11. Single Variable Optimization

The objective here is to develop methods for finding points at which a function achieves its maximum value or its minimum value. Many problems in economics take this form: utility maximization, profit maximization, etc. The following example provides an outline of the kind of problem to be solved.

<u>Example 11.1</u> A firm has a production function $F(K)$ which shows how much of its product is produced at various levels of capital K. The firm is a price taker and faces per unit price p for its product, and its cost function is $C(K)$.

The firm's profit function is $\pi(K) = pF(K) - C(K)$. It chooses levels K^* of capital to maximize its profits. We call the firm's profit function the **objective function**. The variable K is **choice variable**. The "variable" p is a **parameter** of the problem. For this maximization problem it is taken to be a fixed value.

Qn: What is the firm's **optimal choice** of K ?

Qn: What is the firm's profit at the optimal choice of K ?

Let the optimal choice of K be denoted K^*. This depends on the parameter p. For a given fixed value of p, K^* is constant. However, the usefulness of a theory often lies in questions such as

Qn: How do the firm's optimal choice change as p changes?

The firm's profit level at the optimum is $\pi^* = \pi(K^*)$. This also depends on p.

Qn: How does a profit maximizing firm's profit change as p changes?

The objective here is to help you acquire the tools for answering such questions.

We begin with the simplest case, where the objective function is a function of one variable only, and where in many cases there are no parameters, so that the solutions are numerical constants, as in the example that follows.

Example 11.2 What is the minimum value of $f(x) = e^x - x$, and at what value of x does it occur? (answer: min. value of $f(x)$ is 1, and is obtained when $x = 0$).

We will use simple examples like this to help us understand the definitions and concepts involved in maximization/minimization problems. We then proceed to more sophisticated problems, first including parameters (this chapter) then extending to multivariate contexts without constraints and with constraints (later chapters).

Before getting to definitions and methods directly related to optimization, we first mention a few terms regarding sets of real numbers. Let (a,b) denote the set of all real numbers x such that $a < x < b$. We refer to this as an "open interval". The points a and b are called boundary points. A set is closed if it contains all of its boundary points. Therefore $[a,b]$, the set of all real numbers x such that $a \le x \le b$, is a "closed interval". The set $[a,b)$ is neither open or closed. The entire real line $(-\infty,\infty)$ is an open interval, because it doesn't contain its end points. [It is also a closed interval! It is "automatically" closed since there are no boundary point for it to contain. The entire real line, as an interval, is unusual in that it is both open and closed.]

Definitions

Consider a function f defined over domain D. The point x_0 is a **global maximum point** of f if

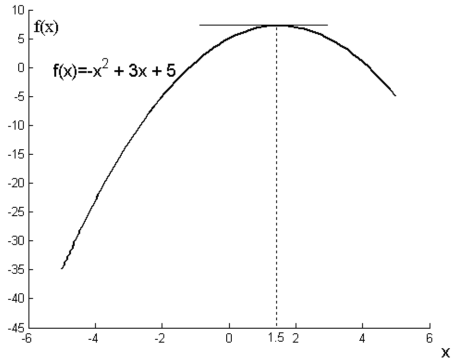
$$f(x) \le f(x_0) \text{ for every } x \in D$$

The value $f(x_0)$ is called the **maximum value** of the function. The point x_0 is a **strict global maximum point** for f if

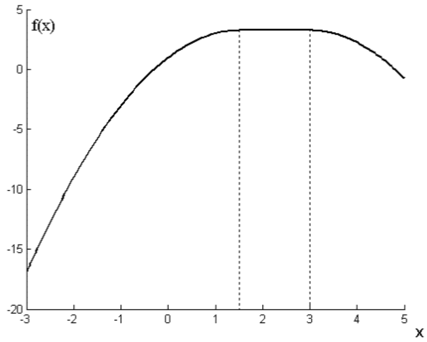
$$f(x) < f(x_0) \text{ for every } x \in D,\ x \ne x_0 .$$

The value $f(x_0)$ is called the **strict maximum value** of the function.

