INSEIKAI Tohoku Spring Camp 2023
Mathematics

TRAN Quang-Thanh (Tedd) [1]

ZHANG Ye (Tony) [2]

Graduate School of Economics and Management
Tohoku University

March 12, 2023

[1] tran.quang.thanh.p1@dc.tohoku.ac.jp
[2] zhang.ye.q8@dc.tohoku.ac.jp

Contents

0.1 Introduction

The course aims to equip learners with the essential mathematical tools to research economics. In that regard, introductory algebra, calculus, matrix algebra, and optimization techniques are the core concepts that every economics student must master, and we must abstract away from advanced topics such as proofs, or control theory. We also try to blend in as many applications of mathematics and programming in economics as possible in the last sessions.

0.1.1 Syllabus

Textbooks used: Simon and Blume (1994), Sydsæter and Hammond (2008), Sydsæter et al. (2008).

(1) Algebra	(2) Equations
(3) Functions of One Variable	(4) Properties of Functions
(5) Differentiation	(6) Differentiation in Use
(7) Single-Variable Optimization	(8) Integration
(9) Functions of Many Variables	(10) Matrix Algebra
(11) Unconstrained Optimization	(12) Constrained Optimization
(13) Applications in Economics	(14) DIY Projects Report

Table 1: Syllabus of the course

0.1.2 Schedule

	Monday	Tuesday	Wednesday	Thursday	Friday
March	6	7	8	9	10
	L(1)	L(2)	L(3)	L(4)	L(5)
March	13	14	15	16	17
	L(6)	L(7)	L(8)	L(9)	L(10)
March	20	21	22	23	24
	L(11)	L(12)	L(12)	L(13)	L(14)
Class hours	Morning	$10:30 - 12:00$			
	Afternoon	$13:30 - 15:00$			

Table 2: Schedule of the course

Greek Letter	LaTeX Code	Greek Letter	LaTeX Code	Greek Letter	LaTeX Code
α	\alpha	β	\beta	γ	\gamma
δ	\delta	ϵ	\epsilon	ζ	\zeta
η	\eta	θ	\theta	ι	\iota
κ	\kappa	λ	\lambda	μ	\mu
ν	\nu	ξ	\xi		
π	\pi	ρ	\rho	σ	\sigma
τ	\tau	υ	\upsilon	ϕ	\phi
χ	\chi	ψ	\psi	ω	\omega

Chapter 1

Algebra

1.1 Exercises

Ex. 1.1 (Expand and simplify).

(a) $\dfrac{24x^3y^2z^3}{4x^2yz^2}$,

(b) $[-(-ab^3)^{-3}(a^6b^6)^2]^3$,

(c) $\dfrac{a^5a^3a^{-2}}{a^{-3}a^6}$,

(d) $\left[\left(\dfrac{x}{2}\right)^3 \cdot \dfrac{8}{x^{-2}}\right]^{-3}$,

(e) $\dfrac{a(a-1)}{b} + \dfrac{(b+1)(b-1)}{a}$,

(f) $(x-3)(x+7)$,

(g) $-\sqrt{3}(\sqrt{3}-\sqrt{6})$,

(h) $\dfrac{p^\gamma(pq)^\sigma}{p^{2\gamma+\sigma}q^{\sigma-2}}$

(i) $\dfrac{2+a}{a^2b} + \dfrac{1-b}{ab^2} - \dfrac{2b}{a^2b^2}$

(k) $\dfrac{x^0x^{-\beta\alpha}}{x^{2\alpha+1}x^{\beta-2}}$

Ex. 1.2 (Expand). Prove the Quadratic Identities

$$(a)\ (a+b)^2 = a^2 + 2ab + b^2,$$
$$(b)\ (a-b)^2 = a^2 - 2ab + b^2$$
$$(c)\ a^2 - b^2 = (a+b)(a-b)$$
$$(e)\ (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3,$$
$$(f)\ (a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3,$$
$$(g)\ a^3 + b^3 = (a+b)(a^2 - ab + b^2),$$
$$(h)\ a^3 - b^3 = (a-b)(a^2 + ab + b^2)$$

Ex. 1.3. The surface area of a sphere with a radius r is

$$S = 4\pi r^2.$$

(a) If the radius is tripled, how much will the area increase?
(b) If the radius increases by 16%, how many % will the surface area increase?
(c) How much must the radius increase for the surface area to increase by 1.5?

Ex.   1.4 (Compound Interest)**.** You have in total $A = \$1000$. The annual interest rate is $i = 5\%$ each year. At time t, you deposit everything at a bank. Answer the following:

(a) How much will you have in your account at time $t + 1$?

(b) How much will you have in your account at time $t + 2$?

(b) How much will you have in your account at time $t + s$? (where s is just a number)

Now, assume that the annual inflation rate is $\pi = 8\%$. You still choose to keep your money at the bank. Calculate:

(c) The real interest rate r. (Hint: it is $i - \pi$)

(d) How much will you have in your account at time $t+1$? at $t+2$? at $t+s$?

(e) Assume that inflation only happens at $t + 3$ and stays that way from that moment onward, how much will you have in your account at $t + 5$? (Hint: inflation rates at $t + 1$ and $t + 2$ are zero)

Ex. 1.5 (Expand. Mind the signs)**.**

$(a)\ p^2(2 - q) + p(q - 1),$ $(b)\ 2p(p + 3) - (q + 1)(q + 2),$

$(c)\ (2pq - 3p^2)(p + 2q) - (q^2 - 2pq)(2p - q),$ $(d)\ -3(n^2 - 2n + 3),$

$(e)\ (a^2b - ab^2)(a + b)$ $(f)\ (x - y)(x - 2y)(x - 3y),$

$(g)\ 6a^2b(5ab - 3ab^2),$ $(h)\ (4n - 3)(n - 2),$

$(i)\ (2 - t^2)(2 + t^2),$ $(k)\ (u - v)^2(u + v)^2,$

$(l)\ (2t - 1)(t^2 - 2t + 1),$ $(m)\ (a + 1)^2 + (a - 1)^2 - 2(a + 1)(a - 1),$

$(n)\ (x + y + z)^2,$ $(o)\ (x + y + z)^2 - (x - y - z)^2,$

$(p)\ (x + 2y)^2,$ $(q)\ \left(\dfrac{1}{x} - x\right)^2,$

Ex. 1.6 (Factoring)**.**

$(a)\ a^3 - a^2b,$ $(b)\ 8x^2y^2 - 16xy,$

$(c)\ 16 - b^2,$ $(d)\ 4t^2s - 8ts^2,$

$(e)\ K^3 - K^2L,$ $(f)\ K^2L - L^2K,$

$(g)\ K^{-4} - LK^{-5},$ $(h)\ -\dfrac{1}{5}x^2 + 2xy - 5y^2$

Ex. 1.7 (Fractions).

(a) $\dfrac{1}{x-2} - \dfrac{1}{x+2}$,

(b) $\dfrac{6x+25}{4x+2} - \dfrac{6x^2+x-2}{4x^2-1}$,

(c) $\dfrac{18b^2}{a^2-9b^2} - \dfrac{a}{a+3b} + 2$,

(d) $\dfrac{1}{8ab} - \dfrac{1}{8b(a+2)}$,

(e) $\dfrac{2t-t^2}{t+2}\left(\dfrac{5t}{t-2} - \dfrac{2t}{t-2}\right)$,

(f) $2 - \dfrac{a(1-\frac{1}{2a})}{1/4}$,

(g) $\dfrac{2}{x} + \dfrac{1}{x+1} - 3$,

(h) $\dfrac{3x}{x+2} - \dfrac{4x}{2-x} - \dfrac{2x-1}{x^2-4}$,

(i) $\dfrac{\frac{1}{x}+\frac{1}{y}}{\frac{1}{xy}}$,

(k) $\dfrac{\frac{1}{x^2}-\frac{1}{y^2}}{\frac{1}{x^2}+\frac{1}{y^2}}$,

(l) $\left(\dfrac{1}{4} - \dfrac{1}{5}\right)^{-2}$,

(m) $n - \dfrac{n}{1-\frac{1}{n}}$,

(n) $\dfrac{1}{1+x^{a-b}} + \dfrac{1}{1+x^{b-a}}$,

(o) $\dfrac{\frac{1}{x-1}+\frac{1}{x^2-1}}{x-\frac{2}{x+1}}$

Ex. 1.8 (Radicals and Roots).

(a) $\dfrac{6}{\sqrt{7}}$, (b) $\dfrac{\sqrt{32}}{\sqrt{2}}$, (c) $\dfrac{\sqrt{54}-\sqrt{24}}{\sqrt{6}}$, &(d) $\dfrac{2}{\sqrt{3}\sqrt{8}}$, (e) $16^{3/2}$, (f) $(1/27)^{-2/3}$

(g) $\dfrac{a^{3/8}}{a^{1/8}}$, (h) $(x^{1/2}x^{3/2}x^{-2/3})^{3/4}$,

(i) $\left(\dfrac{10p^{-1}q^{2/3}}{80p^2q^{-7/3}}\right)^{-2/3}$,

(k) $\dfrac{x\sqrt{y}-y\sqrt{x}}{x\sqrt{y}+y\sqrt{x}}$,

(l) $\dfrac{h}{\sqrt{x+h}-\sqrt{x}}$

(m) $(((a^{1/2})^{2/3})^{3/4})^{4/5}$, &(n) $a^{1/2}a^{2/3}a^{3/4}a^{4/5}$,

(o) $(((3a)^{-1})^{-2}.(2a^{-2})^{-1})/a^{-3}$,

(p) $\dfrac{\sqrt[3]{a}a^{1/12}\sqrt[4]{a^3}}{a^{5/12}\sqrt{a}}$

Ex. 1.9 (Economic problems). Answer the following

(a) The population of a nation increased from 40 mil. to 60 mil. in 12 years. What is the yearly percentage rate of the population growth p?

(b) An item initially costs \$2. Its price is then increased by 5%, and afterward, it decreased by 5%. What is the final price?

Ex. 1.10 (Inequalities)**.** Determine x such that

(a) $3x - 5 > x - 3$,

(b) $\dfrac{2x - 3}{x - 1} > 3 - x$

(d) $\dfrac{x}{24} - (x + 1) + \dfrac{3x}{8} < \dfrac{5}{12}(x + 1)$,

(f) $\dfrac{2x + 1}{x - 3} > 1$,

(h) $x^2 - 2x \leq 0$,

(k) $\dfrac{-x - 2}{x + 2} > 2$,

(m) $(x - 1)(x + 4) > 0$,

(o) $\dfrac{3x - 1}{x} > x + 3$,

(r) $1 \leq 2x - 1 < 16$,

(b) $(x - 1)(3 - x) > 0$,

(c) $\dfrac{2x - 4}{3} \leq 7$

(e) $\dfrac{x + 2}{x - 1} < 0$,

(g) $5x^2 \leq 125$,

(i) $\dfrac{1}{x - 2} + \dfrac{3}{x^2 - 4x + 4} \geq 0$,

(l) $x^4 < x^2$,

(n) $(x - 1)^2(x + 4) > 0$,

(p) $-5 < \dfrac{1}{x} < 0$,

(s) $\dfrac{\dfrac{1}{x} - 1}{\dfrac{1}{x} + 1} \geq 1$

Ex. 1.11 (Absolute Value)**.** Determine x

(a) $|x - 2|$ for $x = -3, -1, 0, 2, 4$

(b) $|3 - 2x| = 5$,

(d) $|x - 2| \leq 1$,

(f) $|x| > \sqrt{2}$,

(c) $|x| \leq 2$,

(e) $|3 - 8x| \leq 5$,

(g) $|x^2 - 2| \leq 1$

1.2 Summary

Arithmetic Operations

The denominator must always be $\neq 0$

$$\frac{p}{0} \text{ is undefined} \qquad ab + bc = (a + c)b,$$

$$a\left(\frac{b}{c}\right) = \frac{ab}{c}, \qquad \frac{a/b}{c} = \frac{a}{bc},$$

$$\frac{a}{b/c} = \frac{ac}{b}, \qquad \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$

$$\frac{a + b}{c} = \frac{a}{c} + \frac{b}{c},$$

Exponent Properties

$$a^{-n} = \frac{1}{a^n},$$

$$a^0 = 1 (a \neq 0), \qquad a^n a^m = a^{n+m},$$

$$\frac{a^n}{a^m} = a^{n-m} = \frac{1}{a^{m-n}}, \qquad (ab)^n = a^n b^n,$$

$$\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}, \qquad \left(\frac{a}{b}\right)^{-n} = \left(\frac{b}{a}\right)^n = \frac{b^n}{a^n},$$

$$a^{\frac{m}{n}} = (a^{1/m})^n = (a^n)^{1/m}$$

Properties of Radicals

$$\sqrt[n]{a} = a^{1/n}, \qquad \sqrt[n]{ab} = \sqrt[n]{a}\,\sqrt[n]{b},$$

$$\sqrt[m]{\sqrt[n]{a}} = \sqrt[nm]{a}, \qquad \sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}}$$

Some Properties

If $a < b$ then

$$a + c < b + c$$
$$a - c < b - c$$

If $a < b$ and $c > 0$ then

$$ac < bc$$
$$\frac{a}{c} < \frac{b}{c}$$

If $a < b$ and $c < 0$ then

$$ac > bc$$
$$\frac{a}{c} > \frac{b}{c}$$

If $\dfrac{a}{b} = \dfrac{c}{d}$ then

$$ad = bc$$

If $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ are 2 points, the distance between them is

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Properties of Absolute Value

$$|a| = \begin{cases} a \text{ if } a \geq 0 \\ -a \text{ if } a < 0 \end{cases}, \qquad\qquad\qquad\qquad |a| \geq 0,$$

$$|-a| = |a|, \qquad\qquad\qquad\qquad\qquad |ab| = |a||b|,$$

$$\left|\dfrac{a}{b}\right| = \dfrac{|a|}{|b|}, \qquad\qquad |a + b| \leq |a| + |b| \textbf{ (triangle inequality)}$$

1.3 Summation Notation

1.3.1 Basics

Shorthand for

$$x_1 + x_2 + \cdots + x_n = \sum_{i=1}^{n} x_i$$

Ex. 1.12. Evaluate:

$$(a) \ \sum_{i=1}^{10} i^2 \qquad\qquad\qquad (b) \ \sum_{j=1}^{4} \frac{j+1}{j}$$

$$(c) \ \sum_{j=0}^{2} \frac{(-1)^j}{(j+1)(j+3)} \qquad\qquad (d) \ \sum_{k=3}^{6} (5k - 3)$$

Ex. 1.13. Expand

$$(a) \ \sum_{i=1}^{5} p_i q_i \qquad\qquad\qquad (b) \ \sum_{i=1}^{N} (x_{ij} - \bar{x}_j)^2$$

$$(c) \ \sum_{k=0}^{2} 2\sqrt{k+2} \qquad\qquad (d) \ \sum_{k=1}^{n} a_{ki} b^{k+1}$$

Ex. 1.14. Express in summation

$$(a) \ 1^3 + 2^3 + \cdots + n^3 \qquad (b) \ 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + (-1)^n \frac{1}{2n+1}$$

1.3.2 Properties

Additivity property

$$\sum_{i=1}^{n}(a_i + b_i) = \sum_{i=1}^{n}a_i + \sum_{i=1}^{n}b_i$$

Homogeneity property

$$\sum_{i=1}^{n}ca_i = c\sum_{i=1}^{n}a_i$$

Ex. 1.15. Evaluate

$$\sum_{m=2}^{n}\frac{1}{(m-1)m} = \frac{1}{1\times 2} + \frac{1}{2\times 3} + \cdots + \frac{1}{(n-1)n}$$

knowing that

$$\frac{1}{(m-1)m} = \frac{1}{m-1} - \frac{1}{m}$$

Ex. 1.16. The mean $\bar{x}$ of T numbers $x_1, x_2, \ldots, x_n$ is the sum of all numbers dividded by T:

$$\bar{x} = \frac{1}{T}\sum_{i=1}^{T}x_i$$

Prove that

$$(a)\ \sum_{i=1}^{T}(x_i - \bar{x}) = 0$$

$$(b)\ \sum_{i=1}^{T}(x_i - \bar{x})^2 = \sum_{i=1}^{T}x_i^2 - T\bar{x}^2$$

1.3.3 Double Sums

$$\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m}a_{ij}$$

Ex. 1.17. Compute

$$(a)\ \sum_{i=1}^{3}\sum_{j=1}^{4}(i+2j) \qquad\qquad (b)\ \sum_{i=1}^{3}\sum_{j=1}^{4}i\cdot 3^j$$

$$(c)\ \sum_{s=0}^{2}\sum_{r=2}^{4}\left(\frac{rs}{r+s}\right)^2 \qquad\qquad (d)\ \sum_{i=1}^{m}\sum_{j=1}^{n}(i+j^2)$$

$$(e)\ \sum_{i=1}^{m}\sum_{j=1}^{2}i^j$$

Chapter 2

Equations

Solving an equation is finding all variable values to satisfy a mathematical expression.

2.1   Simple Equations

Example 2.1. Find all x such that

- (a) $3x + 10 = x + 4$

- (b) $\dfrac{x + 2}{x - 2} - \dfrac{8}{x^2 - 2x} = \dfrac{2}{x}$

- (c) A firm manufactures a good that costs \$20 per unit to produce. In addition, the firm has fixed costs of \$2000. Each unit is sold for \$75. How many units must be sold for the firm to meet its target profit of \$14,500?

A linear equation has the form

$$ax + b = 0, \tag{2.1}$$

where x is a variable, $a \neq 0$. To find the solutions, follow these steps

1. Identify the domain (For instance, in a fraction, the denominator must $\neq 0$; the expression inside square root $\sqrt{u}$ must ≥ 0)

2. Transform any equation to match the form (2.1)

3. Put all variables on the LHS, all parameters on the RHS so that $ax = -b$

4. Solutions: $x = -b/a$ and cross out values that do not belong to the domain

To make the expression on both sides equivalent, you can

- add or subtract the same number on both sides

- multiply both sides by the same number $\neq 0$x

Ex. 2.1. Solve:

$(a)\ \dfrac{x-3}{4}+2=3x$

$(b)\ 3x=\dfrac{1}{4}x-7$

$(c)\ \dfrac{1}{2x+1}=\dfrac{1}{x+2}$

$(d)\ \sqrt{2x+14}=16$

$(e)\ \dfrac{x-3}{x+3}=\dfrac{x-4}{x+4},$

$(f)\ \dfrac{3}{x-3}-\dfrac{2}{x+3}=\dfrac{9}{x^2-9},$

$(g)\ 4x+2(x-4)-3=2(3x-5)-1,$

$(h)\ \dfrac{6x}{5}-\dfrac{5}{x}=\dfrac{2x-3}{3}+\dfrac{8x}{15}$

$(i)\ \dfrac{2-\dfrac{x}{1-x}}{1+x}=\dfrac{6}{2x+1}$

$(j)\ \dfrac{1}{2}\left(\dfrac{x}{2}-\dfrac{3}{4}\right)-\dfrac{1}{4}\left(1-\dfrac{x}{3}\right)-\dfrac{1}{3}(1-x)=-\dfrac{1}{3}$

2.2   Equations with Parameters

The form:

$$y = ax + b$$

Example 2.2. Basic macroeconomic model

$$\begin{cases} Y = C + I, \\ C = a + bY \end{cases} \tag{2.2}$$

where Y is GDP, C is consumption, I is investment, a, b are positive parameters ($b < 1$). By substituting, we can express (2.2) as

$$Y = a + bY + I$$

so that

$$Y = \frac{a}{1-b} + \frac{1}{1-b}I, \tag{2.3}$$

where Y is the **endogenous variable** and I is the **exogenous variable**. Given a, b, and I, we can find any value of Y. Both (2.2) and (2.3) are equivalent. But (2.2) is called the **structural form** of the model while (2.3) is called the **reduced form** of the model

Ex. 2.2. Find the reduced form of C from (2.2).

Ex. 2.3. Suppose the total demand for money of the economy

$$M = \alpha Y + \beta(r - Y)^{-\delta}$$

where M is quantity of money, Y is GDP, r is interest rate, the rest are positive parameters.

- Solve the equation for r

- Estimated parameters for the US during 1929–1952 is $\alpha = 0.14, \beta = 76.03, \gamma = 2, \delta = 0.84$, find r

Ex. 2.4. Solve for x and K

$(a)\ \dfrac{1}{ax} + \dfrac{1}{bx} = 2$

$(b)\ \dfrac{ax+b}{cx+d} = A$

$(c)\ \dfrac{1}{2}px^{-1/2} - w = 0$

$(d)\ \dfrac{ax}{\sqrt{1+x}} + \sqrt{1+x} = 0$

$(e)\ a^2x^2 - b^2 = 0$

$(f)\ (3+a^2)^x = 1,$

$(g)\ \alpha x - \alpha = \beta x - \beta$

$(h)\ \sqrt{qx} - 3q = 5,$

$(i)\ x = 94 + 0.3(x - (20 + 0.5x))$

$(j)\ K^{1/2}\left(\dfrac{1}{2}\dfrac{r}{w}K\right)^{1/4} = Q$

$(k)\ \dfrac{\frac{1}{2}L^{-1/2}K^{1/4}}{\frac{1}{4}K^{-3/4}L^{1/2}} = \dfrac{r}{w}$

$(l)\ \dfrac{1}{2}pK^{-1/4}\left(\dfrac{1}{2}\dfrac{r}{w}\right)^{1/4} = r$

$(m)\ \dfrac{x - 2y + xz}{x - z} = 4y$

$(n)\ Y = C\left(1 - \dfrac{K}{N}\right)$

2.3 Quadratic Equations

The form:
$$ax^2 + bx + c = 0 \tag{2.4}$$

Example 2.3. Solve
$$x^2 + 8x - 9 = 0$$

1. Method 1: Completing the square by adding 16 to both sides

$$x^2 + 8x + 16 = 9 + 16$$
$$\Leftrightarrow (x+4)^2 = 25$$
$$\Leftrightarrow x + 4 = \pm 5$$
$$\Leftrightarrow x = 1 \text{ or } x = -9$$

2. Method 2: Factorization (the solutions often are divisors of c)

$$x - x + 9x - 9 = 0$$
$$\Leftrightarrow (x-1)(x+9) = 0$$

3. Method 3: (more general, works in all cases) find roots with the quadratic formula

$$\triangle = b^2 - 4ac$$

If $\triangle \geq 0$ and $a \neq 0$, then there are 2 solutions

$$x = \frac{-b \pm \sqrt{\triangle}}{2a}$$

If $\triangle = 0$, there is only 1 solution

$$x = -\frac{b}{2a}$$

If $\triangle < 0$, there is no **real** solution.

Ex. 2.5. Verify Method 3 for the general form (2.4).

Ex. 2.6. Solve for x

$(a)\ 15x - x^2 = 0$ $(b)\ x^2 - 16 = 0$

$(c)\ (x - 3)(x + 4) = 0$ $(d)\ x^2 - 4x + 4 = 0$

$(e)\ x^2 - 5x + 6 = 0$ $(f)\ x^2 - x - 12 = 0$

$(g)\ -\dfrac{1}{4}x^2 + \dfrac{1}{2}x + \dfrac{1}{2} = 0$ $(h)\ x^2 + 11x - 26 = 0$

$(i)\ x^3 - 4x = 0$ $(k)\ x^4 - 5x^2 + 4 = 0$

$(l)\ x^{-2} - 2x^{-1} - 15 = 0$ $(m)\ 3x^2 = 5x - 1$

2.4   System of 2 Linear Equations

The form

$$\begin{cases} a_1 x + b_1 y = c_1 \\ a_2 x + b_2 y = c_2 \end{cases}$$

Example 2.4. Find x, y that satisfy both equations

$$\begin{cases} 2x + 3y = 18 \\ 3x - 4y = -7 \end{cases}$$

There are 2 methods: substitution and elimination.
Method 1 From the first equation, we have:

$$y = 6 - \frac{2}{3}x$$

Plugging into the second equation yields

$$3x - 4\left(6 - \frac{2}{3}x\right) = -7$$
$$\Leftrightarrow x = 3$$

which yields $y = 4$.
   Method 2 Multiplying the second equation with $2/3$ [1], the system becomes

$$\begin{cases} 2x + 3y = 18 \\ 2x - \frac{8}{3}y = -\frac{14}{3} \end{cases}$$

Subtract the second equation from the first one yields

$$\begin{cases} 2x + 3y = 18 \\ \frac{17}{3}y = \frac{68}{3} \end{cases}$$

which also yields $y = 4, x = 3$.

---

[1] A smarter way would be to multiply the first equation by 4 and the second equation by 3

Ex. 2.7. Solve the following system for x, y

(a) $\begin{array}{l} x - y = 5 \\ x + y = 11 \end{array}$ (b) $\begin{array}{l} 4x - 3y = 1 \\ 2x + 9y = 4 \end{array}$

(c) $\begin{array}{l} 23x + 45y = 181 \\ 10x + 15y = 65 \end{array}$ (d) $\begin{array}{l} 0.01x + 0.21y = 0.042 \\ -0.25x + 0.55y = -0.47 \end{array}$

Ex. 2.8. Answer the following questions by formulating the right equations

(a) 5 tables and 20 chairs cost \$1800, but 2 tables and 3 chair cost \$420. What are the unit prices for a chair and a table?

(b) A firm makes 2 goods, A and B. The estimated output of A is 50% higher than that of B. The profit per unit sold is \$300 for A and \$200 for B. If the profit target is \$13,000, how many units of A and B should the firm produce?

(c) At the beginning of the year, a person has a total of \$10,000 in 2 accounts. The first account yields an interest rate of 5% and the other yields 7.2% per year. At the end of the year, the total interest earned from both accounts is \$ 676. What was the initial balance in each of the 2 accounts?

Ex. 2.9. Harder problems

(a) $\begin{array}{l} \frac{2}{x} + \frac{3}{y} = 4 \\ \frac{3}{x} - \frac{2}{y} = 19 \end{array}$ (b) $\begin{array}{l} 3\sqrt{x} + 2\sqrt{y} = 2 \\ 2\sqrt{x} - 3\sqrt{y} = 1/4 \end{array}$

(c) $\begin{array}{l} x^2 + y^2 = 13 \\ 4x^2 - 3y^2 = 24 \end{array}$

2.5 Nonlinear Equations

In general
$$ab = ac \Leftrightarrow a = 0 \text{ or } b = c$$

Example 2.5.

(a) $x^3\sqrt{x+2} = 0$

(b) $x(y+3)(z^2+1)\sqrt{w-3} = 0$ (c) $x^2 - 3x^3 = 0$

(d) $x(x+a) = x(2x+b)$ (e) $xy^2(1-y) - 2\lambda(y-1) = 0$

(f) $\dfrac{1-K^2}{\sqrt{1+K^2}} = 0$ (g) $\dfrac{45 + 6r - 3r^2}{(r^4+2)^{2/3}} = 0$

(h) $\dfrac{x^2 - 5x}{\sqrt{x^2 - 25}} = 0$

Ex. 2.10. Solve

(a) $\dfrac{5 + x^2}{(x-1)(x+2)} = 0$

(b) $1 + \dfrac{2x}{x^2+1} = 0$ (c) $\dfrac{(x+1)^{1/3} - \frac{1}{3}x(x+1)^{-2/3}}{(x+1)^{2/3}} = 0$

(d) $\dfrac{x}{x-1} + 2x = 0$ (e) $(x^2 - 4)\sqrt{5-x} = 0$

(f) $(x^4+1)(4+x) = 0$ (g) $k^\alpha = \left(\dfrac{w}{r}\right)^{\frac{\sigma}{1-\sigma}}$

Chapter 3

Functions of One Variable

3.1   The Basics

Notation

$$y = f(x)$$

A function assigns a unique real number (y) to each number x in the domain **D**, based on some specific **rules** described in the function. The set of all possible values of y is called the range of f. We often use letters f, g, h, u, φ for functions.

Example 3.1.

$(a)\ f(x) = x^3,$ $(b)\ f(x) = 2x + 1$

$(c)\ C(x) = 100x\sqrt{x} + 500,$ $(d)\ C(x) = Ax^2 - B$

$(e)\ f(x) = \dfrac{1}{x+3}\ \&(f)\ g(x) = \sqrt{2x+4}$

$(g)\ f(k) = Ak^\alpha$ $(h)\ u(c) = c^{1/2}$

Ex. 3.1. Find the domain of (e), (f), (g), (h) (The values for x so that its function is defined).

Ex. 3.2. Let
$$f(x) = x^2 + 1$$

(a) Compute $f(0), f(-1), f(1/2), f(\sqrt{2})$.
(b) For what values of x, it is true that

1. f(x) = f(-x)?

2. f(x+1) = f(x) + f(1)?

3. f(2x) = 2f(x)?

(c) Suppose $f(x) = a^2 - (x - a)^2$ where a is a constant, compute

1. $f(0), f(a), f(-a), f(2a)$

2. $3f(a) + f(-2a)$

Ex. 3.3. Find the domain and range of the functions

$$(a)\ y = \sqrt{5 - x} \qquad\qquad\qquad (b)\ y = \frac{2x - 1}{x^2 - x}$$

$$(c)\ y = \sqrt{\frac{x - 1}{(x - 2)(x + 3)}} \qquad\qquad (d)\ y = 1 - \sqrt{x + 2}$$

3.2   Graphs

One of the most important aspects of a function is to graph because graphs provide some visual cues and intuition before we want to prove something analytically. For example, the following are the graphs of $y = x^2 - 2x - 1$ and $y = x^2 + 2x + 1$. From here, we can visually see that there is no value of x such that $x^2 + 2x + 1 = 0$. Also, both functions have a global minimum. The function is decreasing before it reaches the minimum and then turns to an increasing function.
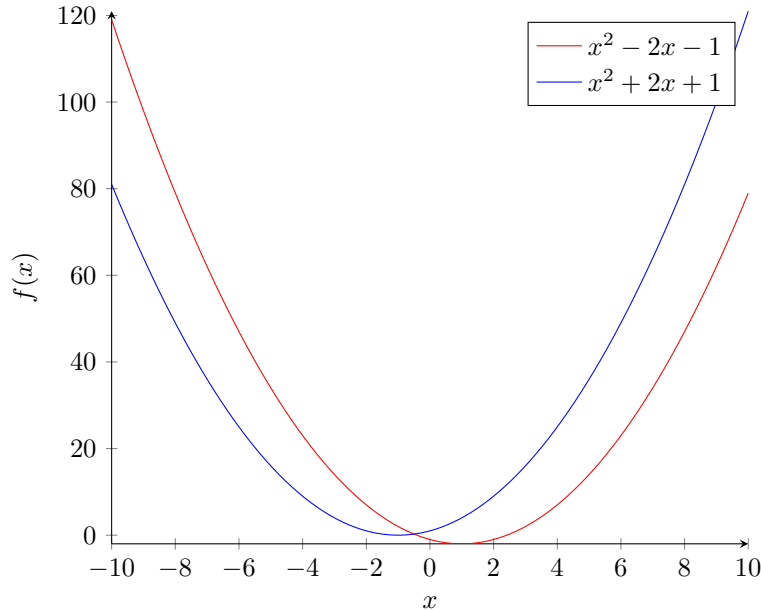
Fig. 3.1. Examples of graphs

3.3   Linear Functions

The form

$$y = f(x) = ax + b$$

The graph is a **straight line**.

How much would the function changes when we increase x by 1 unit?

$$\triangle f = f(x + 1) - f(x) = a(x + 1) + b - (ax + b) = a$$

i.e., f changes by a unit. For this reason, the number a is the **slope** of the line (or the function).
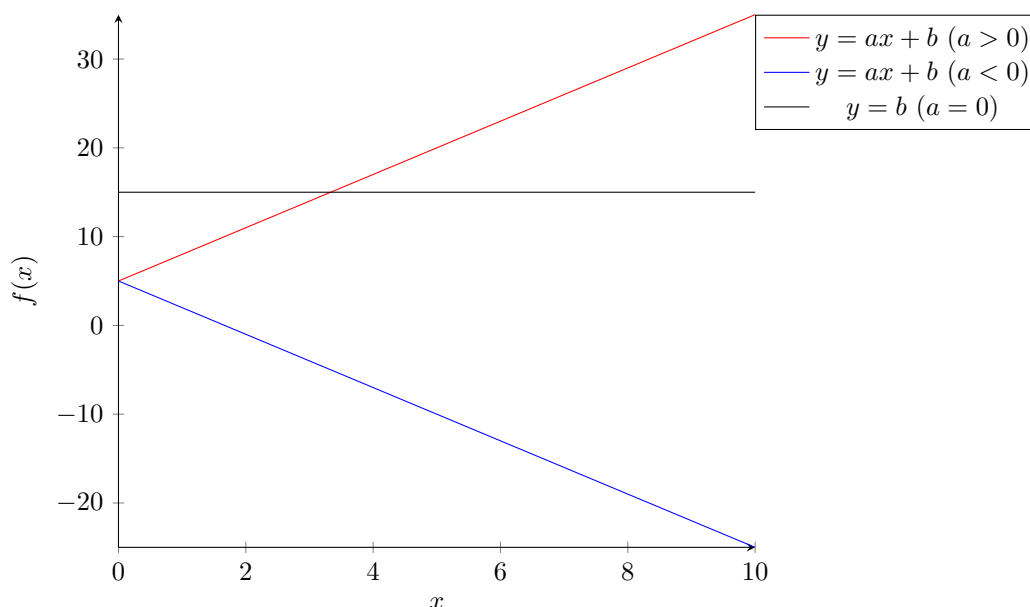
Fig. 3.2. Slope is positive (a¿0), the line slants upward right. The line slants downward if the slope is negative (a¡0). The line is horizontal when there is no steepness (a=0). The point where $x = 0$ is called **y-intercept**

The slope of a straight line L is

$$a = \frac{y_2 - y_1}{x_2 - x_1} \quad (x_1 \neq x_2)$$

where $(x_1, y_1), (x_2, y_2)$ are 2 distinct points on L.

To draw the graph of a linear function, given 2 points [1] on that line

1. First, calculate the slope

$$a = \frac{y_2 - y_1}{x_2 - x_1}$$

2. Plug into the following formula

$$y - y_1 = a(x - x_1)$$

then rearrange to get the form $y = ax + b$.

Ex. 3.4. Draw graphs of the following

$$(a)\ 3x + 4y = 12 \ \&\ (b)\ \frac{x}{10} - \frac{y}{5} = 1$$

$$(c)\ x = 3$$

---

[1] Normally, the easiest 2 points are $(0, f(0))$ and $(x, f(x) = 0)$

3.4   Quadratic Functions

The form

$$y = f(x) = ax^2 + bx + c \quad (a \neq 0)$$

In general, the graph of a quadratic function is a parabola.
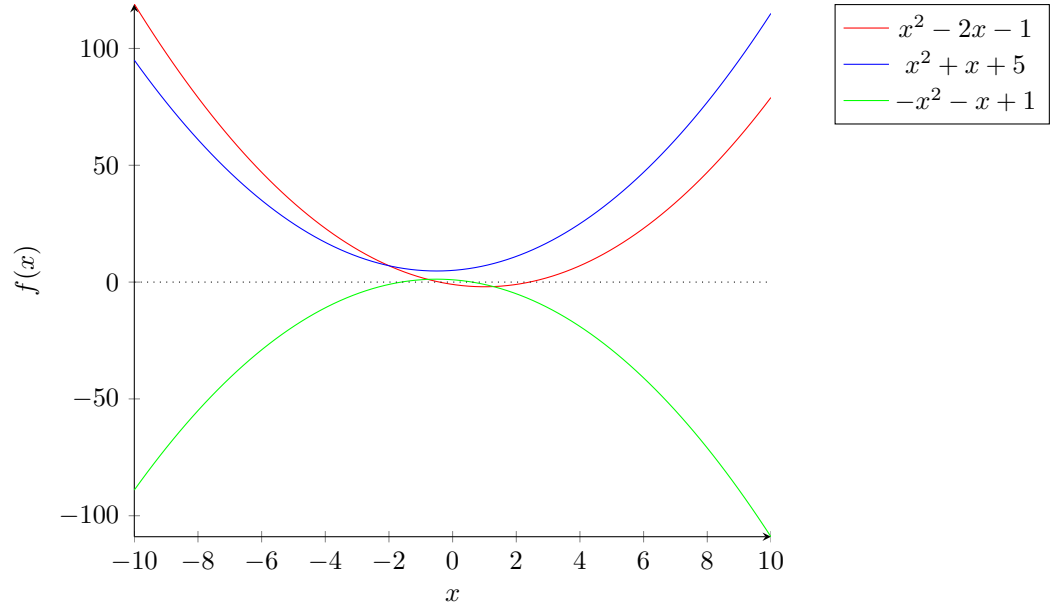
Fig. 3.3. Some graphs of different quadratic functions

It's easy to see that [blue] and [red] graphs have a min, while [green] has a max.

Ex. 3.5. Verify that: For a quadratic function with the form $f(x) = a^2 + bx + c$,

- If $a > 0$, the minimum is at $x = -b/2a$

- If $a < 0$, the maximum is at $x = -b/2a$

In economics, the cost function usually is a quadratic function.

Example 3.2. A firm set the price as follows

$$P = 102 - 2Q$$

where P is the unit price and Q is quantity. The cost of producing Q units is

$$C = \frac{1}{2}Q^2 + 2Q$$

Suppose the profit is $\pi = PQ - C$, find the optimal Q that maximizes profit.

Ex. 3.6. Sandmo model of efficient loan markets

$$U(x) = 72 - (4 + x^2)^2 - (4 - rx^2)$$

where r is a constant. Find x that maximizes $U(x)$.

3.5 Polynomials

A cubic function has the form

$$y = f(x) = ax^3 + bx^2 + cx + d \quad (a \neq 0)$$

From here, we can generalize a polynomial function of degree n as

$$y = f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

Linear, quadratic, and cubic functions are polynomial functions of degrees 1, 2, and 3 respectively.

To find the integer roots of a polynomial $f(x)$ is to find the value of x such that $f(x) = 0$. Then, if such integer roots exist, it must be a factor of a_0.

Ex. 3.7. Find integer roots of

$$x^3 - 2x^2 + x - 2 = 0$$

All integer solutions of this equation must be factors of 2. The candidates are $\pm 1, \pm 2$. Simple checks show that $x = 2$ is the only integer solution. Knowing this root, we can factorize the equation

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

Polynomial Division

Example 3.3. Divide

$$(a) \ (-x^3 + 4x^2 - x - 6) \div (x - 2)$$
$$(b) \ (x^4 + 3x^2 - 4) \div (x^2 + 2x)$$

Results:

$$(a) \ -x^2 + 2x + 3 \& \text{remainder: } 0$$
$$(b) \ x^2 - 2x + 7 \& \text{remainder: } -14x - 4$$

Ex. 3.8. Find integer roots then factorize the following equations

$$(a) \ x^4 - x^3 - 7x^2 + x + 6 = 0 \qquad (b) \ 2x^3 + 11x^2 - 7x - 6 = 0$$

$$(c) \ x^4 + x^3 + 2x^2 + x + 1 = 0 \qquad (d) \ \frac{1}{4}x^3 - \frac{1}{4}x^2 - x + 1 = 0$$

$$(e) \ x^5 - 4x^3 - 3 = 0$$

Ex. 3.9. Divide

$$(a) \ (2x^3 + 2x - 1) \div (x - 1) \qquad (b) \ (x^6 + x^3 + x^2 + x) \div (x^2 + x)$$
$$(c) \ (x^5 - 3x^4 + 1) \div (x^2 + x + 1) \qquad (d) \ (3x^8 + x^2 + 1) \div (x^3 - 2x + 1)$$

Ex. 3.10. In demand theory, we have the function

$$E = \alpha \frac{x^2 - \gamma x}{x + \beta}$$

Perform the division $(x^2 - \gamma x) \div (x + \beta)$ and express E as a sum of a linear function and a fraction. (Hint: $x - (\beta + \gamma)$ remainder $\beta(\beta + \gamma)$.)

3.6   Power Functions

In general form (say: x **raised to the power of** r)

$$f(x) = Ax^r \quad (x > 0, \quad r \text{ and } A \text{ are constants})$$

The graph of power functions. Note that the signs of the **exponent** r influence the shape of the function.
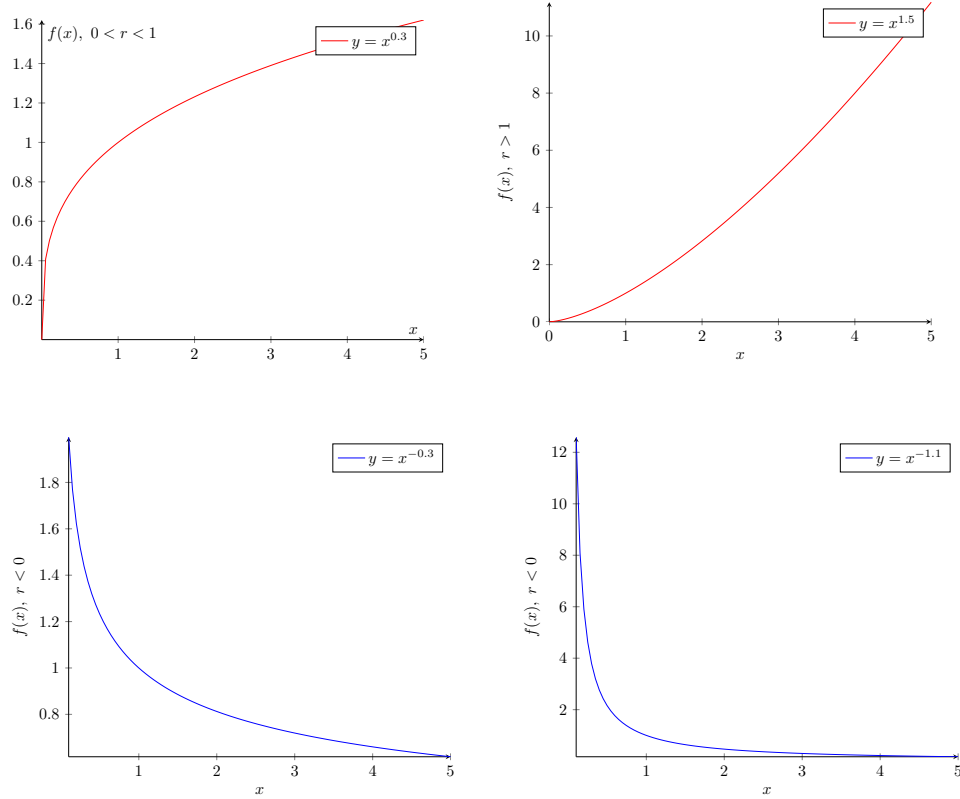
Ex. 3.11. Sketch the graph

$$(a)\ y = x^{-3} \qquad\qquad (b)\ y = x^{-1/2}$$

Ex. 3.12. Solving for x

$$(a)\ 2^{2x} = 8 \qquad\qquad (b)\ 3^{3x+1} = 1/81$$

$$(c)\ 10^{x^2-2x+2} = 100 \qquad\qquad (d)\ 3^{5x}9^x = 27$$

The power function is important since most of the utility and production-per-capita functions normally take this form. For instance, $U(c) = c^{1-\sigma}/(1-\sigma)$ or $y = k^\alpha$. The shapes of these functions (which depends on the sign of the exponent) are also worth paying attention to.

3.7 Exponential Functions

In general form:

$$y = f(x) = Aa^x \quad (A, a \text{ are positive integer })$$

Exponential functions appear in many important topics of economics such as economic and population growth, compound interest, exponential decay, etc.
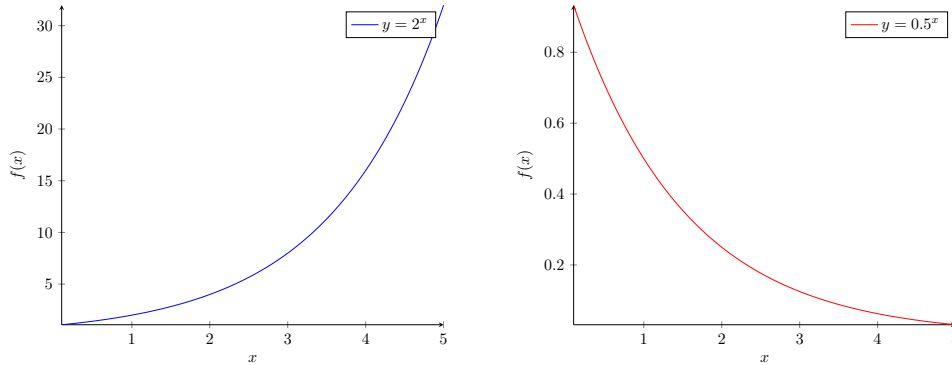
We can consider a the factor by which $f(x)$ changes when x increases by 1.

- If $a = 1 + p(\%)$, where $0 < p < 1$, $A > 0$, then $f(x)$ will increase by $p(\%)$ for each unit increase in x.

- If $a = 1 - p(\%)$, where $0 < p < 1$, $A > 0$, then $f(x)$ will decrease by $p(\%)$ for each unit increase in x.

Example 3.4. (Population Growth)

In the year 1960, the population of Europe is 641 mil. people. If the population grows at the annual rate of $g = 0.72\%$. How is the population in the year 2000? Plot the graph. Compared to a linear function $P = 641 + 6.4t$, which one do you think is a more reasonable estimation?

