离散数学讲义

李雅樵

2022 年 6 月 17 日

目录

引言

离散是和连续对比，离散和连续就好像阴和阳。自然世界和人类生活中既有连续的对象，也有离散的对象。比如流水，风，电流等都是连续的对象。分子的结构，DNA 的排列，地图上的城市，用微信发送的汉语消息，网络购物的一个个商品，大千世界中一个个的人，等等，都是离散的对象。研究连续对象的基础学科是微积分，研究离散对象的基础学科就是离散数学。

在人类发展历史中，连续和离散的数学都在不断发展。无论连续还是离散的数学，最基本的问题都是计算。人类的生活离不开计算。随着时代的发展，人类能计算的东西越来越多。由最初简单的结绳记事，计算一小块土地面积，到计算地球的半径，太阳的温度，到核能发电，基因测序，再到今天的阿尔法狗围棋打败所有人类最杰出的棋手，都与计算问题密切相关。人类发展史的一个重要部分就是计算发展的历史。比如，结绳记事就是最原始的离散数学，用来数那些我们可以一个一个数的东西，比如三只羊，两颗桃子，等等。而对土地面积的计算，就有赖于连续数学的发展。不过，连续和离散并不是两个独立的世界。比如，在微积分中我们学过

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots=\frac{\pi^2}{6}.$$

在左边的每一项都是自然数平方的倒数，自然数当然是离散的，但是右边出现了 π：它与圆的面积有关，而且是一个无理数，是一个连续的对象。这个等式表明离散和连续由紧密的联系。另一个类似的等式是

$$e^{i\pi}+1=0.$$

这个等式中 0 和 1 自然是离散的对象，但是 e 和 π 都是无理数，是连续的对象，而且 i 甚至不是实数而是一个虚数。

事实上，随着科学的发展，以前我们认为是连续的对象可能其实是离散的，或者也有离散的一面。比如水我们觉得当然是连续的，但化学告诉我们每个水分子是由两个氢原子和一个氧原子组成的，因此水分子有离散的结构。又比如，物体的质量，我们认为自然是连续的。然而，量子物理告诉我们，能量是离散的而不是连续的。因此，根据爱因斯坦的质能方程

$$E=mc^2$$

我们就能得出结论：质量也是离散的。又比如，光，我们直觉上觉得它是连续的。但是现代物理告诉我们，光具有波粒二象性，就是说，光有时候表现出波的特点（比如光的衍射），有时候表现出粒子的特点（比如光的直线传播）。波自然是连续的，而粒子则是离散的。再比如，我们的身体，我们觉得也是连续的。然而生物学研究表明 DNA 有双螺旋结构，也是离散的。等等。总之，无论数学还是自然界，离散和连续都是紧密联系的。

从数学的发展史看，连续的数学从 17 世纪的牛顿和莱布尼兹以来取得了重大的发展。而自人类进入 20 世纪以来，以图灵（Alan Turing）和厄多斯（Paul Erdös）等为代表，离散数学也取得了长远的进步，涌现了许多著名的离散数学家。2021 年的数学大奖阿贝尔奖，就授予了两个著名的离散数学家：匈牙利的洛瓦兹（László Lovász）和美国的维格森（Avi Wigderson）。可以说，离散数学在人类历史上从未像今天这么重要，像今天这样发展迅猛。从工程设计，商业拍卖，基因测序，到大数据，人工智能，5G 通信，等等，全都依赖于离散数学的发展。许多科学也越来越受到离散数学的影响，包括物理学（比如量子物理），化学（比如晶体结构），生物（比如基因测序），经济学和社会科学（比如拍卖理论，社会选择理论（投票）），语言文学（比如计算语言学），等等。更不用说计算机科与通信等学科，则几乎完全是建立在离散数学的基础上。在 2000 年，数学家们提出了 21 世纪最重要的尚未解决的七个数学问题，有两个（黎曼猜想，P 与 NP 问题）就与离散数学直接相关。如果说 19 世纪以前的科学是建立在连续数学的基础上，那么对今天的科学和生活而言，离散与连续数学已并驾齐驱，共同成为两座基石。18 世纪，法语曾是欧洲上流社会交流的语言，20 世纪以来，英语成为世界的语言。而从 20 世纪中期以来，尤其是进入 21 世纪，可以说离散数学已经成为隐含在我们生活背后的一门基本语言。这门离散数学的基础课，就是为了介绍这门语言的基本笔画和基本字词。如前所述，掌握了它们，将为我们打开一扇大门通向一个无比宽敞的世界。

最后，引用 Essential discrete mathematics for computer science 书中的一段话作为结束：

> *Diffie 和 Hellman 发表于 1976 年的论文《密码学的新方向》，其影响很少有其他出版物可与之相比。几乎一夜之间，互不熟识的常人之间可以秘密通信。不再需要武装卫士传输有价值的信息。世界上最强大的政府也无法译解其截获的通讯信息。大规模的互联网贸易也从而可能。所有这一切，皆归功于一点简单离散数学的创新应用。Very few publications have had the impact of Diffie and Hellman's 1976 paper New Directions in Cryptography. Almost overnight, secret communications became possible between ordinary people who barely knew each other. No longer were armed guards needed in order to transmit value information from place to place. Even the world's most powerful governments couldn't interpret the communications they intercepted. Secure Internet commerce became possible on a massive scale. All this because of the creative application of some simple discrete mathematics.* ——*Essential discrete mathematics for computer science*

以下是教学参考及教学计划。

教材：（1）讲义。（2）屈婉玲，耿素云，张立昂，离散数学。

参考书：

- L. Lovász and K. Vesztergombi, Discrete Mathematics.
- H. Lewis and R. Zax, Essential Discrete Mathematics for Computer Science.

成绩计算：作业 30%~50%，期末考试 50%~70%。无期中考试。

教学计划：本课共 25 节课，每节课 90 分钟。以下是预计的教学安排，但可能根据教学进度会适当调整。

- 第 1 课: 引言与课程简介，素数，欧几里得算法.
- 第 2 课: 基本证明方法（1）：直接证明，反证法，分类讨论。
- 第 3 课: 基本证明方法（2）：数学归纳法，强归纳法，对称性与不失一般性，鸽笼原理及应用。
- 第 4 课: 计数（1）：二项式定理，组合数大小的估计。
- 第 5 课: 计数（2）：集合、关系、函数。
- 第 6 课: 计数（3）：计数的基本方法：容斥原理，双计数，构造映射（无限集的大小)。
- 第 7 课: 计数（4）：计数的基本方法：容斥原理，双计数，构造映射（无限集的大小)。
- 第 8 课: 图论（1）：哥尼斯堡七桥问题，握手定理。
- 第 9 课: 图论（2）：从树到平面图的欧拉公式，图的同构。
- 第 10 课: how to study and work?
- 第 11 课: 图论（3）：三个简洁快速的算法。
- 第 12 课: 图论（4）：图的矩阵表示及应用：计算通路数，估计正则图的独立集的大小。
- 第 13 课: 图论（5）：图的染色及应用，P 与 NP 问题介绍。
- 第 14 课: 小测验，习题课
- 第 15 课: 逻辑（1）：命题逻辑的基本概念
- 第 16 课: 逻辑（2）：决策树，范式，完备集，消解
- 第 17 课: 逻辑（3）：命题逻辑的推理
- 第 18 课: 逻辑（4）：一阶逻辑的基本概念
- 第 19 课: 逻辑（5）：一阶逻辑的验算与推理
- 第 20 课: 代数（1）：群的定义和例子
- 第 21 课: 代数（2）：群的定义和例子（续)，拉格朗日定理及子群的例子
- 第 22 课: 代数（3）：群的同构与循环群
- 第 23 课: 代数（4）：环与域的概念，公钥加密简介
- 第 24 课: 关系：等价关系，偏序关系，关系的表示与运算
- 第 25-27 课: 习题课，复习

第一章　素数与欧几里得算法

> 上帝懂得算术。 ——高斯
>
> 只有勇于深入探究这一崇高学科的人，才能领略其迷人的魅力。 ——高斯

1.1　素数

素数是数学中最重要的研究对象之一。素数之于自然数就好像元素之于化学分子。

定义 1.1. 自然数集：$0, 1, 2, 3, 4, 5, 6, \ldots$

设 a 和 b 是两个自然数，如果 a 整除 b，则称 a 是 b 的一个因子。数学上用 $a \mid b$ 来表示 a 整除 b。

如果一个自然数 $p \geq 2$ 刚好有 2 个因子 1 和 p，则称 p 是一个素数。

例 1.2. 素数的例子：$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, \ldots$

不是素数的例子：(*1*) 1：只有 1 个因子 1；(*2*) 4：有 3 个因子：$1, 2, 4$；(*3*) 6：有 4 个因子：$1, 2, 3, 6$。

思考： 给一个自然数 $n \geq 2$，如何确定 n 到底是不是素数呢?

根据定义，我们只需要验证在 $2, 3, \ldots, n-1$ 中有没有哪个数可以整除 n。如果有，那么 n 就不是素数，如果没有，那么 n 就是素数。我们可以把这个计算方法写成算法1。

Algorithm 1 计算 n 是否是素数

```
procedure IsPrime
    i ← 2
    while i ≤ n − 1 do
        if i | n then
            Output: n is not a prime. Terminate.
        else
            i ← i + 1
    if i == n then
        Output: n is a prime. Terminate.
```

如果把每次计算 i 是否整除 n 算作一次计算，算法1在最坏的情况下需要做大约 n 次运算，换句话说，算法1最多需要做大约 n 次运算。有没有更快的办法呢? 这是素数研究的重大问题。2002 年，印

度计算机科学家 M. Agrawal, N. Kayal, N. Saxena 取得突破，找到了一个快得多的算法[1]，随后他们的算法被其他的数学家改进到只需要做大概 $(\log_2 n)^6$ 次运算。这一算法被称为 AKS 素数测试算法。

对上面的计算进一步思考，就能够发现如下的定理。

定理 1.3. *算术基本定理：每一个自然数都可以写成素数的乘积。*

我们知道化学元素目前发现的有 100 多种，每个化学分子由不同的化学元素组合而成。因为每个自然数都可以写成素数的乘积，因此对素数的了解就很重要，素数就好比是自然数的 "化学元素"。那么，素数有多少个呢？其实，在人类对素数的研究中，产生了许多重要的问题，有许多至今仍未解决。

问题 1： 素数有多少个呢?

回答 1： 人类两千年前就知道了这个问题的答案。

问题 2： 在 10000 到 100000 之间，有没有素数呢？有的话，有多少个呢?

回答 2： 我们可以用算法1来一个数字一个数字的去验算，这需要验算 $100000 - 10000 = 90000$ 个数字。有没有更好的办法呢？著名数学家高斯花了许多时间研究这个问题。人类在 19 世纪基本解决了这个问题。这就是著名的素数定理。

问题 3： 黎曼猜想在素数定理的基础上，对素数的分布提出了更精确的估计。但黎曼猜想至今还未解决。是 21 世纪七大数学问题之一。

问题 4： 哥德巴赫猜想：每一个大于 2 的偶数都可以写成两个素数的和。

回答 4： 仍未解决。中国数学家陈景润等在这个问题上做出了杰出的贡献。

问题 5： 孪生素数猜想：有无穷多对孪生素数[2]。

回答 5： 仍未解决。2013 年，数学家张益唐在这个问题上做出了杰出的贡献。

与素数相关的问题还有很多。

也许我们会问：研究素数有什么实际用处吗？回答是：我们每天都在用素数！现代加密的基本方法就有赖于对素数的研究。在本课程的最后我们会简介公钥加密法，在那里我们将会看到素数如何发挥作用。

1.2 欧几里得算法

定义 1.4. *设 a 和 b 是两个自然数，设 c 也是一个自然数。如果 $c \mid a$ 并且 $c \mid b$，则称 c 是 a 和 b 的一个公因数。a 和 b 最大的公因数叫做 a 和 b 的最大公因数，用 (a, b) 表示。*

例 1.5. $(4, 6) = 2$, $(9, 0) = 9$, $(12, 36) = 12$, $(24, 35) = 1$, $(18, 48) = 6$.

思考： 为什么我们没有考虑最小公因数?

如何求两个自然数 a 和 b 的最大公因数 (a, b) 呢？假设 $a \leq b$。显然，最大公因数不可能比 a 大。仿照算法1，我们可以这样求：逐个计算数字 $1, 2, 3, \ldots, a$，看哪些是 a 和 b 的公因数，然后挑选出最大的。这个计算方法我们用伪代码写成算法2。

[1] 当时，Kayal 和 Saxena 还是本科生。

[2] 如果 p 和 $p + 2$ 都是素数，则称这两个素数未孪生（双胞胎）素数。比如 3 和 5，5 和 7，11 和 13，17 和 19，等等。

Algorithm 2 计算 a 和 b 的最大公因数 (a, b)

procedure SIMPLEGCD
  $m \leftarrow \min\{a, b\}$
  $i \leftarrow 1$
  $c \leftarrow 1$
  while $i \leq m$ **do**
    if $i \mid a$ 且 $i \mid b$ **then**
      $c \leftarrow i$
    $i \leftarrow i + 1$
  return c

如果我们用算法2来计算 $(18, 48)$，那么需要做 18 次运算。有没有更快的办法呢? 早在两千多年前，希腊的数学家欧几里得就找到了更快的算法。为了介绍这个算法，我们引入一个数学符号。用 $a \mod b$ 来表示 a 被 b 除所得的余数。比如，

$$(5 \mod 3) = 2, \quad (28 \mod 4) = 0, \quad (63 \mod 15) = 3, \quad (37 \mod 58) = 37.$$

算法3是欧几里得算法的伪代码。

Algorithm 3 计算 a 和 b 的最大公因数 (a, b)

procedure EUCLIDGCD
  while b != 0 **do**
    $a' \leftarrow b$
    $b' \leftarrow (a \mod b)$
    $a \leftarrow a'$
    $b \leftarrow b'$
  return a

例 1.6. 用欧几里得算法计算 $(18, 48)$。

$$(18, 48) \to (48, 18) \to (18, 12) \to (12, 6) \to (6, 0).$$

因此，$(18, 48) = 6$。

欧几里得[3]算法是世界上最早的算法之一。

1.3 英文单词

- 离散数学: discrete mathematics
- 连续：continuous

[3]注：欧几里得原本的算法与算法3稍有不同。

- 自然数：natural number
- 整除：divide
- 素数：prime number (or just say prime)
- 定理：theorem；基本定理：fundamental theorem
- 算术：arithmetic
- 算法：algorithm
- 素数定理：prime number theorem
- 黎曼猜想：Riemann hypothesis
- 最大公因子（最大公约数）：greatest common divisor
- 欧几里得：Euclid

第二章　数学证明的基本方法

> 反证法比任何国际象棋的弃子都更为精妙：棋手可以放弃一个兵或棋子，但数学家赌上整个游戏。(*It (proof by contradiction) is a far finer gambit than any chess gambit: a chess player may offer the sacrifice of a pawn or even a piece, but a mathematician offers the game.*)
>
> ——*G. H. Hardy*（哈代）
>
> 令人惊讶的是，鸽笼原理的复杂度本质上依赖于鸽子的数量。(*Surprisingly, the complexity of the pigeonhole principle essentially depends on the number of pigeons*)　——*A. A. Razborov*

要解决一个复杂的数学问题往往需要很多步骤，而每一步可能使用不同的方法。本节我们介绍最基本的证明方法。

2.1　直接证明

直接证明往往是直接根据定义验证，或直接给出问题的计算方法。直接证明往往适用于肯定的陈述。

例 2.1. 11 是素数。

证明. 由直接计算可知

$$11=2\times5+1,\quad 11=3\times3+2,\quad 11=4\times2+3,\quad 11=5\times2+1,\quad 11=6\times1+5,$$
$$11=7\times1+4,\quad 11=8\times1+3,\quad 11=9\times1+2,\quad 11=10\times1+1.$$

即，在 2 到 10 之间没有 11 的因子。因此 11 是素数。 □

例 2.2. 每个奇数都能写成两个平方数的差。

证明. 设 n 是一个奇数，则存在整数 k 使得 $n=2k+1$。由直接计算可知 $(k+1)^2-k^2=2k+1=n$。 □

2.2　反证法

有不少数学问题很难用直接证明，这时候可以考虑反证法。反证法就是先假设要证明的结论不正确，我们想办法推导出某种矛盾。这个矛盾就表明假设不能成立，也就是说，要证明的结论是正确的。

命题 2.3. *设 a 是整数，如果 a^2 是偶数，则 a 是偶数。*

这个陈述看起来很明显。如果我们尝试直接证明的话，根据定义，一个数是偶数就是说它可以被 2 整除，那么因为 a^2 是偶数，所以就存在某个整数 k，使得 $a^2 = 2k$。这样我们得到 $a = \sqrt{2k}$ 或 $-\sqrt{2k}$。但是如何说明 $\sqrt{2k}$ 是偶数呢？看起来很难直接说明。这时候我们可以考虑反证法。

证明. 假设 a 是奇数，则存在整数 t 使得 $a = 2t+1$。因此 $a^2 = (2t+1)^2 = 4t^2+4t+1 = 2(2t^2+2t)+1$。这表明 a^2 是奇数，矛盾。 □

反证法往往适用于根据定义来说是否定的陈述。下面举两个历史上著名的例子。

命题 2.4. *$\sqrt{2}$ 是无理数。*

这个陈述本身 "是无理数" 看起来是肯定的。但我们对无理数的定义是它不是有理数。因此，上面的陈述实际是说：$\sqrt{2}$ 不是有理数。

证明. 假设 $\sqrt{2}$ 是有理数，则存在非零的自然数 a, b 满足 $(a, b) = 1$，使得 $\sqrt{2} = a/b$。因此，

$$a^2 = 2b^2.$$

这表明 a^2 是偶数，根据命题 2.3，所以 a 是偶数。

因为 $(a, b) = 1$，所以 a, b 不可能都是偶数，由于 a 是偶数，因此 b 是奇数。

由于 a 是自然数并且是偶数，因此存在自然数 $c > 0$ 使得 $a = 2c$。所以

$$2b^2 = a^2 = 4c^2.$$

所以 $b^2 = 2c^2$。因此 b^2 是偶数，根据命题 2.3，所以 b 是偶数。但前面我们已说明 b 必须是奇数，矛盾。 □

命题 2.5. *素数有无穷多个。*

这个命题看起来也像是肯定的，但无穷的定义是：不是有限的，所以，命题实际是说：素数不是有限个。

证明. 假设只有有限个素数，我们可以把它们从小到大排列如下：$p_1 = 2, p_2 = 3, p_3 = 5, \ldots, p_n$。现在考虑自然数 $k = p_1 \times p_2 \times p_3 \times \cdots \times p_n + 1$。因为 $k > p_n$，根据假设，那么 k 不是素数。根据算术基本定理，k 能被某个素数整除。但是因为所有的素数就是 $p_1 = 2, p_2 = 3, p_3 = 5, \ldots, p_n$，所以 k 能被某个 p_i 整除。但是根据 k 的定义，k 被 p_i 除得到余数 1，所以 k 不能被 p_i 整除。矛盾。 □

下面我们用反证法证明一个很简单又很重要的定理：鸽笼原理。鸽笼原理在数学中运用很广泛。其内容很简单：如果我们有 10 个笼但有 11 个鸽子，我们要把鸽子放进笼里，那么必定有某一个笼里要放至少 2 个或更多鸽子。更一般地，如果我们有 n 个笼但有 $n+1$ 个鸽子，我们要把这些鸽子放进笼子里，那么必定有某一个笼子里要放至少 2 个或更多鸽子。

鸽笼原理看起来明显是对的。我们怎么用数学的语言来描述鸽笼原理呢，怎么证明它呢？

定理 2.6 (鸽笼原理)**.** *给定 X 和 Y 都是有限非空集合，并且 $|X| > |Y|$。对任意映射 $f : X \to Y$。必定存在 $a, b \in X$，$a \neq b$, 满足 $f(a) = f(b)$。*

证明. 假设结论不成立。那么，则存在某个映射 $f : X \to Y$ 满足对任意的 $a, b \in X$，$a \neq b$, 都有 $f(a) \neq f(b)$。换句话说，在 f 映射下，X 里的每个元素的像都各不相同。假设 $X = \{x_1, x_2, \ldots, x_n\}$。考虑集合

$$\{f(x_1), f(x_2), \ldots, f(x_n)\}.$$

那么，因为 $f(x_i)$ 各不相同，所以

$$|\{f(x_1), f(x_2), \ldots, f(x_n)\}| = n = |X| > |Y|. \tag{2.1}$$

但是，因为 f 是一个映射，根据定义，有

$$\{f(x_1), f(x_2), \ldots, f(x_n)\} \subseteq Y.$$

这又表明

$$|\{f(x_1), f(x_2), \ldots, f(x_n)\}| \leq |Y|. \tag{2.2}$$

这与(2.1)矛盾。 □

2.3 分类讨论法

有的问题本身可以很自然地分解成几种情况，我们可以对每个情况单独分析，单独证明。

例 2.7. *设 n 是自然数，则 $n^2 \mod 4$ 等于 0 或 1。*

证明. 分 n 是偶数或奇数两种情况讨论。

- n 是偶数，则存在自然数 k 使得 $n = 2k$。此时 $n^2 = (2k)^2 = 4k^2$。所以 $n^2 \mod 4$ 等于 0。
- n 是奇数，则存在自然数 k 使得 $n = 2k+1$。此时 $n^2 = (2k+1)^2 = 4k^2 + 4k + 1$。所以 $n^2 \mod 4$ 等于 1。 □

下面是著名的拉姆齐定理，也可以用分类讨论法来证明。拉姆齐定理后来发展成拉姆齐理论，其主要特点是从看似无序的对象中寻找结构，是离散数学的一个重要分支，在计算机学科也有很多应用。

定理 2.8 (拉姆齐定理)**.** *假设[1]人群中的“认识”是相互的，即如果 A 认识 B，那么 B 也认识 A，反过来，如果 A 不认识 B，那么 B 也不认识 A。在这个假设下，任何 6 个人之间，要么有 3 个人相互认识，要么有 3 个人相互不认识。*

证明. 任选一个人，用 X 表示这个人。我们根据 X 认识的人的个数 k 来分类讨论。当然，$0 \leq k \leq 5$。

- 情形一：$k \geq 3$。因此可以假设 X 认识 A, B, C。这时，再细分讨论。

[1]在真实生活中，这一假设一般并不成立：可能 A 认识 B 但 B 并不认识 A。

– A, B, C 相互都不认识。这正好是题目要求的。

– A, B, C 至少有两人相互认识。这时有三种情况。

(i) A, B 相互认识。此时 X, A, B 三人相互认识。

(ii) B, C 相互认识。此时 X, B, C 三人相互认识。

(iii) C, A 相互认识。此时 X, C, A 三人相互认识。

- 情形二: $k \leq 2$。这时, 因为 X 最多认识 2 人, 所以 X 至少有 3 人不认识。假设 X 不认识 A, B, C。可仿照情形一类似地讨论。(略)

□

2.4 数学归纳法与强归纳法

2.4.1 数学归纳法

中学时我们已经学过数学归纳法，下面我们看几个例子，作为复习。

命题 2.9. $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$。

证明. 基础情形：$n = 1$。左边 $= 1$，右边 $= \frac{1\times 2}{2} = 1$。等式成立。

归纳假设：假设等式对 k 成立。

现在证明等式对 $k + 1$ 时也成立。可计算如下：

$$\begin{aligned} 1 + 2 + 3 + \cdots + k + k + 1 &= (1 + 2 + 3 + \cdots + k) + k + 1 \\ &= \frac{k(k+1)}{2} + k + 1 = \frac{k(k+1) + 2(k+1)}{2} = \frac{(k+1)(k+2)}{2}. \end{aligned}$$

其中第二个等号运用了归纳假设。 □

下面我们用归纳法重新证明鸽笼原理，这个证明引自 [1]。

鸽笼原理证明二. 我们对 $|X|$ 的大小采用归纳法。首先，因为 X 和 Y 都不是空集，所以 $|Y| \geq 1$，又因为 $|X| > |Y|$，所以 $|X| \geq 2$。

基础情形：$|X| = 2$。此时，必定有 $|Y| = 1$。此时结论显然成立。

归纳假设：假设结论对 $|X| = k$ 时成立。即：如果 $|X| > |Y|$，且 $|X| = k$，则对任意的 $f : X \to Y$ 都存在 $a, b \in X$，$a \neq b$，满足 $f(a) = f(b)$。

现在证明结论对 $|X| = k + 1$ 时也成立。考虑任意映射 $f : X \to Y$。首先如果 $|Y| = 1$，结论显然成立。因此假设 $|Y| \geq 2$。

任意选取 $y \in Y$，分三种情况讨论。

- 情形一：对任意的 $x \in X$，都有 $f(x) \neq y$。任意选取 $c \in X$, 定义

$$X' = X \setminus \{c\},$$
$$Y' = Y \setminus \{y\}.$$

考虑映射

$$\widetilde{f}: X' \to Y',$$
$$x \mapsto f(x).$$

注意，因为对任意的 $x \in X$ 都有 $f(x) \neq y$，所以特别地，对 $x \in X'$，也有 $f(x) \neq y$，换句话说，$f(x) \in Y'$. 所以映射 $\widetilde{f}$ 的定义是可行的。根据定义，$|X'| = k+1-1 = k$，且 $|X'| > |Y'|$。因此，$\widetilde{f}$ 满足归纳假设的条件。根据归纳假设，存在 $a, b \in X'$，$a \neq b$，满足 $\widetilde{f}(a) = \widetilde{f}(b)$。但是根据我们对 $\widetilde{f}$ 的定义，有

$$f(a) = \widetilde{f}(a) = \widetilde{f}(b) = f(b).$$

这就是所要证明的结论。

- 情形二：存在唯一的一个 $x \in X$ 满足 $f(x) = y$。可仿照情形一证明。（略）
- 情形三：存在 $a, b \in X$，$a \neq b$，满足 $f(a) = f(b)$。这就是所要的结论，因此结论自动成立。 □

2.4.2 强归纳法

我们通过一个例子来看强归纳法怎么使用，本例引自 [1]。

例 2.10. 设 $a_1 = 3, a_2 = 5$，且对于 $n \geq 3$，有 $a_n = 3a_{n-1} - 2a_{n-2}$。那么对所有 $n \geq 1$，有 $a_n = 2^n + 1$。

证明. 基础情形：验证 $n = 1$ 及 $n = 2$ 时，公式成立。

归纳假设：假设公式对于 k 和 $k+1$ 成立。

现在证明公式对于 $k+2$ 成立。可直接计算

$$a_{k+2} = 3a_{k+1} - 2a_k = 3(2^{k+1} + 1) - 2(2^k + 1) = 2^{k+2} + 1.$$ □

一般来说，当我们希望使用归纳法，但归纳假设如果只涉及第 m 项，此时无法完成对第 $m+1$ 项的证明时，就应该考虑使用强归纳法。在强归纳法中，归纳假设不仅可包含两项，也可以包含更多项，但必须注意，需要在基础情形中验证足够多的项成立。如果在基础情形中验证的项目不够，而使用强归纳法，可能得出错误的结果。下例同例2.10，但却得出错误的结论。

