10. The Second-Order Derivative, Concavity and Convexity

The second-order derivative (or simply 'second derivative') is the derivative of the derivative

$$f''(a) = \lim_{h \to 0} \frac{f'(a+h) - f'(a)}{h}.$$

The objective of this section is to provide examples of their application in general, and in economics in particular.

We can also compute 'third-order derivatives', 'fourth-order derivatives', etc. These are useful, although we will be using them less frequently.

Example 10.1 If $f(x) = x^3$, then $f'(x) = 3x^2$, $f''(x) = 6x$, $f'''(x) = 6$.

Notation If $y = f(x)$, then

Second Derivative $f''(x)$, $f^{(2)}(x)$, y'', $\frac{d^2}{dx^2} f(x)$, $\frac{d^2 y}{dx^2}$

n-th ordered derivative $f^{(n)}(x)$, $y^{(n)}$, $\frac{d^n}{dx^n} f(x)$, $\frac{d^n y}{dx^n}$

If y is a function of time, then the second derivative is often written $\ddot{y}$, following the convention that the first-derivative in such applications is written $\dot{y}$.

The most straightforward interpretation of the second-derivative is that it is the rate of change of the rate of change of $f(x)$ as x increases.

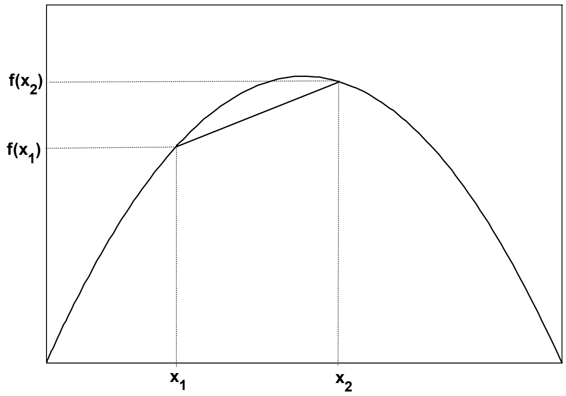
Geometrically, the second derivative can also help to describe the shape of a function. When used in specific disciplines, the first and second derivations can often be given an interpretation appropriate to the context of that discipline. You are probably familiar with the concepts of velocity and acceleration in physics.

Example 10.2 If $x(t)$ gives you the position of an object relative to some reference point $x(0)$, and t is time, then $\dot{x}$ is the velocity of the object, and $\ddot{x}$ is its acceleration.

Geometric Interpretations

We can use the second-derivative to characterize the shape of a function, particularly regarding its curvature. We begin with a geometric definition of the concept of convexity and concavity.

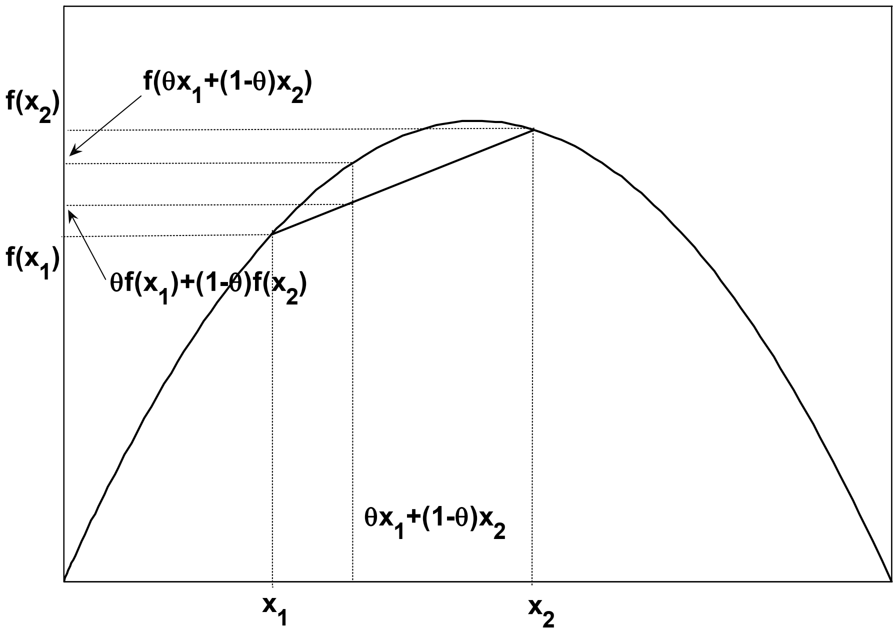
A function is said to be strictly concave if for any $x_1 < x_2$, the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ lies strictly below the function for all values of x strictly between x_1 and x_2, $x_1 < x < x_2$



One way to express this mathematically is given in next definition.

