CHAPTER 1

Static optimization

In this chapter we deal with problems involving the choice of values for a finite number of variables in order to maximize some objective. Sometimes the values the variables may take are unrestricted; at other times they are restricted by equality constraints and also by inequality constraints. In the course of the presentation an important class of functions will emerge; they are called concave functions and are closely associated with “nice” maximum problems. They will be encountered throughout this book. For this reason we weave the concept of concavity of functions through the exposition of maximization problems. This is done to suit our purposes, but concave functions have other important properties in their own right.

The notation we use is fairly standard. If in doubt, the reader should refer to the appendix to this chapter, which also contains a reminder of the basic notions of multivariate calculus and some matrix algebra needed to follow the exposition.

1.1 Unconstrained optimization, concave and convex functions

In what follows we assume all functions to have continuous second-order derivatives, unless otherwise stated. Strictly speaking, all domains of definitions should be open subsets of the multidimensional real space so that no boundary problems arise.

1.1.1 Unconstrained maximization

Consider the problem of finding a set of values $x_1, x_2, \ldots, x_n$ to maximize the function $f(x_1, \ldots, x_n)$. We often write this as

$$\operatorname{Maximize} f(x),$$

(1.1)

where $x$ is understood to be an $n$-dimensional vector. We refer to the problem of (1.1) as an unconstrained maximum because no restrictions are placed on $x$.

Necessary conditions. Suppose we find a solution to this problem and denote the optimal vector by $x^*$. Consider an arbitrarily small deviation
2.1 Static optimization

from \( \mathbf{x}^* \), say \( d\mathbf{x} \). If we have a maximum at \( \mathbf{x}^* \), then \( f \) must not increase for any \( d\mathbf{x} \).

The change in \( f \) is approximated by

\[
df = \sum_i f_{x_i}(\mathbf{x}^*)
\]

Clearly, \( df \leq 0 \) if we have a maximum at \( \mathbf{x}^* \). Furthermore, suppose we found some \( d\mathbf{x} \) vector such that \( df < 0 \); then by using the deviation \((-d\mathbf{x})\) we would obtain an increase in \( f \). Therefore, it must be that for any \( d\mathbf{x} \) vector, \( df \) is equal to zero. The only way this can be achieved for arbitrary deviations is to require each derivative \( f_{x_i}(\mathbf{x}^*) \) to vanish. Formally,

\[
f(\mathbf{x}) \text{ reaches a maximum at } \mathbf{x}^* \text{ implies } f_{x_i}(\mathbf{x}^*) = 0, \quad i = 1, \ldots, n. \quad (1.2)
\]

This is called the first-order condition. Several remarks must now be made. First, the above reasoning, hence (1.2), also applies to minimization problems. Second, we have been lax in defining a maximum. We should have distinguished a global maximum from a local maximum. We say that \( f(\mathbf{x}) \) reaches a global maximum at \( \mathbf{x}^* \) if \( f(\mathbf{x}^*) \geq f(\mathbf{x}) \) for all \( \mathbf{x} \) on its domain of definition (assumed to be an open set). We say \( f(\mathbf{x}) \) reaches a local maximum at \( \mathbf{x}^* \) if \( f(\mathbf{x}^*) \geq f(\mathbf{x}) \) for all \( \mathbf{x} \) “close” to \( \mathbf{x}^* \) (i.e., for all \( \mathbf{x} \) within \( \delta \) units of distance from \( \mathbf{x}^* \), where \( \delta \) is some positive number). The local maximum is a much weaker concept than the global one. However, because our argument relies on arbitrarily small deviations from \( \mathbf{x}^* \), it applies to both cases. The first-order condition (1.2) follows from the existence of a maximum; hence, it is a necessary condition for a maximum, but it is not the only one, as we now show. As we noted previously, condition (1.2) is necessary for a local minimum as well. The following condition, called the second-order necessary condition, takes a different form for a maximum than for a minimum.