Example 11.3 Take the domain to be ℝ. There is a strict global maximum point at $x = 1.5$.



Example 11.1.2 Take the domain to be ℝ. There are many global maximum points (all points x from 1.5 to 3 are maximum points.)



The definition of minimum points is similar, with obvious changes:

The point x_0 is a **minimum point** for f if $f(x) \geq f(x_0)$ for every $x \in D$. The value $f(x_0)$ is called the **minimum value** of the function. The point x_0 is a **strict minimum point** for f if $f(x) > f(x_0)$ for every $x \in D$, $x \neq x_0$. The value $f(x_0)$ is called the **strict minimum value** of the function.

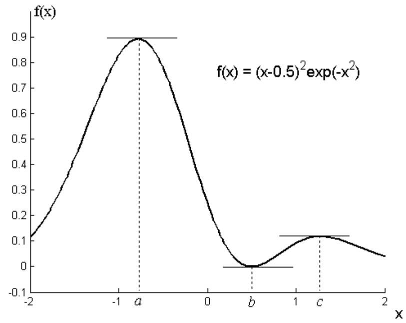
When we do not need to distinguish between a minimum point or a maximum point, we can simply call such points **optimal points**, and their corresponding functions values as **optimal values**.

Local vs Global Optimal Points

Given a function f defined over *D*, the point $x_0 \in D$ is a **local maximum point** if there is an open interval $(x_0 - \varepsilon, x_0 + \varepsilon)$, $\varepsilon > 0$, such that

$$f(x) \le f(x_0) \text{ for every } x \in (x_0 - \varepsilon, x_0 + \varepsilon).$$

The value $f(x_0)$ is called a **local maximum value** of the function. Local minimums are defined in a similar manner.



Example 11.4 The point $x = c$ is a local maximum point. Note that all global maximums are by definition also local maximums, so the points $x = a$ and $x = b$ are, strictly speaking, also local optima.

Our objective will be to develop methods to find optimum points, and to characterize them as maximums or minimums, global or local, unique or not, etc. We will be presenting several different methods, none of them perfect. All apply only in given circumstances, and you will need to know how to adapt when those circumstances are not met. Also, depending on the function to which a method is being applied, the method may give you a less-than-complete answer, compared to when it is applied to other functions.

Optimization is a skill that cannot be fully automated – many tools are available, but the application of the tools and the interpretation of the results require thought.

First Order Condition for (Local or Global) Optima

From the examples, one will suspect that the first derivative of f has a role to play in finding optimal points, as all the optimal points in the examples satisfy $f'(c) = 0$. In fact, the following is true:

Necessary "First-Order" Condition Suppose f is defined over an open interval I and is differentiable over I. Then,

$$x_0 \text{ is an optimum point in } I \;\Rightarrow\; f'(x_0) = 0.$$

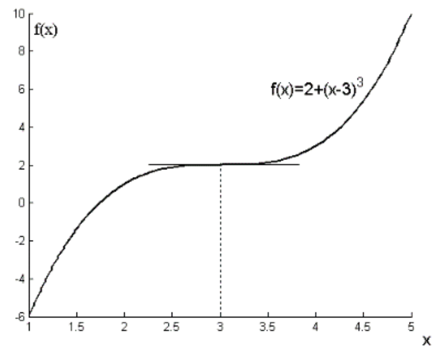
We call x_0 a stationary point.

It is easy to justify this argument. Intuitively, if $f'(x_0) > 0$ or $f'(x_0) < 0$, then we can increase or decrease the value of the function by simply increasing or decreasing x_0 respectively.

Further discussion of the First-Order Condition First, what does it mean for the first-order condition to be a necessary (but not sufficient) condition? "Necessity" means that if all the conditions of the theorem are met, then all optimal points x_0 will have the property that $f'(x_0) = 0$. Finding all the points such that satisfy $f'(x_0) = 0$ will then pick out all the optimum points. However, $f'(x_0) = 0$ is not sufficient to guarantee that x_0 is an optimum point; there may be points that may satisfy $f'(x_0) = 0$ but that are not optimum points. In particular, certain inflection points can also satisfy $f'(x_0) = 0$.

