University of Warwick
Math for Economics, 2009 - 10

Lecture 2: Unconstrained Optimization

1 Infimum and Supremum

Fix a set $Y \subseteq \mathbb{R}$. A number $\alpha \in \mathbb{R}$ is said to be *an upper bound of* Y if $y \leq \alpha$ for all $y \in Y$, and is said to be *a lower bound of* Y if the opposite inequality holds. Number $\alpha \in \mathbb{R}$ is said to be *the least upper bound of* Y, denoted $\alpha = \sup Y$, if: (i) α is an upper bound of Y; and (ii) $\gamma \geq \alpha$ for any other upper bound γ of Y. Analogously, number $\beta \in \mathbb{R}$ is said to be *the greatest lower bound of* Y, denoted $\beta = \inf Y$, if: (i) β is a lower bound of Y; and (ii) if γ is a lower bound of Y, then $\gamma \leq \beta$.

Theorem 1. *$\alpha = \sup Y$ if, and only if, for all $\varepsilon > 0$ it is true that (i) for all $y \in Y$, one has that $y < \alpha + \varepsilon$; and (ii) for some $y \in Y$ one has that $\alpha - \varepsilon < y$.*

The following axiom is crucial in the proof of the Bolzano-Weierstrass Theorem.

Axiom 1 (Axiom of Completeness). *Let $Y \subseteq \mathbb{R}$ be nonempty. If Y is bounded above, then it has a least upper bound.*

2 Maximizers

From now on, maintain the assumptions that set $D \subseteq \mathbb{R}^K$, for a finite K, is nonempty.

Theorem 2. *Let $Y \subseteq \mathbb{R}$ and let $b = \sup Y$. One has that $b \notin Y$ if, and only if, for all $a \in Y$, there is an $a' \in Y$ such that $a' > a$.*

Proof: Sufficiency is left as an exercise. For necessity, note that if there is an $a \in Y$ such that for all $a' \in Y$ it is true that $a' \leq a$, then, by definition, $b \leq a$, whereas $a \leq b$, which implies that $b = a \in Y$, a contradiction. *Q.E.D.*

It follows that we need a stronger concept of extremum, in particular one that implies that the extremum lies within the set. Thus, a point $b \in \mathbb{R}$ is said to be *the maximum of set* $Y \subseteq \mathbb{R}$, denoted $b = \max A$, if $b \in Y$ and for all $a \in Y$ it is true that $a \leq b$. The proofs of the following two results are left as exercises

Theorem 3. *If $\max Y$ exists, then it is unique.*

Theorem 4. *If $\max Y$ exists, then $\sup Y$ exists and $\sup Y = \max Y$. If $\sup Y$ exists and $\sup Y \in Y$, then $\max Y$ exists and $\max Y = \sup Y$.*

Exercise 1. *Given $Y, Y' \subseteq \mathbb{R}$, prove the following:*

1. *If $\sup Y$ and $\inf Y'$ exist, and for all $(a, a') \in Y \times Y'$, one has that $a \leq a'$, then $\sup Y \leq \inf Y'$.*

2. *If* $\sup Y$ *and* $\sup Y'$ *exist,* $\lambda, \lambda' \in \mathbb{R}_{++}$ *and*

$$\tilde{Y} = \{\tilde{a} | \exists (a, a') \in Y \times Y' : \lambda a + \lambda' a' = \tilde{a}\},$$

then $\sup \tilde{Y} = \sup Y + \sup Y'$.

3. *If* $\sup Y$ *and* $\sup Y'$ *exist, and for all* $a \in Y$ *there is an* $a' \in Y'$ *such that* $a \leq a'$*, then* $\sup Y \leq \sup Y'$.

Show also that a strict version of the third statement is not true.

Now, it typically is of more interest in economics to find extrema of functions, rather than extrema of sets. To a large extent, the distinction is only apparent: what we will be looking for are the extrema of the image of the domain under the function. A point $\bar{x} \in D$ is said to be *a global maximizer of* $f : D \to \mathbb{R}$ if for all $x \in D$ it is true that $f(x) \leq f(\bar{x})$. Point $\bar{x} \in D$ is said to be *a local maximizer of* $f : D \to \mathbb{R}$ if there exists some $\varepsilon > 0$ such that for every $x \in B_\varepsilon(\bar{x}) \cap D$ it is true that $f(x) \leq f(\bar{x})$.