Let $t = 2000 - 1960 = 40$, the exponentially growing population is

$$P_{t+40} = P_t(1 + g)^t$$
$$\Rightarrow P_{2000} = P_{1960} (1 + 0.0072)^{40} = 854$$

If we applied the linear function, the population would be 897.
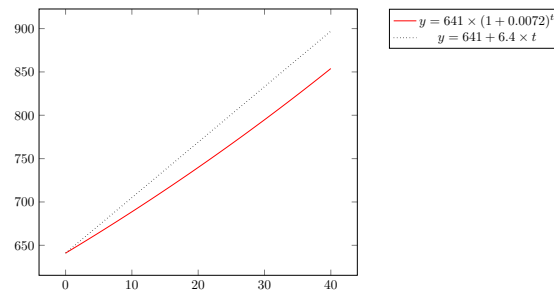
Obviously, the exponential growth function provides more accuracy. Furthermore, a linear function implies that the population is negative for $t \leq 101$.

The Natural Exponential Function

The base e is called the natural exponential function.

$$f(x) = e^x = \exp(x)$$

where e is estimated to be

$$e = 2.718281828459045\ldots$$

e is also known as the Euler number. It's natural for many special reasons

- Many things in nature grow with this base

- Its derivative is its own function $(e^x)' = e^x$.

The functional form of e is actually

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$
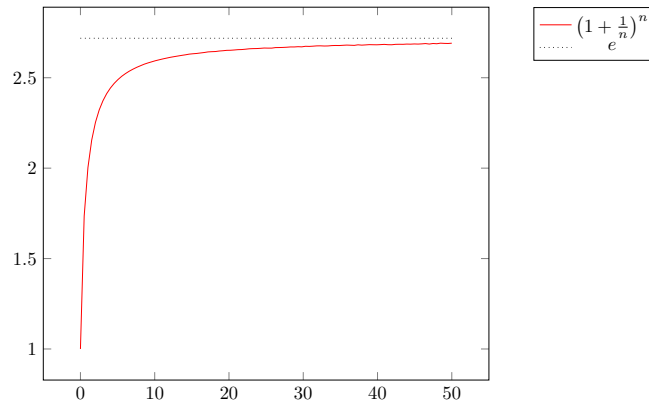
The graphs of $f(x) = e^x$ and $f(x) = e^{-x}$ are
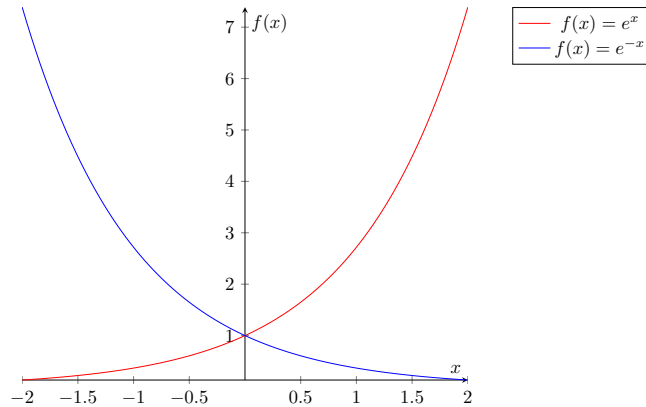
Ex. 3.13. To know why e is important, calculate Example 3.4 using the following formula (other parameters remain the same, $P_t = 641, t = 40, g = 0.72/100$)

$$P_{t+40} = P_t \times e^{gt}$$

Compare the two results.

3.8 Logarithmic Functions

In general, if $a^x = b$ and $a, b > 0$

$$\log_a b = x$$

to **find the solution for a raised to the power of WHAT to get** b.
    Thus, we can derive some important properties.

$$\log_a 1 = 0 \qquad\qquad\qquad \log_a a = 1$$
$$\log_a a^m = m \qquad\qquad\qquad \log_a a^{\log_a n} = n$$
$$\log_a(m \times n) = \log_a m + \log_a n \qquad \log_a\left(\frac{m}{n}\right) = \log_a m - \log_a n$$
$$\log_a b^m = m \log_a b \qquad\qquad\qquad \log_a b = \frac{1}{\log_b a}$$

Change of Base Formula

Ex. 3.14. Prove that for $a, b, c > 0$

$$\log_a b = \frac{\log_c b}{\log_c a}$$

Hint: First, take a raised to both sides, then takes log_c of both sides.

Some special properties of log base e (also written as ln)

$$\ln(1) = 0, \ \ln(e) = 1,$$
$$e^{\ln(x)} = x \ \ (x > 0)$$
$$\ln(x) = \log_e x \Rightarrow \ln(e^x) = x$$

The log of any base can be converted to a natural base log with the Change of Base formula

$$\log_a b = \frac{\ln(b)}{\ln(a)}$$

The graphs of log with different bases.
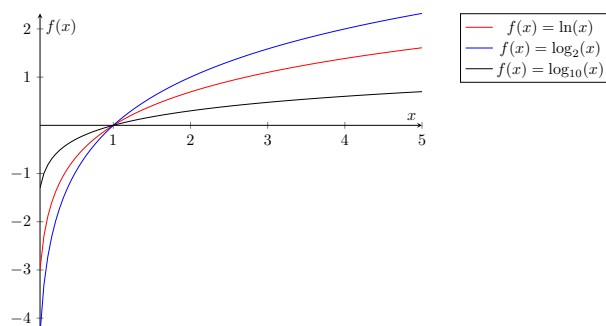
We can see that the result of $x = \log_a b$ for any positive $b < 1$ is an $x < 0$.

Ex. 3.15. Solve for x:

(a) $3^x = 8$, $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (b) $\ln(x) = 3$,

(c) $\ln(x^2 - 4x + 5) = 0$ $\qquad\qquad\qquad\qquad$ (d) $\ln(x(x-2)) = 0$

(e) $\dfrac{x\ln(x+3)}{x^2+1} = 0$ $\qquad\qquad\qquad\qquad\qquad$ (f) $\ln(\sqrt{x} - 5) = 0$

(g) $3^x 4^{x+2} = 8$ $\qquad\qquad\qquad\qquad\qquad$ (h) $3\ln(x) + 2\ln(x^2) = 6$

(i) $4^x - 4^{x-1} = 3^{x+1} - 3^x$ $\qquad\qquad\qquad\qquad$ (k) $\log_2(x) = 2$

(l) $\log_x e^2 = 2$ $\qquad\qquad\qquad\qquad\qquad\qquad$ (m) $\log_3 x = -3$

Ex. 3.16. Prove the Rule of 70 (doubling time)

Given the growth rate $g(\%)$, for an exponential growth function of the form

$$A(t) = A_0 \times (1 + g)^t$$

prove that if $A(t) = 2A_0$, then

$$t = \frac{\ln(2)}{\ln(1 + g)}$$

Note: For small growth g close to 0, why is

$$A_0 \times (1 + g)^t \approx A_0 \times e^{gt} \ ? \qquad\qquad\qquad\qquad (3.1)$$

Proof. We need to show that

$$(1 + g)^t \approx e^{gt}$$

or equivalently

$$\ln((1 + g)^t) = \ln e^{gt}$$
$$\Leftrightarrow t\ln(1 + g) = gt$$

which can be reduced to proving

$$\ln(1 + g) \approx g$$

for small $g \approx 0$.

This can be seen from linear approximation using the first-order Taylor formula. In general, the approximation of $f(x)$ at a is

$$f(x) \approx f(a) + f'(a)(x - a)$$

Let $u = 1 + g$, for very small g, then u is very close to 1. The approximation of $f(u) = \ln(u)$ at 1 is

$$f(1) \approx f(1) + f'(u)(u - 1)$$

We know that

$$f(1) = \ln(1) = 0,$$
$$f'(u) = (\ln(1+g))' = \frac{1}{1+g} \approx 1$$

So that

$$f(1) \approx 0 + 1 \times (u - 1) = 1 + g - 1 = g$$

Thus, for small g close to 0, $\ln(1 + g) \approx g$ and we conclude (3.1). ∎

Chapter 4

Properties of Functions

4.1 Shifting graphs

In the last chapter we learned how to draw the graph of a function. Now we consider how changes in a function relate to shifts in its graph.

The general rules for shifting the graph of $y = f(x)$ is:

(i) If $y = f(x)$ is replaced by $y = f(x) + c$, the graph is moved upwards by c units if $c > 0$; it is moved downwards if $c < 0$.

(ii) If $y = f(x)$ is replaced by $y = f(x + c)$, the graph is moved c units to the left if $c > 0$; it is moved to the right if $c < 0$.

(iii) If $y = f(x)$ is replaced by $y = cf(x)$, the graph is stretched vertically if $c > 0$; it is stretched vertically and reflected about the x-axis if $c < 0$.

(iv) If $y = f(x)$ is replaced by $y = f(-x)$, the graph is reflected about the y-axis.

Example 4.1. Consider the simple example when $c = 2$ for the function $y = \sqrt{x}$, then the 4 transformations above would look like the followings:
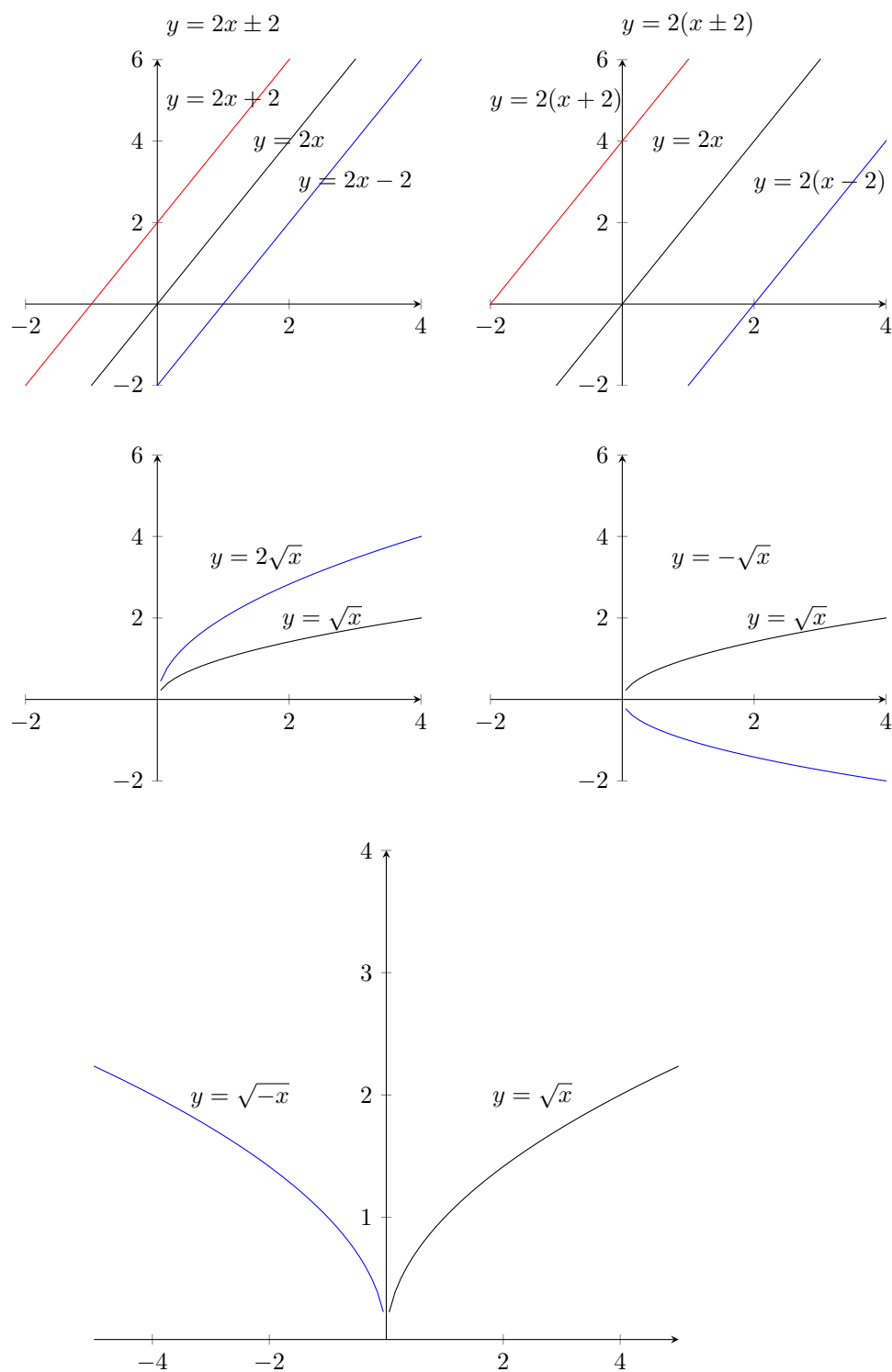
Fig. 4.1. Transformation of graphs

Ex. 4.1. Use the rules for shifting graphs and the following graph of $y = |x|$ to sketch the graph of $y = 2 - |x + 2|$.
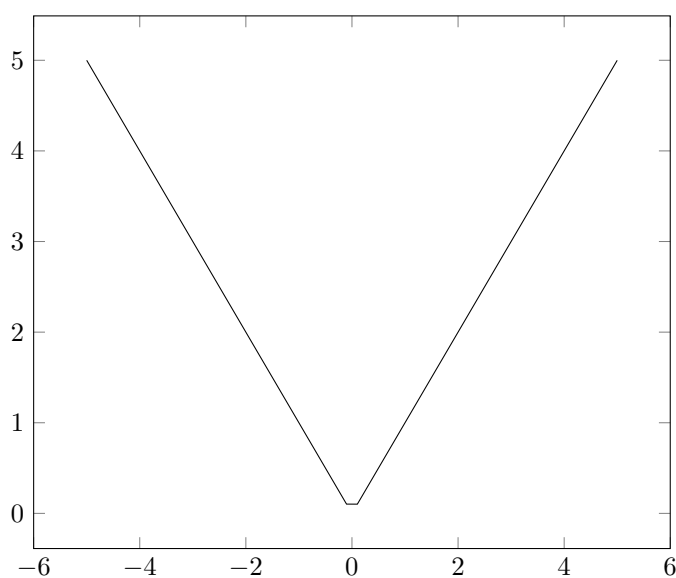
Fig. 4.2. Exercise

Ex. 4.2. Use the rules for shifting graphs to sketch the graphs of the following functions:

(a) $y = x^2 + 1$

(b) $y = (x + 3)^2$

(c) $y = 3 - (x + 1)^2$

Ex. 4.3. Starting from the graph of $f(x) = \frac{1}{x^2}$, sketch the graph of $g(x) = 2 - (x + 2)^{-2}$.

4.2 New functions from the old

4.2.1 Sum and difference

Suppose in general that f and g are functions that are both defined in a set A of real numbers. The function F defined by the formula $F(x) = f(x) + g(x)$ is called the **sum** of f and g, and we write $F = f + g$. The function G defined by $G(x) = f(x) - g(x)$ is called the **difference** between f and g, and we write $G = f - g$.

Example 4.2. Suppose the cost of producing $Q > 0$ units of a commodity is $C(Q)$ defined as:

$$C(Q) = a * Q^3 + b * Q^2 + c * Q + d$$

A commonly used measure in Economics is **average cost**, defined as $A(Q) = C(Q)/Q$:

$$A(Q) = a * Q^2 + b * Q + c + \frac{d}{Q}$$

Thus $A(Q)$ can be expressed as the sum of a quadratic function and a hyperbola. The sum graph is obtained by piling one graph on top of the other as shown in Figure 4.3.
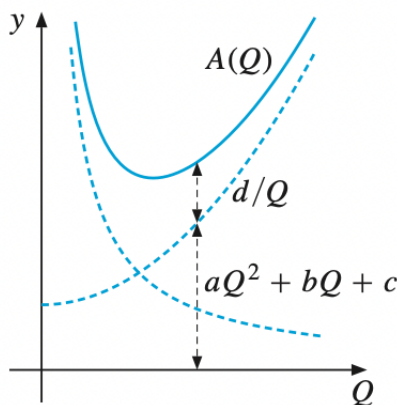
Fig. 4.3. Sum function

Example 4.3. Now let $R(Q)$ denote the revenue obtained by selling Q units. Then, the profit $\pi(Q)$ is given by $\pi(Q) = R(Q) - C(Q)$.
In the case of the firm getting a fixed price p per unit, so that the graph of $R(Q)$ is a straight line through the origin. The graph of $C(Q)$ must be subtracted from that of $R(Q)$. The production level that maximises profit is Q^*.
Similar to the sum function, subtracting the two graphs gives the difference graph as shown in Figure 4.4.

4.2.2 Products and quotients

If f and g are defined in a set A, the function F defined by $F(x) = f(x) \cdot g(x)$ is called the **product** of f and g, and we put $F = f \cdot g$ (or fg).
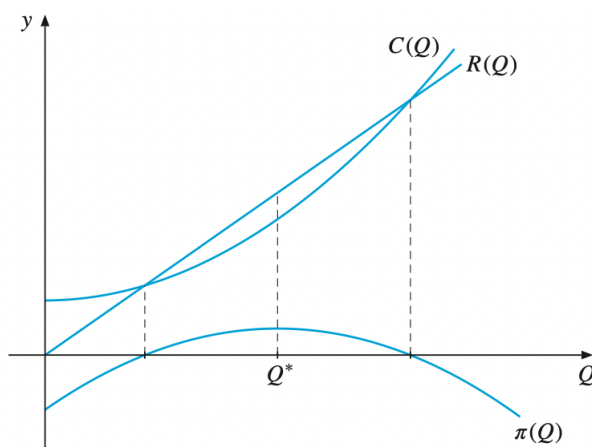
Fig. 4.4. Difference function

The function F defined where $g(x) \neq 0$ by $F(x) = f(x)/g(x)$ is called the **quotient** of f and g, and we write $F = f/g$.

4.2.3 Composite functions

Suppose the demand for a commodity is a function x of its price p. Suppose that price p is not constant, but depends on time t. Then it is natural to regard x as a function of t.

In general, if y is a function of u, and u is a function of x, then y can be regarded as a function of x. We call y a **composite function** of x. If we denote the two functions involved by f and g, with $y = f(u)$ and $u = g(x)$, then we can replace u by $g(x)$ and so write y in the form:

$$y = f(g(x))$$

When computing y, we first apply g to x to obtain $g(x)$, and then we apply f to $g(x)$. Here $g(x)$ is called the **kernel**, or **interior function**, while f is called the **exterior function**.

The function that maps x to $f(g(x))$ is often denoted by $f \circ g$. This is read as "f of g" or "f after g", and is called the **composition** of f with g. Correspondingly, $g \circ f$ denotes the function that maps x to $g(f(x))$. Thus, we have:

$$(f \circ g)(x) = f(g(x)) \quad and \quad (g \circ f)(x) = g(f(x))$$

Note that $f \circ g$ and $g \circ f$ are usually quite different functions. For instance, if $g(x) = 2 - x^2$ and $f(u) = u^3$, then $(f \circ g)(x) = (2 - x^2)^3$, whereas $(g \circ f)(x) = 2 - (x^3)^2 = 2 - x^6$; the two resulting polynomials are not the same.

Example 4.4. Write the followings as composite functions:

(a) $y = (x^3 + x^2)^{50}$

(b) $y = e^{-(x-\mu)^2}$

Solution:

(a) Given a value of x, first compute $x^3 + x^2$, which gives the interior function, $g(x) = x^3 + x^2$. Then take the 50th power of the result, so the exterior function is $f(u) = u^{50}$. Hence,

$$f(g(x)) = f(x^3 + x^2) = (x^3 + x^2)^{50}$$

(b) We can choose the interior function as $g(x) = -(x - \mu)^2$ and the exterior function as $f(u) = e^u$. Alternatively, we could choose $g(x) = (x - \mu)^2$ and $f(u) = e^{-u}$.

4.2.4   Symmetry

The function $f(x) = x^2$ satisfies $f(-x) = f(x)$, as indeed does any even power x^{2n}, with n an integer, positive or negative. So if $f(-x) = f(x)$ for all x in the domain of f, then f is called an **even function**. This condition implies that the graph of f is **symmetric** about the y-axis as shown in Fig. 4.5.
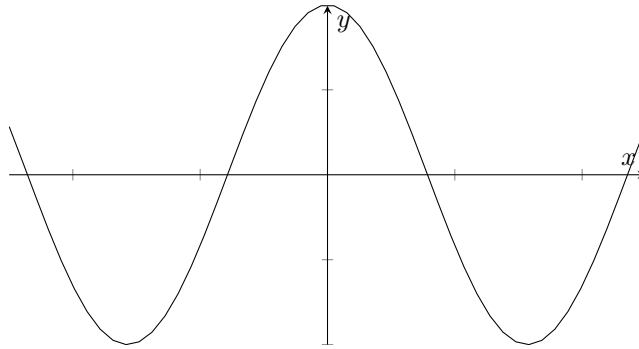
Fig. 4.5. Even function

On the other hand, any odd power x^{2n+1} such as $f(x) =^3$ satisfies $f(-x) = -f(x)$. So if $f(-x) = -f(x)$ for all x in the domain of f, implying that the graph of f is symmetric about the origin, as shown in Fig. 4.6, then f is called an **odd function**.
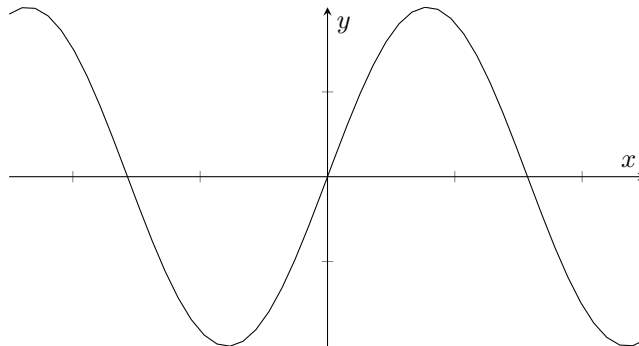
Fig. 4.6. Odd function

Finally, f is symmetric about a if $f(a + x) = f(a - x)$ for all x.The graph of f is then symmetric about the line $x = a$. For example, the quadratic function $f(x) = ax^2 + bx + c$ is symmetric about $x = -b/2a$. The function $y = e^{-(x-\mu)^2}$ is symmetric about $x = \mu$.

Ex. 4.4. Assuming $x > 0$, show graphically how to find the graph of $y = \frac{1}{4}x^2 + \frac{1}{x}$, by adding the graph of 1/x to the graph of $y = \frac{1}{4}x^2$.

Ex. 4.5. Sketch the graphs of the following functions:

(a) $y = \sqrt{x} - x$

(b) $y = e^x + e^{-x}$

(c) $y = e^{-x^2} + x$

Ex. 4.6. If $f(x) = 3x - x^3$ and $g(x) = x^3$, compute:$(f + g)(x)$, $(f - g)(x)$, $(fg)(x)$, $(f/g)(x)$, $f(g(1))$, and $g(f(1))$.

Ex. 4.7. Let $f(x) = 3x + 7$. Compute $f(f(x))$, and find the value x^* when $f(f(x^*)) = 100$.

4.3 Inverse functions

Suppose that the demand quantity D for a commodity depends on the price per unit P according to $D = 30/P^{1/3}$. This formula means that the demand D is corresponding to a given price P.

If we look at the matter from a producer's point of view, however, it may be more natural to treat the output as something it can choose and consider the resulting price. This functional relationship is obtained by solving $D = 30/P^{1/3}$ for P.

The two variables D and P in this example are related in a way that allows each to be regarded as a function of the other. In fact, the two functions:

$$f(P) = 30p^{-1/3} \quad and \quad g(D) = 27000D^{-3}$$

are **inverses** of each other.

In general, Let f be a function with domain A and range B. If and only if f is one-to-one, it has an inverse function g with domain B and range A. The function g is given by the following rule: For each y in B, the value $g(y)$ is the unique number x in A such that $f(x) = y$. Then

$$g(y) = x \Leftrightarrow y = f(x)(x \in A, y \in B)$$

Example 4.5. Solve the following equations for x and find the corresponding inverse functions:

(a) $y = 4x - 3$

(b) $y = \sqrt[5]{x + 1}$

(c) $y = \frac{3x-1}{x+4}$

Solution:

(a) Solving the equation for x, we have the following equivalences:

$$y = 4x - 3 \Leftrightarrow 4x = y + 3 \Leftrightarrow x = \frac{1}{4}y + \frac{3}{4}$$

for all x and y. We conclude that $y = 4x - 3$ and $g(y) = \frac{1}{4}y + \frac{3}{4}$ are inverses of each other.

(b) We begin by raising each side to the fifth power and so obtain the equivalences

$$y = \sqrt[5]{x + 1} \Leftrightarrow y^5 = x + 1 \Leftrightarrow x = y^5 - 1$$

These are valid for all x and all y. Hence, we have shown that $f(x) = \sqrt[5]{x + 1}$ and $g(y) = y^5 - 1$ are inverses of each other.

(c) Here we begin by multiplying both sides of the equation by the denominator and rearranging. Then we will obtain

$$x = \frac{4y + 1}{3 - y}$$

We conclude that $f(x) = (3x - 1)/(x + 4)$ and $g(y) = (4y + 1)/(3 - y)$ are inverses of each other. Observe that f is only defined for $x \neq -4$, and g is only defined for $y \neq 3$.

Graphically, suppose two functions f and g are inverses of each other. Provided that the scales of the coordinate axes are the same, the graphs of $y = f(x)$ and $y = g(x)$ are symmetric about the line $y = x$.

Example 4.6. Consider the functions $f(x) = 4x - 3$ and $g(x) = \frac{1}{4}x + \frac{3}{4}$ that are inverses of each other.
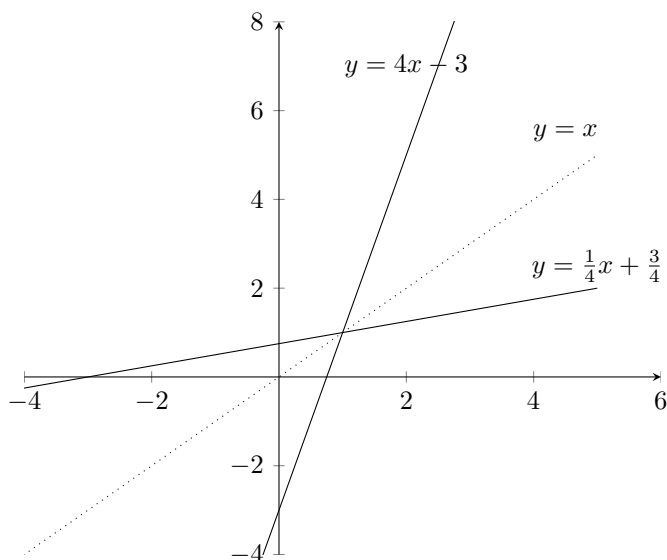
Fig. 4.7. Geometric characterisation of inverse functions

Ex. 4.8. Demand D as a function of price P is given by $D = \frac{32}{5} - \frac{3}{10}P$. Solve the equation for P and find the inverse function.

Ex. 4.9. Find the domains, ranges, and inverses of the functions given by the following formulas:

(a) $y = -3x$

(b) $y = 1/x$

(c) $y = x^3$

(d) $y = \sqrt{\sqrt{x} - 2}$

Ex. 4.10. Why does $f(x) = x^2$, for x in $(-\infty, \infty)$, have no inverse function? Show that f restricted to $[0, \infty)$ has an inverse, and find that inverse.

Ex. 4.11. Find inverses of the following functions, where x is the independent variable:

(a) $f(x) = (x^3 - 1)^{1/3}$

(b) $y = \frac{x+1}{x-2}$

(c) $f(x) = (1 - x^3)^{1/5} + 2$

Ex. 4.12. The functions defined by the following formulas are strictly increasing in their domains. Find the domain of each inverse function, and a formula for the corresponding inverse.

(a) $y = e^{x+4}$

(b) $y = \ln x - 4, \ x > 0$

(c) $y = \ln(2 + e^{x-3})$

Ex. 4.13. Find the inverse of $f(x) = \frac{1}{2}(e^x - e^{-x})$. (Hint: Solve a quadratic equation in $z = e^x$.)

4.4 Graphs of equations

The equations $x\sqrt{y} = 2$, $x^2 + y^2 = 16$, $and \ y^3 + 3x^2y = 13$ are three examples of equations in two variables x and y. A solution of such an equation is an ordered pair (a, b) such that the equation is satisfied when we replace x by a and y by b. The solution set of the equation is the set of all solutions. Representing all pairs in the solution set in a Cartesian coordinate system gives a set called the graph of the equation.

Example 4.7. Find some solutions of the first two of the equations, and try to sketch their graphs.
From $x\sqrt{y} = 2$ we obtain $y = 4/x^2$. Hence it is easy to find corresponding values for x and y as given in Table 4.1:

x	1	2	4	6
y	4	1	1/4	1/9

Table 4.1: Solutions of $x\sqrt{y} = 2$

For $x^2 + y^2 = 16$, if $y = 0$, $x^2 = 16$, so $x = \pm 4$. Thus (4,0) and (-4,0) are two solutions. Table 4.2 combines these with some other solutions.

x	-4	-3	-1	0	1	3	4
y	0	$\pm\sqrt{7}$	$\pm\sqrt{15}$	± 4	$\pm\sqrt{15}$	$\pm\sqrt{7}$	0

Table 4.2: Solutions of $x^2 + y^2 = 16$
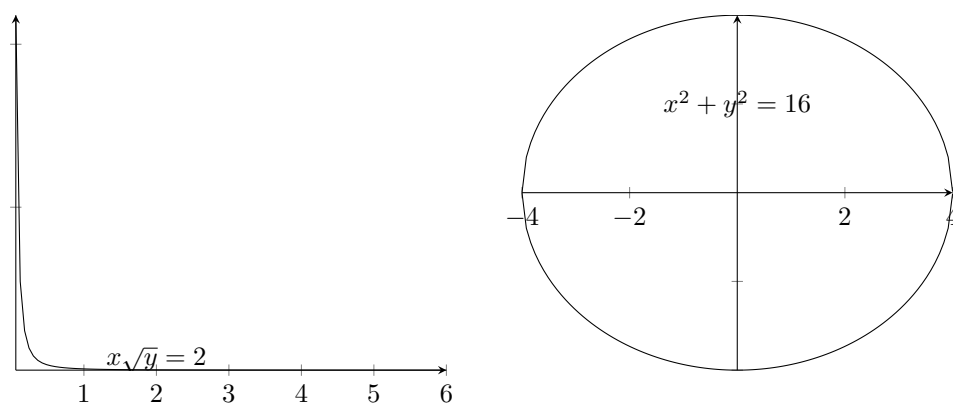
4.4.1   Vertical line test

By definition, a function assigns to each point x in the domain only one y-value. The graph of a function, therefore, has the property that a vertical line through any point on the x-axis has at most one point of intersection with the graph.
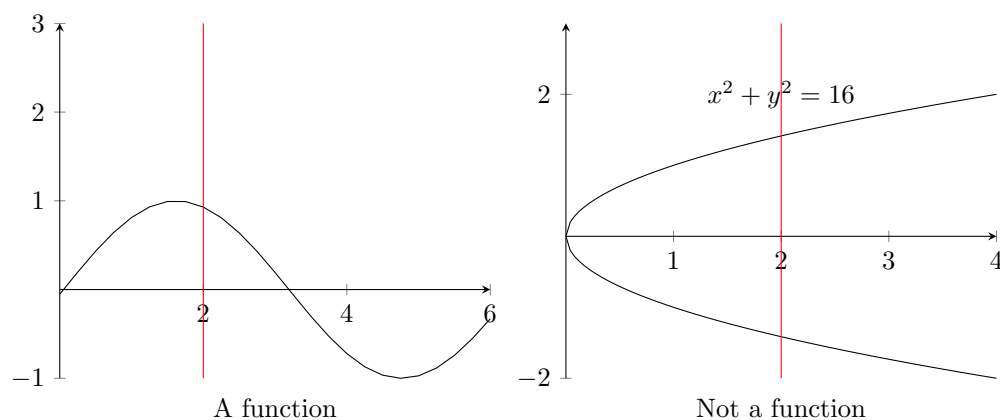
4.4.2 Compound functions

A function may be defined in several pieces, by giving a separate formula for each of a number of disjoint parts of the domain. Two examples of such compound functions are presented.

Example 4.8. Draw the graph of the function f defined by

$$f(x) = \begin{cases} -x \ for \ x \leq 0 \\ x^2 \ for \ 0 < x \leq 1 \\ 1.5 \ for \ x > 1 \end{cases}$$
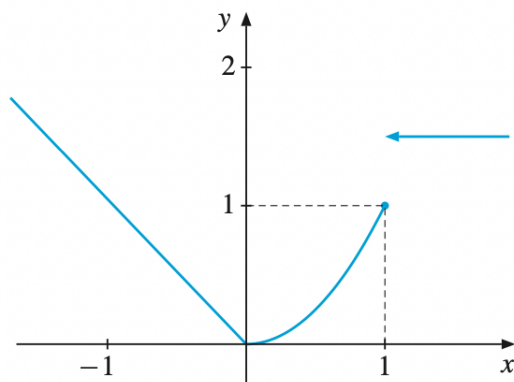
Fig. 4.8. Example of compound function

Ex. 4.14. Find some particular solutions of the following two equations, then sketch their graphs:

$$(a) \ x^2 + 2y^2 = 6 \qquad\qquad (b) \ y^2 - x^2 = 1$$

Ex. 4.15. The function F is defined for all $r \geq 0$ by the following formulas:

$$F(r) = \begin{cases} 0 & for \ r \leq 7500 \\ 0.044(r - 7500) & for \ r > 7500 \end{cases}$$

Compute $F(100000)$, and sketch the graph of F.

4.5   Distance in the plane

Let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ be two points in the xy-plane. Then by Pythagoras's theorem, the distance d between P_1 and P_2 satisfies the equation $d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$.

In other words, the distance between the points (x_1, y_1) and (x_2, y_2) is

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Example 4.9. Find the distance d between $P_1 = (-4, 3)$ and $P_2 = (5, -1)$. **Solution**:

$$d = \sqrt{(5 - (-4))^2 + (-1 - 3)^2} = \sqrt{9^2 + (-4)^2} = \sqrt{97}$$

4.5.1   Circles

The equation of a circle with centre at (a, b) and radius r is $(x-a)^2 + (y-b)^2 = r^2$.

Example 4.10. Find the equation of the circle with centre at $(-4, 1)$ and radius 3. **Solution**: the equation of the circle is:

$$(x + 4)^2 + (y - 1)^2 = 9$$
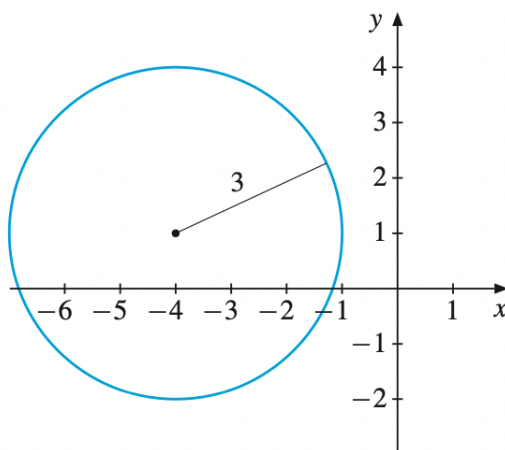
Fig. 4.9. Circle with centre at $(-4, 1)$ and radius 3

4.5.2   Ellipses and hyperbolas

The simplest type of **ellipse** has the equation:

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1$$
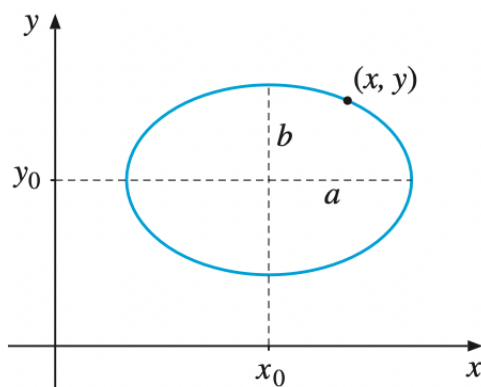
Fig. 4.10. Example of Ellipse

Figure 4.11 shows the graphs of two **hyperbolas**:

$$(a) \quad \frac{(x - x_0)^2}{a^2} - \frac{(y - y_0)^2}{b^2} = 1 \qquad (b) \quad \frac{(x - x_0)^2}{a^2} - \frac{(y - y_0)^2}{b^2} = -1$$
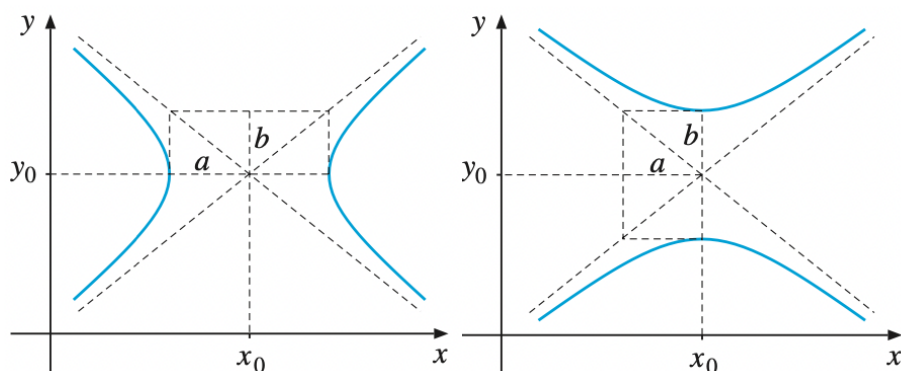
Fig. 4.11. Examples of Hyperbola

These have **asymptotes** which are like tangents at infinity, represented by the same two dashed lines in both figures. Their equations are $y - y_0 = \pm(b/a)(x - x_0)$.

Finally, consider the graph of the general quadratic equation

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

where A, B, and C are not all 0. This will have one of the following shapes:

(i) If $4AC > B^2$, either an ellipse (possibly a circle), or a single point, or empty.

(ii) If $4AC = B^2$, either a parabola, or one line or two parallel lines, or empty.

(iii) If $4AC < B^2$, either a hyperbola, or two intersecting lines.
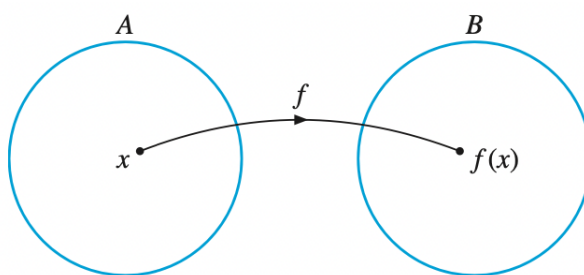
Fig. 4.12. A function from A to B

Ex. 4.16. Determine the distances between the following pairs of points:

$(a)(1,3)$ and $(2,4)$ $\qquad\qquad$ $(b)(-1,2)$ and $(-3,3)$

$(c)(3/2,-2)$ and $(-5,1)$ $\qquad\qquad$ $(d)(x,y)$ and $(2x,y+3)$

$(e)(a,b)$ and $(-a,b)$ $\qquad\qquad$ $(f)(a,3)$ and $(2+a,5)$

Ex. 4.17. The distance between $(2, 4)$ and $(5, y)$ is $\sqrt{13}$. Find y, and explain geometrically why there must be two values of y.

Ex. 4.18. Find the equations of:

(a) The circle with centre at (2, 3) and radius 4.

(b) The circle with centre at (2, 5) and one point at (-1, 3).

Ex. 4.19. Find the centre and the radius of the two circles with equations:

(a) $x^2 + y^2 + 10x - 6y + 30 = 0$

(b) $3x^2 + 3y^2 + 18x - 24y = -39$

4.6   General functions

So far we have studied functions of one variable. Yet a realistic description of many economic phenomena requires considering a large number of variables simultaneously.
An extensive discussion of functions of several variables begins in Chapter 9. This section introduces an even more general type of function.
The general definition of a **function** is a rule which to each element in a set A associates one, and only one, element in a set B.
   The particular value $f(x)$ is often called the **image** of the element x by the function f. If each element of B is the image of at most one element in A, the function f is called **one-to-one**. Otherwise, if one or more elements of B are the images of more than one element in A, the function f is **many-to-one**.

Ex. 4.20. Which of the following rules define functions?

(a) The rule that assigns to each person in a classroom his or her height.

(b) The rule that assigns to each mother her youngest child alive today.

(c) The rule that assigns the perimeter of a rectangle to its area.

(d) The rule that assigns the surface area of a spherical ball to its volume.

(e) The rule that assigns the pair of numbers (x + 3, y) to the pair of numbers (x, y).

4.7 Review Exercises

Ex. 4.21. Use the graph of $y = |x|$ and the rules for shifting graphs to sketch the graphs of the following functions:

(a) $y = |x| + 1$

(b) $y = |x + 3|$

(c) $y = 3 - |x + 1|$

Ex. 4.22. If $f(x) = x^3 - 2$ and $g(x) = (1 - x)x^2$, compute the six functions: $(f + g)(x)$; $(f - g)(x)$; $(fg)(x)$; $(f/g)(x)$; $f(g(1))$; and $g(f(1))$.

Ex. 4.23. Consider the demand and supply curves $D = 150 - 12P$ and $S = 20 + 2P$.

(a) Find the equilibrium price P^*, and the corresponding quantity Q^*.

(b) Suppose a tax of \$2 per unit is imposed on the producer's output. How will this influence the equilibrium price?

(c) Compute the total revenue obtained by the producer before the tax is imposed(R^*) and after($\hat{R}$).

Ex. 4.24. The following functions are strictly increasing in their domains. Find the domains of their inverses and formulas for the inverses, using x as the free variable.

(a) $f(x) = 3 + \ln(e^x - 2)$, for $x > \ln 2$

(b) $f(x) = \frac{a}{e^{-\lambda x} + a}$, where a and λ are positive, for $x \in (-\infty, \infty)$

Ex. 4.25. Determine the distances between the following pairs of points:

$$(a)(-4, 4) \; and \; (3, 8) \qquad\qquad (b)(2a, 3b) \; and \; (2 - a, 3b)$$

Ex. 4.26. Find the equations of the circles with:

(a) centre at (2,-3) and radius 5

(b) centre at (-2,2) and passing through(-10,1)

Ex. 4.27. A point P moves in the plane so that it is always equidistant between the two points A $= (3, 2)$ and B $= (5, -4)$. Find a simple equation that the coordinates (x, y) of P must satisfy. (Hint: Compute the square of the distance from P to the points A and B, respectively.)

Chapter 5

Differentiation

5.1 Slopes of curves

When we study the graph of a function, we would like to have a precise measure of its steepness at a point. We know for linear functions in the form of $y = px + q$, its slope is denoted by p.

However, for an arbitrary function, we define the **slope** of a curve *at a particular point* as the slope of the tangent to the point.

As shown in the graph below, the same curve will have different slopes at different points. Suppose point P has coordinates $(a, f(a)$. The slope of the curve at point P is called the **derivative** of $f(x)$ at $x = a$, denoted by $f'(a)$.

5.2 Tangents and derivatives

Consider a point P on a curve in the xy-plane. Take another point Q on the curve. The entire straight line through P and Q is called a **secant**. If we keep P fixed, but let Q move along the curve toward P, then the secant will rotate around P. The limiting straight line PT toward which the secant tends is called the **tangent** to the curve at P.

Now suppose the x-coordinate of point Q is $a + h$, where h is a small number
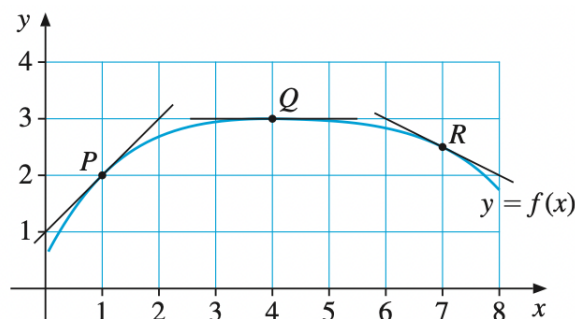
Fig. 5.1. Example of slopes at different points
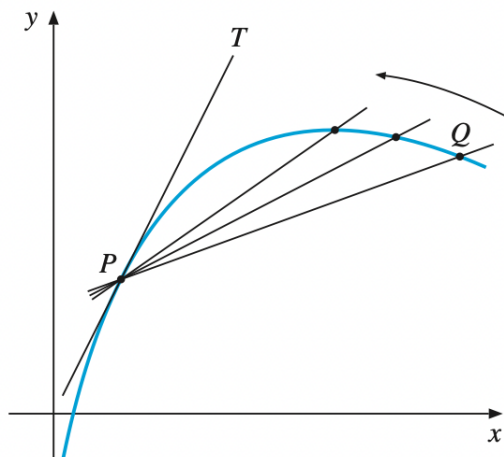
Fig. 5.2. Secants and the tangent

different from zero. The slope of the secant PQ is therefore:

$$\frac{f(a+h)-f(a)}{h}$$

This fraction is called a **Newton quotient** of f. When h tends to zero, the secant PQ tends to the tangent. This suggests that we can define the slope of the tangent at P as the number that the slope of the secant approaches as h tends to 0. Formally:

The derivative of function f at point a, denoted by $f'(a)$, is

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

The equation for the tangent to the graph of $y = f(x)$ at the point $(a, f(a))$ is

$$y - f(a) = f'(a)(x - a)$$

Example 5.1. Use the definition of slope to compute $f'(x)$ when $f(x) = x^2$. Find in particular $f'(1/2)$ and $f'(-1)$. Give geometric interpretations, and find the equation for the tangent at each of the points $(1/2, 1/4)$ and $(-1, 1)$.
Solution:
For $f(x) = x^2$, we have $f(x+h) = (x+h)^2 = x^2 + 2xh + h^2$, so

$$f(x+h) - f(x) = (x^2 + 2xh + h^2) - x^2 = 2xh + h^2$$

For all $h \neq 0$, the Newton quotient is:

$$\frac{f(x+h)-f(x)}{h} = \frac{2xh + h^2}{h} = 2x + h$$

Therefore we can obtain

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0}(2x + h) = 2x$$
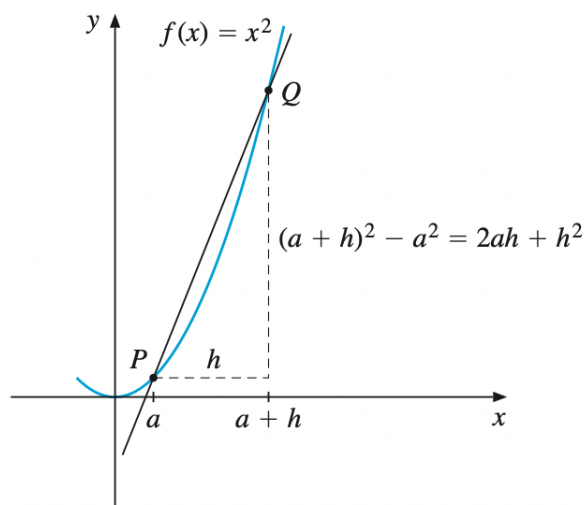
Fig. 5.3. Geometric interpretation

When $x = 1/2$, we obtain $f'(1/2) = 2 \times 1/2 = 1$. Similarly, $f'(-1) = 2 \times (-1) = -2$.