例 2.11. 设 $a_1 = 3, a_2 = 5$，且对于 $n \geq 3$，有 $a_n = 3a_{n-1} - 2a_{n-2}$。那么对所有 $n \geq 1$，有 $a_n = 3$。

错误的证明. 基础情形：验证 $n = 1$ 时，公式成立：$a_1 = 3$。

归纳假设：假设公式对于 k 和 $k+1$ 成立。

现在证明公式对于 $k+2$ 成立。可直接计算

$$a_{k+2} = 3a_{k+1} - 2a_k = 3 \times 3 - 2 \times 3 = 3.$$ □

上面的证明之所以错误，便是因为在归纳假设中使用了两项，但基础情形只验证了一项，因此得出错误的结论。

2.5 对称性与不失一般性

在数学推理中，运用对称性，运用不失一般性，往往可以帮助我们简化推理过程。

如果在我们要考虑的问题中某些概念本身有对称性，那么我们对它进行讨论时，就可以根据其对称性，简化讨论。

例 2.12. 设有一个正方形，边长是 8，证明三角形 A 和 B，如图*2.1*，的面积都是 32。

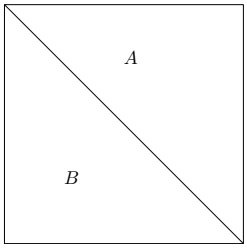

图 2.1: 正方形中两个三角形是对称的。

证明. 计算三角形 A 面积是 32。因为三角形 A 和 B 是对称的，面积相同，所以 B 的面积也是 32。 □

上面的例子用对称性并没有简化证明。我们再看一个例子。回顾拉姆齐定理。定理中有 “认识” 与 “不认识”。定理中假设[2]“认识” 是相互的，即如果 A 认识 B，那么 B 也认识 A，反过来，如果 A 不认识 B，那么 B 也不认识 A。因此，‘认识” 与 “不认识” 具有对称性。我们在推理中就可以使用这种对称性。具体地，在定理2.8的证明中，当我们分类讨论时，情形二与情形一其实是对称的。回忆我们有

- 情形一：$k \geq 3$，即 X 至少认识 3 人。
- 情形二：$k \geq 2$，即 X 至多认识 2 人。

我们可以等价地改写如下：

- 情形一：X 至少 “认识”3 人。
- 情形二：X 至少 “不认识”3 人。

这样我们可以清楚地看出，因为 ‘认识” 与 “不认识” 是对称的，所以情形一和情形二也是对称的，因此，当我们讨论完了情形一，情形二就不需要讨论了。在书写证明的时候，可以如下书写。

[2]注意，在真实的人群中，可能 A 认识 B 但 B 不认识 A，所以认识与不认识不具有对称性。

- 情形一：$k \geq 3$，即 X 至少认识 3 人。给出证明……
- 情形二：$k \geq 2$，即 X 至多认识 2 人，也即 X 至少不认识 3 人。因为认识与不认识具有对称性，情形一已经证明成立，故情形二也成立。

在问题本身有对称性的情况下，我们也常常用“不失一般性，假设……”来简化要讨论的不同情形。

例 2.13. 设有三只鸽子 a, b, c，两个笼子 X, Y。则必有一个笼子里至少有两只鸽子。

证明. 不失一般性，假设鸽子 a 放在笼子 X 中。下面讨论 b 和 c 的情况。如果 b 和 c 中至少有一个放在 X 中，则 X 中至少有两只鸽子，否则 b 和 c 都放在 Y 中，此时 Y 中有两只鸽子。 □

鸽子 a 有可能放在 X 中，也可能放在 Y 中。我们可以用分情况讨论法来一一证明。但是，因为两种情况其实具有对称性，因此，“不失一般性”，只需要讨论其中一种情况。但是，并不是所有的分情况讨论都可以用“不失一般性”来减少讨论的情况，比如例2.7，那里的两种情况各不相同，不具有对称性，因此每一种情况都必须单独讨论。

在第2.6节，我们会再次看到如何使用不失一般性简化推理。

2.6 鸽笼原理的应用

前已证明鸽笼原理，并提及鸽笼原理有广泛的应用，下面我们看几个例子。

例 2.14 ([1])**.** 从 2 到 40 之间任意选择 13 个不同的整数，那么其中必存在两个数 a, b 满足 $(a, b) > 1$。

验证你的直觉：如果要直接证明（验证）这个问题，猜猜大概需要进行多少次计算呢?
（A）几十次；（B）几百次；（C）几千次；（D）几万次；（E）几亿次。

如果要直接证明这个问题，2 到 40 之间总共有 39 个数字，从中选出 13 个数，总共有 $\binom{39}{13} = 8122425444$，也就是说有 81 亿多种选法。对每一组 13 个数字验证是否存在题目要求的 a, b，最坏的情况下需要验证 $\binom{13}{2} = 78$ 种不同的 a, b。而对每一个 a, b，计算其最大公因子，即使使用欧几里得算法，也需要做好几次运算。因此，完成题目的要求可能需要上千亿次运算！这就是数学证明的意义：我们可以把几千亿次的计算简化成几行推理。

证明. 利用素数表（或者直接计算）可知 2 到 40 之间的素数如下：$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37$，恰好有 12 个。记 Y 为这 12 个素数的集合。

记从 2 到 40 之间任意选取的 13 个数字的集合为 X。现在定义一个映射[3]如下：

$$f : X \to Y$$
$$k \mapsto \text{能整除 } k \text{ 的最小素数.}$$

[3]此处，每个 X 中的数字相当于一个鸽子，Y 中的每个素数相当于一个笼子。

例如，如果 $17, 35 \in X$，则 $f(17) = 17$，$f(35) = 5$，等等。注意：因为对任意的 $k \in X$，都有 $2 \le k \le 40$，所以能整除 k 的最小素数必定是 Y 中的某一个素数，所以 f 的定义是合理的。

因为 $|X| = 13$ 且 $|Y| = 12$，所以 $|X| > |Y|$。根据鸽笼原理，必定存在两个数 $a, b \in X$ 且 $a \neq b$ 满足 $f(a) = f(b)$。设 $f(a) = f(b) = t$，则因为 $t \in Y$，所以 $t \ge 2$。又根据 f 的定义，有 $t \mid a$ 且 $t \mid b$，所以，$(a, b) \ge t \ge 2$。 □

回忆鸽笼原理说如果我们要把 11 个鸽子放进 10 个笼中，那么必定有某个笼里要放至少两个鸽子。那么如果我们要把 21 个鸽子放在 10 个笼中呢？想一想，就会发现，必定有某个笼里要放至少三个鸽子。类似地，如果我们要把 31 个鸽子放在 10 个笼中，则必定有某个笼里要放至少四个鸽子。等等。我们可以把这个思考总结为下面的定理。

定理 2.15 (鸽笼原理的推广形式)**.** *给定 X 和 Y 都是有限非空集合，并且 $|X| > |Y|$。那么，对任意映射 $f : X \to Y$，必定存在 $k = \left\lceil \frac{|X|}{|Y|} \right\rceil$ 个各不相同的 X 中的元素 $x_1, x_2, \ldots, x_k$ 满足 $f(x_1) = f(x_2) = \cdots = f(x_k)$。*

证明. 可仿照鸽笼原理的证法，略。 □

以后无论鸽笼原理还是其推广形式，我们统一称呼为鸽笼原理。

下面是一个有趣的应用。

例 2.16. *地球上任意 9 人，必有至少 6 人处于同一个半球（包含边界）。*

证明. 想像每个人是球面上的一点。过球面上任意两点可画一个大圆，把球面分为两个半球，剩下的 7 个点，根据鸽笼原理[4]，必有 4 个点处于同一半球。这 4 个点连同大圆上的那两个点，总共 6 个，处于同一半球（包含边界大圆）。 □

之前我们用分类讨论证明了拉姆齐定理，现在我们用鸽笼原理的推广形式来重新证明拉姆齐定理。

拉姆齐定理证明二. 任选一个人，用 X 表示这个人，其余的 5 人记为 $\{A, B, C, D, E\}$。定义如下映射

$$f : \{A, B, C, D, E\} \to \{0, 1\},$$

其定义为，如果 A 认识 X，则定义 $f(A) = 1$，如果 A 不认识 X，则定义 $f(A) = 0$。类似地定义 $f(B), f(C), f(D), f(E)$。

因为 $|\{A, B, C, D, E\}| = 5$ 且 $|\{0, 1\}| = 2$，根据鸽笼原理[5]，必有至少 3 人在 f 下有相同的像。

不失一般性，假设 $f(A) = f(B) = f(C) = 0$，即 A, B, C 都不认识 X。此时分两种情况：

- A, B, C 相互都认识，此时结论成立。
- A, B, C 不是相互都认识，那么必有至少两人不认识。*不失一般性，假设* A, B 不认识，则 A, B, X 三人相互不认识，结论成立。 □

比较拉姆齐定理的两个证明，可以看出实质上非常相似。但是通过使用 “不失一般性”，证明的推理简化了，书写也简化了。

[4]这里 7 个点相当于鸽子，两个半球相当于笼子。

[5]这里 A, B, C, D, E 相当于鸽子，$0, 1$ 相当于笼子。

2.7 英文单词

- 证明: proof
- 直接证明: direct proof
- 简接证明: indirect proof
- 反证法: proof by contradiction
- 分类讨论法: case analysis
- 数学归纳法: mathematical induction
- 基础情形: base case
- 归纳假设: induction hypothesis
- 无穷：infinite
- 拉姆齐理论：Ramsey theory
- 鸽笼原理：pigeonhole principle
- 强归纳法：strong induction
- 对称性：symmetry
- 不失一般性：without loss of generality

第三章　计数

> 有的无穷比别的无穷更大。(*Some infinites are bigger than other infinites.*)
>
> ——*G. Cantor*（康托尔）
>
> 提出恰当的问题比回答它更难。(*To ask the right question is harder than to answer it.*)
>
> ——*G. Cantor*（康托尔）

数我们关心的对象有多少个，称作计数，这是最古老也是最重要的问题。人类早期就懂得结绳记事。今天，计数问题尤其重要。比如：一个算法运行需要多少时间？需要多少内存？传输一张图片，对图片能进行多少压缩？从 A 城去 B 城，共有多少不同的路线？等等。这些都与计数问题有关。本章我们学习计数方面的一些基础知识。

3.1　组合数的大小

设 n, m 都是自然数。在中学我们学过排列与组合。给定 n 个对象，其不同的排列个数是 $n! = n \times (n-1) \times \cdots \times 1$，从 n 个对象种选取 m 个对象的不同选取方法共有 $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ 这么多种。注意如下特殊情形和等式：

- $0! = 1$；
- $\binom{n}{0} = 1$；
- $\binom{n}{m} + \binom{n}{m+1} = \binom{n+1}{m+1}$；
- $\binom{n}{m} = \binom{n}{n-m}$；
- 如果 $m > n$，那么 $\binom{n}{m} = 0$。

3.1.1　二项式定理

我们也学习过如下等式：$(x+y)^2 = x^2 + 2xy + y^2$ 以及 $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$。一般地，有如下二项式定理。

定理 3.1 (二项式定理)**.** $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i$.

不难看出，我们可以尝试用归纳法来证明。

证明. 基础情形：$n=1$，左边等于 $x+y$，右边计算如下，

$$\binom{1}{0}x^{1-0}y^0+\binom{1}{1}x^{1-1}y^1=x+y.$$

归纳假设：设定理对 $n=k$ 成立。

现证明定理对 $n=k+1$ 成立。直接计算，并根据归纳假设有，如下：

$$(x+y)^{k+1}=(x+y)(x+y)^k=(x+y)\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^i. \tag{3.1}$$

现在对此式展开，继续计算，

$$\begin{aligned}(x+y)\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^i&=\sum_{i=0}^{k}\binom{k}{i}x^{k+1-i}y^i+\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^{i+1}\\&=\binom{k}{0}x^{k+1-0}y^0+\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i+\sum_{i=0}^{k-1}\binom{k}{i}x^{k-i}y^{i+1}+\binom{k}{k}x^{k-k}y^{k+1}\\&=x^{k+1}+\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i+\sum_{i=0}^{k-1}\binom{k}{i}x^{k-i}y^{i+1}+y^{k+1}.\end{aligned} \tag{3.2}$$

对第三项做一个变量替换：设 $j=i+1$，等价地，$i=j-1$，则因为第三项的求和是 $i=0,\dots,k-1$；因此 $j=1,\dots,k$。于是，对第三项有，

$$\sum_{i=0}^{k-1}\binom{k}{i}x^{k-i}y^{i+1}=\sum_{j=1}^{k}\binom{k}{j-1}x^{k-(j-1)}y^{(j-1)+1}=\sum_{j=1}^{k}\binom{k}{j-1}x^{k+1-j}y^j. \tag{3.3}$$

现在继续计算，则有，

$$\begin{aligned}\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i+\sum_{i=0}^{k-1}\binom{k}{i}x^{k-i}y^{i+1}&=\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i+\sum_{j=1}^{k}\binom{k}{j-1}x^{k+1-j}y^j\\&=\sum_{i=1}^{k}\binom{k}{i}x^{k+1-i}y^i+\sum_{i=1}^{k}\binom{k}{i-1}x^{k+1-i}y^i\\&=\sum_{i=1}^{k}\left(\binom{k}{i}+\binom{k}{i-1}\right)x^{k+1-i}y^i\\&=\sum_{i=1}^{k}\binom{k+1}{i}x^{k+1-i}y^i.\end{aligned} \tag{3.4}$$

结合(3.2)及(3.4)，得到

$$\begin{aligned}(x+y)\sum_{i=0}^{k}\binom{k}{i}x^{k-i}y^i&=x^{k+1}+\sum_{i=1}^{k}\binom{k+1}{i}x^{k+1-i}y^i+y^{k+1}\\&=\sum_{i=0}^{k+1}\binom{k+1}{i}x^{k+1-i}y^i.\end{aligned} \tag{3.5}$$

注意(3.5)的计算结果就是定理当 $n=k+1$ 时的右边，因此定理就证明完毕。 □

二项式定理是数学中一个重要的基本定理。$\binom{n}{i}$ 也称为二项式系数。

3.1.2 组合数的大小估计

在计数的过程中，当我们用一个公式表达了计数的结果之后，我们通常希望计算或估计这个公式所表达数字的大小。一般地，我们希望知道，比如，这个公式表达的值是一个常数大小，是线性大小，还是多项式大小，或者是指数级大小?

假设我们用公式 $f(n)$ 表达了一个计数的结果（当然，因为 $f(n)$ 是代表一个计数的结果，因此 $f(n) \geq 0$)。设有另外一个函数 $g(n)$。为了准确地描述这些数量，我们介绍如下的数学记号。

- $f(n) \leq O\big(g(n)\big)$：存在常数 $c > 0$，使得对任意的 n，$f(n) \leq c \cdot g(n)$ 成立；
- $f(n) \geq \Omega\big(g(n)\big)$：存在常数 $c > 0$，使得对任意的 n，$f(n) \geq c \cdot g(n)$ 成立；
- $f(n) = \Theta\big(g(n)\big)$：指 $f(n) = O\big(g(n)\big)$ 及 $f(n) = \Omega\big(g(n)\big)$ 同时成立；
- $f(n) = o\big(g(n)\big)$：指 $\lim_{n\to\infty} \frac{f(n)}{g(b)} = 0$。

例 3.2.

- $\frac{(n-1)n}{2} \leq O(n^2)$；
- $2^n - n^{74} + \sqrt{n} - 48 \geq \Omega\big(2^n\big)$；
- $\frac{(n-1)n}{2} = \Theta(n^2)$；
- $\sqrt{n} + n^{1/3} + 48 = o\big(n\big)$；
- $\sum_{n=1}^{\infty} \frac{1}{n^2} \leq O(1)$。

运用这些符号，我们就可以说 $O(1), O(n), O(n^{23})$ 分别最多是常数级大小，线性大小，多项式大小，等等；而 $\Omega(n^3), \Omega(2^n)$ 分别最少是三次多项式大小，指数级大小，等等。

很多时候，当我们有了公式 $f(n)$，我们往往需要对 $f(n)$ 的上界和下界进行估计。我们先来估计阶乘的大小。

命题 3.3. $\left(\frac{n}{e}\right)^n \leq n! \leq (n+1) \cdot \left(\frac{n}{e}\right)^n \leq O\left(n \cdot \left(\frac{n}{e}\right)^n\right)$.

证明. 根据定义，$n! = n \times (n-1) \times \cdots \times 1$，取对数，得

$$\ln n! = \ln 1 + \ln 2 + \cdots + \ln n.$$

根据积分的定义，有

$$\int_1^n \ln x dx \leq \ln 1 + \ln 2 + \cdots + \ln n \leq \int_1^{n+1} \ln x dx.$$

计算积分可得，

$$\int_1^n \ln x dx = [x \ln x - x]|_1^n = n \ln n - n + 1 \geq \ln n^n - n,$$

及

$$\int_1^{n+1} \ln x dx = [x \ln x - x]|_1^{n+1} = (n+1)\ln(n+1) - n = \ln(n+1)^{n+1} - n.$$

因此，

$$\left(\frac{n}{e}\right)^n \le n! \le \frac{(n+1)^{n+1}}{e^n} = \frac{(n+1)^{n+1}}{e^n} \cdot \frac{n^n}{n^n} \le (n+1) \cdot \left(\frac{n}{e}\right)^n. \qquad \square$$

根据微积分中的 Stirling 公式可以得到更准确的估计，$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

下面再讨论组合数的大小。根据二项式定理，有

$$2^n = (1+1)^n = \sum_{i=0}^{n} \binom{n}{i}.$$

因此，对任意的 $0 \le m \le n$，都有，$\binom{n}{m} \le 2^n$. 因此，我们得到了一个上界。那么下界是多大呢？如果 m 是常数，那么

$$\binom{n}{0} = 1 = \Theta(1), \quad \binom{n}{1} = n = \Theta(n), \quad \binom{n}{2} = \frac{n(n-1)}{2} = \Theta(n^2), \quad \ldots$$

一般地，如果 $m = O(1)$（即，m 是一个常数），那么就有 $\binom{n}{m} = \Theta(n^m)$. 下面考虑 m 不是常数的情况。

命题 3.4. 如果 $1 \le m \le n/2$，则 $\binom{n}{m-1} \le \binom{n}{m}$.

证明. 直接计算。 $\square$

命题3.4表明：当 n 是偶数时，$\binom{n}{n/2}$ 是最大的二项式系数，当 n 是奇数时，$\binom{n}{\frac{n-1}{2}} = \binom{n}{\frac{n+1}{2}}$ 是最大的二项式系数。

以下我们讨论 n 是偶数的情况。我们先看 $\binom{n}{n/2}$。

验证你的直觉： 猜猜 $\binom{n}{n/2}$ 大概是多大呢?
(A) $\Theta(n^{n/2})$;　(B) $\Theta(2^{n/2})$;　(C) $\Theta(2^n)$;　(D) $\Theta((n/2)^{n/2})$.

首先我们尝试用命题3.3。根据定义，则有（假设 n 是偶数）

$$\binom{n}{n/2} = \frac{n!}{(n/2)! \cdot (n/2)!} \le \frac{O\left(n \cdot \left(\frac{n}{e}\right)^n\right)}{\left((n/2e)^{n/2}\right)^2} = \frac{O\left(n \cdot \left(\frac{n}{e}\right)^n\right)}{(n/2e)^n} = O(n \cdot 2^n).$$

这个上界没有任何用处，因为我们已经知道 $\binom{n}{n/2} \le 2^n$.

再考虑下界，仍然用命题3.3。先计算

$$(n/2)! \cdot (n/2)! \le O\left(\left(\frac{n}{2} \cdot (\frac{n}{2e})^{n/2}\right)^2\right) \le O(n^2 \cdot (\frac{n}{2e})^n).$$

于是，有

$$\binom{n}{n/2} = \frac{n!}{(n/2)! \cdot (n/2)!} \ge \frac{(n/e)^n}{O(n^2 \cdot (\frac{n}{2e})^n)} \ge \Omega(2^n/n^2).$$

这样我们得到了一个有意义的下界

$$\binom{n}{n/2} \ge \Omega(2^n/n^2).$$

换一种方法，下面我们将得到更好的下界，还可以得到上界。

引理 3.5. 对任意 $k \geq 1$ 有，$\frac{k-1}{k} \cdot \frac{k}{k+1} \leq (\frac{k}{k+1})^2 \leq \frac{k}{k+1} \cdot \frac{k+1}{k+2}$.

命题 3.6. 设 n 是偶数，则 $2^n/\sqrt{2n} \leq \binom{n}{n/2} \leq 2^n/\sqrt{n}$，即 $\binom{n}{n/2} = \Theta\big(2^n/\sqrt{n}\big)$.

证明. 根据定义，

$$\frac{1}{2^n} \cdot \binom{n}{n/2} = \frac{1 \cdot 2 \cdot 3 \cdot \cdots \cdot n}{(2 \cdot 4 \cdot 6 \cdot \cdots \cdot n) \cdot (2 \cdot 4 \cdot 6 \cdot \cdots \cdot n)} = \frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (n-1)}{2 \cdot 4 \cdot 6 \cdot \cdots \cdot n}. \tag{3.6}$$

我们用引理3.5中的不等式，则有

$$\left(\frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (n-1)}{2 \cdot 4 \cdot 6 \cdot \cdots \cdot n}\right)^2 \leq \frac{1 \cdot 2}{2 \cdot 3} \cdot \frac{3 \cdot 4}{4 \cdot 5} \cdot \cdots \cdot \frac{(n-1) \cdot n}{n \cdot (n+1)} = \frac{1}{n+1} \leq \frac{1}{n}, \tag{3.7}$$

及

$$\left(\frac{1 \cdot 3 \cdot 5 \cdot \cdots \cdot (n-1)}{2 \cdot 4 \cdot 6 \cdot \cdots \cdot n}\right)^2 \geq \frac{1 \cdot 1}{2 \cdot 2} \cdot \frac{2 \cdot 3}{3 \cdot 4} \cdot \cdots \cdot \frac{(n-2) \cdot (n-1)}{(n-1) \cdot n} = \frac{1}{2n}. \tag{3.8}$$

根据(3.6)，(3.7)，(3.8)，即得所证。 □

对一般的 m 呢? 比如 $m = n/3$，那么 $\binom{n}{n/3}$ 有多大呢? 我们给出如下估计，不加证明。定义熵函数

$$h : [0,1] \to [0,1], \quad x \mapsto h(x) = -x \log_2 x - (1-x) \log_2 (1-x).$$

命题 3.7. 如果 $m = xn$ 且 $0 \leq x \leq 1/2$，则有

$$\Omega\left(\frac{2^{h(x)n}}{n}\right) \leq \frac{2^{h(x)n}}{n+1} \leq \binom{n}{m} \leq 2^{h(x)n}.$$

而且有，

$$\sum_{i=0}^{m} \binom{n}{i} \leq 2^{h(x)n}.$$

3.2 集合，关系，与函数

3.2.1 集合

我们简单回顾集合的基本概念和运算。

把一些东西放在一起，就组成了一个集合，那每一个东西叫做一个元素。

比如自然数集 $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$. 素数集,$\mathbb{P} = \{2, 3, 5, 7, 11, \ldots\}$. 整数集 $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}$. $\mathbb{R}$ 表示实数集等。

设 A 和 B 是两个集合，a 是某个元素。集合的一般描述方式为

$$A = \{a : a\text{满足某种性质}\}.$$

基本运算:

- 元素属于一个集合: $a \in A$;

- 元素不属于一个集合：$a \notin A$；
- 子集，即一个集合的元素全部在另一个集合里面：$B \subseteq A$；真子集：$B \subsetneq A$；
- 两个集合的交集：$A \cap B = \{x : x \in A \text{ and } x \in B\}$；
- 两个集合的并集：$A \cup B = \{x : x \in A \text{ or } x \in B\}$；
- 两个集合的差集：$A \setminus B = A - B = \{x : x \in A \text{ and } x \notin B\}$.
- 补集，如果 $B \subseteq A$，则 B 在 A 种的补集定义为：$B^c = A - B$.
- 两个集合的对称差集：$A \Delta B = (A - B) \cup (B - A)$；
- 空集：$\emptyset = \{\}$；
- 笛卡尔乘积：$A \times B = \{(a, b) : a \in A, b \in B\}$；
- 幂集，由某个集合的所有子集组成的集合：$\mathcal{P}(A)$ 或 2^A，即 $\mathcal{P}(A) = \{B : B \subseteq A\}$；
- 集合的势（或者叫集合的大小），即集合中元素的个数：$|A|$。

使用韦恩图往往可以帮助对集合进行运算（见《离散数学》p.97）。
下面看一些例子和性质。

- 交换律：$A \cup B = B \cup A$；$A \cap B = B \cap A$；
- 结合律：$(A \cup B) \cup C = A \cup (B \cup C)$；$(A \cap B) \cap C = A \cap (B \cap C)$；
- 分配律：$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$；$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$；
- 笛卡尔乘积 $A \times B$ 里的元素 (a, b) 是一个有序对，即，a 与 b 在这个地方的顺序是重要的。因此，一般来说，$A \times B \neq B \times A$。换句话说，笛卡尔乘积不具有交换律。例如：$A = \{1, 2\}$，$B = \{2, 3, 4\}$，则

$$A \times B = \{(1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4)\}$$
$$B \times A = \{(2, 1), (2, 2), (3, 1), (3, 2), (4, 1), (4, 2)\}.$$

  因此，$A \times B \neq B \times A$。不过，当 A 和 B 都是有限集时，仍然有 $|A \times B| = |B \times A| = |A| \cdot |B|$.
- 空集：注意：$\emptyset \neq \{\emptyset\}$. $|\emptyset| = 0$，$|\{\emptyset\}| = 1$.
- 幂集的例子。例如：$A = \{1, 2\}$，则

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

  基本性质：$|\mathcal{P}(A)| = 2^{|A|}$.
- $\{1\} \neq \{\{1\}\}$.

