APPENDIX A: MATHEMATICAL FOUNDATIONS

A.1 IS IT REALLY TRUE?

Key ideas: direct proof, proof by induction, proving the contrapositive, proof by contradiction

The whole point of building a mathematical model is to be able to use the power of mathematical analysis to draw conclusions about economic phenomena. Thus it is essential to have on hand a good toolkit of mathematical propositions. For example suppose that you wish to analyze the choice of a profit maximizing firm that sells in a market where the price of the commodity q is p. For any quantity, price pair (q, p) there is some profit $f(q, p)$. To work with this model it is useful to have in the toolkit a clear notion of the rate of change in the function as output changes. Two possible functions are depicted below.

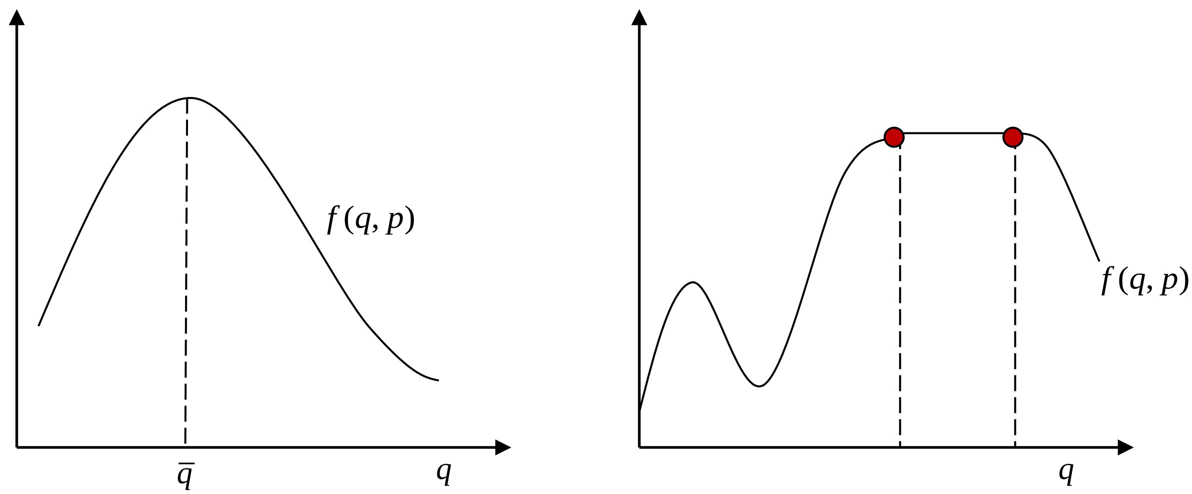


Figure A.1-1: The profit of a firm

Clearly it is easier to characterize the profit maximizing output in the first case, because the top of the profit hill is the unique point where the slope is zero. In the second case there are lots of outputs where the slope is zero but not all of them yield the profit maximum. Thus, to simplify the model, it is useful to seek plausible conditions under which the profit curve has a shape that is simple to analyze. This is relatively easy if the firm is choosing a single output. But what if q represents a vector of outputs of different commodities? Taking the example a little further, suppose that $q(p)$ is a solution to the maximization problem for each output price p. To analyze the firm's behavior it is simpler if this output varies smoothly with the output price. But is this a reasonable assumption?

One answer to such questions might be that the key is to put together a toolkit of mathematical methods and results that is large enough to deal with all the basics of economic modeling. However, while this approach works quite well at an introductory level, it is ultimately inadequate for someone wanting a deeper understanding of economic models. Without a clear appreciation of why results hold and how techniques work, it is extremely difficult deciding how to approach a problem. For this reason understanding the derivation of results and how techniques are developed is very important.

How does one go about proving some proposition, that is, demonstrate convincingly that the proposition is true? First there must be some common set of principles or results that are sufficiently familiar that they do not need reiterating or reproving. Mathematicians then employ several approaches. We discuss each in turn.

Direct proof

The straightforward approach is built on other accepted propositions. Consider, for example the present value of an investment yielding a constant payment of 1 dollar in each of *n* periods. With an interest rate *r*, the present value, discounted to the time of the first payment is

$$V_n = 1 + \frac{1}{1+r} + \ldots + \left(\frac{1}{1+r}\right)^{n-1}.$$

This is an example of a geometric sum $S_n = 1 + a + a^2 + \ldots + a^{n-1}$. The well known formula for this sum is

$$S_n = \frac{1-a^n}{1-a}, \quad a \neq 1.$$

The direct proof is short. We note that $aS_n = a(1 + a + a^2 + \ldots + a^{n-1})$ and hence, $1 + aS_n = (1 + a + a^2 + \ldots + a^{n-1}) + a^n$. But the term in the parentheses is S_n. Therefore

$S_n(1-a) = 1-a^n$. As long as $a \neq 1$, we can divide by $1-a$ to obtain $S_n = \dfrac{1-a^n}{1-a}$.

Quod erat demonstrandum (QED)

(Translation: This is what was to be demonstrated).

Note that even this simple proof assumes a common understanding of the rules of algebra.

Proof by induction

Proof by induction is often helpful if the goal is to establish that some proposition P_n is true for all integers $n = 1,2,\ldots$. The first step is to establish that the proposition is true for $n=1$. The second step is to show that if the proposition holds for all integers up to k, it must hold for $n=k+1$. If both can be proved then we are finished because if P_1 is true then, by the second step P_2 is true. Then, by the second step, because P_1 and P_2 are both true so is P_3. Because this argument can be repeated over and over again it follows that for all n P_n is true.

Consider the following example. Let S_n be the sum of the first n integers, that is, $S_n = 1+ ... + n$. Note that the sequence S_1, S_2, S_3, S_4, S_5 is 1,3,6,10,15. Someone might note that the formula $S_n = \frac{1}{2}n(n+1)$ fits for all 5 elements of this sequence and conjecture (correctly) that the formula must hold for all n. To prove this by induction we begin by supposing that the conjecture holds for all integers up to some integer k and seek to show that it must then hold for the integer $k+1$.

For our example, suppose $S_k = \frac{1}{2}k(k+1)$. We wish to show that

$S_k + (k+1) = \frac{1}{2}(k+1)(k+2)$.

Substituting for S_k,

$$S_k + (k+1) = \tfrac{1}{2}k(k+1) + (k+1) = (\tfrac{1}{2}k+1)(k+1) = \tfrac{1}{2}(k+1)(k+2).$$

Thus if the proposition is true for $n = k$ it is true for $n = k + 1$. We have already seen that the conjecture is true for n = 1 thus it is true for all n.

Proof of the contrapositive

Suppose we would like to prove that if the statement A is true then the statement B must also be true. That is, any event in which A is true is also an event in which B is true.[1]

Consider the three Venn diagrams below. In each case the box is the set of all possible events. In the left diagram the heavily shaded region is the set of events in which A is true. The set of events in which A is false is the set $\sim A$.

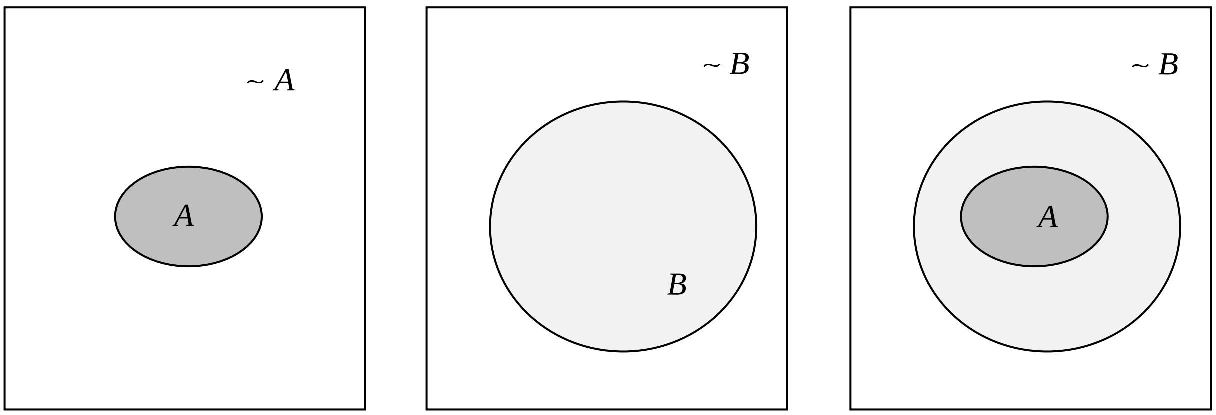


Figure A.1-2: The Contrapositive

Similarly in the middle diagram, the lightly shaded region is the set of events in which B is true. The two sets are superimposed in the right hand diagram. As drawn, A is a subset of B. Therefore, if A is true then B must be true as well. Thus the goal is to prove that A is a subset of B $(A \subset B)$. Note that if this is true, it is also true that any event not in B is also not in A $(\sim B \subset \sim A)$. That is, the two statements are equivalent:

$$A \subset B \text{ if and only if } \sim B \subset \sim A.$$

[1] Equivalently, condition B is a necessary condition for condition A.

Thus instead of attempting a direct proof that *A* implies *B*, we can appeal to the contrapositive and attempt to prove that if *B* is false, then *A* must also be false.

A good example of this approach can be found in section A.5 where we show how to construct a proof of the following statement.

"If the smooth function *f* takes on a maximum at $\bar{x}$, then the slope is zero at $\bar{x}$."

The contrapositive of this statement is that if the slope of a function is not zero at $\bar{x}$, then the function does not take on a maximum at $\bar{x}$. The proof then follows from an examination of the formal definition of the "slope" of a function.

Proof by contradiction

Last but by no means least, we often prove that a statement is true by deriving the implications that follow if the statement is false. Suppose, by a combination of luck and cunning, we find an implication that is impossible. Then we know that the statement cannot be false and so it must be true!