Example 11.5 *Why is the FOC not sufficient?*

Because the candidate optimal point might be an inflection point. In example on the right, the domain is ℝ. We have $f'(3) = 0$, but 3 is clearly not an optimal point. In fact, there are no optimal points.



Nonetheless, the FOC remains a very useful tool, because it helps us narrow down the set of 'candidate points', which we can then sort through using other tools. For example, suppose we are trying to find the optimum points of the function

$$f(x) = \frac{2x^2}{x^4+1}, \; x \in \mathbb{R}$$

We note this function is defined for all x, i.e., the function is defined over the open interval $(-\infty, \infty)$, and is differentiable over the entire interval. The first derivative is

$$f'(x) = \frac{4x(1+x^2)(1+x)(1-x)}{(x^4+1)^2},$$

so the point $x = -1$, 0, and 1, satisfy $f'(x) = 0$. These three points are therefore candidate optimum points.

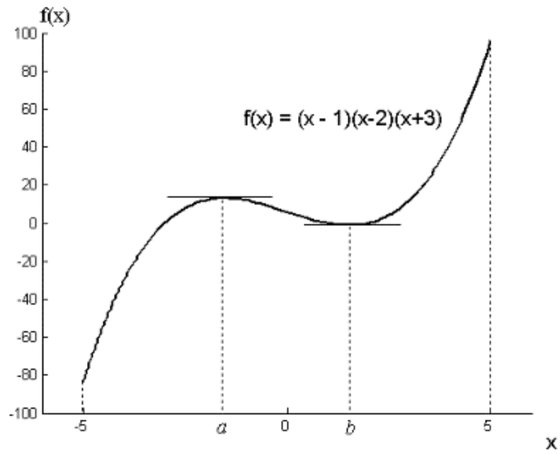
Our analysis up to this stage cannot say more about the points. Some of the points may be max, or min, or neither. But we have reduced the problem substantially, since we know that all other points are **not** optimum points.

It is also essential to remember that the FOC works well only when the function is defined over an open interval, and is differentiable over that open interval. In particular, the FOC does not apply to boundary points, and points of non-differentiability. The following example illustrates what can happen when either of the two conditions are not met.

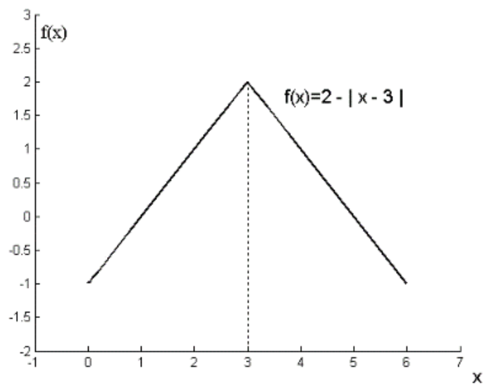
Example 11.6 Why did we require I to be open? Consider an example where f is not defined over an open interval. Here

$$f(x)=(x-1)(x-2)(x+3)$$

with $x \in [-5,5]$. The FOC picks out the two (local) optimum points, but misses the global max and min which occur at 5 and -5 respectively. Note that if this function was defined over $(-5,5)$, then we would not have this problem, since $x=-5$ and $x=5$ would no longer be optimum points.



Example 11.7 *Why must f be differentiable over the interval?* In the example on the right, there is a strict maximum at $x=3$. However, $f'(3) \neq 0$. In fact, the function is not differentiable at this point. The FOC would have missed this point.



The problem is that (i) the optimum point might occur at a boundary point, and (ii) the optimum point might occur at points of non-differentiability, but the first-order condition applies only at interior points that are also differentiable.

We next turn to the next steps after finding the candidate optimum points via the FOC. Several methods are presented. Remember, when you would use one method over another depends on the situation.