When $\bar{x} \in D$ is a local (global) maximizer of $f : D \to \mathbb{R}$, the number $f(\bar{x})$ is said to be *a local (the global) maximum of* f. Notice that, in the latter case, $f(\bar{x}) = \max f[D]$, although more standard notation for $\max f[D]$ is $\max_D f$ or $\max_{x \in D} f(x)$.[1] Notice that there is a conceptual difference between maximum and maximizer! Also, notice that a function can have only one global maximum even if it has multiple global maximizers, but the same is not true for the local concept. The set of maximizers of a function is usually denoted by $\text{argmax}_D f$.

By analogy, $b \in \mathbb{R}$ is said to be *the supremum of* $f : D \to \mathbb{R}$, denoted $b = \sup_D f$ or $b = \sup_{x \in D} f(x)$, if $b = \sup f[D]$. Importantly, note that there is no reason why $\exists x \in D$ such that $f(x) = \sup_D f$ even if the supremum is defined.

3 Existence

Theorem 5 (Weierstrass). *Let* $C \subseteq D$ *be nonempty and compact. If the function* $f : D \to \mathbb{R}$ *is continuous, then there are* $\bar{x}, \underline{x} \in C$ *such that for all* $x \in C$ *it is true that* $f(\underline{x}) \leq f(x) \leq f(\bar{x})$.

Proof: Let $(y_n)_{n=1}^\infty$ be a sequence in $f[C]$ and such that $y_n \to y$. Fix $(x_n)_{n=1}^\infty$ in C such that $f(x_n) = y_n$. Since C is bounded, there exists a subsequence $(x_{n_m})_{m=1}^\infty$ that converges to some x, with $x \in C$ because C is closed. By continuity, $y = \lim_{m \to \infty} y_{n_m} = \lim_{m \to \infty} f(x_{n_m}) = f(x)$, so $y \in f[C]$.

Now, suppose that for all $\Delta \in \mathbb{R}$, there is $y \in f[C]$ such that $|y| \geq \Delta$. Then, for all $n \in \mathbb{N}$, there is $x_n \in C$ for which $|f(x_n)| \geq n$. Since C is compact, as before, there exists a subsequence $(x_{n_m})_{m=1}^\infty$ that converges to some $x \in C$. By continuity, $|f(x)| = \lim_{m \to \infty} |f(x_{n_m})| = \infty$, which is impossible.

It follows from the two arguments above that $f[C]$ is compact. Now, let $\bar{y} = \sup f[C]$. By Theorem 1, for all $n \in \mathbb{N}$ there is some $y_n \in f[C]$ for which $\bar{y} - 1/n < y_n < \bar{y}$. Clearly, $y_n \to \bar{y}$, so, since $f[C]$ is closed, $\bar{y} \in f[C]$, and it follows that there is $\bar{x} \in C$ such that $f(\bar{x}) = \bar{y}$. By definition, then, for every $x \in C$ it is true that $f(x) \leq \bar{y} = f(\bar{x})$.

[1] A point $\bar{x} \in D$ is said to be *a local minimizer of* $f : D \to \mathbb{R}$ if there is an $\varepsilon > 0$ such that for all $x \in B_\varepsilon(\bar{x}) \cap D$ it is true that $f(x) \geq f(\bar{x})$. Point $\bar{x} \in D$ is said to be *a global minimizer of* $f : D \to \mathbb{R}$ if for every $x \in D$ it is true that $f(x) \geq f(\bar{x})$. From now on, we only deal with maxima, although the minimization problem is obviously covered by analogy.

Existence of $\underline{x}$ is left as an exercise. *Q.E.D.*

The importance of this result is that when the domain of a continuous function is closed and bounded, then the function does attain its maxima and minima within its domain.

4 Characterizing maximizers

Even though maximization is not a differential problem, when one has differentiability there are results that make it easy to find maximizers. For this section, we take set D to be open.

4.1 Problems in $\mathbb{R}$

For simplicity, we first consider the case $K = 1$.