To illustrate, we prove that the sum of two odd numbers must be even. Any even number can be written as $2n$ where n is an integer and so any odd number can be written as $2n+1$. Suppose that the statement is false so that for some integers, a,b,c

$$(2a+1)+(2b+1)=2c+1. \tag{A.1.1}$$

Rearranging this equation,

$$2(c-a-b)=1.$$

The number $2(c-a-b)$ is divisible by 2 so it cannot be equal to 1. Hence equation (A.1.1) leads to a contradiction. Thus there cannot be any such numbers a,b,c and so the statement must be true.

A.2 MAPPINGS OF A SINGLE VARIABLE[2]

Key Ideas: neighborhood, limit, continuous function, set-valued mappings

A real number is a point in the infinite interval $\mathbb{R} = (-\infty, \infty)$. Throughout this book we will consider sets of real numbers. Typically we will consider sets of points that are intervals, for example,

$$X = \{x \mid x \in \mathbb{R}, a \le x \le b\}.$$

Any set such as X that contains its endpoints is known as a closed interval. The mathematical shorthand for such an interval is $[a,b]$. Similarly we define an open interval as $(a,b) = \{x \mid x \in \mathbb{R},\ a < x < b\}$.

A function maps each of the points in some set X into a unique point in $(-\infty, \infty)$. We write the output of this mapping as $f(x)$. The set X over which the function is defined is the domain of f and we write $X = D_f$.

Example 1: Firm's supply function

Consider a firm producing a single output. Let $C_1(q)$ be the minimized cost for each level of output q. This is a single valued mapping from output to cost so it is the firm's "cost function." Suppose that the cost function is as depicted in Figure A.2-1.

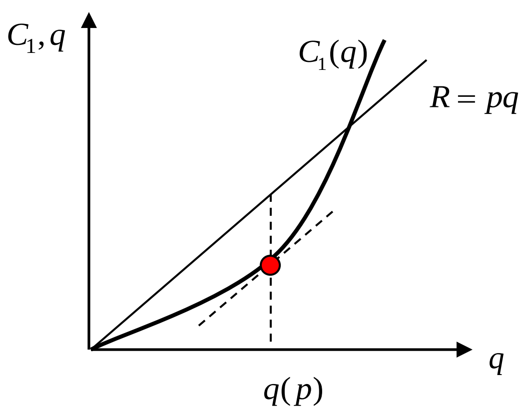


Fig A.2-1: Profit maximizing output

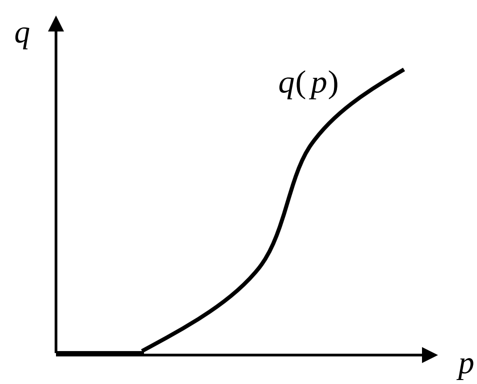


Figure A.2-2: Firm's supply function

[2] This chapter is a review of the mathematical foundations. For those readers who find the going tough, it will be helpful to keep open a basic calculus textbook at the same time.

Given the price *p*, profit is maximized at $q(p)$ where the gap between the cost curve and the revenue line, $R = pq$, is greatest. Thus the profit-maximizing firm supplies $q(p)$ units of output. As long as the slope of the cost curve is everywhere increasing, the profit maximizing output is unique so the mapping $q(p)$ from price to output is also a function. This is depicted in

Figure A.2-2. ■

Suppose we next modify Example 1 and introduce a fixed cost.

Example 2: Discontinuous function

The firm's cost is $C_2(q) = F + C_1(q)$. The new cost curve is depicted below.

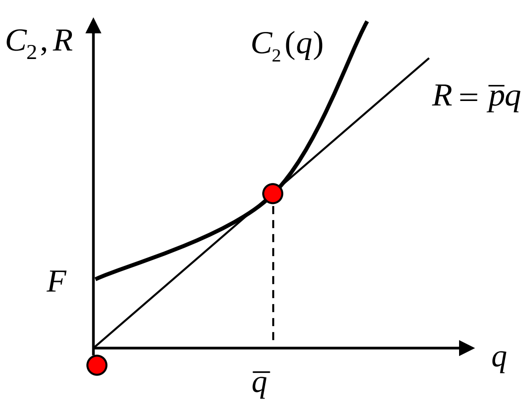


Figure A.2-3: Price yielding a zero profit

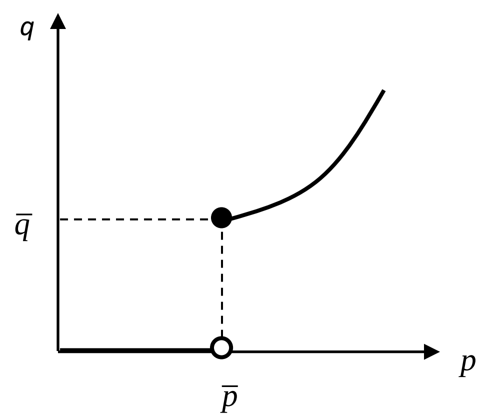


Figure A.2-4: Firm's supply function

Suppose that the price is $\bar{p}$. That is, the revenue line just touches the cost line at $\bar{q}$. The firm can make a profit of zero by either shutting down or producing $\bar{q}$ units of output. Any other output level yields a strictly negative profit. We will assume that in this case it produces $\bar{q}$.[3] At prices higher than $\bar{p}$, the profit maximizing output $q(p)$ is strictly positive, while at prices

[3] The output at price $\bar{p}$ is indicated by the dot in Figure A.1-4. The point indicated by the ring is not on the supply curve. An alternative approach would be to allow either point to be a possible choice by the firm. But then we no longer have a 1-1 mapping from price to output.

lower than $\bar{p}$ the firm shuts down. The mapping from price to output is therefore as depicted in Figure A.2-4. $\blacksquare$

There is an important difference between the supply function in Example 1 (Figure A.2-2) example and the supply function in Example 2 (Figure A.2-4). In the former case, output varies smoothly ("continuously") with price. In the latter, there is a jump in output at $\bar{p}$. The supply function is discontinuous at $\bar{p}$.

To provide a precise definition of continuity we appeal to the following preliminary definitions.

Definition: Neighborhood and deleted neighborhood

The set of points whose distance from x^0 is strictly less than δ is called a neighborhood (or $\delta-$neighborhood) of x^0. When the distance is strictly less then δ but strictly greater than zero (so that x^0 itself is deleted) the set of points is called a deleted neighborhood.

Definition: Limit of a function

The function f has a limit L at x^0 if, for any $\varepsilon > 0$, $|f(x) - L| < \varepsilon$ for all x in some deleted neighborhood of x^0.

Definition: Continuous function

Let f be a function and suppose that $x^0 \in D_f$. Then f is continuous at x^0 if it has a limit point at x^0 equal to $f(x^0)$.

Example 3: Function with no limit at x^0

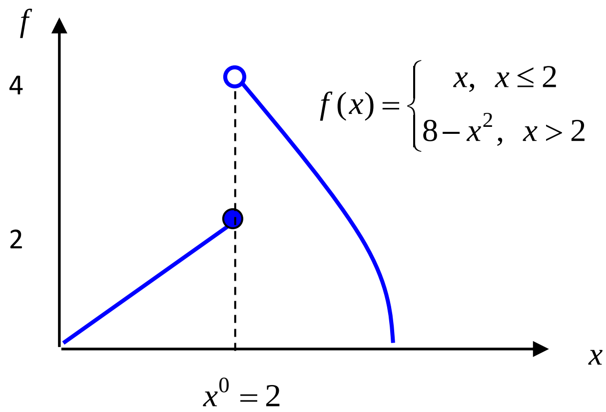


Figure A-2-5: f has no limit at x^0

Consider some small δ. For all x in the interval $(2-\delta, 2)$, $f(x)$ is close to 2. For all x in the interval $(2, 2+\delta)$, $f(x)$ is close to 4. Thus the function does not approach a limit L as x approaches $x^0 = 2$. ■

A function f is discontinuous over its domain if it is not continuous at some $x^0 \in D_f$. Consider the following example.

Example 4: A discontinuous function with a limit point

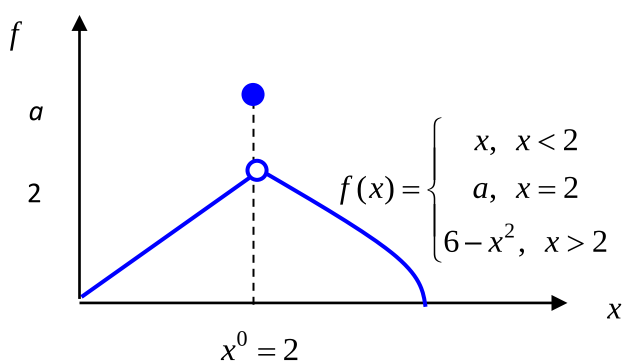


Figure A.2-6: f has a limit point but is not continuous at x^0

For all x close to 2, $f(x)$ is close to 2. Thus the limit of the function is $L = 2$.

Also $f(x^0) = a$. Therefore the function is discontinuous at x^0 unless $a = 2$. $\blacksquare$

Set-valued mappings

In economics, it is sometimes the case that the optimal choice of an economic decision-maker is not unique. For example, there can be prices for which a firm may have two or more profit maximizing outputs. Then the mapping from prices to outputs is not unique. More formally, mappings are sometimes set valued rather than single valued. That is, some or all points in a set *X* are mapped into sets of points on the real line. Such set-valued mappings are called correspondences.

To illustrate, suppose that a firm has two machines. Using machine 1 it can produce up to a units per day at a unit cost of c_1 and using machine 2, up to b units per day at a unit cost of $c_2 > c_1$. The cost curve is depicted in Figure A.2-7.