The Extreme Value Theorem The first method, in fact, deals with the case where the function is defined over a **closed** interval.

Theorem (The Extreme Value Theorem) Suppose f is a continuous function defined over a closed and bounded interval. Then there is a point $x_1 \in [a,b]$ that is a global maximum point, and a point $x_2 \in [a,b]$ that is a global minimum point.
(Proof omitted.)

This theorem says that if the domain of the function is a closed and bounded interval, and if the function is continuous over this domain, then a global minimum and a global maximum will exist. The fact that the global maximum and global minimum exists is very helpful. If we know the global max and global min exists, then the discussion in the previous section suggests that the following steps lead you to them:

1. Find all the points x_0 such that $f'(x_0) = 0$ and compute the value of the function at these points.
2. Find all the points where the function is not differentiable, and compute the value of the function at these points.
3. Find $f(a)$ and $f(b)$, the value of the function at the boundary points.
4. Compare the value of the function at all of these points. The point with the largest function value is the global maximum point, and the point with the lowest function value is the global minimum point.

Example 11.8 Let

$$f(x) = \frac{x^2 - 4x + 4}{1 + |x-1|}, \; x \in [0,5].$$

Find the global maximum and global minimum points.

This function is easier to deal with than it first appears. Notice that the denominator is always greater than zero, so the function is defined at all points in its domain (in fact defined everywhere, but we are only interested in the function over the stated domain). There is on obvious point of non-differentiability, that is at $x = 1$. The function is differentiable everywhere else. The value of the function at the boundary points are

$$f(0) = 2 \text{ and } f(5) = \frac{5^2 - 4(5) + 4}{1 + |5 - 1|} = 9/5.$$

Over the interval $(0,1)$, we have

$$f(x) = \frac{x^2 - 4x + 4}{1 + |x - 1|} = \frac{x^2 - 4x + 4}{1 + 1 - x} = \frac{x^2 - 4x + 4}{2 - x} = 2 - x,$$

so $f'(x) = -1$, and there are no points over $(0,1)$ such that $f'(x) = 0$.

The function is not differentiable at $x = 1$. At this point,

$$f(1) = \frac{1 - 4 + 4}{1} = 1$$

Over the interval $(1,5)$, the function is

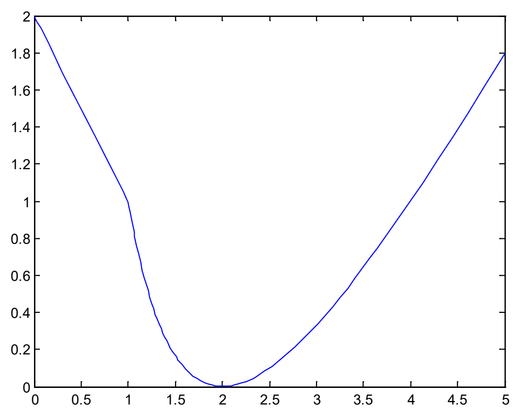
$$f(x) = \frac{x^2 - 4x + 4}{1 + |x - 1|} = \frac{x^2 - 4x + 4}{1 + x - 1} = x - 4 + \frac{4}{x}.$$

The first derivative is

$$f'(x) = 1 - \frac{4}{x^2},$$

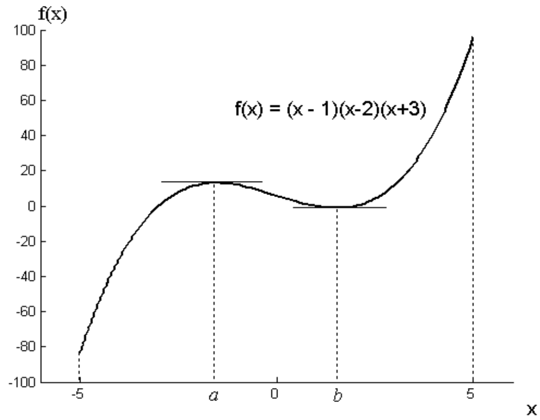
so $f'(x) = 0$ at the point $x = 2$, with $f(2) = 0$.