At $(1/2, 1/4)$, the equation of the tangent is $y - 1/4 = 1 \times (x - 1/2)$ or $y = x - 1/4$. Similarly, the equation of the tangent at point $(-1, 1)$ is $y = -2x - 1$.

A commonly used form of notation for derivatives is the **differential notation** due to Leibniz. If $y = f(x)$, then in place of $f'(x)$, we write $\frac{dy}{dx}$, $\frac{df(x)}{dx}$ or $\frac{d}{dx} f(x)$.

We can think of the symbol "d/dx" as an instruction to differentiate what follows with respect to x. Differentiation occurs so often in mathematics that it has become standard to use w.r.t. as an abbreviation for with respect to.

When we use letters other than f, x, and y, the notation for the derivative changes accordingly. For example: $P(t) = t^2 \Rightarrow P'(t) = 2t$; $Y = K^3 \Rightarrow Y' = 3K^2$; and $A = r^2 \Rightarrow dA = 2r$.

Ex. 5.1. Let $f(x) = 3x^2 + 2x - 1$.

(a) Show that $f(x + h) - f(x) = 6x + 2 + 3h$ for $h \neq 0$, and use this result to find $f'(x)$.

(b) Find in particular $f'(0)$, $f'(-2)$, and $f'(3)$. Find also the equation of the tangent to the graph at the point $(0, -1)$.

Ex. 5.2. The demand function for a commodity with price P is given by the formula $D(P) = a - bP$ and the cost of producing x units of a commodity is given by the formula $C(x) = p + qx^2$. Use the rule of differentiation to find $dD(P)/dP$ and $C'(x)$.

Ex. 5.3. Show that $[f(x + h) - f(x)]/h = -1/x(x + h)$, and use this to show that

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

Ex. 5.4. In each case below, find the slope of the tangent to the graph of f at the specified point:

$(a) f(x) = 3x + 2$ at $(0, 2)$ $\qquad\qquad$ $(b) f(x) = x^2 - 1$ at $(1, 0)$

$(c) f(x) = 2 + 3/x$ at $(3, 3)$ $\qquad\qquad$ $(d) f(x) = x^3 - 2x$ at $(0, 0)$

$(e) f(x) = x + 1/x$ at $(-1, -2)$ $\qquad\qquad$ $(f) f(x) = x^4$ at $(1, 1)$

Ex. 5.5. Figure 5.4 shows the graph of a function f. Determine the sign of the derivative $f'(x)$ at each of the four points a, b, c, and d.
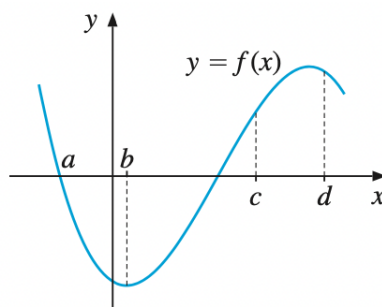
Fig. 5.4. Exercise 5.5

Ex. 5.6. Let $f(x) = \sqrt{x}$.

(a) Show that $(\sqrt{x + h} - \sqrt{x})(\sqrt{x + h} + \sqrt{x}) = h$

(b) Use the result in part (a) to show that the Newton quotient of $f(x)$ is $1/(\sqrt{x + h} + \sqrt{x})$.

(c) Use the result in part (b) to show that for $x > 0$, $f'(x) = \frac{1}{2\sqrt{x}} = \frac{1}{2} x^{-1/2}$

Ex. 5.7. Apply the results of Ex 5.5 to prove first that

$$[(x + h)^{1/3} - x^{1/3}][(x + h)^{2/3} + (x + h)^{1/3} x^{1/3} + x^{2/3}] = h$$

Then follow the argument used to solve Ex 5.5 to show that $f(x) = x^{1/3} \Rightarrow f'(x) = \frac{1}{3} x^{-2/3}$.

5.3   Increasing and decreasing functions

Assume that f is defined in an interval I and that x_1 and x_2 are numbers from that interval.

(i) If $f(x_2) \geq f(x_1)$ whenever $x_2 > x_1$, then f is *increasing* in I.

(ii) If $f(x_2) > f(x_1)$ whenever $x_2 > x_1$, then f is *strictly increasing* in I.

(iii) If $f(x_2) \leq f(x_1)$ whenever $x_2 > x_1$, then f is *decreasing* in I.

(iv) If $f(x_2) < f(x_1)$ whenever $x_2 > x_1$, then f is *strictly decreasing* in I.

We can also test whether a function is increasing or decreasing by checking the sign of its derivative:

$$f'(x) \geq 0 \text{ for all } x \text{ in the interval } I \Leftrightarrow f \text{ is increasing in } I$$
$$f'(x) \leq 0 \text{ for all } x \text{ in the interval } I \Leftrightarrow f \text{ is decreasing in } I$$
$$f'(x) = 0 \text{ for all } x \text{ in the interval } I \Leftrightarrow f \text{ is constant in } I$$

Example 5.2. Find the derivative of $f(x) = \frac{1}{2}x^2 - 2$. Then examine whether f is increasing or decreasing.
Solution:
We find that $f'(x) = x$, which is non-negative for $x \geq 0$, and non-positive if $x \leq 0$, and thus $f'(0) = 0$. We conclude that f is increasing in $[0, \infty)$ and decreasing in $(-\infty, 0]$.

Ex. 5.8. Examine whether $f(x) = x^2 - 4x + 3$ is increasing or decreasing.

Ex. 5.9. Examine whether $f(x) = -x^3 + 4x^2 - x - 6$ is increasing or decreasing.

Ex. 5.10. Show algebraically that $f(x) = x^3$ is strictly increasing by studying the sign of

$$x_2^3 - x_1^3 = (x_2 - x_1)(x_1^2 + x_1 x_2 + x_2) = (x_2 - x_1)\left[\left((x_1 + 2x_2)^2\right) + \frac{3}{4}x_2^2\right]$$

5.4 Rates of change

The derivative can be interpreted in many ways, one of which is the **rate of change**. The Newton quotient can be interpreted as the *average rate of change of f over the interval from a to a+h.*
Taking the limit as h tends to 0 gives the derivative of f at a, which we interpret as the **instantaneous rate of change**.
Sometimes we are interested in studying the proportion $f'(a)/f(a)$. This proportion can be interpreted as the **relative rate of change**.

Example 5.3. Consider a firm producing some commodity in a given period, and let $C(x)$ denote its cost of producing x units. The derivative $C'(x)$ at x is called the **marginal cost** at x. According to the definition, it is equal to

$$C'(x) = \lim_{h \to 0} \frac{C(x + h) - C(x)}{h}$$

When h is small in absolute value, we obtain the approximation

$$C'(x) \approx \frac{C(x + h) - C(x)}{h}$$

The difference $C(x + h) - C(x)$ is called the **incremental cost** of producing h units of extra output. For h small, a linear approximation to this incremental cost is $hC'(x)$, the product of the marginal cost and the change in output. This is true even when $h < 0$, signifying a decrease in output and, provided that $C'(x) > 0$, a lower cost.
Note that putting $h = 1$ makes marginal cost approximately equal to

$$C'x) \approx C(x + 1) - C(x)$$

Marginal cost is then approximately equal to the incremental cost $C(x+1) - C(x)$, that is, the additional cost of producing one more unit than x.

Example 5.4. Let $C(x)$ denote the cost in millions of dollars for removing $x\%$ of the pollution in a lake. Give an economic interpretation of the equality $C'(50) = 3$.

Solution:

Because of the linear approximation $C(50 + h) - C(50) \approx hC'(50)$, the precise interpretation of $C'(50) = 3$ is that, starting at 50%, for each extra 1% of pollution that is removed, the extra cost is about 3 million dollars. Much less precisely, $C'(50) = 3$ means that it costs about 3 million dollars extra to remove 51% instead of 50% of the pollution.

Ex. 5.11. Let $C(x) = x^2 + 3x + 100$ be the cost function of a firm. Show that when x is changed from 100 to $100 + h$, where $h \neq 0$, the average rate of change per unit of output is

$$\frac{C(100 + h) - C(100)}{h} = 203 + h$$

What is the marginal cost $C'(100)$? Use the definition of derivatives to find $C'(x)$ and in particular $C'(100)$.

Ex. 5.12. If the cost function of a firm is $C(x) = \bar{C} + cx$, give economic interpretations of the parameters c and $\bar{C}$.

Ex. 5.13. If the total saving of a country is a function $S(Y)$ of the national product Y, then $S'(Y)$ is called the **marginal propensity to save**, or MPS. Find the MPS for the following functions:

$$(a) S(Y) = \bar{S} + sY \qquad\qquad (b) S(Y) = 100 + 0.1Y + 0.0002Y^2$$

5.5   Limits

Suppose, in general, that a function f is defined for all x near a, but not necessarily at $x = a$. Then we say that the number A is the limit of $f(x)$ as x tends to a if $f(x)$ tends to A as x tends to (but is not equal to) a. We write:

$$\lim_{x \to a} f(x) = A, \quad or \ f(x) \to A \ as \ x \to a$$

Consider the function

$$F(x) = \frac{e^x - 1}{x}$$

The function is not defined for $x = 0$ but we can still find the values of $F(x)$ when x is close to zero.

x	-1	-0.1	-0.001	-0.0001	0	0.0001	0.001	0.1	1
$F(x)$	0.632	0.956	0.999	1.000	Not defined	1.000	1.001	1.052	1.718

Table 5.1: Values of $F(x)$ when x is close to 0

So we can write:

$$\lim_{x \to 0} \frac{e^x - 1}{x} = 1 \ or \ \frac{e^x - 1}{x} \to 1 \ as \ x \to 0$$

.

General rules for limits:

If $\lim_{x \to a} f(x) = A$ and $\lim_{x \to a} g(x) = B$, then:

$$\lim_{x \to a} [f(x) \pm g(x)] = A \pm B$$

$$\lim_{x \to a} [f(x) \cdot g(x)] = A \cdot B$$

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{A}{B}, \text{ if } B \neq 0$$

$$\lim_{x \to a} [f(x)]^r = A^r, \text{ if } A^r \text{ is defined and } r \text{ is a real number}$$

Example 5.5. Use the rules for limits to compute the following limits:

$$(a) \lim_{x \to -2} (x^2 + 5x) \quad (b) \lim_{x \to 4} \frac{2x^{3/2} - \sqrt{x}}{x^2 - 15} \quad (c) \lim_{x \to a} Ax^n$$

Solution:

(a) The limit can be expressed as:

$$\lim_{x \to -2} x \cdot \lim_{x \to -2} x + \lim_{x \to -2} 5 \cdot \lim_{x \to -2} x$$

So

$$\lim_{x \to -2} (x^2 + 5x) = (-2)(-2) + 5(-2) = -6$$

(b) Similar to (a), the limit can be expressed as:

$$\lim_{x \to 4} \frac{2x^{3/2} - \sqrt{x}}{x^2 - 15} = \frac{2 \lim_{x \to 4} x^{3/2} - \lim_{x \to 4} \sqrt{x}}{\lim_{x \to 4} x^2 - 15} = \frac{2 \cdot 4^{3/2} - \sqrt{4}}{4^2 - 15} = 14$$

(c) Finally,

$$\lim_{x \to a} Ax^n = \lim_{x \to a} A \cdot \lim_{x \to a} x^n = A \cdot (\lim_{x \to a} x)^n = A \cdot a^n$$

where n is a natural number.

Ex. 5.14. Determine the following by using the rules for limits:

$$(a) \lim_{x \to 0} (3 + 2x^2) \qquad (b) \lim_{x \to -1} \frac{3 + 2x}{x - 1} \qquad (c) \lim_{x \to 2} (2x^2 + 5)^3$$

$$(d) \lim_{t \to 8} (5t + t^2 - \frac{1}{8} t^3) \qquad (e) \lim_{y \to 0} \frac{(y + 1)^5 - y^5}{y + 1} \qquad (f) \lim_{z \to -2} \frac{1/z + 2}{z}$$

Ex. 5.15. If $f(x) = x^2 + 2x$, compute the following limits:

$$(a) \lim_{x \to 1} \frac{f(x) - f(1)}{x - 1} \qquad (b) \lim_{h \to 0} \frac{f(2 + h) - f(2)}{h} \qquad (c) \lim_{h \to 0} \frac{f(a + h) - f(a - h)}{h}$$

Ex. 5.16. Compute the following limits, where in part (c) n denotes any natural number:

$$(a) \lim_{x \to 2} \frac{x^2 - 2x}{x^3 - 8} \qquad (b) \lim_{h \to 0} \frac{\sqrt[3]{27 + h} - 3}{h} \qquad (c) \lim_{x \to 1} \frac{x^n - 1}{x - 1}$$

5.6   Simple rules for differentiation

If f is a constant function, then its derivative is 0:

$$f(x) = A \Rightarrow f'(x) = 0$$

When taking derivatives, additive constants disappear while multiplicative constants are preserved:

$$y = A + f(x) \Rightarrow y' = f'(x)$$
$$y = Af(x) \Rightarrow y' = Af'(x)$$

The **power rule** is that given any constant a,

$$f(x) = x^a \Rightarrow f'(x) = ax^{a-1}$$

Example 5.6. Use the power rule to compute:

$$(a)\frac{d}{dx}(\frac{x^{100}}{100}) \qquad (b)\frac{d}{dx}(x^{-0.33}) \qquad (c)\frac{d}{dr}(-5r^{-3})$$

$$(d)\frac{d}{dp}(Ap^\alpha + B) \qquad (e)\frac{d}{dx}(\frac{A}{\sqrt{x}})$$

Solution:

(a) $\frac{d}{dx}(\frac{x^{100}}{100}) = \frac{1}{100}100x^{100-1} = x^{99}$

(b) $\frac{d}{dx}(x^{-0.33}) = -0.33x^{-0.33-1} = -0.33x^{-1.33}$

(c) $\frac{d}{dr}(-5r^{-3}) = (-5)(-3)r^{-3-1} = 15r^{-4}$

(d) $\frac{d}{dp}(Ap^\alpha + B) = A\alpha p^{\alpha-1}$

(e) $\frac{d}{dx}(\frac{A}{\sqrt{x}}) = \frac{d}{dx}(Ax^{-1/2}) = A(-\frac{1}{2})x^{-1/2-1} = -\frac{1}{2}Ax^{-3/2}$

Ex. 5.17. Compute the following:

$$(a)\frac{d}{dr}(4\pi r^2) \qquad (b)\frac{d}{dy}(Ay^{b+1}) \qquad (c)\frac{d}{dA}\left(\frac{1}{A^2\sqrt{A}}\right)$$

5.7   Sums, Products, and Quotients

5.7.1   Sums and Differences

Suppose f and g are both defined on a set A of real numbers. If both f and g are differentiable at x, then the sum $f + g$ and the difference $f - g$ are both differentiable at x, with

$$F(x) = f(x) \pm g(x) \Rightarrow F'(x) = f'(x) \pm g'(x)$$

Example 5.7. Compute $\frac{d}{dx}(3x^8 + x^{100}/100)$.
Solution:

$$\frac{d}{dx}(3x^8 + x^{100}/100) = \frac{d}{dx}(3x^8) + \frac{d}{dx}(x^{100}/100) = 24x^7 + x^{99}$$

5.7.2 Products

If both f and g are differentiable at the point x, then so is $F = f \cdot g$, and

$$F(x) = f(x) \cdot g(x) \Rightarrow F'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

Example 5.8. Find $h'(x)$ when $h(x) = (x^3 - x)(5x^4 + x^2)$. Confirm the answer by expanding $h(x)$ as a single polynomial, then differentiating the result.

Solution:

We see that $h(x) = f(x) \cdot g(x)$ with $f(x) = x^3 - x$ and $g(x) = 5x^4 + x^2$. Here $f'(x) = 3x^2 - 1$ and $g'(x) = 20x^3 + 2x$. Thus from the products rule:

$$h'(x) = (3x^2 - 1) \cdot (5x^4 + x^2) + (x^3 - x) \cdot (20x^3 + 2x) = 35x^6 - 20x^4 - 3x^2$$

Alternatively, expanding $h(x)$ as a polynomial gives $h(x) = 5x^7 - 4x^5 - x^3$, which gives the same derivative.

5.7.3 Quotients

If f and g are differentiable at x and $g(x) \neq 0$, then $F = f/g$ is differentiable at x, and

$$F(x) = \frac{f(x)}{g(x)} \Rightarrow F'(x) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2}$$

Example 5.9. Compute $F'(x)$ and $F'(4)$ when

$$F(x) = \frac{3x - 5}{x - 2}$$

Solution:

We apply the quotient rule with $f(x) = 3x - 5$ and $g(x) = x - 2$. Then $f'(x) = 3$ and $g'(x) = 1$. So we obtain, for $x \neq 2$:

$$F'(x) = \frac{3 \cdot (x - 1) - (3x - 5) \cdot 1}{(x - 2)^2} = \frac{-1}{(x - 2)^2}$$

To find $F'(4)$, we put $x = 4$ in the formula and obtain $F'(4) = -1/4$.

Ex. 5.18. Differentiate w.r.t x the following functions:

$$(a)\frac{3}{5}x^2 - 2x^7 + \frac{1}{8} - \sqrt{x} \qquad (b)(2x^2 - 1)(x^4 - 1) \qquad (c)\left(x^5 + \frac{1}{x}\right)(x^5 + 1)$$

Ex. 5.19. Differentiate w.r.t x the following functions:

$$(a)\frac{\sqrt{x} - 2}{\sqrt{x} + 1} \qquad (b)\frac{x^2 - 1}{x^2 + 1} \qquad (c)\frac{x^2 + x + 1}{x^2 - x + 1}$$

Ex. 5.20. For each of the following functions, determine the intervals where it is increasing.

$$(a)y = 3x^2 - 12x + 13 \qquad (b)y = \frac{1}{4}(x^4 - 6x^2) \qquad (c)y = \frac{x^2 - x^3}{2(x + 1)}$$

Ex. 5.21. Find the equations for the tangents to the graphs of the following functions at the specified points:

$(a) y = 3 - x - x^2$ at $x = 1$
$\qquad\qquad$ $(b) y = \dfrac{x^2 - 1}{x^2 + 1}$ at $x = 1$

$(c) y = \left(\dfrac{1}{x^2} + 1\right)(x^2 - 1)$ at $x = 2$
$\qquad$ $(d) y = \dfrac{x^4 + 1}{(x^2 + 1)(x + 3)}$ at $x = 0$

5.8 Chain rule

The **chain rule** states that if y is a differentiable function of u, and u is a differentiable function of x, then y is a differentiable function of x, and

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

The generalised power rule is that if $y = u^a$ and u is a differentiable function of x, then

$$y' = au^{a-1}u'$$

Example 5.10. Find dy/dx when:

(a) $y = u^5$ and $u = (1 - x^3)$

(b) $y = \frac{10}{(x^2+4x+5)^7}$

Solution:

(a) Here we apply the chain rule directly. Since $dy/du = 5u^4$ and $du/dx = -3x^2$, we have

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = 5u^4(-3x^2) = -15x^2u^4 = -15x^2(1 - x^3)^4$$

(b) If we write $u = x^2 + 4x + 5$, then $y = 10u^{-7}$. By the generalised power rule, one has

$$\frac{dy}{dx} = 10(-7)u^{-8}u' = 5u^4(-3x^2) = -70u^{-8}(2x + 4) = \frac{-140(x + 2)}{(x^2 + 4x + 5)^8}$$

5.8.1 Alternative formulation of the chain rule

Consider a composite function $y = f(g(x))$. If g is differentiable at x_0 and f is differentiable at $u_0 = g(x_0)$, then the composite function $F(x) = f(g(x))$ is differentiable at x_0, and

$$F'(x_0) = f'(u_0)g'(x_0) = f'(g(x_0))g'(x_0)$$

Example 5.11. Find the derivative of the compound function $F(x) = f(g(x))$ at $x_0 = -3$ in case $f(u) = u^3$ and $g(x) = 2 - x^2$.

$\qquad$ **Solution**:

$\qquad$ In this case we have $f'(u) = 3u^2$ and $g'(x) = -2x$. So according to the alternative formulation of the chain rule, one has $F'(-3) = f'(g(-3))g'(-3)$. Now $g(-3) = 2 - (-3)^2 = 2 - 9 = -7$; $g'(-3) = 6$; and $f'(g(-3)) = f'(-7) = 3(-7)^2 = 3 \cdot 49 = 147$. So $F'(-3) = f'(g(-3))g'(-3) = 147 \cdot 6 = 882$.

Ex. 5.22. Find the derivatives of the following functions:

$$(a) y = \frac{1}{(x^2 + x + 1)^5} \qquad (b) y = \sqrt{x + \sqrt{x + \sqrt{x}}}$$

Ex. 5.23. Suppose that $C = 20q - 40q(25 - \frac{1}{2}x)^{\frac{1}{2}}$, where q is a constant and $x < 50$. Find dC/dx.

5.9 Higher-order derivatives

The derivative f' of a function f is often called the first derivative of f. If f' is also differentiable, then we can differentiate f' in turn. The result $(f')'$ is called the second derivative, written more concisely as f''.

Similarly, a function $f(x)$ may have third, forth derivatives and so on. In general, the n-th derivative of f at x is expressed as:

$$y^{(n)} = f^{(n)}(x) \ or \ \frac{d^n y}{dx^n}$$

The number n is the **order** of the derivative.

A different form of notation is:

$$f''(x) = \frac{d^2 f(x)}{dx^2} \ or \ y'' = \frac{d^2 y}{dx^2}$$

Example 5.12. Find $f'(x)$ and $f''(x)$ when $f(x) = 2x^5 - 3x^3 + 2x$

Solution:

The rules for differentiating polynomials imply that $f'(x) = 10x^4 - 9x^2 + 2$. Then we differentiate each side of this equality to get $f''(x) = 40x^3 - 18x$.

Recall that the sign of the first derivative determines whether a function is increasing or decreasing on an interval I. Then the second derivative implies:

$$f''(x) \geq 0 \text{ on } I \Leftrightarrow f' \text{ is increasing in } I$$
$$f''(x) \leq 0 \text{ on } I \Leftrightarrow f' \text{ is decreasing in } I$$

Graphical visualisations are:
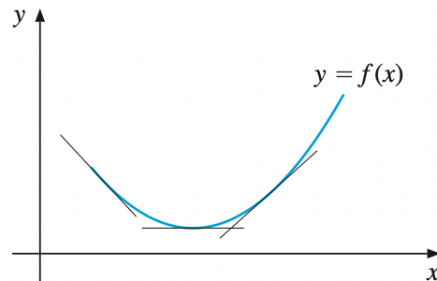
Fig. 5.5. The slope of the tangent line increases as x increases
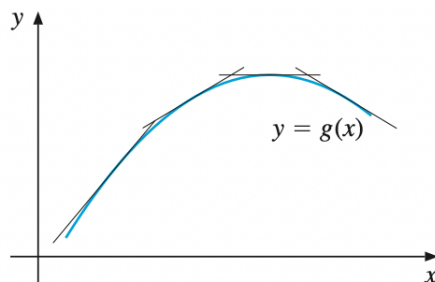
Fig. 5.6. The slope of the tangent line decreases as x increases

Suppose that f is continuous in the interval I and twice differentiable in the interior of I. Then we can introduce the following definitions:

$$f \text{ is } \textit{convex} \text{ on } I \Leftrightarrow f'' \geq 0 \text{ for all } x \text{ in } I$$
$$f \text{ is } \textit{concave} \text{ on } I \Leftrightarrow f'' \leq 0 \text{ for all } x \text{ in } I$$

Example 5.13. Check the convexity/concavity of the following functions:

$$(a) f(x) = x^2 - 2x + 2 \qquad (b) f(x) = ax^2 + bx + c$$

Solution:

(a) Here $f'(x) = 2x - 2$ so $f''(x) = 2$. Because $f''(x) > 0$ for all x, the function f is convex.

(b) Here $f'(x) = 2ax + b$, so $f''(x) = 2a$. If $a = 0$, then f is linear, so it is both concave and convex. If $a > 0$, then $f''(x) > 0$, so f is convex. If $a < 0$, then $f''(x) < 0$, so f is concave.

Ex. 5.24. Compute the second derivative of:

(a) $y = x^5 - 3x^4 + 2$

(b) $y = \sqrt{x}$

(c) $y = (1 + x^2)^{\frac{1}{2}}$

Ex. 5.25. Find $g''(2)$ when $g(t) = \frac{t^2}{t-1}$

5.10   Exponential functions

Derivative of the natural exponential function:

$$f(x) = e^x \Rightarrow f'(x) = e^x$$

Example 5.14. Find the first and second derivatives of:

$$(a) y = x^3 + e^x \qquad (b) y = x^5 e^x \qquad (c) y = e^x / x$$

Solution:

(a) We can easily find that $y' = 3x + e^x$ and $y'' = 6x + e^x$.

(b) By the product rule, $y' = 5x^4 e^x + x^5 e^x$. To find the second derivative, differentiate y' once more to obtain $y'' = 20x^3 e^x + 5x^4 e^x + 5x^4 e^x + x^5 e^x = 20x^3 e^x + 10x^4 e^x + x^5 e^x$.

(c) The quotient rule yields:

$$y' = \frac{e^x x - e^x \cdot 1}{x^2} = \frac{e^x(x-1)}{x^2}$$

Differentiating again gives:

$$y'' = \frac{(e^x x + e^x - e^x)x^2 - (e^x x - e^x)2x}{(x^2)^2} = \frac{e^x(x^2 - 2x + 2)}{x^3}$$

.

Example 5.15. For each of the following functions, find the intervals where they are increasing:

$$(a) y = \frac{e^x}{x} \qquad (b) y = x^4 e^{-2x} \qquad (c) y = xe^{-\sqrt{x}}$$

Solution:

(a) As shown in the previous example, $y' = e^x(x-1)/x^2$, so $y' \geq 0$ if and only if $x \geq 1$. Thus y is increasing in $[1, \infty)$.

(b) We can easily obtain $y' = x^3 e^{-2x}(4 - 2x)$. A sign diagram reveals that y is increasing in [0,2].

(c) The function is only defined for $x \geq 0$. Using the chain rule and product rule, the derivative of y is:

$$y' = 1 \cdot e^{-\sqrt{x}} - \frac{xe^{-\sqrt{x}}}{2\sqrt{x}} = e^{-\sqrt{x}}\left(1 - \frac{1}{2}\sqrt{x}\right)$$

. It follows that y is increasing when $x > 0$ and $1 - \frac{1}{2}\sqrt{x} \geq 0$. Therefore y is increasing in [0,4].

Now consider in general the power function where the base is no longer the constant e. We have the following rule for differentiation:

$$y = a^x = e^{(\ln a)x} \Rightarrow y' = \ln a \, e^{(\ln a)x} = a^x \ln a$$

Example 5.16. Find the derivatives of: (a) $f(x) = 10^{-x}$; and (b) $g(x) = x2^{3x}$.
Solution:

(a) Using the rule above we obtain $f'(x) = -10^{-x} \ln 10$.

(b) Rewrite $y = 2^{3x} = 2^u$, where $u = 3x$. By the chain rule:

$$y' = (2^u \ln 2)u' = (2^{3x} \ln 2) \cdot 3 = 3 \cdot 2^{3x} \ln 2$$

Finally, using the product rule we obtain

$$g'(x) = 1 \cdot 2^{3x} + x \cdot 3 \cdot 2^{3x} \ln 2 = 2^{3x}(1 + 3x \ln 2)$$

Ex. 5.26. Find the first and second derivatives of:

$$(a)y = e^{-3x} \qquad\qquad (b)y = 2e^{x^3}$$
$$(c)y = e^{1/x} \qquad\qquad (d)5e^{2x^2-3x+1}$$

Ex. 5.27. Find the intervals where the following functions are increasing:

$$(a)y = x^3 + e^{2x} \qquad (b)y = 5x^2 e^{-4x} \qquad (c)y = x^2 e^{-x^2}$$

Ex. 5.28. Find:

$$(a)\frac{d}{dx}\left(e^{e^x}\right) \qquad\qquad (b)\frac{d}{dt}\left(e^{t/2} + e^{-t/2}\right)$$
$$(c)\frac{d}{dt}\left(\frac{1}{e^t + e^{-t}}\right) \qquad\qquad (d)\frac{d}{dz}\left((e^{z^3} - 1)^{1/3}\right)$$

5.11   Logarithmic Functions

Derivative of the natural logarithmic function:

$$g(x) = \ln x \ \Rightarrow g'(x) = \frac{1}{x}$$

Example 5.17. Compute y' and y'' when:

$$(a)y = x^3 + \ln x \qquad (b)y = x^2 \ln x \qquad (c)y = \ln x/x$$

Solution:

(a) We find easily that $y' = 3x^2 + 1/x$. Furthermore, $y'' = 6x - 1/x^2$.

(b) The product rule gives $y' = 2x \ln x + x^2(1/x) = 2x \ln x + x$. Then $y'' = 2 \ln x + 2x(1/x) + 1 = 2 \ln x + 3$.

(c) Here we use the quotient rule:

$$y' = \frac{(1/x)x - \ln x \cdot 1}{x^2} = \frac{1 - \ln x}{x^2}$$

Differentiating again gives:

$$y'' = \frac{-(1/x)x^2 - (1 - \ln x)2x}{(x^2)^2} = \frac{2 \ln x - 3}{x^3}$$

In general, suppose that $y = \ln h(x)$, where $h(x)$ is differentiable and positive. Then:

$$y = \ln h(x) \ \Rightarrow y' = \frac{h'(x)}{h(x)}$$

Example 5.18. Find the domains of the following functions and compute their derivatives:
$$(a)y = \ln(1-x) \qquad (b)y = \ln\left(\frac{x-1}{x+1}\right) - \frac{1}{4}x$$

Solution:

(a) $\ln(1-x)$ is defined if $1 - x > 0$, that is if $x < 1$. To find its derivative, we use the rule earlier, with $h(x) = 1 - x$. Then $h'(x) = -1$, and

$$y' = \frac{-1}{1-x} = \frac{1}{x-1}$$

(b) We can write $y = \ln u - \frac{1}{4}x$, where $u = (x-1)/(x+1)$. For the function to be defined, we require that $u > 0$. A sign diagram shows that this is satisfied if $x < -1$ or $x > 1$. Then we obtain:

$$y' = \frac{u'}{u} - \frac{1}{4}$$

where
$$u' = \frac{1 \cdot (x+1) - 1 \cdot (x-1)}{(x+1)^2} = \frac{2}{(x+1)^2}$$

So
$$y' = \frac{2(x+1)}{(x+1)^2(x-1)} - \frac{1}{4} = \frac{9-x^2}{4(x^2-1)} = \frac{(3-x)(3+x)}{4(x-1)(x+1)}$$

Similar to power functions, when the base of the logarithmic function is not e, differentiating gives:

$$y = \log_a x \implies y' = \frac{1}{\ln a}\frac{1}{x}$$

5.11.1   Approximating the number e

If $g(x) = \ln x$, then $g'(x) = 1/x$, and, in particular, $g'(1) = 1$. We use in turn: (i) the definition of $g'(1)$; (ii) the fact that $\ln 1 = 0$; (iii) the rule $\ln x^p = p \ln x$. The result is

$$1 = g'(1) = \lim_{h\to 0}\frac{\ln(1+h) - \ln 1}{h} = \lim_{h\to 0}\frac{1}{h}\ln(1+h) = \lim_{h\to 0}\ln(1+h)^{1/h}$$

Taking exponential of both sides gives:

$$e = \lim_{h\to 0}(1+h)^{1/h}$$

Ex. 5.29. Find the derivatives of:

$$(a)y = x^3(\ln x)^2 \qquad\qquad\qquad (b)y = \frac{x^2}{\ln x}$$

$$(c)y = (\ln x)^{10} \qquad\qquad\qquad (d)y = (\ln x + 3x)^2$$

Ex. 5.30. Determine the domains of the functions defined by:

$$(a)y = \ln(x^2 - 1) \qquad (b)y = \ln(\ln x) \qquad (c)y = \frac{1}{\ln(\ln x) - 1}$$

Ex. 5.31. Prove that if u and v are differentiable functions of x, and $u > 0$, then

$$y = u^v \Rightarrow y' = u^v \left(v' \ln u + \frac{vu'}{u} \right)$$

5.12   Review Exercises

Ex. 5.32. Let $f(x) = x^2 - x + 2$. Show that $[f(x+h) - f(x)]/h = 2x - 1 + h$, and use this result to find $f'(x)$.

Ex. 5.33. Let $C(Q)$ denote the cost of producing Q units per month of a commodity. What is the interpretation of $C'(1000) = 25$? Suppose the price obtained per unit is fixed at 30 and that the current output per month is 1000. Is it profitable to increase production?

Ex. 5.34. For each of the following functions, find the equation for the tangent to the graph at the specified point:

$$(a)y = \sqrt{x} - x^2 \text{ at } x = 4 \qquad (b)y = \frac{x^2 - x^3}{x + 3} \text{ at } x = 1$$

Ex. 5.35. If $R = S^\alpha$, $S = 1 + \beta K^\gamma$, and $K = At^p + B$, find an expression for dR/dt.

Ex. 5.36. Find the intervals where the following functions are increasing:

$$(a)y = \ln(x)^2 - 4 \qquad (b)y = \ln(e^x + e^{-x}) \qquad (c)y = x - \frac{3}{2}\ln(x^2 + 2)$$

Ex. 5.37.   (a) Suppose $\pi(Q) = QP(Q) - cQ$, where P is a differentiable function and c is a constant. Find an expression for dπ/dQ.

  (b) Suppose $\pi(L) = PF(L) - wL$, where F is a differentiable function and P and w are constants. Find an expression for dπ/dL.

Chapter 6

Derivatives in Use

6.1 Implicit differentiation

In the last chapter we learned how to differentiate functions given by explicit formulas like $y = f(x)$. Now we consider how to differentiate functions defined implicitly by an equation such as $g(x, y) = c$, where c is a constant.

Example 6.1. Find the slope of $y^3 + 3x^2y = 13$ at the point (2,1).
   Solution:
   Using the product rule, differentiating gives:

$$3y^2y' + 6xy + 3x^2y' = 0$$

Solving the equation for y' yields

$$y' = \frac{-6xy}{3x^2 + 3y^2} = \frac{-2xy}{x^2 + y^2}$$

For $x = 2$ and $y = 1$, $y' = -4/5$.

   To find y' when an equation relates two variables x and y:

   (i) Differentiate each side of the equation w.r.t. x, considering y as a function of x.

   (ii) Solve the resulting equation for y'.

Example 6.2. The equation $x^2y^3 + (y + 1)e^{-x} = x + 2$ defines y as a differentiable function of x in a neighbourhood of $(x, y) = (0, 1)$. Compute y' at this point.
   Solution:
   Implicit differentiation w.r.t. x gives

$$2xy^3 + x^2 3y^2y' + y'e^{-x} + (y + 1)(-e^{-x}) = 1$$

Inserting $x = 0$ and $y = 1$ yields $y' + 2(-1) = 1$, implying that $y' = 3$.

Ex. 6.1. Find dy/dx and d^2y/dx^2 by implicit differentiation when:

$$(a)x - y + 3xy = 2 \qquad (b)y^5 = x^6$$

Ex.   6.2. Suppose that y is a differentiable function of x that satisfies the equation $2x^2 + 6xy + y^2 = 18$. Find y' and y'' at the point $(x, y) = (1, 2)$.

Ex. 6.3. The elegant curve shown below is known as a lemniscate. In the late 1600s, the Swiss mathematician Johann Bernoulli (1667–1748) discovered that it is the graph of the equation
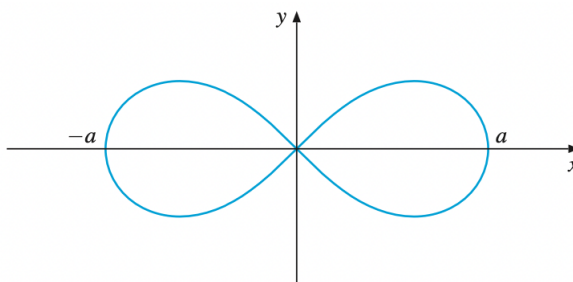
Fig. 6.1. A Lemniscate

$$(x^2 + y^2)^2 = a^2(x^2 - y^2)$$

where a is a positive constant.

(a) Find the slope of the tangent to this curve at any point (x, y) where $y \neq 0$.

(b) Determine those points on the curve where the tangent is parallel to the x-axis.

6.2   Economic examples

Example 6.3. A general version of the standard macroeconomic model for determining national income is that: (i) $Y = C + \bar{I}$; and (ii) $C = f(Y)$. Assume that $f'(Y)$, the **marginal propensity to consume**, exists and lies between 0 and 1.

(a) Suppose first that $C = f(Y) = 95.05 + 0.712Y$, use equations (i) and (ii) to find Y in terms of $\bar{I}$.

(b) Inserting the expression for C from (ii) into (i) gives $Y = f(Y) + \bar{I}$. Suppose that this equation defines Y as a differentiable function of I. Find an expression for $dY/d\bar{I}$.

(c) Assuming that $f''(Y)$ also exists, find $Y'' = d^2Y/d\bar{I}^2$

Solution:

(a) In this case, we find that $Y = 95.05 + 0.712Y + \bar{I}$. Solving for Y yields

$$Y = (95.05 + \bar{I})/(1 - 0.712) \approx 3.47\bar{I} + 330.03$$

In particular, $dY/dI \approx 3.47$, so if I is increased by \$1 billion, then the corresponding increase in GDP is approximately \$3.47 billion.

(b) Differentiating $Y = f(Y) + \bar{I}$ w.r.t. $\bar{I}$, and using the chain rule, we have

$$\frac{dY}{d\bar{I}} = f'(Y)\frac{dY}{d\bar{I}} + 1 \ or \ \frac{dY}{d\bar{I}}[1 - f'(Y)] = 1$$

Solving for $\frac{dY}{d\bar{I}}$ yields

$$\frac{dY}{d\bar{I}} = \frac{1}{1 - f'(Y)}$$

In this model, a \$1 billion increase in investment will always lead to a more than \$1 billion increase in GDP. Also, the greater is $f'(Y)$, the marginal propensity to consume, the smaller is $1 - f'(Y)$, and so the greater is $dY/d\bar{I}$.

(c) Differentiate the first equation implicitly w.r.t. $\bar{I}$:

$$\frac{d}{d\bar{I}}\left(f'(Y)\frac{dY}{d\bar{I}}\right) = f''(Y)\frac{dY}{d\bar{I}}\frac{dY}{d\bar{I}} + f'(Y)\frac{d^2Y}{d\bar{I}^2}$$

Hence,

$$\frac{d^2Y}{d\bar{I}^2} = f''(Y)\left(\frac{dY}{d\bar{I}}\right)^2 + f'(Y)\frac{d^2Y}{d\bar{I}^2}$$

Since $\frac{dY}{d\bar{I}} = 1/(1 - f'(Y))$, simple algebra yields

$$\frac{d^2Y}{d\bar{I}^2} = \frac{f''(Y)}{[1 - f'(Y)]^3}$$

Example 6.4. In the linear supply and demand model, suppose that consumers are required to pay a tax of τ per unit, thus the price they face is $P + \tau$. Then

$$D = a - b(P + \tau), \qquad S = \alpha + \beta P$$

Here a, b, α, and β are positive constants. The equilibrium price is determined by equating supply and demand, so that

$$a - b(P + \tau) = \alpha + \beta P$$

(a) The equilibrium equation implicitly defines the price P as a function of the unit tax τ. Compute $dP/d\tau$ by implicit differentiation. What is its sign? What is the sign of $(d/d\tau)(P+\tau)$? Check the result by first solving the equation for P and then finding $dP/d\tau$ explicitly.

(b) Compute tax revenue T as a function of τ. For what value of τ does the quadratic function T reach its maximum?

(c) Generalise the foregoing model by assuming that $D = f(P + \tau)$ and $S = g(P)$, where f and g are differentiable functions with $f' < 0$ and $g' > 0$. The equilibrium condition $f(P + \tau) = g(P)$ defines P implicitly as a differentiable function of τ. Find an expression for $dP/d\tau$ by implicit differentiation. Illustrate geometrically.

Solution:

(a) Differentiating w.r.t. τ yields $-b\left(\dfrac{dP}{d\tau} + 1\right) = \beta\frac{dP}{d\tau}$. Solving for $\frac{dP}{d\tau}$ gives

$$\frac{dP}{d\tau} = \frac{-b}{b + \beta}$$

We see that $dP/d\tau$ is negative. Because P is the price received by the producer, this price will go down if the tax rate τ increases. But $P + \tau$ is the price paid by the consumer. Because

$$\frac{d}{d\tau}(P + \tau) = \frac{dP}{d\tau} + 1 = \frac{-b}{b + \beta} + 1 = \frac{\beta}{b + \beta}$$

we see that $0 < d(P+\tau)/d\tau < 1$. Thus, the consumer price $P+\tau$ increases, but by less than the increase in the tax.

If we solve the equilibrium for P, we obtain

$$P = \frac{a - \alpha}{b + \beta} - \frac{b}{b + \beta}\tau$$

This equation shows that the equilibrium producer price P is a linear function of τ , the tax per unit, with slope $-b/(b + \beta)$.

(b) The total tax revenue is $T = S\tau = (\alpha + \beta P)\tau$, where P is the equilibrium price. Thus,

$$T = \left[\alpha + \beta\left(\frac{a - \alpha}{b + \beta} - \frac{b}{b + \beta}\tau\right)\right]\tau = \frac{-b\beta}{b + \beta}\tau^2 + \frac{\alpha b + \beta a}{b + \beta}\tau$$

This quadratic function has its maximum at $\tau = (\alpha b + \beta a)/2b\beta$.

(c) Differentiating the equation $f(P + \tau) = g(P)$ w.r.t. τ yields

$$f'(P + \tau)\left(\frac{dP}{d\tau} + 1\right) = g'(P)\frac{dP}{d\tau}$$

Solving for $\frac{dP}{d\tau}$ gives

$$\frac{dP}{d\tau} = \frac{f'(P + \tau)}{g'(P) - f'(P + \tau)}$$

Because $f' < 0$ and $g' > 0$, we see that $dP/d\tau$ is negative in this case as well. Moreover,

$$\frac{d}{d\tau}(P + \tau) = \frac{dP}{d\tau} + 1 = \frac{f'(P + \tau)}{g'(P) - f'(P + \tau)} + 1 = \frac{g'(P)}{g'(P) - f'(P + \tau)}$$

Again, because $f' < 0$ and $g' > 0$, this implies that $0 < d(P+\tau)/d\tau < 1$.
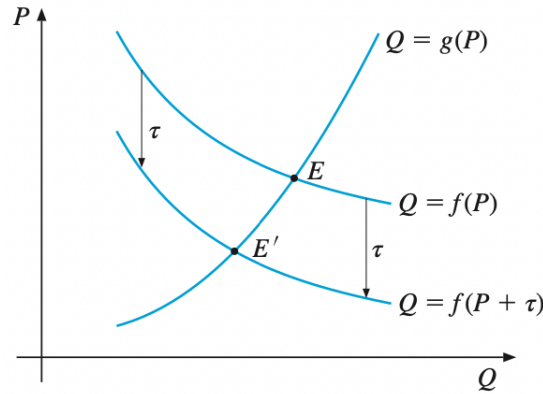
Fig. 6.2. Shift in Demand curve

Figure 6.2 has a graph which illustrates this answer. As usual in economics, we have quantity on the horizontal axis, and price on the vertical axis. The demand curve with the tax is represented by the curve $Q = f(P + \tau)$. Its graph is obtained by shifting the graph of $Q = f(P)$, or equivalently, the graph of the inverse demand curve $P = f^{-1}(Q)$, down by τ units, so it becomes $P = f^{-1}(Q) - \tau$, or $Q = f(P + \tau)$. The figure confirms that, when τ increases, the new equilibrium E' corresponds to a decreased price P. Nevertheless, $P + \tau$ increases because the decrease in P is smaller than the increase in τ.