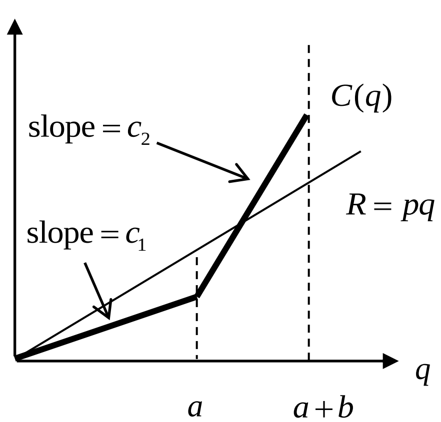


Figure A.2-7: Cost function

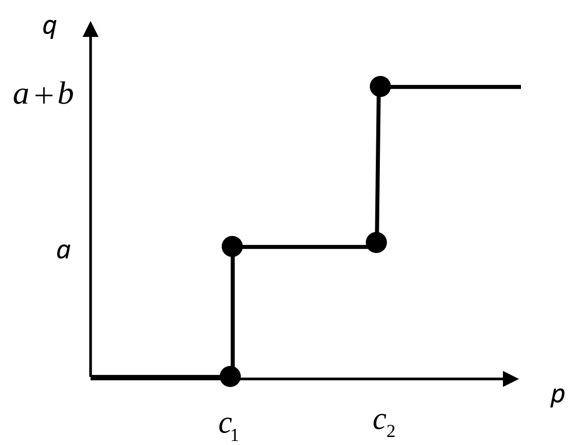


Figure A.2-8: Firm's supply

If, as depicted, the output price lies in the open interval (c_1, c_2), the vertical distance between the revenue line and the total cost function is greatest at $q = a$. That is, profit is maximized by fully utilizing the first machine. If the output price is below c_1 the profit maximizing output is zero. If the price is above c_2 it is profitable to fully utilize both machines and produce an output of $a+b$. Finally, if the price is c_1 or c_2, there is an interval of profit-maximizing outputs. The mapping from price to profit-maximizing output is depicted in Figure A.2-8.

In analyzing supply and demand, economists almost always follow Marshall and place the price on the vertical axis. This is depicted below. Note that the "inverse mapping" from *q* to *p* is also multi-valued. Also shown in the figure are four demand curves. While a formal mathematical analysis of demand shifts is somewhat more complex with a supply "correspondence" rather than a supply function, the graphical analysis of such shifts is essentially unaffected. If the demand curve shifts from D_0 to D_1, the market clearing price rises, but output remains at *a*. If the demand curve rises from D_2 to D_3, output increases but the market clearing price remains at c_2.

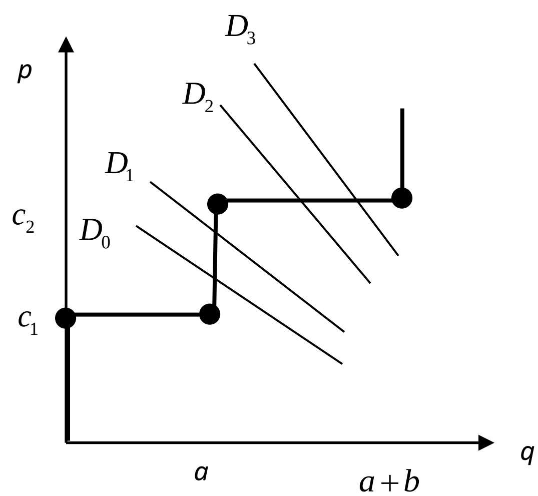


Figure A.2-9: Supply and Demand

Exercise A.2-1: Rules for limits

Suppose that the functions f_1 and f_2 have limits L_1 and L_2 at x^0.

(a) Show that the sum of the two functions $g = f_1 + f_2$ has a limit of $L_1 + L_2$.

(b) Show also that $g = f_1 f_2$ has a limit of $L_1 L_2$.

A.3 DERIVATIVES AND INTEGRALS

Key Ideas: rules of derivatives, increasing function, integral as anti-derivative

The slope of a function

Consider the function depicted below in Figure A.3-1. In graphical terms, the slope of the function at x^0 is the slope of the line just touching the graph of the function at $(x^0, f(x^0))$. This slope is called th4e derivative of the function art x^0. Analysis of models is further simplified if we assume that each function has an everywhere well-defined derivative and that this derivative varies continuously. Such functions are called "continuously differentiable." As a shorthand we write $f \in \mathbb{C}^1$.

Consider the function f depicted in Figure A.3-1. We now provide a formal definition of the derivative of the function at x^0. Consider the point x^0 and any neighboring point $x^0 + h$ and draw the chord AB connecting the two points on the curve. Then we define the slope of the curve at x^0 to be the limiting value of the slope of the chord as h approaches zero, that is, the slope of the tangent line AC.

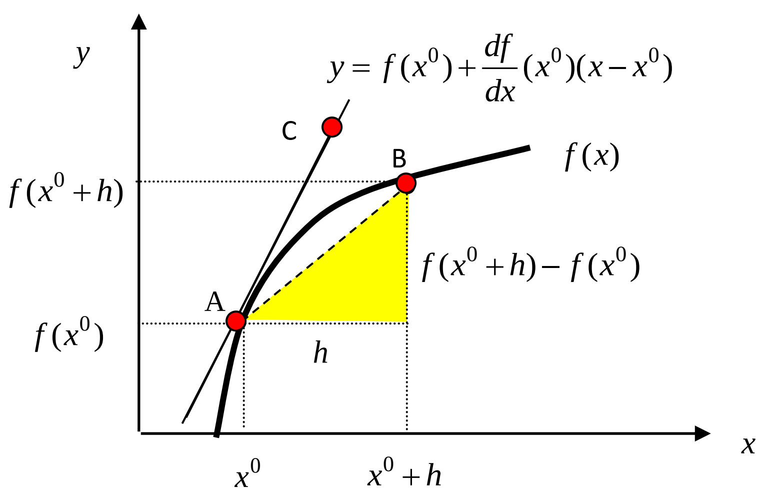


Figure A.3-1: Derivative of a function

Definition: Derivative of a function[4]

The derivative $\frac{df}{dx}(x^0)$ of the function *f* is the limit point at x^0 of the

ratio $g(x) = \frac{f(x) - f(x^0)}{x - x^0}$.

In the figure, the line through A with slope $\frac{df}{dx}(x^0)$ is also depicted. The equation of this line is

$$y = f(x^0) + \frac{df}{dx}(x^0)(x - x^0).$$

Because the line and curve take on the same value and have the same derivative at x^0, the line approximates the curve in the neighborhood of x^0. Because the linear approximation has the same first derivative it is also known as the first-order approximation of f at x^0.

Rules of differentiation

Appealing to the definition of a derivative as a limit, we have the following basic rules of differentiation.

Chain rule

Suppose that $y = f(x)$ and $z = g(y) = g(f(x))$. Then

$$\frac{d}{dx}g(f(x)) = \frac{dg}{dy}(y)\frac{dy}{dx}(x).$$

[4] It is very common to write the derivative of the function $f(x)$ as $f'(x)$ and in later chapters we will often follow this alternative convention. The disadvantage is that it does not naturally generalize to the case of two or more variables.

Product rule

$$\frac{d}{dx}[f(x)g(x)]=\frac{df}{dx}(x)g(x)+f(x)\frac{dg}{dx}(x).$$

Quotient rule

$$\frac{d}{dx}\frac{f(x)}{g(x)}=\frac{\frac{df}{dx}(x)g(x)-f(x)\frac{dg}{dx}(x)}{g(x)^2}.$$

Each of these properties is said to hold over the domain of a function if it holds at every point in the domain.

Often economic models include functions that everywhere increase (or decrease) with the underlying variable.

Definition: Increasing[5] and strictly increasing function

A function f is increasing at x^0 if $\frac{f(x)-f(x^0)}{x-x^0}\geq 0$ for all x in some deleted neighborhood of x^0.

A function *f* is strictly increasing at x^0 if $\frac{f(x)-f(x^0)}{x-x^0}>0$, for all x in some deleted neighborhood of x^0.

[5] We will make no distinction between "increasing" and "non-decreasing." Similarly, if $x_1 \geq 0$ we will say either that x_1 is positive or that x_1 is non-negative.

If a function f is strictly increasing, it is very tempting to conclude that the map of the function must have a strictly positive slope. Indeed if you were to draw some strictly increasing functions you might be able to convince yourself that this was "obvious." But while pictures are often extremely helpful, they are only stepping stones towards proofs. One way to prove that a statement is false is to find a counterexample.

Counterexample: cubic function

To see that the slope is not necessarily strictly positive, consider the cubic function $f(x) = x^3$. Then

$$\frac{f(z)-f(x)}{z-x} = \frac{z^3-x^3}{z-x} = z^2 + xz + x^2 = (z+\tfrac{1}{2}x)^2 + \tfrac{3}{4}x^2 . \qquad \text{(A.3-1)}$$

This is strictly positive for all x and $z \neq x$. Therefore the cubic is a strictly increasing function over the real line. However,

$$\frac{df}{dx}(x) = \lim_{z\to x}\frac{f(z)-f(x)}{z-x} = \lim_{z\to x}(z^2+zx+x^2) = 3x^2 .$$

Hence $\frac{df}{dx}(0) = 0$. Thus the slope is zero at $x = 0$. $\blacksquare$

Suppose that the function f is differentiable at x^0. We have just argued that the statement

$A = \{f$ is strictly increasing at $x^0\}$ does not imply the statement $B = \{\frac{df}{dx}(x^0) > 0\}$.

However the converse statement is true.[6]

[6] In mathematical shorthand, $A \not\Rightarrow B$ however $B \Rightarrow A$.