Definition A function is **strictly concave** over its domain if for any two values of x, say x_1 and x_2 with $x_1 < x_2$, and for all $0 < \theta < 1$,

$$\theta f(x_1) + (1-\theta) f(x_2) \;<\; f(\theta x_1 + (1-\theta) x_2)$$



If we allow the chord to lie on the function at some points, then the function is said to be weakly concave, or simply 'concave'. A function is **concave** if for any two values of x, say x_1 and x_2 with $x_1 < x_2$, and for all $0 \le \theta \le 1$, we have

$$\theta f(x_1) + (1-\theta) f(x_2) \le f(\theta x_1 + (1-\theta) x_2).$$

A function is **strictly convex** if for any two values of x, say x_1 and x_2 with $x_1 < x_2$, and for all $0 < \theta < 1$, we have

$$\theta f(x_1) + (1-\theta) f(x_2) > f(\theta x_1 + (1-\theta) x_2).$$

A function is **convex** if for any two values of x, say x_1 and x_2 with $x_1 < x_2$, and for all $0 < \theta < 1$, we have

$$\theta f(x_1) + (1-\theta) f(x_2) \ge f(\theta x_1 + (1-\theta) x_2), \ 0 \le \theta \le 1.$$

Note that strictly concave functions are concave; the set of strictly concave functions are a subset of the set of concave functions. Likewise, strictly convex functions are convex. Note also that strictly concave functions can be strictly increasing or strictly decreasing.

Example 10.3 $f(x) = x^{1/2}$ is strictly concave, and strictly increasing.

The definitions of concavity given above are very general, and can apply to functions that may be non-differentiable at certain points. For differentiable functions, however, it is often easier to make use of the second derivations to show concavity/convexity. From these definitions given earlier, it is possible to show that

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

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

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

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

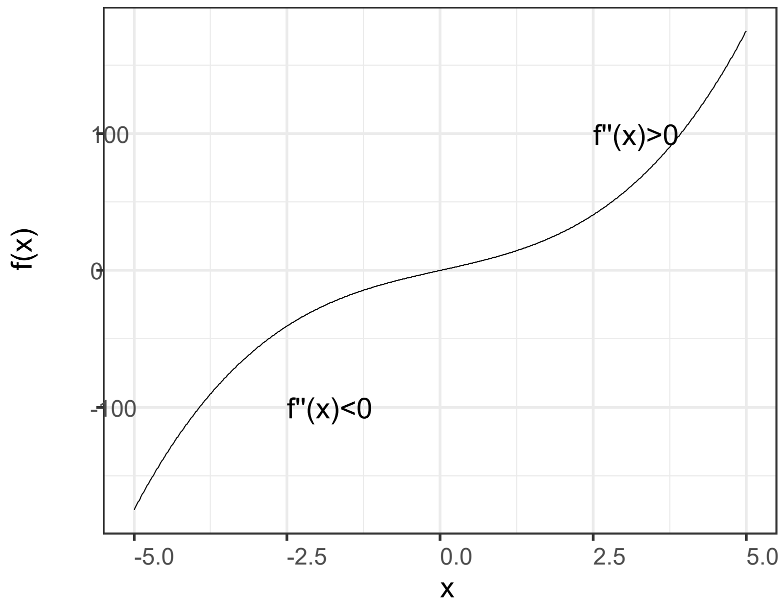
We omit proofs. The results are sufficiently intuitive for our purpose. For a function to be concave, it is apparent that the slope must never increase as x increases. It is also apparent that if the slope of the function always decreases as x increases, i.e., if $f''(x) < 0$, then the function will have a strictly concave shape.

The only aspect of definitions (1) to (4) that may not be obvious is that it is possible for a function to be strictly concave over an interval I without $f''(x) < 0$ for all x in I.

Example 10.4 The function $f(x) = -x^4$ is strictly concave, but it is not true that $f''(x) < 0$ for all x. In particular, $f''(x) = -12x^2 = 0$ at $x = 0$.

Convexity and concavity will play an important role in function optimization. For instance, if a strictly concave function $f(x)$ has a point x^* at which $f(x^*) = 0$, then x^* must be an optimal point.

A function may be concave in some portions of its domain, and convex in other parts of it. The point where a function switches from concavity to convexity, or the other way around, is called an **inflection point**. The plot below is of a function $f(x) = x^3 + 10x$, its first derivative is $f'(x) = 3x^2 + 10$ and its second derivative is $f''(x) = 6x$. It is concave when $x < 0$ and convex when $x > 0$.