Ex. 6.4. Consider a profit-maximising firm producing a single commodity. If the firm gets a fixed price P per unit sold, its profit from selling Q units is $\pi(Q) = PQ - C(Q)$, where $C(Q)$ is the cost function. Assume that $C'(Q) > 0$ and $C''(Q) > 0$. It can be shown that $Q = Q^* > 0$ maximises profits w.r.t. Q provided that

$$P = C'(Q^*)$$

Thus, at the optimum, marginal cost must equal the price per unit.

(a) By implicitly differentiating the equation above w.r.t. P, find an expression for dQ^*/dP.

(b) Comment on the sign of dQ^*/dP.

Ex. 6.5. Extending the standard macroeconomic model for an economy open to international trade gives:

$$(i)\, Y = C + \bar{I} + X - M; \qquad (ii)\, C = f(Y), \text{ with } 0 < f'(Y) < 1; \qquad (iii)\, M = g(Y)$$

Here X is an exogenous constant that denotes exports, whereas M denotes the volume of imports. The function g is called an **import function**, which is assumed to satisfy $0 < g'(Y) < f'(Y)$.

(a) By inserting (ii) and (iii) into (i), obtain an equation that defines Y as a function of exogenous investment $\bar{I}$.

(b) Find an expression for $dY/d\bar{I}$ by implicit differentiation. Discuss the sign of $dY/d\bar{I}$.

(c) Find an expression for $d^2Y/d\bar{I}^2$

6.3   Differentiating the inverse

If f is differentiable and strictly increasing (strictly decreasing) in an interval I, then f has an inverse function g, which is strictly increasing (strictly decreasing) in the interval $f(I)$. If x_0 is an interior point of I and $f'(x_0) \neq 0$, then g is differentiable at $y_0 = f(x_0)$, and

$$g'(y_0) = \frac{1}{f'(x_0)}$$

It is common to present the theorem in the deceptively simple way:

$$\frac{dx}{dy} = \frac{1}{dy/dx}$$

Example 6.5. Suppose the function f is defined for all real x by the following formula: $f(x) = x^5 + 3x^3 + 6x - 3$. Show that f has an inverse function g, and then, given that $f(1) = 7$, find $g'(7)$.

   Solution:

   Differentiating $f(x)$ yields $f'(x) = 5x^4 + 9x^2 + 6$. Clearly, $f'(x) > 0$ for all x, so f is strictly increasing and consequently it is one-to-one. It therefore has an inverse function g. To find $g'(7)$, we use theorem with $x_0 = 1$ and $y_0 = 7$. Since $f'(1) = 20$, we obtain $g'(7) = 1/f'(1) = 1/20$.

Ex. 6.6. The function f is defined,for $-2 \leq x \leq 2$, by the formula $f(x) = \frac{1}{3}x^3\sqrt{4 - x^2}$.

   (a) Find the intervals where f increases, and the intervals where f decreases, then sketch its graph.

   (b) Explain why f has an inverse g on $[0, \sqrt{3}]$, and find $g'(\frac{1}{3}\sqrt{3})$.   (Hint: $f(1) = \frac{1}{3}\sqrt{3}$.)

Ex. 6.7. Let f be defined by $f(x) = \ln(2 + e^{x-3})$, for all x.

   (a) Show that f is strictly increasing and find the range of f.

   (b) Find an expression for the inverse function, g, of f . Where is g defined?

   (c) Verify that $f'(3) = 1/g'(f(3))$.

Ex. 6.8. Find dx/dy when:

$$(a) y = e^{-x-5} \qquad (b) y = \ln(e^{-x} + 3) \qquad (c) xy^3 - x^3 y = 2x$$

6.4   Linear approximations

Consider a function $f(x)$ that is differentiable at $x = a$. Suppose we approximate the graph of f by its tangent line at $(a, f(a))$, as shown in Fig. 6.3. This tangent line is the graph of the function $y = p(x) = f(a) + f'(a)(x - a)$.
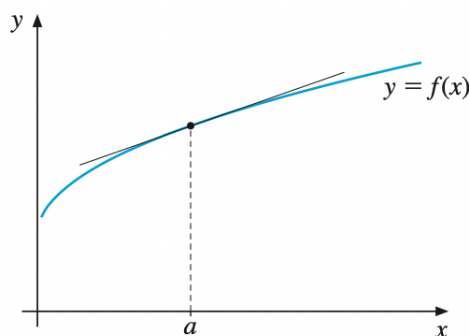
Fig. 6.3. Approximation of a function by its tangent

For x close to a,

$$f(x) \approx f(a) + f'(a)(x - a)$$

Example 6.6. Find the linear approximation to $f(x) = \sqrt[3]{x}$ about $x = 1$.

Solution:

We have $f(x) = \sqrt[3]{x} = x^{1/3}$, so $f(1) = 1$, and $f'(x) = \frac{1}{3}x^{-2/3}$, implying that $f'(1) = \frac{1}{3}$. Inserting these values into the linear approximation formula when $a = 1$ yields

$$\sqrt[3]{x} \approx f(1) + f'(1)(x - 1) = 1 + \frac{1}{3}(x - 1)$$

For example, $\sqrt[3]{1.03} \approx 1 + \frac{1}{3}(1.03 - 1) = 1.01$. The correct value to four decimals is 1.0099.

Consider a differentiable function $f(x)$, and let dx denote an arbitrary change in the variable x. In this notation, dx is a single symbol representing the change in the value of x. The expression $f'(x)dx$ is called the **differential** of $y = f(x)$, and it is denoted by dy (or df), so that

$$dy = f'(x)dx$$

Now, if x changes by dx, then the corresponding change in $y = f(x)$ is

$$\Delta y = f(x + dx) - f(x)$$

In the linear approximation, suppose we replace x by $x + dx$ and a by x. The result is $f(x + dx) \approx f(x) + f'(x)dx$. Using the definitions of dy and Δy above, we get $\Delta y \approx dy = f'(x)dx$.

The notation $(d/dx)(\cdot)$ calls for the expression in parentheses to be differentiated with respect to x. For example, $(d/dx)(x^3) = 3x^2$. In the same way, we let $d(\cdot)$ denote the differential of whatever is inside the parentheses.

Example 6.7. (a) $d(Ax^a + B)$, where A, B, and a are constants

(b) $d(f(K))$, where f is a differentiable function of K

Solution:

(a) Putting $f(x) = Ax^a + B$, we get $f'(x) = Aax^{a-1}$, so $d(Ax^a + B) = Aax^{a-1}dx$.

(b) $d(f(K)) = f'(K)dK$.

Let f and g be differentiable functions of x, and let a and b be constants. Then the following rules hold:

$$d(af + bg) = adf + bdg$$

$$d(fg) = g\,df + f\,dg$$

and, if $g \neq 0$,

$$d\left(\frac{f}{g}\right) = \frac{g\,df - f\,dg}{g^2}$$

Example 6.8. Consider again the model in Example 6.3. Find the differential dY, expressed in terms of $d\bar{I}$. If employment is also a function $N = g(Y)$ of Y, find the differential dN expressed in terms of $d\bar{I}$.

Solution:

Taking the differential of (i) in Example 6.3, we obtain $dY = dC + d\bar{I}$. Doing the same for (ii) gives $dC = f'(Y)dY$. Substituting dC from the latter into the former, and solving for dY yields

$$dY = \frac{1}{1 - f'(Y)}d\bar{I}$$

which is the same formula found previously. From $N = g(Y)$ we get $dN = g'(Y)dY$, so

$$dN = \frac{g'(Y)}{1 - f'(Y)}d\bar{I}$$

Economists usually claim that employment increases as GDP increases ($g'(Y) > 0$), and that the marginal propensity to consume, $f'(Y)$, is between 0 and 1. From the above formula for dN, these claims imply that if investment increases, then employment increases as well.

Ex. 6.9. Find the linear approximations to the following functions about $x = 0$:

$$(a)f(x) = (1+x)^{-1} \qquad (b)f(x) = (1+x)^5 \qquad (c)f(x) = (1-x)^{1/4}$$

Ex. 6.10. Find the following differentials:

$(a)\ d(10x^3)$ $\qquad\qquad\qquad$ $(b)\ d(5x^3 - 5x^2 + 5x + 5)$

$(c)\ d(1/x^3)$ $\qquad\qquad\qquad\qquad$ $(d)\ d(\ln x)$

Ex. 6.11. The equation $3xe^{xy^2} - 2y = 3x^2 + y^2$ defines y as a differentiable function of x about the point $(x, y) = (1, 0)$.

(a) Find the slope of the graph at this point by implicit differentiation.

(b) What is the linear approximation to y about $x = 1$?

Ex. 6.12. If an amount K is charged to a credit card on which interest is $p\%$ per year, then unless some payments are made beforehand, after t years the balance will have grown to $K_t = K(1 + p/100)^t$ (even without any penalty charges). Using the approximation $\ln(1 + p/100) \approx p/100$, prove that $\ln K_t \approx \ln K + pt/100$. Find the percentage interest rate p at which the balance doubles after t years.

6.5 Polynomial approximations

Quadratic approximation to $f(x)$ about $x = a$:

For x close to a,

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

Example 6.9. Find the quadratic approximation to $y = y(x)$ about $x = 0$ when y is defined implicitly as a function of x near $(x, y) = (0, 1)$ by $xy^3 + 1 = y$.

Solution:

Implicit differentiation w.r.t. x yields

$$y^3 + 3xy^2 y' = y'$$

Substituting $x = 0$ and $y = 1$ gives $y' = 1$. Differentiating again w.r.t x now yields

$$3y^2 y' + (3y^2 + 6xyy')y' + 3xy^2 y'' = y''$$

Substituting $x = 0$, $y = 1$, and $y' = 1$, we obtain $y'' = 6$. Hence,

$$y(x) \approx y(0) + y'(0)x + \frac{1}{2}y''(0)x^2 = 1 + x + 3x^2$$

In general, suppose we want to approximate a function $f(x)$ over an interval centred at $x = a$ (**Taylor Approximation**):

$$f(x) \approx f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^n(a)}{n!}(x - a)^n$$

Example 6.10. Find the third-order Taylor approximation to $f(x) = \sqrt{1 + x}$ about $x = 0$.

Solution:

We write $f(x) = \sqrt{1 + x} = (1 + x)^{1/2}$. Then

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

$$f''(x) = \frac{1}{2}\left(-\frac{1}{2}\right)(1 + x)^{-3/2}$$

$$f'''(x) = \frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)(1 + x)^{-5/2}$$

Putting $x = 0$ gives $f(0) = 1$, $f'(0) = 1/2$, $f''(0) = (1/2)(-1/2) = -1/4$, and finally $f'''(0) = (1/2)(-1/2)(-3/2) = 3/8$. Hence, by Taylor Approximation for the case $n = 3$, we have

$$f(x) \approx 1 + \frac{1}{1!}\frac{1}{2}x + \frac{1}{2!}\left(-\frac{1}{4}\right)x^2 + \frac{1}{3!}\frac{3}{8}x^3 = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3$$

Example 6.11. For any natural number n, write the nth order Taylor approximation to $f(x) = e^x$ about $x = 0$.

Solution:

All derivatives of f are equal to e^x, and thus $f^{(k)}(0) = 1$ for all $k = 1, 2, \ldots, n$. Hence,

$$e^x \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

.

Ex. 6.13. Find the fifth-order Taylor approximation to $f(x) = \ln(1+x)$ about $x = 0$.

Ex. 6.14. Establish the approximation $e^{\sigma\sqrt{t/n}} \approx 1 + \sigma\sqrt{t/n} + \sigma^2 t/2n$.

6.6   Taylor's formula

Consider the Taylor approximation about $x = 0$:

$$f(x) \approx f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^n(0)}{n!}x^n$$

Now we denote the **remainder** after n terms by $R_{n+1}(x)$, then

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^n(0)}{n!}x^n + R_{n+1}(x)$$

The following theorem gives an important explicit formula for the remainder. Its proof is deferred to Chapter 7.

Lagrange's form of the remainder:

Suppose f is differentiable $n + 1$ times in an interval that includes 0 and x. Then the remainder $R^{n+1}(x)$ can be written as

$$R^{n+1}(x) = \frac{1}{(n+1)!}f^{(n+1)}(z)x^{n+1}$$

for some number z between 0 and x.

Using this theorem, we obtain

Taylor's Formula:

Suppose f is differentiable $n + 1$ times in an interval that includes 0 and x. Then

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^n(0)}{n!}x^n + \frac{1}{(n+1)!}f^{(n+1)}(z)x^{n+1}$$

for some number z between 0 and x.

Example 6.12. Use first order Taylor's Formula to approximate the function

$$f(x) = \sqrt{25 + x} = (25 + x)^{1/2}$$

Use the result to estimate $\sqrt{25.01}$, with a bound on the absolute value of the remainder.

Solution:

If we put $n = 1$ in the Taylor's Formula, we obtain that, for some z between 0 and x,

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(z)x^2$$

Differentiate $f(x)$ to obtain

$$f'(x) = \frac{1}{2}(25 + x)^{-1/2} \text{ and } f''(x) = \frac{1}{2}\left(-\frac{1}{2}\right)(25 + x)^{-3/2}$$

Then $f(0) = 5$, whereas $f'(0) = 1/2 \cdot 1/5 = 1/10$ and $f''(z) = -(1/4)(25 + z)^{-3/2}$. So there exists z between 0 and x such that

$$\sqrt{25 + x} = 5 + \frac{1}{10}x + \frac{1}{2}\left(-\frac{1}{4}\right)(25 + z)^{-3/2}x^2 = 5 + \frac{1}{10}x - \frac{1}{8}(25 + z)^{-3/2}x^2$$

In order to estimate 25.01, we set $x = 0.01$. Then z lies between 0 and 0.01, so $25 + z > 25$. Then $(25 + z)^{-3/2} < (25)^{-3/2} = 1/125$, so the absolute value of the remainder is

$$|R_2(0.01)| = \left| -\frac{1}{8}(25 + z)^{-3/2}\left(\frac{1}{100}\right)^2 \right| < \frac{1}{80000} \cdot \frac{1}{125} = 10^{-7}$$

We conclude that $\sqrt{25.01} \approx 5 + 1/10 \cdot 1/100 = 5.001$, with an error less than 10^{-7}.

Ex. 6.15. Write Taylor's formula with $n = 2$ for $f(x) = \ln(1 + x)$.

Ex. 6.16. Use the approximation $(1 + x)^m \approx 1 + mx + \frac{1}{2}m(m - 1)x^2$ and the fact that $\sqrt[3]{25} = 3(1 - 2/27)^{1/3}$ in order to find values of: (a) $\sqrt[3]{25}$; and (b) $\sqrt[5]{33}$. Then check these approximations by using a calculator.

Ex. 6.17. Let $g(x) = \sqrt[3]{1 + x}$.

(a) Find the Taylor polynomial of $g(x)$ of order 2 about the origin.

(b) For $x \geq 0$ show that $|R_3(x)| \leq 5x^3/81$.

(c) Find $\sqrt[3]{1003}$ to seven decimal places.

6.7 Elasticity

Price Elasticity of Demand

The elasticity of the demand function $D(p)$ (with respect to the price p) is

$$\frac{p}{D(p)}\frac{dD(p)}{dp}$$

Usually, we get a good approximation to the elasticity by letting $\Delta p/p = 1/100 = 1\%$ and computing $p\Delta x/(x\Delta p)$.

The General Definition of Elasticity

If f is differentiable at x and $f(x) \neq 0$, the elasticity of f w.r.t. x is

$$El_x f(x) = \frac{x}{f(x)}f'(x)$$

Example 6.13. Find the elasticity of $f(x) = Ax^b$, where A and b are constants, with $A \neq 0$.

Solution:

In this case, $f'(x) = Abx^{b-1}$. Hence, $El_x(Ax^b) = (x/Ax^b)Abx^{b-1} = b$, so

$$f(x) = Ax^b \Rightarrow El_x f(x) = b$$

This function has constant elasticity.

Example 6.14. Assume that the quantity demanded of a particular commodity is given by $D(p) = 8000p^{-1.5}$. Compute the elasticity of $D(p)$ and find the exact percentage change in quantity demanded when the price increases by 1% from $p = 4$.

   Solution:

   Using the result from the previous example, we find that $El_p D(p) = -1.5$, so that an increase in the price of 1% causes the quantity demanded to decrease by about 1.5%. In this case, we can also compute the decrease in demand exactly. When the price is 4, the quantity demanded is $D(4) = 8000 \cdot 4^{-1.5} = 1000$. If the price $p = 4$ is increased by 1%, the new price will be $4 + 4/100 = 4.04$, so that the change in demand is

$$\Delta D = D(4.04) - D(4) = 8000 \cdot 4.04^{-1.5} - 1000 = -14.81$$

The percentage change in demand from $D(4)$ is $-(14.81/1000) \cdot 100 = -1.481\%$.

Example 6.15. Let $D(P)$ denote the demand function for a product. By selling $D(P)$ units at price P, the producer earns revenue $R(P) = P \cdot D(P)$. The elasticity of $R(P)$ w.r.t. P is

$$El_P R(P) = \frac{P}{R(P)} \frac{d}{dP}[P \cdot D(P)] = \frac{1}{D(P)}[D(P) + P \cdot D'(P)] = 1 + El_P D(P)$$

Economists sometimes use the following terminology:

1. If $|El_x f(x)| > 1$, then f is elastic at x.

2. If $|El_x f(x)| = 1$, then f is unit elastic at x.

3. If $|El_x f(x)| < 1$, then f is inelastic at x.

4. If $|El_x f(x)| = 0$, then f is perfectly inelastic at x.

5. If $|El_x f(x)| = \infty$, then f is perfectly elastic at x.

If $y = f(x)$ has an inverse function x = g(y), then the Inverse Function Theorem implies that

$$El_y(g(y)) = \frac{y}{g(y)} g'(y) = \frac{f(x)}{x} \frac{1}{f'(x)} = \frac{1}{El_x f(x)}$$

So it follows:

$$El_y x = \frac{1}{El_x y}$$

Sometimes it is useful to take natural logarithm of each side of the equation. It can be proved using chain rule that:

$$El_x y = \frac{x}{y} \frac{dy}{dx} = \frac{d(\ln y)}{d(\ln x)}$$

Ex. 6.18. A study of transport economics uses the relation $T = 0.4K^{1.06}$, where K is expenditure on building roads, and T is a measure of traffic volume. Find the elasticity of T w.r.t. K. In this model, if expenditure increases by 1%, by what percentage (approximately) does traffic volume increase?

Ex. 6.19. Use the definition of elasticity to find $El_x y$ for the following functions, where a and p are constants:

(a) $y = e^{ax}$ (b) $y = \ln x$

(c) $y = x^p e^{ax}$ (d) $y = x^p \ln x$

Ex. 6.20. Prove that if f and g are positive-valued differentiable functions of x and A is a constant, then the following rules hold. Here we write, for instance, $El_x f$ instead of $El_x f(x)$.

(a) $El_x A = 0$ (b) $El_x(fg) = El_x f + El_x g$

(c) $El_x(f/g) = El_x f - El_x g$ (d) $El_x(f + g) = \dfrac{f El_x f + g El_x g}{f + g}$

(e) $El_x(f - g) = \dfrac{f El_x f - g El_x g}{f - g}$ (f) $El_x(f(g(x))) = El_u f(u) El_x g$ $(u = g(x))$

Ex. 6.21. Use the rules of Exercise 6.20 to evaluate the following:

(a) $El_x(-10x^{-5}) = 0$ (b) $El_x(x + x^2)$

(c) $El_x(x^3 + 1)^{10}$ (d) $El_x(El_x 5x^2)$

(e) $El_x(1 + x^2)$ (f) $El_x(\dfrac{x-1}{x^5 + 1})$

6.8   Continuity

Roughly speaking, a function $y = f(x)$ is continuous if small changes in the independent variable x lead to small changes in the function value y. Geometrically, a function is continuous on an interval if its graph is connected.
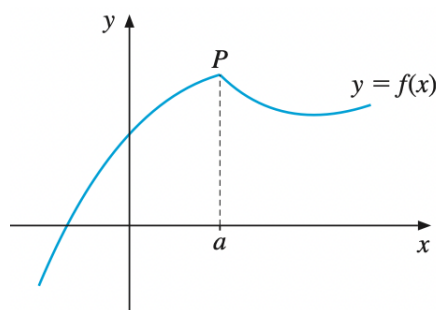
Fig. 6.4. A continuous function
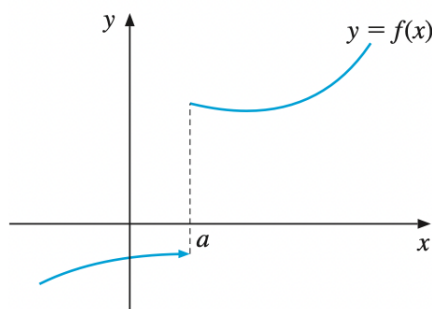
Fig. 6.5. A discontinuous function

Formally, the function f is **continuous** at $x = a$ iff

$$\lim_{x \to a} f(x) = f(a)$$

Hence, we see that in order for f to be continuous at $x = a$, the following three conditions must all be fulfilled:

(i) The function f must be defined at $x = a$.

(ii) The limit of $f(x)$ as x tends to a must exist.

(iii) This limit must be equal to $f(a)$.

Unless all three of these conditions are satisfied, we say that f is **discontinuous** at a.

Properties of continuous functions

If f and g are continuous at a, then:

(a) $f + g$ and $f - g$ are continuous at a.

(b) fg and, in case $g(a) \neq 0$, the quotient f/g are continuous at a.

(c) $[f(x)]^r$ is continuous at a, if $[f(a)]^r$ is defined, where r is a real number.

(d) If f has an inverse on the interval I, then its inverse f^{-1} is continuous on $f(I)$.

Example 6.16. Determine for which values of x the functions f and g are continuous:

$$(a) f(x) = \frac{x^4 + 3x^2 - 1}{(x - 1)(x + 2)} \qquad (b) g(x) = (x^2 + 2)\left(x^3 + \frac{1}{x}\right)^4 + \frac{1}{\sqrt{x + 1}}$$

Solution:

(a) This is a rational function that is continuous at all x, except where the denominator $(x - 1)(x + 2)$ vanishes. Hence, f is continuous at all x different from 1 and -2.

(b) This function is defined when $x \neq 0$ and $x + 1 > 0$. By properties (a), (b), and (c), it follows that g is continuous in the domain $(-1, 0) \cup (0, \infty)$.

Ex. 6.22. Let f and g be defined for all x by

$$f(x) = \begin{cases} x^2 - 1, \text{ for } x \geq 0 \\ -x^2, \text{ for } x > 0 \end{cases}$$

Draw a graph of each function. Is f continuous at $x = 0$? Is g continuous at $x = 2$?

Ex. 6.23. For what value of a is the following function continuous for all x?

$$f(x) = \begin{cases} ax - 1 \text{ for } x \leq 1 \\ 3x^2 + 1 \text{ for } x > 1 \end{cases}$$

6.9 Limits

Example 6.17. Examine $\lim_{x \to -2} \frac{1}{(x+2)^2}$ using a calculator.

Solution:

At $x = -2$, the fraction is not defined. As x gets closer and closer to -2, we see that the value of the fraction becomes larger and larger. For example at $x = -1.9999$ and $x = -2.0001$, the value becomes 100 million.

Figure 6.6 shows the graph of $f(x) = 1/(x+2)^2$. The line $x = -2$ is called an **asymptote** for the graph of f.
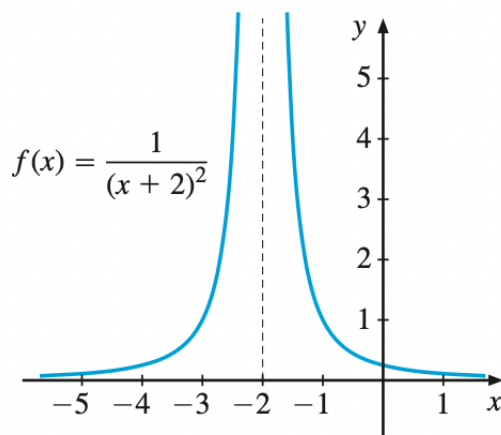
Fig. 6.6. $f(x) \to \infty$ as $x \to -2$

6.9.1 One sided limits

The function whose graph is shown in Fig. 6.7 also fails to have a limit as x tends to a. However, it seems from the figure that if x tends to a with values less than a, then $f(x)$ tends to the number B. We say, therefore, that the limit of $f(x)$ as x tends to a from below is B, and we write

$$\lim_{x \to a^-} f(x) = B \text{ or } f(x) \to B \text{ as } x \to a^-$$

Analogously, we say that the limit of $f(x)$ as x tends to a from above is A, and we write

$$\lim_{x \to a^+} f(x) = A \text{ or } f(x) \to A \text{ as } x \to a^+$$

We call these **one-sided limits**, the first from below and the second from above. They can also be called the left limit and right limit, respectively.
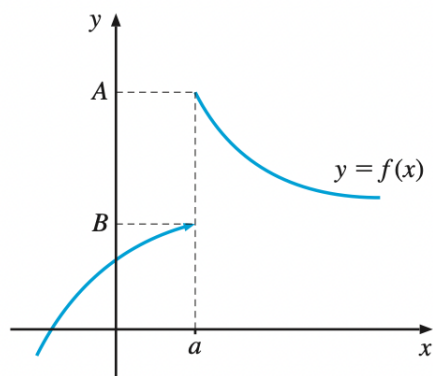
Fig. 6.7. Limit does not exist

A necessary and sufficient condition for the limit to exist is that the two one-sided limits of f at a exist and are equal:

$$\lim_{x \to a} f(x) = A \iff \left[\lim_{x \to a^-} f(x) = A \text{ and } \lim_{x \to a^+ f(x) = A} \right]$$

Example 6.18. Figure 6.8 shows the graph of a function f defined on $[0, 9]$. Use the figure to check that the following limits seem correct:

$$\lim_{x \to 2} f(x) = 3 \qquad \lim_{x \to 4^-} f(x) = 1/2 \qquad \lim_{x \to 4^+} f(x) = 3$$

$$\lim_{x \to 6^-} f(x) = \lim_{x \to 6^-} f(x) = -\infty$$
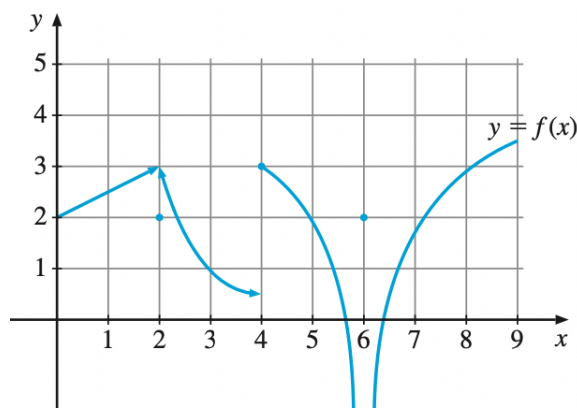
Fig. 6.8. Example

6.9.2 One sided continuity

The introduction of one-sided limits allows us to define one-sided continuity. Suppose f is defined on the half-open interval $(c, a]$. If $f(x)$ tends to $f(a)$ as x tends to a^-, we say that f is left continuous at a. Similarly, if f is defined on $[a, d)$, we say that f is right continuous at a if $f(x)$ tends to $f(a)$ as x tends to a^+. Therefore, a function f is continuous at a if and only if f is both left and right continuous at a.

6.9.3 Limits at infinity

Let f be defined for arbitrarily large positive numbers x. We say that $f(x)$ has the limit A as x tends to infinity if $f(x)$ can be made arbitrarily close to A by making x sufficiently large. We write

$$\lim_{x \to \infty} f(x) = A \ or \ f(x) \to A \ as \ x \to \infty$$

Similarly,

$$\lim_{x \to -\infty} f(x) = B \ or \ f(x) \to B \ as \ x \to -\infty$$

The horizontal line $y = A$ is a horizontal asymptote for the graph of f as x tends to ∞, whereas $y = B$ is a horizontal asymptote for the graph as x tends to $-\infty$.
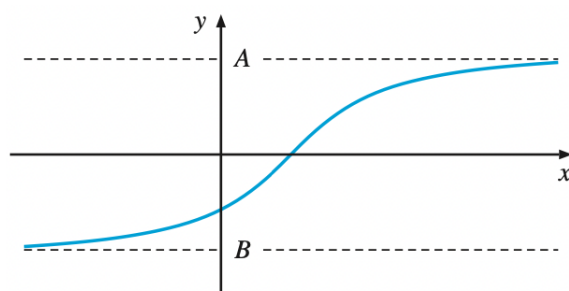
Fig. 6.9. Horizontal asymptotes

Example 6.19. Examine the following functions as $x \to \infty$ and as $x \to -\infty$:

Example 6.20. Let $f(x) = 1/x^2$ and $g(x) = 1/x^4$. As $x \to 0$, so $f(x) \to \infty$ and $g(x) \to \infty$. Examine the limits as $x \to 0$ of the following expressions:

$$f(x) - g(x), \ g(x) - f(x), \ \frac{f(x)}{g(x)}, \ and \ \frac{g(x)}{f(x)}$$

Solution:

$f(x) - g(x) = (x^2 - 1)/x^4$. As $x \to 0$, the numerator tends to -1 and the denominator to 0, so the fraction tends to $-\infty$. For the other three limits we have:

$$g(x) - f(x) = \frac{1 - x^2}{x^4} \to \infty, \ \frac{f(x)}{g(x)} = x^2 \to 0, \ and \ \frac{g(x)}{f(x)} = \frac{1}{x^2} \to \infty$$

6.9.4   Continuity and differentiability

Consider the function f graphed in Fig. 6.10. At point $(a, f(a))$ the graph does not have a unique tangent. Thus f has no derivative at $x = a$, but f is continuous at $x = a$. So a function can be continuous at a point without being differentiable at that point. On the other hand, differentiability implies continuity:
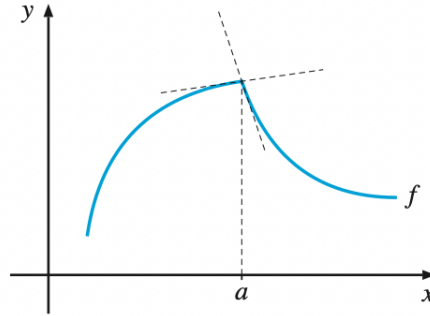
Fig. 6.10. Continuous but not differentiable function

If function f is differentiable at $x = a$, then it is continuous at $x = a$.

6.9.5   Rigorous definition of limits

We say that $f(x)$ has limit A as x tends to a if, for each number $\epsilon > 0$, there exists a number $\delta > 0$ such that $|f(x) - A| < \epsilon$ for every x with $0 < |x - a| < \delta$. When this holds, we write

$$\lim_{x \to a} f(x) = A$$

and also say that $f(x)$ tends to A as x tends to a.
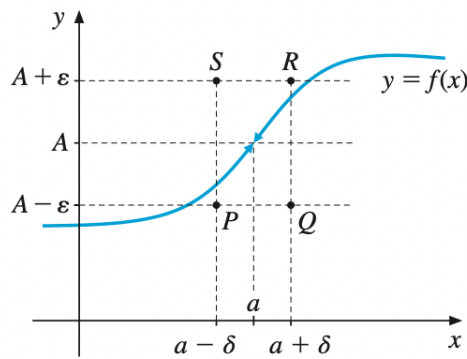
Fig. 6.11. Definition of limit

This definition is the basis of all mathematically rigorous work on limits. It is illustrated in Fig. 6.11. The definition implies that the graph of f must remain within the rectangular box $PQRS$, for all $x \neq a$ in $(a - \delta, a + \delta)$.

Ex. 6.24. Evaluate the following limits:

$$(a)\ \lim_{x\to 0^+}(x^2+3x-4) \qquad\qquad (b)\ \lim_{x\to 0^-}\frac{x+|x|}{x}$$

$$(c)\ \lim_{x\to 0^+}\frac{x+|x|}{x} \qquad\qquad (d)\ \lim_{x\to 0^+}\frac{-1}{\sqrt{x}}$$

$$(e)\ \lim_{x\to 3^+}\frac{x}{x-3} \qquad\qquad (f)\ \lim_{x\to 3^-}\frac{x}{x-3}$$

Ex. 6.25. Evaluate

$$(a)\ \lim_{x\to\infty}\frac{x-3}{x^2+1} \qquad (b)\ \lim_{x\to-\infty}\sqrt{\frac{2+3x}{x-1}} \qquad (c)\ \lim_{x\to\infty}\frac{(ax-b)^2}{(a-x)(b-x)}$$

Ex. 6.26. Find asymptotes for the graph of each of the following functions of x:

$$(a)\frac{x^2}{x+1} \qquad\qquad (b)\frac{2x^3-3x^2+3x-6}{x^2+1}$$

$$(c)\frac{3x^2+2x}{x-1} \qquad\qquad (d)\frac{5x^4-3x^2+1}{x^3-1}$$

Ex. 6.27. Consider the function f defined by the formula

$$f(x)=\frac{3x}{-x^2+4x-1}$$

Compute $f'(x)$ and use a sign diagram to determine where the function increases. (The function is not defined for $x=2\pm\sqrt{3}$.)

6.10   The Intermediate Value Theorem and Newton's Method

6.10.1   Intermediate Value Theorem

Let f be a function which is continuous in the interval $[a,b]$.

  (i) If $f(a)$ and $f(b)$ have different signs, then there is at least one c in (a,b) such that $f(c)=0$.

 (ii) If $f(a)\neq f(b)$, then for every intermediate value y in the open interval between $f(a)$ and $f(b)$ there is at least one c in (a,b) such that $f(c)=y$.

   This theorem is important in assuring the existence of solutions to some equations that cannot be solved explicitly.

Example 6.21. Prove that the equation

$$2x-5e^{-x}(1+x^2)=0$$

has a unique solution, which lies in the interval $(0,2)$.

Solution:

Define $g(x) = 2x - 5e^{-x}(1 + x^2)$. Then $g(0) = -5$ and $g(2) = 4 - 25/e^2$. In fact $g(2) > 0$ because $e > 5/2$. According to the intermediate value theorem, therefore, the continuous function g must have at least one zero in $(0, 2)$. Moreover, note that $g'(x) = 2 + 5e^{-x}(1 + x^2) - 10xe^{-x} = 2 + 5e^{-x}(1 - 2x + x^2) = 2 + 5e^{-x}(x - 1)^2$.But then $g'(x) > 0$ for all x, so g is strictly increasing. It follows that g can have only one zero.

6.10.2   Newton's method

The intermediate value theorem can often be used to show that an equation $f(x) = 0$ has a solution in a given interval, but it says nothing more about where to find this zero. In this subsection we shall explain an effective method for finding a good approximate solution. The method was first suggested by Isaac Newton. It has an easy geometric explanation.

Consider the graph of the function $y = f(x)$ shown in Fig. 6.12. It has a zero at $x = a$, but this zero is not known. To find it, start with x_0 as an initial estimate of a. It is usually better to start with x_0 not too far from a, if possible.
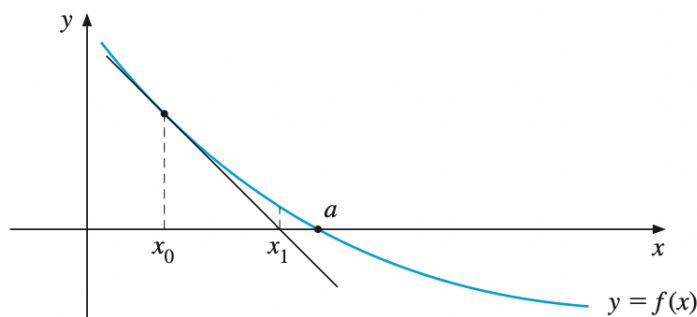
Fig. 6.12. Newton's method

Usually x_1 is a significantly better estimate of a than x_0 was. After having found x_1 , repeat the procedure by constructing the tangent line to the curve at the point $(x_1, f(x_1))$. Denote by x_2 the point where this new tangent line crosses the x-axis. Repeating this procedure, we obtain a sequence of points which usually converges very quickly to a.

The slope of the tangent at x_0 is $f'(x_0)$. According to the point–slope formula, the equation for the tangent line through the point $(x_0, f(x_0))$ with slope $f'(x_0)$ is given by

$$y - f(x_0) = f'(x_0)(x - x_0)$$

At the point where this tangent line crosses the x-axis, we have $y = 0$ and $x = x_1$. Hence

$$-f(x_0) = f'(x_0)(x_1 - x_0)$$

Solving this equation for x_1, we get the first new approximation

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Similarly, given x_1, the formula for the second approximation is

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

In general, one has the following formula for the $(n+1)$-th approximation x_{n+1}, expressed in terms of the n-th approximation x_n:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \ f'(x_n) \neq 0 \text{ and } n = 0, 1, \dots$$

Example 6.22. Consider the function

$$f(x) = x^6 + 3x^2 - 2x - 1$$

Use Newton's method once to find an approximate value for the zero of f in the interval $[0, 1]$.

Solution:

Choose $x_0 = 1$. Then $f(x_0) = f(1) = 1$. Because $f'(x) = 6x^5 + 6x - 2$, we have $f'(1) = 10$. Hence, using Newton's method for $n = 0$ yields

$$x_1 = 1 - \frac{f(1)}{f'(1)} = 1 - \frac{1}{10} = 0.9$$

In most cases, Newton's method is very efficient, but it can happen that the sequence x_n does not converge. Figure 6.13 shows an example where x_1 is a much worse approximation to a than x_0 was. Usually, Newton's method fails only if the absolute value of $f'(x_n)$ becomes too small, for some n.
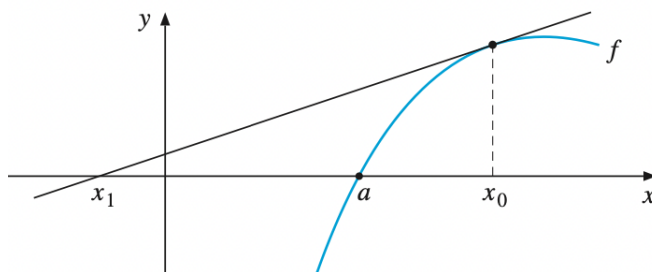
Fig. 6.13. Newton's method

Ex. 6.28. Show that each of the following equations has at least one root in the given interval.

$(a) x^7 - 5x^5 + x^3 - 1 = 0$, in $(-1, 1)$ $\qquad$ $(b) x^3 + 3x - 8 = 0$, in $(1, 3)$

$(c) \sqrt{x^2 + 1} = 3x$, in $(0, 1)$ $\qquad\qquad$ $(d) e^{x-1} = 2x$, in $(0, 1)$

Ex. 6.29. Find a better approximation to $\sqrt[3]{17} \approx 2.5$ by using Newton's method once.

Ex. 6.30. The equation $x^4 + 3x^3 - 3x^2 - 8x + 3 = 0$ has an integer root. Find it. The three additional roots are close to -1.9, 0.4, and 1.5. Find better approximations by using Newton's method once for each root that is not an integer.

6.11   L'Hôpital's Rule

We often need to examine the limit as x tends to a of a quotient in which both numerator and denominator tend to 0:

$$\lim_{x \to a} \frac{f(x)}{g(x)} = "0/0"$$

We call such a limit an **indeterminate form** of type $0/0$.

We start with the simple case of an indeterminate form $f(x)/g(x)$, where f and g are differentiable and $f(a) = g(a) = 0$. When $x \neq a$ and $g(x) \neq g(a)$, then some routine algebra allows us to write

$$\frac{f(x)}{g(x)} = \frac{[f(x) - f(a)]/(x - a)}{[g(x) - g(a)]/(x - a)}$$

The right-hand side is the ratio of two Newton quotients. Taking the limit as $x \to a$, we see that provided $g'(a) \neq 0$, this ratio tends to $f'(a)/g'(a)$. This gives the following result:

If $f(a) = g(a) = 0$ and $g'(a) \neq 0$, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}$$

Example 6.23. Assuming that $x > 0$ and $y > 0$, compute

$$\lim_{\lambda \to 0} \frac{x^\lambda - y^\lambda}{\lambda}$$

Solution:

Define $f(\lambda) = x^\lambda - y^\lambda$ and $g(\lambda) = \lambda$. Then $f(0) = g(0) = 0$. Using the rule $(d/dx)a^x = a^x \ln a$, we obtain $f'(\lambda) = x^\lambda \ln x - y^\lambda \ln y$, so that $f'(0) = \ln x - \ln y$. Moreover, $g'(\lambda) = 1$, so $g'(0) = 1$. Using l'Hôpital's rule,

$$\lim_{\lambda \to 0} \frac{x^\lambda - y^\lambda}{\lambda} = \frac{\ln x - \ln y}{1} = \ln \frac{x}{y}$$

In particular, if $y = 1$, then

$$\lim_{\lambda \to 0} \frac{x^\lambda - 1}{\lambda} = \ln x$$

Example 6.24. Consider the "constant elasticity of substitution", or CES, function

$$F(K, L) = A(aK^{-\rho} + (1 - a)L^{-\rho})^{-1/\rho}$$

where $A > 0$, $K > 0$, $L > 0$, $a \in (0, 1)$, and $\rho \neq 0$. Keeping A, K, L, and a fixed, apply l'Hôpital's rule to $\ln[F(K, L)/A]$ as $\rho \to 0$ in order to show that $F(K, L)$ converges to the Cobb–Douglas function $AK^a L^{1-a}$.

Solution:

We take natural log of both sides:

$$\ln F = \ln(aK^{-\rho} + (1-a)L^{-\rho})^{-1/\rho} = -\ln(aK^{-\rho} + (1-a)L^{-\rho})/\rho \to 0/0 \ as \ \rho \to 0$$

Because $(d/d\rho)K^{-\rho} = -K^{-\rho}\ln K$ and $(d/d\rho)L^{-\rho} = -L^{-\rho}\ln L$, applying l'Hôpital's rule gives

$$\lim_{\rho}\ln F = \lim_{\rho\to 0}\left[\frac{aK^{-\rho}\ln K + (1-a)L^{-\rho}\ln L}{aK^{-\rho} + (1-a)L^{-\rho}}\right] \div 1$$
$$= a\ln K + (1-a)\ln L$$
$$= \ln K^a L^{1-a}$$

Hence $e^{\ln F} \to K^a L^{1-a}$. Therefore it follows that $F(K, L) \to AK^a L^{1-a}$ as $\rho \to 0$.

In general, given $a > 1$, for any fixed positive number p,

$$\lim_{x\to\infty}\frac{x^p}{a^x} = 0$$

Ex. 6.31. Use l'Hôpital's rule to find the following limits:

(a) $\lim\limits_{x\to 1}\dfrac{x-1}{x^2-1}$

(b) $\lim\limits_{x\to -2}\dfrac{x^3 + 3x^2 - 4}{x^3 + 5x^2 + 8x + 4}$

(c) $\lim\limits_{x\to 2}\dfrac{x^4 - 4x^3 + 6x^2 - 8x + 8}{x^3 - 3x^2 + 4}$

(d) $\lim\limits_{x\to 1}\dfrac{\ln x - x + 1}{(x-1)^2}$

(e) $\lim\limits_{x\to 1}\dfrac{1}{x-1}\ln\left(\dfrac{7x+1}{4x+4}\right)$

(f) $\lim\limits_{x\to 1}\dfrac{x^x - x}{1 - x + \ln x}$

Ex. 6.32. The family of CES utility functions is given by

$$u(c) = \begin{cases} \frac{c^{1-\rho}-1}{1-\rho}, & \text{if } \rho \neq 1 \\ \ln c, & \text{if } \rho = 1 \end{cases}$$

for all $c > 0$. Use l'Hôpital's rule to show that $\lim_{\rho\to 1}\frac{c^{1-\rho}-1}{1-\rho} = \ln c$.

In this sense, the family is "continuous in ρ".

6.12 Review exercises

Ex. 6.33. The graph of the equation $x^3 + y^3 = 3xy$ passes through the point (3/2, 3/2). Find the slope of the tangent line to the curve at this point. This equation has a nice graph, called Descartes's folium, which appears in Fig. 6.14.
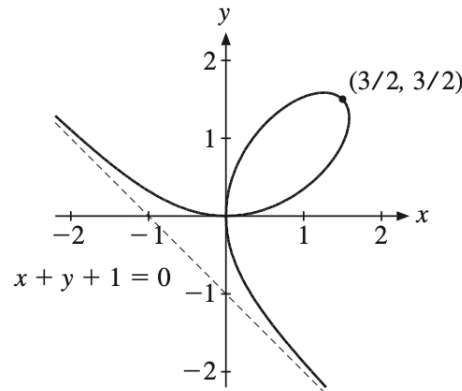
Fig. 6.14. Descartes's folium

Ex. 6.34. If $K^{1/3}L^{1/3} = 24$, compute dL/dK by implicit differentiation.

Ex. 6.35. Consider the following macroeconomic model:

$$(i)Y = C + I \qquad (ii)C = f(Y - T) \qquad (iii)T = \alpha + \beta Y$$

where Y is GDP, C is consumption, T denotes taxes, and α and β are constants. Assume that $f' \in (0,1)$ and $\beta \in (0,1)$.

  (a) From equations (i)–(iii) derive the equation $Y = f((1 - \beta)Y - \alpha) + I$.

  (b) Differentiate the equation in (a) implicitly w.r.t. I and find an expression for dy/dI.

  (c) Examine the sign of dy/dI.