To establish it we must take a Taylor’s expansion (with remainder) of the function \( f \) about the point \( \mathbf{x}^* \):

\[
f(\mathbf{x}^* + d\mathbf{x}) = f(\mathbf{x}^*) + \sum_{i=1}^{n} f_{x_i}(\mathbf{x}^*)(d\mathbf{x}_i)
\]

\[+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{x_ix_j}(\mathbf{x}^*)(d\mathbf{x}_i)(d\mathbf{x}_j) + \cdots + R, \quad (1.3a)
\]

or in vector notation (see the Appendix for details),

\[
f(\mathbf{x}^* + d\mathbf{x}) = f(\mathbf{x}^*) + (d\mathbf{x})' \cdot f_x(\mathbf{x}^*) + \frac{1}{2} (d\mathbf{x})' \cdot f_{xx}(\mathbf{x}^*) \cdot (d\mathbf{x}) + \cdots + R,
\]

(1.3b)

where \( d\mathbf{x} \) is small enough (i.e., \( |d\mathbf{x}| < \delta \)) that higher-order terms vanish relative to second-order terms.
1.1 Unconstrained optimization

Suppose again that we have a (at least local) maximum, that is, \( f(x^*) \geq f(x^* + dx), \forall dx, \|dx\| < \delta \). Then \( f_x(x^*) = 0 \), and neglecting terms higher than the second order we have

\[
f(x^* + dx) - f(x^*) = \frac{1}{2}(dx)' \cdot f_{xx}(x^*) \cdot dx \\
\leq 0, \quad \text{because } x^* \text{ is a maximum.}
\]

Since \((dx)' \cdot f_{xx}(x^*) \cdot (dx)\) is negative or zero for all small deviation vectors \(dx\), the Hessian matrix of \(f\) evaluated at \(x^*\) must be negative-semidefinite. This is the second-order necessary condition:

\[f(x)\] reaches a maximum at \(x^*\) implies \(f_{xx}(x^*)\) is negative-semidefinite.

(1.4)

Again, (1.4) applies to global as well as local maxima.

Sufficient conditions (for a local maximum). It is unfortunately not possible to state conditions that are both necessary and sufficient for a function to reach a maximum. We can, however, easily provide sufficient conditions:

If \(f_x(x^*) = 0, i = 1, \ldots, n,\) and \(f_{xx}(x^*)\) is negative-definite,

then \(f(x)\) reaches a local maximum at \(x^*\).

(1.5)

To prove this we shall consider again Taylor’s expansion in (1.3) and let \(dx \to 0\), so that the second-degree term dominates those of higher order while the first-degree term vanishes; we obtain \(f(x^* + dx) < f(x^*)\), thus establishing \(x^*\) as a local maximum.

1.1.2 Global results and concave functions

When we seek a maximum in an economic problem, it is most often a global one. Indeed, it is little comfort to know that we are doing the best we can but only if considering policies which differ minutely from the current one (local optimum). It is also clear that we will not be able to characterize a global maximum with conditions on the values of the function and its derivatives at the maximum itself; we will need to place restrictions on the overall shape of the function, restrictions that apply everywhere on the domain of definition, which we denote by \(X\).

Consider the exact form of Taylor’s expansion to the second degree: there exists a point \(x_t\) on the line segment between \(x\) and \(\bar{x}\) such that

\[
f(x) = f(\bar{x}) + (x - \bar{x})' \cdot f_x(\bar{x}) + \frac{1}{2}(x - \bar{x})' \cdot H(x_t) \cdot (x - \bar{x}),
\]

(1.6)

where \(H(x_t)\) denotes the Hessian matrix of \(f\), evaluated at the point \(x_t\). If we were to restrict our attention to functions with a negative-semidefinite
4  Static optimization

matrix everywhere on its domain of definition, then the last term of (1.6) would be guaranteed to be nonpositive for any \( x \), and the requirement that \( x \) be a global maximum (i.e., \( f(x) - f(x) \leq 0 \ \forall x \in X \)) would be equivalent to the first-order condition \( f'(x) = 0 \). We now formalize this argument.