Proposition A.3-1: Sufficient condition for a function to be strictly increasing/decreasing at a point

If $\frac{df}{dx}(x^0) > 0$, then f is strictly increasing at x^0. If $\frac{df}{dx}(x^0) < 0$, then f is strictly decreasing at x^0.

Proof: Suppose that $\frac{df}{dx}(x^0) > 0$. Choose $\varepsilon = \frac{1}{2}\frac{df}{dx}(x^0)$. Because $\varepsilon > 0$, it follows from the definition of a derivative that there exists a $\delta > 0$ such that if $x^0 - \delta < x < x^0 + \delta$ and $x \neq x^0$, then

$$\varepsilon > \frac{f(x) - f(x^0)}{x - x^0} - \frac{df}{dx}(x^0) > -\varepsilon = -\frac{1}{2}\frac{df}{dx}(x^0).$$

Rearranging, it follows that if $x \in (x^0 - \delta, x^0 + \delta)$ and $x \neq x^0$, then

$$\frac{f(x) - f(x^0)}{x - x^0} > \frac{1}{2}\frac{df}{dx}(x^0).$$

Thus if $\frac{df}{dx}(x^0) > 0$, then $\frac{f(x) - f(x^0)}{x - x^0} > 0$ and so *f* is strictly increasing at x^0. A symmetric argument establishes that if the derivative is negative at x^0, f is strictly decreasing at x^0.

QED

Elasticity of a function

Economists often find it useful to examine proportional changes in a function $y = f(x)$ as the variable *x* changes. This is the elasticity of a function. For finite changes, the (arc) elasticity is the ratio of the proportional changes in *y* and *x*, that is

$$\frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}} = \frac{x}{y}\frac{\Delta y}{\Delta x} \quad .$$

Taking the limit as $\Delta x \to 0$, we define the (point) elasticity

$$\mathcal{E}(y,x)=\frac{x}{y}\frac{dy}{dx}.$$

Note that $\frac{d}{dx}\ln y=\frac{1}{y}\frac{dy}{dx}$. Therefore the point elasticity of demand can be written as follows.

$$\mathcal{E}(y,x)=x\frac{d}{dx}\ln y.$$

Example: Elasticity of demand

Suppose that demand for a product is $y=f(p)=ap^{-\nu}$. Taking the logarithm,

$$\ln y=\ln a-\nu\ln p.$$

Then $\frac{1}{y}\frac{dy}{dp}=\frac{-\nu}{p}$ and so $\mathcal{E}(y,p)=-\nu$.

Rules for Elasticities

Product Rule: $\mathcal{E}(fg,x)=\mathcal{E}(f,x)+\mathcal{E}(g,x)$.

Quotient Rule: $\mathcal{E}(f/g,x)=\mathcal{E}(f,x)-\mathcal{E}(g,x)$.

To derive the first rule note that $\mathcal{E}(fg,x)=x\frac{d}{dx}\ln fg$. Also $\ln fg=\ln f+\ln g$. Then

$$\mathcal{E}(fg,x)=x\frac{d}{dx}(\ln f+\ln g)=x\frac{d}{dx}\ln f+x\frac{d}{dx}\ln g=\mathcal{E}(f,x)+\mathcal{E}(g,x).$$

The derivation of the quotient rule is almost identical.

Integral of a function

Consider the continuous function f depicted below. The integral of a function over some interval $[\alpha,\beta]$ is the area under the curve. One way to compute this area is to divide the interval $[\alpha,\beta]$ into n equal segments of length $\delta x=\frac{\beta-\alpha}{n}$. Label the intermediate points $x_1,x_2,...,x_{n-1}$

and then add up the areas of each of the *n* bars depicted in the figure. That is, we compute the following summation.

$$I_n = \sum_{i=0}^{n-1} f(x^i)\delta x\,.$$

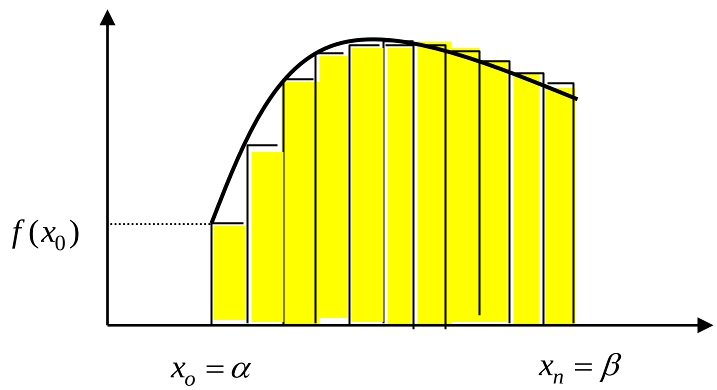


Figure A.3-2: Approximating the area under a curve

As *n* increases, the step function approximates the curve better and better. Therefore one way to determine the area under the curve is to take the limit of these approximations.

$$I = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(x^i)\delta x\,.$$

In mathematical terms, this is the integral of the function *f* over $[\alpha, \beta]$. Rather than write the expression as a limit, it is convenient to use the following short hand.

$$I = \int_{\alpha}^{\beta} f(x)dx\,.$$

In principal, for any particular function, it is possible to find an expression for this limit by following the above steps. However, in many cases, there is a much simpler way to proceed using the rules of differentiation.

Consider the change in the integral as we increase β to $\beta+h$. Because we are varying the right end-point we write the integral as $I(\beta)$. Then the change in the area under the curve is $I(\beta+h)-I(\beta)$. For concreteness, suppose that the function *f* is increasing, as depicted below.

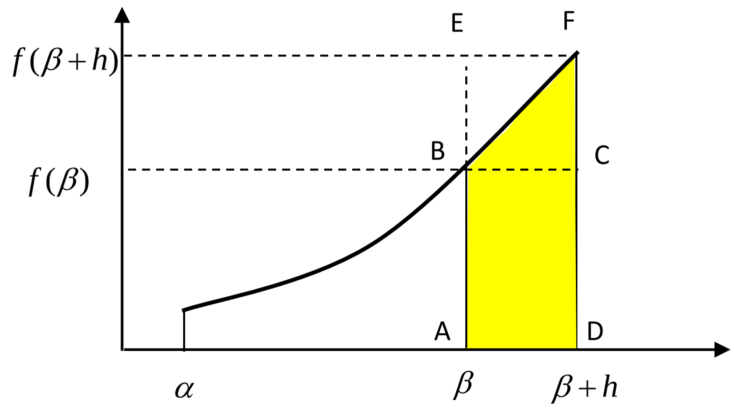


Figure A.3-3: Change in the area under the curve

Then the change in the area under the curve (the shaded area) is bounded from below by the area of the rectangle ABCD and bounded from above by the rectangle AEFD. That is,

$$f(\beta)h \le I(\beta+h)-I(\beta) \le f(\beta+h)h .$$

Dividing by *h*,

$$f(\beta) \le \frac{I(\beta+h)-I(\beta)}{h} \le f(\beta+h) .$$

Note that in the limit as $h \to 0$, the right-hand expression approaches $f(\beta)$. Also the limiting value of the middle term is the derivative of I. That is,

$$\frac{dI}{d\beta}(\beta) = f(\beta) .$$

Because this argument holds for any value β, we may write

$$\frac{dI}{dx}(x) = f(x).$$

Thus the integral of a function is the "anti-derivative" of the function.

Note that the anti-derivative is not unique. We can add any constant to it without affecting its derivative. For this reason, the anti-derivative is called the indefinite integral.

The area under the curve over an interval $[\alpha, \beta]$ is called the definite integral. For some constant *k*, the area is

$$\int_\alpha^\beta f(x)dx = I(\beta) + k.$$

Note that as $\beta \to \alpha$, the area under the curve must approach zero, that is $I(\alpha) + k = 0$. Then the definite integral

$$\int_\alpha^\beta f(x)dx = I(\beta) - I(\alpha).$$

Example: Integral as anti-derivative

Suppose that $f(x) = x^n$. Because the derivative of x^{n+1} is $(n+1)x^n$, the derivative of $\frac{1}{n+1}x^{n+1}$ is x^n. Then $\frac{1}{n+1}x^{n+1}$ is the anti-derivative of x^n and we may write

$$I(x) = \frac{1}{n+1}x^{n+1}.$$

The definite integral over the interval $[\alpha, \beta]$ is therefore $\frac{1}{n+1}(\beta^{n+1} - \alpha^{n+1})$. $\blacksquare$

One very useful trick for analyzing the integral of a function *f* is to see if it can be written in the form $f(x) = u(x)v(x)$ where the function $u(x)$ has a known integral $U(x)$. The integral of *f* can then be expressed as follows.

Proposition: A.3-2 Integration by parts

Suppose $f(x)=u(x)v(x)$ where $u(x)$ has integral $U(x)$. Then

$$\int_{\alpha}^{\beta} f(x)dx = U(\beta)v(\beta) - U(\alpha)v(\alpha) - \int_{\alpha}^{\beta} U(x)\frac{dv}{dx}dx .$$

Proof: By the Product rule,

$$\frac{d}{dx}U(x)v(x) = u(x)v(x) + U(x)\frac{dv}{dx}.$$

Rearranging this expression,

$$u(x)v(x) = \frac{d}{dx}U(x)v(x) - U(x)\frac{dv}{dx}.$$

The Proposition then follows by integrating both side and noting that the integral of the derivative of the function $U(x)v(x)$ is just the function.

QED

Exercise A.3-1 Rules of Differentiation

(a) Define $y(x)=f(x)g(x)$. Confirm that

$$\frac{y(x+h)-y(x)}{h} = f(x+h)\frac{g(x+h)-g(x)}{h} + g(x)\frac{f(x+h)-f(x)}{h}.$$

(b) Take the limit as $h \to 0$ and so sketch a proof of the Product rule

(c) Define $y = f(x)$ and $y+k = f(x+h)$. Confirm that

$$\frac{g(f(x+h))-g(f(x))}{h} = \frac{g(y+k)-g(y)}{k}\,\frac{f(x+h)-f(x)}{h}.$$

Hence sketch a proof of the Chain rule.