Ex. 6.36.   (a) Find y' when y is given implicitly by the equation

$$x^2 - xy + 2y^2 = 7$$

  (b) Find the points where the graph has horizontal tangent and the points where it has vertical tangent. Do your results accord with Fig. 6.15, which shows the graph of the equation?
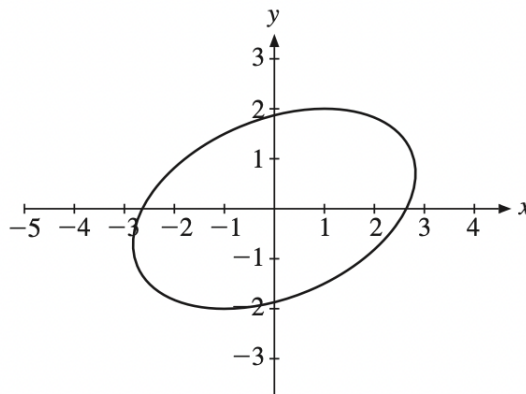
Fig. 6.15. Exercise 6.36

Ex. 6.37. The graph of the equation $x^2 y - 3y^3 = 2x$ passes through the point $(x, y) = (-1, 1)$.

 (a) Find the slope of the graph at this point.

 (b) Find the points at which the curve has vertical tangent. Show that no point on the curve has a horizontal tangent. Do your results accord with Fig. 6.16, which shows the graph of the equation?
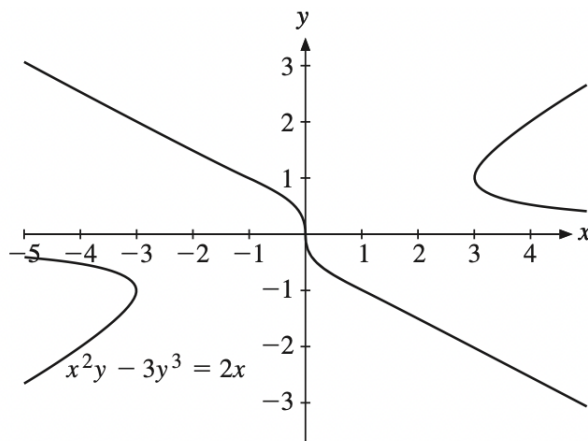
Fig. 6.16. Exercise 6.37

Ex. 6.38. Let $f(x)$ be defined for all $x > 0$ by $f(x) = (\ln x)^3 - 2(\ln x)^2 + \ln x$.

 (a) Compute $f(e^2)$ and find the zeros of $f(x)$.

 (b) Prove that $f(x)$ defined on $[e, \infty)$ has an inverse function h, then determine $h'(2)$.

Ex. 6.39. The demands for margarine (marg) and for meals away from home (mah) in the UK during the period 1920–1937, as functions of personal income r, were estimated to be $D_{marg} = Ar^{-0.165}$ and $D_{mah} = Br^{2.39}$, respectively, for suitable constants A and B. Find and interpret the (Engel) elasticities of D_{marg} and D_{mah} w.r.t. r.

Ex. 6.40. Prove that $f(x) = e^{\sqrt{x}} - 3$ has a unique zero in the interval $(1, 4)$. Find an approximate value for this zero by using Newton's method once, with $x_0 = 1$.

Ex. 6.41. Evaluate the limits:

 (a) $\displaystyle\lim_{x \to 3^-} (x^2 - 3x + 2)$

 (b) $\displaystyle\lim_{x \to -2^+} \frac{x^2 - 3x + 14}{x + 2}$

 (c) $\displaystyle\lim_{x \to -1} \frac{3 - \sqrt{x + 17}}{x + 1}$

 (d) $\displaystyle\lim_{x \to 0} \frac{(2 - x)e^x - x - 2}{x^3}$

 (e) $\displaystyle\lim_{x \to 3} \left(\frac{1}{x - 3} - \frac{5}{x^2 - x - 6} \right)$

 (f) $\displaystyle\lim_{x \to 4} \frac{x - 4}{2x^2 - 32}$

 (g) $\displaystyle\lim_{x \to 2} \frac{x^2 - 3x + 2}{x - 2}$

 (h) $\displaystyle\lim_{x \to -1} \frac{4 - \sqrt{x + 17}}{2x + 2}$

 (i) $\displaystyle\lim_{x \to \infty} \frac{(\ln x)^2}{3x^2}$

Chapter 7

Single-Variable Optimization

7.1 Extreme points

If $f(x)$ has domain D, then

$$c \in D \text{ is a \textbf{maximum point} for } f \iff f(x) \leq f(c) \text{ for all } x \in D$$

$$d \in D \text{ is a \textbf{minimum point} for } f \iff f(x) \geq f(d) \text{ for all } x \in D$$

7.1.1 First-order condition

Suppose that a function f is differentiable in an interval I and that c is an interior point of I. For $x = c$ to be a maximum or minimum point for f in I, a necessary condition is that it is a critical point for f —i.e., $x = c$ is a solution of

$$f'(x) = 0$$

7.1.2 Second-order condition

Let f be a twice differentiable function in an interval I, and let c be an interior point of I.

(a) If $f'(c) = 0$ and $f''(c) < 0$, then $x = c$ is a strict local maximum point.

(b) If $f'(c) = 0$ and $f''(c) > 0$, then $x = c$ is a strict local minimum point.

(c) If $f'(c) = 0$ and $f''(c) = 0$, then the character of $x = c$ remains undetermined.

7.2 Test for extreme points

Suppose that f is a function defined in an interval I and that c is a critical point for f in the interior of I.

(a) If f is concave, then c is a maximum point for f in I.

(b) If f is convex, then c is a minimum point for f in I.

Example 7.1. Consider the function f defined for all x by $f(x) = e^{x-1} - x$. Show that f is convex and find its minimum point. Sketch the graph.

    Solution:

    $f'(x) = e^{x-1} - 1$ and $f''(x) = e^{x-1} > 0$, so f is convex. Note that $f'(x) = e^{x-1} - 1 = 0$ for $x = 1$. It follows that $x = 1$ minimises f. See Fig. 7.1 for the graph of f, which confirms the result.
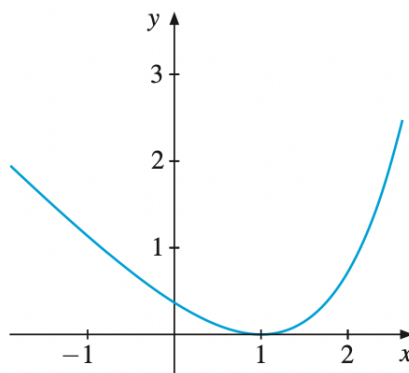
Fig. 7.1. Example 7.1

Ex. 7.1. Find the extreme points of the function h, defined for all x by the formula $h(x) = 8x/(3x^2 + 4)$.

Ex. 7.2. Find the values of x that maximises/minimises the functions given by the following formulas and confirm they are maxima/minima.

$$(a) y = e^x + e^{-2x} \qquad (b) y = \ln x - 5x$$

7.3   Mean value theorem

If f is continuous in the closed and bounded interval [a, b], and differentiable in the open interval (a, b), then there exists at least one point x^* in (a, b) such that

$$f'(x^*) = \frac{f(b) - f(a)}{b - a}$$

Example 7.2. Test the mean value theorem on the function $f(x) = x^3 - x$, defined over [0,2].

    Solution:

    We find that $[f(2) - f(0)]/(2 - 0) = 3$ and $f'(x) = 3x^2 - 1$. The equation $f'(x) = 3$ has two solutions, $x = \pm 2\sqrt{3}/3$. The positive root $x^* = 2\sqrt{3}/3 \in (0, 2)$, and

$$f'(x^*) = \frac{f(2) - f(0)}{2 - 0}$$

This confirms the mean value theorem in this case.

Ex. 7.3. For the following functions determine all numbers x^* in the specified intervals such that $f'(x^*) = \frac{f(b)-f(a)}{b-a}$:

$$(a) f(x) = x^2, \text{ in } [1,2] \qquad (b) f(x) = \sqrt{1-x^2}, \text{ in } [0,1]$$

$$(c) f(x) = 2/x, \text{ in } [2,6] \qquad (d) f(x) = \sqrt{9+x^2}, \text{ in } [0,4]$$

7.4 Inflection points

If the function f is twice differentiable, the point c is called an **inflection point** for f if there exists an interval (a, b) about c such that:

(a) $f''(x) \geq 0$ in (a,c) and $f''(x) \leq 0$ in (c,b); or

(b) $f''(x) \leq 0$ in (a,c) and $f''(x) \geq 0$ in (c,b).
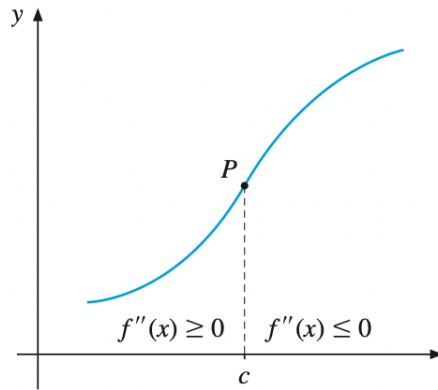
Fig. 7.2. Example of inflection point

Let f be a function with a continuous second derivative in an interval I, and let c be an interior point of I.

(a) If c is an inflection point for f, then $f''(c) = 0$.

(b) If $f''(c) = 0$ and f'' changes sign at c, then c is an inflection point for f.

7.4.1 Alternative definition of concave and convex functions

A function f is called concave if the line segment joining any two points on the graph is below the graph, or on it. It is called convex if any such line segment lies above, or on the graph.
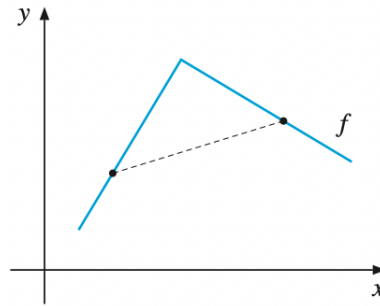
Fig. 7.3. f is concave
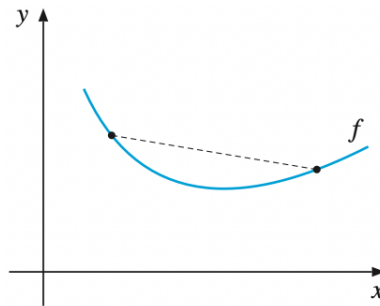
Fig. 7.4. f is convex

7.5   Review exercises

Ex. 7.4. A firm's production function is $Q(L) = 12L^2 - \frac{1}{20}L^3$, where L denotes the number of workers, with $L \in [0, 200]$.

  (a) What size of the work force maximises output $Q(L)$?

  (b) What size of the work force maximises output per worker, $Q(L)/L$? Letting L^* denote such size, note that $Q'(L^*) = Q(L^*)/L^*$. Is this a coincidence?

Ex. 7.5. By producing and selling Q units of some commodity a firm earns a total revenue $R(Q) = -0.0016Q^2 + 44Q$ and incurs a cost of $C(Q) = 0.0004Q^2 + 8Q + 64000$.

  (a) What production level maximises profits?

  (b) The elasticity $El_Q C(Q) \approx 0.12$ for $Q = 1000$. Interpret this result.

Ex. 7.6. Let $g(x) = x - 2\ln(x + 1)$.

  (a) Where is g defined?

  (b) Find $g'(x)$ and $g''(x)$.

  (c) Find possible extreme points and inflection points, and sketch the graph.

Ex. 7.7. Let $f(x) = \ln(x+1) - x + \frac{1}{2}x^2 - \frac{1}{6}x^3$.

(a) Find the domain of the function and prove that, for x in the domain:

$$f'(x) = \frac{x^2 - x^3}{2(x+1)}$$

(b) Find possible extreme points and inflection points.

(c) Check $f(x)$ as $x \to (-1)^+$, and sketch the graph on the interval (-1,2].

Ex. 7.8. Consider the function defined, for all x, by $h(x) = e^x/(2 + e^{2x})$.

(a) Where is h increasing/decreasing? Find possible maximum and minimum points for h.

(b) If one restricts the domain of h to $(-\infty, 0]$, it has an inverse. Why? Find an expression for the inverse function.

Ex. 7.9. Let $f(x) = (e^{2x} + 4e^{-x})^2$.

(a) Find $f'(x)$ and $f''(x)$.

(b) Determine where f is increasing/decreasing, and show that f is convex.

(c) Find possible global extreme points for f .

Ex. 7.10. Letting $a > 0$, consider the function

$$f(x) = \frac{x}{\sqrt[3]{x^2 - a}}$$

(a) Find the domain D_f of f and the intervals where $f(x)$ is positive. Show that the graph of f is symmetric about the origin.

(b) Where is f increasing and where is it decreasing? Find possible local extreme points.

(c) Find possible inflection points for f .

Chapter 8

Integration

8.1 Indefinite integrals

Suppose we do not know the function F, but we do know that its derivative is $F'(x)$. What is F?

We pass from F to F' by taking the derivative, so the reverse process of passing from F' to F could appropriately be called taking the **antiderivative** or **integration**. Following usual mathematical practice, we call F an **indefinite integral** of F' over the interval I, and denote it by $\int f(x)dx$. Two functions having the same derivative throughout an interval must differ by a constant, so:

If $F'(x) = f(x)$, then

$$\int f(x)dx = F(x) + C$$

where C is an arbitrary constant.

The symbol $\int$ is the **integral sign**, and the function $f(x)$ is the **integrand**. Then we write dx to indicate that x is the **variable of integration**. Finally, C is a **constant of integration**.

Integration and differentiation cancel each other out:

$$\frac{d}{dx}\int f(x)dx = f(x)$$

8.1.1 Some important integrals

If $a \neq -1$, then

$$\int x^a dx = \frac{1}{a+1}x^{a+1} + C$$

$$\int \frac{1}{x}dx = \ln|x| + C$$

If $a \neq 0$, then

$$\int e^{ax}dx = \frac{1}{a}e^{ax} + C$$

When $a > 0$ and $a \neq 1$,

$$\int a^x dx = \frac{1}{\ln a}a^x + C$$

8.1.2 General rules

$$\int af(x)dx = a\int f(x)dx, \text{ where } a \neq x \text{ is a constant}$$

$$\int [f(x) + g(x)]dx = \int f(x)dx + \int g(x)dx$$

Example 8.1. Evaluate:

(a) $\int(3x^4 + 5x^2 + 2)dx$

(b) $\int(\frac{3}{x} - 8e^{-4x})dx$

Solution:

(a)

$$\int(3x^4 + 5x^2 + 2)dx = \frac{3}{5}x^5 + \frac{5}{3}x^3 + 2x + C$$

(b)

$$\int \left(\frac{3}{x} - 8e^{-4x} \right) dx = 3\ln|x| + 2e^{-4x} + C$$

Ex. 8.1. In the manufacture of a product, the marginal cost of producing x units is $C'(x)$ and fixed costs are $C(0)$. Find the total cost function $C(x)$ when:

$$C'(x) = 3x + 4 \text{ } and \text{ } C(0) = 40$$

Ex. 8.2. Find the following integrals:

(a)

$$\int \frac{(y-2)^2}{\sqrt{y}}dy$$

(b)

$$\int \frac{x^3}{x+1}$$

(c)

$$\int x(1+x^2)^{15}$$

Ex. 8.3. Show that

(a)

$$\int x^2 \ln x dx = \frac{1}{3}x^3 \ln x - \frac{1}{9}x^3 + C$$

(b)

$$\int \sqrt{x^2+1}dx = \frac{1}{2}x\sqrt{x^2+1} + \frac{1}{2}\ln\left(x + \sqrt{x^2+1}\right) + C$$

Ex. 8.4. Suppose that $f''(x) = x^{-2} + x^3 + 2$ for $x > 0$, and $f(1) = 0$, $f'(1) = 1/4$. Find $f(x)$.

8.2 Area and definite integrals

The problem to be considered and solved in this section is illustrated in Fig. 8.1:
How do we compute the area A under the graph of a continuous and nonnegative
function f over the interval $[a, b]$?
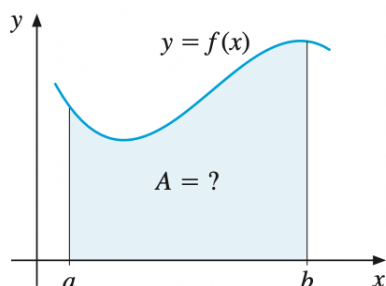
Fig. 8.1. Area A under the graph

Let t be an arbitrary point in $[a, b]$, and let $A(t)$ denote the area under the
curve $y = f(x)$ over the interval $[a, t]$, as shown in Fig. 8.2. It cannot be larger
than the area of the rectangle with base Δt and height $f(t + \Delta t)$, nor smaller
than the area of the rectangle with base Δt and height $f(t)$. Hence, for all
$\Delta t > 0$, one has
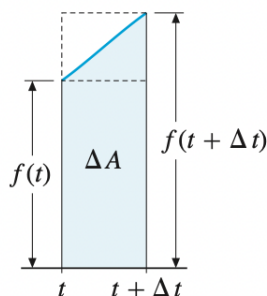
Fig. 8.2. Approximating ΔA

$$f(t)\Delta t \leq A(t + \Delta t) - A(t) \leq f(t + \Delta t)\Delta t$$

Because $\Delta t > 0$, this implies

$$f(t) \leq \frac{A(t + \Delta t) - A(t)}{\Delta t} \leq f(t + \Delta t)$$

As $\Delta t \to 0$, we have

$$A'(t) = f(t), \text{ for all } t \text{ in (a,b)}$$

Therefore the area under the curve is the **definite integral**:

$$A(t) = \int_a^b f(x)dx = F(b) - F(a), \text{ where } F(x) \text{ is the integral of } f(x)$$
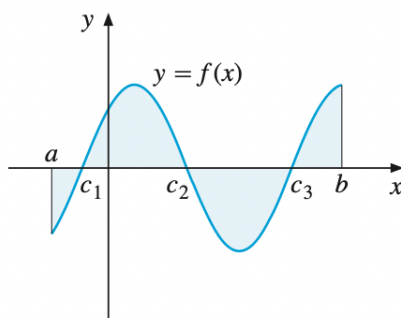
Fig. 8.3. A function that takes positive and negative values

Example 8.2. Evaluate the definite integrals:

(a)

$$\int_2^5 e^{2x}dx$$

(b)

$$\int_{-2}^2 (x - x^3 - x^5)dx$$

Solution:

(a)

$$\int_2^5 e^{2x}dx = \left.\frac{1}{2}e^{2x}\right|_2^5 = \frac{1}{2}e^{10} - \frac{1}{2}e^4$$

(b)

$$\int_{-2}^2 (x-x^3-x^5)dx = \left.(\frac{1}{2}x^2 - \frac{1}{4}x^4 - \frac{1}{6}x^6)\right|_{-2}^2 = (2-4-\frac{64}{6})-(2-4-\frac{64}{6}) = 0$$

In the example we see that the second integral is zero. This is because the graph is symmetric about the origin. The area below the x-axis is negative, which offsets the positive area above. Suppose the function f is defined and continuous in $[a,b]$, and that it is positive in some subintervals, negative in others, as shown in Fig. 8.3. Let c_1, c_2, c_3 denote three roots of the equation $f(x) = 0$ - that is, three points where the graph crosses the x-axis. The definite integral $\int_b^a f(x)dx$ is the sum of the shaded areas above the x-axis, minus the sum of the shaded areas below the x-axis. The total area bounded by the graph of f, the x-axis, and the lines $x = a$ and $x = b$, on the other hand, is calculated by computing the positive areas in each subinterval $[a,c_1]$, $[c_1,c_2]$, $[c_2,c_3]$, and $[c_3,b]$ in turn according to the previous definitions, and then adding these areas. Specifically, the total shaded area is

$$-\int_a^{c_1} f(x)dx + \int_{c_1}^{c_2} f(x)dx - \int_{c_2}^{c_3} f(x)dx + \int_{c_3}^b f(x)dx$$

This illustrates a general result: the area between the graph of a function f and the x-axis is given by the definite integral $\int |f(x)|dx$ of the absolute value of the integrand $f(x)$, which equals the area under the graph of the nonnegative-valued function $|f(x)|$.

Ex. 8.5. Compute the area bounded by the graph of $f(x) = 1/x^3$, the x-axis, and the two lines $x = -2$ and $x = -1$. Make a drawing. (Hint: $f(x) < 0$ in $[-2, -1]$.)

Ex. 8.6. Let $f(x) = x(x-1)(x-2)$.

(a) Calculate $f'(x)$. Where is $f(x)$ increasing?

(b) Sketch the graph and calculate $\int_0^1 f(x)dx$.

Ex. 8.7. The profit of a firm as a function of its output $x > 0$ is given by

$$f(x) = 4000 - x - \frac{3000000}{x}$$

(a) Find the level of output that maximises profit. Sketch the graph of f.

(b) The actual output varies between 1000 and 3000 units. Compute the average profit

$$I = \frac{1}{2000} \int_{1000}^{3000} f(x)dx$$

If f is a continuous function in an interval that contains the points a, b, and c, then

$$\int_a^b f(x)dx = - \int_b^a f(x)dx$$

$$\int_a^a f(x)dx = 0$$

$$\int_a^b \alpha f(x)dx = \alpha \int_a^b f(x)dx, \text{ where } \alpha \text{ is an arbitrary number}$$

$$\int_b^a f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

Suppose that $F'(x) = f(x)$ for all x in an open interval (a, b). Suppose that $a < t < b$. It follows that $\int_a^t f(x)dx = |_a^t F(x) = F(t) - F(a)$, so

$$\frac{d}{dt} \int_a^t f(x)dx = F'(t) = f(t)$$

In general, if $a(t)$ and $b(t)$ are differentiable and $f(x)$ is continuous, then

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x)dx = f(b(t))b'(t) - f(a(t))a'(t)$$

Ex. 8.8. Evaluate the following integrals

(a)
$$\int_0^3 \left(\frac{1}{3}e^{3t-2} + (t+2)^{-1}\right) dt$$

(b)
$$\int_0^1 (x^2 + 2)^2 dx$$

(c)
$$\int_0^1 \frac{x^2 + x + \sqrt{x+1}}{x+1} dx$$

Ex. 8.9. Find the area between the two parabolas defined by the equations $y + 1 = (x-1)^2$ and $3x = y^2$. (Hint: The points of intersection have integer coordinates.)

Ex. 8.10. Consider the function f defined, for $x > 0$, by $f(x) = 4\ln(\sqrt{x+4} - 2)$.

(a) Show that f has an inverse function g, and find a formula for g.

(b) Draw the graphs of f and g in the same coordinate system.

(c) Give a geometric interpretation of $A = \int_5^{10} 4\ln(\sqrt{x+4}-2)dx$, and explain why
$$A = 10a - \int_0^a (e^{x/2} + 4e^{x/4})dx$$
where $a = f(10)$. Use this equality to express A in terms of a.

8.3   Integration by parts

The rule for differentiating a product allows us to derive an important and useful rule for integrating products. The product rule for differentiation states that

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

Now take the indefinite integral of each side, and then use the rule for integrating a sum. The result is

$$f(x)g(x) = \int f'(x)g(x)dx + \int f(x)g'(x)dx$$

Rearranging this last equation yields the following formula:

$$f(x)g'(x) = f(x)g(x) - \int f'(x)g(x)dx$$

Example 8.3. Use integration by parts to evaluate $\int xe^x dx$.
    Solution:
    Let $f(x) = x$ and $g'(x) = e^x$, implying that $g(x) = e^x$. Then $f(x)g'(x) = xe^x$, so

$$\int xe^x dx = xe^x - \int 1 \cdot e^x dx = xe^x - \int e^x dx = xe^x - e^x + C$$

In the example, if we interchange f and g, then it will be difficult to evaluate. In general, there is an order for choosing f:

1. Inverse functions (e.g. $arcsinx$)

2. Logarithmic functions (e.g. $\ln x$)

3. Power functions (e.g. x^a)

4. Trigonometric functions (e.g. $\sin x$)

5. Exponential functions (e.g. e^x)

Example 8.4. Evaluate the following:

(a)
$$I = \int (1/x) \ln x \, dx$$

(b)
$$J = \int x^3 e^{2x} dx$$

Solution:

(a) Choosing $f(x) = \ln x$ and $g(x) = 1/x$.
$$I = \ln x \cdot \ln x - \int \frac{1}{x} \cdot \ln x \, dx$$

The last integral is equivalent to I. Therefore we have
$$I = \frac{1}{2}(\ln x)^2 + C$$

(b) We choose $f(x) = x^3$ and $g(x) = e^{2x}$. Integrating by parts once we obtain
$$J = \int x^3 e^{2x} dx = x^3 \left(\frac{1}{2}e^{2x}\right) - \int (3x^2) \left(\frac{1}{2}e^{2x}\right) dx = \frac{1}{2}x^3 e^{2x} - \frac{3}{2} \int x^2 e^{2x} dx$$

Doing the same process two more times yields
$$J = \frac{1}{2}x^3 e^{2x} - \frac{3}{4}x^2 e^{2x} + \frac{3}{4}xe^{2x} - \frac{3}{8}e^{2x} + C$$

For definite integrals, the rule is very similar:
$$\int_a^b f(x)g'(x)dx = \left. f(x)g(x) \right|_a^b - \int_a^b f'(x)g(x)dx$$

Example 8.5. Evaluate $\int_0^{10}(1 + 0.4t)e^{-0.05t}dt$.
    Solution:
    Choosing $f(t) = 1 + 0.4$ and $g'(t) = e^{-0.05t}$, we obtain
$$\int_0^{10}(1 + 0.4t)e^{-0.05t}dt = \left. (1 + 0.4t)(-20)e^{-0.05t} \right|_0^{10} - \int_0^{10}(0.4)(-20)e^{-0.05t}dt$$
$$= -100e^{-0.5} + 20 + 8\int_0^{10} e^{-0.05t}dt$$
$$= -100e^{-0.5} + 20 - 160(e^{-0.5} - 1) \approx 22.3$$

Ex. 8.11. Use integration by parts to evaluate the following:

$$(a) \int xe^{-x}dx \qquad (b) \int 3xe^{4x}dx$$

$$(c) \int (1+x^2)e^{-x}dx \qquad (d) \int x\ln x dx$$

Ex. 8.12. Use integration by parts to evaluate the following:

$$(a) \int_{-1}^{1} x\ln(x+2)dx \qquad (b) \int_{0}^{2} x2^x dx$$

$$(c) \int_{0}^{1} x^2 e^x dx \qquad (d) \int_{0}^{3} x\sqrt{1+x}dx$$

In part (d) you should graph the integrand and decide if your answer is reasonable.

8.4   Integration by substitution

In the previous section, we reversed the product rule. In this section we try to reverse the chain rule.

Example 8.6. Evaluate the integrals:

(a)

$$\int (x^2 + 10)^{50} 2x dx$$

(b)

$$\int_{0}^{a} xe^{-cx^2} dx, \text{ where } c \neq 0$$

Solution:

(a) Let $u = x^2 + 10$. Differentiation yields $du = 2xdx$. Inserting these into the integral yields $\int u^{50}du = \frac{1}{51}u^{51} + C$. Therefore

$$\int (x^2 + 10)^{50} 2x dx = \frac{1}{51}(x^2 + 10)^{51} + C$$

(b) Let $u = -cx^2$. Then $du = -2cx\,dx \Rightarrow xdx = -du/2x$. Therefore

$$\int xe^{-cx^2}dx = \int -\frac{1}{2c}e^u du = -\frac{1}{2c}e^u + C = -\frac{1}{2c}e^{-cx^2} + C$$

The definite integral is

$$\int_{0}^{a} xe^{-cx^2} = -\frac{1}{2c} \times \left. e^{-cx^2}\right|_{0}^{a} = \frac{1}{2c}(1 - e^{-ca^2})$$

In general we have the following rule:

Suppose that g is continuously differentiable, and $f(u)$ is continuous at all points u belonging to the relevant range of g. Then,

$$\int f(g(x))g'(x)dx = \int f(u)du$$

where $u = g(x)$.

Example 8.7. Consider a more complicated integral $\int \frac{x-\sqrt{x}}{x+\sqrt{x}}dx$, $x > 0$.

Solution:

Use the substitution $u = \sqrt{x}$. Then $x = u^2$ and $dx = 2u\,du$. So

$$\int \frac{x-\sqrt{x}}{x+\sqrt{x}}dx = \int \frac{u^2-u}{u^2+u}2udu = 2\int\left(u-2+\frac{2}{u+1}\right)du$$

$$= u^2 - 4u + 4\ln|u+1| + C$$

$$= x - 4\sqrt{x} + 4\ln(\sqrt{x}+1) + C$$

We can summarise the method of substitution as:

In order to find $\int G(x)dx$

1. Pick out a "part" of $G(x)$ and introduce this "part" as a new variable, $u = g(x)$.

2. Compute $du = g'(x)dx$.

3. Using the substitution $u = g(x)$, $du = g'(x)dx$, transform, if possible, $\int G(x)dx$ to an integral of the form $\int f(u)du$.

4. Find, if possible, $\int f(u)du = F(u) + C$.

5. Replace u by $g(x)$.

   Then the final answer is $\int G(x)dx = F(g(x)) + C$.

8.4.1   Integrating rational functions and partial fractions

Example 8.8. Calculate the integral $\int \frac{x^4+3x^2-4}{x^2+2x}dx$.

Solution:

We apply polynomial division to the integrand, which yields

$$\frac{x^4 + 3x^2 - 4}{x^2 + 2x} = x^2 - 2x + 7 - \frac{14x+4}{x^2+2x}$$

The last term can be expressed as the sum of two **partial fractions**:

$$\frac{14x+4}{x^2+2x} = \frac{A}{x} + \frac{B}{x+2}$$

Solving for simultaneous equations we obtain $A = 2$ and $B = 12$. Therefore the integral becomes

$$\int \frac{x^4+3x^2-4}{x^2+2x} = \int x^2 - 2x + 7 - \frac{2}{x} + \frac{12}{x+2}dx$$

$$= \frac{1}{3}x^3 - x^2 + 7x - 2\ln|x| - 12\ln|x+2| + C$$

Ex. 8.13. Find the following integrals:

$$(a) \int_0^1 x\sqrt{1+x^2}dx \qquad (b) \int_0^e \frac{\ln y}{y}dy$$

$$(c) \int_1^3 \frac{1}{x^2}e^{2/x}dx (d) \int_5^8 \frac{x}{x-4}dx$$

Hint: In (d), use both integration by substitution and expansion of partial fractions, as alternative methods to find the integral.

Ex. 8.14. Calculate the following integrals:

(a)
$$\int_0^1 (x^4 - x^9)(x^5 - 1)^{12}dx$$

(b)
$$\int (\ln x/\sqrt{x})dx$$

(c)
$$\int_0^4 \frac{1}{\sqrt{1+\sqrt{x}}}$$

Ex. 8.15. Calculate the following integrals:

(a)
$$\int_1^4 \frac{e^{\sqrt{x}}}{\sqrt{x}(1+e^{\sqrt{x}})}dx$$

(b)
$$\int_0^{1/3} \frac{1}{e^x + 1}dx$$

(c)
$$\int_{8.5}^{41} \frac{1}{\sqrt{2x-1} - \sqrt[4]{2x-1}}dx$$

Ex. 8.16. Use the method of partial fractions to find the integrals:

$$(a) \int \frac{x}{(x+1)(x+2)}dx \qquad (b) \int \frac{1-2x}{x^2 - 2x - 15}dx$$

8.5   Infinite intervals of integration

In general, suppose f is a function that is continuous for all $x \geq a$. Then $\int_a^b f(x)dx$ is defined for each $b \geq a$. If the limit of this integral as $b \to \infty$ exists (and is finite), then we say that f is integrable over $[a, \infty)$, and define

$$\int_a^\infty f(x)dx = \lim_{b \to \infty} \int_a^b f(x)dx$$

The improper integral $\int_a^\infty f(x)dx$ is then said to converge. If the limit does not exist, then it is said to diverge.

Example 8.9. The exponential distribution in statistics is defined by the density function $f(x) = \lambda e^{-\lambda x}$, where $x \geq 0$ and λ is a positive constant. Show that the area below the graph of f over $[0, \infty)$ is equal to 1.

For $b > 0$, the area below the graph of f over $[0, b]$ is equal to

$$\int_0^b \lambda e^{-\lambda x} dx = \Big|_0^b (-e^{-\lambda x}) = e^{-\lambda b} + 1$$

As $b \to \infty$, so $e^{-\lambda b} + 1$ approaches 1. Therefore

$$\int_0^\infty \lambda e^{-\lambda x} dx = \lim_{b \to \infty} \int_0^b \lambda e^{-\lambda x} dx = \lim_{b \to \infty} (-e^{-\lambda b} + 1) = 1$$

8.5.1 Test for convergence

Suppose that for all $x \geq a$, f and g are continuous, and $|f(x)| \leq g(x)$. If $\int_a^\infty g(x) dx$ converges, then $\int_a^\infty f(x) dx$ also converges, and

$$\left| \int_a^\infty f(x) dx \right| \leq \int_a^\infty g(x) dx$$

Example 8.10. The function $f(x) = e^{-x^2}$ is extremely important in statistics. When multiplied by a suitable constant, $1/\sqrt{\pi}$, it is the density function associated with a **Gaussian**, or **normal**, distribution. We want to show the integral

$$\int_{-\infty}^\infty e^{-x^2} dx$$

converges. Because $f(x)$ is symmetric about the y-axis, one has $\int_{-\infty}^\infty e^{-x^2} dx = 2 \int_0^\infty e^{-x^2} dx$, so it suffices to prove that $\int_0^\infty e^{-x^2} dx$ converges. To show this, subdivide the interval of integration so that

$$\int_0^\infty e^{-x^2} dx = \int_0^1 e^{-x^2} dx + \int_1^\infty e^{-x^2} dx$$

$\int_0^1 e^{-x^2} dx$ presents no problem because it is the integral of a continuous function over a bounded interval. For $x \geq 1$, one has $x^2 \geq x$ and so $0 \leq e^{-x^2} \leq e^{-x}$. Now $\int_1^\infty e^{-x} dx$ converges (to $1/e$), so according to the test for convergence, the integral $\int_1^\infty e^{-x^2} dx$ must also converge.

In fact, one can use polar substitution to show that

$$\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$$

Ex. 8.17. In connection with Example 8.9, find the following:

(a)

$$\int_0^\infty x \lambda e^{-\lambda x} dx$$

(b)

$$\int_0^\infty (x - 1/\lambda^2) \lambda e^{-\lambda x} dx$$

(c)
$$\int_0^\infty (x - 1/\lambda^3)\lambda e^{-\lambda x} dx$$

The three numbers you obtain are called respectively the **expectation**, the **variance**, and the **third central moment** of the exponential distribution.

Ex. 8.18. The function f is defined for $x > 0$ by $f(x) = (\ln x)/x^3$.

(a) Find the maximum and minimum points of f, if any.

(b) Examine the convergence of $\int_0^1 f(x)dx$ and $\int_1^\infty f(x)dx$.

Ex. 8.19. Show that

$$\int_{-2}^3 \left(\frac{1}{\sqrt{x+2}} + \frac{1}{\sqrt{3-x}} \right) dx = 4\sqrt{5}$$

Ex. 8.20. In statistics, the normal, or Gaussian, density function with mean μ and variance σ^2 is defined by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp[-(x - \mu)^2/2\sigma^2]$$

in the interval $(-\infty, \infty)$. Prove that

(a)
$$\int_{-\infty}^\infty f(x)dx = 1$$

(b)
$$\int_{-\infty}^\infty xf(x)dx = \mu$$

(c)
$$\int_{-\infty}^\infty (x - \mu)^2 f(x)dx = \sigma^2$$

8.6   Separate and linear differentiation equations

8.6.1   Separate equations

A differential equation of the type

$$\dot{x} = f(t)g(x)$$

is called **separable**.

The following general method for solving separable equations is as follows:

(i) Write the differential equation as

$$\frac{dx}{dt} = f(t)g(x)$$

(ii) Separate the variables

$$\frac{1}{g(x)}dx = f(t)dt$$

(iii) Integrate each side

$$\int \frac{1}{g(x)}dx = \int f(t)dt$$

(iv) Evaluate the two integrals.

Example 8.11. Solve the differential equation

$$\frac{dx}{dt} = e^t x^2$$

and find the solution curve, which is also called the **integral curve**, that passes through the point $(t, x) = (0, 1)$.

Solution:

We observe first that $x(t) \equiv 0$ is one (trivial) solution. To find the other solutions, we follow the last three parts of the recipe:

Separate: $(1/x^2)dx = e^t dt$;

Integrate: $\int (1/x^2)dx = \int e^t dt$;

Evaluate: $-1/x = e^t + C$.

It follows that

$$x = \frac{-1}{e^t + C}$$

To find the integral curve through (0, 1), we must determine the correct value of C. Because we require $x = 1$ for $t = 0$, it follows that $1 = -1/(1 + C)$, so $C = -2$. Thus, the integral curve passing through (0, 1) is $x = 1/(2 - e^t)$.

Example 8.12. (**Economic growth**) Let $X = X(t)$ denote the national product, $K = K(t)$ the capital stock, and $L = L(t)$ the number of labour in a country at time t. Suppose that, for all $t \geq 0$,

$$(a) X = \sqrt{K}\sqrt{L} \qquad (b) \dot{K} = 0.4X \qquad (c) L = e^{0.04t}$$

Derive from these equations a single differential equation for $K = K(t)$, and find the solution of that equation when $K(0) = 10000$.

Solution:

From equations (a)–(c), we derive the single differential equation

$$\dot{K} = \frac{dK}{dt} = 0.4\sqrt{K}\sqrt{L} = 0.4e^{0.02t}\sqrt{K}$$

Solve the separable equation

$$2\sqrt{K} = 20e^{0.02t} + C$$

If $K = 10000$ for $t = 0$, then $2\sqrt{10000} = 20 + C$, so $C = 180$. Then $\sqrt{K} = 10e^{0.02t} + 90$, and so the required solution is

$$K(t) = (10e^{0.02t} + 90)^2 = 100(e^{0.02t} + 9)^2$$

The capital–labour ratio has a somewhat bizarre limiting value in this model: as $t \to \infty$, so

$$\frac{K(t)}{L(t)} = 100 \times \frac{(e^{0.02t} + 9)^2}{e^{0.04t}} = 100 \left[\frac{e^{0.02t+9}}{e^{0.02t}}\right]^2 = 100(1 + 9e^{-0.02t})^2 \to 100$$

8.6.2   First order differential equations

A first-order linear differential equation is one that can be written in the form

$$\dot{x} + a(t)x = b(t)$$

where $a(t)$ and $b(t)$ denote known continuous functions of t in a certain interval, and $x = x(t)$ is the unknown function.

To find the solution, we multiply both sides by an **integrating factor** $e^{\int a(t)dt}$. This gives

$$\dot{x}e^{\int a(t)dt} + a(t)e^{\int a(t)dt} = b(t)e^{\int a(t)dt}$$

Then the LHS happens to be the exact derivative of the product xe^{at}. Thus

$$\frac{d}{dt}(xe^{\int a(t)dt}) = b(t)e^{\int a(t)dt}$$

Hence the solution is

$$x = Ce^{-\int a(t)dt} + e^{-\int a(t)dt}\int e^{\int a(t)dt}b(t)dt$$

Example 8.13. (**Economic growth**) Consider the following model of economic growth:

$$(a) X(t) = 0.2K(t) \qquad (b)\dot{K}(t) = 0.1X(t) + H(t) \qquad (c)N(t) = 50e^{0.03t}$$

$X(t)$ is annual GDP, $K(t)$ is capital stock, $H(t)$ is the net inflow of foreign investment per year, and $N(t)$ is the size of the population, all measured at time t.

Assume that $H(t) = 10e^{0.04t}$ and derive from these equations a differential equation for $K(t)$. Find its solution given that $K(0) = 200$. Find also an expression for $x(t) = X(t)/N(t)$, which is domestic product per capita.

Solution:

From (a) and (b), it follows that $K(t)$ must satisfy the linear equation

$$\dot{K}(t) - 0.02K(t) = 10e^{0.04t}$$

Solve the first order differential equation

$$K(t) = Ce^{0.02t} + e^{0.02t}\int e^{-0.02t}10e^{0.04t}dt = Ce^{0.02t} + 10e^{0.02t}\int e^{0.02t}dt$$

$$= Ce^{0.02t} + (10/0.02)e^{0.04t} = Ce^{0.02t} + 500e^{0.04t}$$

For $t = 0$, $K(0) = 200 = C + 500$, so $C = -300$. Thus, the solution is

$$K(t) = 500e^{0.04t} - 300e^{0.02t}$$

Per capita production is $x(t) = X(t)/N(t) = 0.2K(t)/50e^{0.03t} = 2e^{0.01t} - 1.2e^{-0.01t}$.

Ex. 8.21. Solve the following differential equations:

$$(a)\dot{x} = e^{2t}/x^2 \qquad (b)\dot{x} = e^{-t+x} \qquad (c)\dot{x} - 3x = 18$$

$$(a)\dot{x} = (1+t)^6/x^6 \qquad (b)\dot{x} - 2x = -t \qquad (c)\dot{x} + 3x = te^{t^2-3t}$$

Ex. 8.22. Suppose $Y = Y(t)$ is GDP, $C(t)$ is consumption, and I is investment, which is constant. Suppose $\dot{Y} = \alpha(C + \bar{I} - Y)$ and $C = aY + b$, where a, b, and α are positive constants with $a < 1$.

(a) Derive a differential equation for Y.

(b) Find its solution when $Y(0) = Y_0$ is given. What happens to $Y(t)$ as $t \to \infty$?

8.7 Review exercises

Ex. 8.23. Find the following integrals:

$$(a) \int_1^\infty \frac{5}{x^5}dx \qquad (b) \int_0^1 x^3(1+x^4)^4 dx \qquad (c) \int_0^\infty \frac{-5t}{e^t}dt \qquad (d) \int_1^e (\ln x)^2 dx$$

$$(e) \int_0^2 x^2\sqrt{x^3+1}dx \qquad (f) \int_{-\infty}^0 \frac{e^{3z}}{e^{3z}+5}dz \qquad (g) \int_{1/2}^{e/2} x^3 \ln(2x)dx \qquad (h) \int_1^\infty \frac{e^{-\sqrt{x}}}{\sqrt{x}}dx$$

Ex. 8.24. Find the following integrals:

$$(a) \int_0^{25} \frac{1}{9+\sqrt{x}}dx \qquad (b) \int_2^7 t\sqrt{t+2}dt \qquad (c) \int_0^1 57x^2 \sqrt[3]{19x^3+8}dx$$

Ex. 8.25. For the following two cases, find the equilibrium price and quantity and calculate the consumer and producer surplus when the inverse demand curve is $f(Q)$ and the inverse supply curve is $g(Q)$:

(a) $f(Q) = 100 - 0.05Q$ and $g(Q) = 10 + 0.1Q$.

(b) $f(Q) = \frac{50}{Q+5}$ and $g(Q) = 4.5 + 0.1Q$.

Ex. 8.26. Define f for $t > 0$ by $f(t) = 4(\ln t)^2/t$.

(a) Find $f'(t)$ and $f''(t)$.

(b) Find possible local extreme points, and sketch the graph of f.

(c) Calculate the area below the graph of f over the interval $[1, e^2]$.

Ex. 8.27. Solve the following differential equations:

$$(a)\dot{x} = tx^2 \qquad (b)2\dot{x} + 3x = -15 \qquad (c)\dot{x} - 3x = 30$$

$$(a)\dot{x} + 5x = 10t \qquad (b)\dot{x} + \frac{1}{2}x = e^t \qquad (c)\dot{x} + 3x = t^2$$

Ex. 8.28. Let $V(x)$ denote the number of litres of fuel left in an aircraft's fuel tank if it has flown x km. Suppose that $V(x)$ satisfies the following differential equation: $V'(x) = -aV(x) - b$. Here, the fuel consumption per km is a constant $b > 0$. The term $-aV(x)$, with $a > 0$, is due to the weight of the fuel.

(a) Find the solution of the equation with $V(0) = V_0$.

(b) How many km, x^*, can the plane fly if it takes off with V_0 litres in its tank?

(c) What is the minimum number of litres, V_m, needed at the outset if the plane is to fly $\hat{x}$ km?

(d) Let $b = 8$, $a = 0.001$, $V_0 = 12000$, and $\hat{x} = 1200$. Find x^* and V_m in this case.

Chapter 9

Functions of Many Variables

9.1 Functions of 2 Variables

One prime example is the Cobb-Douglas production function

$$z = F(x, y) = Ax^a y^b, \qquad (9.1)$$

where A, a, b are constants, $x, y > 0$.

Ex. 9.1. For (9.1), verify that

$$F(2x, 2y) = 2^{a+b} F(x, y)$$
$$F(tx, ty) = t^{a+b} F(x, y)$$
$$F(x + h, y) - F(x, y) = Ay^b[(x+h)^a - x^a]$$

and give economic interpretations.

Ex. 9.2. Compute $F(2x, 2y), F(tx, ty)$ for an arbitrary $t > 0$, and $F(x+h, y) - F(x, y)$ for each of the following functions

(a) $F(x, y) = x + 2y$
(b) $F(x, y) = xy^2$

(c) $F(x, y) = 10x^{1/2} y^{1/3}$
(d) $F(x, y) = \sqrt{x^2 + y^2}$

The **Domain** of a function of 2 variables $F(x, y)$ is the set of values of x, y in which the function is defined.