Comparing $f(0) = 2$, $f(1) = 1$, $f(2) = 0$, and $f(5) = 9/5$ we find that the global maximum point occurs at the boundary point $x = 0$ with $f(0) = 2$, and the global minimum occurs at the (interior, differentiable) point $x = 2$ with $f(2) = 0$. The function is plotted for you on the right.



Example 11.2.2 (revisited) Let

$f(x) = (x-1)(x-2)(x+3)$, with $x \in [-5,5]$.



For this example, using the procedure outlined above, you will find two interior stationary points [points where $f'(x) = 0$]. Comparing the value of the function at these points with the value of the function at the boundary, you will come to the conclusion that the global min occurs at $x = -5$ and the global max occurs at $x = 5$.

Notice that this method will be silent on the two interior stationary point. It cannot tell you that the points $x = a$ and $x = b$ in the figure are local max and local min respectively.

For the extreme value theorem to guarantee the existence of a global min and a global max, we must have f to be continuous, and the domain to the closed and bounded. If f is not continuous, or the domain not closed and bounded, (or both), then the either the global min, or the global max (or both) might not exist (they might exist, or they might not).

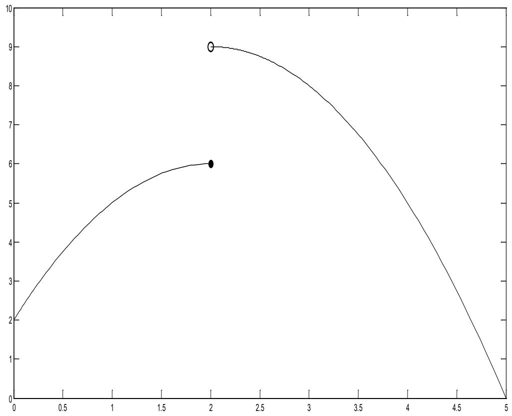
Suppose the function above was defined over $(-5,5)$. Then there are no global maximum and global minimum points. Using the procedure above, you would have identified two points $x = a$ and $x = b$ as candidate points, and determined by comparing values that $x = a$ is the global maximum, and $x = b$ is the global minimum point. This would be incorrect. They are merely local optima.

The next example shows what might happen if $f(x)$ is not continuous.

Example 11.9 Let

$$f(x) = \begin{cases} 4x - x^2 + 2, & 0 \le x \le 2 \\ 4x - x^2 + 5, & 2 < x \le 5 \end{cases}$$

This function is defined over the closed and bounded interval $[0,5]$ but it is not continuous. The global minimum still exists (at $x=5$), and the procedure described above will pick this out. However, the procedure will have picked out $x=2$ as the global maximum point. This would be incorrect. The point $x=2$ is not a global maximum point (it is not even a local minimum point). The problem here is, as it turns out, that a global maximum point does not exist.



We return to differentiable functions defined over open intervals.

The First Derivative Test We can use the first derivative more extensively to get further information about the nature of any stationary points.

<u>The First-Derivative Test for Optima</u> Suppose f is defined on an open interval I and is differentiable on I. Let x_0 be a stationary point.

(a) If $f'(x) \geq 0$ for all x in I such that $x \leq x_0$, and $f'(x) \leq 0$ for all $x \geq x_0$, then x_0 is a maximum point over the interval I.

(b) If $f'(x) \leq 0$ for all x in I such that $x \leq x_0$, and $f'(x) \geq 0$ for all $x \geq x_0$, then x_0 is a minimum point over the interval I.

If the interval I is the entire domain of the function, then x_0 is a global optimum. If the interval I is some small interval centered at x_0, then the condition guarantees x_0 is a local optimum. If the inequalities are changed to strict inequalities, the optima are strict. If the sign does not change across a stationary point, you have an inflection point.

Example 11.10 Let $f(x) = e^{2x} - 5e^x + 4$, $x \in \mathbb{R}$.