3.2.2 关系

定义 3.8. 从 A 到 B 的二元关系：笛卡尔积 $A \times B$ 的任意一个子集都叫做一个二元关系。

关系在生活与应用中无处不在。

例 3.9. 前面拉姆齐定理中人与人的认识关系，就是一个二元关系。比如

$$A = \{\text{刘备}, \text{关羽}, \text{张飞}, \text{孙悟空}, \text{猪八戒}, \text{沙和尚}, \text{林冲}, \text{武松}, \text{鲁智深}, \text{贾宝玉}, \text{林黛玉}, \text{薛宝钗}\}.$$

如果我们考虑集合 A 里的人的“认识”这种关系，就可以用 $A \times A$ 上的二元关系来描述。比如 (林冲, 武松) 和 (林黛玉, 薛宝钗) 都属于这个二元关系。但是 (张飞, 孙悟空) 就不属于这个二元关系。问：这个二元关系是一个从 A 到 A 的函数吗？

数学中的许多常见关系也可以用二元关系来描述。比如，整数集 $\mathbb{Z}$ 上的数字大小关系：

$$\{(a, b) \in \mathbb{Z} \times \mathbb{Z} : a \leq b\}.$$

任意一个集合 A 里的子集的包含关系：

$$\{(B, C) \in \mathcal{P}(A) \times \mathcal{P}(A) : B \subseteq C\}.$$

等等。

某些数学对象其实也是一个二元关系。比如：圆。我们知道，单位圆的方程是 $x^2 + y^2 = 1$. 因此，单位圆也可以如下描述：

$$\{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}.$$

如果某个问题涉及到一系列集合 $A_1, A_2, \ldots, A_k$，则它们笛卡尔积 $A_1 \times A_2 \times \cdots \times A_k$ 的子集就叫做 k-元关系。比如，数据库的数据集，或者机器学习中的数据集，都往往涉及到许多不同的 features，因此，那里的数据集，一般都是一个多元关系。比如 Iris flower dataset （鸢尾花数据集）。

3.2.3 函数

定义 3.10. 从 A 到 B 的函数 (映射)：函数是一种特殊的关系。要求满足如下条件：对每一个 $a \in A$，都有且仅有唯一的 $b \in B$，满足 (a, b) 在这个二元关系中。

下面简单复习函数 $f : A \to B$ 的基本概念。

- 定义域：A；
- 设 $a \in A$，则 $f(a)$ 叫做 a 在函数 f 下在 B 中的像；
- 像集：$f(A) = \{f(a) : a \in A\}$，即所有像组成的集合；当然有 $f(A) \subseteq B$ 成立，但是 $f(A)$ 不一定等于 B；
- f 是单射：$a, a' \in A$，若 $a \neq a'$ 则 $f(a) \neq f(a')$；
- f 是满射：$f(A) = B$；

- f 是双射：既是单射，又是满射。当 f 是双射时，可以定义逆函数：

$$f^{-1}: B \to A, \quad b \mapsto a,$$

满足 $f(a) = b$.

- 复合函数：若 $f: A \to B$，且 $g: B \to C$，则可以定义复合函数

$$g \circ f: A \to C, \quad a \mapsto c,$$

其中 $c = g(b)$ 而 $b = f(a)$. 注意：给定 a 后，b 和 c 都唯一确定，因此复合函数 $g \circ f$ 的定义是正确的。

- 函数的相等：两个函数 $f: A \to B$ 与 $g: A \to B$ 相等，当且仅当，对每一个 $a \in A$，都有 $f(a) = g(a)$。

- 当 f 是双射时，有：

$$f^{-1} \circ f: A \to A, \quad a \mapsto a.$$

而且

$$f \circ f^{-1}: A \to A, \quad a \mapsto a.$$

所以此时，$f^{-1} \circ f = f \circ f^{-1}$.

- 给定集合 A 和 B，从 A 到 B 的所有函数的集合一般记作 B^A.

例 3.11. 用符号 $[n]$ 代表集合 $[n] = \{1, 2, \ldots, n\}$。问：$\{0,1\}^{[3]}$ 是什么？

根据定义，这是从 $[3]$ 到 $\{0,1\}$ 的所有函数的集合。具体写出来如下：

$$\begin{aligned}
&f_0: [3] \mapsto \{0,1\}, \quad 1 \mapsto 0, \quad 2 \mapsto 0, \quad 3 \mapsto 0;\\
&f_1: [3] \mapsto \{0,1\}, \quad 1 \mapsto 0, \quad 2 \mapsto 0, \quad 3 \mapsto 1;\\
&f_2: [3] \mapsto \{0,1\}, \quad 1 \mapsto 0, \quad 2 \mapsto 1, \quad 3 \mapsto 0;\\
&f_3: [3] \mapsto \{0,1\}, \quad 1 \mapsto 0, \quad 2 \mapsto 1, \quad 3 \mapsto 1;\\
&f_4: [3] \mapsto \{0,1\}, \quad 1 \mapsto 1, \quad 2 \mapsto 0, \quad 3 \mapsto 0;\\
&f_5: [3] \mapsto \{0,1\}, \quad 1 \mapsto 1, \quad 2 \mapsto 0, \quad 3 \mapsto 1;\\
&f_6: [3] \mapsto \{0,1\}, \quad 1 \mapsto 1, \quad 2 \mapsto 1, \quad 3 \mapsto 0;\\
&f_7: [3] \mapsto \{0,1\}, \quad 1 \mapsto 1, \quad 2 \mapsto 1, \quad 3 \mapsto 1.
\end{aligned}$$

可以看出，其实，$\{0,1\}^{[3]}$ 这个从 $[3]$ 到 $\{0,1\}$ 的所有函数的集合，就是所有长度是 3 的 0、1 字符串的集合。因此，$\{0,1\}^{[n]}$ 常常写成 $\{0,1\}^n$. 容易看出 $|\{0,1\}^n| = 2^n$.

一般地，如果 A 与 B 都是有限集，则有 $|B^A| = |B|^{|A|}$. 这是因为，函数 $f: A \to B$ 其实就是对每个 $a \in A$，选取一个 $b \in B$ 与 a 对应。因为对于每个 $a \in A$ 都有 $|B|$ 种不同的选法，因此，总共不同的函数个数就是 $|A|$ 个 $|B|$ 相乘：$|B| \times \cdots \times |B| = |B|^{|A|}$. 从这里可以明白，为什么从 A 到 B 的所有函数的集合记号用 B^A，而不用 A^B.

如上我们讨论的都是一元函数 $f : A \times B$，即变量只有一个的函数。某些问题有多个变量，就需要考虑多元函数 $f : A_1 \times \cdots \times A_k \times B$. 从集合与函数的观点来看，一元函数与多元函数没有本质的区别，因为定义域不管是 A，还是 $A_1 \times \cdots \times A_k$，都是一个集合而已（后者是 k 个集合的笛卡尔积，但仍然是一个集合）。所以，数学抽象的意义之一就在于，一些看起来不同的对象，在抽象的观点下，都是同一类对象。当然，在具体分析函数时，多元函数的分析往往复杂得多。

3.3 计数的基本方法

前面学过的鸽笼原理，也可以看作一种基本的计数方法。不过鸽笼原理一般只用于得到某种下界（即：某个笼子种至少有多少只鸽子）。如果还要得到上界，或者精确的计算，有时需要别的方法。本节介绍一些最基本的方法。

3.3.1 计数方法一：容斥原理

在计数时，如果问题涉及好几个集合，并且这些集合都是有限集，则往往可以使用容斥原理。比如，如下最基本的问题：如果知道了集合 A 和集合 B 的大小，那么 $A \cup B$ 的大小是多少呢？$A \cap B$ 的大小是多少呢？稍加思考（或观察韦恩图），可得如下等式：

$$|A| + |B| = |A \cup B| + |A \cap B|.$$

等价地，

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

类似地，如果有三个集合 A, B, C，那么（通过观察韦恩图可得）

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

定理 3.12 (容斥原理)**.** 设有 n 个有限集 $A_1, A_2, \ldots, A_n$，则

$$\begin{aligned} |\cup_{i=1}^{n} A_i| = &\sum_{i=1}^{n} |A_i| - \sum_{1 \le i < j \le n} |A_i \cap A_j| + \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| \\ &+ (-1)^{t-1} \sum_{1 \le i_1 < i_2 < \cdots < i_t \le n} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_t}| \\ &+ (-1)^{n-1} |A_1 \cap A_2 \cap \cdots \cap A_n|. \end{aligned} \tag{3.9}$$

下面看几个例子。

例 3.13. 计算 $\{1, 2, \ldots, 100\}$1 里不能被 2、3 或 5 整除的数的个数。

解答. 设 $A = \{1, 2, \ldots, 100\}$. 令

$$\begin{aligned} B &= \{x \in A : 2 \mid x\}, \quad |B| = 50, \\ C &= \{x \in A : 3 \mid x\}, \quad |C| = 33, \\ D &= \{x \in A : 5 \mid x\}, \quad |D| = 20, \end{aligned}$$

所求的答案是 $|B \cup C \cup D|$. 根据容斥原理，有

$$|B \cup C \cup D| = |B| + |C| + |D| - |B \cap C| - |B \cap D| - |C \cap D| + |B \cap C \cap D|.$$

注意有，

$$\begin{aligned} B \cap C &= \{x \in A : 6 \mid x\}, \quad |B \cap C| = 16, \\ B \cap D &= \{x \in A : 10 \mid x\}, \quad |B \cap D| = 10, \\ C \cap D &= \{x \in A : 15 \mid x\}, \quad |C \cap D| = 6, \\ B \cap C \cap D &= \{x \in A : 30 \mid x\}, \quad |C \cap D| = 3. \end{aligned}$$

因此，$|B \cup C \cup D| = 50 + 33 + 20 - 16 - 10 - 6 + 3 = 26$. □

下面是一个运用容斥原理的经典的例子。

例 3.14 (错位排列)**.** 设一副牌有 n 张，每张牌上标有一个 1 到 n 之间不同的数字。洗几次牌后，如果对每一个 i，第 i 张牌的标号都不是 i，我们就说牌形成了一个错位排列，换句话说：每张牌的位置都是错的。问：在所有可能的牌的排列中，每张牌的位置都是错的的可能性是多少？

验证你的直觉： 每张牌的位置都是错的的可能性是多少?
(A) 小于 50%　　(B) 大于 50%.

解答. 用 P_i 表示标号为 i 的牌在正确的位置（即在第 i 个位置）的所有排列的集合。那么，至少有一张牌在正确的位置的所有排列数是：$|\cup_{i=1}^n P_i|$。故，每张牌的位置都是错的的排列数是：$n! - |\cup_{i=1}^n P_i|$。

利用容斥原理3.12计算 $|\cup_{i=1}^n P_i|$。我们看其中的一般项为

$$(-1)^{t-1} \sum_{1 \le i_1 < i_2 < \cdots < i_t \le n} |P_{i_1} \cap P_{i_2} \cap \cdots \cap P_{i_t}|. \tag{3.10}$$

其中，$P_{i_1} \cap P_{i_2} \cap \cdots \cap P_{i_t}$ 表示第 $i_1, i_2, \ldots, i_t$ 张牌都在正确的位置。由对称性，$|P_{i_1} \cap P_{i_2} \cap \cdots \cap P_{i_t}|$ 不依赖于 $i_1, i_2, \ldots, i_t$ 的值[1]。易见：

$$|P_{i_1} \cap P_{i_2} \cap \cdots \cap P_{i_t}| = (n-t)!$$

又注意到，(3.10)中一共有 $\binom{n}{t}$ 项。因此，

$$(-1)^{t-1} \sum_{1 \le i_1 < i_2 < \cdots < i_t \le n} |P_{i_1} \cap P_{i_2} \cap \cdots \cap P_{i_t}| = (-1)^{t-1} \binom{n}{t} (n-t)! = (-1)^{t-1} \frac{n!}{t!}. \tag{3.11}$$

因此，用容斥原理公式(3.9)得，

$$|\cup_{i=1}^n P_i| = \sum_{t=1}^{n} (-1)^{t-1} \frac{n!}{t!}$$

[1]比如：第 1、3、27 张牌在正确位置的排列个数，和第 7、15、19 张牌在正确位置的排列个数，两者相等。

从而得到，每张牌都在错误位置的可能性是，

$$\frac{n! - |\cup_{i=1}^{n} P_i|}{n!} = \frac{n! - \sum_{t=1}^{n}(-1)^{t-1}\frac{n!}{t!}}{n!} = 1 - \sum_{t=1}^{n}(-1)^{t-1}\frac{1}{t!} = \sum_{t=0}^{n}(-1)^{t}\frac{1}{t!}.$$

根据 e^x 的泰勒展开式，$e^x = \sum_{t=0}^{\infty}\frac{x^t}{t!}$，有：当 $n \to \infty$ 时，$\lim_{n\to\infty}\sum_{t=0}^{n}\frac{(-1)^t}{t!} = e^{-1} \approx 37\%$. □

你猜对了吗?

3.3.2 计数方法二：双计数

双计数就如同生活中我们看问题时，站在不同的角度，就能看到同一事物的不同方面。生活中，尝试用另一个角度看问题，往往能有所收获。计数也是如此。

双计数操作如下：对一个恰当的（往往是二元的）集合用两种不同的方式计数。下面看两个例子[2]。

例 3.15. *正 20 面体（每个面都是三角形）有多少条边？*

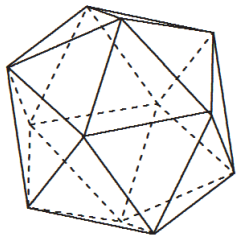

图 3.1: 正 20 面体。

当然我们可以找一个正 20 面体，如图3.1，去数它的边，但有没有不用那么费劲的方法呢?

解答. 我们要计数的对象是边的个数，考虑与边密切联系的面。因此，令

$$A = \text{正 20 面体所有边的集合}, \quad B = \text{正 20 面体所有面的集合}.$$

我们知道 $|B| = 20$，想问 $|A| =$?

考虑如下集合

$$C = \{(e, f) : e \in A, f \in B, e \in f\} \subseteq A \times B.$$

也就是说，我们考虑（边，面）这样的二元对，满足边在面上。

用两种观点来看 C 的大小 $|C|$。

- 从边的观点：每条边都在两个面上。因此，$|C| = 2|A|$.
- 从面的观点：因为每个面都是三角形，所以每个面都对应三条边。因此，$|C| = 3|B|$.

[2]在课堂上我们也将用图的语言来描述双计数

从而得

$$2|A| = |C| = 3|B|.$$

因此，$|A| = 3|B|/2 = 30$. □

例 3.16. 定义 $[n]^{(r)} = \{T \subseteq \{1, 2, \ldots, n\} : |T| = r\}$. 设 $\mathcal{A} \subseteq [n]^{(r)}$. 设 $s > r$。定义

$$\mathcal{B} = \{B \in [n]^{(s)} : \exists\ A \in \mathcal{A},\ s.t.\ A \subseteq B\}.$$

$|\mathcal{B}|$ 有多大呢（当然 $|\mathcal{B}|$ 依赖于 $|\mathcal{A}|$）？

解答. 上界：根据定义 $\mathcal{B} \subseteq [n]^{(s)}$，所以 $|\mathcal{B}| \leq \left|[n]^{(s)}\right| = \binom{n}{s}$.

下界：考虑如下集合

$$M = \{(A, B) \in \mathcal{A} \times \mathcal{B} : A \subseteq B\}.$$

分别采用 $\mathcal{A}$ 和 $\mathcal{B}$ 的观点来计算 $|M|$。

- 从 $\mathcal{A}$ 的观点：对每个 $A \in \mathcal{A}$，A 包含在 $\binom{n-r}{s-r}$ 个不同的 B 中。因此，$|M| = |\mathcal{A}|\binom{n-r}{s-r}$.
- 从 $\mathcal{B}$ 的观点：对每个 $B \in \mathcal{B}$，B 最多包含 $\binom{s}{r}$ 个不同的 A。因此，$|M| \leq |\mathcal{B}|\binom{s}{r}$.

从而有，

$$|\mathcal{A}|\binom{n-r}{s-r} = |M| \leq |\mathcal{B}|\binom{s}{r}.$$

所以，$|\mathcal{B}| \geq \frac{\binom{n-r}{s-r}}{\binom{s}{r}}|\mathcal{A}| = \frac{\binom{n}{s}}{\binom{n}{r}}|\mathcal{A}|$.

综合上、下界，我们得到

$$\frac{|\mathcal{A}|}{\binom{n}{r}}\binom{n}{s} \leq |\mathcal{B}| \leq \binom{n}{s}. \tag{3.12}$$

□

体会一下(3.12)中下界的直观含义。

3.3.3 计数方法三：构造映射

这一节我们用构造映射的方法来比较集合的大小。特别是对于无限集。

验证你的直觉： 下面哪些正确?

(A) $|\mathbb{N}| < |\mathbb{Z}| < |\mathbb{Q}| < |\mathbb{R}|$　　(B) $|\mathbb{Z}| = 2|\mathbb{N}|$　　(C) $|\mathbb{Q}| = |\mathbb{Z}|$　　(D) $|\mathbb{R}| > |(0, 1)|$

给定一个有限的集合，$|A|$ 就是 A 里的元素个数。如果 A 是无限的集合，我们还没有准确定义 $|A|$ 的含义。尽管如此，我们先使用 $|A|$ 来表示（无论有限还是无限）集合 A 的 “大小”。对有限的集合来说，如果 $A \subsetneq B$，那么 $|A| < |B|$. 但是对于无限的集合呢？我们如何比较无限的集合的大小?

命题 3.17. 设 A 和 B 是两个集合，则，

- 若存在某个单射函数 $f: A \to B$，则 $|B| \geq |A|$；
- 若存在某个满射函数 $f: A \to B$，则 $|A| \geq |B|$；
- 若存在某个双射函数 $f: A \to B$，则 $|A| = |B|$.

命题 3.18. *(1)* $|\mathbb{N}| = |\mathbb{Z}|$,

(2) $|\mathbb{N}| = |\mathbb{Q}|$,

(3) $|(-1,1)| = |\mathbb{R}|$,

(4) $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$.

证明. 分别构造如下的双射：

(1) 考虑映射

$$f: \mathbb{Z} \to \mathbb{N},$$
$$0 \mapsto 0, \quad 1 \mapsto 1, \quad -1 \mapsto 2, \quad 2 \mapsto 3, \quad -2 \mapsto 4, \quad \cdots$$

容易看出，f 是双射。

(2) 类似（1）可证。

(3) 函数 $f(x) = \tan(\pi x/2)$ 是双射.

(4) 考虑映射

$$f: \mathbb{N} \times \mathbb{N} \to \mathbb{N},$$
$$(0,0) \mapsto 0, \quad (0,1) \mapsto 1, \quad (1,0) \mapsto 2, \quad (0,2) \mapsto 3, \quad (1,1) \mapsto 4, \quad (2,0) \mapsto 5, \quad \cdots$$

容易看出，f 是双射。

□

命题3.18告诉我们，对于无穷集合的大小，不能简单地像有限集合那样，根据 $A \subsetneq B$，便得出 $|A| < |B|$。比如，上面已经证明，虽然 $\mathbb{N} \subsetneq \mathbb{Z}$，但却有 $|\mathbb{N}| = |\mathbb{Z}|$。那么，$\mathbb{R}$ 和 $\mathbb{N}$ 比较又如何呢？从十九世纪下半叶到二十世纪初期，数学家康托尔在研究这个问题的过程中，发明了如下著名的对角线构造方法。

命题 3.19. *(1)* $|\mathbb{N}| < |[0,1]|$,

(2) 对任意集合 A，都有 $|A| < |\mathcal{P}(A)|$.

证明. 我们仅证明（1），(2) 可以类似地证明。

首先，容易看出 $|\mathbb{N}| \leq |[0,1]|$. 因为可以构造单射函数 $g: \mathbb{N} \to [0,1]$，定义为 $g(0) = 0$，对 $k \geq 1$，$g(k) = 1/k$。

现在用反证法来证明（1）。假设（1）不成立，那么因为我们已证明 $|\mathbb{N}| \le |[0,1]|$ 始终成立，故，若（1）不成立，则必有 $|\mathbb{N}| = |[0,1]|$. 所以，$[0,1]$ 与 $|\mathbb{N}|$ 存在双射。设 f 是一个双射如下

$$f : \mathbb{N} \to [0,1]$$
$$0 \mapsto x_1, \quad 1 \mapsto x_2, \quad 2 \mapsto x_3, \quad 3 \mapsto x_4, \quad \cdots$$

换句话说，$[0,1]$ 可以写成

$$[0,1] = \{x_1, x_2, x_3, x_4, \cdots\}. \tag{3.13}$$

我们如下来表示 x_k:

$$x_k = 0.x_{k1}x_{k2}x_{k3}x_{k4}\cdots$$

其中 $x_{ki} \in \{0, 1, \ldots, 9\}$. 现在构造如下表格

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{44} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

我们关注其对角线，并考察数字 y 如下

$$y = 0.y_1y_2y_3y_4\cdots$$

满足

$$y_1 \ne x_{11}, \quad y_2 \ne x_{22}, \quad y_3 \ne x_{33}, \quad y_4 \ne x_{44}, \quad \cdots$$

显然，$y \in [0,1]$。然而，根据构造，对所有的 $k \in \mathbb{N}$ 都有 $y \ne x_k$. 因此，根据(3.13)又有 $y \notin [0,1]$，这与 $y \in [0,1]$ 矛盾。 □

对角线构造法 + 反证法，是离散数学和计算机科学中一个很基本的方法。

根据命题3.19就得到如下推论。

推论 3.20. *(1)* $|\mathbb{N}| < |\mathbb{R}|$.

(2) $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| < |\mathcal{P}(\mathcal{P}(\mathbb{N}))| < \cdots$.

证明. 对（1），由 $[0,1] \subseteq \mathbb{R}$ 可得 $|[0,1]| \le |\mathbb{R}|$. 因此，由命题3.19得 $|\mathbb{N}| < |[0,1]| \le |\mathbb{R}|$.

对（2），在命题3.19的（2）中令 $A = \mathbb{N}$ 则得到 $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$. 再令 $A = \mathcal{P}(\mathbb{N})$ 则得到 $|\mathcal{P}(\mathbb{N})| < |\mathcal{P}(\mathcal{P}(\mathbb{N}))|$. 等等。 □

特别地，推论3.20种的（2）告诉我们：有无穷多种不同的 “无穷大”。

根据上面对集合大小的讨论，可以总结成如下的定义。

定义 3.21. 一个集合 A 如果满足 $|A| \le |\mathbb{N}|$，则 A 叫做是可数集，否则，A 叫做不可数集.

如上我们证明了，有理数，整数，及 $\mathbb{N}\times\mathbb{N}$ 等都是可数集。而实数集或任何开区间 (a,b) 或闭区间 $[a,b]$ 均是不可数集。实数集的大小，一般用 $|\mathbb{R}|=\aleph_0$ 表示。特别地，实数集包含的数字比自然数“多”，即 $|\mathbb{R}|=\aleph_0>|\mathbb{N}|$，而有理数则和自然数“一样多”。因此，这也就说明：无理数比有理数“多”。数学后来进一步严格地发展了集合及测度等概念，就可以进一步说明，无理数不仅比有理数多，而且“多得多”。

总之，以康托尔为代表的数学家对无限这一概念的深入研究，大大加深了人类对无限的理性认识。这是人类理性认识自然的一个里程碑。

3.4 英文单词

- 二项式定理：binomial theorem
- 熵：entropy
- 集合：set
- 元素：element
- 子集，真子集：subset, proper subset
- 交集：intersection
- 并集：union
- 差集：difference set
- 补集：complement
- 对称差集：symmetric difference
- 空集：empty set
- 笛卡尔乘积: Descartes product
- 幂集：power set
- 韦恩图：venn diagram
- 容斥原理：inclusion–exclusion principle
- 错位排列：derangement
- 双计数：double counting (or, counting in two ways)
- 交换：commutative
- 结合：associative

- 分配：distributive
- 关系：relation
- 有序对：ordered pair
- 单射、满射、双射：injective, surjective, bijection
- 康托尔：Cantor
- 对角线方法：diagonalization
- 可数：countable
- 不可数：uncountable

第四章　图论

> 读读欧拉，读读欧拉，他是我们大家的老师。 ——拉普拉斯
>
> 世间一切事物的意义，无不在于追求某种最大或最小。 ——欧拉

解决实际问题往往需要一个恰当的数学模型。甚至解决有些数学问题，也需要有一个恰当的模型或观点。比如：假设要证明三角形的三条高线相交于同一点，在平面几何的框架下，如何证明呢？这看起来需要很努力地思考。然而，如果我们采用解析几何的观点，这个问题就迎刃而解。如图4.1。假设顶点 $A=(a,b), B=(0,0), C=(c,d)$. 那么，我们可以直接计算三条高线 AD, BE, CF 的方程。然后可以算出 AD 和 BE 的交点 G，最后再计算验证 G 也在直线 CF 上即可。注意：如上的过程完全是机械化的，不需要思考，只需要计算。比如，用计算机可以很容易完成这样的计算。为什么原本看起来比较难的需要努力思考的问题，变成了用简单计算就可以解决的问题呢？因为我们把问题放在了一个更恰当的数学模型下来考虑。在这个例子中，就是我们不像平面几何里那样，仅仅把三角形看成一个孤零零的图形，而是把它放在坐标系中来研究。

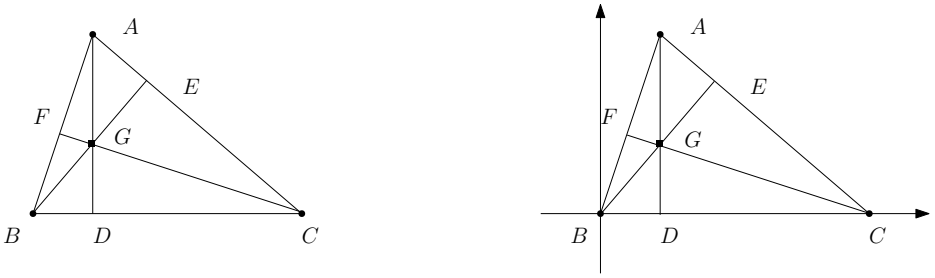

图 4.1: 从平面几何到解析几何

本章我们学习图这一数学模型。图论就是图的理论研究。许多实际的或数学的问题都可以用图这个模型来表示。一方面，当把某些实际问题用图的模型来表示后，问题往往就变得更容易解决（因为我们去掉了无关的信息，凸显了问题的本质）。另一方面，图论研究也大大地丰富和加深了人类对自然的认识。总之，在当今的科学和信息时代，图论的运用无所不在。因此，图是一个基本的科学语言和工具，是离散数学的一个主要研究对象。

4.1 哥尼斯堡七桥问题

传说，十八世纪初期，哥尼斯堡（今天俄罗斯的加里宁格勒）的人们很热衷于一个问题：能不能有一条散步的路线，经过哥尼斯堡的所有七座桥，见图4.2，且每座桥仅经过一次？很长时间人们都没有找到这样的散步路线。幸运地是，十八世纪最伟大的数学家欧拉思考了这个问题，并在他 29 岁时（1736 年）给出了优美且深刻的解答。欧拉的解答开创了图论这个学科，同时也预示了拓扑学的发展。事实上，欧拉提供的是一个如何看待这个问题的观点。一旦我们采取了欧拉的观点，问题就并不太难。欧拉的发明，就是用图（点和线）来给这个问题提供一个简单的模型，去掉无关的信息，而专注于其本质。