Ex. 9.3. Find the domain of the following functions

(a) $f(x, y) = \sqrt{x - 1} + \sqrt{y}$
(b) $f(x, y) = \dfrac{2}{(x^2 + y^2 - 4)^{1/2}} + \sqrt{9 - (x^2 + y^2)}$

(c) $f(x, y) = \dfrac{1}{e^{x+y} - 3}$
(d) $f(x, y) = \ln(x - a)^2 + \ln(y - b)^2$

(e) $f(x, y) = 2\ln(x - a) + 2\ln(y - b)$
(f) $f(x, y) = \sqrt{1 - xy}$

(g) $f(x, y) = \ln(2 - (x^2 + y^2))$
(h) $f(x, y) = \sqrt{y - x^2} - \sqrt{\sqrt{x} - y}$

[*] Draw the domain in the xy-plane.

9.2 Partial Derivatives

Definition 9.1. If $z = f(x, y)$, then

1. $\partial f / \partial x$ denotes the derivative of $f(x, y)$ wrt x when y is held constant.

2. $\partial f / \partial y$ denotes the derivative of $f(x, y)$ wrt y when y is held constant.

Formal definition

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}$$

By taking the partial derivative of x, we know the rate of change of the function wrt to a change in x.

Notation 9.1. The partial derivative of f wrt x can be written as

$$\frac{\partial f}{\partial x} = f'_x(x, y)$$

This is known as **first-order derivatives**.

Example 9.1. The production function

$$F(K, L) = K^\alpha L^\beta \quad (\alpha, \beta > 0) \tag{9.2}$$

The partial derivative of F wrt to K

$$\frac{\partial F}{\partial K} = \alpha K^{\alpha-1} L^\beta$$

The partial derivative of F wrt to L

$$\frac{\partial F}{\partial L} = \alpha K^\alpha L^{\beta-1}$$

Since F'_K, F'_L are positive, an increase in K or L leads to an increase in F. As an economic interpretation, provided that the economy is currently at $F(1, 1)$ and $\alpha = 0.3, \beta = 0.7$, holding L fixed at 1, F will increase by $\alpha K^{\alpha-1} L^\beta = 0.3$ units per unit increase in K.

Second-order Derivatives The derivative of the derivative is called the second-order derivative.

Notation 9.2. The second-order partial derivative of f wrt x is

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} = f''_{xx}$$

Other partials

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x} = f''_{xy}, \quad \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y} = f''_{yx}, \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial^2 y} = f''_{yy}$$

Ex. 9.4. Find the first and second-order partial derivatives of

$$(a) \ z = x^3 y^2 \qquad\qquad (b) \ z = x^5 - 3x^2 y + y^6$$

$$(c) \ z = \frac{x - y}{x + y} \qquad\qquad (d) \ z = \sqrt{x^2 + y^2}$$

9.2.1 Economic Interpretation

Marginal Products

Let $Q = F(K, L)$ be a production function. Assume that the firm currently employs K^* units of capital and L^* units of labor. The rate at which output changes with respect to capital K, keeping L fixed is:

$$\triangle Q \approx \frac{\partial F}{\partial K}(K^*, L^*) \cdot \triangle K.$$

This is known as the marginal product of capital.

Elasticity

If $Q_1 = Q_1(P_1, P_2, I)$ represents the demand for good 1 in terms of the prices of goods 1 and 2 and income, then $\partial Q_1 / \partial P_1$ is the rate of change of demand wrt own price. For a small $\triangle P_1$, the demand for good 1 will change by:

$$\triangle Q_1 \approx \frac{\partial Q_1}{\partial P_1} \triangle P_1.$$

In percentage terms:

$$\varepsilon_1 = \frac{\%\text{change in demand}}{\%\text{change in own price}} = \frac{\triangle Q_1 / Q_1}{\triangle P_1 / P_1} = \frac{P_1}{Q_1} \cdot \frac{\triangle Q_1}{\triangle P_1}.$$

Since:

$$\frac{\triangle Q_1}{\triangle P_1} = \frac{Q_1(P_1 + \triangle P_1) - Q_1(P_1)}{P_1 + \triangle P_1} \approx \frac{\partial Q_1}{\partial P_1},$$

For small $\triangle P_1$ ($\triangle P_1 \to 0$), the elasticity can be expressed as:

$$\varepsilon_1 = P_1^* \frac{\frac{\partial Q_1}{\partial P_1}(P_1^*, P_2^*, I^*)}{Q_1(P_1^*, P_2^*, I^*)}.$$

It's usually negative so $\varepsilon \in (-\infty, 0]$. If $0 \le |\varepsilon_1| < 1$, the good is inelastic. If $|\varepsilon_1| > 1$, it is elastic. A small percentage change in price results in a larger percentage change in quantity demanded.

What happens to the demand for good 1 when the price of good 2 changes?

To study such sensitivity, we use cross-price elasticity of demand.

$$\varepsilon_{Q_1, P_2} = \frac{\triangle Q_1 / Q_1}{\triangle P_2 / P_2}.$$

In calculus terms:

$$\varepsilon_{Q_1, P_2} = \frac{P_2^* \frac{\partial Q_1}{\partial P_2}(P_1^*, P_2^*, I^*)}{Q_1(P_1^*, P_2^*, I^*)}$$

The signs are ambiguous. If ε_{P_1, Q_2} and ε_{P_2, Q_1} are both **positive**, they are **substitutes**. The price of good 1 increases making consumers demand more of good 2. If they are both **negative**, the goods are **complements**.

Finally, the sensitivity of demand to changes in income is:

$$\varepsilon_{Q_1, I} = \frac{\triangle Q_1 / Q_1}{\triangle I / I}.$$

In calculus:

$$\varepsilon_{Q_1, I} = \frac{I}{Q_1} \frac{\partial Q_1}{\partial I},$$

evaluated at (P_1^*, P_2^*, I^*).

9.3   Geometric Representation

9.3.1   The Graph of a Function of 2 Variables

Suppose $z = f(x, y)$. For instance, let's consider the function

$$x^2 + y^2 = z$$

Then its graph can be represented by a 3d surface.
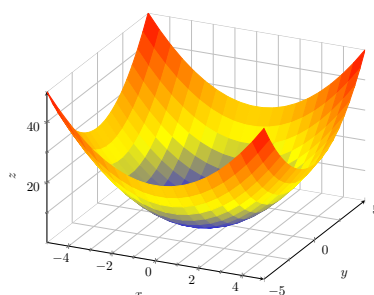
Another more convenient representation is to use the level curves. The idea is to use a plane and cut the surface into different altitudes. The (x, y) points belonging to the same altitude should produce the same z, hence the name *level curves* since the (x, y) points on this curve produce the same level of z.
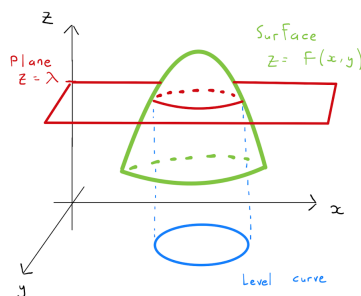
Applying to $z = x^2 + y^2$

Ex. 9.5. Suppose a Cobb-Douglas production function has the following form

$$F(K, L) = AK^\alpha L^\beta = c$$

Assuming that $\alpha = 0.3, \beta = 0.7, A = 0.4$, we can generate a geometric representation of this production function. The curve is called an isoquant ("equal quantity").

(a) in 3D **(b)** in 2D

Fig. 9.1. $x^2 + y^2 = c$

(a) in 3D **(b)** in 2D
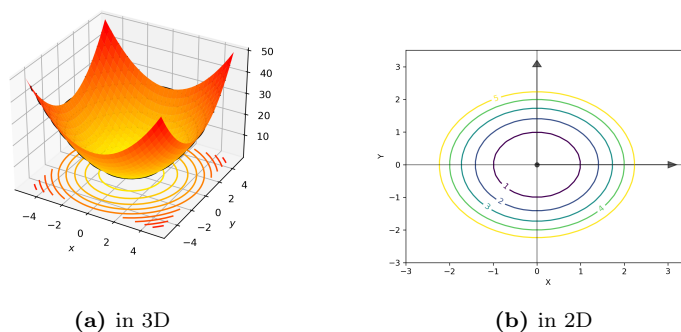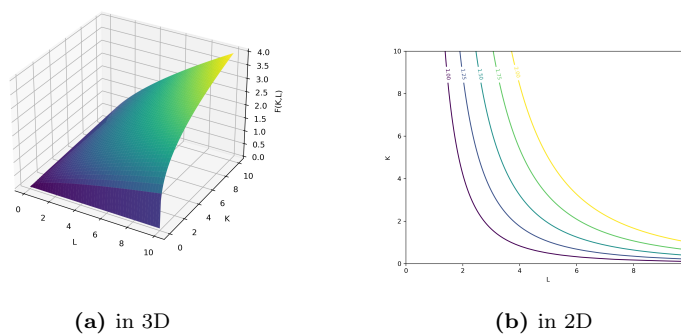
Fig. 9.2. $Y = F(K, L)$

9.3.2 Geometric Representation of Partial Derivatives

A function $z = f(x, y)$ can be geometrically represented by drawing its graph in $\mathbb{R}^3$ or its level curves in $\mathbb{R}^2$.

The graph of $x \longmapsto f(x, b)$ on the slice $\{y = b\}$. The graph of $y \longmapsto f(a, y)$ on the slice $\{x = a\}$.
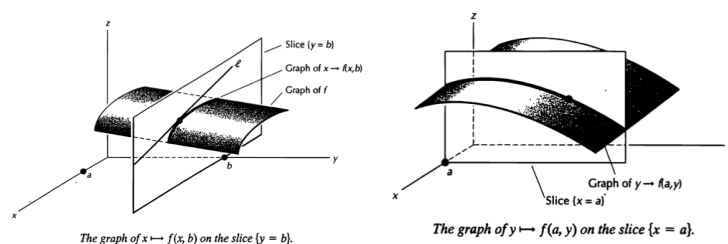
Fig. 9.3. The mapping of $x \mapsto f(x, b)$ (left) and $y \mapsto f(a, y)$ (right)

When we study $\frac{\partial f}{\partial x}(a, b)$, we are holding y constant at b and looking at variations of x around $x = a$. That means, we slice f by a two-dimensional slice $\{y = b\}$ in $\mathbb{R}^3$ and see how f moves when x changes (left)

On this slide, the graph of f is a curve – the graph of the function of **one**

variable $x \mapsto f(x, b)$. The partial derivative $\partial f / \partial x(a, b)$ is the slope of the tangent line ℓ to this graph on this slice. By keeping x fixed and letting y vary only, we are measuring the rate of change of f in the vertical direction.

Similarly, if we hold x constant at $x = a$ and let y change, we have the graph 9.4. Let us now focus on the study of the level curves $z = f(x, y)$. Then, $\partial f / \partial x(a, b)$ is the (infinitesimal) variation of f restricted to the vertical line $x = a$. Some variations are depicted in the below picture. The closer the level curves are to each other, the more f changes as y increases by one unit, and the larger is $(\partial f / \partial y)$.

$(\partial f / \partial y)(a, b)$ measures the change in f on $\{x = a\}$.
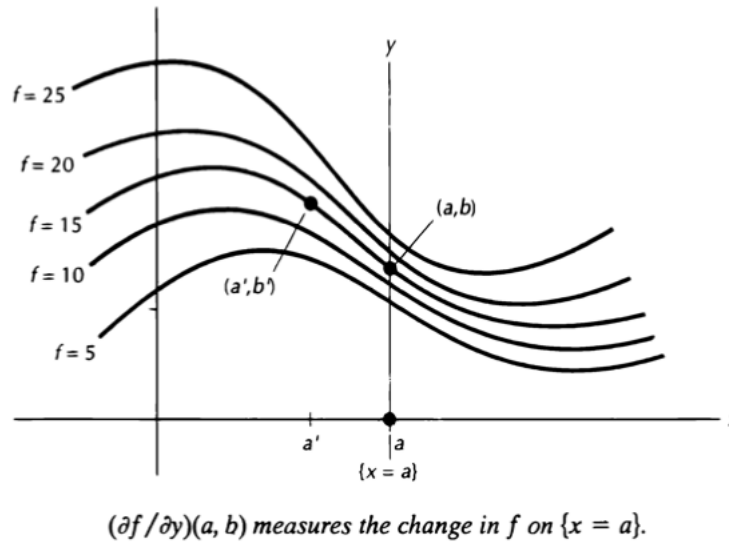
Fig. 9.4. Level curve at $\{x = a\}$

9.4 Functions of More than 2 Variables

We can extend the concept of partial derivative to functions of more than two variables

Definition 9.2. If $z = f(x) = f(x_1, x_2, \ldots, x_n)$, then $\partial f / \partial x_i$ for $i = 1, 2, \ldots, n$ means the partial derivatives of f wrt x_i, when all other variables $x_{j \neq i}$ are held constant.

Ex. 9.6. In a study of milk production, Hjelm and Sandquist found that

$$y = 2.9 x_1^{0.015} x_2^{0.25} x_3^{0.35} x_4^{0.408} x_5^{0.03}$$

where y is the output of milk, and $x_1, \ldots, x_5$ are the quantities of 4 different kinds of inputs, such as workers' effort, grass consumption, etc. If all factors of production were doubled, what would happen to y? Write the relation in log-linear form (by taking logs of both sides).

Ex. 9.7. Consider a Cobb-Douglas production function that consists of 3 input factors: K, L, T, which stand for capital, labor, and land.

$$Y = AK^\alpha L^\beta T^\theta \quad (A, \alpha, \beta, \theta > 0) \tag{9.3}$$

1. Find the marginal products of each factor.

2. Find the direct second-order partials $F''_{KK}, F''_{LL}, F''_{TT}$

3. Find the cross second-order partials $F''_{KL}, F''_{KT}, F''_{LT}, F''_{LK}, F''_{TK}, F''_{TL}$.

4. Find the elasticity of Y wrt to each factors. [1]

Ex. 9.8. For which pairs of (x, y) that the following functions are defined?

(a) $3xy^3 - 45x^4 - 3y, \& (b) \ln(2 - (x^2 + y^2))$,

(c) $\sqrt{1 - xy}, \& (d) \dfrac{1}{\sqrt{x + y - 1}}$,

(e) $\sqrt{x^2 - y^2} + \sqrt{x^2 + y^2 - 1}, \& (f) \sqrt{y - x^2} - \sqrt{\sqrt{x} - y}$

Ex. 9.9. Find the first-order partial derivatives and second-order partial derivatives of

(a) $F(K, L) = (K^{1/2} + L^{1/2})^2$,   (b) $g(t, w) = \dfrac{3t}{dw} + wt^2$,

(c) $g(x_1, x_2, x_3) = (x_1^2 + x_2^2 + x_3^2)^{1/2}$ (d) $f(x, y, z) = 2x^2yz - y^3 + x^2z^2$

9.4.1 Directional Derivatives and Gradients

We want to compute the rate of change of a function $F(x_1, \ldots, x_n)$ at a given point $\mathbf{x}^*$ in any given direction $\mathbf{v} = (v_1, \ldots, v_n)$. First, we parameterize the direction $\mathbf{v}$ from the point $\mathbf{x}^*$ by writing a parameterized equation of the line through $\mathbf{x}^*$ in the direction $\mathbf{v}$:

$$\mathbf{x} = \mathbf{x}^* + t\mathbf{v}.$$

To see how F changes along that line, first evaluate F along this line

$$g(t) \equiv F(\mathbf{x}^* + t\mathbf{v}) = F(x_1^* + tv_1, \ldots, x_n^* + v_n)$$

and then use the Chain Rule to take the derivative of g at $t = 0$:

$$g'(0) = \frac{\partial F}{\partial x_1}(\mathbf{x}^*)v_1 + \cdots + \frac{\partial F}{\partial x_n}(\mathbf{x}^*)v_n$$

or in matrix notation,

$$\left(\frac{\partial F}{\partial x_1}(\mathbf{x}^*) \quad \cdots \quad \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \right) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = DF_{\mathbf{x}^*} \cdot \mathbf{v}. \tag{9.4}$$

---

[1] The formula for elasticity of y wrt x is

$$\varepsilon_x y = \frac{\partial y/y}{\partial x/x}, \quad \text{or equivalently} \quad = \frac{\partial \ln y}{\partial \ln x}$$

Expression 9.4 is called **the derivative of F at x^* in the direction v**. Some other notation for this directional derivative:

$$\frac{\partial F}{\partial \mathbf{v}}(\mathbf{x}^*), \ \mathbf{or} \ \ D_{\mathbf{v}}F(\mathbf{x}^*).$$

If we want to know how F changes at $\mathbf{x}^*$ in the direction $\mathbf{e}_1 = (1, 0, \ldots, 0)$ parallel to the x_1 axis, we compute

$$DF_{\mathbf{x}^*}(\mathbf{e}_1) = \left(\frac{\partial F}{\partial x_1}(\mathbf{x}^*) \quad \frac{\partial F}{\partial x_2}(\mathbf{x}^*) \quad \ldots \quad \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \right) \cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \frac{\partial F}{\partial x_1}(\mathbf{x}^*).$$

Example 9.2. Consider the production function:

$$Q = F(K, L) = 4K^{3/4}L^{1/4}$$

at $(K_0, L_0) = (5, 3)$. The derivative of F at $(5, 3)$ in the direction $(1, 1)$ is

$$\frac{\partial F}{\partial K}(K_0, L_0) \times 1 + \frac{\partial F}{\partial L}(K_0, L_0) \times 1 = 3K_0^{-1/4}L_0^{1/4} \times 1 + K_0^{3/4}L_0^{-3/4} \times 1 = 2.64 + 1.47 = 4.11$$

9.4.2   Gradient Vector

So far, the derivative of $y = F(x_1, \ldots, x_n)$ at a point $\mathbf{x}^*$ has been written in row matrix:

$$DF_{\mathbf{x}^*} = \left(\frac{\partial F}{\partial x_1}(\mathbf{x}^*) \quad \ldots \quad \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \right)$$

Let's write it in a column matrix form:

$$\begin{pmatrix} \frac{\partial F}{\partial x_1}(\mathbf{x}^*) \\ \vdots \\ \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}$$

We think of this column matrix as a **vector in $\mathbb{R}^n$** with tail at $\mathbf{x}^*$. We write it as $\nabla F(\mathbf{x}^*)$. This is the **gradient** or **gradient vector** of F at $\mathbf{x}^*$.

The important characteristics of a vector are its length and direction.

$$\nabla F(\mathbf{x}^*) \cdot \mathbf{v} = \begin{pmatrix} \frac{\partial F}{\partial x_1}(\mathbf{x}^*) \\ \vdots \\ \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(\mathbf{x}^*) v_i.$$

Let's ignore the length of $\mathbf{v}$ ($||\mathbf{v}||$) and focus on the direction of $\mathbf{v}$. To do so, let's assume $\mathbf{v}$ has length 1. Then, the $\nabla F(\mathbf{x}^*) \cdot \mathbf{v}$ measures the rate at which F rises or falls as one moves it out from $\mathbf{x}^*$ in the direction $\mathbf{v}$. By the usual property of dot product, it should be:

$$\nabla F(\mathbf{x}^*) \cdot \mathbf{v} = ||\nabla F(\mathbf{x}^*)|| \ ||\mathbf{v}|| \cos \theta = ||\nabla F(\mathbf{x}^*)|| \cos \theta, \qquad (9.5)$$

because $||\mathbf{v}|| = 1$, and θ is the angle between vectors $\nabla F(\mathbf{x}^*)$ and $\mathbf{v}$ at the base point $\mathbf{x}^*$, as depicted in Fig. 9.5 (a) below.

At this point, we can ask **in what direction does F increase most rapidly?**. Since $\cos \theta$ varies between -1 and 1, Eq.9.5 implies that $\nabla F(\mathbf{x}^*) \cdot \mathbf{v}$ will be largest among vectors v of unit length when $\cos \theta = 1$, which means, when $\theta = 0^o$. Therefore, $\nabla F(\mathbf{x}^*) \cdot \mathbf{v}$ is largest as $\mathbf{v}$ varies among vectors of unit length, when $\mathbf{v}$ points in THE SAME DIRECTION as $\nabla F(\mathbf{x}^*)$.

(a) dot prod: $\nabla F(\mathbf{x}^*)$ and $\mathbf{v}$. **(b)** Production Function
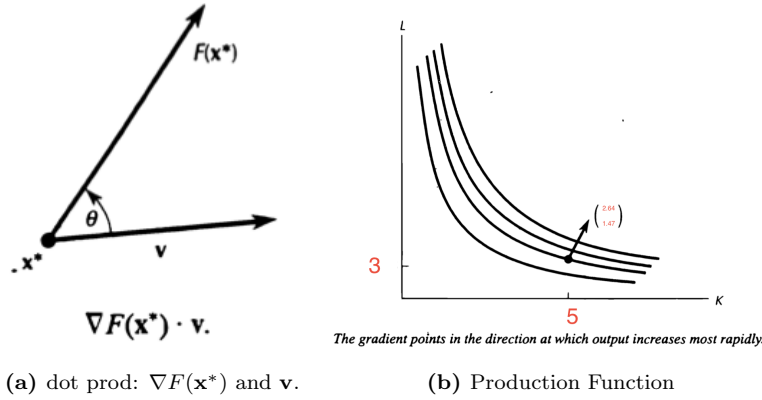
Fig. 9.5. Gradient Vector $\nabla F(\mathbf{x}^*) \cdot \mathbf{v}$.

Theorem 9.1. *Let $F : \mathbb{R}^n \mapsto R^1$ be a C^1 function. At any point $\mathbf{x}$ in the domain of F at which $\nabla F(\mathbf{x}) \neq 0$, the gradient vector $\nabla F(\mathbf{x})$ points at $\mathbf{x}$ into the direction in which F increases most rapidly.*

To sum up, the Gradient ∇ tells us the direction that the curve increases the most, and its magnitude $|\nabla|$ is the rate of change in that direction.

Example 9.3. Going back to the production function:

$$Q = F(K, L) = 4K^{3/4}L^{1/4}$$

at $(K_0, L_0) = (5, 3)$. If we want to know in what proportions to add K and L to $(5, 3)$ to increase production most rapidly, we compute the gradient vector

$$\nabla F(5, 3) = \begin{pmatrix} 3K^{-1/4}L^{1/4} \\ K^{3/4}L^{-3/4} \end{pmatrix} |_{(5,3)} = \begin{pmatrix} 2.64 \\ 1.47 \end{pmatrix}$$

which implies that, by adding K and L at the ratio $2.64/1.47$, we get the proportion necessary to increase production most rapidly, as seen in Fig 9.5 (b).

One interesting question arises is that, **what happens to the function if its gradient is zero?**

The gradient of a function measures the rate and direction of change of the function. It is a vector that points in the direction of the steepest ascent or descent of the function at a particular point. When the gradient of a function is zero, it means that the function is not changing in any direction at that point.

Intuitively, if the gradient of a function is zero at a particular point, it means that the function has reached a critical point, which can be a minimum, maximum, or saddle point. At a minimum or maximum, the function has stopped increasing or decreasing in any direction, while at a saddle point, the function changes in different directions.

For example, if we consider the function representing the height of a hill, the gradient of the function at a particular point gives the direction of the steepest ascent or descent. At the top of the hill, the gradient is zero because the hill has

reached its highest point and there is no direction in which the height increases. Similarly, at the bottom of a valley, the gradient is zero because there is no direction in which the height decreases.

To determine the nature of a critical point, we need to analyze the Hessian matrix, which is a matrix of second partial derivatives of the function with respect to its input variables. The Hessian matrix provides information about the curvature of the function at a critical point, and it can be used to classify the critical point as a minimum, maximum, or saddle point.

If the Hessian matrix is positive-definite at a critical point, then the critical point is a local minimum. If it is negative-definite, then the critical point is a local maximum. If it has both positive and negative eigenvalues, then the critical point is a saddle point.

If the Hessian matrix is indefinite or has zero eigenvalues, then the nature of the critical point is more complex, and additional analysis is needed.

9.5   Linear Approximation

Linear approximation is an important technique to find values for an arbitrary point near a fixed point along a nonlinear curve. In some cases, it is difficult to compute the number of a function (e.g. $\sqrt[3]{5}$).

Definition 9.3 (2-variable)**.** The linear approximation of $f(x, y)$ about the point (x_0, y_0) is

$$f(x, y) \approx f(x_0, y_0) + f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0)$$

[multivariable] The linear approximation of $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ about $(\mathbf{x}^0) = (x_1^0, \ldots, x_n^0)$ is

$$f(\mathbf{x}) \approx f(\mathbf{x}) + \sum_{i=1}^{n} f'_i(\mathbf{x}^0)(x_i - x_i^0)$$
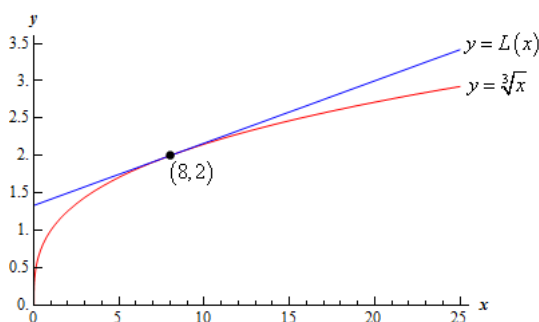
Fig. 9.6. Linear Approximation of a function near the point a:
$L(x) \approx f(a) + f'(a)(x - a)$.

Example 9.4. Find the linear approximation to $f(x, y) = e^{x+y}(xy - 1)$ about $(0, 0)$.

Observe that $f(0,0) = -1$. The partial derivatives are:

$$f'_x = e^{x+y}(xy - 1) + e^{x+y}y,$$
$$f'_y = e^{x+y}(xy - 1) + e^{x+y}x.$$

At the point $(0,0)$, we have $f'_x(0,0) = -1$ and $f'_y(0,0) = -1$. Hence, the linear approximation near $(0,0)$ is

$$L(x,y) \approx -1 + (-1)(x - 0) + (-1)(y - 0) = -1 - x - y$$

Ex. 9.10. Find the linear approximation about $(0,0)$ for the following

$$(a)\ f(x,y) = (x+1)^5(y+1)^6,$$
$$(b)\ f(x,y) = \sqrt{1 + x + y},$$
$$(c)\ f(x,y) = e^x \ln(1 + y)$$

Ex. 9.11. Suppose

$$g(\mu, \varepsilon) = [(1 + \mu)(1 + \varepsilon)^\alpha]^{1/(1-\beta)-1}$$

with α, β as constants. Show that if μ, ε are close to 0, then

$$g(\mu, \varepsilon) \approx (\mu + \alpha\varepsilon)/(1 - \beta)$$

Ex. 9.12. Let

$$f(x,y) = 3x^2y + 2y^3.$$

Given the point $(1, -11)$ where $f(1, -1) = -5$, find the estimated value of $f(0.98, -1, 01)$. How large is the error?

9.6 Implicit Function Theorem

Functions where exogenous variables x can be separated from endogenous variables y are **explicit functions**.

$$y = F(x_1, \ldots, x_n)$$

However, functions where exogenous and endogenous variables are mixed together, are more difficult to solve, and are called **implicit functions**

$$G(x_1, x_2, \ldots, x_n, y) = 0.$$

Still, we want to know **how a small change in one of the exogenous variables affects the value of the endogenous variable**.

Example 9.5. Consider a profit-maximizing firm:

$$\Pi(x) = pf(x) - wx.$$

The choice of x is the solution to:

$$pf'(x) - w = 0.$$

We can think of p and w as exogenous variables. Under some class of model, $f'(x)$ can be explicitly defined as a function of p, w. However, it can be defined as an implicit function as well.

Yet, we still want to know how the optimal choice of input x changes as p or w increases. If there are multiple solutions of x given p, w, we can choose among the candidates by using second-order conditions for a maximum or looking for the global maximizer.

If a function is explicit, finding the derivatives is easy. What if a function is implicit? Suppose that

$$F(x, y) = c$$
$$y = f(x)$$

If f is differentiable, what is the derivative of $y = f(x)$? If the graph of f looks like the figure below, the problem is to find the slope of the graph at an arbitrary point on the curve like P
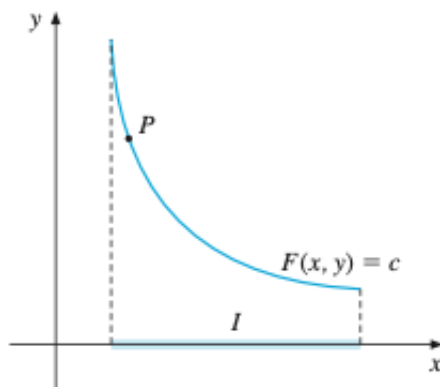
Let function u be defined for all x in I such that

$$u(x) = F(x, f(x))$$

Then, applying the chain rules yields

$$u'(x) = F'_x(x, f(x)) \cdot 1 + F'_y(x, f(x)) \cdot f'(x)$$

Since $u(x) = c$ (constant) for all x in I, its derivative is equal to zero and we have

$$0 = F'_x(x, f(x)) + F'_y(x, f(x)) \cdot f'(x)$$

Rearranging yields

$$y' = f'(x) = -\frac{F'_x(x, y)}{F'_y(x, y)}$$

A condition is that $F'_y \neq 0$.

This is known as the Implicit Function Theorem and is a very important one.

Theorem 9.1. *(Implicit Function Theorem)*
  Let $G(x, y)$ be a C^1 function on a ball around $(x_0, y_0) \in \mathbb{R}^2$. Suppose that $G(x, y) = c$ and consider the expression

$$G(x, y) = c.$$

If $(\partial G/\partial y)(x_0, y_0) \neq 0$, then there exists a C^1 function $y = y(x)$ defined on an interval I about the point x_0 such that:

  1. $G(x, y(x)) \equiv c$ for all x in I.

  2. $y(x_0) = y_0$,

  3. and $y'(x_0) = -\dfrac{\dfrac{\partial G}{\partial x}(x_0, y_0)}{\dfrac{\partial G}{\partial y}(x_0, y_0)}$

Example 9.6. Find $y'(x)$ when $xy = 5$.
  Let $F(x, y) = xy$. Then $F'_x = y, F'_y = x$. For $x \neq 0$, the IFT says that

$$y' = -\frac{F'_x}{F'_y} = -\frac{y}{x}$$

  We can use this theorem for comparative statics, to see the relations between 2 variables that are implicitly defined, and also to find the second derivative.

Ex. 9.13. Suppose

$$e^{ty^2} - 2x - 4y = c$$

implicitly defines y as a differentiable function of x, $y = f(x)$. Find a value of c such that $f(0) = 1$, and find y' at $(x, y) = (0, 1)$.

  More than 2 variables
  Suppose $F(x_1, x_2, \ldots, x_n, y) = c$ implicitly defines $y = f(x_1, \ldots, x_n)$, the IFT says that

$$\frac{\partial y}{\partial x_i} = -\frac{\partial F/\partial x_i}{\partial F/\partial y} \left(\equiv -\frac{F'_{x_i}}{F'_y} \right)$$

  Finding the Second Derivatives
  Suppose we have an implicit function y of x, $y = f(x)$ such that

$$F(x, y) = c$$

Say we want to know the concavity of f, just one needs to verify f' and f''. By the IFT, we obtain

$$y' = -\frac{F'_x}{F'_y} \equiv \frac{g(x, y)}{h(x, y)} \tag{9.6}$$

To find the second-order derivatives, we use the quotient rule to obtain

$$y'' = -\frac{g'(x,y)h(x,y) - g(x,y)h'(x,y)}{h^2(x,y)} \tag{9.7}$$

Note that g, h are composite functions, so the chain rule implies

$$g'(x,y) = \frac{\partial g}{\partial x} = F''_{xx} + F''_{xy}y',$$

$$h'(x,y) = \frac{\partial h}{\partial x} = F''_{yx} + F''_{yy}y',$$

plugging back to (9.7) yields

$$y'' = -\frac{(F''_{xx} + F''_{xy}y')F'_y - (F''_{yx} + F''_{yy}y')F'_x}{F'^2_y}$$

Ex. 9.14. Substituting (9.6), show that

$$y'' = -\frac{1}{F'^3_y}\left(F''_{xx}F'^2_y - 2F''_{xy}F'_xF'_y + F''_{yy}F'^2_x\right)$$

where we apply Young's theorem for $f(x,y)$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

implying that the order of differentiation does not matter so $F''_{xy} = F''_{yx}$

Ex. 9.15. Using the IFT, find y', y'' of

(a) $x^2y = 1$,   (b) $x - y + 3xy = 2$

(c) $y^5 - x^6 = 0$,   (d) $x^3 + y^3 + z^3 - 3z = 0$

For (d), find $z'_x, z'_y, z''_{xx}, z''_{xy}, z''_{yy}$.

Chapter 10

Matrix Algebra

This chapter reviews some basic concepts and operations concerning matrix algebra.

10.1 Matrices and Matrix Operations

A matrix is simply a rectangular array of numbers considered as one mathematical object, denoted by a bold letter. In general, an $m \times n$ matrix is

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

Some special matrices note mentioning (all of them are square matrices of order-n, aka, $n \times n$ matrices).

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \quad \textbf{(identity matrix)}$$

For any square matrix $\mathbf{A}$ of order-n

$$\mathbf{I} \times \mathbf{A} = \mathbf{A} \times \mathbf{I} = \mathbf{A}$$
$$\mathbf{A} \times \mathbf{A}^{-1} = \mathbf{I}.$$

$$\mathbf{O} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} \quad \textbf{(zero matrix)}$$

$$\mathbf{Ł} = \begin{pmatrix} x & 0 & \dots & 0 \\ x & x & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x & x & \dots & x \end{pmatrix} \quad \textbf{(Lower triangle matrix)}$$

$$\mathbf{U} = \begin{pmatrix} x & x & \dots & x \\ 0 & x & \dots & x \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x \end{pmatrix} \quad \textbf{(Upper triangle matrix)}$$

Definition 10.1 (Rules of Addition and Scalar Multiplication). Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ be $m \times n$ matrices, α, β be real numbers. Then:

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$$
$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$
$$\mathbf{A} + \mathbf{O} = \mathbf{A}$$
$$\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$$
$$(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$$
$$\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$$

Ex. 10.1. Evaluate $\mathbf{A} + \mathbf{B}, \mathbf{A} - \mathbf{B}, 5\mathbf{A} - 3\mathbf{B}$ where

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & -1 \\ 2 & 3 & 7 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & -1 & 5 \\ 0 & 1 & 9 \end{pmatrix}$$

10.2   Matrix Multiplication

Definition 10.2. Let $\mathbf{A} = (a_{ij})_{m \times n}, \mathbf{B} = (b_{ij})_{n \times p}$, then the product of $\mathbf{AB}$ is a matrix $\mathbf{C} = (c_{ij})_{m \times p}$ where

$$\underbrace{\mathbf{A}}_{(m \times n)} \times \underbrace{\mathbf{B}}_{(n \times p)} = \underbrace{\mathbf{C}}_{(m \times p)}$$

such that

$$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{jr}$$

of the ith row of $\mathbf{A}$ and the jth row of $\mathbf{B}$.

Example 10.1.

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 3 & 1 \\ 4 & -1 & 6 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 3 & 2 \\ 1 & 0 \\ -1 & 1 \end{pmatrix}$$

1. Compute **AB**. Result:

$$\mathbf{C} = \begin{pmatrix} -1 & 2 \\ 8 & 5 \\ 5 & 14 \end{pmatrix}$$

2. The product **BA** is undefined because of the mismatch in rows – columns.

Ex. 10.2. Compute the multiplication of these matrices

$$(a) \begin{pmatrix} 0 & -2 \\ 3 & 1 \end{pmatrix} \& \begin{pmatrix} -1 & 4 \\ 1 & 5 \end{pmatrix}$$

$$(b) \begin{pmatrix} 8 & 3 & -2 \\ 1 & 0 & 4 \end{pmatrix} \& \begin{pmatrix} 2 & -2 \\ 4 & 3 \\ 1 & -5 \end{pmatrix}$$

$$(c) \begin{pmatrix} -1 & 0 \\ 2 & 4 \end{pmatrix} \& \begin{pmatrix} 3 & 1 \\ -1 & 1 \\ 0 & 2 \end{pmatrix}$$

$$(d) \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix} \& \begin{pmatrix} 0 & -2 & 3 \end{pmatrix}$$

Ex. 10.3. Compute $\mathbf{A} + \mathbf{B}$, $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{A}(\mathbf{BC})$, $(\mathbf{AB})\mathbf{C}$, given that

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -3 \\ 5 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 3 & -1 & 2 \\ 4 & 2 & 5 \\ 2 & 0 & 3 \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 4 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & -2 & 3 \end{pmatrix},$$

Ex. 10.4. Compute $\mathbf{A} + \mathbf{B}$, $\mathbf{A} - \mathbf{B}$, $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{A}(\mathbf{BC})$, $(\mathbf{AB})\mathbf{C}$, given that

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & -2 \\ 3 & 4 & 5 \\ -6 & 7 & 15 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 0 & -5 & 3 \\ 5 & 2 & -1 \\ -4 & 2 & 0 \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 6 & -2 & -3 \\ 2 & 0 & 1 \\ 0 & 5 & 7 \end{pmatrix},$$

10.3   Transpose

A transpose $\mathbf{A}^T$ of a $k \times n$ matrix is the $n \times k$ matrix obtained by interchanging rows and columns of $\mathbf{A}$. Notation: $\mathbf{A}^T$ or $\mathbf{A}'$.

Ex. 10.5. If

$$\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 2 & 3 \\ 5 & -1 \end{pmatrix}$$

Its transpose is

$$\mathbf{A} = \begin{pmatrix} -1 & 2 & 5 \\ 0 & 3 & -1 \end{pmatrix}$$

Definition 10.3 (Rules for Transposition)**.**

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T,$$
$$(\mathbf{A} - \mathbf{B})^T = \mathbf{A}^T - \mathbf{B}^T,$$
$$(\mathbf{A}^T)^T = \mathbf{A},$$
$$(r\mathbf{A})^T = r\mathbf{A}^T,$$
$$(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T.$$

A matrix is **symmetric** iff $\mathbf{A} = \mathbf{A}'$

Ex. 10.6. Show that these 2 matrices are symmetric

$$A = \begin{pmatrix} 3 & 2 & -3 \\ 2 & -1 & 1 \\ 3 & 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 4 & 8 \\ 4 & 0 & 13 \\ 8 & 13 & 0 \end{pmatrix}$$

Ex. 10.7. Determine the values of a so that the following matrix is symmetric

$$\begin{pmatrix} a & a^2 - 1 & -3 \\ a + 1 & 2 & a^2 + 4 \\ -3 & 4a & -1 \end{pmatrix} \tag{10.1}$$

Ex. 10.8. Find a, b, x such that

$$\begin{pmatrix} a & b \\ x & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ x & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 4 \end{pmatrix}$$

Ex. 10.9. We have

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{pmatrix}$$

Compute

$$(a)\ (\mathbf{A} - \mathbf{B})'$$
$$(b)\ (\mathbf{C}'\mathbf{A}')\mathbf{B}'$$
$$(c)\ \mathbf{C}'(\mathbf{A}'\mathbf{B}')$$
$$(d)\ \mathbf{D}'\mathbf{D}$$

10.4 Determinants

10.4.1 Determinants of order 2

We have the matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then its determinant is a number such that

$$|\mathbf{A}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb$$

Geometrically, the determinant of a matrix shows the factor by which areas are scaled by this matrix (during some linear transformations). The signs of the determinant show the direction. If it's negative, it means the space reversed its orientation. If a matrix has a determinant of 0, it's non-invertible. A matrix being non-invertible means that the transformation the matrix represents cannot be undone or reverted. If we only know how determinants are computed and nothing about their geometric meaning, justifying this fact is tough. In contrast, using our freshly established intuition about determinants explaining why this is true becomes not that hard:

Let's say we have a matrix with determinant 0. This means that the matrix scales all areas by a factor of 0, which in turn means that all areas become 0 after the transformation. This can only happen if the matrix squishes the whole space into a lower dimension. For example, the two-dimensional space would be squished into a single line or point and such a transformation cannot be undone.
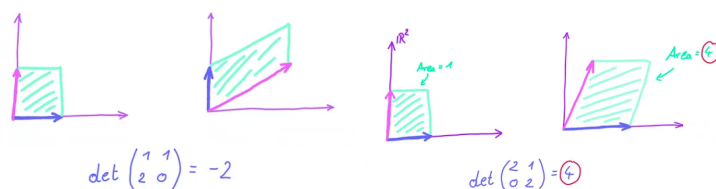
Fig. 10.1. Positive and negative determinant

Ex. 10.10. Evaluate the determinants of

$$(a)\begin{vmatrix} 5 & -2 \\ 3 & -2 \end{vmatrix} \qquad (b)\begin{vmatrix} 1 & a \\ a & 1 \end{vmatrix} \qquad (c)\begin{vmatrix} (a+b)^2 & a-b \\ (a-b)^2 & a+b \end{vmatrix} \qquad (d)\begin{vmatrix} 1-c & 2 \\ 2 & 4-c \end{vmatrix}$$

If a matrix $\mathbf{A}$ has its determinant $|\mathbf{A}| \neq 0$, then $\mathbf{A}$ is called a nonsingular matrix, there is row (column) independence in matrix $\mathbf{A}$.

10.4.2 Determinant of higher order

The formula for the determinant of order -3

$$|\mathbf{A}| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Definition 10.4 (Minor and Cofactor). Let A be a $n \times n$ matrix. Let A_{ij} be the $(n-1) \times (n-1)$ submatrix obtained by deleting row i and column j from A. Then, the scalar

$$M_{ij} = \det(A_{ij})$$

is called the $(i,j)^{th}$ **minor** of A and the scalar

$$C_{ij} = (-1)^{i+j} M_{ij}$$

is called the $(i,j)^{th}$ **cofactor** of A. A cofactor is just a **signed minor**.

Note that $M_{ij} = C_{ij}$ if $(i+j)$ is even and $M_{ij} = -C_{ij}$ if $(i+j)$ is odd.

These definitions give us a general formula to calculate the **determinant**

Definition 10.5. The **determinant** of an $n \times n$ matrix A is given by:

$$
\begin{aligned}
\det A &= a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} \\
&= a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n}.
\end{aligned}
$$

As a result, we can apply to a 3×3 matrix:

Definition 10.6. The determinant of a 3×3 matrix is given by:

$$
\begin{aligned}
|A| = \det &\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\
&= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\
&= a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13} \\
&= a_{11}[a_{22}a_{33} - a_{32}a_{23}] - a_{12}[a_{21}a_{33} - a_{31}a_{23}] + a_{13}[a_{21}a_{32} - a_{31}a_{22}]
\end{aligned}
$$

Theorem 10.1. *The determinant of a lower-triangular, upper-triangular, or diagonal matrix is simply* **the product of its diagonal entries**.

Thus, if we reduce a matrix to its row echelon form, then applying the above theorem helps us calculate the det. of a matrix more easily.

Ex. 10.11. Evaluate the determinants of

(a) $\begin{vmatrix} 8 & 1 & 3 \\ 4 & 0 & 1 \\ 6 & 0 & 3 \end{vmatrix}$, $\quad (b)$ $\begin{vmatrix} 4 & 0 & 2 \\ 6 & 0 & 3 \\ 8 & 2 & 3 \end{vmatrix}$ (c) $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$ $\quad (d)$ $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & 5 \\ 3 & 6 & 9 \end{vmatrix}$ (e) $\begin{vmatrix} 1 & 1 & 4 \\ 8 & 11 & -2 \\ 0 & 4 & 7 \end{vmatrix}$ $\quad (f)$ $\begin{vmatrix} x & 5 & 0 \\ 3 & y & 2 \\ 9 & -1 & 8 \end{vmatrix}$

10.5   System of Linear Equations

For example, given the following system:

$$
\begin{aligned}
x_1 - 0.4x_2 - 0.3x_3 &= 130 \\
-0.2x_1 + 0.88x_2 - 0.14x_3 &= 74 \\
-0.5x_1 - 0.2x_2 + 0.95x_3 &= 95,
\end{aligned}
$$

Normally, you would solve this system of linear equations by successive substitution. However, we can solve it by using matrix operations.

$$
\begin{aligned}
x_1 - \qquad 0.4x_2 - \qquad 0.3x_3 &= 130 \\
0.8x_2 - \qquad 0.2x_3 &= 100 \\
0.7x_3 &= 210.
\end{aligned}
$$

The next step is to apply Gauss-Jordan elimination. After reaching the upper triangle system, multiply each equation by a scalar so that **the first nonzero coefficient is 1**:

$$
\begin{aligned}
x_1 - \quad 0.4x_2 - \quad 0.3x_3 &= 130 \\
x_2 - \quad 0.25x_3 &= 125 \\
x_3 &= 300.
\end{aligned}
$$

Finally, use Gaussian elimination from the bottom to the top. For example, adding 0.25×300 to both sides of equation (c) can eliminate x_3 from the system, leaving x_2 exposed. In equation (a), to eliminate x_2 and x_3, add $0.4 \times x_2 + 0.3 \times x_3$ to both sides. The system should look like this:

$$
\begin{aligned}
x_1 \qquad\qquad &= 300 \\
x_2 \qquad &= 200 \\
x_3 &= 100.
\end{aligned}
$$

The augmented matrix simplifies the notation. To the system in the example, one can reach:

$$
\left(\begin{array}{ccc|c}
1 & -0.4 & -0.3 & 130 \\
0 & 0.8 & -0.2 & 100 \\
0 & 0 & 0.7 & 210
\end{array}\right),
$$

Each row begins with more zeros than the previous row does. Such a matrix is said to be in **row echelon form.**

Some examples:

Row echelon matrices: $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 4 \\ 0 & 1 & 6 \end{pmatrix}, and \begin{pmatrix} 2 & 3 \\ 0 & 6 \\ 0 & 0 \end{pmatrix}$

Not row echelon matrices: $\begin{pmatrix} 1 & 5 & 2 \\ 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 7 \\ 9 & 0 \\ 0 & 2 \end{pmatrix}.$

For the system above, if we can reduce it to:

$$
\left(\begin{array}{ccc|c}
1 & -0.4 & -0.3 & 130 \\
0 & 0.8 & -0.2 & 100 \\
0 & 0 & 0.7 & 210
\end{array}\right)
$$

multiply row 3 with $1/0.7$, row 2 with $1/0.8 \rightarrow$
$\left(\begin{array}{ccc|c}
1 & -0.4 & -0.3 & 130 \\
0 & 1 & -0.25 & 125 \\
0 & 0 & 1 & 300
\end{array}\right)$

add $0.25\times$ row 3 to row 2, add $0.3\times$ row 3 to row 1 $\rightarrow$
$\left(\begin{array}{ccc|c}
1 & -0.4 & 0 & 220 \\
0 & 1 & 0 & 200 \\
0 & 0 & 1 & 300
\end{array}\right)$

add $0.4\times$ row 2 to row 1 $\rightarrow$
$\left(\begin{array}{ccc|c}
1 & 0 & 0 & 300 \\
0 & 1 & 0 & 200 \\
0 & 0 & 1 & 300
\end{array}\right)$

This system is called a **reduced row echelon form**.