Then $f'(x) = 2e^{2x} - 5e^x = e^x(2e^x - 5) = 0$ when $2e^x - 5 = 0$, i.e., $x = \ln(5/2)$ is a stationary point. Furthermore, e^x is always positive, and $2e^x - 5$ is increasing, therefore $f'(x) < 0$ for all $x < \ln(5/2)$, and $f'(x) > 0$ for all $x > \ln(5/2)$, so $x = \ln(5/2)$ is a strict minimum point.

Example 11.11 Let

$$f(x) = \frac{2x^2}{x^4 + 1},$$

then $f'(x) = \dfrac{4x(1+x^2)(1+x)(1-x)}{(x^4+1)^2}$. We have the following "sign" diagram

	---------- -1 ------------	0 --------------	$+1$ ---------------	
$4x$	$-$	$-$	$+$	$+$
$(1+x)$	$-$	$+$	$+$	$+$
$(1-x)$	$+$	$+$	$+$	$-$
$f'(x)$	$+$	$-$	$+$	$-$

so $x = -1$ and $x = 1$ are local maximums, and $x = 0$ is a local minimum of $f(x)$. Note that this analysis does not say that the stationary points are not also global optima. In fact, the stationary points here are also global minimums. To show this requires further argument, however. The first derivative test here guarantees only local minimization.

Maximum/Minimum for Concave/Convex Functions The class of concave and convex functions are much easier to deal with. Recall that for twice-differentiable functions

f is convex on an interval I $\Leftrightarrow$ $f''(x) \geq 0$ for all x in I

f is concave on an interval I $\Leftrightarrow$ $f''(x) \leq 0$ for all x in I

$f''(x) > 0$ for all x in I $\Rightarrow$ f is strictly convex on an interval I

$f''(x) < 0$ for all x in I $\Rightarrow$ f is strictly concave on an interval I

If $f'(x_0) = 0$, and $f''(x) \geq 0$ for all x in some interval I, i.e., $f'(x)$ is non-decreasing over the interval, then it must be that $f'(x) \geq 0$ for all $x > x_0$ in I, and $f'(x) \leq 0$ for all $x < x_0$ in I. In other words, x_0 is a minimum over the interval I.

A similar argument shows that if $f'(x_0) = 0$ for $x_0 \in I$, and $f''(x) > 0$ over I, then we have a strict minimum over the interval I.

Example 11.12 Let $f(x) = e^{x-1} - x$. Then

$f'(x) = e^{x-1} - 1 = 0$ when $x = 1$, and

$f''(x) = e^{x-1} > 0$ for all x.

Therefore, $f(x)$ is strictly convex, and $x = 1$ is a strict global minimum point.

There are times when it is easier to use the first derivative test to find maximum and minimum points (such as when the function's second derivative is messy, or when the function is not twice differentiable everywhere). At other times, it is easier to use second derivative (such as in Example 11.12). Note, however, that many economics problem fall into the latter category. Also, it is often easier to use the second derivative if our intention is to characterize an optimum point, rather than to find its value.

Economists often call the condition that $f'(x_0) = 0$ for a stationary point the "First-Order Condition" of "FOC", and the second-derivative condition the "Second-Order Condition", or "SOC".

The following is a problem with parameters.

<u>Example 11.13</u> Let $\pi(x) = P\sqrt{x} - 2x$, $x > 0$. Find the optimal points.

Here is a problem with a parameter P, so the solution will no longer be a numerical constant, but will depend on P:

FOC: $\pi'(x) = \frac{1}{2}\frac{P}{\sqrt{x}} - 2 = 0$ which gives $x^* = P^2/16$.

SOC: $\pi''(x) = -\frac{1}{4}Px^{-3/2} < 0$ so the solution is a global maximum.

