inequality is strict, then we have a **strict global minimum.**

Optimization problems are often written in the form

$$\max_{x \in X} f(x)$$

In this notation max refers to the global maximum of f on X. The point at which the maximum is achieved is called the maximizer of f on X and usually denoted x^* or $\text{argmax}_{x \in X} f(x)$. If there are multiple global maxima of f on X, then $\text{argmax}_{x \in X} f(x)$ denotes the whole set of them.

In the following we seek conditions whereby we can tell whether maxima or minima exist, and if a point $x^* \in X$ is a local maximum or minimum.

2 Weierstrass Theorem

The Weierstrass Theorem is one of the most important in economics. It states conditions under which we are guaranteed to find a global maximum. We state it here without proof.

Weierstrass Theorem Let $D \subset \mathbb{R}^N$ be a compact set, and $f : D \to \mathbb{R}$ a continuous function. Then f attains a (global) maximum and a (global) minimum on D, i.e. $\exists\, z_1, z_2 \in D$ such that $f(z_1) \geq f(x) \geq f(z_2)\ \forall\, x \in D$.

The crucial part of this theorem is that the set D has to be compact, that is, bounded and closed. f attains its maximum either at the boundary of D or in the interior. The following discussion will focus on extreme points to be found in the interior of a set (in fact, we will usually let the domain of our function be an open set). But we should not forget that if a function's domain is compact, the extreme points can also be attained on the boundary of the domain. These extreme points cannot be found with the first order conditions that apply for open domains.

3 First Order Conditions

Let $F : U \to \mathbb{R}^1$ be a continuously differentiable function defined on U an open subset of $\mathbb{R}^n$. If $\mathbf{x}^*$ is a local maximum or minimum of F in U, then

$$DF(\mathbf{x}^*) = \mathbf{0}.$$

Notice that the converse is not true, namely "if $DF(\mathbf{x}^*) = \mathbf{0}$, then $\mathbf{x}^*$ is a local maximum or a local minimum". For example, consider the function $f(x) = x^3$ on $\mathbb{R}^1$. Then $Df = 3x^2$, which implies that when $x = 0$, $Df(0) = 0$. However, $x = 0$ is not a local maximum or minimum since if $x > 0$ in any ϵ-ball about 0 then $f(x) > f(0)$, and if $x < 0$ in any ϵ-ball about 0 then $f(x) < f(0)$.

A condition which only goes in the $\Rightarrow$ direction such as this is called a **necessary condition.** A condition which only goes in the $\Leftarrow$ direction is called a **sufficient condition.** Therefore, if a condition goes in both directions, we say it is a necessary and sufficient condition. Note that our first order condition for maxima or minima is a necessary condition, but not sufficient.

Examples

1. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x^3 - 3x^2$.
   Then $Df(x) = 6x^2 - 6x = 6x(x-1)$, which implies that the only candidates for a maximum or minimum are $x = 0$ and $x = 1$. Without further conditions, however, we cannot say whether these are actual maxima or minima.

2. Let $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^3 - y^3 + 9xy$.
   Then $DF(x) = (3x^2 + 9y, -3y^2 + 9x)$, which implies that the only candidates for a maximum or minimum are when $3x^2 + 9y = 0$ and $-3y^2 + 9x = 0$. Solving the first equation for y yields $y = -\frac{1}{3}x^2$. Plugging this into the other equation we have:

   $$0 = -3y^2 + 9x = -3\left(-\frac{1}{3}x^2\right)^2 + 9x = -\frac{1}{3}x^4 + 9x.$$

   This equation can be re-written as:

   $$-\frac{1}{3}x^4 + 9x = 27x - x^4 = x(27 - x^3),$$

   which implies that $x = 0$ and $x = 3$ are possible solutions. Plugging the x solutions into the equation for y gives $y = 0$ and $y = -3$ respectively. Therefore, the only possible optima are at $(x, y) = (0, 0)$ and $(x, y) = (3, -3)$. Without further conditions, however, we cannot say whether these are actual maxima or minima.

4 Second Order Conditions

The following theorem provides a sufficient condition for finding local maxima or minima:

Let $F : U \to \mathbb{R}$ be twice continuously differentiable, where U is an open subset of $\mathbb{R}^n$, and the first order condition holds for some $\mathbf{x}^* \in U$:

1. If the Hessian matrix $D^2F(\mathbf{x}^*)$ is a negative definite matrix, then $\mathbf{x}^*$ is a strict local maximum of F.