**Definition 1.1.1.** A function with continuous second-order derivatives defined on a convex set \( X \) is concave if and only if its Hessian matrix is negative-semidefinite everywhere on its domain of definition \( X \).

**Theorem 1.1.1.** Let \( f(x) \) be a concave function; then it reaches a global maximum at \( x \) if and only if \( f'(x) = 0 \).

Definition 1.1.1 applies only to functions with continuous second-order derivatives. It is useful to have a more general definition of concavity that does not require this assumption.

**Definition 1.1.2.** A function \( f(x) \) with continuous first-order derivatives defined on a convex set \( X \) is concave if and only if

\[
 f(x_2) - f(x_1) \leq (x_2 - x_1)' \cdot f'(x_1),
\]

for all \( x_1, x_2 \) on \( X \).

Note that Definition 1.1.2 is less stringent than Definition 1.1.1 in terms of differentiability restrictions, since it requires continuity only for the first derivatives; this is the only difference between the two definitions. Indeed, if we assume that the function has continuous second-order derivatives, we can see that the two definitions are equivalent simply by writing down the exact form of Taylor’s expansion. Given two arbitrary points \( x_1 \) and \( x_2 \), there exists a point \( x \) between them such that

\[
 f(x_2) = f(x_1) + (x_2 - x_1)' \cdot f'(x_1) + \frac{1}{2} (x_2 - x_1)' \cdot H(x_1) \cdot (x_2 - x_1),
\]

\[
 f(x_2) - f(x_1) - (x_2 - x_1)' \cdot f'(x_1) = \frac{1}{2} (x_2 - x_1)' \cdot H(x_1) \cdot (x_2 - x_1) \leq 0.
\]

The geometric interpretation is simply that a tangent plane to the graph of \( f(x) \) must remain everywhere above the graph, the equation for the tangent plane at \( x_1 \) being

\[
 y = f(x_1) + (x - x_1)' \cdot f'(x_1).
\]

This is illustrated in Figure 1.1a for functions of one variable. Definition 1.1.2 does not cover functions that have “kinks” and as such are not differentiable everywhere. To admit this case, a more general definition is needed.
1.1 Unconstrained optimization

Figure 1.1

Definition 1.1.3. A function $f(x)$ defined on a convex set $X$ is concave if and only if

$$f(x_t) \geq tf(x_1) + (1-t)f(x_2), \quad 0 \leq t \leq 1,$$

all $x_1, x_2$ in $X$,

where $x_t = tx_1 + (1-t)x_2$.

If a function satisfies Definition 1.1.2, it also satisfies Definition 1.1.3. To see this we state Definition 1.1.2 in two instances:
1 Static optimization

\[ f(x_2) - f(x_1) \leq (x_2 - x_1)' \cdot f_s(x_1) \]

and

\[ f(x_1) - f(x_t) \leq (x_1 - x_t)' \cdot f_s(x_t), \]

where

\[ x_t = tx_1 + (1-t)x_2 \text{ for some } t, \quad 0 \leq t \leq 1. \]

Since

\[ x_2 - x_t = t(x_2 - x_1) \quad \text{and} \quad x_1 - x_t = -(1-t)(x_2 - x_1), \]

we have

\[ f(x_2) - f(x_t) \leq t(x_2 - x_1)' \cdot f_s(x_t), \]

\[ f(x_1) - f(x_t) \leq -(1-t)(x_2 - x_1)' \cdot f_s(x_t). \]

Multiplying the first inequality by \((1-t)\), the second by \(t\), and adding yields (with \(0 \leq t \leq 1\))

\[ t f(x_1) + (1-t) f(x_2) - f(x_t) \leq 0, \]

which was to be proved.