LEMMA 1. *Suppose that $f : D \to \mathbb{R}$ is differentiable. Let $\bar{x} \in X$. If $f'(\bar{x}) > 0$, then there is some $\delta > 0$ such that for every $x \in B_\delta(\bar{x}) \cap D$ we have $f(x) > f(\bar{x})$ if $x > \bar{x}$, and that $f(x) < f(\bar{x})$ if $x < \bar{x}$.*

Proof: By assumption, we have $f'(\bar{x}) \in \mathbb{R}_{++}$. Then, by definition, there is some $\delta > 0$ such that for any $x \in B'_\delta(\bar{x}) \cap D$,

$$|\frac{f(x) - f(\bar{x})}{x - \bar{x}} - f'(\bar{x})| < f'(\bar{x}),$$

and, since $f'(\bar{x}) > 0$, $(f(x) - f(\bar{x}))(x - \bar{x}) > 0$. *Q.E.D.*

The analogous result for the case of a negative derivative is presented, without proof, as the following corollary.

COROLLARY 1. *Suppose that $f : D \to \mathbb{R}$ is differentiable. Let $\bar{x} \in D$. If $f'(\bar{x}) < 0$, then there is some $\delta > 0$ such that for every $x \in B_\delta(\bar{x}) \cap D$ we have $f(x) < f(\bar{x})$ if $x > \bar{x}$, and that $f(x) > f(\bar{x})$ if $x < \bar{x}$.*

THEOREM 6. *Suppose that $f : D \to \mathbb{R}$ is differentiable. If $\bar{x} \in \operatorname{int}(D)$ is a local maximizer of f then $f'(\bar{x}) = 0$.*

Proof: Suppose not: $f'(\bar{x}) \neq 0$. If $f'(\bar{x}) > 0$, then, by Lemma 1, there is $\delta > 0$ such that for all $x \in B_\delta(\bar{x}) \cap D$ satisfying $x > \bar{x}$ we have that $f(x) > f(\bar{x})$. Since $\bar{x}$ is a local maximizer of f, then there is $\varepsilon > 0$ such that for all $x \in B_\varepsilon(\bar{x}) \cap D$ it is true that $f(x) \leq f(\bar{x})$. Since $\bar{x} \in \operatorname{int}(D)$, there is $\gamma > 0$ such that $B_\gamma(\bar{x}) \subseteq D$. Let $\beta = \min\{\varepsilon, \delta, \gamma\} > 0$. Clearly, $(\bar{x}, \bar{x}+\beta) \subset B'_\beta(\bar{x}) \neq \varnothing$ and $B'_\beta(\bar{x}) \subseteq D$. Moreover, $B'_\beta(\bar{x}) \subseteq B_\delta(\bar{x}) \cap D$ and $B'_\beta(\bar{x}) \subseteq B_\varepsilon(\bar{x}) \cap D$. This implies that for some x one has $f(x) > f(\bar{x})$ and $f(x) \leq f(\bar{x})$, an obvious contradiction. A similar contradiction appears if $f'(\bar{x}) < 0$, by Corollary 1. *Q.E.D.*

THEOREM 7. *Let $f : D \to \mathbb{R}$ be of class $\mathbf{C}^2$. If $\bar{x} \in \operatorname{int}(D)$ is a local maximizer of f then $f''(\bar{x}) \leq 0$.*

Proof: Since $\bar{x} \in \operatorname{int}(D)$, there is a $\varepsilon > 0$ for which $B_\varepsilon(\bar{x}) \subseteq D$. For every $h \in B_\varepsilon(0)$, since f is twice differentiable, by Taylor's Theorem, there is some x_h^* in the interval joining $\bar{x}$ and $\bar{x} + h$, such that

$$f(\bar{x} + h) = f(\bar{x}) + f'(\bar{x})h + \frac{1}{2} f''(x_h^*)h^2$$

Since $\bar{x}$ is a local maximizer, there is a $\delta > 0$ such that $x \in B_\delta(\bar{x}) \cap D$ implies $f(x) \leq f(\bar{x})$. Let $\beta = \min\{\varepsilon, \delta\} > 0$. By construction, for any $h \in B'_\beta(0)$ one has that

$$f'(\bar{x})h + \frac{1}{2}f''(x_h^*)h^2 = f(\bar{x}+h) - f(\bar{x}) \leq 0.$$

By Theorem 6, since f is differentiable and $\bar{x}$ is a local maximizer, $f'(\bar{x}) = 0$, from where $h \in B'_\beta(0)$ implies that $f''(x_h^*)h^2 \leq 0$, and hence that $f''(x_h^*) \leq 0$. Now, letting $h \to 0$, we get that $\lim_{h\to 0} f''(x_h^*) \leq 0$, and hence that $f''(\bar{x}) \leq 0$, since f'' is continuous and each x_h lies in the interval joining $\bar{x}$ and $\bar{x}+h$. *Q.E.D.*

Notice that the last theorems only give us necessary conditions:[2] this is not a tool that tells us which points are local maximizers, but it tells us what points are not. A complete characterization requires both necessary and sufficient conditions. We now find sufficient conditions.