Ex. 10.12. Use Gaussian Elimination on the following systems

$$
(a)\ \begin{array}{rrrl} x & +2y & +z & = 4 \\ x & -y & +z & = 5 \\ 2x & +3y & -z & = 1 \end{array} \quad (b)\ \begin{array}{rrrl} 2x & +2y & -z & = 2 \\ x & -3y & +z & = 0 \\ 3x & +4y & -z & = 1 \end{array}
$$

Ex. 10.13. Find all solutions to the following system

$$
\frac{3}{4}x + y + \frac{7}{4}z = b_1
$$
$$
x + 2y + z = b_2
$$
$$
3x + 4y + 7z = b_3
$$

Ex. 10.14. Find c so that the following system has a solution

$$
2x + y + 4z + 3t = 1
$$
$$
x + 3y + 2z - t = 3c
$$
$$
x + y + 2z + 2t = c^2
$$

10.6   Inverse Matrix (self-read)

Summary from Chiang and Wainwright (1984). Given a matrix $\mathbf{A}$, we say $\mathbf{X}$ is an inverse of $\mathbf{A}$ if there exists a matrix $\mathbf{X}$ such that

$$
\mathbf{AX} = \mathbf{XA} = \mathbf{I}.
$$

Then $\mathbf{A}$ is said to be invertible. Because $\mathbf{XA} = \mathbf{AX} = \mathbf{I}$, the matrix $\mathbf{A}$ is also an inverse of $\mathbf{X}$—that is, $\mathbf{A}$ and $\mathbf{X}$ are inverses of each other. Note that the two matrix products $\mathbf{AX}$ and $\mathbf{XA}$ are defined and equal only if $\mathbf{A}$ and $\mathbf{X}$ are square matrices of the same order. Thus, only square matrices can have inverses. But not even all square matrices have inverses.

10.6.1   Find the Inverse

As we know how to calculate the determinant, it's time to apply it to find the inverse of a matrix. Steps:

1. Find det $|A|$ of A

2. Build the co-factor matrix $\mathbf{C}$, then transpose it to get $\mathbf{C}^T$, which is the adjoint matrix of A called $adj\,A$

3. Calculate:

$$
A^{-1} = \frac{1}{|A|} \times adj\,A. \tag{10.2}
$$

Step 1: Determinant $|A|$ (Reprise from Chiang)
For a 2×2 matrix, $\det(A)$ is a scalar such that:

$$
A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow \det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}
$$

- If $\det(A) \neq 0$, then A is nonsingular and thus is invertible.

- If $\det(A) = 0$, then A is singular and thus is not invertible.

For a 3×3 matrix, $\det(A)$ is a scalar such that:

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \Rightarrow \det(A) = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

For n-order det, we must use the Laplace expansion of the determinant.
For the 3-order det:

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{10.3}$$

$|M_{ij}|$ is the minor of the element a_{ij}, obtained by deleting the ¡em¿i¡/em¿-th row and ¡em¿j¡/em¿-th column of a given det.
So:

$$|M_{11}| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

$$|M_{12}| = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

$$|M_{13}| = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

The cofactor of a minor is a minor with a prescribed algebraic sign attached to it.

$$|C_{ij}| = (-1)^{i+j}|M_{ij}| \tag{10.4}$$

In general, the value of a determinant $|A|$ of order n can be found by the Laplace expansion of ANY ROW or ANY COLUMN. In other words, starting row/column doesn't matter.

$$|A| = \sum_{j=1}^{n} a_{ij}|C_{ij}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}| \quad \text{expansion by the i-th row}$$

$$= \sum_{i=1}^{n} a_{ij}|C_{ij}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}| \quad \text{expansion by the j-th column}$$

Step 2: Adjoint of A
The adjoint of a matrix A is the transpose of the cofactor matrix of A.
Consider a $n \times n$ matrix A as follows.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Each element of A has a corresponding cofactor matrix, which can be used to create a cofactor matrix $\mathbf{C} = [|C_{ij}|]$. Specifically:

$$\mathbf{C} = \begin{bmatrix} |C_{11}| & |C_{12}| & \dots & |C_{1n}| \\ |C_{21}| & |C_{22}| & \dots & |C_{2n}| \\ \vdots & \vdots & \ddots & \vdots \\ |C_{n1}| & |C_{n2}| & \dots & |C_{nn}| \end{bmatrix}$$

Transpose this matrix to get the adjoint $adj\,A$ $(n \times n)$

$$adj\,A = \mathbf{C}' = \begin{bmatrix} |C_{11}| & |C_{21}| & \dots & |C_{n1}| \\ |C_{12}| & |C_{22}| & \dots & |C_{n2}| \\ \vdots & \vdots & \ddots & \vdots \\ |C_{1n}| & |C_{2n}| & \dots & |C_{nn}| \end{bmatrix}$$

Step 3: The Inverse formula

To understand how the formula is obtained, premultiply the adjoint with the original matrix.

$$\begin{aligned} A\mathbf{C}' &= \begin{bmatrix} \sum_{j=1}^{n} a_{1j}|C_{1j}| & \sum_{j=1}^{n} a_{1j}|C_{2j}| & \dots & \sum_{j=1}^{n} a_{1j}|C_{nj}| \\ \sum_{j=1}^{n} a_{2j}|C_{1j}| & \sum_{j=1}^{n} a_{2j}|C_{2j}| & \dots & \sum_{j=1}^{n} a_{2j}|C_{nj}| \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} a_{nj}|C_{1j}| & \sum_{j=1}^{n} a_{nj}|C_{2j}| & \dots & \sum_{j=1}^{n} a_{nj}|C_{nj}| \end{bmatrix} \\ &= \begin{bmatrix} |A| & 0 & 0 & \dots & 0 \\ 0 & |A| & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |A| \end{bmatrix} \\ &= |A|I_n \end{aligned} \qquad (10.5)$$

The reason why the non-diagonal elements in AC' result in 0 is that they are the "alien cofactor", namely the cofactor of a wrong row or column.

Consider a 3×3 matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The cofactor of a_{11} is:

$$|C_{11}| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

Other cofactors such as

$$|C_{12}| = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix},$$

$$|C_{13}| = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

are the "alien cofactors" and their determinants are zero. Observe that:

$$\sum_{j=1}^{3} a_{1j}|C_{2j}| = a_{11}|C_{21}| + a_{12}|C_{22}| + a_{13}|C_{23}|$$

$$= -a_{11}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{13}\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= -a_{11}a_{12}a_{33} + a_{11}a_{13}a_{32} + a_{12}a_{11}a_{33} - a_{12}a_{13}a_{31} - a_{13}a_{11}a_{32} + a_{13}a_{12}a_{31}$$

$$= 0.$$

In general

$$\sum_{j=1}^{n} a_{ij}|C_{i'j}| = 0 \quad (i \neq i')$$

$$\sum_{i=1}^{n} a_{ij}|C_{ij'}| = 0 \quad (j \neq j')$$

Back to the result at 10.5. Since $|A| \neq 0$, we divide both sides to $|A|$ to get:

$$\frac{AC'}{|A|} = I.$$

Then premultiply both sides with A^{-1} to get:

$$\frac{C'}{|A|} = A^{-1}$$

or equivalently:

$$A^{-1} = \frac{1}{|A|}adj\,A$$

Cramer's Rule

Given the system $A\mathbf{x} = \mathbf{b}$, let x^* be the solution and can be written as:

$$x^* = A^{-1}\mathbf{b} = \frac{1}{|A|}adj\,A\mathbf{b}.$$

Assume A is nonsingular. Using the adjoint formula, we obtain:

$$x^* = \begin{bmatrix} x^*_1 \\ x^*_2 \\ \vdots \\ x^*_n \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} |C_{11}| & |C_{21}| & \dots & |C_{n1}| \\ |C_{12}| & |C_{22}| & \dots & |C_{n2}| \\ \vdots & \vdots & \ddots & \vdots \\ |C_{1n}| & |C_{2n}| & \dots & |C_{nn}| \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$= \frac{1}{|A|}\begin{bmatrix} b_1|C_{11}| + b_2|C_{21}| + \dots + b_n|C_{n1}| \\ b_1|C_{12}| + b_2|C_{22}| + \dots + b_n|C_{n2}| \\ \dots \\ b_1|C_{1n}| + b_2|C_{2n}| + \dots + b_n|C_{nn}| \end{bmatrix}$$

$$= \frac{1}{|A|}\begin{bmatrix} \sum_{i=1}^{n} b_i|C_{i1}| \\ \sum_{i=1}^{n} b_i|C_{i2}| \\ \vdots \sum_{i=1}^{n} b_i|C_{in}| \end{bmatrix},$$

so that, for instance, the first and second components of $\mathbf{x}$ can be expressed as:

$$x^*{}_1 = \frac{1}{|A|} \sum_{i=1}^{n} b_i |C_{i1}|$$

$$x^*{}_2 = \frac{1}{|A|} \sum_{i=1}^{n} b_i |C_{i2}|,$$

Therefore:

$$x^*{}_1 = \frac{1}{|A|} |A_1|$$

where $|A_1|$ is the Laplace expansion but the first column has been replaced by b. In other words, we are replacing the first column of A with the column vector $\mathbf{b}$ but keeping all other columns intact. In a more general form:

$$x^*{}_j = \frac{|A_j|}{|A|} = \frac{1}{|A|} \underbrace{\begin{vmatrix} a_{11} & a_{12} & \dots & b_1 & \dots a_{1n} \\ a_{21} & a_{22} & \dots & b_2 & \dots a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & b_n & \dots a_{nn} \end{vmatrix}}_{j^{th} \text{column replaced by } \mathbf{b}}$$

Note that the matrix inversion method yields the solution values of **all** the endogenous variables at one (x^* as a vector), while the Cramer's Rule can give us the solution value of **only a single** endogenous variable at a time ($x^*{}_j$ is a scalar).

Theorem 10.1. *Let A be a nonsingular matrix. Then,*

a) *Its **inverse** is:*

$$A^{-1} = \frac{1}{\det A} adj A$$

where
$$adj A = \mathbf{C}' , \quad \mathbf{C} = [\, |C_{ij}| \,]$$

b) *(**Cramer's Rule**) The unique solution $\mathbf{x} = (x_1, \dots, x_n)$ of the $n \times n$ system $A\mathbf{x} = \mathbf{b}$ is:*

$$x_i = \frac{\det B_i}{\det A}, \quad for\, i = 1, \dots, n,$$

where B_i is the matrix A with the right hand side $\mathbf{b}$ replacing the ith column of A.

Chapter 11

Unconstrained Optimization

11.1 Two Choice Variables: Necessary Conditions

> **Theorem 11.1** (First-order necessary conditions). *A differentiable function $z = f(x, y)$ can have a max or min at an interior point (x_0, y_0) of its domain only if it is a critical point – that is, if the point (x_0, y_0) satisfies the two equations*
>
> $$f'_x(x, y) = 0, \ \& f'_y(x, y) = 0.$$
>
> *These are known as first-order conditions (FOCs).*

Example 11.1 (Profit-Maximization)**.** Suppose the production function F is twice differentiable

$$Q = F(K, L),$$

where K, L are capital and labor. The input price is p, the capital rental cost is r, and the wage rate is w such that $p, r, w > 0$. The firm chooses the optimal input factors (K, L) to maximize the following profit

$$\max_{K,L} \pi = pF(K, L) - rK - wL$$

The FOCs are:

$$\pi'_K = pF'_K - r = 0,$$
$$\pi'_L = pF'_L - w = 0.$$

which implies the optimal choices of K^*, L^* as

$$F'_K(K^*, L^*) = r/p,$$
$$F'_L(K^*, L^*) = w/p.$$

This is a system of 2 equations of 2 unknowns, so there exists a unique solution for K^*, L^*.

Ex. 11.1 (Profit Maximizing Firm)**.** Solve the profit maximization problem for a firm with the following production function

$$Q = F(K, L) = 12K^{1/2}L^{1/4}$$

with $p = 1, w = 0.6, r = 1.2$.
    [Answers: $K = L = 625$.]

Ex. 11.2 (Discriminating Monopolist)**.** A monopolist producing a SINGLE output has 2 types of customers. It produces Q_1 for customers of type 1, who are willing to pay with the price of $p_1 = 50 - 5Q_1$. The monopolist produces Q_2 for customers of type 2 under the price $p_2 = 100 - 10Q_2$. The cost of manufacturing is $90 + 20Q$. How should the monopolist produce for each market?

11.2   Two Choice Variables: Sufficient Conditions

The sufficient condition ensures that the optimal values found are indeed the maximizers or minimizers. In 1-variable calculus, you need to test the second-order derivatives. If $f''(x) \leq 0$, then the function is concave and your critical point is a MAX. For a 2-variable function, we need to rely on the Hessian matrix to test its concavity. The premise is still the same, the solution is indeed a MAX if the function is concave down, and MIN if the function is concave up.

Theorem 11.1. *Let f be a C^2 function defined in a convex set S in $\mathbb{R}^2$. Suppose (x^*, y^*) is an interior critical point, then the Hessian matrix of f at this critical point is*

$$D^2 F(x^*, y^*) = \begin{pmatrix} \dfrac{\partial^2 F}{\partial x^2}(x^*, y^*) & \dfrac{\partial^2 F}{\partial y \partial x}(x^*, y^*) \\ \dfrac{\partial^2 F}{\partial x \partial y}(x^*, y^*) & \dfrac{\partial^2 F}{\partial y^2}(x^*, y^*) \end{pmatrix}$$

shorthand

$$D^2 F(x^*, y^*) = \begin{pmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{pmatrix}$$

- *If the leading principal minors alternate in sign such that*

$$|f''_{xx}| = f''_{xx} \leq 0, \ \& \ |f''_{yy}| = f''_{yy} \leq 0$$

$$\begin{vmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{vmatrix} = f''_{xx}f''_{yy} - f''_{yx}f''_{xy} \geq 0$$

at the point (x^*, y^*), then the Hessian is negative definite, and (x^*, y^*) is a MAXIMUM point of f in S.

- *If the leading principal minors are all positive such that*

$$|f''_{xx}| = f''_{xx} \geq 0, , \ \& \ |f''_{yy}| = f''_{yy} \leq 0$$

$$\begin{vmatrix} f''_{xx} & f''_{yx} \\ f''_{xy} & f''_{yy} \end{vmatrix} = f''_{xx}f''_{yy} - f''_{yx}f''_{xy} \geq 0$$

at the point (x^*, y^*), then the Hessian is negative definite, and (x^*, y^*) is a MINIMUM point of f in S.

Example 11.2. Show that we have indeed found a maximum in Ex.11.1.
   For $K, L > 0$, We find that

$$\pi''_{KK} < 0, \qquad\qquad \pi''_{LL} < 0,$$
$$\det(D^2\pi) > 0$$

So the critical point is indeed the maximizer.

Ex. 11.3. A firm producing 3 items x, y, z has the following cost function:

$$C_1(x) = 200 + \frac{1}{100}x^2,$$

$$C_2(y) = 200 + y + \frac{1}{300}y^3,$$

$$C_3(z) = 200 + 10z,$$

$$x + y + z = 2000.$$

1. Find the optimal values of x, y, z to minimize cost

$$\min_{x,y,z} C(x, y, z) = C_1(x) + C_2(y) + C_3(z)$$

2. Show that the values are indeed optimal.

Ex. 11.4. A firm producing 2 items, x units of A and y units of B has the following cost function:

$$C(x, y) = 2x^2 - 4xy + 4y^2 - 40x - 20y + 514$$

The unit price of A is \$24 and B is \$12.

1. Find the optimal values of x, y to maximize profit.

2. The firm is required to produce exactly 54 units of the combined products. What is the optimal production now?

Ex. 11.5. Profit function

$$\pi(x, y) = px + qy - \alpha x^2 - \beta y^2$$

Find x^*, y^* to maximize profits. Verify using the second-order conditions.

11.3 General Case

Say we want to find the solutions of n choice variables ($\mathbf{x} = (x_1, \ldots, x_n)$)

$$\max_{\mathbf{x}} F(\mathbf{x})$$

11.3.1   Necessary Conditions

This condition requires that the solution $\mathbf{x}^*$ must be a critical point of f, that is $f'(\mathbf{x}^*) = 0$. $\mathbf{x}^*$ will not be the endpoint of the interval under consideration, which means it lies in the INTERIOR of the domain of f.

Theorem 11.1. *Let $F : U \mapsto \mathbb{R}^1$ be a C^1 function defined on a subset U of $\mathbb{R}^n$. If $\mathbf{x}^*$ is a local max or min of F in U, and if $\mathbf{x}^*$ is an interior point of U, then*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \ \text{for } i = 1, \ldots, n$$

Basically, the FOCs to every choice variable must be 0.

11.3.2   Sufficient Conditions

We need to use a condition on the second derivatives of F to determine whether the critical point is a max or a min. A C^2 function of n variables has n^2 second-order partial derivatives at each point in its domain. We combine them into a $n \times n$ matrix called the **Hessian** of F

$$D^2 F(\mathbf{x}^*) = \begin{pmatrix} \dfrac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \cdots & \dfrac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}^*) & \cdots & \dfrac{\partial^2 F}{\partial x_n^2}(\mathbf{x}^*) \end{pmatrix}$$

The Hessian is always a symmetric matrix. Whether the critical point is a min or max depends on the definiteness of the Hessian matrix at that point.

Theorem 11.2. *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^n$. Suppose that $\mathbf{x}^*$ is a critical point of F, then*

1. *If the Hessian $D^2 F(\mathbf{x}^*)$ is a NEGATIVE DEFINITE symmetric matrix, then $\mathbf{x}^*$ is a strict LOCAL MAX of F,*

2. *If the Hessian $D^2 F(\mathbf{x}^*)$ is a POSITIVE DEFINITE symmetric matrix, then $\mathbf{x}^*$ is a strict LOCAL MIN of F,*

3. *If the Hessian $D^2 F(\mathbf{x}^*)$ is INDEFINITE, then $\mathbf{x}^*$ is neither a local max nor a local min of F.*

How to test for definiteness of the Hessian?

Now, to determine the definiteness of the Hessian matrix, we rely on the following theorems.

Theorem 11.3 (Sufficient Conditions for a MAX)**.** *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^1$. Suppose that*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \ \text{for } i = 1, \ldots, n$$

and that the n leading principal minors of $D^2F(\mathbf{x}^*)$ <u>*alternate in sign*</u>

$$|F_{x_1\ x_1}| < 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} \\ F_{x_2\ x_1} & F_{x_2\ x_2} \end{vmatrix} > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} & F_{x_3\ x_1} \\ F_{x_2\ x_1} & F_{x_2\ x_2} & F_{x_3\ x_2} \\ F_{x_3\ x_1} & F_{x_2\ x_3} & F_{x_3\ x_3} \end{vmatrix} < 0, \dots$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local max of F.

Theorem 11.4 (Sufficient Conditions for a MIN). *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^1$. Suppose that*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \ for \ i = 1, \dots, n$$

and that the n leading principal minors of $D^2F(\mathbf{x}^*)$ <u>*all positive*</u>

$$|F_{x_1\ x_1}| > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} \\ F_{x_2\ x_1} & F_{x_2\ x_2} \end{vmatrix} > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} & F_{x_3\ x_1} \\ F_{x_2\ x_1} & F_{x_2\ x_2} & F_{x_3\ x_2} \\ F_{x_3\ x_1} & F_{x_2\ x_3} & F_{x_3\ x_3} \end{vmatrix} > 0, \dots$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local min of F.

Fig. 11.1. Principal Minors

11.3.3   Sufficient Conditions: Eigenvalues

Another way to test for concavity using the Hessian matrix is to evaluate the Eigenvalues of the Hessian at critical points. If the Hessian at a given point has **all positive eigenvalues**, it is said to be positive-definite, meaning the function is **concave up (convex)** at that point. If all the **eigenvalues are negative**, it is said to be a negative-definite, equivalent to **concave down**.

When a random matrix $\mathbf{A}$ acts as a scalar multiplier on a vector $\mathbf{x}$, then that vector is called an eigenvector of $\mathbf{x}$. The value of the multiplier is known as an eigenvalue. For the purpose of analyzing Hessians, the eigenvectors are not important, but the eigenvalues are. Because the Hessian of an equation is a square matrix, its eigenvalues can be found. Because Hessians are also sym-

metric [1] (the original and the transpose are the same (by Young's theorem)), they have a special property that their eigenvalues will always be real numbers.

Steps in finding the Eigenvectors and Eigenvalues

$$\mathbf{Ax} = \lambda\mathbf{x}$$
$$\Rightarrow (\mathbf{A} - \lambda)\mathbf{x} = 0$$
$$\Rightarrow (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$$
$$\iff \det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

Solving this expression gives us the Eigenvalues λ and the eigenvector $\mathbf{x}$.

Checks when you found the eigenvalues

1. (trace) the sum of all the eigenvalues will be the sum of the diagonal of $\mathbf{A}$

2. (determinants) the product of all the eigenvalues is the determinant

The eigenvectors tell you the directions that do not change during some linear transformation, while the eigenvalues tell you the scaling vector of these eigenvectors. As learned earlier, the gradient of a multivariable function can be regarded as the first derivative of a function, and the Hessian matrix is the second derivative. If the eigenvalues of the Hessian matrix are **positive** at the critical point we found, it means that the gradient of a function at that point is increasing for any point outside the critical point, so the critical point should be a local MIN. Similarly, if the eigenvalues of the Hessian matrix are negative, then the critical point attains a local MAX. If the eigenvalues are mixed with 1 positive, and 1 negative, then we have a saddle point. The function is concave up in one direction and concave down in the other, about the critical point. If the eigenvalues are zero, that means the determinant of the Hessian is zero, and we cannot conclude the nature of the critical point.

(A) positive defnite (min)   **(B) negative definite (max)**   **(C) (inconclusive) saddle point**
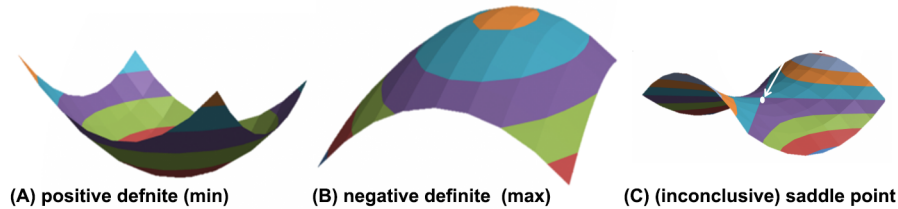
Fig. 11.2. The definiteness of the Hessian

Ex. 11.6. Find the eigenvalues of

$$(a)\begin{pmatrix} 3 & 8 \\ 4 & -1 \end{pmatrix}, \quad (b)\begin{pmatrix} 2 & 6 \\ -1 & 3 \end{pmatrix}$$

---

[1] In fact, we can write the Hessian in 2 ways and they are the same:

$$D^2 F(\mathbf{x}^*) = \begin{pmatrix} \dfrac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \cdots & \dfrac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}^*) & \cdots & \dfrac{\partial^2 F}{\partial x_n^2}(\mathbf{x}^*) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \cdots & \dfrac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}^*) & \cdots & \dfrac{\partial^2 F}{\partial x_n^2}(\mathbf{x}^*) \end{pmatrix}$$

Example 11.3. Suppose $\mathbf{A}$ is a square 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then finding the eigenvalues is to solve

$$\left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = \left| \begin{matrix} a - \lambda & b \\ c & d - \lambda \end{matrix} \right| = 0$$

And the eigenvalues λ are the solutions of

$$(a - \lambda)(d - \lambda) - bc = 0$$

Find the eigenvalues of

$$\begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$$

We need to solve

$$\left| \begin{matrix} 2 - \lambda & 3 \\ 5 & 4 - \lambda \end{matrix} \right| = 0$$

which yields $\lambda = 7$ or -1. These are the eigenvalues. We can do a quick check.

1. (trace) $\lambda_1 + \lambda_2 = 6 = 2 + 4$ (diagonal of $\mathbf{A}$)

2. (determinants) $\lambda_1 \times \lambda_2 = -7 = \det \mathbf{A} (= 2 \times 4 - 5 \times 3)$.

Ex. 11.7. Find the eigenvalues and evaluate them at given points, and determine whether the matrix is negative-definite, positive-definite, or indefinite.

$$(a) \begin{pmatrix} 12x^2 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{at } (3, 1),$$

$$(b) \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix} \quad \text{at } (-1, 2),$$

$$(c) \begin{pmatrix} -2y^2 & -4xy \\ -4xy & -2x^2, \end{pmatrix} \quad \text{at } (1, -1) \text{ and } (1, 0)$$

Ex. 11.8. Determine the concavity of

$$f(x, y) = x^3 + 2y^3 - xy$$

at the following points
$(0, 0), \ (3, 3), \ (3, -3), \ (-3, 3), \ (-3, -3)$.

Ex. 11.9. Find the critical points of the following functions, and use the eigenvalues calculations to determine the Hessian to determine whether the point is a min, max, or neither

$$(a) \ f(x, y) = 4x + 2y - x^2 - 3y^2,$$

$$(b) \ f(x, y) = x^4 + y^2 - xy.$$

11.4　Local Extreme Points

A (strict) local minimum point is defined in an obvious way, and it should also be clear what we mean by local maximum and minimum values, local extreme points, and local extreme values. Note how these definitions imply that a global extreme point is also a local extreme point; the converse is not true, of course.

These first-order conditions are necessary for a differentiable function to have a local extreme point. However, a critical point does not have to be a local extreme point. A critical point (x_0, y_0) of f which, like point R in Fig. 11.3 is neither a local maximum nor a local minimum point, is called a saddle point of f. Hence: A saddle point (x_0, y_0) is a critical point with the property that there exist points (x, y) arbitrarily close to (x_0, y_0) with $f(x, y) < f(x_0, y_0)$, and there also exist such points with $f(x, y) > f(x_0, y_0)$.
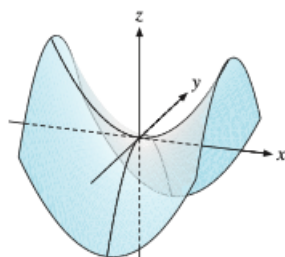
Figure 13.3.1 $z = x^2 - y^2$, with saddle point at $(0, 0)$
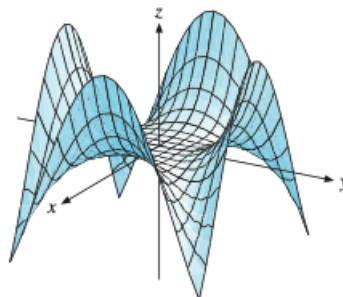
Figure 13.3.2 $z = x^4 - 3x^2y^2 + y^4$, with saddle point at $(0, 0)$

Fig. 11.3. Saddle Points

The following theorem can be used to determine the nature of the critical points in most cases.

Theorem 11.1 (Second-order derivatives Test for Local Extrama). *Suppose that $f(x, y)$ is a C^2 function in a domain S, and let (x_0, y_0) be an interior critical point of S, the Hessian is*

$$H = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix}$$

and its determinant is $\det(H) = f''_{xx} f''_{yy} - f''^2_{xy}$, then

　1. *if $f''_{xx} < 0$, and $\det(H) > 0$, then (x_0, y_0) is a strict local MAX.*

　2. *if $f''_{xx} > 0$, and $\det(H) > 0$, then (x_0, y_0) is a strict local MIN.*

　3. *if H violates the previous 2, then (x_0, y_0) is a saddle point.*

　4. *if $\det(H) = 0$, then (x_0, y_0) could be a local max, local min, or a saddle point.*

Ex. 11.10. A function $f(x, y) = x^2 + 2xy^2 + 2y^2$.

1. Find its partial derivatives of first and second order.

2. Show that its critical points are $(0,0), (-1,1), (-1,-1)$, classify them.

Figure 13.5.1 Open and closed sets

Fig. 11.4. Sets' Nature.

The analysis of local extreme points brings about the topic of a true extreme value. While we can show a point is a local MAX, there is no guarantee that it is truly the MAX of a given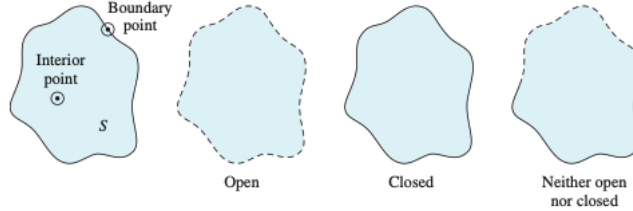 set S. Suppose the set S is bounded and closed, then we also need to evaluate the boundary points of S, and compare its value to the critical points to find the maximum/minimum on the set.

Theorem 11.2 (Extreme Value Theorem). *Suppose the function $f(x, y)$ is continuous throughout a **nonempty, closed, and bounded** set S in the plane. Then there exists both a point (a, b) in S where f has a minimum and a point (c, d) in S where it has a maximum—that is,*

$$f(a, b) \leq f(x, y) \leq f(c, d) \qquad (11.1)$$

for all (x, y) in S.

This theorem helps prove the existence of an extreme value. Although we don't go so deep to prove the closedness/ boundedness of a set, it is still good to know about the theorem.

11.5 Comparative Statics & Envelope Theorem

Optimization problems in economics usually involve maximizing or minimizing functions that depend not only on endogenous variables one can choose but also on one or more exogenous parameters like prices, tax rates, income levels, etc. Although these parameters are held constant during the optimization, they vary according to the economic situation. For example, we may calculate a firm's profit-maximizing input and output quantities while treating the prices it faces as parameters. But then we may want to know how the optimal quantities respond to changes in those prices, or in whatever other exogenous parameters affect the problem we are considering.

Consider first the following simple problem. A function f depends on a single variable x as well as on a single parameter r. We wish to maximize $f(x, r)$ w.r.t. x while keeping r constant

$$\max_x f(x, r)$$

The value of x that maximizes f will usually depend on r, so we denote it by $x^*(r)$. Inserting $x^*(r)$ into $f(x, r)$, we obtain the **value function**:

$$f^*(r) = f(x^*(f), r)$$

What happens to the value function as r changes? Assuming that $f^*(r)$ is differentiable, the chain rule yields

$$\frac{df^*(r)}{dr} = \underbrace{f'_x(x^*)\frac{dx}{dr}}_{=0 \text{ in FOC}} + f'_r(x^*) \tag{11.2}$$

If f achieves a maximum at an interior point $x^*(r)$ in the domain of variation for x, then the FOC $f'_x(x^*(r), r) = 0$ is satisfied. It follows that

$$\frac{df^*(r)}{dr} = f'_r(x^*) \tag{11.3}$$

Note that when r is changed, then $f^*(r)$ changes for 2 reasons. First, a change in r changes the value of f^* directly because r is the second variable in $f(x, r)$. Second, a change in r changes the value of the function $x^*(r)$, and hence $f^*(x^*(r), r)$ is changed indirectly. (11.3) shows that the total effect is simply found by computing the partial derivative of $f^*(x^*(r), r)$ w.r.t. r, ignoring entirely the indirect effect of the dependence of x^* on r.

On further reflection, however, you may realize that the first-order condition for $x^*(r)$ to maximize f w.r.t. x implies that any small change in x, whether or not it is induced by a small change in r, must have a negligible effect on the value of f. The collection of all the value functions of $f^*(r)$ when r changes constitutes the **Envelope Theorem**.

Theorem 11.1 (Envelope Theorem)**.** *Assume that $f(\mathbf{x}, \mathbf{r})$ is differentiable. If $f^*(\mathbf{r}) = \max_{\mathbf{x}} f(\mathbf{x}, \mathbf{r})$ and if $\mathbf{x}^*(\mathbf{r})$ is the value of $\mathbf{x}$ that maximizes $f(\mathbf{x}, \mathbf{r})$ then*

$$\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \frac{\partial f(\mathbf{x}^*(\mathbf{r}), \mathbf{r})}{\partial r_j}$$

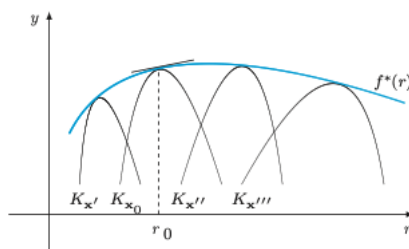*for $j = 1, \ldots, m$ provided that the partial derivative exists.*

Figure 13.7.1 The curve $y = f^*(r)$ is the envelope of all the curves $y = f(\mathbf{x}, r)$

Fig. 11.5. Envelope shape

In Economics, these results are crucial for a comparative statics analysis of the steady state when we change some parameters. It also provides a shortcut.

Example 11.4. Consider the maximization problem

$$\max\ f(x) = -x^2 + 2ax + 4a^2$$

What is the effect of an increase in a on the maximal value of $f(x, a)$?

First, Find the critical point. Taking FOC yields $x^* = a$

1. (Direct Solution) Inserting to the objective function $f(a) = 5a^2$ and so the effect of a on f is $df/da = 10a$.

2. (Envelop Theorem)

$$f'(a) = \frac{\partial f}{\partial a} = 2x + 8a$$

evaluate at $x = a$ also yields $10a$.

11.6 Review Exercises

Borrowed from http://www.columbia.edu/ md3405/Unconstrained_Optimization_Solutions.pdf and https://s3.wp.wsu.edu/uploads/sites/289/2018/09/Problem-set-1-unconstrained-opt-f-2018-ans-key.pdf.

Ex. 11.11. For the following functions, find the critical points (FOC) and determine whether they are the local maximum, local minimum, saddle points, or undetermined using the Second-order conditions.

$(a)\ f(x, y) = -2x^2 + 2xy - y^2 + 18x - 14y + 4$

$(b)\ f(x, y) = 4x^2 - xy + y^2 - x^3,$

$(c)\ f(x, y) = 3x^4 + 3x^2 y - y^3,$

$(d)\ f(x, y) = x^4 + x^2 - 6xy + 3y^2$

$(e)\ f(x, y) = x^2 - 6xy + 2y^2 + 10x + 2y - 5$

$(f)\ f(x, y) = xy^2 + x^3 y - xy$

$(g)\ f(x, y, z) = 2x^2 + y^2 + 4z^2 - x + 2z$

$(h)\ f(x, y, z) = x^2 + 6xy + y^2 - 3yz + 4x^2 - 10x - 5y - 21z.$

Example 11.5. Suppose

$$f(x, y) = x^4 + y^2 - xy$$

The critical point is found by

$$(x) : \frac{\partial f}{\partial x} = 4x^3 - y = 0 \iff y = 4x^3,$$

$$(y) : \frac{\partial f}{\partial y} = 2y - x = 0 \iff y = x/2.$$

Solving for x yields the following critical points

$$(x^*, y^*) = (0, 0),\ (\frac{1}{2\sqrt{2}}, \frac{1}{4\sqrt{2}}),\ (-\frac{1}{2\sqrt{2}}, -\frac{1}{4\sqrt{2}})$$

To verify the extremum, we evaluate the Hessian matrix at the critical points

$$H = \begin{pmatrix} 12x^2 & -1 \\ -1 & 2 \end{pmatrix}$$

The first principal minor is $12x^2$. The second principal minor is H itself, its determinant is

$$|H| = 24x^2 - 1$$

Applying Theorem 11.1, we conclude that $(0,0)$ is a saddle point. The other 2 critical points $(\frac{1}{2\sqrt{2}}, \frac{1}{4\sqrt{2}})$, $(-\frac{1}{2\sqrt{2}}, -\frac{1}{4\sqrt{2}})$ are minima.
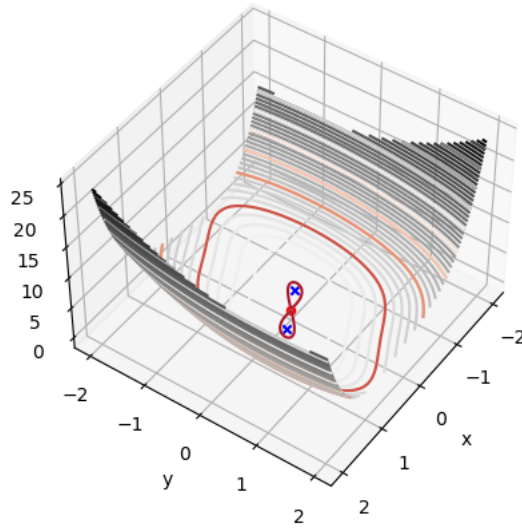
Chapter 12

Constrained Optimization

Let's start with a simple 2 variables optimization problem with 1 equality constraint.

$$\max_{x_1, x_2} f(x_1, x_2)$$
$$\textbf{s.t.} \ \ h(x_1, x_2) = c.$$

12.1 Equality Constraints (Lagrangian)

12.1.1 First-order Necessary Conditions

Geometrically, we have the constraint set C in the $x_1 x_2$-plane and the level curves of the objective function f. The maximizer is the highest level curve of f to touch (be tangent to) C.
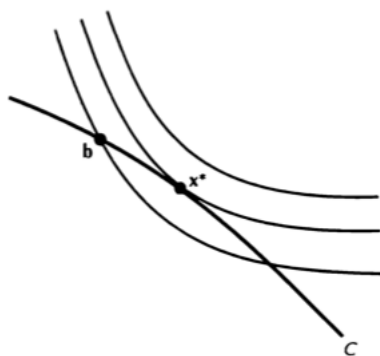
At the constrained max $\mathbf{x}^*$, the highest level curve of f is tangent to the constraint set C.

Fig. 12.1. The highest level curve of f tangent to the constraint set C

The fact that the level curve of f is tangent to the constraint set C at the constrained maximizer $\mathbf{x}^*$ means that the slope of the level set of f equals the slope of the constraint curve C at $\mathbf{x}^*$.

Using the Implicit Function Theorem, the slope of the level set at $\mathbf{x}^*$ is

$$-\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}.$$

The slope of the constraint set at $\mathbf{x}^*$ is

$$-\frac{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}.$$

These two slopes are equal at $\mathbf{x}^*$ implies

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)} = \frac{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}.$$

Rearranging

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}. \tag{12.1}$$

To avoid working with (possibly) zero denominators, let λ denote the common value of the 2 quotients

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)} = \lambda.$$

Rewrite the above as the 2 equations

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) - \lambda \frac{\partial h}{\partial x_1}(\mathbf{x}^*) = 0,$$

$$\frac{\partial f}{\partial x_2}(\mathbf{x}^*) - \lambda \frac{\partial h}{\partial x_2}(\mathbf{x}^*) = 0.$$

To solve for 3 unknowns (x_1, x_2, λ) we need 3 equations. Aside from the above 2, the third equation is the budget constraint.

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) - \lambda \frac{\partial h}{\partial x_1}(\mathbf{x}^*) = 0,$$

$$\frac{\partial f}{\partial x_2}(\mathbf{x}^*) - \lambda \frac{\partial h}{\partial x_2}(\mathbf{x}^*) = 0,$$

$$h(x_1, x_2) - c = 0.$$

The above system can be written in a convenient way, that is the **Lagrangian function**

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda\left[h(x_1, x_2) - c\right]$$

and solve for $\mathbf{x}^*$ by computing $\partial \mathcal{L}/\partial x_1 = \partial \mathcal{L}/\partial x_2 = \partial \mathcal{L}/\partial \lambda = 0$, which gives us exactly the above results.

This technique will not work if both $\partial h/\partial x_1$ and $\partial h/\partial x_2$ were zero at the maximizer $\mathbf{x}^*$ in Eq. (12.1). For this reason, we need to assume that $\partial h/\partial x_1$ or $\partial h/\partial x_2$ (or both) is not zero at the constrained maximizer. This is called a **constraint qualification**. If the constraint is LINEAR, this constraint is automatically satisfied.

Theorem 12.1. *Let f and h be C^1 functions of 2 variables. Suppose that* $\mathbf{x}^* = (x_1^*, x_2^*)$ *is a solution of the problem*

$$\max f(x_1, x_2)$$
$$s.t. \ h(x_1, x_2) = c$$

Suppose further that (x_1^*, x_2^*) is not a critical point of h. Then, there is a real number λ^* such that $(x_1^*, x_2^*, \lambda^*)$ is a critical point of the Lagrangian function

$$\mathcal{L}(x_1, x_2, \lambda) \equiv f(x_1, x_2) - \lambda[h(x_1, x_2) - c].$$

In other words, at $(x_1^*, x_2^*, \lambda^*)$ we can obtain the First-order conditions

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0,$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = 0,$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0,$$

The above theorem provides us with a cookbook method to solve a constrained optimization problem with 2 variables and 1 equality constraint. λ is interpreted as the Lagrangian Multiplier.

Example 12.1. The problem is

$$\max_{x,y} \ f(x, y) = xy$$
$$s.t. \ 2x + y = 100$$

Method 1: Lagrangian Method
The Lagrangian is

$$\mathcal{L} = xy - \lambda(2x + y - 100)$$

The FOCs:

$$\mathcal{L}'_x = y - 2\lambda = 0,$$
$$\mathcal{L}'_y = x - \lambda = 0,$$
$$\mathcal{L}'_\lambda = 2x + y - 100 = 0,$$

which yields the solution $(x^*, y^*) = (25, 50)$.

Method 2: Substitution Method or "Naive" Method
We can turn the constrained optimization problem into an unconstrained problem. From the constraint, we have $y = 100 - x$, the problem becomes

$$\max_x x(100 - 2x)$$

The FOC is

$$100 - 4x = 0,$$

which also gives $x^* = 25, y^* = 50$.

Ex. 12.1. Use the Lagrangian method for these problems
(a) $\max_{x,y} \quad f(x,y) = x^{1/2}y^{1/3} \quad$ s.t. $\quad 2x + 4y = m$
(b) $\max_{x,y} \quad f(x,y) = \log(x) + \beta \log(y) \quad$ s.t. $\quad x + y = c$
(c) $\max_{x,y} \quad f(x,y) = \log(x) + \gamma \log(1-h) \quad$ s.t. $\quad (1-\tau)x + c = hw$

12.1.2 Sufficient Conditions

The Lagrangian method gives the necessary conditions for the local solution of constrained optimization problems. To confirm that we have found the solution, however, a more careful check is needed.

Theorem 12.2. *Suppose (x_0, y_0) is a critical point for the Lagrangian $\mathcal{L}$*

- *If the Lagrangian is concave, then (x_0, y_0) is a MAX*

- *If the Lagrangian is convex, then (x_0, y_0) is a MIN*

To show concave, show that $\mathcal{L}' > 0, \mathcal{L}'' < 0$. To show convex, show that $\mathcal{L}' > 0, \mathcal{L}'' > 0$.

To verify whether the FOCs from Lagrangian form a MAX or MIN, we can also use the second-order sufficient conditions by constructing a bordered Hessian matrix. This is the most common method but cumbersome. We provide here sufficient conditions for a problem of **2 variables and 1 constraint**.

Theorem 12.3. *Let f, h be C^2 functions on $\mathbb{R}^2$. Consider the problem*

$$\max_{x,y} f(x,y) \quad \textbf{s.t.} \ h(x,y) = c \ \text{ for } c \in C_h (\text{constraint set})$$

The Lagrangian is

$$\mathcal{L}(x,y,\lambda) = f(x,y) - \lambda(h(x,y) - c)$$

Suppose that (x^*, y^*, λ^*) satisfies the following FOCs

$$\mathcal{L}'_x = \mathcal{L}'_y = \mathcal{L}'_\lambda = 0 \quad \text{at } (x^*, y^*, \lambda^*)$$

and the bordered Hessian matrix is

$$H = \begin{pmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{pmatrix}$$

if $\det(H) > 0$ at (x^*, y^*), then (x^*, y^*) is the local max of f on C_h.