* ***Exercise A.3-2: Discontinuously differentiable function***[7]

Consider the following function.

$$f(x) = \begin{cases} x^2(\sin\frac{1}{x} - 1), & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

(a) Because $\frac{d}{dx}\sin x = \cos x$, apply the Chain and Product rules to show that for all $x \neq 0$,

$$\frac{df}{dx}(x) = 2x(\sin\frac{1}{x} - 1) - \cos\frac{1}{x}.$$

(b) Explain why the slope changes more and more rapidly with *x*, as *x* approaches zero.

(c) Explain why $-2x^2 \leq f(x) \leq 0$ and depict the map of the function in a neat figure. Appeal to the definition of the derivative as a limit to show that $\frac{df}{dx}(0) = 0$.

Remark: This is an example of a function that is everywhere differentiable but is not continuously differentiable at $x = 0$.

Exercise A.3-3 Integrating by parts

(a) Use the facts that $\ln x = u(x)\ln x$, where $u(x) = 1$ and $\frac{d}{dx}\ln x = \frac{1}{x}$, to show that

$$\int_\alpha^\beta \ln x dx = \beta(\ln\beta - 1) - \alpha(\ln\alpha - 1).$$

(b) Obtain an expression for $\int_\alpha^\beta x^\gamma \ln x dx, \ \gamma > 0$.

[7] An asterisk (*) indicates a somewhat harder question. This exercise is designed for students with strong mathematical backgrounds.

(c) The exponential function $f(x) = e^x$ has a derivative equal to the value of the function hence the indefinite integral $\int e^x dx = e^x$. Obtain an expression for $\int_\alpha^\beta xe^x dx$.

A.4 OPTIMIZATION

Key Ideas: first- and second- order conditions, approximating a function, elasticity

Optimizing behavior underlies almost all economic modeling. In explaining how a model works, it is often extremely helpful to present the implications of optimizing in a relatively informal manner. However, the model itself needs to be built on solid mathematical foundations. Suppose a downward shift in demand has left the profit function of a firm as depicted below. Profit $f(q)$ is a decreasing function and is negative for all strictly positive outputs. Thus the firm's profit maximizing strategy is to close down.

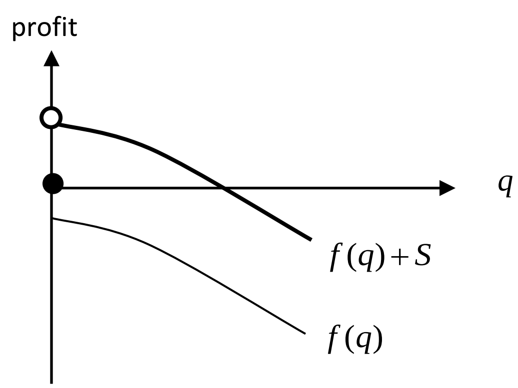


Figure A.4-1: Discontinuous profit function

To help the domestic industry, the government intervenes and announces that every firm that sells this commodity will receive a subsidy S. As a result, the profit function is

$$\Pi(q)=\begin{cases} 0 & ,\ q=0 \\ f(q)+S, & q>0 \end{cases}.$$

From the figure, the smaller the output of the firm, the greater is the firm's profit. However, if the firm produces nothing, its profit is zero. As a practical matter the firm's output will be "very small." But from a mathematical perspective, there is no solution to the following maximization problem

$$\underset{q}{Max}\{\Pi(q) \,|\, q \geq 0\},$$

so there is no profit-maximizing output.

Intuitively, we can be assured that there is a solution to an optimization problem if we are willing to assume that the function to be optimized is continuous over some interval $[a,b]$. This intuition is confirmed by the following fundamental theorem.

Extreme Value Theorem

If the function *f* is continuous over a closed interval, then *f* attains its maximizing and minimizing value at some numbers in this interval.

We now seek to characterize the maximizing value of a differentiable (and hence continuous) function. Suppose that $D_f = \mathbb{R}$ and consider any point x^0 where the slope is strictly positive. By Proposition A.3-1, if $\frac{df}{dx}(x^0) > 0$, *f* is strictly increasing at x^0. Thus there exists $x^1 < x^0$ such that $f(x^1) < f(x^0)$ and $x^2 > x^0$ such that $f(x^2) > f(x^0)$. Thus *f* does not have maximum or a minimum at x^0. A symmetric argument applies if $\frac{df}{dx}(x^0) < 0$. We have therefore proved the following proposition.

Proposition A.4-1: First Order Condition (FOC)

Suppose that f takes on its maximum value over $\mathbb{R}$ at x^0. If f is differentiable at x^0 then $\frac{df}{dx}(x^0) = 0$.

In economics it is often natural to restrict the domain of a function. In particular, many variables (icecream, hours worked etc.) only have meaning as positive (equivalently, non-negative) numbers. The FOC must be modified to reflect this. Consider the following modified problem

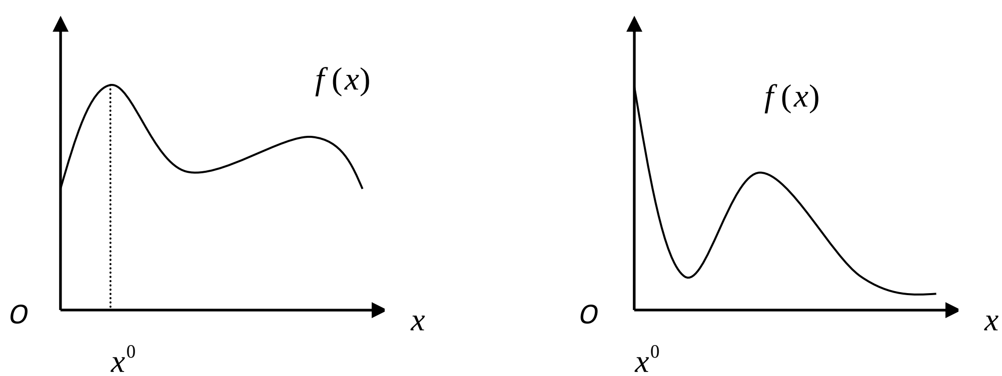
$$\underset{x}{Max}\{ f(x) \mid x \in \mathbb{R}^{+} \}.$$

Here we have used the shorthand $\mathbb{R}^{+}$ to denote the positive real numbers. If the maximum occurs at a point x^0 in the interior of the domain, the FOC is unchanged. If the maximum occurs at the endpoint, $x^0 = 0$ then it continues to be the case that f cannot be strictly increasing at $x = 0$. However, as depicted in Figure A.4-2, f may be strictly decreasing.

The modified FOC then has two parts:

(i) If $x^0 > 0$ then $\dfrac{df}{dx}(x^0) = 0$.

(ii) If $x^0 = 0$ then $\dfrac{df}{dx}(x^0) \le 0$.



Interior Solution Boundary Solution

Figure A.4-2: FOC

This is typically written more compactly as follows:

$$\frac{df}{dx}(x^0) \le 0, \text{ with equality if } x^0 > 0 .$$

Higher-Order Derivatives

If the function f is differentiable over D_f then the derivative is itself a function. Thus, appealing to the definition of a derivative as a limit, we can define the derivative of $\frac{df}{dx}(x)$ at x^0. Because this is the derivative of the derivative of the function f, it is called the second derivative of f. More generally we define $\frac{d^n f}{dx^n}(x^0)$ to be the *n*th derivative of f at x^0.

Proposition A.4-2: Second-Order Condition

Suppose that f is differentiable on $\mathbb{R}$ and is twice differentiable at x^0. If f takes on its maximum value at x^0 then $\frac{d^2 f}{dx^2}(x^0) \le 0$.

Proof: Suppose the proposition is false, that is, f takes on its maximum value at x^0 so that $\frac{df}{dx}(x^0) = 0$ and $\frac{d^2 f}{dx^2}(x^0) = \frac{d}{dx}(\frac{df}{dx})(x^0) > 0$. By Proposition A.3-1, if the function g has a strictly positive derivative at x^0 then g is strictly increasing at x^0. Then, because $\frac{d}{dx}(\frac{df}{dx})(x^0) > 0$, $\frac{df}{dx}(x)$ is strictly increasing at x^0. That is, in some deleted $\delta-$neighborhood of x^0,

$$\frac{\frac{df}{dx}(x) - \frac{df}{dx}(x^0)}{x - x^0} > 0 .$$

Hence $\frac{df}{dx}(x) > 0$ over some interval $(x^0, x^0 + \delta)$. Again appealing to Proposition A.3-1, f is strictly increasing at each point in the interval $(x^0, x^0 + \delta)$ and hence in the interval $(x^0, x^1]$ where $x^1 = x^0 + \frac{1}{2}\delta$.

Because f takes on its maximum at x^0, $f(x^1) \le f(x^0)$. Moreover, because f is strictly increasing at x^1 there exists $\hat{x} \in (x^0, x^1)$ such that $f(\hat{x}) < f(x^1)$. Appealing to the Extreme Value Theorem, f must take on its minimum at some point $c \in (x^0, x^1)$. From the FOC, $\frac{df}{dx}(c) = 0$. But this is impossible because we have already argued that $\frac{df}{dx}(x) > 0$ over some interval $(x^0, x^0 + \delta)$. Then it cannot be the case that $\frac{d^2 f}{dx^2}(x^0) > 0$ after all.

QED.

The geometry of the proof is depicted below.