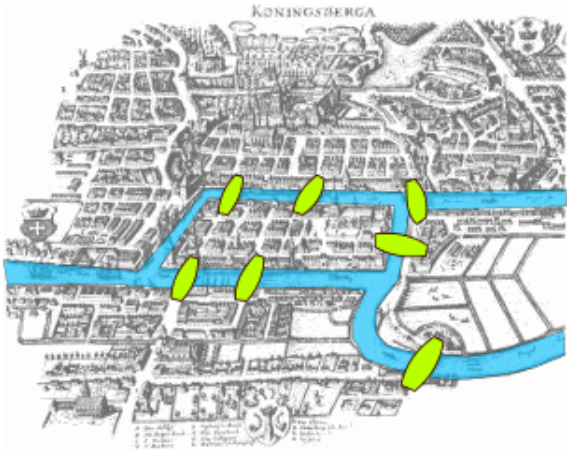

图 4.2: 欧拉时代的哥尼斯堡七桥（图片引自 wikipedia），哥尼斯堡在今天俄罗斯的加里宁格勒。

具体地，欧拉用图4.3来表示哥尼斯堡的城市和七座桥的关系，图中的七条边就代表了七座桥。

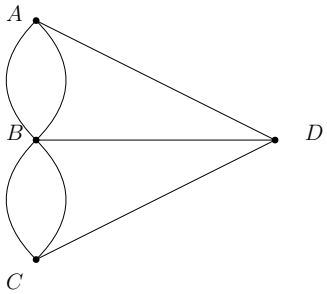

图 4.3: 哥尼斯堡七桥问题用图来建模。

当我们用图来表达哥尼斯堡七桥时，原来的问题就变成如下的图上的问题。我们可以根据散步路线是否要回到出发点，把问题分为两类。

定义 4.1. 欧拉通路：经过图中的每一条边一次且仅一次。

欧拉回路：经过图中的每一条边一次且仅一次，且回到出发点。

图4.4分别给出了欧拉通路和欧拉回路的例子。注意：这里仅要求边不重复通过，顶点是可以重复经过的。

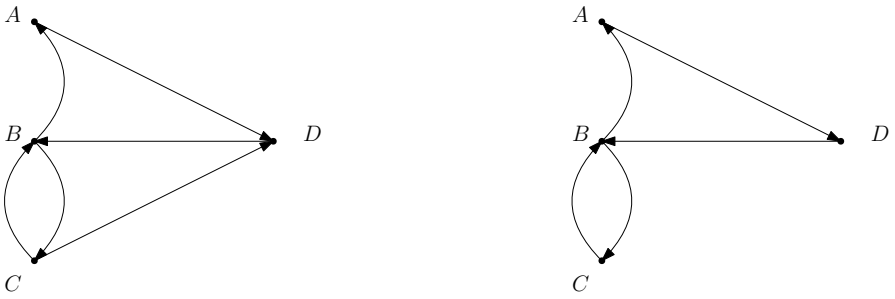

图 4.4: 欧拉通路 $C \to B \to A \to D \to B \to C \to D$，和欧拉回路 $C \to B \to A \to D \to B \to C$。

现在我们来研究哥尼斯堡七桥问题。我们先考虑图4.3是否存在欧拉通路。在图4.4我们已经看到，如果去掉 A 与 B 之间左边的边，就存在欧拉通路。看起来，图4.3中并不存在欧拉通路。我们能不能证明这一点呢? 如何思考这个问题? 首先我们观察到，如果存在一条欧拉通路，那么在这条通路上，每条边一定有一个确定的方向（因为每条边只通过一次）。现在假设图中有一条欧拉通路，考虑通路中的中间的某个点 v，即，v 不是起点也不是终点。我们来考虑进入 v 的边与从 v 出去的边。因为 v 是一个中间点，所以每一条进入 v 的边，都必定有对应的一条从 v 出发的边，见图4.5。于是，我们就证明了下面历史上第一个关于图的结论。

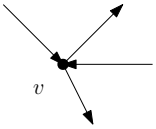

图 4.5: 欧拉通路之中某个中间点的边的进出情况。

定义 4.2. 设 v 是图中的某个顶点，与 v 连接的边的数目叫做 v 的度数，记作 $d(v)$。

命题 4.3. 对任意的图 G，如果图 G 中存在一条欧拉通路，则最多只有两个点的度数是奇数。

证明. 前面的分析证明了欧拉通路中每一个中间的点都应该有偶数条边与它连接，所以最多只有欧拉通路的起点和终点的度数可以是奇数。 □

我们回头来看图4.4中的欧拉通路，可以看出，只有起点 C 和终点 D 的度数是奇数（都是 3），而中间的点 B, A, D 的度数都是偶数。一旦我们发现了定理4.3，哥尼斯堡七桥问题就很容易回答了：在图4.3中，A, B, C, D 的度数分别是 $3, 5, 3, 3$，全是奇数，因此，不存在只通过每座桥一次且仅一次的散步路线。

解决了欧拉通路的问题后，稍微思考一下，就可以类似回答欧拉回路的问题。

命题 4.4. 对任意的图 G，如果图 G 中存在一条欧拉回路，则每个点的度数都是偶数。

所以，对哥尼斯堡七桥问题而言，自然也没有欧拉回路，即：没有散步路线能通过每座桥刚好一次并且回到出发点。解决了这些问题之后，我们再回顾解决的过程，可以体会到数学推导其实并不太难（就是认识到图4.5的现象而已），而解决问题的第一步在于，当我们面对真实的地图图4.2时，如何用恰

当的数学语言来表达这个问题。更进一步，事实上我们不只是解决了哥尼斯堡七桥问题，当我们发现了定理4.3和定理4.4之后，我们可以判定很多图有没有欧拉通路或回路。比如假设一个图有 6 个顶点，其度数分别是 $2,2,3,3,3,3$，那么我们立即知道这个图上既没有欧拉回路，也没有欧拉通路。注意，我们甚至根本没有看过，也不需要知道这个图到底是什么样子，就已经判断了它不存在欧拉回路！多么精彩！这不就是艺术上所谓的 “源于生活，高于生活” 吗?

如果我们进一步思考，会发现定理4.3和定理4.4只能用来判断一个图没有欧拉通路或者欧拉回路。比如，一个图有 6 个顶点，其度数分别是 $2,2,3,3,4,4$，它有没有欧拉通路呢? 一个图有 6 个顶点，其度数分别是 $2,2,4,4,4,4$，它有没有欧拉回路呢? 进一步思考后，可以证明新的定理。首先，一个图要有欧拉通路，那么任意两点之间必须有一条路线连着，这在图中被称为连通性。从这个考虑出发，下面我们稍微严格地定义一些图的基本概念（我们已经有了图的直觉，但严格定义对应的概念有助于我们更准确地把握我们在思考什么，没有思考什么，等等)。

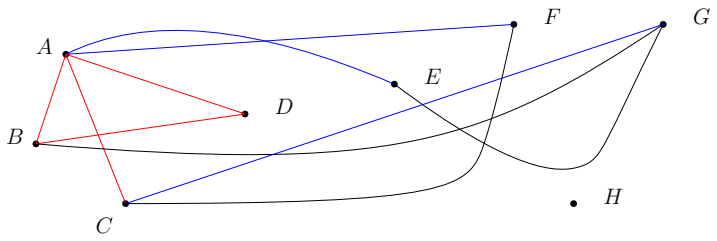

图 4.6: 图的例子。

定义 4.5. 一个（无向，即边没有方向）图 G 由点和边组成，点的集合通常用 V 表示，边的集合用 E 表示。图 G 一般写作 $G=(V,E)$.

- 点的相邻：如果两点之间有边，则说它们相邻；
- 边的相邻：如果两边有共同顶点，则说它们相邻；
- 点的度数：与点连接的边的数目；
- 简单图：两点之间最多一条边；
- 通路：从一点到另一点由边连接的序列（点和边均允许重复)；
- 路径：点不重复的通路；
- 回路：起点和终点相同的通路；
- 连通图：任意两点之间均有通路连接；
- 子图：由图的某些点和这些点上某些边组成的图；
- 导出子图：由图的某些点和这些点上所有的边组成的图；
- 连通分支：极大的连通子图，即，多加任何一点，得到的新的子图就不连通了。

例 4.6. 在图*4.6*中，$G=(V,E)$，其中

$$V=\{A,B,C,D,E,F,G,H\},$$
$$E=\{(A,B),(A,C),(A,D),(A,E),(A,F),(B,D),(B,G),(C,F),(C,G),(E,G)\}.$$

- 点的相邻：比如 A 与 B 相邻，C 与 G 相邻，但 B 与 F 不相邻，等等；
- 边的相邻：比如边 AC 与 CG 相邻，AC 与 BG 不相邻；
- 点的度数：可以写出点的度数序列如下

$$d(A),d(B),d(C),d(D),d(E),d(F),d(G),d(H)=5,3,3,2,2,2,3,0.$$

- 简单图：图*4.6*是简单图，哥尼斯堡七桥图*4.3*不是简单图；
- 通路：在图*4.6*中，$A(A,B)B(B,G)G(G,C)C(C,A)A(A,F)F$ 是一条通路。
- 路径：在图*4.6*中，$B(B,D)D(D,A)A(A,C)C(C,F)F$ 是一条路径，但 B 到 F 的最短的路径就是 $B(B,A)A(A,F)F$;
- 回路：在图*4.6*中，$A(A,B)B(B,G)G(G,C)C(C,A)A(A,F)F(F,C)C(C,F)F$ 是一条回路。
- 连通图：图*4.6*不是连通图，哥尼斯堡七桥图*4.3*是连通图；
- 子图：图*4.6*中蓝色的边是点集 A,C,E,F,G 上的子图，但不是 A,C,E,F,G 上的导出子图。
- 导出子图：图*4.6*中红色的边是点集 A,B,C,D 上的导出子图；
- 连通分支：图*4.6*有两个连通分支，一个是点集 A,B,C,D,E,F,G 上的导出子图，一个是点 H 上的导出子图（没有任何边）。图*4.6*中蓝色的边作为点集 A,C,E,F,G 上的子图也不是连通图，也有两个连通分支。

有了这些概念，我们可以叙述如下的一般定理，它更加完整地回答了如何判断一个图是否有欧拉通路和回路。

定理 4.7. 对任意的连通图 G，

- 图 G 中存在一条欧拉通路，当且仅当要么每个点的度数都是偶数，要么恰有两点的度数是奇数。
- 图 G 中存在一条欧拉回路，当且仅当每个点的度数都是偶数。

根据定理4.7，可以判定如下：图4.6不连通，所以肯定不存在欧拉通路或欧拉回路；如果只考虑其在 A,C,E,F,G 上的连通分支，因为度数序列是 $5,3,3,2,2,2,3$，有 3 个奇数的度数，因此不存在欧拉通路或欧拉回路；如果考虑在 A,B,C,D 上的导出子图（红色的边），在导出子图的度数序列是 $3,2,1,2$，恰有两个奇数点，因此存在欧拉通路（比如：CABD），但不存在欧拉回路。

上面我们定义并讨论了一些无向图，如果每一条边都标有方向，就叫做有向图。我们主要讨论无向图。只要掌握了无向图研究的基本方法，也可以类似地研究有向图。无向图和有向图的研究也不是相互独立的，有时能互相帮助。比如，上面我们在讨论哥尼斯堡七桥图这个有向图是否有欧拉通路时，思考的过程运用了有向图（图4.5）的观点。

4.2 一些基本定理

从本节开始，如无特别说明，图均指无向、简单图。本节我们继续介绍一些图的基本定理和基本概念。

4.2.1 握手定理

我们先用人的相互认识关系来看一个例子。假设人的认识关系是相互的，那么任何一群 25 个人的聚会，一定至少有一个人认识的人的个数是偶数！任何一群 37 个人的聚会，一定也至少有一个人认识的人的个数是偶数！这看起来有些玄妙。下面我们用图的语言来描述这种现象，并用我们之前学过的计数方法来讨论，就会豁然开朗。

定理 4.8. *对任意一个图 $G=(V,E)$，有 $\sum_{v\in V} d(v)=2|E|$.*

图中有点和边，点和边有连接关系。我们可以对此用双计数法。

证明. 考虑如下点和边的相邻关系，

$$A=\{(v,e)\in V\times E : v\text{是边}e\text{的一个顶点}\}.$$

对 A 用双计数。

- 从点的观点来看，因为点 v 是 $d(v)$ 条边的顶点，所以 $|A|=\sum_{v\in V} d(v)$.
- 从边的观点来看，因为每条边刚好有两个顶点，所以 $|A|=2|E|$.

两者结合，即得结论。 □

定理4.8又叫握手定理。

推论 4.9. *对任意一个图 $G=(V,E)$，度数是奇数的顶点必定是偶数个。*

证明. 用反证法，假设某个图的度数是奇数的顶点是奇数个，则此图的所有顶点的度数和是奇数。但定理4.8表明，所有顶点的度数和等于 $2|E|$，是偶数。这是矛盾的。 □

回头再来看前面的例子。为什么 25 人的聚会中必定至少有一人认识的人的个数是偶数呢？用图的观点来看，人的相互认识构成一个图。如果所有人认识的人的个数都是奇数，也就等于说图中所有顶点度数都是奇数，因为总人数是 25 个，那么所有顶点的度数之后就是一个奇数，这与定理4.8矛盾。因此，必定至少有一个顶点的度数是偶数。即，至少有一个人认识的人的个数是偶数个。这不是很有趣吗？

4.2.2 从树到平面图的欧拉公式

环与树

在计算机、生物等学科中常见的一种图叫做树，从数学上来说，树也是最简单的图之一。为了准确地定义树，我们需要说明什么是环这个概念。

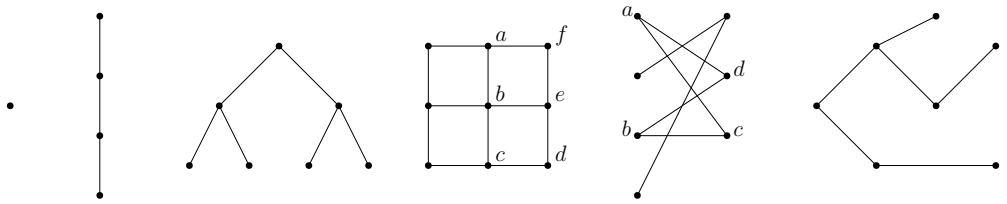

图 4.7: 树的例子：第 1，2，3，6 图都是树，第 4，5 不是。第 4 个图中有很多环，比如，$abef$ 是一个长度是 4 的环，$abcdef$ 是一个长度是 6 的环。第 5 个图中有一个长度是 4 的环 $acbd$。

定义 4.10. 边不重复的回路叫做环。如果一个简单图是连通的且没有环，则叫做一棵树。

引理 4.11. 设图 $T = (V, E)$ 是一个树。则必存在 $v \in V$ 满足 $d(v) = 1$.

证明. 证明大意：用反证法。假设每个点的度数都是 ≥ 2。则任选一点从它开始可选择与它连接的一条边与下一个顶点相连，对下一个顶点，因为其度数 ≥ 2，因此对它也可以选择一条边与另一个顶点相连。如此不断继续，则可不断构造越来越长的路径，经过越来越多的点。然而，因为点的个数有限，因此在某一步时，选择的边的端点必然要回到之前经过的某个点，从而形成一个环，这与 T 是树矛盾。 □

定义 4.12. 树中度数是 1 的顶点叫做树叶。

引理4.11表明：任何树都有树叶（在生活中，这句话当然是不对的，树的树叶可以都掉光）。

定理 4.13. 设图 $T = (V, E)$ 是一个树。则 $|E| = |V| - 1$.

证明. 对 $|V|$ 用归纳法。

基础情形：$|V| = 1$，此时没有边，即 $|E| = 0$，结论成立。

归纳假设：假设结论对 $|V| = k$ 成立。即，任意的树 $T = (V, E)$，如果树的顶点个数是 k，则边的条数是 $k - 1$。

现在考虑 $|V| = k+1$ 的情形。根据引理4.11，树 T 中存在度数是 1 的顶点。设 $v \in V$ 满足 $d(v) = 1$，用 e 表示以 v 为顶点的唯一的边。考虑 T 的子图 $T' = (V', E')$，其中 $V' = V - \{v\}$，$E' = E - \{e\}$. 注意到 T' 是一个有 $|V'| = |V| - 1 = k + 1 - 1 = k$ 个顶点的树，根据归纳假设 $|E'| = k - 1$. 但是因为 $E' = E - \{e\}$，所以 $|E'| = |E| - 1$，故，$|E| = |E'| + 1 = k - 1 + 1 = k$. 证毕。 □

平面图的欧拉公式

欧拉在研究多面体过程中发现了著名的多面体的欧拉公式。回忆之前我们用双计数方法考察了正 20 面体，见图3.1。用 V, E, F 分别表示正 20 面体的顶点集，边集，面集。我们知道 $|F| = 20$，并用双计数方法证明了 $|E| = 30$。正 20 面体有多少顶点呢? 当然我们也可以去一个个数它。不过从数学的角度，我们总是想找到别的方法来计算，而不是用眼睛来数。如图3.1，可以看出正 20 面体的每个顶点度数都是 5。那么，根据定理4.8，有 $\sum_{v \in V} 5 = 2 \times |E| = 60$. 因此，$|V| = 12$. 类似地，可以计算其他正多面体的顶点数，边数，和面数，见表4.1。

表 4.1: 正多面体的顶点数，边数，和面数。

多面体	$\|V\|$	$\|E\|$	$\|F\|$	$\|V\|-\|E\|+\|F\|$
正 4 面体	4	6	4	2
正 6 面体（立方体）	8	12	6	2
正 8 面体	6	12	8	2
正 12 面体	20	20	12	2
正 20 面体	12	30	20	2

表4.1的数据表明，对正多面体而言，$|V|-|E|+|F|$ 总是等于 2，即，不论顶点数，边数，和面数如何变化，公式 $|V|-|E|+|F|$ 的值都不变，恒等于 2，这种不变的量，在数学上被称为不变量。对不变量的研究，是数学一个重要的内容，在变化中寻求不变，也就是去发现研究对象的内在性质。

我们将用图的语言来描述并证明欧拉公式。当用图的语言来描述欧拉公式后，欧拉公式的应用范围也大大扩大，从多面体到平面图。

定义 4.14. 如果一个图能画在平面上使得所有边除了顶点外都不相交，这样的图叫做平面图。

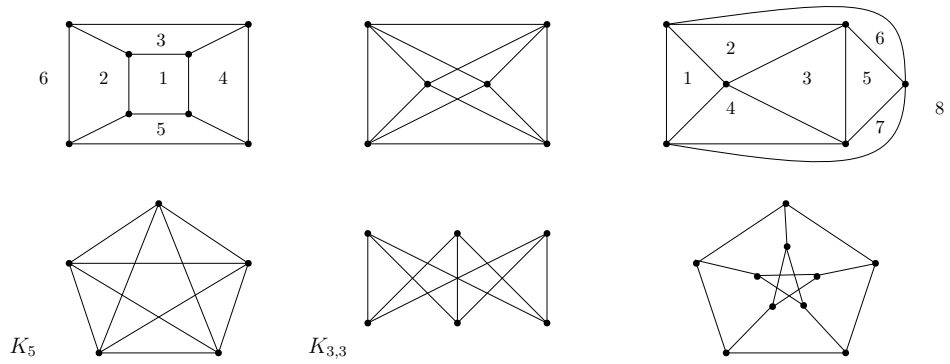

图 4.8: 第一行都是平面图，第二行都不是平面图。第一个图有 6 个面。第二个图与第三个图其实是一样的，只是画法改变了一下。第三个图有 8 个面。

欧拉公式包含顶点，边，和面的关系。为了考察平面图的欧拉公式，我们需要说明平面图确定了多少个 “面” 这一概念。首先把平面图画成每条边都不相交的图形，此时，每个环都围成一块区域，如果这块区域里不包含更小的环，则称为一个面。注意：整个图形外面的部分也是一个面。见图4.8.

定理 4.15. 对任意的连通、平面、简单图 $G=(V,E)$，如果把 G 画在平面上使得边（除了顶点外）都不相交，以 F 记此时得到的面的集合，则有欧拉公式 $|V|-|E|+|F|=2$.

下面将采用归纳法给出两个稍微不同的证明，分别用两个不同的参数来做归纳法。（感谢王文静和杨坤在课堂上的讨论）

证明一. 对图中面的个数 $F(G)$ 做归纳法。注意 $F(G)\geq 1$。

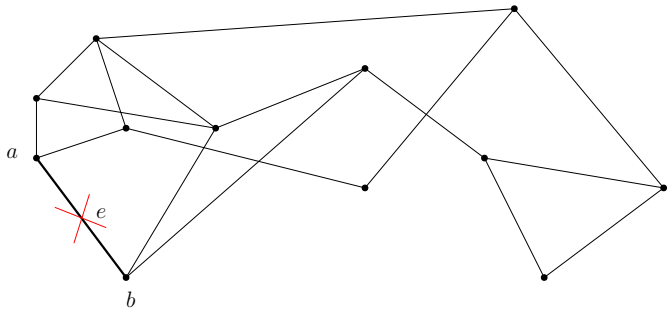

图 4.9: 欧拉公式定理4.15证明的示例。

基础情形：$F(G)=1$. 此时图 G 中必定没有环（因为有环则至少有两个面），因此图 G 是一个树。因此，根据定理4.13，有 $|V|-|E|+|F|=|V|-(|V|-1)+1=2$. 结论正确。

归纳假设：假设图包含 $F(G)=k$ 个面时结论成立。即，如果图 $F(G)=k$，则 $|V(G)|-|E(G)|+|F(G)|=2$.

现在考虑 $F(G)=k+1$ 的情形。设 $G=(V,E)$ 包含 $F(G)=k+1$ 个面. 设 F 表示 G 的面的集合。注意 $k+1\geq 2$。因此图 G 中必定有环。任选一个环，考虑环中的任意一条边 e，如图4.9所示。考虑去掉边 e（边 e 的两个端点 a,b 仍然保留）后得到的子图 $G'=(V,E-\{e\})$. 设 F' 表示图 G' 的面的集合，因为边 e 只是两个面的公共边，则有 $|F'|=|F|-1=k+1-1=k$。因此，可以对 G' 用归纳假设，有 $V(G')-E(G')+F(G')=2$. 因此，

$$2=V(G')-E(G')+F(G')=|V|-|E-\{e\}|+|F'|=|V|-(|E|-1)+(|F|-1)=|V|-|E|+|F|$$

即得所证。 □

证明二. 对图中环的个数 $c(G)$ 做归纳法。注意 $c(G)\geq 0$。

基础情形：$c(G)=0$. 此时图 G 是一个树。注意，树恰好有 1 个面。因此，根据定理4.13，有 $|V|-|E|+|F|=|V|-(|V|-1)+1=2$. 结论正确。

归纳假设：假设图包含 $c(G)\leq k$ 个环时结论成立。即，如果图 G 包含的环数不超过 k 个，则 $|V(G)|-|E(G)|+|F(G)|=2$.

现在考虑 $c(G)=k+1$ 的情形。设 $G=(V,E)$ 包含 $c(G)=k+1$ 个环. 设 F 表示 G 的面的集合。任选一个环，考虑环中的任意一条边 e，如图4.9所示。考虑去掉边 e（边 e 的两个端点 a,b 仍然保留）后得到的子图 $G'=(V,E-\{e\})$. 设 F' 表示图 G' 的面的集合. 注意到如下两点：

- $|F'|=|F|-1$，（因为边 e 只是两个面的公共边）
- $c(G')\leq c(G)-1=k+1-1=k$. （去掉边 e，环至少减少 1 个）

因为 $c(G')\leq k$，可以对 G' 用归纳假设，有 $V(G')-E(G')+F(G')=2$. 因此，

$$2=V(G')-E(G')+F(G')=|V|-|E-\{e\}|+|F'|=|V|-(|E|-1)+(|F|-1)=|V|-|E|+|F|$$

即得所证。 □

推论 4.16. *若 $G=(V,E)$ 是至少有 3 个顶点的连通、平面、简单图，则 $|E|\le 3|V|-6$.*

欧拉公式是一个包含点、边、面的等式。为了得到点和边大小的不等式，我们需要把面的个数从欧拉公式中去掉。为此，我们考虑边和面的关系，并用双计数方法。

证明. 对图 G 的边与面的关系的双计数。设

$$A=\{(e,f)\in E\times F : e\text{是}f\text{的一个边}\}.$$

类似于例3.15，对 A 用双计数。

- 从边的观点：每条边都在两个面上。因此，$|A|=2|E|$.
- 从面的观点：因为每个面至少包含 3 条边，故，$|A|\ge 3|F|$.

因此，$2|E|\ge 3|F|$，即，$|F|\le 2|E|/3$. 把此不等式代入欧拉公式，得到

$$2=|V|-|E|+|F|\le |V|-|E|/3.$$

从而得 $|E|\le 3|V|-6$. □

推论 4.17. *若 $G=(V,E)$ 是连通、平面、简单图，则 G 中必有一个顶点度数 ≤ 5.*

证明. 用反证法。假设对任意的 $v\in V$,都有 $d(v)\ge 6$. 根据握手定理,即定理4.8,有 $2|E|=\sum_{v\in V} d(v)\ge 6|V|$，即 $|E|\ge 3|V|$. 但这与推论4.16矛盾。 □

推论4.16告诉我们，连通平面图的边的条数不能太多。这样我们就得到判断一个图不是平面图的方法。比如，图4.8中的第四个图 K_5，有 5 个顶点，10 条边，不满足推论4.16中的不等式，从而我们知道它必定不是平面图。不过，第五个图 $K_{3,3}$ 有 6 个顶点，9 条边，仍然满足推论4.16中的不等式，但第五个图也不是平面图。因此，推论4.16中的不等式只是平面图的必要条件，并不是充分条件。第六图也不是平面图，不过它也满足推论4.16中的不等式。有没有判断一个图是不是平面图的充要条件呢？数学家经过研究，确实找到了这样的充要条件。但定理内容的准确陈述比较复杂，我们仅大略陈述如下。

定理 4.18 (Kuratowski)**.** *设图 G 是连通简单图，则 G 是平面图当且仅当 G 不包含 K_5 和 $K_{3,3}$，并且不包含这个两个图的“某种拷贝”。*

有兴趣了解细节的同学可以参考中文教材上的相关内容。

4.2.3 图的同构

图是用来模拟某些具体问题的。对有些问题而言，当把其用图的语言来描述时，顶点各自有不同的含义。而对有些问题而言，我们只关心边的相互关系，并不对顶点做过多区分。为了做出这种区别，在数学上我们考虑一个新的概念，叫做同构。

定义 4.19. *设 $G=(V,E)$，$H=(V',E')$ 是两个图。如果存在一个双射 $f:V\to V'$ 满足：*

- *如果 $(a,b)\in E$，则 $(f(a),f(b))\in E'$,*

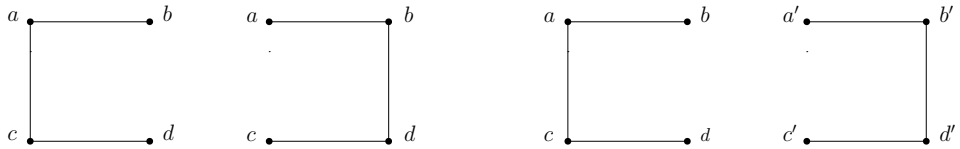

图 4.10: 图的同构。第一和第二个图是不同的图。当考虑同构时，它们是同构的。

- 如果 $(a,b) \notin E$，则 $(f(a),f(b)) \notin E'$，

则说图 G 和图 H 同构。此时，f 是 G 和 H 之间的一个同构映射。

如图4.10。如果我们认为 a,b,c,d 四个顶点都有重要的含义，不能混为一谈，这种情况下，第一个图和第二个图就是不同的图。比如：如果 a,b,c,d 代表四个城市的名字，边代表它们之间有没有直飞的航线，那么第一个图表明 a 与 c 之间有直飞的航线，而 b 与 d 之间没有；第二个图表明 a 与 c 之间没有直飞的航线，而 b 与 d 之间有。因此，两个图表达的意义完全不同。

另一方面，当我们不关心图中每个顶点的具体含义时，不同的图则可以看作是 “相同” 的，即同构。直观上来说，同构就是说我们忽略掉图上不同顶点的区别，只关注图的边的关系看起来是不是一样的。仍然考虑图4.10。我们看第三和第四个图（其实就是第一和第二个图，只是为了避免混淆，我们给最后一个图的顶点重新命名），它们是同构的。因为我们可以定义如下的同构映射，

$$f(a)=b', \quad f(b)=a', \quad f(c)=d', \quad f(d)=c'.$$

容易验证，f 是一个同构映射，满足定义4.19中的条件。比如，在第三个图中有边 (a,c)，而对应的 $(f(a),f(c)) = (b',d')$ 也是第四个图的一条边；在第三个图中没有边 (b,d)，而对应的 $(f(b),f(d)) = (a',c')$ 在第四个图中也不是一条边。

许多时候，当我们研究图的某些性质时，都是在同构的意义下考察图。比如，当我们问一个图是否是连通图，显然此时我们没有必要区分不同的顶点。其实，在哥尼斯堡七桥问题中，我们问图中是否存在一个欧拉回路，也是在同构的意义下考察图的性质，如图4.11. 当考虑欧拉回路时，按第一个图来考虑，和按第二个图来考虑，问题没有区别。

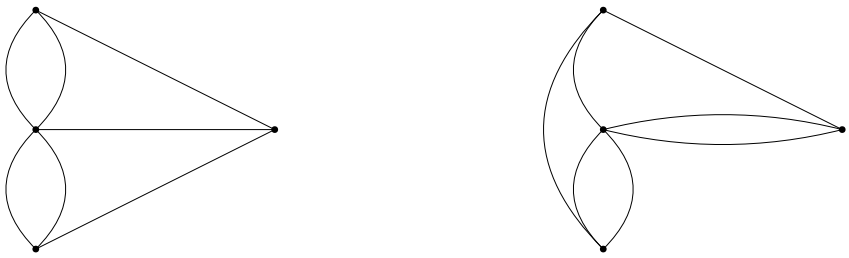

图 4.11: 哥尼斯堡七桥问题中图的同构。第二个图和第一个图是同构的。哥尼斯堡七桥问题对两个图而言都是同一个问题。

同构的图有许多相同的图的性质。比如，同构的图

- 有相同的顶点数，相同的边数；

- 有相同的顶点度数序列；
- 有相同的连通性；
- 有相同个数的环；
- 要么都有欧拉回路，要么都没有欧拉回路。

等等。

4.3 三个简洁快速的算法

The following is from wikipedia:

In 1951, David A. Huffman and his MIT information theory classmates were given the choice of a term paper or a final exam. The professor, Robert M. Fano, assigned a term paper on the problem of finding the most efficient binary code. Huffman, unable to prove any codes were the most efficient, was about to give up and start studying for the final when he hit upon the idea of using a frequency-sorted binary tree and quickly proved this method the most efficient.