THEOREM 8. *Suppose that $f : D \to \mathbb{R}$ is twice differentiable. Let $\bar{x} \in \mathrm{int}(D)$. If $f'(\bar{x}) = 0$ and $f''(\bar{x}) < 0$, then $\bar{x}$ is a local maximizer.*

Proof: Since $f : D \to \mathbb{R}$ is twice differentiable and $f''(\bar{x}) < 0$, we have, by Corollary 1, that for some $\delta > 0$ it is true that whenever $x \in B_\delta(\bar{x}) \cap D$ we have that (i)$f'(x) < f'(\bar{x}) = 0$, when $x > \bar{x}$; and (ii)$f'(x) > f'(\bar{x}) = 0$, when $x < \bar{x}$. Since $x \in \mathrm{int}(D)$, there is an $\varepsilon > 0$ such that $B_\varepsilon(\bar{x}) \subseteq D$. Let $\beta = \min\{\delta, \varepsilon\} > 0$. By the Mean Value Theorem, we have that for all $x \in B_\beta(\bar{x})$,

$$f(x) = f(\bar{x}) + f'(x^*)(x - \bar{x})$$

for some x^* in the interval between $\bar{x}$ and x (why?). Thus, if $x > \bar{x}$, we have $x^* \geq \bar{x}$, and, therefore, $f'(x^*) \leq 0$, so that $f(x) \geq f(\bar{x})$. On the other hand, if $x < \bar{x}$, then $f'(x^*) \geq 0$, so that $f(x) \leq f(\bar{x})$. *Q.E.D.*

Notice that the sufficient conditions are stronger than the set of necessary conditions: there is a little gap that the differential method does not cover.

4.2 Higher-dimensional problems

We now allow for functions defined on higher-dimesional domains (namely $K \geq 2$). The results of the one-dimensional case generalize as follows.

THEOREM 9. *If $f : D \to \mathbb{R}$ is differentiable and $x^* \in D$ is a local maximizer of f, then $Df(x^*) = 0$.*

THEOREM 10. *If $f : D \to \mathbb{R}$ is of class $\mathbf{C}^2$ and $x^* \in D$ is a local maximizer of f, then $D^2f(x^*)$ is negative semidefinite.*

As before, these conditions do not tell us which points are maximizers, but only which ones are not. Before we can argue sufficiency, we need to introduce the following lemma.

THEOREM 11. *Suppose that $f : D \to \mathbb{R}$ is of class $\mathbf{C}^2$ and let $\bar{x} \in D$. If $Df(\bar{x}) = 0$ and $D^2f(\bar{x})$ is negative definite, then $\bar{x}$ is a local maximizer.*

[2] And there are further necessary conditions.

5 Maxima and concavity

For this section, we take $D \subseteq \mathbb{R}^K$, $K \in \mathbb{N}$, $D \neq \varnothing$ and drop the openness assumption.

Note that the results that we obtained in the previous sections hold only locally. We now study the extent to which local extrema are, in effect, global extrema.

THEOREM 12. *Suppose that D is a convex set and $f : D \to \mathbb{R}$ is a concave function. Then, if $\bar{x} \in D$ is a local maximizer of f, it is also a global maximizer.*

Proof: We argue by contradiction: suppose that $\bar{x} \in D$ is a local maximizer of f, but it is not a global maximizer. Then, there is $\varepsilon > 0$ such that for every $x \in B_\varepsilon(\bar{x}) \cap D$, $f(x)(\bar{x})$; and there is $x^* \in D$ such that $f(x^*) > f(\bar{x})$. Clearly, then, $x^* \notin B_\varepsilon(\bar{x})$, which implies that $\|x^* - \bar{x}\| \geq \varepsilon$. Now, since D is convex and f is concave, we have that for $\theta \in [0,1]$,

$$f(\theta x^* + (1-\theta)\bar{x}) \geq \theta f(x^*) + (1-\theta) f(\bar{x}),$$

but, since $f(x^*) > f(\bar{x})$, we further have that if $\theta \in (0,1]$, then $\theta f(x^*) + (1-\theta) f(\bar{x}) > f(\bar{x})$, so that $f(\theta x^* + (1-\theta)\bar{x}) > f(\bar{x})$.