Definition The point x_0 is called an **inflection point** of a function $f(x)$ if over some interval (a,b) containing x_0, we have

$$f''(x) \leq x \text{ over } (a, x_0) \text{ and } f''(x) \geq 0 \text{ over } (x_0, b)$$

or

$$f''(x) \geq x \text{ over } (a, x_0) \text{ and } f''(x) \leq 0 \text{ over } (x_0, b)$$

If x_0 is an inflection point, it must be that $f''(x_0)=0$. But note that $f''(x_0)=0$ does not imply that x_0 is an inflection point. (example: $f(x)=x^4$). However, if $f''(x_0)=0$ and $f''(x_0)$ changes sign at x_0, then x_0 is an inflection point. Note also that $f'(x_0)$ need not be zero for x_0 to be an inflection point, as in the function plotted above.

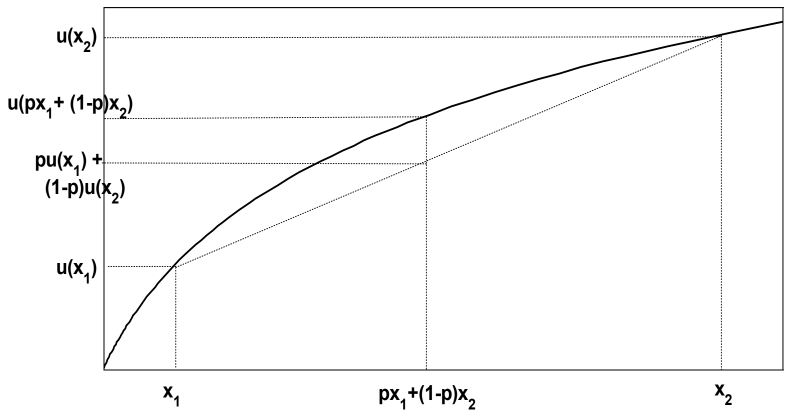
Some Economic Applications

The second derivative is useful in economics not only because of its role in optimization, but also because it can often be given economic interpretations.

Example 10.5 *Diminishing Marginal Productivity* As a firm increases labor input starting from low levels, its production capacity generally increases. However, the increase in productive capacity, as more labor is added, tends to get smaller: increasing labor from 100 to 101 might increase productive capacity by 10 units; increasing labor from 101 to 102 might increase productive capacity by 7 units, say. That is, the rate of increase in productive capacity falls as more labor is added. Suppose we represent the firm's production function by $f(L)$ where L is labor input. We can impose increasing productivity be requiring the condition $f'(L)>0$. We can impose "diminishing marginal productivity" by requiring $f''(x)<0$.

Example 10.6 *Diminishing Marginal Utility* Suppose $u(x)$ represents the utility that a person gets from consuming x amount of a good. If $u'(x)>0$ then more of the good is preferred to less of it. If $u''(x)<0$, then every additional unit of the good generates less additional utility for the consumer than the previous additional unit.

Example 10.7 *Risk aversion* Suppose again that $u'(x)>0$ and $u''(x)<0$. Such a function has a shape like in the figure below.



Suppose the person with utility function $u(x)$ is given a choice. He can either accept a lottery that pays x_1 with probability p, and x_2 with probability $1-p$, where $0<p<1$, or accept $px_1+(1-p)x_2$ with certainty. In the first case, the person's expected utility is $pu(x_1)+(1-p)u(x_2)$. In the second case, the person gets utility $u(px_1+(1-p)x_2)$.

In the figure we see that because of the shape of the function,

$$pu(x_1)+(1-p)u(x_2) \; < \; u(px_1+(1-p)x_2)$$

This person would rather have the certain amount. He would avoid the lottery if given a choice. He is 'risk-averse'. Contrast this with the fellow for whom

$$pu(x_1)+(1-p)u(x_2) \; > \; u(px_1+(1-p)x_2) .$$

The person prefers the lottery. He is 'risk-loving'.

Exercise

1. Show that the function $f(x)=e^{x-1}-x$ is strictly convex.
2. Find the regions of x for which $f(x)=(x-1)(x-2)(x+3)$ is (i) strictly convex, (ii) strictly concave.
3. Find the first and second derivatives of each of the following functions.

   (a) $f(\mu) = \sum_{i=1}^{n}(x_i-\mu)^2$ where x_i, $i=1,2,...,n$ is a set of numbers. Find the intervals over which the function is (strictly) increasing, decreasing, concave, or convex.

   (b) $y = f(x) = \log_x e$. What is the largest possible domain of this function? Find the intervals over which the functions are (strictly) increasing, decreasing, concave, or convex. Sketch the graph