The following is from wikipedia:

What is the shortest way to travel from Rotterdam to Groningen, in general: from given city to given city. It is the algorithm for the shortest path, which I designed in about twenty minutes. One morning I was shopping in Amsterdam with my young fiancée, and tired, we sat down on the café terrace to drink a cup of coffee and I was just thinking about whether I could do this, and I then designed the algorithm for the shortest path. As I said, it was a twenty-minute invention. In fact, it was published in '59, three years later. The publication is still readable, it is, in fact, quite nice. One of the reasons that it is so nice was that I designed it without pencil and paper. I learned later that one of the advantages of designing without pencil and paper is that you are almost forced to avoid all avoidable complexities. Eventually, that algorithm became to my great amazement, one of the cornerstones of my fame.

—*Edsger Dijkstra*

本节我们学习三个比较简单，非常著名，极其重要，被广泛应用的与图有关的算法。

4.3.1 最小生成树的 Kruskal 算法：贪心算法

前面我们学习过，树就是连通无环（简单）图。容易看出，如果从树中去掉任意一条边（但仍保留顶点），则得到的图就是不连通的了。从这个意义上说，树就是最小的连通图。许多实际问题都对应某个连通的图。比如，要在一个城市铺设某种网络，使得不同地方都可以互通。网络路线有不同的选择，

成本和资源消耗也因此不同，为了节约资源和成本，需要寻找某个最经济的连通子图，这就是最小生成树的概念。

定义 4.20. 设 $G=(V,E)$ 是一个连通图，$w: E \to \mathbb{R}^+$ 是一个定义在边集上的函数。对任意的边 $e \in E$，$w(e)$ 叫做边 e 的权，图 G 因此也叫带权图。设 $T=(V,E')$ 是 G 的一个子图，如果 T 满足

- T 是一棵树，
- $\sum_{e\in E'} w(e)$ 是所有树中最小的，

则称 T 是 G 的一个最小生成树。

给定一个连通的带权图，如何求出它的一棵最小生成树？有好一些不同的算法。下面介绍著名的 Kruskal 算法，这是一个典型的贪心算法。*贪心算法*是算法设计的一个最基本的方法，其基本思想是：根据问题的目标，一步一步地选择局部最优解，以期达到全局的最优解。对某些问题，贪心算法给出的解是最优的，比如最小生成树问题。而对另一些问题，贪心算法给出的解则不一定是最优的，如后面我们将讨论的最短路问题。对于什么样的问题，贪心算法给出的解是最优的呢？可以用拟阵（matroid，是矩阵 matrix 的一种推广）的理论给出回答，不过本课不予讨论。

Algorithm 4 给定连通带权图 $G=(V,E)$ 和权函数 $w: E \to \mathbb{R}^+$，求 G 的一个最小生成树 T

procedure MST
  $V_T = \emptyset, E_T = \emptyset$
  将边按权的大小从小到大排序，设：$w(e_1) \le w(e_2) \le \cdots \le w(e_m)$
  $i = 1$
  while $V_T \neq V$ **do**
    if e_i 和 E_T 中的边不形成环 **then**
      $E_T = E_T \cup \{e_i\}$
      $V_T = V_T \cup \{\text{two vertices of } e_i\}$
      $i = i + 1$
    else
      $i = i + 1$
  return $T = (V_T, E_T)$

图4.12举例说明了 Kruskal 算法4的运行过程。从算法的运行可以看出，最小生成树并不一定是唯一的：图4.12中（7）和（7'）都是最小生成树，对应的权的和是 $2+3+5+7+9=26$.

4.3.2 Huffman 编码与香农熵

信息传输的一个基本问题是编码。以文本传输为例。一般地，如果传输的信息对应的不同字母个数是 k 个，那么总可以用 $\lceil \log_2 k \rceil$ 这么多位二进制对每个字母进行编码。比如，如果一段文本只涉及字母 $\{a,b,c,d,e,f,g\}$，因为只有 7 个字母，所以可以用 $\lceil \log_2 k \rceil = 3$ 位二进制来给每个字母编码，如表4.2. 在这种编码下，如果传输一段文本有 1000 个字母组成，那么传输的二进制总共就是 $1000\lceil \log_2 k \rceil$ 这么多位。有没有可能采用不同的编码方式来减少传输量呢？

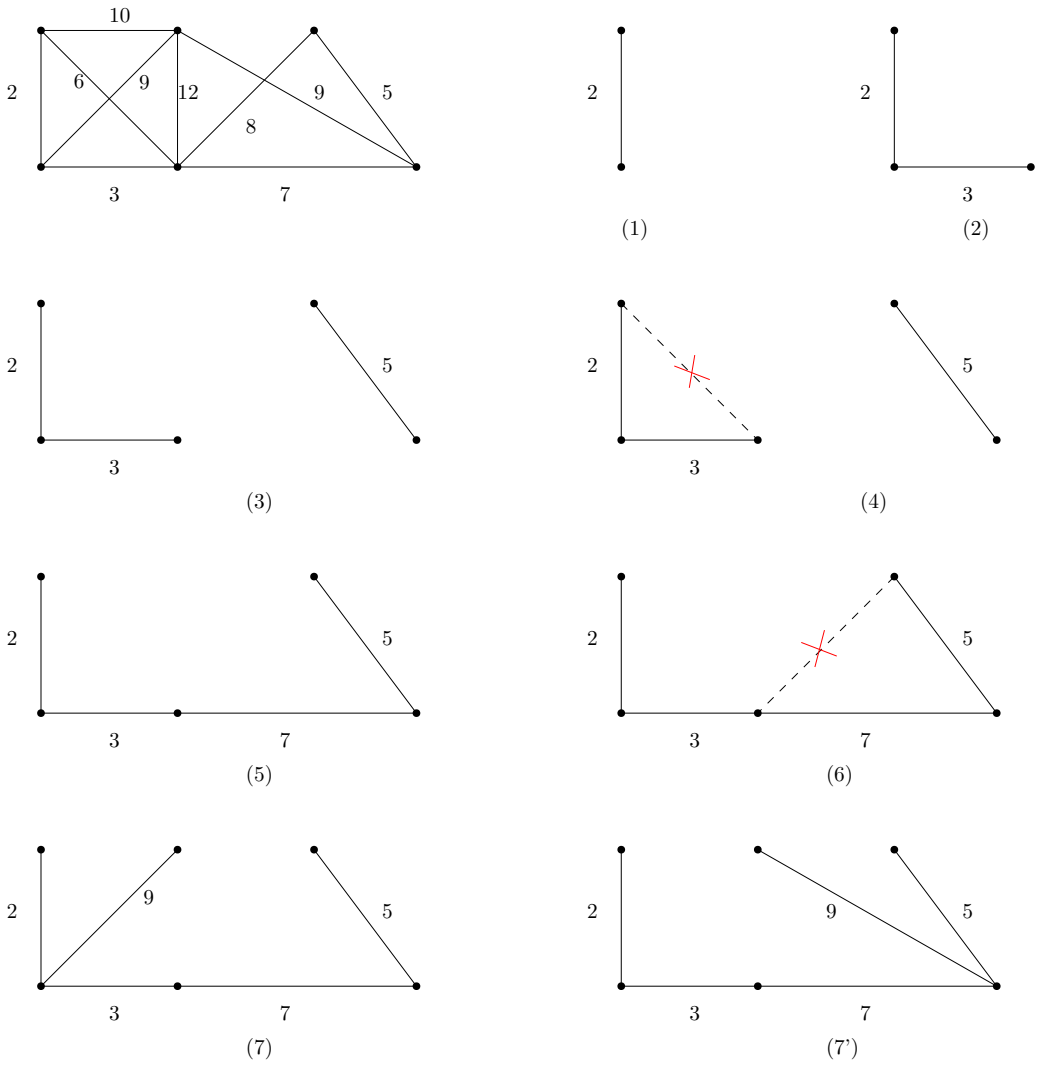

图 4.12: 最小生成树的 Kruskal 算法。

表 4.2: 用 $\lceil \log_2 k \rceil$ 位二进制对 k 个字母编码

a	b	c	d	e	f	g
000	001	010	011	100	101	110

稍加观察就会发现，自然语言的文本每个字母出现的频率有不小的差异，有的字母出现的频率很高，有的很低。比如，在英语单词中，字母 e 的出现频率约是字母 q 的 56 倍！如果我们对出现频率较高的字母用比较短的编码，而对出现频率低的字母允许长一些的编码，那么传输一段比较长的文本时，可能总的传输量就能得到减少：因为高频出现的字母的编码较短。这就是 Huffman 编码的基本思想：依据字母出现的频率来编码。具体地，Huffman 编码可以用构造二叉树的方式来构造。

下面举例说明。假设字母 $\{a, b, c, d, e, f, g\}$ 出现的频率如表4.3所示。图4.13中给出了 Huffman 编

Algorithm 5 给定字母表 $S=\{x_1,\ldots,x_m\}$ 及频率（或权重）表 $\{w(x_1),\ldots,w(x_m)\}$，求每个字母的 Huffman 编码

procedure HUFFMANCODE
  $S=\{x_1,\ldots,x_m\}$
  $i=1$
  while $|S|>1$ **do**
    从 S 中选择两个权重最小的字母，设为 a,b
    构造一个新的顶点 y_i，y_i 分别连接 a 和 b
    $w(y_i)=w(a)+w(b)$
    $S=S-\{a,b\}+\{y_i\}$
    $i=i+1$
  return 字母的 Huffman 编码（根据构造的二叉树而得）

码的运行过程。直接计算可得，Huffman 编码得平均长度是

$$0.12\times 3+0.20\times 2+0.08\times 4+0.08\times 4+0.24\times 2+0.12\times 3+0.16\times 3=2.72.$$

表4.2中的编码平均长度自然是 3。因此，利用字母出现频率的高低进行编码可以缩短编码的平均长度。

香农在研究信息论时定义了香农熵这一概念，此后这成为信息社会的一个基本概念，极其重要，后来在纯粹数学中也得到许多应用。给定一个离散概率分布 $p_1,p_2,\ldots,p_m$，满足 $\sum_i p_i=1$. 其对应的香农熵定义为 $\sum_i p_i\log_2\frac{1}{p_i}$. 直接计算表4.3中 7 个字母概率分布的香农熵是

$$\begin{aligned}&0.12\times\log_2\frac{1}{0.12}+0.20\times\log_2\frac{1}{0.20}+0.08\times\log_2\frac{1}{0.08}+0.08\times\log_2\frac{1}{0.08}\\&+0.24\times\log_2\frac{1}{0.24}+0.12\times\log_2\frac{1}{0.12}+0.16\times\log_2\frac{1}{0.16}\approx 2.70.\end{aligned}$$

比较可见，香农熵和 Huffman 编码的平均长度非常接近。直观上来说，香农熵从理论上给出了编码的平均长度的最优值，而 Huffman 编码用一个具体（而且简洁的）方法构造了一种编码，其平均长度非常接近香农熵。

表 4.3: 字母出现有不同频率时的 Huffman 编码

字母	a	b	c	d	e	f	g
权重	3	5	2	2	6	3	4
频率	12%	20%	8%	8%	24%	12%	16%
Huffman 编码	100	11	0000	0001	01	101	001

4.3.3 最短路的 Dijkstra 算法

给定一个带权连通图中的两点，求这两点之间最短的路径，这就是最短路问题。很明显，这是一个非常基本的具有重要实际意义的问题。我们首先尝试看贪心算法能否给出我们最优的路线。以图4.15为

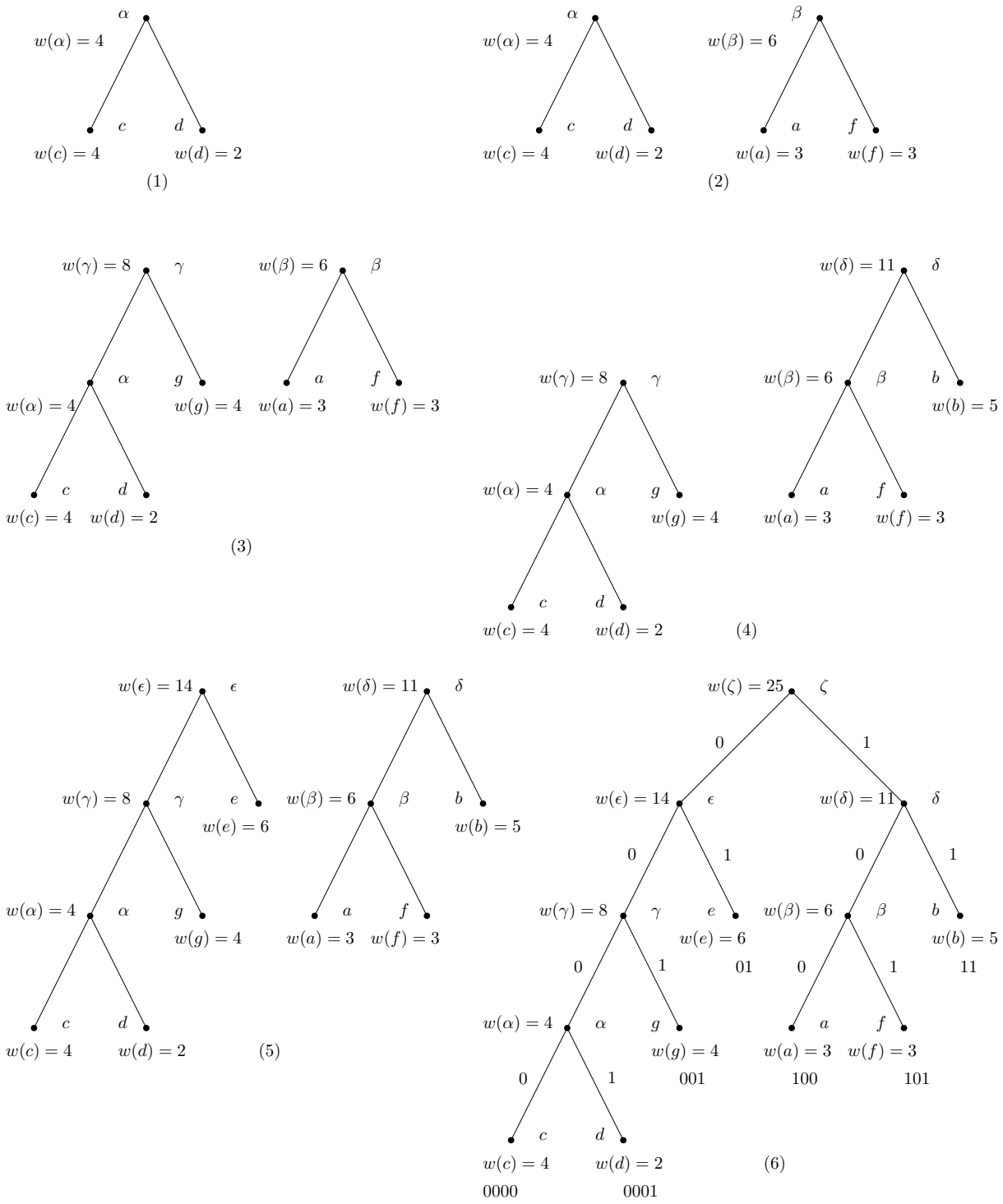

图 4.13: Huffman 编码示例。

例。假设我们要求从 a 到 f 的最短路。直接应用贪心算法的思想可以给出如下算法：从 a 开始，每一步选择到达某个新顶点的最短边，依次前进直到达到顶点 f。按照这样的贪心算法，在图4.15中我们会得

到两条可能的路线，如图4.15- (i) (ii) 所示，路线距离分别是 $3+2+3+2=10$ 及 $3+2+3+2+2=12$. 然而，容易看出，图4.15中 a 到 f 的最短路为红色标注的部分，距离是 $5+2+2=9$. 因此，贪心算法此时未能给出最短路线。稍加思考，可以知道，能够构造适当的图，使得贪心算法给出的距离比最优距离要大很多。

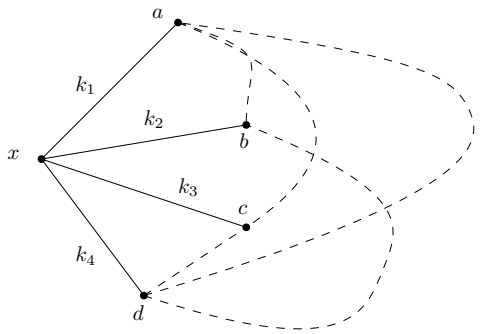

图 4.14: Dijkstra 算法的局部思路。

下面我们介绍 Dijkstra 给出的求最短路的算法。注意到如下事实：如果某条从 x 到 y 的路线是 x 到 y 的最短路，那么对路线上的任意一点 z，从 x 到 z 沿此路线也是从 x 到 z 的最短路。因此，为了求出从 x 到 y 的最短路，也必须求出从 x 到 z 的最短路。由于事先我们并不知道哪条路线是从 x 到 y 的最短路，因此，看起来也许我们应该把从 x 到任意一点的最短路都求出来。有了这个想法后，我们先来考虑 x 的邻居（即与 x 直接相邻的点），如图4.14. 稍加观察，可以发现，一般而言，x 与其邻居直接相连的边并不一定是从 x 到该邻居的最短路。例如，图4.15中 a 到 c 的最短路是经过 b 到达 c，而非直接由 a 到 c，即边 ac 的长度并不是从 a 到 c 的最短距离。但是，稍加思考容易发现如下事实。如图4.14所示，设 x 的邻居是 a,b,c,d，边长分别是 k_1,k_2,k_3,k_4. 那么，对这些边中最短的边对应的顶点而言，从 x 到该顶点的最短路一定就是对应的边。比如，如果 k_1 是 k_1,k_2,k_3,k_4 中最小的，那么从 x 到 a 的最短路一定就是 k_1. 也就是说，通过考察 x 的邻居的边长，我们至少可以确定 x 的某个邻居的最短路。比如在上例中，当我们确定了从 x 到 a 的最短路后，我们就可以从 a 出发，来继续考察 a 的邻居。如此等等。这就是 Dijkstra 算法的基本思想。具体算法见算法6. 图4.15给出了 Dijkstra 算法在一个图上执行的逐步示例。

4.4 图的矩阵表示及其应用

截至目前，我们认识图都是直接画出图的顶点和边。为了设计图的某些算法，为了研究图的某些性质等，把图用别的方式表示出来将会更加方便。下面我们介绍图的两种矩阵表示及其应用。我们将会看到，图的矩阵表示对于研究图的性质有很大的帮助。我们仅介绍无向图的表示，有向图的表示可以类似考虑和研究。

Algorithm 6 给定连通带权图 $G=(V,E)$ 和权函数 $w:E\to\mathbb{R}^+$，求 G 中从 x 出发到其他点的最短路

```
procedure DIJKSTRA
    for v ∈ V do
        d(v) = ∞
        IsFinished(v) = 0
    d(x) = 0
    curvtx = x
    while IsFinished(curvtx) == 0 do
        N = curvtx 的所有邻居
        for v ∈ N do
            if IsFinished(v) == 0 && d(curvtx) + w(curvtx, v) < d(v) then
                d(v) = d(curvtx) + w(curvtx, v)
        IsFinished(curvtx) = 1
        if 所有顶点都 finished then
            break
        else
            curvtx = 还没有 finished 的顶点中 d(v) 值最小的那个
    return
```

4.4.1 关联矩阵

关联矩阵 M 的行对应图的顶点，列对应图的边。给定一个顶点 v 和一条边 e，$M(v,e)$ 表示边 e 与 v 的关联次数。以图4.16为例说明，图（1）对应的关联矩阵如下。

$$M=\begin{pmatrix}2&1&1&0&0&1&0&0\\0&1&1&1&1&0&1&0\\0&0&0&1&1&0&0&1\\0&0&0&0&0&1&1&1\end{pmatrix}.$$

其中，M 的行分别对应顶点 u,v,w,x，列分别对应边 a,b,c,d,e,f,g,h. 比如，$M(u,b)=1$，因为边 b 与顶点 u 关联 1 次。$M(u,g)=0$，因为边 g 与顶点 u 不相关。注意，$M(u,a)=2$，因为边 a 的两个端点都是 u，所以边 a 与顶点 u 的关联次数是 2.

容易观察到如下事实：

- 每一列的和都是 2：因为每条边与两个顶点关联；
- 每一行的和就是对应的顶点的度数。

根据如上观察，如果我们计算关联矩阵里所有的数字和，那么按列计算得到 $2|E|$，按行计算得到 $\sum_{v\in V}d(v)$。于是我们重新（也是用双计数方法）证明了握手定理 $2|E|=\sum_{v\in V}d(v)$.