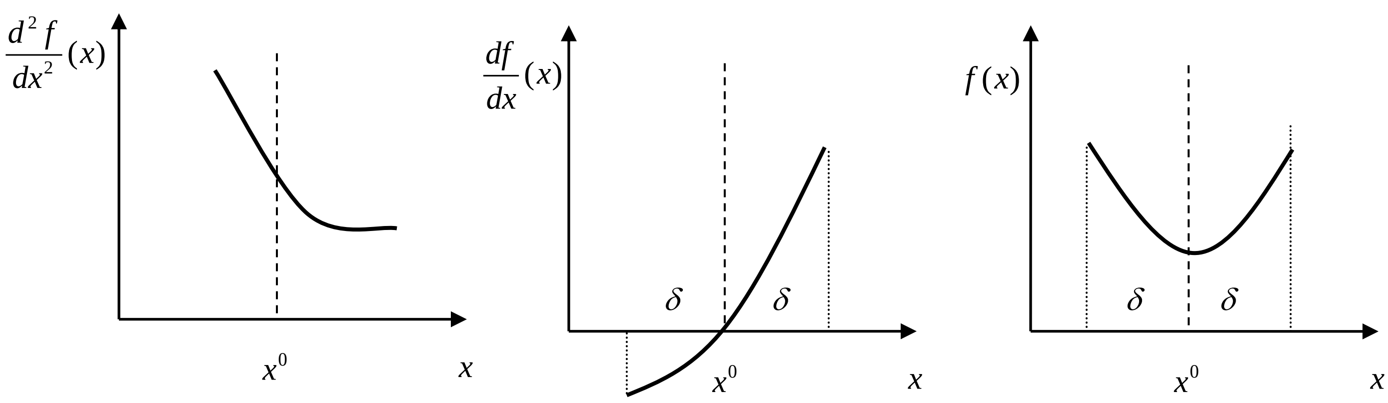


Figure A.4-3: Second-Order Condition

Suppose that the second derivative is positive at x^0, so that the slope of the function is strictly increasing over some interval $(x^0-\delta, x^0+\delta)$. Also, the slope is zero at x^0. Thus the slope is strictly negative for $x \in (x^0-\delta, x^0)$ and strictly positive for $x \in (x^0, x^0+\delta)$. This is depicted in the middle diagram. It follows that *f* must have a local minimum at x^0.

Finally we note that many of the implications of optimization depend only on first and second derivatives of functions. To understand these results it is usually enough to consider a quadratic approximation of the function.

Approximating a function

Suppose we approximate a function in the neighborhood of x^0 using the following quadratic function:

$$f_a(x) = \alpha_0 + \alpha_1(x-x^0) + \alpha_2(x-x^0)^2 .$$

Setting $x = x^0$, $f_a(x^0) = \alpha_0$. Thus the approximating function takes on the same value at x^0 if $\alpha_0 = f(x^0)$. Taking the derivative,

$$\frac{df_a}{dx}(x) = \alpha_1 + 2\alpha_2(x-x^0) .$$

Thus the approximating function has the same slope if $\alpha_1 = \frac{df}{dx}(x^0)$.

Differentiating again,

$$\frac{d^2 f_a}{dx^2}(x^0) = 2\alpha_2 .$$

Thus the approximating function has the same second derivative at x^0 if $\alpha_2 = \frac{1}{2}\frac{d^2 f}{dx^2}(x^0)$.

Collecting these results, we can write the quadratic (or "second order") approximation of the function f at x^0 as follows:

$$f_a(x) = f(x^0) + \frac{df}{dx}(x^0)(x - x^0) + \frac{1}{2}\frac{d^2 f}{dx^2}(x^0)(x - x^0)^2 .$$

Arguing in exactly the same manner we can define higher order approximations as well. However it is the first- and second-order approximations that typically prove most useful.

Consider once more the optimization problem $\underset{x}{Max}\{f(x) \mid x \in \mathbb{R}\}$. If f takes on its maximum value at x^0 then we can appeal to the FOC and thus write the second-order approximation as follows:

$$f_a(x) = f(x^0) + \frac{1}{2}\frac{d^2 f}{dx^2}(x^0)(x - x^0)^2 .$$

The three possible approximating functions are depicted below. In the first two cases, the approximating function takes on its maximum at x^0. In the third case the approximating function takes on its minimum at x^0. As we have shown, if case three holds then the original function f cannot have a maximum at x^0.

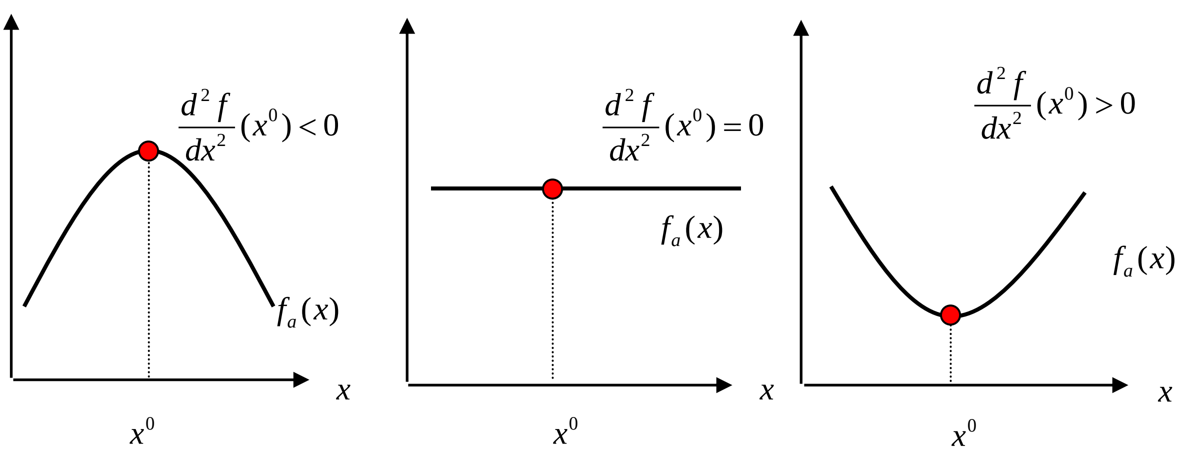


Figure A.4-4: Approximating quadratic functions

Elasticity of a function

Suppose that the function $x(p)$ represents the solution to some optimization problem given the parameter *p*. In economics it is often very useful to summarize the sensitivity of the response function in proportional terms. Let the initial price be p^0 and the final price be p^1. Define $x^i = x(p^i)$, $i = 0,1$. Then the proportional changes as *p* changes from p^0 to p^1 are

$$\frac{\Delta p}{p^0} = \frac{p^1 - p^0}{p^0} \text{ and } \frac{\Delta x}{x^0} = \frac{x^1 - x^0}{x^0}.$$

We define the arc elasticity of *x* with respect to *p* to be the ratio of these changes, that is

$$\frac{\frac{\Delta x}{x^0}}{\frac{\Delta x}{x^0}} = \frac{p^0}{x^0}\frac{\Delta x}{\Delta p}.$$

Taking the limit, the point elasticity of *x* at p^0 is

$$\mathcal{E}(x, p^0) = \frac{p^0}{x(p^0)}\frac{dx}{dp}(p^0).$$

Because $\frac{d}{dx}\ln x = \frac{1}{x}$, it is sometimes helpful to write point elasticity as follows:

$$\mathcal{E}(x, p) = p\frac{d}{dp}\ln x. \qquad \text{(A.4-1)}$$

Point elasticities have very nice properties. For example, the elasticity of a product is the sum of the elasticities

$$\mathcal{E}(xy, p) = \mathcal{E}(x, p) + \mathcal{E}(y, p).$$

This follows from the properties of the logarithmic function. Appealing to (A.4-1),

$$\mathcal{E}(xy,p) = p\frac{d}{dp}\ln xy$$

$$= p\frac{d}{dp}\ln x + p\frac{d}{dp}\ln y$$

$$= \mathcal{E}(x,p) + \mathcal{E}(y,p)\,.$$

Exercise A.4-1: First-order conditions

Solve the following problem and provide conditions under which $x^* = 0$ and $x^* = b$.

$$\underset{x}{Max}\{\, f(x) = \alpha + \beta x - x^2 \mid x \in [0,b]\}\,.$$

Exercise A.4-2: Consumer choice

Suppose that a consumer can purchase units of commodity x at a price of p and commodity y at a price of 1. His utility function is $u(x,y) = B(x) + y$, where B is concave. If he purchases x units and has an income of I, his consumption of the other commodity is $y = I - px$. We can therefore substitute for y in his utility function and write

$$U(x) = B(x) + I - px\,.$$

The consumer cannot spend more than his income on x so $0 \le x \le \frac{I}{p}$.

(a) Write down the FOC for this example by completing the following statements.

If $x^* \in (0,\frac{I}{p})$ then…..If $x^* = 0$ then…. If $x^* = \frac{I}{p}$ then….

(b) Solve for the consumer's utility maximizing choice of x if (i) $B(x) = 10\ln x$ and if (ii) $B(x) = 10\ln(5 + x)$.

Exercise A.4-3: Robinson Crusoe

Robinson has an endowment of ω units of coconuts. He invests z units in order to produce coconuts next period and consumes $x = \omega - z$ this period. Output of next year's coconuts is $y = a\sqrt{z}$. Robinson's two period preferences are represented by the utility function $u(x, y) = x + b\ln y$. The parameters a and b are both positive.

(a) Express Robinson's utility as a function of his input choice z.

(b) Under what conditions will Robinson invest less than his endowment (and so enjoy some current consumption)? If this condition is satisfied how much will he invest?

(c) Under what conditions (if any) will Robinson invest nothing?

Exercise A.4-4: Elasticities

(a) Show that $\mathcal{E}(ax, by) = \mathcal{E}(x, y)$, where a and b are parameters.

(b) Show that the elasticity of the ratio of x and y is the difference in elasticities of x and y.

(c) Show that $\mathcal{E}(x, y)\mathcal{E}(y, z) = \mathcal{E}(x, z)$.

(d) Hence or otherwise show that $\mathcal{E}(y, \frac{1}{x}) = -\mathcal{E}(y, x)$.

