Chapter 13 Optimization

Two Variables

We can easily extend the problem of optimization of a function with one variable to optimization of a function with two variables. A differential function $z = f(x, y)$ can have a maximum or a minimum at an interior point (x_0, y_0) in a set S only if it is a stationary point. In other words, it is a necessary condition for an interior extreme point to satisfy the *first order conditions (FOCs)* to be an optimum:

$f_1'(x,y) = 0 \quad and \quad f_2'(x,y) = 0$

There are also conditions for a potential maximum or a minimum that are sufficient, but not necessary:

- If for all (x, y) in S the following inequalities hold, then function is concave and the point (x_0, y_0) is a maximum point of the function:
  - $f_{11}''(x,y) \leq 0$
  - $f_{22}''(x,y) \leq 0$
  - $f_{11}''(x,y)f_{22}''(x,y) - \left(f_{12}''(x,y)\right)^2 \geq 0$
- If for all (x, y) in S the following inequalities hold, then the function is convex and the point (x_0, y_0) is a minimum point of the function:
  - $f_{11}''(x,y) \geq 0$
  - $f_{22}''(x,y) \geq 0$
  - $f_{11}''(x,y)f_{22}''(x,y) - \left(f_{12}''(x,y)\right)^2 \geq 0$

Local Optima

At any local extreme point in the interior of a domain of a differentiable function, the function must be stationary, which means that all its first-order conditions need to be equal to zero. A point is said to be a local maximum point when $f(x,y) \leq f(x_0, y_0)$ for all pairs of x and y that lie sufficiently close to that particular point (x_0, y_0). For a local minimum the reasoning is the other way around. Note that a global maximum or a global minimum are also a local maximum or local minimum, but the other way around is not true.

A so-called *saddle point* is a stationary point with the property that there exist points (x, y) arbitrarily close to (x_0, y_0) with $f(x, y) < f(x_0, y_0)$, but also points with $f(x,y) > f(x_0, y_0)$. Therefore a stationary point, when the FOCs are equal to zero, can be a local maximum, a local minimum and a saddle point. To determine the nature of such a stationary point we need the *second-order conditions.* Again, these conditions are sufficient but not necessary.

- (x_0,y_0) is a strict local maximum if
  - $f''_{11}(x_0,y_0) < 0$
  - $f''_{11}(x_0,y_0)f''_{22}(x_0,y_0) - \left(f''_{12}(x_0,y_0)\right)^2 > 0$
- (x_0,y_0) is a local minimum if
  - $f''_{11}(x_0,y_0) > 0$
  - $f''_{11}(x_0,y_0)f''_{22}(x_0,y_0) - \left(f''_{12}(x_0,y_0)\right)^2 > 0$
- (x_0,y_0) is a saddle point if
  - $f''_{11}(x_0,y_0)f''_{22}(x_0,y_0) - \left(f''_{12}(x_0,y_0)\right)^2 < 0$
- (x_0,y_0) could be any of the three options if
  - $f''_{11}(x_0,y_0)f''_{22}(x_0,y_0) - \left(f''_{12}(x_0,y_0)\right)^2 = 0$

The Extreme Value Theorem

Consider a function $f(x,y)$ that is continuous throughout a nonempty, closed and bounded set. Then there exist both a point (a,b) in the set where the function has a minimum and a point (c,d) in the set where the function has a maximum. In a more formal way:

$$f(a,b) \leq f(x,y) \leq f(c,d)$$

A point (a,b) is called an interior point of a set in the plane if there exists a circle centered at this point such that all points inside this circle are still part of the set. This circle can be extremely small. An *open* set consists only of such interior points. A set is called a *closed* set when it contains all its *boundary points*. A point is a boundary point if every circle (every size) contains points of the set as well as points of its complement. In other words, a set is closed if and only if its complement is open. A set can also be neither open nor closed, when a set contains some boundary points but not all. Illustrations you can find on page 483.

A set is *bounded* if the whole set is contained within a sufficiently large circle. Otherwise a set is unbounded. A set that is both closed and bounded is called a *compact* set.

The following steps need to be taken to find the maxima and minima of a differentiable function defined on a compact set:

1. First find the stationary points of the function $f(x,y)$ by finding the FOCs.
2. Then find the largest and smallest value of the function on the boundary of the set.
3. Compute the values of the function at all the points found in 1 and 2. The largest value is the maximum and the smallest value is the minimum.

Three or More Variables

The optimization problem can be generalized to problems of three or more variables. The first order conditions are formulated as $f_i'(\boldsymbol{x}) = 0$. The extreme value theorem is also valid if the function is continuous throughout a nonempty and compact set: $f(\boldsymbol{d}) \leq f(\boldsymbol{x}) \leq f(\boldsymbol{c})$ for all **x** in the set.

An important and useful result is that maximizing a function is equivalent to maximizing a strictly increasing transformation of that same function. Formally speaking, if $g(\boldsymbol{x}) = F(f(\boldsymbol{x}))$, then

- If F is increasing and $\boldsymbol{c}$ maximizes (or minimizes) the function f over the set, then $\boldsymbol{c}$ also maximizes (or minimizes) g over the set.
- If F is strictly increasing, then $\boldsymbol{c}$ maximizes (or minimizes) f over the set if and only if $\boldsymbol{c}$ maximizes (or minimizes) g over the set.

For example, differentiating $f(x,y) = e^{x^2+2xy^2-y^3}$ has exactly the same solutions as differentiating $g(x,y) = x^2 + 2xy^2 - y^3$.

Envelope theorem

In economics many problems depend on changing parameters that are held constant in the process of optimization, such as prices, tax rates, income levels etc. We may want to know however how optimal values change when these parameters change.

Consider the optimization problem of a function $\max_x f(x,r)$ that we can rewrite as $f^*(r) = f(x^*(r),r)$ because the optimum value of x, x*, depends on parameter r. This last function is called the *value function*. Differentiation to r gives us the following equation:

$$\frac{df^*(r)}{dr} = f_1'(x^*(r),r)\frac{dx^*(r)}{dr} + f_2'(x^*(r),r)$$
$$\frac{df^*(r)}{dr} = f_2'(x^*(r),r)$$

The generalization of this result is called the envelope theorem. If $f^*(\boldsymbol{r}) = \max_{\boldsymbol{x}} f(\boldsymbol{x},\boldsymbol{r})$ and if $\boldsymbol{x}^*(\boldsymbol{r})$ is the value of **x** that maximizes the function $f(\boldsymbol{x},\boldsymbol{r})$, then:

$$\frac{\partial f^*(\boldsymbol{r})}{\partial r_j} = \left[\frac{\partial f(\boldsymbol{x},\boldsymbol{r})}{\partial r_j}\right]_{\boldsymbol{x}=\boldsymbol{x}^*(\boldsymbol{r})}$$