关联矩阵有 $|V|$ 行 $|E|$ 列，因此，一般地，关联矩阵并不是方阵。

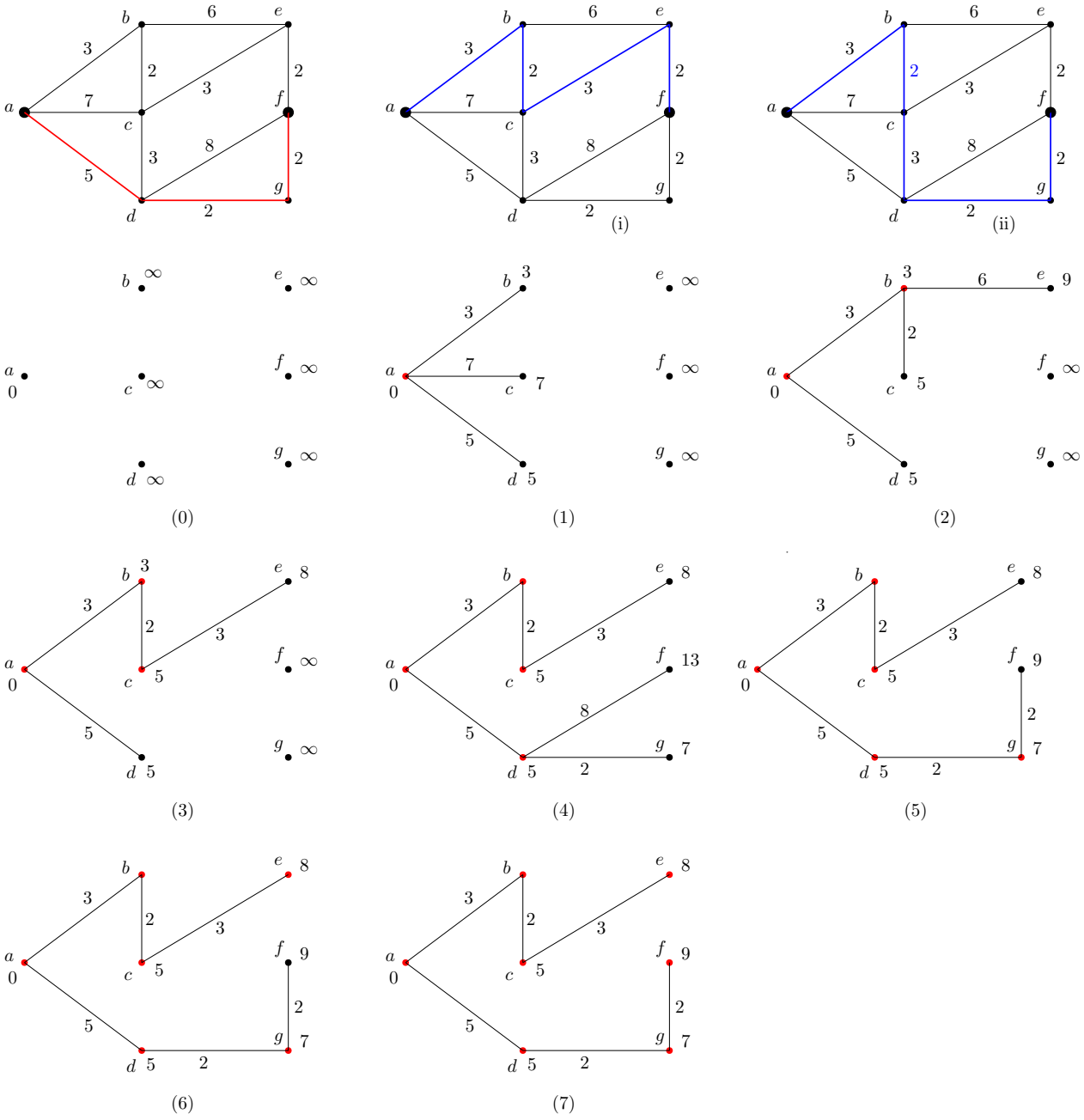

图 4.15: Dijkstra 算法示例。

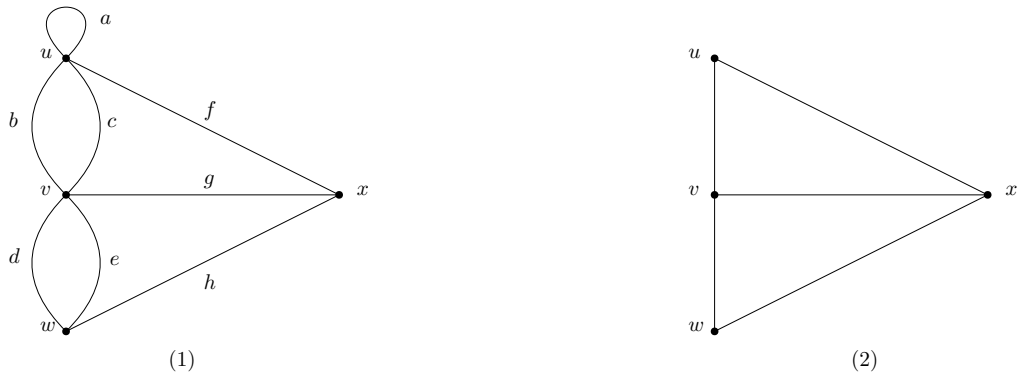

图 4.16: 图的矩阵表示例子。

4.4.2 邻接矩阵

我们知道方阵可以进行更多的计算，比如计算矩阵的幂，求特征值和特征向量等等。图的邻接矩阵，就是一个方阵。仍以图4.16为例。图（1）的邻接矩阵表示如下。

$$A = \begin{pmatrix} 2 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

具体地，以图4.16-（1）为例，邻接矩阵 A 的行分别表示顶点 u, v, w, x，同样地，A 的列也分别表示顶点 u, v, w, x。当 $u \neq v$ 时，$A(u, v)$ 表示顶点 u 到 v 之间的边的条数。在非简单图中，有时一个顶点有回到自己的边，比如图（1）中的边 a，这种边称为一个 loop。在邻接矩阵中，每个 loop 以 2 计算。比如在上面的例子中，$A(u, u) = 2$。图4.16-（2）是一个简单图，其邻接矩阵表示如下。

$$B = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

容易观察到如下事实：

- 邻接矩阵的每一行的和（或每一个列的和）等于对应顶点的度数；
- （无向图的）邻接矩阵是一个对称矩阵；
- 简单图的邻接矩阵里的元素要么是 0 要么是 1；
- 简单图的对角线元素全都是 0，等等。

下面我们介绍邻接矩阵的一些应用。

计算图中顶点之间的通路数

前面我们用 B 表示图4.16-（2）的邻接矩阵。直接计算可得

$$B^2 = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix}, \quad B^3 = \begin{pmatrix} 2 & 5 & 2 & 5 \\ 5 & 4 & 5 & 5 \\ 2 & 5 & 2 & 5 \\ 5 & 5 & 5 & 4 \end{pmatrix}, \dots$$

因为邻接矩阵 B 表示了一个图，所以 B^2 和 B^3 可能也有某种含义。稍加观察，不难发现，$B^t(i,j)$ 表示了从顶点 i 到 j 长度为 t 的通路有多少条。具体地：

- $B(i,j)$ 表示从 i 到 j 长度为 1 的通路有多少条，即 i 与 j 之间是否有边直接连接；
- $B^2(i,j)$ 表示从 i 到 j 长度为 2 的通路有多少条；比如：$B^2(v,v)=3$，图4.17-（1）中画出了对应的 3 条从 v 出发又回到 v 的通路；
- $B^3(i,j)$ 表示从 i 到 j 长度为 3 的通路有多少条；比如：$B^2(u,x)=5$，图4.17-（2）中画出了对应的 5 条从 u 出发到 x 的通路。

图 4.17: 计算图上的通路数。

进一步，设 $C = \sum_{1\le t\le k} B^t$。则 $C(i,j)$ 表示从 i 到 j 长度不超过 k 的通路数。比如，由前面的计算可知，从 u 到 x 的长度不超过 3 的通路有 $1+1+5=7$ 条。可以体会一下，通过计算矩阵的幂来得到图中的通路数（而不用直接盯着图去数），是多么简单优美！尤其是当图的顶点和边都很多的时候。

独立集的上界

定义 4.21. *设 $S \subseteq V$ 是图 $G=(V,E)$ 顶点集的一个子集。如果对任意的 $i,j \in S$，$(i,j) \notin E$，则 S 叫做 G 的一个（点）独立集.*

如果我们考虑一个社交网络图，边表示人之间的相互认识关系，那么独立集就是一群相互全都不认识的陌生人。在图4.16-（2）中，容易看出，$\{u,w\}$ 是一个最大的独立集，包含两个顶点。一般地，图的独立集能有多大呢？下面我们通过研究邻接矩阵给出独立集大小的一个上界。

首先我们把独立集表示为一个向量。给定独立集 $S \subseteq V$，定义函数

$$f_S : V \to \mathbb{R},$$
$$i \mapsto \begin{cases} 1, & i \in S, \\ 0, & i \notin S. \end{cases}$$

注意，f_S 也可以看成一个向量 $f_S \in \mathbb{R}^{|V|}$. 例如，在图4.16-（2）中，$S = \{u, w\}$ 是一个独立集，其对应的向量 $f_S = (1, 0, 1, 0) \in \mathbb{R}^4$.

现在考虑任意一个（无向简单）图 $G = (V, E)$，设 $S \subseteq V$ 是顶点集的一个子集，设 A 是 G 的邻接矩阵。

引理 4.22. *S 是一个独立集当且仅当 $f_S A f_S^T = 0$.*

证明. 直接计算可得，

$$f_S A f_S^T = \sum_{i \in V} f_S(i)(A f_S^T)(i) = \sum_{i \in V} f_S(i) \sum_{j \in V} A(i,j) f_S(j) = \sum_{i \in V} \sum_{j \in V} f_S(i) A(i,j) f_S(j).$$

注意，$f_S(i) \in \{0, 1\}$，$A(i,j) \in \{0, 1\}$，$f_S(j) \in \{0, 1\}$. 因此，$f_S A f_S^T = 0$ 当且仅当对任意 $(i,j) \in V \times V$，都有 $f_S(i) A(i,j) f_S(j) = 0$。

因此，若 $f_S A f_S^T = 0$，则当 $f_S(i) = 1$ 及 $f_S(j) = 1$ 时，即当 $i, j \in S$ 时，必定有 $A(i,j) = 0$，即 $(i,j) \notin E$。这就是说 S 是一个独立集。

反之，当 S 是一个独立集时，则有当 $f_S(i) = 1$ 及 $f_S(j) = 1$ 时，即 $i, j \in S$ 时，必定有 $(i,j) \notin E$，也即 $A(i,j) = 0$。因此，$f_S(i) A(i,j) f_S(j) = 0$ 总成立。因此，$f_S A f_S^T = 0$。 □

下面我们再计算与 f_S 有关的内积。回忆给定两个向量 $g, h \in \mathbb{R}^n$，其内积定义为

$$\langle g, h \rangle = \sum_{i=1}^{n} g(i) h(i). \tag{4.1}$$

用 $\vec{1}$ 表示全 1 的向量，即 $\vec{1} = (1, \ldots, 1)$.

引理 4.23. *对任意 $S \subseteq V$，有 $\left\langle f_S, \vec{1} \right\rangle = |S|$ 及 $\langle f_S, f_S \rangle = |S|$.*

证明. 直接计算. □

定义 4.24. *如果图 $G = (V, E)$ 的每个顶点度数都是 r，则图 G 叫做一个 r-正则图。*

定理 4.25 (Hoffman)**.** *设 $G = (V, E)$ 是一个 r-正则图，设 $S \subseteq V$ 是 G 的一个独立集。设 A 是 G 的邻接矩阵，$\lambda_{\max}$ 和 $\lambda_{\min}$ 分别是矩阵 A 的最大和最小的特征值。则*

$$|S| \leq \frac{\lambda_{\min}}{\lambda_{\min} - \lambda_{\max}} \cdot |V|.$$

证明. 为了符号上简化，设 $|V| = n$. 因此 $A_{n\times n}$ 是一个 $n \times n$ 的实对称矩阵。直接计算可得

$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = r \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

这表明 r 是 A 的一个特征值，其对应的特征向量是 $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$，将其单位化（按照内积(4.1)）后得的向量记为 U_1，即

$$U_1 = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n.$$

可以证明，r 事实上是 A 的最大的特征值（略）。用 $r = \lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = \lambda_{\min}$ 表示 A 的 n 个特征值，用 $U_1, U_2, \ldots, U_n$ 表示对应的单位正交化后的特征向量。即 $\langle U_i, U_i\rangle = 1$，当 $i \neq j$ 时，$\langle U_i, U_j\rangle = 0$。因为 A 是实对称矩阵，所以 $U_1, U_2, \ldots, U_n$ 是 $\mathbb{R}^n$ 的一组基。因此，向量 $f_S \in \mathbb{R}^n$ 可用 $U_1, U_2, \ldots, U_n$ 线性表示。设

$$f_S = \sum_{i=1}^{n} a_i U_i. \tag{4.2}$$

现在我们用 f_S 的线性表示(4.2)来重新计算引理 4.22和引理4.23的内容。计算如下。由引理 4.22，

$$0 = f_S A f_S^T = \left(\sum_{i=1}^{n} a_i U_i\right) A \left(\sum_{j=1}^{n} a_j U_j^T\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j U_i A U_j^T = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j U_i \lambda_j U_j^T = \sum_{i=1}^{n} a_i^2 \lambda_i. \tag{4.3}$$

由引理4.23，

$$|S| = \left\langle f_S, \vec{1}\right\rangle = \left\langle f_S, \sqrt{n}U_1\right\rangle = \sqrt{n} \cdot \langle f_S, U_1\rangle = \sqrt{n} \cdot \left\langle \sum_{i=1}^{n} a_i U_i, U_1 \right\rangle = \sqrt{n} \cdot a_1. \tag{4.4}$$

还有，

$$|S| = \langle f_S, f_S\rangle = \left\langle \sum_{i=1}^{n} a_i U_i, \sum_{j=1}^{n} a_j U_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j U_i U_j^T = \sum_{i=1}^{n} a_i^2. \tag{4.5}$$

根据(4.3)，(4.4)，及(4.5)可得

$$0 = \sum_{i=1}^{n} a_i^2 \lambda_i = a_1^2 \lambda_{\max} + \sum_{i=2}^{n} a_i^2 \lambda_i \geq \frac{|S|^2}{n}\lambda_{\max} + \lambda_{\min}\sum_{i=2}^{n} a_i^2 = \frac{|S|^2}{n}\lambda_{\max} + \lambda_{\min}\left(|S| - \frac{|S|^2}{n}\right).$$

注意，$n = |V|$。化简之后，即得所证。 □

像定理4.25这种利用图的矩阵表示对应的一些代数性质（如特征值等）来研究图的性质的方法叫做代数图论的方法，是图论研究的一个重要方向。

4.5 图的染色

我们先看下面一个具体的例子（引自 [1, Chapter 18]）。假设 A, B, C, D, E, F 6 个学生共用一个实验室，实验室里有 5 台设备，标号为 1,2,3,4,5，每个学生要做 1 个小时自己的实验，需要用到不同的设备，如表4.4所示。假设每台设备同时只能被一个学生使用，那么应该如何安排学生做实验的时间，总共需要多长时间能让所有学生都把实验做完呢?

表 4.4: 学生对设备的需求

	1	2	3	4	5
A	✓				
B	✓	✓	✓		
C	✓	✓	✓		
D		✓		✓	
E			✓		✓
F				✓	✓

我们可以把上面的问题用一个图来表示：每个同学用一个点表示，如果两个同学之间有设备冲突的情况，就在他们之间连一条边，如图4.18所示。因为每台设备同时只能被一个学生使用，因此，有设备冲突的同学不能同时做实验。换句话说，能同时做实验的同学应该是一个独立集。因此问题就是把同学按图4.18划分成不同的独立集，每个独立集里的同学可以同时做实验，总共需要的时间就是独立集的个数。显然，我们希望独立集的个数最少。图4.18给出了一个划分：有相同颜色的顶点划分为同一个独立集，总共有 3 个独立集，因此所有实验能在 3 小时内完成。具体调度安排为：第 1 小时 A，D，E 同时做实验，第 2 小时 B，F 同时做实验，第 3 小时 C 做实验。

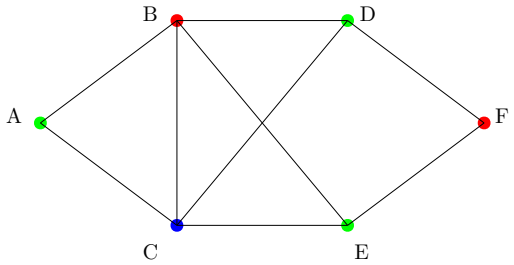

图 4.18: 用图染色解决资源冲突的问题。

以上的例子提示我们可以如下定义一般的图的（顶点）染色的问题。

定义 4.26. 给定图 $G = (V, E)$。如果映射 $\phi : V \to \mathbb{N}$ 满足条件：若 $(i, j) \in E$ 则 $\phi(i) \neq \phi(j)$，那么 ϕ 就叫做图 G 的一个染色。$|\phi(V)|$ 就是 ϕ 所用的颜色数。用 $\chi(G)$ 表示图 G 的染色所需要的最小的颜色数。

根据如上的定义，$\phi^{-1}(k)$ 是 G 的一个独立集，图的染色问题其实就是把图划分成互不相交的独立集。$\chi(G)$ 就是在这样的划分下独立集的最少个数。图4.18的染色数是 3. 前面的例子也告诉我们，比如

解决资源调度与冲突的问题中，求图 G 的染色数 $\chi(G)$ 是一个很重要的问题。历史上有著名的四色猜想，即：任意（平面）地图都可以用最多 4 种颜色染色。经过 100 多年的研究，该问题最终在上个世纪被解决，其证明依赖于大量的计算机计算。四色猜想（现已是四色定理）本身是要求对地图上的城市或国家染色，即对图的面染色。不过，给定一个地图后，我们可以把每个面转化为一个点，如果两个面相邻，则对应的两点之间连一条边。这样，问题就转化为给图的顶点的染色问题。因此，四色定理说明，任何一个平面图，都可以用 4 种颜色染色。

下面我们再看一些图染色的简单例子，如图4.19。

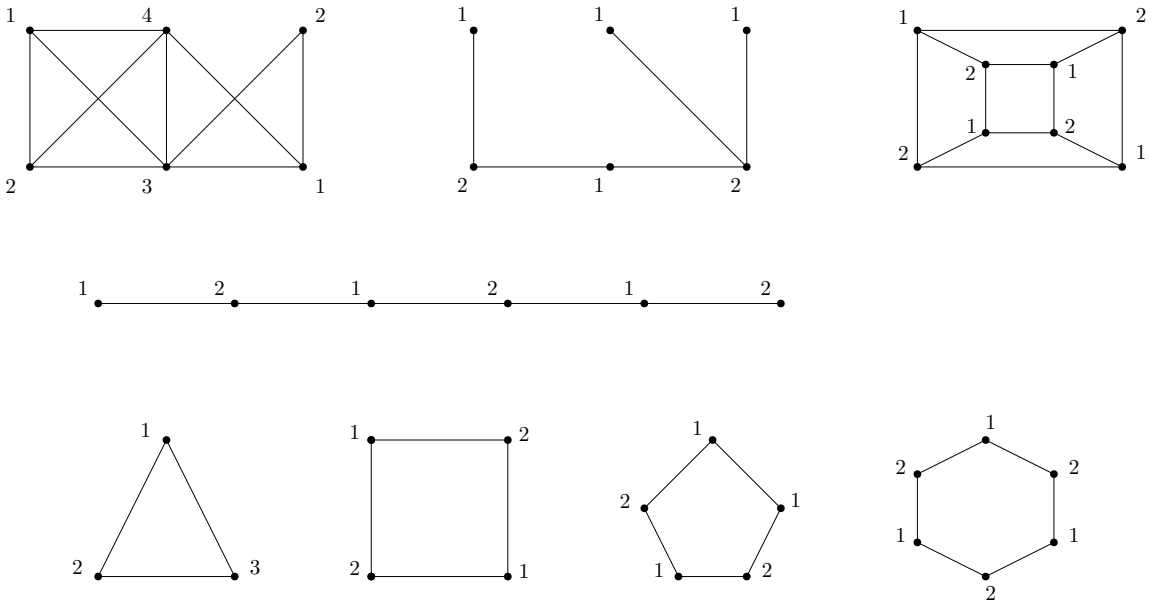

图 4.19: 图染色的例子。

定义 4.27. 如果图 $G=(V,E)$ 的任意两点之间都有边，则 G 叫做一个完全图。有 n 个顶点的完全图一般用 K_n 表示。如果图 G 的一个子图 H 是一个完全图，则 H 叫做 G 的一个团，H 的顶点个数叫做团的大小。用 $\omega(G)$ 表示图 G 的最大团的大小。

比如，在图4.19的第一个图中，有一个大小是 4 的团，即第一个图包含 K_4。根据图染色的定义，容易看出 $\chi(K_n)=n$.

定义 4.28. 用 $\Delta(G)$ 表示图 G 中顶点的最大度数。

命题 4.29. $\omega(G)\le\chi(G)\le\Delta(G)+1$.

证明. 若 $\omega(G)=t$，则 G 包含 K_t。因为 $\chi(K_t)=t$，因此 $\chi(G)\ge t$.

上界对图 G 的顶点个数用归纳法可证明。简述思路如下：任选图的一个顶点 v，考虑 G 去掉 v（及与 v 相连的边）后的子图 G'，很明显 $\Delta(G')\le\Delta(G)$. 根据归纳假设，$\chi(G')\le\Delta(G')+1\le\Delta(G)+1$. 也就是说，$G'$ 可用最多 $\Delta(G)+1$ 个颜色染色。现在假设 G' 已经染色完成。再考虑 G。G 只比 G' 多一个顶点 v，因此只需要考虑 v 的染色。因为 $d(v)\le\Delta(G)$，所以 v 的邻居最多用了 $\Delta(G)$ 这么多种

不同颜色，因为我们允许用 $\Delta(G)+1$ 这么多种颜色，因此还有一个颜色可选用来染色 v. 这样就完成了 G 的染色。 □

一般来说，给定图 G，要确定 $\chi(G)$ 是多少是很困难的。稍加思考，可知 $\Delta(G)$ 容易计算。因此，根据命题4.29，便容易得到图的染色数的一个上界。不过一般而言，$\omega(G)$ 也是很难计算的。

某些比较简单的图容易确定 $\chi(G)$。比如，从图4.19可以看出：有奇数个顶点的环的染色数是 3，有偶数个顶点的环的染色数是 2。

定义 4.30. 定义 $\chi(G)\leq 2$ 的图叫做二部图。

因为有奇数个顶点的环的染色数是 3，因此，二部图不可能包含有奇数个顶点的环。很棒的是，用这个条件就可以判定图是否是二部图。图4.19种的第 3 个图就是包含有偶数个顶点的环的二部图，注意它不包含奇数个顶点的环。

定理 4.31. 一个图是二部图当且仅当它不包含奇数个顶点的环。

证明. 前已说明二部图不可能包含奇数个顶点的环。仅需要证明如果一个图不包含奇数个顶点的环，那么它就是二部图，此留作练习。 □

判断一个图是否是二部图是比较容易的。然而，一般地，计算一个图的 $\chi(G)$ 是非常困难的。

回忆根据 $\chi(G)$ 的定义，图的染色把图的顶点分成互不相交的独立集。因此，由定理4.25可立即得出如下推论。

推论 4.32. 设 G 是正则图。在定理*4.25*的条件下，用与定理*4.25*一样的符号，则有 $\chi(G)\geq \frac{\lambda_{\min}-\lambda_{\max}}{\lambda_{\min}}$.

矩阵的特征值一般不难计算(或计算其近似值)。因此，根据推论4.32，我们也可以计算得到正则图的染色数的一个下界。根据命题4.29及推论4.32，就能比较快速的计算出正则图的染色数的上界 $\Delta(G)+1$ 和下界 $\frac{\lambda_{\min}-\lambda_{\max}}{\lambda_{\min}}$。

4.6 P 与 NP 问题

在 1999 年到 2000 年的世纪之交和千年之交的时候，数学家确定了新的时代最重大的七个数学问题，被称作千禧年问题，P 与 NP 问题便是其中之一。因此，可以说 P 与 NP 问题是离散数学中最重大的问题的之一。下面我们粗略地解释 P 与 NP 问题及其含义。

定义 4.33. **P** 是如下问题的集合：问题的输入的大小是 n，问题能在 n 的多项式 (比如：$3n^2+2n-5$) 时间内求解。

NP 是如下问题的集合：问题的输入的大小是 n，假设给出了问题的一个答案，该答案能在 n 的多项式时间内被验证。

根据定义，有 $\mathbf{P}\subseteq\mathbf{NP}$。直观地说，**P** 代表那些我们认为“容易”的问题，而 $\mathbf{NP}\setminus\mathbf{P}$ 代表那些我们认为“困难”的问题。注意，这里的容易和困难，指的是计算时间的长短，短时间内就可以计算出结果的，我们叫做容易，反之，则叫做困难。在理论上，如果计算的时间是输入大小 n 的多项式函数（比如

$3n^2+2n-5$）则认为是短时间，如果不是多项式函数，则认为是很长的时间，比如 $n^{\log n}$，或 2^n，或 $n!$ 等等。

下面是一些例子。

例 4.34. P 集合中一些问题的例子：

- n 个数字的排序：输入大小是 n，计算时间 $O(n^2)$；
- 判定一个数字 k 是否是素数：输入大小是 $t=\lceil \log_2 k\rceil$ 比特，*AKS* 算法（见第*1.1*章）运行时间是 $O((\lceil \log_2 k\rceil)^6)=O(t^6)$；
- 图的最小生成树：设图有 n 个顶点 m 条边，*Kruskal* 算法的运行时间是 $O(m^2+nm)$；
- 最短路问题：设图有 n 个顶点 m 条边，*Dijkstra* 算法的运行时间是 $O(nm)$；
- 构造 *Huffman* 编码：假设输入是 n 个字母和对应的权重，*Huffman* 编码的运行时间是 $O(n^3)$；
- 判定图是否存在欧拉回路：设图有 n 个顶点 m 条边，根据定理*4.7*，只需要检查图的顶点度数是否都是偶数，因此运行时间是 $O(nm)$；
- 判定图是否是二部图：设图有 n 个顶点 m 条边，根据定理*4.31*，计算时间是 $O(nm)$。

下面我们介绍 **NP** 集合中的一些问题。

例 4.35. NP 集合中的一些问题：

- 判断图中是否存在 *Hamilton* 回路（即，经过每个顶点刚好一次并且回到出发点）？如图*4.20*所示，红色与蓝色是两条 *Hamilton* 回路；
- 旅行商问题：对带权的图，给定一个长度 ℓ，判断是否存在长度不超过 ℓ 的 *Hamilton* 回路？如图*4.20*所示，如果 $\ell=38$，那么红色 *Hamilton* 回路满足条件（其长度是 37）；

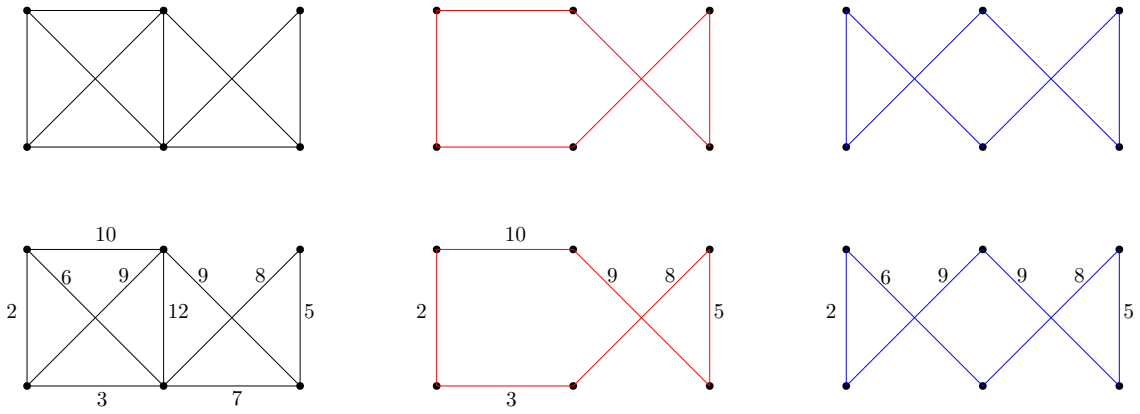

图 4.20: Hamilton 回路与旅行商问题。

- 图的最大独立集的大小：给定一个有 n 个顶点的图，给定数字 k，判断图中是否存在大小是 k 的独立集？

- 图的最大团的大小：给定一个有 n 个顶点的图，给定数字 k，判断图中是否存在大小是 k 的团？
- 图的染色问题：给定一个有 n 个顶点的图，给定数字 k，判断图中是否存在用 k 个颜色的染色方案？

注意，例4.35中的这些问题，如果给出一个答案，都容易验证答案正确与否，因此它们都属于 **NP**。

P 与 NP 问题： P = NP?