(e) Show that $\mathcal{E}(\frac{1}{y}, \frac{1}{x}) = \mathcal{E}(x, y)$.

A.5 SUFFICIENT CONDITIONS FOR A MAXIMUM

Key ideas: concave and convex functions, quasi-concave function

Characterizing the solution to a maximization problem is especially easy if all we have to do is examine the FOC. This is the case if a function is concave.

Definition: Concave and Strictly Concave Function

A function *f* is concave on the interval $[a,b]$ if, for any points x^0 and x^1 in this interval,

and any convex combination $x^\lambda = (1-\lambda)x^0 + \lambda x^1,\ 0 < \lambda < 1$,

$$f(x^\lambda) \geq (1-\lambda)f(x^0) + \lambda f(x^1).$$

The function is strictly concave if the inequality is always strict.

In geometrical terms, the graph of the function *f* between any two points x^0 and x^1 lies above the chord connecting the points $(x^0, f(x^0))$ and $(x^1, f(x^1))$.

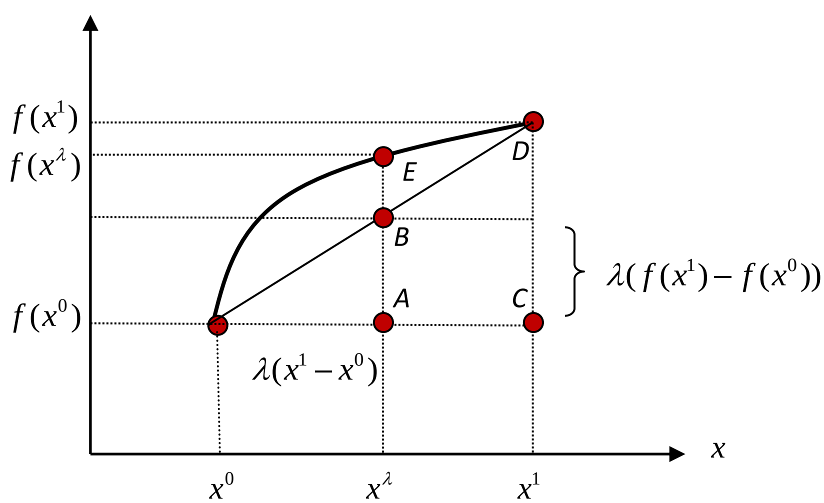


Figure A.5-1: Concave function

To see that this is the case, note that

$$x^{\lambda} - x^{0} = \lambda(x^{1} - x^{0}). \qquad \text{(A.5-1)}$$

That is, the distance between x^0 and the convex combination x^{λ} is a fraction λ of the distance between x^0 and x^1. Then the vertical distance AB is a fraction λ of the vertical distance *CD*, so $AB = \lambda(f(x^1) - f(x^0))$. If *f* is concave, $f(x^{\lambda}) \geq (1-\lambda)f(x^0) + \lambda f(x^1)$ so

$$f(x^{\lambda}) - f(x^{0}) \geq \lambda(f(x^{1}) - f(x^{0})). \qquad \text{(A.5-2)}$$

Thus $AE \geq AB$.

If the function is differentiable, the tangent line at any point x^0 must lie above the graph of the function. This is depicted below.

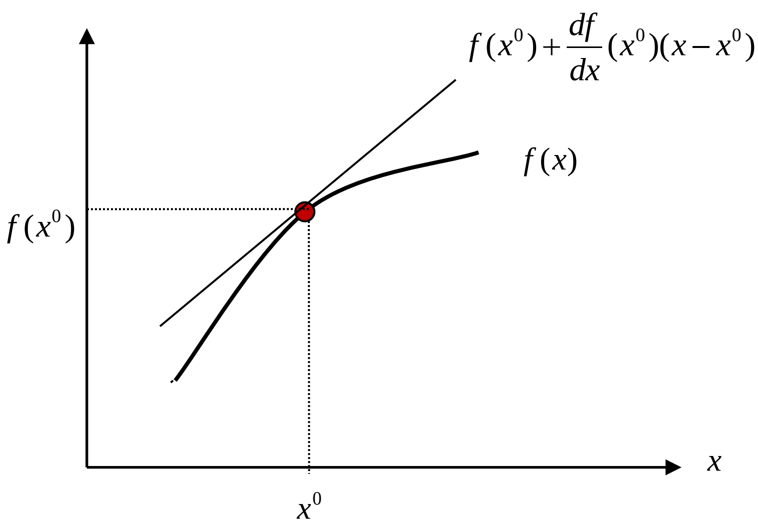


Figure A.5-2 Tangent line

Formally we have the following proposition.

Proposition A.5-1: If *f* is concave and differentiable at x^0 and x^1, then

$$\frac{df}{dx}(x^0)(x_1 - x_0) \geq f(x^1) - f(x^0) \geq \frac{df}{dx}(x^1)(x_1 - x_0).$$

Proof: Appealing to equation (A.5-1) $\lambda = \dfrac{x^\lambda - x^0}{x^1 - x^0}$. Substituting for λ in inequality (A.5-2),

$f(x^\lambda) - f(x^0) \geq (\dfrac{x^\lambda - x^0}{x^1 - x^0})(f(x^1) - f(x^0))$. Rearranging this inequality,

$$\frac{f(x^\lambda) - f(x^0)}{x^\lambda - x^0} x^1 - x^0 \geq f(x^1) - f(x^0).$$

The first inequality then follows by taking the limit as $\lambda \to 0$.

Similarly $x^1 - x^\lambda = (1-\lambda)(x^1 - x^0)$ and, from the definition of a concave function,

$$f(x^1) - f(x^\lambda) \leq (1-\lambda)(f(x^1) - f(x^0)).$$

Substituting for $(1-\lambda)$ from the previous expression,

$$f(x^1) - f(x^0) \geq \frac{f(x^1) - f(x^\lambda)}{x^1 - x^\lambda}(x^1 - x^0).$$

The second inequality then follows by taking the limit as $\lambda \to 1$.

QED

Note that an immediate implication is that if $x^1 > x^0$, then the slope of the function is lower at x^1. For a strictly concave function we now show that the slope is strictly lower.

Corollary A.5-2: If f is strictly concave on some interval X and is differentiable at $x^0 \in X$ then

$$\frac{df}{dx}(x^0)(x^1 - x^0) > f(x^1) - f(x^0).$$

Moreover, if f is also differentiable at $x^1 > x^0$, then $\frac{df}{dx}(x^0) > \frac{df}{dx}(x^1)$.

Proof: Because Proposition A.5-1 holds for any $x^1 \neq x^0$, the first inequality holds for all convex combinations, that is

$$\frac{df}{dx}(x^0)(x^\lambda - x^0) \geq f(x^\lambda) - f(x^0).$$

If f is strictly concave, $f(x^\lambda) - f(x^0) > \lambda(f(x^1) - f(x^0))$. Combining these two inequalities,

$$\frac{df}{dx}(x^0)(x^\lambda - x^0) > \lambda(f(x^1) - f(x^0)).$$

Substituting for $x^\lambda - x^0$, by appealing to (A.5-1), it follows that

$$\frac{df}{dx}(x^0)(x^1 - x^0) > f(x^1) - f(x^0).$$

Thus if f is differentiable at x^0 and x^1, we can appeal to the second inequality of Proposition A.5-1 to conclude that

$$\frac{df}{dx}(x^0)(x^1 - x^0) > \frac{df}{dx}(x^1)(x^1 - x^0).$$

QED

A further implication of Proposition A.5-1 is that if a concave function is everywhere differentiable, all the tangent lines lie above the graph of the function. We now show that the converse is also true.

Proposition A.5-3: A differentiable function f is concave on the interval X if and only if for any x^0 and $x^1 \in X$,

$$f(x^1) \leq f(x^0) + \frac{df}{dx}(x^0)(x^1 - x^0). \qquad \text{(A.5-3)}$$

Proof: We have already established necessity. To demonstrate sufficiency, consider any x^0, x^1 and convex combination $x^\lambda = (1-\lambda)x^0 + \lambda x^1$. Appealing to (A.5-3),

$$f(x^0) \le f(x^\lambda) + \frac{df}{dx}(x^\lambda)(x^0 - x^\lambda)$$

and

$$f(x^1) \le f(x^\lambda) + \frac{df}{dx}(x^\lambda)(x^1 - x^\lambda).$$

Multiplying the first inequality by $(1-\lambda)$, the second by λ and then adding the two inequalities, it follows that

$$(1-\lambda)f(x^0) + \lambda f(x^1) \le f(x^\lambda).$$

We have also seen that if a function is concave and differentiable, the slope of the function declines as x increases. Thus if a concave function is twice differentiable, the second derivative must be negative. The converse of this statement is also true.

Proposition A.5-4: A twice differentiable function f is concave on the interval X if and only if, for all $x \in X$, $\frac{d^2 f}{dx^2}(x) \le 0$.

Proof: To demonstrate sufficiency, define $g(\lambda) = f(x^\lambda) = f(x^0 + \lambda(x^1 - x^0))$ then

$$\frac{dg}{d\lambda}(\lambda) = (x^1 - x^0)\frac{df}{dx}(x^\lambda) \text{ and so } \frac{d^2 g}{d\lambda^2}(\lambda) = (x^1 - x^0)^2 \frac{d^2 f}{dx^2}(x^\lambda).$$

Because the second derivative of f is negative, it follows that $\frac{dg}{d\lambda}(\lambda)$ is decreasing over the interval $(0,1)$.

Then

$$f(x^1) - f(x^0) = g(1) - g(0) = \int_0^1 \frac{dg}{d\lambda}(\lambda) d\lambda$$

$$\leq \int_0^1 \frac{dg}{d\lambda}(0) d\lambda \text{, because the derivative is everywhere decreasing.}$$

Taking $\frac{dg}{d\lambda}(0)$ outside the integral, and noting that $\frac{dg}{d\lambda}(0) = (x^1 - x^0)\frac{df}{dx}(x^0)$, it follows that

$$f(x^1) - f(x^0) \leq \frac{df}{dx}(x^0)(x^1 - x^0).$$

QED

Maximization

Suppose that the function *f* is differentiable on $\mathbb{R}$ and there is a point x^0 satisfying the FOC $\frac{df}{dx}(x^0) = 0$. While this might be a maximum there are three possibilities, all depicted below.