Note that no differentiability properties are required in Definition 1.1.3. The geometric interpretation of this definition is that a line (or chord) joining two points of the graph always lies below the graph, since the left-hand side of the inequality represents the value of \(f\) at a convex combination of \(x_1\) and \(x_2\) and the right-hand side is the same convex combination of the values of the function at \(x_1\) and \(x_2\) – hence the height of the point on the chord above \(x_t\). This is illustrated in Figure 1.1b for functions of one variable.

Concave functions have many notable properties; Theorem 1.1.2 lists some of the most useful ones.

**Theorem 1.1.2**

(i) Let \(f(x)\) be a concave function and \(k \geq 0\) a constant; then \(kf(x)\) is a concave function.

(ii) Let \(f(x)\) and \(g(x)\) be concave functions; then \(f(x) + g(x)\) is itself a concave function.

(iii) Let \(f(x)\) be a concave function; then the upper contour set defined by \(B(\bar{x}) = \{x \in \mathbb{R}_n \mid f(x) \geq f(\bar{x})\}\) is a convex set.

(The converse of (iii) is not true!)

The proofs of these results are straightforward; for instance, (iii) requires that we show that if \(f(x_1) \geq f(\bar{x})\) and \(f(x_2) \geq f(\bar{x})\), it follows that \(f(x_t) \geq f(\bar{x})\); this is obvious from Definition 1.1.3.
1.1 Unconstrained optimization

Strictly concave functions: unique global maximum. While concave functions have the property that a solution of the first-order condition yields a global maximum, this does not ensure the uniqueness of that solution: a concave function may reach its global maximum at several points. For example, the following function is concave, but the first-order condition admits as a solution any point between 1 and 2; thus, the function reaches a global maximum at any $x^*$ such that $1 \leq x^* \leq 2$.

$$f(x) = \begin{cases} 
  x - 0.5x^2, & x < 1, \\
  0.5, & 1 \leq x \leq 2, \\
  (x-1) - 0.5(x-1)^2, & 2 < x.
\end{cases}$$

Other examples will be encountered in Section 1.1.5.

It is sometimes desirable to place more restrictions on the function so that if a maximum exists, it is the unique global maximum. We use this as a means of introducing a subclass of concave functions called strictly concave functions. Definitions 1.1.2 and 1.1.3 are adapted by simply requiring strict inequalities.

**Definition 1.1.3′.** A function $f(x)$ defined on a convex set $X$ is strictly concave if and only if

$$f(tx_1) > tf(x_1) + (1-t)f(x_2), \quad 0 < t < 1,$$

for all $x_1, x_2$ in $X$, where $x_1 \neq x_2$ and $x_t = tx_1 + (1-t)x_2$.

**Definition 1.1.2′.** A function $f(x)$ with continuous first-order derivatives defined on a convex set $X$ is strictly concave if and only if

$$f(x_2) - f(x_1) < (x_2 - x_1)^T f'(x_1)$$

for all $x_1$ and $x_2$ in $X$, where $x_1 \neq x_2$.

It is obvious from Definition 1.1.2′ that $f'_x(x_1) = 0$ is necessary and sufficient for $x_1$ to be the unique global maximum of that function $f$.

We cannot claim that functions with continuous second-order derivatives are strictly concave if and only if their Hessian matrix is negative-definite, because some strictly concave functions have a Hessian matrix which becomes negative-semidefinite at some points. One instance is $f(x_1, x_2) = -(x_1)^4 - (x_2)^2$, which is negative-definite everywhere but at $x_1 = 0$, when it is negative-semidefinite. We must be content with the following theorem.

**Theorem 1.1.3.** A function that is defined on a convex set $X$ and has a negative-definite Hessian matrix everywhere on $X$ is strictly concave.