直观地说，**P** 与 **NP** 问题就是问，那些我们 "认为" 困难的问题是否是真的很困难，还是只是因为我们目前还没有找到快速的计算方法，如果继续研究就能找到在多项式时间内计算的方法？目前，大部分科学家倾向于猜测 $\mathbf{P} \neq \mathbf{NP}$，也有一部分科学家猜测 $\mathbf{P} = \mathbf{NP}$。

定义 4.36. **NP**-完全问题集是如下问题的集合：**NP** 中最难的那些问题。

NP-难问题集是如下问题的集合：该问题的难度不低于任何 **NP**-完全的问题。

图4.21给出了这些问题集的相互关系。

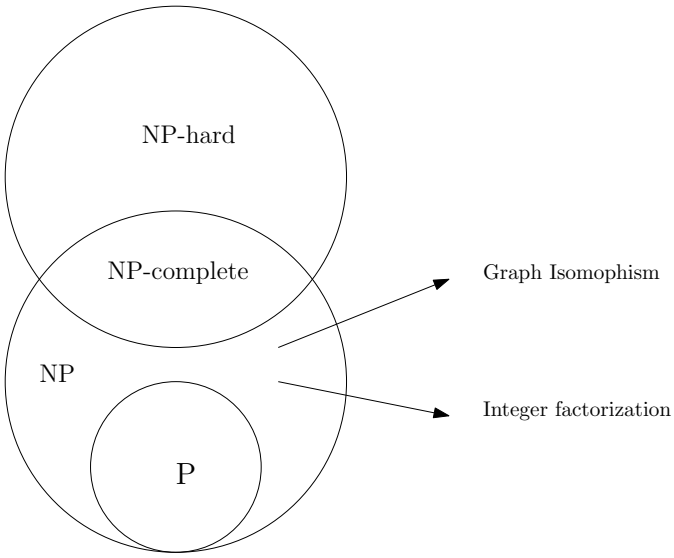

图 4.21: P 与 NP，NP-完全，及 NP-难。

事实上，例4.35中的那些问题都是 **NP**-完全问题。其对应的优化问题都是 **NP**-难问题。以旅行商问题为例，其对应的优化问题是指：找出带权图的最短的 Hamilton 回路。以染色数为例，其对应的优化问题是指：计算图的最少的染色数，即 $\chi(G)$ 的值。这些问题都是 **NP**-难问题。直观上说，就是它们（可能）都不能在多项式时间内求解。

几十年来，科学家们已经发现在数学、科学、工程等等领域许多的问题都是 **NP**-难问题。这表明，这些问题要精确求解可能都是非常困难的。因此，对这些问题，往往我们只能考虑求它们的近似解。对有的问题（比如最大独立集问题），甚至是求近似解都是很难的。**NP**-难的问题虽然本身很难精确求解，

但是我们却可以利用其本身难求解的特点来达成别的目标，比如加密。在实际工作中面临某个问题时，首先确定这个问题是否是 **NP**-难的，对我们考虑如何解决这个问题具有指导性的意义。因为这些原因，回答 **P** 是否等于 **NP** 是数学、科学、及工程中的一个具有基本重要性的问题。

前面举例说明了有许多问题属于 **P**，也有许多问题是 **NP**-完全问题。但是，迄今为止，科学家仍然不能确定给定两个图，判断其是否同构，这一问题到底是属于 **P**，还是一个 **NP**-完全问题。同样地，给定一个整数，将其分解为素数的乘积这一问题，我们也不知道它到底是属于 **P**，还是一个 **NP**-完全问题。判定不同的图是否同构在图像处理，蛋白质结构，社会网络等问题中均有重要有用。把整数分解为素数乘积也是一个很重要的问题，在密码学中被广泛使用。麻省理工大学的科学家 Shor 在上世纪 90 年代证明，使用量子计算机，可以在多项式时间内把整数分解为素数的乘积。因此，一旦有了可进行大规模计算的量子计算机，那么今天的许多密码都能被破解。

4.7 英文单词

- 图: graph，无向：undirected，有向：directed
- 相邻: adjacent
- 点：vertex，边：edge，面：face
- 度数：degree
- 简单图：simple graph
- 通路：walk
- 路径：path
- 回路：circuit
- 连通图：connected graph
- 子图：subgraph
- 导出子图：induced subgraph
- 连通分支：connected component
- 握手定理：handshaking theorem
- 树：tree
- 平面图：planar graph
- 环：cycle
- 权：weight
- 最小生成树：minimal spanning tree
- 贪心算法：greedy algorithm
- Huffman 编码：Huffman coding
- 香农熵：Shannon entropy
- 二叉树：binary tree
- 最短路：shortest path
- 关联矩阵：incidence matrix
- 邻接矩阵：adjacency matrix

- 独立集：independent set
- 正则图：regular graph
- 图染色：graph coloring
- 染色数：chromatic number
- 完全图：complete graph
- 团：clique
- 旅行商问题：Travelling salesman problem
- **NP**-完全：NP-complete
- **NP**-难：NP-hard

第五章　逻辑

> 自亚里士多德之后，逻辑学迄无新意。(*Nothing new had been done in Logic since Aristotle!*)
>
> ——哥德尔
>
> 所有错误皆由外因（如情感与教育）所致；推理本身并不出错。(*But every error is due to extraneous factors (such as emotion and education); reason itself does not err.*)
>
> ——哥德尔

5.1　命题逻辑

命题是指真或假唯一确定的陈述。命题逻辑研究把自然语言所表示的命题用数学符号表示并进行推理。当把最简单的不可分割的命题用数学符号如 p, q 等表示之后，可用表5.1中的联结词把简单的命题变量如 p, q 等结合在一起，就成为一个公式。公式本身是一个更为复杂的命题，我们需要通过计算或推理来判断其真假，这就是命题逻辑研究的基本问题。

表 5.1: 命题的联结词 $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \rightarrow q$	$p \leftrightarrow q$
0	0	1	0	0	1	1
0	1	1	0	1	1	0
1	0	0	0	1	0	0
1	1	0	1	1	1	1

例 5.1. 设 p: 海水是咸的; q: 苹果是固体; r: 牛顿相信上帝。则公式 $(p \wedge \neg r) \rightarrow q$ 翻译为自然语言如下: 如果海水是咸的并且牛顿不相信上帝，则苹果是固体。上述单个命题的真值是: $p = 1, q = 1, r = 1$, 因此 $p \wedge \neg r = 0$, 所以 $(p \wedge \neg r) \rightarrow q = 1$. 所以 "如果海水是咸的并且牛顿不相信上帝，则苹果是固体。"这句话在逻辑上是正确的（虽然从自然语言的角度看起来没有意义）。

5.1.1 公式与布尔函数

表5.1中的联结词都可以看成函数。其中，$\neg$ 是一个一元函数

$$\begin{aligned} f : \{0,1\} &\to \{0,1\} \\ 0 &\mapsto 1, \\ 1 &\mapsto 0. \end{aligned} \tag{5.1}$$

其余四个联结词都是二元函数。比如，$\to$ 是如下二元函数，

$$\begin{aligned} g : \{0,1\} \times \{0,1\} &\to \{0,1\} \\ (0,0) &\mapsto 1, \\ (0,1) &\mapsto 1, \\ (1,0) &\mapsto 0, \\ (1,1) &\mapsto 1. \end{aligned} \tag{5.2}$$

如果某个变量只取集合 $\{0,1\}$ 中的值，即变量的值只能是 0 或者 1，则这样的变量叫做*布尔变量*。我们也经常把 $\{0,1\} \times \{0,1\}$ 写成 $\{0,1\}^2$. 一般地，$f : \{0,1\}^n \to \{0,1\}$ 叫做*n-元布尔函数*。因为定义域 $|\{0,1\}^n| = 2^n$，值域 $|\{0,1\}| = 2$，所以，n-元布尔函数有 2^{2^n} 个. 本章后面提到函数时，一般均指布尔函数。

特别地，一元布尔函数有 4 个，二元布尔函数有 16 个。为什么表5.1中只有 1 个一元函数对应的联结词，只有 4 个二元函数对应的联结词呢？如果把 4 个一元布尔函数都写下来，是下面这 4 个：

$$\begin{array}{llll} 0 \mapsto 0, & 0 \mapsto 0, & 0 \mapsto 1, & 0 \mapsto 1, \\ 1 \mapsto 0, & 1 \mapsto 1, & 0 \mapsto 0, & 0 \mapsto 1. \end{array}$$

分别是如下 4 个函数（每一列表示一个函数）：$f(x) = 0$（恒等于 0 的常数函数），$f(x) = x$，$f(x) = \neg x$，$f(x) = 1$（恒等于 1 的常数函数）。因此，只需要一个符号 $\neg$. 二元函数有 16 个，不过，事实上，所有的二元函数，甚至所有的 n 元函数都可以用 $\neg, \wedge, \vee$ 来表示。因此，二元联结词其实用 $\wedge$ 和 $\vee$ 两个就够了。为了推理的方便和直观，我们也使用 $\to$（蕴含），$\leftrightarrow$（等价）这两个二元联结词。

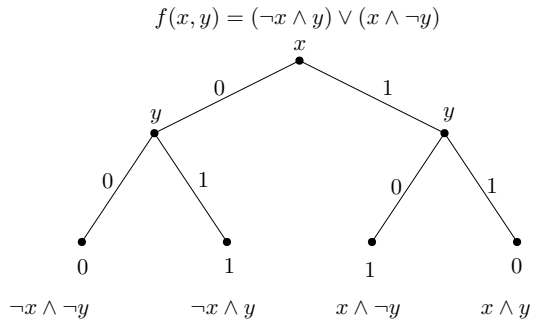

图 5.1: 把布尔函数用 $\neg, \wedge, \vee$ 表示

例 5.2. *考虑布尔函数 $f:\{0,1\}^2\to\{0,1\}$, 其定义为 $f(0,0)=0$, $f(0,1)=1$, $f(1,0)=1$, $f(1,1)=0$. 则 $f(x,y)$ 可表达如下: $f(x,y)=(\neg x\wedge y)\vee(x\wedge\neg y)$. 如图5.1所示。一般地, 任意 n-元布尔函数均可用如图5.1所示的方法用 $\neg,\wedge,\vee$ 来表示。*

5.1.2 公式的类型与真值表及逻辑计算

前面提到公式由一些布尔变量和联结词组成, 因次, 公式也可以看成布尔函数。如果一个公式不管其中的变量取什么值, 公式的值始终为真（函数值始终为 1), 这样的公式叫做*永真式*, 如果公式的值始终为假（函数值始终为 0), 这样的公式叫做*永假式*, 如果公式不是永假式, 则叫做*可满足式*。判断一个公式是否是可满足式是计算机科学中一个重大的基本问题。很明显, 因为公式可以看成一个函数, 我们可以把这个函数的每个值都计算出来, 这样就可以判断公式是哪种类型。比如, 对例5.1中的公式 $(p\wedge\neg r)\to q$ 直接计算可得表5.2。像表5.2这样的把公式的所有值都列出来的表叫做公式的*真值表*。根据真值表5.2可知, 公式 $(p\wedge\neg r)\to q$ 是可满足式。该公式有 3 个变量, 因此有 $2^3=8$ 个值。一般地, 如果公式有 n 个变量, 那么真值表就有 2^n 个值。有没有别的办法可以更简洁地判断一个公式是否是哪种类型呢? 事实上, 可以证明判断一个公式是否是可满足式是也是一个 **NP**-完全问题。换句话说, 迄今科学家还没有找到比计算真值表简单许多的方法, 而且大部分科学家倾向于认为并不存在特别简单的方法。

表 5.2: $(p\wedge\neg r)\to q$ 的真值表

$pqr=$???	000	001	010	011	100	101	110	111
$(p\wedge\neg r)\to q$	1	1	1	1	0	1	1	1

除了计算真值表外, 有时也可以通过逻辑计算来判断一个公式的类型。另外, 有时我们也经常需要判断两个公式是否是一样的, 也就是说, 两个公式是否表达同一个函数。显然, 我们也可以把两个公式的真值表各自计算出来比较即可。但有时候用逻辑计算就可以得到结果。常用的逻辑计算（推理）可参见中文课本 p.21, 其中, 应该熟练掌握德摩根律

$$\neg(p\wedge q)=\neg p\vee\neg q,\quad \neg(p\vee q)=\neg p\wedge\neg q,$$

及如何去掉蕴含

$$p\to q=\neg p\vee q.$$

比如, 我们可以把例5.1中的公式 $(p\wedge\neg r)\to q$ 简化成只用 $\neg,\wedge,\vee$ 来表达:

$$(p\wedge\neg r)\to q=\neg(p\wedge\neg r)\vee q=\neg p\vee r\vee q.$$

这同时也表明, 公式 $(p\wedge\neg r)\to q$ 和公式 $\neg p\vee r\vee q$ 是等价的。注意, 从后者可以直接看出, 只有当 $pqr=100$ 时, 公式的值为 0, 而从最初的公式则并不那么容易直接看出这一点。

5.1.3 联结词的完备集

回顾例5.2, 用图5.1我们表明了任何布尔函数都可以用 $\neg,\wedge,\vee$ 来表达。如果一组联结词能用来表达任意的布尔函数, 则称其为一个完备集。因此, $\{\neg,\wedge,\vee\}$ 是一个完备集。根据德摩根律, $\wedge$ 总可以转

换为 $\neg$ 及 $\vee$，比如 $p \wedge q = \neg(\neg p \vee \neg q)$. 因此，当用 $\neg, \wedge, \vee$ 表示了某个函数后，我们可以继续利用德摩根律把其中的 $\wedge$ 都去掉，这样得到的新的公式就只有 $\neg, \vee$. 因此，$\{\neg, \vee\}$ 也是一个完备集。类似地（或者根据 $\wedge$ 与 $\vee$ 的对称性），$\{\neg, \wedge\}$ 也是一个完备集。

我们看到，从完备集 $\{\neg, \wedge, \vee\}$ 中去掉 $\wedge$ 或 $\vee$ 后仍然是完备集。那么去掉 $\neg$ 呢？即，$\{\wedge, \vee\}$ 是完备集吗？可以证明，它不是完备集。事实上，$\wedge$ 与 $\vee$ 两个函数本身都是单调递增函数，因此，由它们组成的函数也都是单调递增函数。所以，比如，例5.2中的函数就不能指用 $\wedge$ 和 $\vee$ 来表达。

容易看出，单独的 $\neg$ 明显不是完备集（因为它只是一个一元函数）。那么，有没有可能只用一个二元函数的完备集呢？答案是肯定的。我们已经知道 $\{\neg, \wedge\}$ 是一个完备集，因此只要找到一个二元函数，用它既可以表达 $\neg$，也可以表达 $\wedge$ 就可以了。考虑如下二元函数 $f(x, y) = \neg(x \wedge y)$，用符号 $x \uparrow y$ 来表达这个函数。用它可以表达 $\neg$ 和 $\wedge$，如下：

$$\neg x = \neg(x \wedge x) = x \uparrow x, \quad x \wedge y = \neg(\neg(x \wedge y)) = \neg(x \uparrow y) = (x \uparrow y) \uparrow (x \uparrow y). \tag{5.3}$$

根据(5.3)，$\{\uparrow\}$ 是一个完备集。类似地，如果定义 $x \downarrow y = \neg(x \vee y)$，也可以证明 $\{\downarrow\}$ 也是一个完备集。$\uparrow, \downarrow$ 分别称为与非和或非。这里的讨论表明，用与非，或者或非，仅仅一个二元函数，就可以表示任意布尔函数。这一点在逻辑电路设计中非常有用。当然，如果完备集太小，函数的表达形式可能就较为复杂，如(5.3)。因此，有时需要在选取较小的完备集及函数的表达式不过与复杂之间取得平衡。

5.1.4 析取范式与合取范式

回顾例5.2，用图5.1我们表明了任何布尔函数都可以用 $\neg, \wedge, \vee$ 来表达，而且，所得到的表达式是一种特殊类型：每一个括号里都是用 $\neg, \wedge$ 表达的，而括号之间用 $\vee$ 联结。这种公式叫做*析取范式*。比如，下面的公式也是析取范式：

$$(x \wedge \neg y \wedge z) \vee (\neg x \wedge y \wedge z) \vee (x \wedge y) \vee (\neg z).$$

每个括号里的部分称作一个句子（clause）。析取范式就是说每个句子都是用 $\neg, \wedge$ 来表达的，而不同句子用 $\vee$ 联结起来。把公式写成析取范式的好处是可以立即看出函数在变量取什么值时函数值是 1。比如，在上面的例子中，当

$$xyz = 101, \quad xyz = 011, \quad xy = 11, \quad z = 0$$

时，函数的值为 1.

相反地，形如

$$(x \vee \neg y \vee z) \wedge (\neg x \vee y \vee z) \wedge (x \vee y) \wedge (\neg z)$$

的公式称作*合取范式*。把公式写成合取范式的好处是可以立即看出函数在变量取什么值时函数值是 0。比如，在上面的例子中，当

$$xyz = 010, \quad xyz = 100, \quad xy = 00, \quad z = 1$$

时，函数的值为 0.

因此，把公式改写成析取范式或合取范式，也有助于帮助我们判断公式是哪种类型（永真式，永假式，还是可满足式）。

5.1.5 消解法

最后，我们介绍用消解法来简化某些公式。消解法基于如下基本的事实.

命题 5.3. 如果 $(x \vee A) \wedge (\neg x \vee B)$ 可满足，则 $A \vee B$ 也可满足。反之，如果 $A \vee B$ 可满足，则 $(x \vee A) \wedge (\neg x \vee B)$ 也可满足。

证明. 假设 $A \vee B$ 不可满足，则 $A = 0, B = 0$. 此时，$(x \vee A) \wedge (\neg x \vee B) = x \wedge \neg x = 0$，也不可被满足。

反之，如果 $A \vee B$ 可满足，则 A, B 中至少有一个为 1。不失一般性，假设 $A = 1$. 此时，取 $x = 0$，则有 $(x \vee A) \wedge (\neg x \vee B) = 1$. 也就是说，公式 $(x \vee A) \wedge (\neg x \vee B)$ 也能被满足。 □

消解规则经常写成如下形式：

$$\begin{array}{c} x \vee A \\ \neg x \vee B \\ \hline A \vee B \end{array} \tag{5.4}$$

也就是说，可以把变量 x 消去。

注 5.4. 注意：用消解规则得到的新的公式与原来的公式并不表示同一个函数，不过，它们有相同的可满足性。因此，消解法一般用来判断一个公式是否是可满足式。如果通过消解法能推导出永假式，则原公式也是永假式，即不可被满足。如果通过消解法不能推到出永假式，则原公式是可满足式。

例 5.5. *(1)* 判断 $(\neg p \vee q) \wedge (p \vee q) \wedge (\neg q)$ 是否是可满足式。

运用消解法计算如下：

$$\begin{array}{c} \neg p \vee q \\ p \vee q \\ \hline q \vee q = q \\ \neg q \\ \hline 0. \end{array}$$

因为最后我们得到 0，即永假式，因此原式不可满足，也是永假式。

(2) 判断 $p \wedge (p \vee q) \wedge (p \vee \neg q) \wedge (q \vee \neg r) \wedge (q \vee r)$ 是否是可满足式。

运用消解法计算如下：

$$\begin{array}{c} p \vee q \\ p \vee \neg q \\ \hline p \vee p = p. \end{array}$$

再运用消解法有

$$\begin{array}{c} p \vee \neg q \\ q \vee \neg r \\ \hline p \vee \neg r. \end{array}$$

再运用消解法有

$$\begin{array}{c} p \vee \neg q \\ q \vee r \\ \hline p \vee r. \end{array}$$

再运用消解法有

$$\begin{array}{c} q \vee \neg r \\ q \vee r \\ \hline q \vee q = q. \end{array}$$

再运用消解法有

$$\begin{array}{c} p \vee \neg r \\ p \vee r \\ \hline p \vee p = p. \end{array}$$

再运用消解法有

$$\begin{array}{c} p \vee \neg r \\ q \vee r \\ \hline p \vee q. \end{array}$$

现在已经没有可再用的消解，注意我们没有得到任何永假式，因此原公式是可满足的。

5.1.6 推理

to write

5.2 一阶逻辑

to write

5.3 英文单词

- 逻辑：logic
- 布尔函数：Boolean function
- 真值表：truth table
- 永真式：tautology，永假式：contradiction (unsatisfiable)，可满足：satisfiable
- 析取范式：disjunctive normal form(DNF)，合取范式：conjunctive normal form(CNF)
- 消解：resolution

第六章　代数结构

数学家专注于寻找代数方程的求根公式，有些人已尝试证明其不存在。然而，如果我没弄错的话，这一尝试尚未成功。因此，我斗胆希冀数学家以善意接收这份报告，其目的是填补代数方程理论的这一空白。(*The mathematicians have been very much absorbed with finding the general solution of algebraic equations, and several of them have tried to prove the impossibility of it. However, if I am not mistaken, they have not as yet succeeded. I therefore dare hope that the mathematicians will receive this memoir with good will, for its purpose is to fill this gap in the theory of algebraic equations.*)

——*Niels Henrik Abel*（阿贝尔）

写作者对读者伤害最甚者莫过于把困难隐而不宣。(*un auteur ne nuit jamais tant à ses lecteurs que quand il dissimule une difficulté*)

——*Evariste Galois*（伽罗瓦）

6.1　群

6.1.1　定义及例子

我们先给出群的定义，然后从诸多的例子中体会群的含义。

定义 6.1. 设 G 是一个集合，$*$ 是 G 上的一个二元运算，即，对任意的 $a,b\in G$，有 $a*b\in G$. 如果 $(G,*)$ 还满足以下条件，则 $(G,*)$ 叫做一个群：

- 运算的结合律：对任意的 $a,b,c\in G$，有 $(a*b)*c=a*(b*c)$；
- 存在单位元，即，存在一个元素 $e\in G$，满足：对任意的 $a\in G$，都有 $a*e=e*a=a$；
- 每个元素都存在逆元，即，对任意的 $a\in G$，都有另外一个元素 $b\in G$ 满足 $ab=ba=e$，一般把 b 记作 a^{-1}.

如果运算 $*$ 还满足交换律，即，对任意的 $a,b\in G$，都有 $a*b=b*a$，则群叫做交换群（也叫阿贝尔群），否则叫非交换群。

如果 G 的元素个数是无限的，则叫无限群，否则 G 是有限群，G 的元素个数 $|G|$ 叫做群的阶。

定义 6.2. 设 $(G,*)$ 是一个群，$H\subseteq G$ 是一个子集. 如果 $(H,*)$ 也是一个群，则 $(H,*)$ 叫做 $(G,*)$ 的子群，记作：$(H,*)\leq(G,*)$.

例 6.3. $(\mathbb{Z},+)$ 是一个群，叫做整数的加法群。容易验证其满足定义*6.1*的要求：

- $\forall\ a,b\in\mathbb{Z}$，$a+b\in\mathbb{Z}$；
- $\forall\ a,b,c\in\mathbb{Z}$，$(a+b)+c=a+(b+c)$；
- 单位元 $e=0$；
- $\forall\ a\in\mathbb{Z}$，a 的逆元是 $-a$.

容易验证，$(\mathbb{Z},-)$，$(\mathbb{Z},\times)$，$(\mathbb{Z},\div)$ 都不是群。这说明，同一个集合是否构成群，非常依赖于其上的运算是什么。

另外，容易验证，$(\mathbb{Q}-\{0\},\times)$ 也是一个群，即有理数上的乘法群。因此，虽然 $(\mathbb{Z},\times)$ 不是群，$(\mathbb{Z}-\{0\},\times)$ 也不是群，但是我们把集合从 $\mathbb{Z}-\{0\}$ 扩大至 $\mathbb{Q}-\{0\}$ 后，在乘法下，$(\mathbb{Q}-\{0\},\times)$ 就是一个群。容易验证，$(\mathbb{R}-\{0\},\times)$ 也是一个群。

这个例子的群都是交换群，也都是无限群，并且 $(\mathbb{Q}-\{0\},\times)\leq(\mathbb{R}-\{0\},\times)$.

例 6.4. 定义记号

$$M(r,s)=\text{所有 } r\times s \text{ 的实矩阵.}$$

则，$(M(r,s),+)$ 是一个交换群，其中运算 $+$ 是矩阵的加法，这个群的单位元是 $r\times s$ 的 0 矩阵。

考虑 $M(r,r)$，这些方阵可以相乘，用 $\times$ 表示矩阵的乘法，注意到 $(M(r,r),\times)$ 不是群：在矩阵乘法下，单位元存在，即 $r\times r$ 的单位矩阵，但是并非所有 $r\times r$ 的矩阵都有逆元（即，并非所有 $r\times r$ 的矩阵都可逆）。

定义记号

$$N(r,r)=\text{所有 } r\times r \text{ 的实可逆矩阵.}$$

则，$(N(r,r),\times)$ 是一个群。注意，因为矩阵乘法一般不满足交换律，因此，$(N(r,r),\times)$ 是非交换群

定义记号

$$N'(r,r)=\{A\in N(r,r):|A|=1,\forall\ 1\leq i,j\leq r,A(i,j)\in\mathbb{Z}\}.$$

则，$(N'(r,r),\times)$ 也是一个非交换群（请自行验证，确定你明白为什么 $(N'(r,r),\times)$ 也是一个群），这个群在数学中非常重要，通常记作 $SL(r,\mathbb{Z})$，叫做一般线性群，并且有 $SL(r,\mathbb{Z})\leq N(r,r)$.

例 6.5. 用记号

$$\mathbb{Z}_k=\{0,1,2,\ldots,k-1\}.$$

用 $+$ 表示模 k 的加法，用 $\times$ 表示模 k 的乘法。容易验证 $(\mathbb{Z}_k,+)$ 是群，其单位元是 0，i 的逆元是 $k-i$。这是一个 k 阶有限交换群。不过，$(\mathbb{Z}_k,\times)$ 不是群：它有单位元 1，但是有些元素并没有逆元，特别地，0 没有逆元。

思考：如果去掉 0，$(\mathbb{Z}_k-\{0\},\times)$ 一定是一个群吗？

例 6.6. 考虑三角形的沿中心旋转，及沿着高线反射，如图*6.1* 所示。设

$$G=\{r_1,r_2,r_3,s_1,s_2,s_3\}.$$

下面定义一个 G 上的运算 $\circ$ 如下：对任意 $a,b\in G$，定义 $a\circ b$ 为先做 b 再做 a. 下面我们来看运算的结果。例如，如图*6.2*，$r_1\circ s_2=s_1$，$s_1\circ s_2=r_1$，等。通过具体计算，可以写出下面的运算表。容易看出，$e=r_3$ 是单位元。

表 6.1: $(G,\circ)$ 的运算，$e=r_3$

$a\circ b$	r_1	r_2	e	s_1	s_2	s_3
r_1	r_2	e	r_1	s_3	s_1	s_2
r_2	e	r_1	r_2	s_2	s_r	s_1
e	r_1	r_2	e	s_1	s_2	s_3
s_1	s_2	s_3	s_1	e	r_1	r_2
s_2	s_3	s_1	s_2	r_2	e	r_1
s_3	s_1	s_2	s_3	r_1	r_2	e

通过表中的运算，容易验证，结合律成立。比如

$$(r_1\circ s_2)\circ s_3=s_1\circ s_3=r_2=r_2\circ r_1=r_1\circ(s_2\circ s_3),$$
$$(s_1\circ r_2)\circ r_1=s_3\circ r_1=s_1=s_1\circ e=s_1\circ(r_1\circ r_1),$$

等等。单位元是 $e=r_3$。表*6.1*也给出了 G 中每个元素的逆元：

$$r_1^{-1}=r_2,\quad r_2^{-1}=r_1,\quad r_3^{-1}=r_3,\quad s_1^{-1}=s_1,\quad s_2^{-1}=s_2,\quad s_3^{-1}=s_3.$$

综上，我们说明了 $(G,\circ)$ 是一个 6 阶非交换群。

事实上，$(G,\circ)$ 可以用置换的语言来更简单地表示。记 $[n]=\{1,2,\dots,n\}$. 置换就是一个从 $[n]$ 到它自身的双射 (一一映射)。由组合知识知道，这样的双射共有 $n!$ 个。用 S_n 表示所有这 $n!$ 个置换。比如 $|S_3|=3!=6$. 用如下的记号来表示这 6 个置换：

$$\rho_1=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\quad \rho_2=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\quad \rho_3=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},$$

及

$$\sigma_1=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\quad \sigma_2=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\quad \sigma_3=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}.$$

比如，置换 ρ_1 就是表示如下双射函数：$1\mapsto 2,2\mapsto 3,3\mapsto 1$. 因为置换就是双射函数，所以我们可以考虑置换的复合，即复合函数，这就定义了 S_n 上的一个运算。比如，

$$\rho_1\sigma_2=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\sigma_1.$$

又比如，

$$\sigma_1\sigma_2=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=\rho_1.$$

事实上，如果我们定义如下集合间的双射：$\phi : G \to S_n$，

$$r_1 \mapsto \rho_1, \quad r_2 \mapsto \rho_2, \quad r_3 \mapsto \rho_3, \quad s_1 \mapsto \sigma_1, \quad s_2 \mapsto \sigma_2, \quad s_3 \mapsto \sigma_3,$$

则可以验证如下性质：

$$\forall\ a, b \in G, \quad \phi(a \circ b) = \phi(a)\phi(b).$$

换句话说，G 上的运算 $\circ$ 与 S_3 上的置换的复合运算是“一样的”。如果我们用 $\cdot$ 来表示 S_3 上的复合运算，那么这就说明 $(S_3, \cdot)$ 也是一个 6 阶非交换群，叫做 6 阶置换群。置换群在群论中非常重要，一般地，$(S_n, \cdot)$ 都是群，一般仅用记号 S_n 代表这个群. S_n 是一个 $n!$ 阶的非交换群.