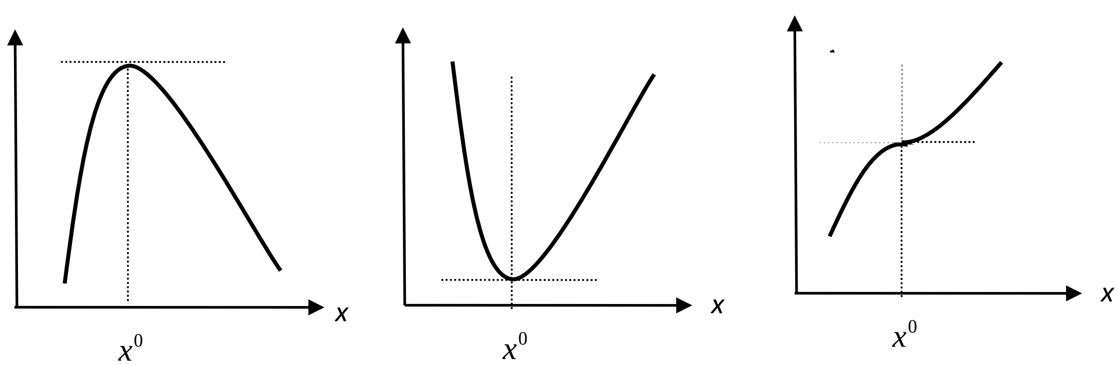


Figure A.5-2: (a) Maximum (b) Minimum (c) Point of Inflection

However, if f is concave, then it is maximized at x^0. This follows from Proposition A.5-4. Because $\frac{df}{dx}(x^0)=0$, for any $x^1 \neq x^0$, $f(x^1) \leq f(x^0) + \frac{df}{dx}(x^0)(x^1 - x^0) = f(x^0)$.

We have therefore proved the following result.

Proposition A.5-5: Sufficient Conditions for a Maximum

Suppose f is concave on $[a,b]$ and for some $x^0 \in (a,b)$, $\frac{df}{dx}(x^0)=0$. Then

$$f(x) \leq f(x^0), \text{ for all } x \in [a,b]$$

While concavity assures that the FOC is both necessary and sufficient for a maximum, this is an undesirably strong assumption for many economic applications.

Definition: Local property of a function

A function f satisfies the property P locally at x^0 if the property P holds in some neighborhood of x^0.

It follows from Corollary A.5-2 that if the FOC holds at x^0 and the function is locally strictly concave, then $f(x) < f(x^0)$ for all $x \neq x^0$ in some neighborhood of x^0. In this case f is said to have a local maximum at x^0.

A local maximum must be the "global" maximum if the function only has one turning point. This is the case if the function is "quasi-concave."

Definition: Quasi-concave function

The function f is quasi-concave over the interval $[a,b]$ if, for any x^0 and x^1 in this interval such that $f(x^1) \geq f(x^0)$,

$f(x^{\lambda}) \geq f(x^0)$ where $x^{\lambda} = (1-\lambda)x^0 + \lambda x^1$ and $\lambda \in (0,1)$.

Proposition A.5-6: Quasi-concavity and Sufficient Conditions for a Maximum

If $\frac{df}{dx}(x^0) = 0$, f is locally strictly concave at $x^0 \in [a,b]$ then f has a local maximum at x^0. In addition, if f is quasi-concave on this interval, then $f(x) < f(x^0)$ for all $x \in [a,b]$.

Proof: We have just argued that the first statement follows from Corollary A.5-2. That is, there is some neighborhood $N(x^0, \delta)$ within which

$$f(x) < f(x^0), \text{ for all } x \neq x^0. \qquad \text{(A.5-4)}$$

Suppose that there is some point x^1 such that $f(x^1) > f(x^0)$. If f is quasi-concave, then for all convex combinations, $f(x^{\lambda}) \geq f(x^0)$. In particular, this must be true in $N(x, \delta)$. But this contradicts (A.5-4).

QED

Minimization

We can appeal to the sufficient conditions for a maximum to obtain similar sufficient conditions for the FOC to yield the minimum. We simply note that if f takes on its minimum at x^0, so that $f(x) \geq f(x^0)$, then $-f(x) \leq -f(x^0)$. Thus one sufficient condition is that –f must be concave. That is, for any x^0 and $x^1 \neq x^0$, and any convex combination x^{λ},

$$-f(x^{\lambda}) \geq (1-\lambda)(-f(x^0)) + \lambda(-f(x^1)).$$

Multiplying the inequality by –1 changes the sign so

$$f(x^{\lambda}) \leq (1-\lambda) f(x^0) + \lambda f(x^1).$$

This is depicted in the figure below. The map of the function lies everywhere below the chord. Such functions are called convex functions.

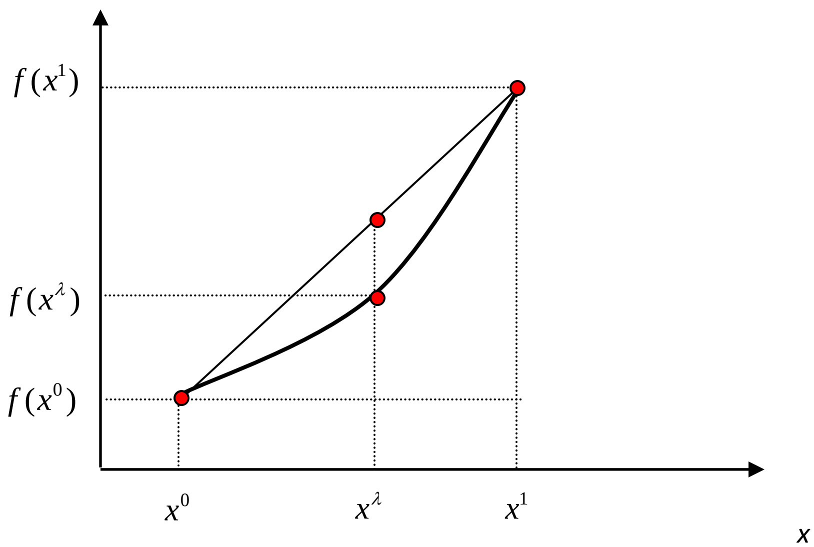


Figure A.5-3: Convex Function

Definition : Convex and Strictly Convex Function

A function f is convex on the interval $[a,b]$ if for any points x^0, x^1 in this interval

and any convex combination $x^\lambda = (1-\lambda)x^0 + \lambda x^1,\ 0 < \lambda < 1$,

$$f(x^\lambda) \leq (1-\lambda) f(x^0) + \lambda f(x^1).$$

The function is strictly convex if the inequality is always strict.

Proposition A.5-7: Sufficient Conditions for a Minimum

Suppose $D_f = \mathbb{R}$ and f is convex. If the FOC holds at x^0 then

$$f(x) \geq f(x^0) \text{ for all } x \in \mathbb{R}.$$

While this condition is stated for completeness, it is certainly not important to remember it. For, as we have just argued, we can always convert any minimization problem into a maximization problem. This is the approach that we will usually adopt.

Exercise A.5-1: Profit maximizing firm

A firm hiring x units of labor can produce an output $q(x) = 10x^{\frac{1}{2}}$. The price of output is p and the wage rate is w.

(a) Show that the profit function of the firm $f(x) = pq(x) - wx$ is concave.

(b) Solve for the profit maximizing labor demand and supply of the firm as a function of the wage and price.

Exercise A.5-2: Cost minimizing inputs

A firm has a production function $q = 4K + \sqrt{L}$. The wage rate is w and the cost of renting capital equipment is r.

(a) Write down an expression for the capital equipment requirements if L units of labor are hired and the objective is to produce q units of output. Hence obtain an expression for total cost $C(L)$ as a function of L and the three parameters, w, r and q. Explain why $L \in [0, q^2]$.

(b) Show that $-C(L)$ is a concave function of L.

(c) Use the first order condition for a maximum to solve for the optimal demand for labor.

(d) For what parameter values is the demand for both labor and capital strictly positive?

Exercise A.5-3: Properties of concave and quasi-concave functions

Use the primary definition of concave and quasi-concave functions to establish the following:

(a) The sum of n concave functions is concave. (Start with $n = 2$.)

(b) An increasing function of a quasi-concave function is quasi-concave.

(c) An increasing concave function of a concave function is concave.

(d) If f is concave, show that it is quasi-concave.

Exercise A.5-4: Family of concave functions

A widely used family of utility functions satisfy the following condition:

$$\frac{d}{dx}\ln\frac{dU}{dx}=\frac{\dfrac{d^2U}{dx^2}}{\dfrac{dU}{dx}}=-\frac{1}{a+bx} \text{ where } a\geq 0,\; b\geq 0\,.$$

In each of the following cases, integrate to obtain the implied utility function (i) $a>0=b$, (ii) $b=1$, and (iii) $b>0,\; b\neq 1$.

Exercise A.5-5: Quasi-concavity

(a) Depict some quasi-concave functions.

(b) Draw the following function in a neat figure.

$$f(x)=\begin{cases} 8-(x-2)^2, & x<2 \\ 8 \quad , & 2\leq x\leq 3. \\ 8+(x-3)^3, & x>3 \end{cases}$$

Confirm that f is concave on $[2,3]$ and is quasi-concave on $\mathbb{R}^+$.

(c) For what values of *x* are the first- and second-order necessary conditions satisfied? Show that the sufficient conditions for a local maximum fail at each of these points.

INDEX

P

Q

S