The reader is invited to prove this result using Definition 1.1.2′.
8 1 Static optimization

1.1.3 Unconstrained minimization and convex functions

Results for minimization problems are just mirror images of those for maximization problems and are obtained by replacing $f(x)$ by $-f(x)$. Thus, the first-order necessary condition for a local minimum at $x^*$ is

$$ f_i(x^*) = 0, \quad i = 1, \ldots, n, $$

(1.7)

and the second-order necessary condition is

$$ f_{xx}(x^*) \text{ is positive-semidefinite}. $$

(1.8)

The sufficient conditions for a local minimum at $x^*$ are

$$ f_x(x^*) = 0 \quad \text{and} \quad f_{xx}(x^*) \text{ is positive-definite}. $$

(1.9)

Similarly, we have to define convex functions in order to obtain global results on minimization. Corresponding to Definitions 1.1.1, 1.1.2, and 1.1.3 we now have the following (results on strictly convex functions are indicated in parentheses).

**Definition 1.1.4.** A function with continuous second-order derivatives defined on a convex set is (strictly) convex if and only if its Hessian matrix is positive-semidefinite (if its Hessian matrix is positive-definite).

**Definition 1.1.5.** A function $f(x)$ with continuous first-order derivatives defined on a convex set $X$ is (strictly) convex if and only if

$$ f(x_2) - f(x_1) \geq (x_2 - x_1)' f_x(x_1) \quad \text{for all } x_1, x_2 \text{ in } X, $$

$$(f(x_2) - f(x_1)) > (x_2 - x_1)' f_x(x_1) \quad \text{for all } x_1, x_2 \text{ in } X, \text{ where } x_1 \neq x_2. $$

**Definition 1.1.6.** A function $f(x)$ defined on a convex set $X$ is (strictly) convex if and only if

$$ f(x_1) \leq tf(x_1) + (1-t)f(x_2), \quad 0 \leq t \leq 1, \text{ all } x_1, x_2 \text{ in } X, $$

$$(f(x_1) < tf(x_1) + (1-t)f(x_2), \quad 0 < t < 1, \text{ all } x_1, x_2 \text{ in } X, \text{ } x_1 \neq x_2. $$

**Theorem 1.1.4**

(i) Let $f(x)$ be a convex function and $k \geq 0$ a constant; then $kf(x)$ is a convex function.

(ii) Let $f(x)$ and $g(x)$ be convex functions; then $f(x) + g(x)$ is itself a convex function.

(iii) Let $f(x)$ be a convex function; then the lower contour set defined by $W(\bar{x}) = \{ x \in R^n | f(x) \leq f(\bar{x}) \}$ is a convex set. (The converse of (iii) is not true!)

(iv) Let $f(x)$ be a (strictly) convex function; then $-f(x)$ is a (strictly) concave function.
1.1 Unconstrained optimization

![Graph](image)

Figure 1.2

(v) A linear function is both convex and concave but not strictly either.

Definitions 1.1.5 and 1.1.6 are illustrated in Figure 1.2 for convex functions of one variable.

1.1.4 Geometric representation

Figures 1.3a and 1.3b represent the graphs of a concave and a convex function, respectively. It is important to realize that a concave function
Figure 1.3

need not have a maximum, nor a convex function a minimum. If they
do, then one is somewhat dome-shaped and the other bowl-shaped. It is
then obvious that a rod connecting two points of the dome remains under
it (Definition 1.1.3), while such a rod connecting two points of the bowl
remains above its walls (Definition 1.1.6). It is clearly inconvenient to rely
on three-dimensional diagrams; instead, we most often use level curves.
We know that if a function is concave, its upper contour sets are convex
sets. We use this information in Figure 1.4a to draw some level curves of
a concave function, where the arrows indicate directions of increase of
the function and one convex upper contour set is hatched. We can also
verify that Definition 1.1.3 is satisfied: the function takes on the value $c$
at points $A$ and $B$; thus, it takes on a higher value at a point between