Now, consider $\theta^* \in (0, \varepsilon/\|x^* - \bar{x}\|)$. Clearly, $\theta^* \in (0,1)$, so $f(\theta^* x^* + (1-\theta^*)\bar{x}) > f(\bar{x})$. However, by construction,

$$\|(\theta^* x^* + (1-\theta^*)\bar{x}) - \bar{x}\| = \theta^* \|x^* - \bar{x}\| < (\frac{\varepsilon}{\|x^* - \bar{x}\|})\|x^* - \bar{x}\| = \varepsilon,$$

which implies that $(\theta^* x^* + (1-\theta^*)\bar{x}) \in B_\varepsilon(\bar{x})$, and, moreover, by convexity of D, we have that $(\theta^* x^* + (1-\theta^*)\bar{x}) \in B_\varepsilon(\bar{x}) \cap D$. This contradicts the fact that $f(x) \leq f(\bar{x})$ for all $x \in B_\varepsilon(\bar{x}) \cap D$. *Q.E.D.*

THEOREM 13. *Suppose that D is convex, $f : D \to \mathbb{R}$ is of class $\mathbf{C}^2$ and for each $x \in D$, $Df''(x)$ is negative definite. Then, there exists at most one point $\bar{x} \in D$ such that $Df(\bar{x}) = 0$. If such point exists, it is a global maximizer.*

Proof: We first prove the last part of the theorem. Suppose that there is $\bar{x} \in D$ such that $Df(\bar{x}) = 0$. By assumption, $Df''(\bar{x})$ is negative definite, and therefore, by Theorem 11, $\bar{x}$ is a local maximizer. Since $Df(x)$ is negative definite everywhere, we have that f is concave and, therefore, by Theorem 12, $\bar{x}$ is a global maximizer.

We must now show that there cannot exist more than one such point. We argue by contradiction: suppose that there are distinct $\bar{x}_1, \bar{x}_2 \in D$ such that $f'(\bar{x}_1) = f'(\bar{x}_2) = 0$. By our previous argument, both $\bar{x}_1$ and $\bar{x}_2$ are global maximizers, so that $f(\bar{x}_1) = f(\bar{x}_2)$. Now, since $D^2 f(x)$ is negative definite everywhere, we have that f is strictly concave, and

$$f(\frac{1}{2}\bar{x}_1 + \frac{1}{2}\bar{x}_2) > \frac{1}{2} f(\bar{x}_1) + \frac{1}{2} f(\bar{x}_2) = f'(\bar{x}_1) = f'(\bar{x}_2),$$

contradicting the fact that both $\bar{x}_1$ and $\bar{x}_2$ are global maximizers, since D is convex. *Q.E.D.*

For the sake of practice, it is a good idea to work out an exercise like Exercise 3.9 in pages 57 and 58 of Simon and Blume.

6 Parametric Programming

For the purposes of this section, fix a non-empty set $\Omega \subset \mathbb{R}^M$, for a positive integer M.

6.1 Continuity

Theorem 14 (Theorem of the Maximum). *Let $f : D \times \Omega \to \mathbb{R}$ be continuous. The* value function $v : \Omega \to \mathbb{R}$*, defined by*

$$v(\omega) = \max_{x \in D} f(x, \omega),$$

is continuous.

6.2 Differentiability

Suppose furthermore that sets D and Ω are both open. Suppose that $f : D \times \Omega \to \mathbb{R}$ is of class $\mathbb{C}^2$ and the partial Hessian with respect to x, matrix $D^2_{x,x} f(x, \omega)$ is negative definite at all $(x, \omega) \in D \times \Omega$. Suppose, moreover that function $x^* : \Omega \to D$ captures the maximizers of f given omega: for each $\omega \in \Omega$, point $x^*(\omega)$ is the unique solution to problem

$$\max_{x \in D} f(x, \omega).$$

Theorem 15 (The Envelope Theorem). *Under the assumptions of this subsection, functions x^* and v are continuously differentiable, and $Dv(\bar{\omega}) = D_\omega f(x^*(\bar{\omega}), \bar{\omega})$.*