2. If the Hessian matrix $D^2F(\mathbf{x}^*)$ is a positive definite matrix, then $\mathbf{x}^*$ is a strict local minimum of F.

3. If the Hessian matrix $D^2F(\mathbf{x}^*)$ is an indefinite matrix, then $\mathbf{x}^*$ is neither a local maximum nor a local minimum of F In this case $\mathbf{x}^*$ is called a saddle point.

Notice again, however, that this proof does not go both ways. For example, it is not true that all local minima have positive definite Hessian matrices. For example, take the function $f(x) = x^4$, which has a local minimum at $x = 0$, but its Hessian at $x = 0$ is $D^2f(x) = 0$, which is not positive definite.

The following theorem provides weaker necessary conditions on the Hessian for a local maximum or minimum:

1. Let $F : U \to \mathbb{R}$ be twice continuously differentiable, where U is an open subset of $\mathbb{R}^n$, and $\mathbf{x}^*$ is a local maximum of F on U. Then $DF(\mathbf{x}^*) = 0$, and $D^2F(\mathbf{x}^*)$ is negative semidefinite.

2. Let $F : U \to \mathbb{R}$ be twice continuously differentiable, where U is an open subset of $\mathbb{R}^n$, and $\mathbf{x}^*$ is a local minimum of F on U. Then $DF(\mathbf{x}^*) = \mathbf{0}$, and $D^2F(\mathbf{x}^*)$ is positive semidefinite.

According to the weaker necessary conditions, if we can find $\mathbf{x}^*$ such that $DF(\mathbf{x}^*) = \mathbf{0}$ and $D^2F(\mathbf{x}^*) = \mathbf{0}$ is either negative (or positive) semidefinite, then that $\mathbf{x}^*$ is a candidate for a local maximum (or minimum). However, we cannot know for sure without further inspection.

Examples

1. Recall the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 2x^3 - 3x^2$ has $DF(x) = 0$ when $x = 0$ or $x = 1$. We can calculate that $D^2F(x) = 12x - 6$. When $x = 0$, then $D^2F(x) = -6$ which is negative definite, so we can be sure that $x = 0$ is a local maximum. However, when $x = 1$, then $D^2F(x) = 6$ which is positive definite, so we can be sure that $x = 1$ is a local minimum.

2. Recall the function $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^3 - y^3 + 9xy$ has $DF(x, y) = 0$ when $(x, y) = (0, 0)$ or $(x, y) = (3, -3)$. We can calculate that

$$D^2F(x) = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}.$$

When $(x, y) = (0, 0)$, then

$$D^2F(x) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix}.$$

In this case, the first order leading principal minor (the determinant of the matrix left after we delete the last row and column, or the determinant of the top left element) is 0, and the second order principal minor (the determinant of the whole matrix) is -81. Therefore, this matrix is indefinite, and $(x, y) = (0, 0)$ is neither a maximum or minimum.

When $(x, y) = (3, -3)$, then

$$D^2F(x) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}.$$

In this case, the first order leading principal minor is 18, and the second order principal minor is 243. Therefore, this matrix is positive definite, and $(x, y) = (0, 0)$ is a strict local minimum.

5 Concavity, Convexity, and Global Optima

Let $F : U \to \mathbb{R}$ be twice continuously differentiable, and U an open set. Then the function F is concave iff $D^2F(\mathbf{x})$ is negative semidefinite for all $\mathbf{x} \in U$, and is convex iff $D^2F(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in U$.

If F is a concave function and $DF(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in U$, then $\mathbf{x}^*$ is a global maximum of F on U. If F is a convex function and $DF(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x}^* \in U$, then $\mathbf{x}^*$ is a global minimum of F on U.

Example

1. Inspect the function $F(x, y) = x^2 + y^2$, and find any maxima or minima. Can you tell whether they are local or global?

   $$DF(x, y) = (2x, 2y) \Rightarrow DF(x, y) = \mathbf{0} \text{ for } x = 0 \text{ and } y = 0$$

   The Hessian is

   $$D^2F(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

   Since this is a diagonal matrix with only strictly positive entries on the diagonal, it is positive definite, independent of x and y. Therefore, it is positive definite at the critical point $(0, 0)$, and $(0, 0)$ is a strict local min (second order condition). But we can say even more: The Hessian is positive definite at all points, hence also positive semi-definite (a weaker condition) at all points. Therefore $F(x, y)$ is convex. By the theorem above, the point $(0, 0)$ is a strict global minimum.