12.1.3 General Case of n choice variables and k constraints

Move on to the general case. Here, we borrow the lecture note from Tornike Kadeishvili (https://rmi.tsu.ge/ kade/LecturesT.Kadeishvili/MathEconomics/Term4/). For a detailed exposition, see Simon and Blume, 1994 (Chapter 19.3)

Let us consider an optimization problem for n variables with k constraints such that

$$\max_{\mathbf{x}}(\min) f \underbrace{(x_1, \ldots, x_n)}_{\mathbf{x}}$$

$$s.t. \& h_i(\mathbf{x}) = c_i \quad \text{for} \quad i = 1, \ldots, k.$$

There will be k Lagrangian multipliers λ_i for $i = 1, \ldots, k$. The Lagrangian is

$$\mathcal{L}(\mathbf{x}, \lambda_i) = f(\mathbf{x}) - \sum_{i=1}^{k} \lambda_i(h_i(\mathbf{x}) - c_i)$$

As usual, the FOC is just

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0$$

For sufficient conditions, we need to use the notion of Bordered Hessian Matrix. The construction of such a matrix is

$$H = \left(\begin{array}{ccc|ccc} 0 & \ldots & 0 & B_{11} & \ldots & B_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & B_{m1} & \ldots & B_{mn} \\ \hline B_{11} & \ldots & B_{m1} & a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{1n} & \ldots & B_{mn} & a_{1n} & \ldots & a_{nn} \end{array} \right)$$

in short, it looks like this

$$H = \begin{pmatrix} 0 & \mathbf{B} \\ \mathbf{B}^T & \mathbf{L} \end{pmatrix}$$

where $\mathbf{B}$ is the matrix of derivatives of constraints h_i wrt to $\mathbf{x}$, and L is the matrix of second-order derivatives of the Lagrangian $\mathcal{L}$ wrt to $\mathbf{x}$.

In our case, it looks like this

$$H = \left(\begin{array}{ccc|ccc} 0 & \ldots & 0 & \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \frac{\partial h_k}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_n} \\ \hline \frac{\partial h_1}{\partial x_1} & \frac{\partial h_k}{\partial x_1} & & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \cdots & \frac{\partial \mathcal{L}}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2} \end{array} \right)$$

This $(k + n) \times (k + n)$ matrix has $k + n$ leading principal minors (the biggest one is H itself). The first m matrices $H_1, \ldots, H_k$ are zero matrices. The next $k - 1$ matrices $H_{k+1}, \ldots, H_{2k-1}$ have zero determinant.

The determinant of the next minor H_{2k} is $\pm \det(H')^2$ where H' is the upper $k \times k$ minor of H after block of zeros, so $\det H_{2k}$ does not contain information about f.

And only the determinants of the last $n - k$ leading principal minors

$$H_{2k+1}, H_{2k+2}, \ldots, H_{2k+(n-k)=k+n} \equiv H$$

carry information about both, the objective function f and the constraints h_i. Exactly these minors are essential for the following sufficient condition for constraint optimization.

Theorem 12.4 (Second-order Conditions). *Suppose $\mathbf{x}^*$ satisfies the FOCs.*

1. *For the bordered Hessian matrix H, the last $n - k$ leading principal minors*

$$H_{2k+1}, H_{2k+2}, \ldots, H_{2k+(n-k)=k+n} \equiv H$$

evaluated at the critical point ALTERNATE in signs where the last minor $H_{n+k} \equiv H$ has the sign as $(-1)^n$, then $\mathbf{x}^*$ is a LOCAL MAX.

2. *For the bordered Hessian matrix H, the last $n - k$ leading principal minors*

$$H_{2k+1}, H_{2k+2}, \ldots, H_{2k+(n-k)=k+n} \equiv H$$

evaluated at the critical point have the SAME sign where the last minor $H_{n+k} \equiv H$ has the sign as $(-1)^k$, then $\mathbf{x}^*$ is a LOCAL MIN.

---

which can be summarized as

	H_{2k+1}	H_{2k+2}	$\ldots$	H_{k+n-1}	$H_{k+n} \equiv H$
max	$(-1)^{k+1}$	$(-1)^{k+2}$	$\ldots$	$(-1)^{n-1}$	$(-1)^n$
min	$(-1)^k$	$(-1)^k$	$\ldots$	$(-1)^k$	$(-1)^k$

The case of 3 choice variables and 1 constraint

In this case, $n = 3, k = 1$ so the second-order conditions check is to check the last 2 principal minors.

$$H = \begin{pmatrix} 0 & dh/dx & dh/dy & dh/dz \\ dh/dx & d^2L/dx^2 & d^2L/dydx & d^2L/dzdx \\ dh/dy & d^2L/dxdy & d^2L/dy^2 & d^2L/dzdy \\ dh/dz & d^2L/dxdz & d^2L/dydz & d^2L/dz^2 \end{pmatrix}$$

1. For the problem to be a max, the sign of $\det H_4 \equiv \det H$ must be $(-1)^n = (-1)^3 < 0$, and the second-last principal minors $\det H_3$ must > 0.

2. For the problem to be a min, the sign of the last 2 principal minors must be negative as $(-1)^1 < 0$.

Example 12.2. Find extremum of

$$F(x, y) = xy$$

$$s.t.\& \ h(x, y) = x + y = 6.$$

The Lagrangian is

$$L(x, y) = xy - \lambda(x + y - 6)$$

The FOCs are

$$(x) : \frac{\partial L}{\partial x} = y - \lambda = 0,$$

$$(y) : \frac{\partial L}{\partial y} = x - \lambda = 0,$$

$$(\lambda) : \frac{\partial L}{\partial \lambda} = x + y - 6 = 0,$$

which gives

$$x^* = y^* = 3, \ \lambda = 3$$

To tell the nature of its extremum, we test the second-order conditions. The bordered Hessian is

$$H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

We have $n = 2, k = 1$ so we have to check the $n - k = 1$ last leading principal minors, a.k.a, H itself. Calculation shows that $\det H = 2 > 0$ has the same sign with $(-1)^n = (-1)^2 > 0$ so our critical point is a MAX.

Example 12.3. Find extremum of

$$F(x, y, z) = x^2 + y^2 + z^2$$

$$s.t.\& \ h_1(x, y, z) = 3x + y + z = 5,$$

$$h_2(x, y, z) = x + y + z = 1$$

The Lagrangian is

$$L(x, y) = x^2 + y^2 + z^2 - \lambda_1(3x + y + z - 5) - \lambda_2(x + y + z - 1)$$

The FOCs are

$$(x) : \frac{\partial L}{\partial x} = 2x - 3\lambda_1 - \lambda_2 = 0,$$

$$(y) : \frac{\partial L}{\partial y} = 2y - \lambda_1 - \lambda_2 = 0,$$

$$(z) : \frac{\partial L}{\partial z} = 2z - \lambda_1 - \lambda_2 = 0,$$

$$(\lambda_1) : \frac{\partial L}{\partial \lambda_1} = 3x + y + z - 5 = 0,$$

$$(\lambda_2) : \frac{\partial L}{\partial \lambda_2} = x + y + z - 1 = 0$$

which gives

$$x^* = 2, \ y^* = -1/2, \ z^* = -1/2, \ \lambda_1 = 5/2, \ \lambda_2 = -7/2$$

To tell the nature of its extremum, we test the second-order conditions. The bordered Hessian is

$$
H = \left(\begin{array}{cc|ccc}
0 & 0 & 3 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 \\
\hline
3 & 1 & 2 & 0 & 0 \\
1 & 1 & 0 & 2 & 0 \\
1 & 1 & 0 & 0 & 2
\end{array}\right)
$$

We have $n = 3, k = 2$ so we have to check the $n - k = 1$ last leading principal minors, a.k.a, H itself. Calculation shows that $\det H = 16 > 0$ has the same sign with $(-1)^k = (-1)^2 > 0$ so our critical point is a MIN.

12.1.4 Review Exercises

Ex. 12.2. Find the extremum of f wrt h. Verify whether they are a min or max.

$$(a)\ f(x,y) = xy, \qquad\qquad\qquad h(x,y) = x + 2y = 2;$$
$$(b)\ f(x,y) = x(y + 4), \qquad\qquad h(x,y) = x + y = 8;$$
$$(c)\ f(x,y) = x - 3y - xy, \qquad\quad h(x,y) = x + y = 6;$$
$$(d)\ f(x,y) = 7 - y - x^2, \qquad\quad h(x,y) = x + y = 0;$$
$$(e)\ f(x,y,z) = x^2 y^2, \qquad h(x,y,z) = x^2 + y^2 + z^2 = 3;$$

Ex. 12.3. Some problems with 3 choice variables and 2 constraints.

(a) Minimize $x_1^2 + x_2^2 + x_3^2$ subject to $x_1 + x_2 + x_3 = 1$ and $x_1^2 + x_2^2 = 0.25$.

(b) Maximize $x_1 x_2 x_3$ subject to $x_1 + x_2 + x_3 = 1$ and $x_1^2 + x_2^2 = 0.25$.

(c) Minimize $x_1 + x_2 + x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$ and $x_1 + x_2 + x_3 = 2$.

(d) Max $x_1^2 + x_2^2 + x_3^2$ subject to $x_1 + x_2 + x_3 = 1$ and $x_1 x_2 + x_1 x_3 + x_2 x_3 = 0$.

(e) Minimize $x_1 + x_2 + x_3$ subject to $x_1 + x_2 + x_3 = 1$ and $x_1^3 + x_2^3 + x_3^3 = 0.1$.

Ex. 12.4. Some problems with 3 choice variables and 1 constraint.

(a) Minimize $x_1^2 + x_2^2 + x_3^2$ subject to $x_1 + x_2 + x_3 = 1$.

(b) Maximize $x_1 x_2 x_3$ subject to $x_1 + x_2 + x_3 = 10$.

(c) Minimize $x_1 + x_2 + x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$.

(d) Maximize $x_1^2 + x_2^2 + x_3^2$ subject to $x_1 + x_2 + x_3 = 5$.

(e) Minimize $x_1 + x_2 + x_3$ subject to $x_1 x_2 x_3 = 1$.

Ex. 12.5. Some problems with utility functions subject to budget constraints.

(a) Maximize $u(x_1, x_2, x_3) = x_1^{0.5} x_2^{0.3} x_3^{0.2}$ subject to $p_1 x_1 + p_2 x_2 + p_3 x_3 = I$ and $p_1 x_1 + p_2 x_2 = 0.8I$.

(b) Maximize $u(x_1, x_2, x_3) = \frac{x_1^{0.5} x_2^{0.3}}{x_3^{0.2}}$ subject to $p_1 x_1 + p_2 x_2 + p_3 x_3 = I$ and $p_1 x_1 + p_3 x_3 = 0.9I$.

(c) Maximize $u(x_1, x_2, x_3) = \ln(x_1) + 2\ln(x_2) + 3\ln(x_3)$ subject to $p_1 x_1 + p_2 x_2 + p_3 x_3 = I$ and $p_1 x_1 + p_2 x_2 = 0.8I$.

12.2 Inequality Constraints (KKT)

12.2.1 KKT Conditions

In this branch of problems, the constraint has inequality signs.

$$\max f(x,y) \quad s.t. \quad g(x,y) \leq c.$$

We solve this problem by employing the cookbook method called KKT conditions (Karush-Kuhn-Tucker).

Theorem 12.1 (The KKT Conditions). *Suppose we have 2 choice variables and 1 inequality constraint.*

$$\max f(x,y) \ s.t. \ g(x,y) \leq c$$

1. *Construct the Lagrangian*

$$\mathcal{L}(x,y) = f(x,y) - \lambda(g(x,y) - c)$$

2. *FOCs*

$$\mathcal{L}'_x = f'_x - \lambda g'_x = 0,$$
$$\mathcal{L}'_y = f'_y - \lambda g'_y = 0,$$
$$\lambda(g(x,y) - c) = 0,$$
$$\lambda \geq 0,$$
$$g(x,y) \leq c$$

3. *Complimentary slackness condition*

$$\lambda > 0, \ the \ constraint \ binds \ so \ that \ g(x,y) = c$$
$$\lambda = 0, \ the \ constraint \ does \ not \ bind \ so \ that \ g(x,y) < c$$

4. *For a minimum problem, the FOCs are the same, except that $\lambda \leq 0$.*

The two inequalities $\lambda \geq 0$ and $g(x,y) \leq c$ are complementary in the sense that at most one can be "slack" – that is, at most one can hold with inequality. Equivalently, at least one must be an equality. Failure to observe that it is possible to have both $\lambda = 0$ and $g(x,y) = c$ in the complementary slackness condition is the most common error when solving nonlinear programming problems.

Theorem 12.2 (Sufficient Conditions). *Suppose (x_0, y_0) satisfies the KKT conditions, then consider the Lagrangian*

$$\mathcal{L}(x,y) = f(x,y) - \lambda(g(x,y) - c)$$

> If the Lagrangian is concave, then (x_0, y_0) is the MAX. If the Lagrangian is convex, then (x_0, y_0) is the MIN.

Example 12.4. Solve the problem

$$\max f(x, y) = x^2 + y^2 + y - 1,$$
$$s.t. \& g(x, y) = x^2 + y^2 \leq 1$$

The Lagrangian is

$$\mathcal{L} = x^2 + y^2 + y - 1 - \lambda(x^2 + y^2 - 1)$$

The FOCs are

$$\mathcal{L}'_x = 2x - 2\lambda x = 0 \iff 2x(1 - \lambda) = 0,$$
$$\mathcal{L}'_y = 2y + 1 - 2\lambda y = 0 \iff 2y(1 - \lambda) + 1 = 0,$$
$$\lambda(x^2 + y^2 - 1) = 0,$$
$$\lambda \geq 0 \quad (*),$$
$$x^2 + y^2 \leq 1 \quad (**)$$

The last condition implies that

$$\begin{cases} \lambda > 0, x^2 + y^2 = 1 & \text{or} \\ \lambda = 0, x^2 + y^2 < 1 \end{cases} \tag{12.2}$$

(i) Suppose the constraint binds such that

$$x^2 + y^2 = 1$$

From $\mathcal{L}'_x = 0$, that means $x = 0$, and so $y = \pm 1$. Try $y = 1$, then $\mathcal{L}'_y = 0$ implies $\lambda = 3/2$, which satisfies 12.2, so $(0, 1)$ is a candidate for optimality. Next, try $y = -1$ then $\mathcal{L}'_y = 0$ implies $\lambda = 1/2$, which satisfies 12.2, so $(0, -1)$ is another candidate for optimality.

(ii) Suppose the constraint does not bind (or slacks) such that

$$x^2 + y^2 < 1$$

Then, $\lambda = 0$, the FOC $\mathcal{L}'_x = 0$ implies that $x = 0$, and therefore, $-1 < y < 1$. Since $\lambda = 0$, the FOC $\mathcal{L}'_y = 0$ yields $y = -1/2$. Since these values satisfy 12.2, $(0, -1/2)$ is another candidate.

In summary, the 3 candidates for optimality are

$$f(0, 1) = 1, \quad f(0, -1) = -1, \quad f(0, -1/2) = -5/4.$$

Since the function is continuous over a closed, bounded set, by the Extreme Value Theorem, a solution exists. It is easy to see from the value functions that $(x, y) = (0, 1)$ is the solution to maximize f.

Ex. 12.6. Use the KKT conditions to find solutions to

(a) $\max\ -x^2 - y^2\ \ s.t.\ \ x - 3y \le 10$,

(b) $\max\ x^{1/2} + \sqrt{y}\ \ s.t.\ \ px + qy \le m$,

(c) $\max\ 4 - \dfrac{1}{2}x^2 - 4y\ \ s.t.\ \ 6x - 4y \le a$

Write $V(a)$ as the value function, verify that $V'(a) = \lambda$.

(d) $\max\ x^2 + 2y^2 - x\ \ s.t.\ \ x^2 + y^2 \le 1$,

(e) $\max\ f(x, y) = 2 - (x - 1)^2 - e^{y^2}\ \ s.t.\ \ x^2 + y^2 \le 1/2$

Let $f^*(a)$ as the value function, verify that $df^*(a)/da = \lambda$.

12.2.2 Multiple Inequality Constraints

Consider an optimization problem of n choice variables and m inequality constraints

$$\max\ \& f\ \underbrace{(x_1, \ldots, x_n)}_{\mathbf{x}} \tag{12.3}$$

$$s.t.\ \& g_1(\mathbf{x}) \le c_1,$$

$$\ldots,$$

$$g_m(\mathbf{x}) \le c_m$$

Theorem 12.3 (KKT Formulation). *Steps in solving the problem* (12.3)

1. *Write down the Lagrangian*

$$\mathcal{L}(\mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\mathbf{x}) - c_j)$$

2. *FOCs:*

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0,$$

for each $i = 1, \ldots, n$.

3. *Complementary slackness*

$$\lambda_j \ge 0, g_j(\mathbf{x}) = c_j\ \ \boldsymbol{or}\ \ \lambda_j = 0, g_j(\mathbf{x}) < c_j,$$

$$g_j(\mathbf{x}) < c_j$$

for $j = 1, \ldots, m$

> 4. *Find all* $\mathbf{x} = (x_1, \ldots, x_n)$ *associated with their* $\lambda_1, \ldots, \lambda_m$ *that satisfy FOCs and the complementary slackness. These are the solution candidates, and at least 1 of them solves the problem if it has a solution.*

Example 12.5. The problem is

$$\max x + 3y - 4e^{-x-y}$$

$$s.t. \begin{cases} 2 - x \ge 2y \\ x - 1 \le -y \end{cases}$$

Write the problem as

$$\max x + 3y - 4e^{-x-y}$$

$$s.t. \begin{cases} x + 2y \le 2 \\ x - 1 \le -y \end{cases}$$

The Lagrangian is

$$\mathcal{L}(x,y) = x + 3y - 4e^{-x-y} - \lambda_1(x + 2y - 2) - \lambda_2(x + y - 1)$$

KKT conditions

$$(i)\ \mathcal{L}'_x = 1 + 4e^{-x-y} - \lambda_1 - \lambda_2 = 0,$$
$$(ii)\ \mathcal{L}'_y = 3 + 4e^{-x-y} - 2\lambda_1 - \lambda_2 = 0,$$
$$(iii)\ \lambda_1 \ge 0\ (\lambda_1 = 0 \iff x + 2y < 2),$$
$$(iv)\ \lambda_2 \ge 0\ (\lambda_2 = 0 \iff x + y < 1)$$

Since $\mathcal{L}$ is concave (prove this by evaluating the Hessian matrix), the KKT conditions are both necessary and sufficient for optimality.

From (ii), (i), we get $\lambda_1 = 2 > 0$, thus making (iii) binds such that $x + 2y = 2$. Suppose $\lambda_2 = 0$, there is a contradiction. Suppose $\lambda_2 > 0$, from (iv) we deduce $x + y = 1$. Using (i) and (ii), we can find $\lambda_2 = e^{-1}(4 - e) > 0$. Thus, the solution is

$$(x^*, y^*, \lambda_1, \lambda_2) = (0,\ 1, 2\ ,\ e^{-1}.(4 - e))$$

A general method for finding all candidates for optimality in a nonlinear programming problem with two constraints can be formulated as follows: First, examine the case where both constraints bind. Next, examine the two cases where only one constraint binds. Finally, examine the case where neither constraint binds. In each case, find all vectors $\mathbf{x}$, with associated non-negative values of the Lagrange multipliers, that satisfy all the relevant conditions—if any do. Then calculate the value of the objective function for these values of $\mathbf{x}$, and retain that $\mathbf{x}$ with the highest values. Except for perverse problems, this procedure will find the optimum.

12.3 Nonnegativity Constraints

Most oftentimes, in economics, we want to restrict the choice variables to take nonnegative values. Consider the problem

$$\max \& f(x, y)$$
$$s.t. \& g(x, y) \le c,$$
$$x \ge 0,$$
$$y \ge 0$$

Rewrite the problem to

$$\max \& f(x, y)$$
$$s.t. \& g(x, y) \le c,$$
$$-x \le 0,$$
$$-y \le 0$$

The Lagrangian is

$$\mathcal{L}(x, y) = f(x, y) - \lambda[g(x, y) - c] - \mu_1(-x) - \mu_1(-y)$$

The KKT conditions

(i) $\mathcal{L}'_x = f'_x - \lambda g'_x + \mu_1 = 0$
(ii) $\mathcal{L}'_y = f'_y - \lambda g'_y + \mu_2 = 0$
(iii) $\lambda \ge 0$, with $\lambda = 0$ if $g(x, y) < c$
(iv) $\mu_1 \ge 0$, with $\mu_1 = 0$ if $x > 0$
(v) $\mu_2 \ge 0$, with $\mu_2 = 0$ if $y > 0$

Combining (i) and (iv) yields

$$f'_x - \lambda g'_x \le 0, \text{ **with equality if** } x > 0$$

Combining (ii) and (v) yields

$$f'_y - \lambda g'_y \le 0, \text{ **with equality if** } y > 0$$

So the KKT conditions are reduced to just

$$f'_x - \lambda g'_x \le 0, \text{ **with equality if** } x > 0$$
$$f'_y - \lambda g'_y \le 0, \text{ **with equality if** } y > 0$$
$$\lambda \ge 0, \text{ with } \lambda = 0 \text{ if } g(x, y) < c$$

Example 12.6. Consider the utility maximization problem

$$\max \& x + \ln(1 + y)$$
$$s.t. \& px + y \le m, \ x \ge 0, \ y \ge 0.$$

12.4   The Meaning of the Lagrangian Multiplier

Consider again the problem

$$\max(\min) \; f(x,y) \;\; s.t. \;\; g(x,y) = c$$

The FOCs

$$\frac{\partial \mathcal{L}}{\partial x_1} = f'_x - \lambda g'_x = 0 \iff f'_x = \lambda g' x,$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = f'_y - \lambda g'_y = 0 \iff f'_y = \lambda g'_y$$

Suppose that x^*, y^* are the values of x, y that solve the problem. In general, x^* and y^* depend on c, so we write $x^* = x(c)$ and $y^* = y(c)$. We assume that these solutions are differentiable functions of c. The associated value of $f(x,y)$ is then also a function of c, with

$$f^*(c) = f(x^*(c), \; y^*(c))$$

Here $f^*(c)$ is called the (optimal) value function for the problem. Of course, the associated value of the Lagrange multiplier also depends on c, in general, so we write $\lambda(c)$. Taking the total differentiation of the value function yields

$$\begin{aligned}
df^*(c) &= f'_x(x^*, y^*)dx^* + f_y(x^*, y^*)dy^* \\
&= \lambda(c)g'_x(x^*, y^*)dx^* + \lambda(c)g'_y(x^*, y^*)dy^* \\
&= \lambda(c) \underbrace{(g'_x(x^*, y^*)dx^* + g'_y(x^*, y^*)dy^*)}_{dg(x^*, y^*)}
\end{aligned}$$

By identity $g(x,y) = c$ so $dg \equiv dc$, which implies that

$$df^*(c) = \lambda(c) \; dc$$

Thus, we have a remarkable result that

$$\frac{df^*(c)}{dc} = \lambda(c)$$

Thus, the Lagrange multiplier $\lambda = \lambda(c)$ is the rate at which the optimal value of the objective function changes with respect to changes in the constraint constant c. In particular, if dc is a small change in c, then

$$f^*(c + dc) - f^*(c) \approx \lambda(c)dc$$

In economic applications, c often denotes the available stock of some resource, and $f(x,y)$ denotes utility or profit. Then $\lambda(c)dc$ measures the approximate change in utility or profit that can be obtained from dc units more (or $-dc$ unit less when $dc < 0$). That is why economists call λ a **shadow price** of the resource constraint.

12.5 Review Exercises

Ex. 12.7. Solve the following problems with KKT necessary conditions and verify the solutions using sufficient conditions.

1. $\max_{x,y} f(x,y) = xy$ subject to $x^2 + y^2 \le 1$.

2. $\max_{x,y} f(x,y) = x^2 + 3y^2$ subject to $2x + y \le 4$.

3. $\max_{x,y} f(x,y) = x^2 + 2xy + y^2$ subject to $x + y \le 2$.

4. $\max_{x,y} f(x,y) = 2x^2 + y^2$ subject to $x + y \le 3$.

5. $\max_{x,y} f(x,y) = x^3 + 3xy^2$ subject to $x^2 + y^2 \le 4$.

6. $\max_{x,y} f(x,y) = x^2 + 2xy + y^2$ subject to $x + y \ge 2$.

7. $\max_{x,y} f(x,y) = x^3 + 3xy^2$ subject to $x + y \le 3$.

8. $\max_{x,y} f(x,y) = x^2 + 2y^2$ subject to $x + 2y \le 3$.

9. $\max_{x,y} f(x,y) = x^3 + 2xy + y^3$ subject to $x + y \le 2$.

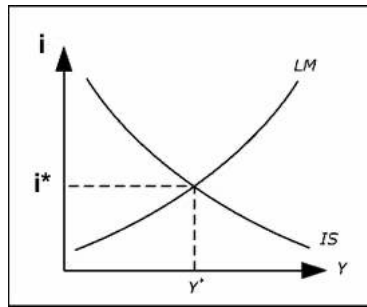10. $\max_{x,y} f(x,y) = 2x^2 + 3y^2$ subject to $x^2 + y^2 \le 1/2$.

Ex. 12.8. Solve the problems for 3 variables using KKT

1. $\min_{x,y,z} f(x,y,z) = -x^3 + y^2 - 2xz^2$ subject to $2x + y^2 + z - 5 = 0$, $5x - y^2 - z \ge 2$, and x, y, z all ≥ 0. Is (1,0,3) a solution?

2. $\min_{x,y,z} = x^2 + y^2 + z^2$ subject to $-x + y - z \ge -10$, and $x + y + 4z \ge 20$. Find all solutions.

Chapter 13

Applications in Economics

13.1 The IS-LM Analysis

A tool in macroeconomics. The IS curve represents Investment-Saving (IS) that governs the goods market. The LM curve represents Liquidity-Money (LM) that governs the financial market. Their intersection is the equilibrium point, which tells us how GDP is connected with interest rate.

The IS curve is formalized by analyzing

$$AD = C + I$$
$$C = c_0 + c_1 Y$$
$$I = b_0 - b_1 i$$
$$AD = Y \quad \text{(in equilibrium)} ,$$

where AD is the Aggregate Demand, C is consumption, I is the investment, i is the interest rate, and Y is output (GDP). We abstract away government consumption (G) and export-import net (NX). Show that

$$Y = AD = \frac{1}{1 - c_1}(c_0 + b_0 - b_1 i) \tag{13.1}$$

The LM curve is formalized by analyzing the money market

$$M^d = m_1 Y - m_2 i$$
$$M^s = \bar{M}/P$$
$$M^d = M^s \quad \text{(in equilibrium)} ,$$

where M^d is the money demand, M^s is the real money supply, $\bar{M}$ is the amount of nominal money and P is the price level. Show that

$$i = \frac{1}{m_2}(m_1 Y - \frac{\bar{M}}{P}) \tag{13.2}$$

Ex. 13.1. Derive the Equilibrium by substituting (13.2) into (13.1).

Ex. 13.2 (Adding Government & Trade)**.** The economy is given by

$$
\begin{aligned}
\text{(aggregate demand)} \qquad & AD = C + I + G + (X - M) \\
\text{(consumption)} \qquad & C = 50 + 0.9\,(Y - T), \\
\text{(investment)} \qquad & I = 150 - 5i, \\
\text{(government spending)} \qquad & G = 100, \\
\text{(export)} \qquad & X = 20, \\
\text{(import)} \qquad & M = 10 + 0.1Y, \\
\text{(money demand)} \qquad & M^d = 0.2Y - 10i, \\
\text{(money supply)} \qquad & \bar{M} = 110, \\
\text{(price level)} \qquad & P = 1.1
\end{aligned}
$$

1. Show the IS, LM equations of this economy (knowing that $G = T, AD = Y, M^s = \bar{M}/P = M^d$.

2. Find the equilibrium output (Y^*) and interest rate (i^*)

3. Find the Balance of Trade $(X - M)$. Does this economy has a trade deficit or trade surplus?

4. If the government increases its expenditure to 120, what would happen to the equilibrium output and interest rate?

(Hint): (2) $Y^* = 900, i^* = 8$; (4) -80.
(Source): https://www.economicsdiscussion.net/is-lm-curve-model/algebraic-analysis-of-is-lm-model-with-numerical-problems/10609

13.2   Growth Accounting

This section shows how we can use math to examine the contributions of different factors to economic growth.

13.2.1   Technique

This section explains the technique "take logs and differentiate".
   First, "dot notation". Say $Y_{2000} = 1, Y_{2010} = 1.5$, so the change is

$$\triangle Y = Y_{2010} - Y_{2000}$$

which is 0.5. Then, what is the average change across 1 year? Since there are 10 years during this interval, the average change in 1 year is

$$\triangle Y / \triangle t,$$

which is 0.05. Imagine we want to know the average change across an even smaller interval, such as 1 quarter, 1 week, 1 day, or even 1 hour, the time interval must shrink smaller and smaller. By squeezing the interval, we have

$$\lim_{\triangle t \to 0} \frac{\triangle Y}{\triangle t} = \frac{dY}{dt}.$$

which follows the definition of derivatives. This is the derivatives wrt time, so it is denoted by a dot and is called **instantaneous change**:

$$\dot{Y} = \frac{dY}{dt}$$

Second, what is the rate of growth? The growth rate of Y is

$$g_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \frac{\triangle Y_t}{Y_{t-1}}$$

If $Y_{2000} = 1, Y_{2010} = 1.5$, then it has grown 50% from 2000 to 2010. But, what is the average growth during that time? Let the initial value $Y_{t-1} = Y$, then

$$\tilde{g} = \frac{g_t}{\triangle t} = \frac{\triangle Y}{\triangle t}\frac{1}{Y}$$

Let $\triangle t \to 0$, we can write the **instantaneous growth rate** as

$$\tilde{g} = \frac{dY}{dt}\frac{1}{Y} = \frac{\dot{Y}}{Y}$$

Third, the trick to simplify this expression is by reviewing the log property

$$y(t) = \log x(t)$$

then differentiate y to time, using the Chain Rule gives

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = \frac{1}{x}\dot{x}$$

The derivative wrt to time of the log of some variable is the growth rate of that variable. Thus

$$\frac{d \log Y}{dt} = \frac{\dot{Y}}{Y} = \tilde{g}$$

In conclusion, this method shows us a way to calculate the instantaneous growth rate, by taking logs and calculating the difference in logs, divided by the time span.

13.2.2   Hicks Neutral

Consider a production function

$$Y = AK^{\alpha}L^{1-\alpha}$$

where Y is the real GDP, K is real capital stock, L is labor service, α is capital share, A is total factor productivity (TFP).

Example 13.1. (Using data for Belgium, what are the main factors of growth? Answer by performing a growth accounting)

1. take logs on both sides.

2. take the derivatives wrt time to show the growth rate of factors

3. show the average growth rate every 10 years since 1959. Interpret the results.

 Answers:

1. Taking logs on both sides yields

$$\ln Y = \ln A + \alpha \ln K + (1 - \alpha) \ln L$$

2. Taking the derivatives wrt time yields

$$\underbrace{\frac{d \ln Y}{dt}}_{g_Y} = \underbrace{\frac{d \ln A}{dt}}_{g_A} + \alpha \underbrace{\frac{d \ln K}{dt}}_{g_K} + (1 - \alpha) \underbrace{\frac{d \ln L}{dt}}_{g_L}$$

3. You have the data, take logs and calculate the difference in the log between years, divided by the year interval.

   Check the results here: https://rpubs.com/thanhqtran/775009

13.3   Solow Model

The economy has the following production function:

$$Y_t = F(K_t, L_t) = AK_t^\alpha L_t^{1-\alpha},$$

where Y, A, K, L are output, technology, capital, and labor. The price for each unit of labor is $w > 0$, and the rental price for each unit of capital is $r > 0$. The saving rate of this economy is $s \in (0, 1)$ and in each period, the capital depreciates at the rate of $\delta \in (0, 1)$. No population or technological growth. Assume that

$$\lim_{K \to 0} F_K = \infty, \ \lim_{K \to \infty} F_K = 0,$$
$$\lim_{L \to 0} F_L = \infty, \ \lim_{L \to \infty} F_L = 0,$$

1. (Preliminary) Show that the production function is twice differentiable and exhibits diminishing marginal products in K and L. Show that the production function exhibits constant returns to scale.

2. (Optimization) The economy is maximizing profit, how much K, L it should employ? Verify using second-order conditions.

3. In economics, we care more about wealth per person. Let k_t be K_t/L_t, rewrite the production function, and question 1 in terms of k_t.

4. We want to know the dynamics of k_t over time. Capital is accumulated according to the following

$$K_{t+1} = sY_t + (1 - \delta)K_t$$

Rewrite the accumulation function in terms of k_t.

5. To see the amount of capital accumulated, we construct the difference

$$\triangle k_t = k_{t+1} - k_t$$

An equilibrium is a state where the variable stays the same over time. In other words, $\triangle k_t = 0$. Find the equilibrium capital per labor k^*.

6. (Comparative Statics) What would happen to the steady state k^* if δ increases? if s increases?

7. Use Python, write a simple code to show the dynamics of k, given $\delta = 0.05, \alpha = 0.33, s = 0.5$.

13.4 Overlapping Generations Model

13.4.1 Benchmark

A 2-period lived representative Household

$$\max_{c_1, d_{t+1}} \ln c_t + \beta \ln d_{t+1}$$

$$s.t. \& c_t + s_t = w_t,$$
$$d_{t+1} = R_{t+1}s_t$$

Firms' production

$$Y_t = K_t^\alpha L_t^{1-\alpha}$$

and profit

$$\pi_t = Y_t - R_t K_t - w_t L_t$$

The population N_t grows at the rate of n.

1. (Household problem) Solve for c_t, d_{t+1} using either Lagrangian or substitution method (in this case, find s_t). Verify the results were optimal.

2. (Firms) Rewrite the problem in terms of k_t where

$$k_t = \frac{K_t}{L_t}$$

Solve the problems for profit-maximizing firms.

3. (Equilibrium) The labor market clears s.t. the demand for labor equals the supply of labor

$$L_t = N_t$$

The capital market clears s.t. the demand for capital equals the supply of capital

$$K_{t+1} = s_t N_t$$

Write the capital accumulation function per labor

$$k_{t+1} = g(k_t)$$

4. Find the steady-state capital-labor ratio k^*.

[Hint: $k^* = \left(\dfrac{\beta(1-\alpha)}{(1+\beta)(1+n)} \right)^{1/(1-\alpha)}$].

5. (Comparative Statics) What would happen to k^* if β increases?

13.4.2   Numerical Simulations

Use Python and the following parameters, write a code to see the progression of k_t over time.

Parameters	Values	Meaning
α	0.33	capital share
β	$0.99^{30 \times 4}$	subjective discount (.99 per quarter)
n	0.3	population growth
K_0	0.01	initial capital stock
L_0	1	initial population stock

All simulation codes can be found here:
https://github.com/thanhqtran/spring23_camp/tree/main/math

13.5   DIY Projects

Project 1 (Do your own country growth accounting). Pick a country of your own and perform Growth Accounting every 10 years, dating as far back as possible. Then interpret the results.

Data source:
https://www.rug.nl/ggdc/productivity/pwt/?lang=en
Machine-readable:
https://raw.githubusercontent.com/thanhqtran/dataset/main/pwt/pwt10_raw.csv
You can do it with R, Python, or Excel. If you want a longer time span, check this data source
Maddison Project
https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2020?lang=en

Project 2. Calibrate K, L, n, α, β so that it matches your country's data. Perform a simulation to see the long-run dynamics of capital/output per labor in your country using the OLG model.

13.6 Optional Problems

13.6.1 Endogenous Retirement

We expand the OLG model to allow endogenous retirement.

$$\max_{c_1, d_{t+1}} \ln c_t + \pi(\beta \ln d_{t+1} + \ln l_{t+1})$$

$$s.t. \& c_t + s_t = w_t,$$

$$d_{t+1} = R_{t+1}s_t + (1 - l_{t+1})\varepsilon w_{t+1},$$

where we added π as survival rate, l_{t+1} is time spent for true retirement (leisure), and $\varepsilon \in (0,1)$ is productivity when old.

The production side is the same

$$Y_t = K_t^\alpha L_t^{1-\alpha}$$

1. Find optimal choices of c_t, s_t, l_{t+1} using KKT formulation.

$$\mathcal{L} = \ln c_t + \pi(\beta \ln d_{t+1} + \ln l_{t+1}) + \lambda_1(c_t + \frac{d_{t+1}}{R_{t+1}} - (1 - l_{t+1})\varepsilon w_{t+1} + w_t) + \lambda_2(1 - l_{t+1})$$

2. (Comparative Statics) What happens to l_{t+1} if π, ε change?

3. Find k^*

An application of this model can be found in Tran (2022).

13.6.2 Endogenous Education and Fertility

We expand the OLG model to allow endogenous education and fertility.

$$\max_{c_1, d_{t+1}} \ln c_t + \beta \ln d_{t+1} + \ln(n_t \Pi_t)$$

$$s.t. \& c_t + s_t = w_t(1 - \phi n_t),$$

$$\Pi_t = (e_t + \theta)^\eta,$$

$$d_{t+1} = R_{t+1}s_t,$$

where we added n, e, as parents' choice of fertility and educational investment per child, Π is the education-to-skill conversion algorithm. Parameters ϕ, η govern the time-cost of raising children and returns to education.

1. Find optimal choices of c_t, s_t, n_t, e_t. Can e_t equal zero? on which condition? (hint: there is a threshold level of w where too poor parents cannot afford children's education)

2. What can be said about the children's quantity-quality trade-off?

3. (Comparative Statics) What happens to n_t, e_t if ϕ, η change?

An application of this model can be found in de la Croix (2012).

13.6.3   Ordinary Least-Squares (OLS)

We briefly review the Ordinary Least-Squares in Linear Regression. The fit function we seek to find is

$$y = \alpha + \beta x$$

Suppose the error term of each data point is e, then the actual function is

$$y = \alpha + \beta x_t + e_t$$

for $t = 1, 2, \ldots, T$ data points. We need to minimize the errors. To prevent error canceling-out problem, we square them and obtain the **loss function** containing the mean square error

$$L(\alpha, \beta) = \frac{1}{T} \sum_{t=1}^{T} e_t^2 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \alpha - \beta x_t)^2$$

Expansion gives

$$L(\alpha, \beta) = \frac{1}{T} \sum_{t=1}^{T} (y_t^2 + \alpha^2 + \beta^2 x_t^2 - 2\alpha y_t - 2\beta x_t y_t + 2\alpha\beta x_t)$$

Denote the statistical means of x_t and y_t as

$$\mu_x = \frac{1}{T} \sum_{t=1}^{T} x_t, \qquad\qquad \mu_y = \frac{1}{T} \sum_{t=1}^{T} y_t$$

Statistical variances are

$$\sigma_{xx} = \frac{1}{T} \sum_{t=1}^{T} (x_t - \mu_t)^2, \quad \sigma_{yy} = \frac{1}{T} \sum_{t=1}^{T} (y_t - \mu_y)^2 \quad \sigma_{xy} = \frac{1}{T} \sum_{t=1}^{T} (x_t - \mu_x)(y_t - \mu_y)$$

Assume that x_t are not all equal, so $\sigma_{xx} > 0$. We can verify that

$$\sigma_{xx} = \frac{1}{T} \sum_{t=1}^{T} x_t^2 - \mu_x^2, \quad \sigma_{yy} = \frac{1}{T} \sum_{t=1}^{T} y_t^2 - \mu_y^2 \quad \mu_{xy} = \frac{1}{T} \sum_{t=1}^{T} x_t y_t - \mu_x \mu_y$$

And then rewrite the loss function as

$$L(\alpha, \beta) = (\sigma_{yy} + \mu_y)^2 + \alpha^2 + \beta^2 (\sigma_{xx} + \mu_x)^2 - 2\alpha\mu_y = 2\beta(\sigma_{xy} + \mu_x\mu_y) + 2\alpha\beta\mu_x$$
$$= \alpha^2 + \mu_y^2 + \beta^2 \mu_x^2 - 2\alpha\mu_y - 2\beta\mu_x\mu_y + 2\alpha\beta\mu_x + \beta^2 \sigma_{xx} - 2\beta\sigma_{xy} + \sigma_{yy}$$

FOCs

$$L'_\alpha = 2\alpha - 2\mu_y + 2\beta\mu_x = 0,$$
$$L'_\beta = 2\beta\mu_x^2 - 2\mu_x\mu_y + 2\alpha\mu_x + 2\beta\sigma_{xx} - 2\sigma_{xy} = 0$$

Note that

$$L'_\beta = \mu_x L'_\alpha + 2\beta\sigma_{xx} - 2\sigma_{xy}$$

So there is a set of unique solutions of $(\hat{\alpha}, \hat{\beta})$ such that

$$\hat{\beta} = \frac{\sigma_{xy}}{\sigma_{xx}}, \qquad \hat{\alpha} = \mu_y - \hat{\beta}\mu_x = \mu_y - \left(\frac{\sigma_{xy}}{\sigma_{xx}}\right)\mu_x$$

Sufficient conditions: Furthermore, $L''_{\alpha\alpha} = 2, L''_{\alpha\beta} = 2\mu_x, L''_{\beta\beta} = 2\mu_x^2 + 2\sigma_{xx}$, thus $L''_{\alpha\alpha} \geq 0, L''_{\beta\beta} \geq 0$, and the det of the Hessian is

$$L''_{\alpha\alpha}L''_{\beta\beta} - (L''_{\alpha\beta})^2 = 2(2\mu_x^2 + 2\sigma_{xx}) - (2\mu_x)^2 = 4\sigma_{xx} = 4\frac{1}{T}\sum_{t=1}^{T}(x_t - \mu_x)^2 \geq 0$$

The Hessian matrix is positive-definite, thus justifies the solutions.

13.6.4   Markov Chain

[**2-state Markov**] Assume that each period, some portion of the population gets unemployed/employed, according to the following chart. This is called a Markov Chain.
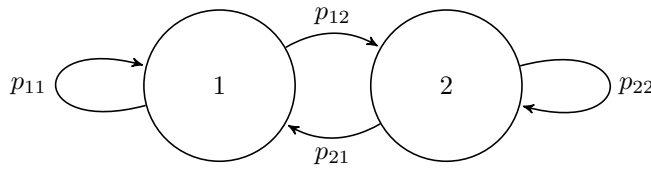

where the transition matrix is

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix}$$

The sum of all rows of a Markov transition matrix must be equal to 1.

$$p_{11} + p_{21} = 1, \qquad\qquad p_{12} + p_{22} = 1$$

It guarantees that the Markov chain will always be in some known state at each time step. If the sum of any row is not equal to 1, it means that there is a probability of transitioning to a state that is not in the set of possible states, which violates the definition of a Markov chain.

Let x, y are the state of employed and unemployed, and q, p are the transition probabilities associated with each state.
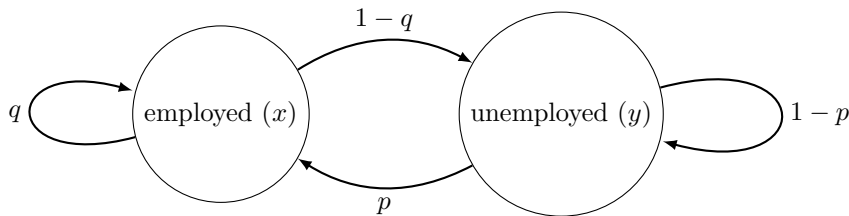
This transition says that

$$qx_t + py_t = x_{t+1},$$
$$(1-q)x_t + (1-p)y_t = y_{t+1}$$

This transition can be written in matrix form as

$$\underbrace{\begin{pmatrix} x_t & y_t \end{pmatrix}}_{\mathbf{x}_t} \underbrace{\begin{pmatrix} q & 1-q \\ p & 1-p \end{pmatrix}}_{\mathbf{P}} = \underbrace{\begin{pmatrix} x_{t+1} & y_{t+1} \end{pmatrix}}_{\mathbf{x}_{t+1}}$$

The dynamics is

$$\mathbf{x}_{t-1}\mathbf{P} = \mathbf{x}_t \Rightarrow \mathbf{x}_1\mathbf{P}^{t-1} = \mathbf{x}_t$$

We wonder what happens if $t \to \infty$. Does the labor market converge? To solve that, assume the steady state $\mathbf{x}^*$ exists, then at that stage, it must be that

$$\mathbf{x}^* = \mathbf{x}^*\mathbf{P}$$

Let $\mathbf{x}^*$ be a 2×2 matrix, then

$$\begin{pmatrix} x^* & y^* \end{pmatrix} = \begin{pmatrix} x^* & y^* \end{pmatrix} \begin{pmatrix} q & 1-q \\ p & 1-p \end{pmatrix}$$

where we can solve for x^*, y^* as

$$x^* = \frac{p}{1+p-q}, \qquad\qquad y^* = \frac{1-q}{1+p-q}.$$

Ex. 13.3. Given the transition matrix the US's employment in 1966

$$\mathbf{P} = \begin{pmatrix} 0.998 & 0.002 \\ 0.136 & 0.864 \end{pmatrix}$$

Find the steady state.

3-state Markov

Ex. 13.4. The transitional matrix is now

$$\mathbf{P} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$$

The stationary condition:
$$\mathbf{x} = \mathbf{P}\mathbf{x}$$

Find the steady-state $\mathbf{x}^*$.

Bibliography

Chiang, AC and K Wainwright (1984). Fundamental methods of mathematical economics (3rd. ed).
    MacGraw-Hill.

de la Croix, David (2012). Fertility, Education, Growth, and Sustainability. Cam-
    bridge University Press.

Simon, Carl P and Lawrence Blume (1994). Mathematics for economists. Nor-
    ton New York.

Sydsæter, Knut and Peter J Hammond (2008). Essential mathematics for economic analysis.
    Pearson Education.

Sydsæter, Knut et al. (2008). Further mathematics for economic analysis. Pear-
    son education.

Tran, Quang-Thanh (2022). "The aging tax on potential growth in Asia". In:
    Journal of Asian Economics 81, p. 101495. ISSN: 1049-0078.