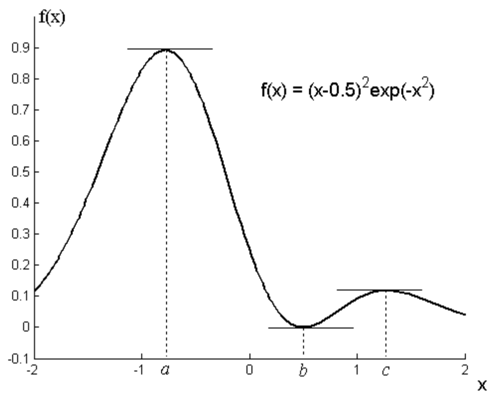
The maximum value is $\pi^* = \pi(x^*) = P\sqrt{x^*} - x^* = \frac{P^2}{4} - \frac{P^2}{16} = \frac{3P^2}{16}$.

In problems involving parameters, we are often interested not only in finding the optimal points and values of the objective function at the optima, but also in finding out how these changes as the parameters change. In this example, we have

$$\frac{dx^*}{dP} = \frac{P}{8} > 0 \quad \text{and} \quad \frac{d\pi^*}{dP} = \frac{3P}{8} > 0.$$

We have shown that if we can show that the function to be optimized belongs to a certain class, then we can assert that the first-order condition delivers the global optimal point.

Remember, however, that there are many functions that have global optima but are neither concave nor convex. In this case, you won't be able to use the SOC to come up with a statement concerning global optimality. The function on the right is not concave or convex over its entire domain, but it most certainly has a global maximum (and a global minimum).



The function in example 11.10 is another example. That function is not convex or concave everywhere – the second derivative doesn't have a fixed sign. Yet it has a strict global minimum, as we showed earlier. In this sense, the second order condition is sufficient, but not necessary.

Using the Second Derivative to Characterize Local Optimal Points How do we use the second derivative to characterize <u>local</u> optimal points? We develop rules for functions f where f'' is continuous;

1) $f'(x_0) = 0$ and $f''(x_0) < 0 \quad \Rightarrow \quad f$ has a strict local maximum at $x = x_0$.

2) $f'(x_0) = 0$ and $f''(x_0) > 0 \quad \Rightarrow \quad f$ has a strict local minimum at $x = x_0$.

3) $f'(x_0) = 0$ and $f''(x_0) = 0 \quad \Rightarrow \quad ?$

<u>Proof</u> If $f'(x_0) = 0$ and $f''(x_0) < 0$, then

$$f''(x_0) \;=\; \lim_{h\to 0}\frac{f'(x_0+h) - f'(x_0)}{h} \;=\; \lim_{h\to 0}\frac{f'(x_0+h)}{h} \;<\; 0$$

Therefore there is a small interval $(-\varepsilon,\varepsilon)$ such that $f'(x_0+h)/h < 0$ for all $h \in (-\varepsilon,\varepsilon)$. For $h \in (0,\varepsilon)$, we have $f'(x_0+h) < 0$ because h is positive. For $h \in (-\varepsilon, 0)$, $f'(x_0+h) > 0$. By the first-derivative test, x_0 is a strict local minimum point. Proof of part (2) is similar.

To show (3), we show examples where $f'(x_0) = 0$ and $f''(x_0) = 0$, but where x_0 might be neither a local maximum or a local minimum:

Example 11.14 Define for all $x \in \mathbb{R}$,

$$f(x) = x^4, \ g(x) = x^3, \text{ and } h(x) = -x^4.$$

We have $f'(x) = 4x^3$, $g'(x) = 3x^2$, and $h'(x) = -4x^3$ so the only stationary point of each of these functions is $x = 0$. In addition

$$f''(x) = 12x^2, \ g''(x) = 6x, \text{ and } h''(x) = -12x^2$$

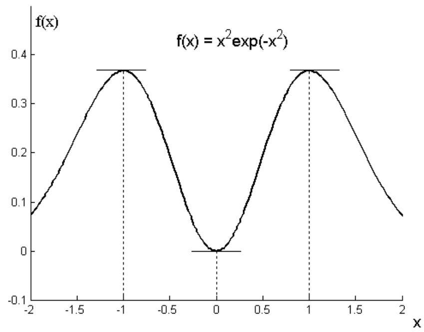
so $f''(0) = 0$, $g''(0) = 0$, and $h''(0) = 0$. However, $x = 0$ is a min. point of $f(x)$, neither a min. point nor a max. point for $g(x)$, and a max. point for $h(x)$.