2. Consider the function $f(x) = x^4$. Find the critical values and classify them.
   $f'(x) = 4x^3$ so the only critical point is $x^* = 0$. The Hessian is $f''(x) = 12x^2$. Evaluated at x^*, the Hessian is $f''(x^*) = 0$, so the second order condition does not tell us whether x^* is a maximum or a minimum. However, looking at the Hessian for all points in the domain gives us more information: $f''(x) = 12x^2 \geq 0$ for all x, so the Hessian is positive semi-definite on the whole domain and f is convex. Therefore, x^* has to be a global minimum.

3. Consider $F(x, y) = x^2y^2$. Find the critical values and classify them.

   $$DF(x, y) = (2xy^2, 2x^2y) \Rightarrow DF(x, y) = \mathbf{0} \text{ for } x = 0 \text{ or } y = 0$$

   We have infinitely many critical points which are of three different forms:

   $$z_1 = (x, 0) \text{ with } x \neq 0, \quad z_2 = (0, y) \text{ with } y \neq 0, \quad z_3 = (0, 0)$$

   The Hessian is

   $$D^2F(x, y) = \begin{pmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{pmatrix}$$

   First of all, the Hessian is not always positive semidefinite or always negative definite (first oder principal minors are ≥ 0, second order principal minor is ≤ 0), so F is neither concave nor convex.
   Let's determine the definiteness of $D^2F(x, y)$ at critical points of the form $(x, 0)$ with $x \neq 0$. In this case the Hessian is NOT negative semidefinite, so these points cannot be local maxima. The Hessian is positive semidefinite, but that is not sufficient to conclude anything else.
   Similarly, at critical points of the form $(0, y)$ with $y \neq 0$ the Hessian is NOT negative definite, so these points cannot be local maxima either, but we can't say more.
   At the critical point $(0, 0)$, the Hessian is positive semidefinite and negative semidefinite, so we can't say anything.
   This is an example where the conditions on the Hessian do not provide a lot of information. But we can use "common sense" instead: $F(x, y) \geq 0$ for all x, y, so points at which $F(x, y) = 0$ have to be global minima. Therefore, all the critical points are global minima.

6 Homework

1. For each function, determine whether the it definitely has a maximum, definitively does not have a maximum, or that there is not enough information to tell, using the Weierstrass Theorem. If it definitely has a maximum, prove that this is the case.

   (a) $f : \mathbb{R} \to \mathbb{R},\ f(x) = x.$

   (b) $f : [-1, 1] \to \mathbb{R},\ f(x) = x$

   (c) $f : (-1, 1) \to \mathbb{R},\ f(x) = x$

   (d) $f : [-1, 1] \to \mathbb{R},\ f(x) = \begin{cases} 0 \text{ if } x = 1 \\ x \text{ otherwise} \end{cases}$

   (e) $f : \mathbb{R}_{++} \to \mathbb{R},\ f(x) = \begin{cases} 1 \text{ if } x = 5 \\ 0 \text{ otherwise} \end{cases}$

2. Consider the standard utility maximization problem

$$\max_{x \in B(p,I)} U(x), \text{ where } B(p, I) = \{x \in \mathbb{R}^n_+ | \ p \cdot x \leq I\}$$

   Prove a solution exists for any $U(x)$ continuous, $I > 0$ and $p \in \mathbb{R}^n_{++}$. Show a solution may not necessarily exist if $p \in \mathbb{R}^n_+$.

3. Search for local maxima and minima in the following functions. More specifically, find the points where $DF(\mathbf{x}) = 0$, and then classify then as a local maximum, a local minimum, definitely not a maximum or minimum, or can't tell. Also, check whether the functions are concave, convex, or neither. The answers (except for the concavity/convexity part) are found in the back of Simon and Blume, Exercises 17.1 - 17.2.

   (a) $F(x, y) = x^4 + x^2 - 6xy + 3y^2$

   (b) $F(x, y) = x^2 - 6xy + 2y^2 + 10x + 2y - 5$

   (c) $F(x, y) = xy^2 + x^3y - xy$

   (d) $F(x, y) = 3x^4 + 3x^2y - y^3$

   (e) $F(x, y, z) = x^2 + 6xy + y^2 - 3yz + 4z^2 - 10x - 5y - 21z$

   (f) $F(x, y, z) = \left(x^2 + 2y^2 + 3z^2\right) e^{-\left(x^2+y^2+z^2\right)}$