The next example uses the second derivative to characterize local optimality.

Example 11.15 Let

$$f(x) = x^2 \exp(-x^2).$$

The stationary points $f(x)$ are the values of x which satisfy $f'(x) = 2e^{-x^2}(x - x^3) = 0$, and these are $x = -1, 0, 1$. Because $f''(x) = 2e^{-x^2}[1 - 5x^2 + 2x^4]$, we have $f''(-1) = f''(1) < 0$, and $f''(0) = 2 > 0$, therefore this function has local maxima at $x = -1$ and $x = 1$, and a local minimum at $x = 0$.



Exercises

1. (a) Use the first derivative test to find the extreme points of

$$\text{(i)} \quad f(x) = x^3 - x^2 - 2x + 3, \qquad \text{(ii)} \quad f(x) = e^x + e^{-2x}$$

Characterize the extreme points as much as you can (local, max or min, etc.).

(b) Is $f(x) = e^x + e^{-2x}$ convex or concave (or neither?) Use your result to determine whether the stationary point is a strict maximum or minimum point.

2. Find all local maximum and minimum points of the following functions, using the second derivative test:

$$\text{(a)} \quad f(x) = x^3 - x^2 - 2x + 3. \qquad \text{(b)} \quad f(x) = \frac{2x^2}{x^4 + 1}$$

3. Find all optimum points (local or global) of each of the following functions

(a) $f(x) = 4x^2 - 4 + 1$ (b) $f(x) = x^2 e^{-x^2}$

(c) $f(x) = \dfrac{x^3}{3} - \dfrac{3x^2}{2} + 2x + 6$ (d) $f(x) = \dfrac{x}{x^2+1}$

4. Find the global maximum point x^* of the function $y = 9 - (x-1)^2 - 2(x-b)^2$ and the value of y at that point, y^*? How does x^* and y^* change with b? Plot $x^*(b)$ and $y^*(b)$.

5. Find all global and local maximums and minimums, if any:

(a) $f(x) = \dfrac{3x}{\sqrt{4x^2+1}}$, $x \in [-1,1]$ (b) $f(x) = \dfrac{|x|}{|x+1|}$

(c) $f(x) = \begin{cases} 4x-2, & x<1 \\ (x-2)(x-3), & x \geq 1 \end{cases}$ (d) $f(x) = x^3 e^{-x}$

6. Let

$$f(x) = e^{2x} - 5e^x + 4.$$

The first-order condition $f'(x) = 2e^{2x} - 5e^x = 0$ produces $x^* = \ln(5/2)$ as the only stationary point. In the notes, we used the first-derivative test to show that this point is a global minimum point.

Find $f''(x)$. Is it true that $f''(x) > 0$ for all x? That is, does $f(x)$ satisfy the (global) second order condition for a global minimum? If yes, show it; if not, find the intervals over which $f''(x) > 0$, and the intervals over which $f''(x) < 0$. Plot the graph of $f(x)$. [*Hint for plotting the graph: what is* $\lim_{x\to-\infty} f(x)$*? What is* $\lim_{x\to\infty} f(x)$*?*]

7. What is wrong with the following argument:

> "Given a function $f(x)$, we find that $f'(-1) = f'(0) = f'(1) = 0$. In addition, we find that $f'(x) > 0$ over $(-\infty,-1)$ and $(0,1)$, and $f'(x) < 0$ over $(-1,0)$ and $(1,\infty)$. Therefore $x=-1$ and $x=1$ are local maximum points, and $x=0$ is a local minimum point. Because $x=0$ is the only local minimum point, it must also be a global minimum point."

Point out the error in the argument, giving examples to illustrate the error.

8. What is the difference between the first order condition and the first derivative test?