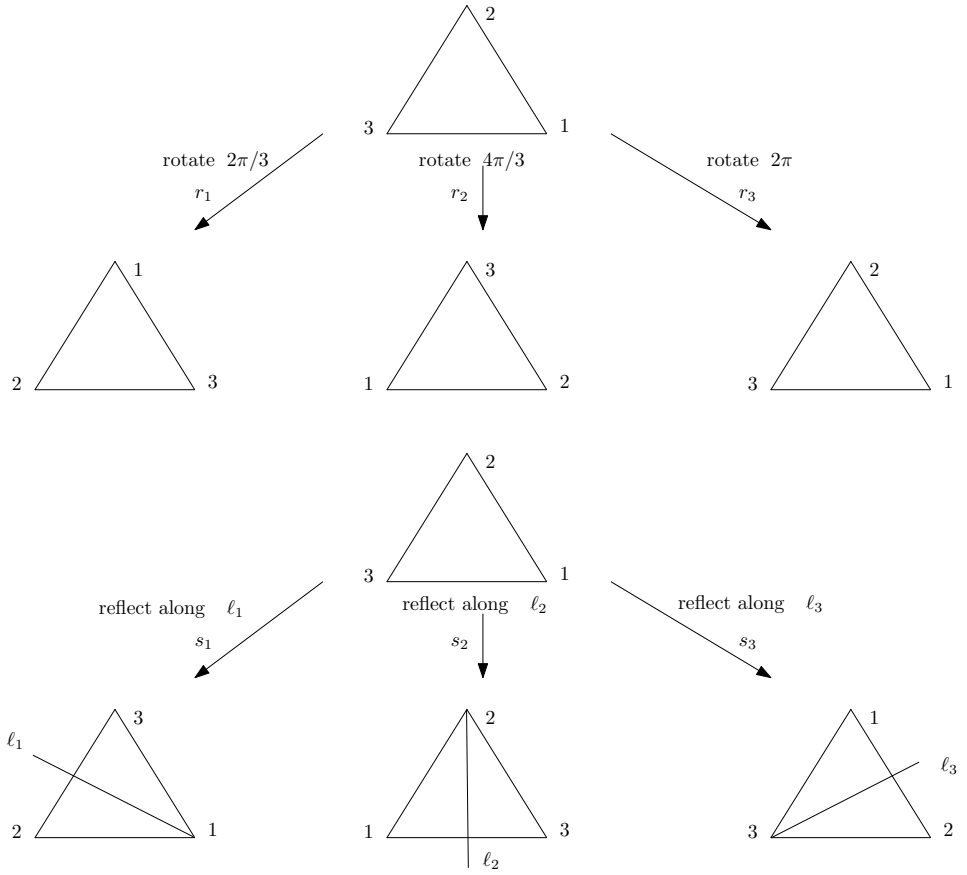

图 6.1: 三角形的旋转与反射

例 6.7. 用 P_d 表示所有次数不超过 d 的整系数多项式，那么 $(P_d, +)$ 是一个交换群，其中 $+$ 是指多项式的加法，即，如果 $f(x) = \sum_{i=0}^{d} a_i x^i$，$g(x) = \sum_{i=0}^{d} b_i x^i$，则

$$(f + g)(x) = \sum_{i=0}^{d} (a_i + b_i) x^i.$$

这个群的单位元是什么？f 的逆元是什么？

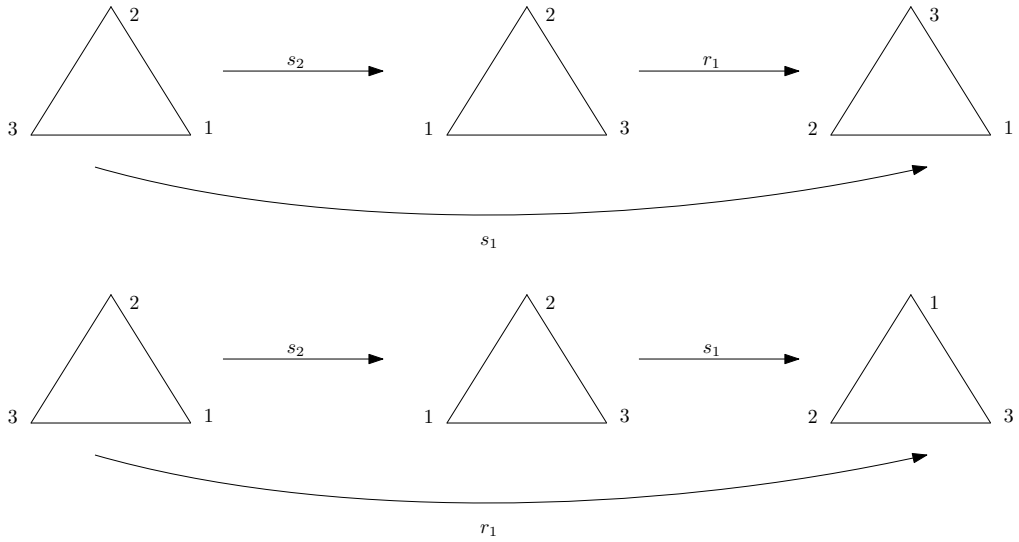

图 6.2: 旋转与反射的复合运算

例 6.8. 定义

$$C_0 = \{f : \mathbb{R} \to \mathbb{R}, f\text{是连续函数}\}.$$

那么 $(C_0, +)$ 是一个交换群，其中 $+$ 是指函数的加法。

例 6.9. 考虑布尔变量 $\mathbb{Z}_2 = \{0, 1\}$. 用 $+$ 表示布尔运算，回忆布尔运算如下

$$0 + 0 = 0, \quad 0 + 1 = 1 + 0 = 11 + 1 = 0.$$

因此，$(\mathbb{Z}_2, +)$ 是一个 2 阶交换群.

一般地，考虑 $\mathbb{Z}_2^n = \{0, 1\}^n$，其上的加法 $+$ 类似定义为逐位按照布尔变量的加法相加。例如，设 $n = 3$,

$$101 + 011 = 110, \quad 110 + 011 = 101,$$

则，$(\mathbb{Z}_2^n, +)$ 也是一个交换群，其阶是 2^n.

例 6.10. 前面我们学习过图的同构。设 $G = (V, E)$ 是一个图，如果 $\phi : G \to G$ 是一个从 G 到其自身的同构，则称 ϕ 是图 G 的一个自同构。设

$$Aut(G) = \{\phi : G \to G, \phi \text{ 是图 } G \text{ 的一个自同构}\}.$$

因为同构是一个映射，所以可以考虑映射的复合运算。考虑 $Aut(G)$ 上的映射的复合运算 $\circ$，即，若 $\phi, \varphi \in Aut(G)$，则 $(\phi \circ \varphi)(v) = \phi(\varphi(v))$. 可以验证（留作练习），$(Aut(G), \circ)$ 也是一个群，叫做图 G 的自同构群。

例 6.11. 用 $\mathbb{C}$ 表示所有复数的集合。设 $\phi : \mathbb{C} \to \mathbb{C}$ 是一个双射，并且满足如下条件：

- $\forall\ a, b \in \mathbb{C}$，$\phi(ab) = \phi(a)\phi(b)$,

- $\forall\ a,b\in\mathbb{C},\ \phi(a+b)=\phi(a)+\phi(b)$.

如果 ϕ 满足如上性质，则叫做 $\mathbb{C}$ 上的一个自同构。可以证明如下结论。

引理 6.12. 如果 ϕ 是 $\mathbb{C}$ 上的一个自同构，则，$\forall\ a\in\mathbb{Q}$，有，$\phi(a)=a$.

证明. 留作练习。 □

命题 6.13. 设 $P(x)=\sum_{i=0}^{n}a_ix^i$ 是一个整系数的 n 次多项式，即 $\forall\ i=0,1,\ldots,n$，都有 $a_i\in\mathbb{Z}$. 设 ϕ 是 $\mathbb{C}$ 上的一个自同构，则如果 t 是 $P(t)$ 的一个根，那么 $\phi(t)$ 也是 $P(t)$ 的一个根.

证明. 要说明 $\phi(t)$ 也是 $P(t)$ 的一个根，就是要说明 $P(\phi(t))=0$. 根据 ϕ 是 $\mathbb{C}$ 上的一个自同构，可得 $\phi(t^i)=\phi(t)^i$. 又因为 $a_i\in\mathbb{Z}$，由引理6.12得 $\phi(a_i)=a_i$. 下面直接计算

$$P(\phi(t))=\sum_{i=0}^{n}a_i\phi(t)^i=\sum_{i=0}^{n}\phi(a_i)\phi(t^i)=\sum_{i=0}^{n}\phi(a_it^i)=\phi\left(\sum_{i=0}^{n}a_it^i\right)=\phi(0)=0,$$

最后一步 $\phi(0)=0$ 也是用了引理6.12. □

代数基本定理告诉我们，n 次多项式 $P(x)$ 在 $\mathbb{C}$ 上必定有 n 个根。设这 n 个根是

$$x_1,x_2,\ldots,x_n.$$

设 ϕ 是 $\mathbb{C}$ 上的一个自同构，命题*6.13*表明，

$$\phi(x_1),\phi(x_2),\ldots,\phi(x_n)$$

也是多项式 $P(x)$ 的根。注意，自同构是双射，因此 $\phi(x_1),\phi(x_2),\ldots,\phi(x_n)$ 也就是多项式 $P(x)$ 的 n 个根. 换句话说，如果仅考虑 ϕ 在 $\{x_1,x_2,\ldots,x_n\}$ 上的映射，则

$$\phi:\{x_1,x_2,\ldots,x_n\}\to\{x_1,x_2,\ldots,x_n\}$$

是一个置换。这样，对 n 次多项式 P，可以定义如下一个集合

$$G_p=\{\sigma\in S_n:\sigma\text{ 是某个 }\mathbb{C}\text{ 上的自同构诱导的 }\{x_1,x_2,\ldots,x_n\}\text{ 上的置换}\}.$$

容易看出，自同构的复合仍然是自同构，恒等映射是自同构，如果某个自同构在 $\{x_1,x_2,\ldots,x_n\}$ 上诱导的置换是 σ，则其逆在 $\{x_1,x_2,\ldots,x_n\}$ 上诱导的置换就是 σ^{-1}. 因此，$(G_P,\cdot)\le(S_n,\cdot)$ 也是一个群，这个群叫做多项式 P 的伽罗瓦（*Galois*）群. 伽罗瓦群的发现是伽罗瓦在证明五次及以上方程没有求根公式的重要一步. 下面我们看一些例子.

- $P(x)=x^2-1$. 方程的两个根是 $-1,1$，都是有理数. 根据引理*6.12*，任何 $\mathbb{C}$ 上的自同构 ϕ 都满足 $\phi(-1)=-1$ 且 $\phi(1)=1$. 因此，

  $$G_P=\left\{\begin{pmatrix}1&2\\1&2\end{pmatrix}\right\},$$

  即，它仅包含恒等置换，$|G_P|=1$.

- $P(x)=x^3+x$. 方程的三个根分别是 $0,i,-i$. 由于任何自同构 ϕ 都满足 $\phi(0)=0$，并且因为自同构是双射，因此，仅有两种选择

$$\phi(i)=i,\quad \phi(-i)=-i,$$

或者

$$\phi(i)=-i,\quad \phi(-i)=i.$$

前者就是恒等置换，后者可由 $\mathbb{C}$ 上的如下同构得到

$$\phi:\mathbb{C}\to\mathbb{C},\quad a+bi\mapsto a-bi,$$

即，ϕ 把每个复数映射到它的共轭. 可以验证，如上共轭映射是一个 $\mathbb{C}$ 上的自同构. 因此，

$$G_P=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\right\}.$$

6.1.2 拉格朗日定理及子群的判定

上一节我们介绍了群的许多例子，并了解了群，无限群，有限群，子群，及群的阶的概念。本节介绍有限群的一个基本定理：拉格朗日定理。为了简便，有时我们不再写出群的运算。

定理 6.14 (拉格朗日)**.** 设 G 是有限群，$H\le G$，则 $|H|$ 整除 $|G|$.

拉格朗日定理可以快速的帮我们判断一个有限群不可能有某些阶的子群，比如，如果 G 是 6 阶群，则 G 没有 4 阶子群.

证明. 为了证明拉格朗日定理，我们考虑陪集这一概念. 对任意 $a\in G$，定义

$$Ha=\{ha:h\in H\},$$

叫做 H 关于 a 的右陪集. 以下我们证明：

(1) $|Ha|=|H|$；

(2) 对任意 $a,b\in G$，要么 $Ha=Hb$，要么 $Ha\cap Hb=\emptyset$.

注意，因为 $e\in H$，所以 $a\in Ha$. 因此 $G=\bigcup_a Ha$. 从而再根据（2）可知，互不相交的 Ha 构成了 G 的一个划分，又由（1），这些陪集大小都相同（都等于 $|H|$），因此可得 $|H|$ 必须整除 $|G|$.

下面先证明（1）. 很明显 $|Ha|\le|H|$. 只需要证明 $|H|\le|Ha|$，而这只需要证明对任意 $h,h'\in H$，若 $h\ne h'$，则 $ha\ne h'a$. 用反证法，假设 $ha=h'a$. 因为 $a\in G$，而 G 是群，因此 a 有逆 a^{-1}. 从而有

$$h=h(aa^{-1})=(ha)a^{-1}=(h'a)a^{-1}=h'(aa^{-1})=h',$$

这与条件 $h\ne h'$ 矛盾.

再证明（2）. 假设 $Ha\ne Hb$，我们需要证明 $Ha\cap Hb=\emptyset$. 仍用反证法. 假设有 $\alpha\in Ha\cap Hb$. 则存在 $r\in H,s\in H$ 满足：

$$ra=\alpha=sb. \tag{6.1}$$

由(6.1)可得 $a = r^{-1}sb$. 因此,

$$Ha = \{ha : h \in H\} = \{hr^{-1}sb : h \in H\} \subseteq Hb,$$

注意,在最后一步我们用到了因为 H 是子群,并且 $r \in H, s \in H$,所以对任意的 $h \in H$,都有 $hr^{-1}s \in H$. 类似地, 由(6.1)可得 $b = s^{-1}ra$. 从而同理可证

$$Hb = \{hb : h \in H\} = \{hs^{-1}ra : h \in H\} \subseteq Ha.$$

由以上两式 $Ha \subseteq Hb$ 及 $Hb \subseteq Ha$ 可得 $Ha = Hb$, 这与假设 $Ha \neq Hb$ 矛盾. □

下面的命题可以帮助判断 $H \subseteq G$ 是否是一个子群。

命题 6.15. $\emptyset \neq H \subseteq G$ 是子群当且仅当 $\forall\ a, b \in H$, 有 $ab^{-1} \in H$.

证明. $\Longrightarrow$: 如果 $H \leq G$ 是子群, 则 $b \in H \Longrightarrow b^{-1} \in H$. 从而又因为 $a \in H$, 所以 $ab^{-1} \in H$.

$\Longleftarrow$: 假设 $\forall\ a, b \in H$, 有 $ab^{-1} \in H$, 要证明 $H \leq G$ 是子群.

- 首先, 因为 $H \neq \emptyset$, 任选 $a \in H$, 则有 $e = aa^{-1} \in H$. 因此 H 包含单位元.
- 由 $e, b \in H$, 得 $b^{-1} = eb^{-1} \in H$. 因此对每个 H 中的元素, H 也包含其逆元.
- 现在, 对任意 $a, b \in H$, 由于已证明 $b^{-1} \in H$, 故有 $ab = a(b^{-1})^{-1} \in H$. 因此群运算在 H 中也是封闭的.
- 最后, 结合律自然在 H 中也成立。

综上可得, H 也是一个群, 因此 $H \leq G$ 是一个子群. □

例 6.16. S_3 的子群有哪些呢?首先由拉格朗日定理, 可知子群的阶必须整除 $|S_3| = 6$, 因此, 子群的阶只可能是 $1, 2, 3, 6$. 设 d 是子群 $H \leq S_n$ 的阶, 下面分别分析:

- $d = 1$: 此时, 只有 $H = \{e\}$, 此处 $e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$.
- $d = 2$: 由于子群必须包含 e, 因此, 除 e 外值包含另外一个元素. 设 $H = \{e, a\}$, 且 $a \neq e$. 则必有 $aa \in H$. 因此 $aa = e$ 或者 $aa = a$, 后者是不可能的, 因为如果 $aa = a$, 则有 $a = e$, 这与 $a \neq e$ 矛盾. 因此 $aa = e$, 即 $a^{-1} = a$. 直接验证可知,

  $$a = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

  因此, 2 阶子群 H 有三种可能:

  $$\left\{e, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\right\}, \quad \left\{e, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}\right\}, \quad \left\{e, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\right\}.$$

  这三个群都是交换群.

- $d=3$：设 $H=\{e,a,b\}$，且 e,a,b 互不相同. 类似 $d=2$ 分析，可得必有 $ab=e$. 即 $b=a^{-1}$. 换言之，a 必须满足 $a^{-1}\neq a$. 从 S_n 中的元素直接验证，可得 $a=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$ 或 $a=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$. 因为这两个置换是互逆的，所以可得 3 阶子群 H 只有一个

$$\left\{e,\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\right\}.$$

  容易验证这也是一个交换群.

- $d=6$：此时，只有 $H=S_3$.

注 6.17.

- 一般地，任意一个群 $(G,*)$ 至少有两个子群：1 阶子群 $(\{e\},*)$，及 G 本身. 这两个子群我们叫做 G 的平凡子群，此外的子群叫非平凡子群. 例*6.16*表明，S_3 有 4 个非平凡子群.

- 例*6.16*中对 $d=2,3$ 的情况的分析实际上说明了任意的 2 阶和 3 阶群都是交换群，也就是说，没有 2 阶或 3 阶的非交换群.

- 根据定义，交换群的子群肯定都是交换群. 然而，例*6.16*表明，非交换群的子群并不一定都是非交换群. 事实上，S_3 的子群除了它自身之外，全都是交换群!

- 在置换 $\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$ 中，1 保持不变，只有 2 和 3 相互交换. 像这种只有两个数字相互交换，其余数字均保持不变的置换叫做对换。上面的对换常常简单记作 (2 3). 类似地，$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$ 记作 (1 3)，$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$ 记作 (1 2). 置换 $\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$ 不是对换，不过由前面的计算可知，它是对换的乘积，即：

$$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=(1\ 3)(1\ 2).$$

  像这样可以写成偶数个对换乘积的置换叫做偶置换. 因此，$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$ 也是偶置换，因为它等于 (1 2)(1 3). 单位元可以看成是 0 个对换的乘积，因此也是偶置换. 因此，在 S_3 中，偶置换有三个，奇置换也有三个. 根据上面对 S_3 子群的分析，可见 S_3 的所有三个偶置换刚好构成 S_3 的三阶子群，这个子群记作 A_3.

  一般地，可以证明置换群 S_n 中的每个置换都可以分解成对换的乘积，根据分解中的对换的个数是奇数还是偶数，把相应的置换叫做奇置换或偶置换. 可以证明，奇置换和偶置换的个数是一样多的，因此，都有 $n!/2$ 个. 由 S_n 的所有偶置换构成的子群记作 A_n，叫做交错群. 当 $n\geq 4$ 时，它是一个阶是 $n!/2$ 的非交换群. 比如 A_4 就是一个阶是 12 的非交换群. 比如，$\begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}\in A_4$，因为它可以分解为：(2 4)(2 3)(1 4)(1 3). 但，$\begin{pmatrix}1&2&3&4\\3&4&2&1\end{pmatrix}\notin A_4$，因为它可以分解为 (1 2)(1 3)(2 4).

6.1.3 群的同构及循环群

我们首先讨论前面的一个例子，群 $(\mathbb{Z}_n,+)$，这里的 $+$ 是指模 n 的加法.

例 6.18. 考虑 $n=4$ 的例子，$\mathbb{Z}_4=\{0,1,2,3\}$. 其模 4 的加法运算表如下.

表 6.2: 模 4 的加法运算

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

根据上表的运算，容易验证 $(\mathbb{Z}_4,+)$ 是一个 4 阶交换群，其单位元是 0，每个元素的逆元如下：0 的逆元是 0（因为 $0+0=0$），1 的逆元是 3（因为 $1+3=3+1=0$），2 的逆元是 2（因为 $2+2=0$），3 的逆元是 1（因为 $1+3=3+1=0$）.

回忆群的阶的意思就是群里元素的个数. 现在我们定义群里面元素的阶的概念.

定义 6.19. 设 $(G,*)$ 是群，$e\in G$ 是 G 的单位元. 对任意元素 $g\in G$. 定义 g 的阶，记作 $|g|$，为最小的整数 k，满足 $g*g*\cdots g=e$，其中等式左边是 k 个 g 做 $*$ 运算. 如果这样的最小整数不存在，就说 g 的阶是无穷大，即 $|g|=\infty$.

例 6.20. 根据例*6.18*，在 4 阶交换群 $(\mathbb{Z}_4,+)$ 中，元素的阶分别如下：

$$|0|=1,\quad |1|=4,\quad |2|=2,\quad |3|=4.$$

根据例*6.6*或例*6.16*，在 6 阶非交换群 S_3 中，元素的阶分别如下：

$$\left|\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\right|=1,\quad \left|\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\right|=3,\quad \left|\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}\right|=3,$$
$$\left|\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\right|=2,\quad \left|\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}\right|=2,\quad \left|\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\right|=2.$$

在整数加法群 $(\mathbb{Z},+)$（这里的加法是普通的整数加法）中，单位元 0 的阶是 1，其他元素的阶都是无穷大.

定义 6.21. 设 $(G,*)$ 是有限群. 如果存在 $g\in G$，满足 $|g|=|G|$，则称 G 是循环群，且称 g 是 G 的一个生成元，记作 $G=\langle g\rangle$.

根据定义，S_3 不是循环群.

由例6.18可见，$(\mathbb{Z}_4,+)$ 是循环群，并且有两个生成元 1 和 3，可以记作 $\mathbb{Z}_4=\langle 1\rangle=\langle 3\rangle$. 注意 2 不是生成元. 事实上，有（注意这里的加法是模 4 的加法）

$$\begin{aligned}
&1=1,\quad 1+1=2,\quad 1+1+1=3,\quad 1+1+1+1=0,\quad \ldots\\
&2=2,\quad 2+2=0,\quad 2+2+2=2,\quad 2+2+2+2=0,\quad \ldots\\
&3=3,\quad 3+3=2,\quad 3+3+3=1,\quad 3+3+3+3=0,\quad \ldots
\end{aligned}$$

换句话说，用 2 自己和自己做模 4 的加法运算，只能得到 0 和 2，不能像 1 或者 3 自己运算那样 “生成” 群 $\mathbb{Z}_4$ 的所有元素 $\{0,1,2,3\}$.

例 6.22. 循环群 $(\mathbb{Z}_4,+)$ 有哪些子群呢？根据拉格朗日定理，我们知道子群的阶必须整除 $|\mathbb{Z}_4|=4$，即，可能是：$1,2,4$. 其中 1 阶和 4 阶子群就是平凡子群，即，1 阶子群就是 $\{0\}$，4 阶子群就是 $\mathbb{Z}_4$ 本身. 只需要讨论 2 阶子群. 类似例*6.16*讨论可得，只有一个 2 阶子群，是 $(\{0,2\},+)$（注意这里的加法仍然是模 4 的加法）. 注意，根据循环群的定义，这个 2 阶子群本身也是一个循环群，有 $(\{0,2\},+)=\langle 2\rangle$. 总之，循环群 $(\mathbb{Z}_4,+)$ 的子群在每个可能的阶 $1,2,4$ 的情况下都有且刚好有一个，并且子群也都是循环群.

例6.22的现象事实上具有一般性，下面我们证明循环群的基本性质.

定理 6.23. 设 $(G,*)$ 是 n 阶的循环群，则

(1) $(G,*)$ 是交换群；

(2) 若 $d\mid n$，则 G 有且刚好有一个 d 阶子群，并且也是循环群.

先证明（1），如下.

证明. 设 $g\in G$ 是循环群 G 的一个生成元，即 $G=\langle g\rangle$. 不妨设 $|G|=n$. 用记号 g^k 表示 $g*g*\cdots g$，其中 g 有 k 个. 则根据循环群的定义有，$g^n=e$. 用记号 g^0 来表示 $g^n=e$. 则根据循环群的定义有，

$$G=\{e=g^0,g,g^2,\ldots,g^{n-1}\}. \tag{6.2}$$

换言之，任意 G 里的元素都是 g 的某个幂次的形式. 为了证明 G 是交换群，考虑任意 G 中两个元素 $a,b\in G$，则可知，存在 $0\le r,s\le n-1$，满足 $a=g^r,b=g^s$. 因此，

$$a*b=g^r*g^s=(g*g*\cdots*g)*(g*g*\cdots*g)=g^{r+s}, \tag{6.3}$$

其中第一个括号中是 r 个 g，第二个括号中是 s 个 g. 利用群的运算的结合律，可得 $a*b=g^{r+s}$. 类似计算可验证有 $b*a=g^{s+r}=g^{r+s}=a*b$. 因此 G 是交换群. □

下面简略地介绍（2）的证明. 为此我们先讨论群的同构这一概念. 之前我们介绍过图的同构，如果两个图同构，那么这两个图就具有完全相同的图的性质，本质上是同一个图. 群的同构也是一样的含义，如果两个群同构，那么它们就具有完全相同的群的性质，也就是说，本质上都是同一个群.

定义 6.24. 设 $(G,*)$ 和 $(H,\cdot)$ 是两个群，设 $\phi: G\to H$ 是一个映射. 如果 ϕ 满足如下性质：对任意的 $a,b\in G$，都有

$$\phi(a*b)=\phi(a)\cdot\phi(b), \tag{6.4}$$

则 ϕ 叫做从群 $(G,*)$ 到群 $(H,\cdot)$ 的一个同态. 如果 ϕ 是一个同态，并且还是一个双射，则 ϕ 叫做从群 $(G,*)$ 到群 $(H,\cdot)$ 的一个同构，记作 $(G,*)\cong(H,\cdot)$. 性质(6.4)也常叫做 ϕ"保持群的运算".

同构的基本涵义就是：同构的群具有完全相同的群的性质（比如：群的阶，是否是交换群，是否是循环群，群里每个元素的阶，等等）. 反之，如果两个群具有某些不同的性质，那么它们一定是不同构的. 比如，不同阶的群肯定不同构，交换群与非交换群也不同构. 等等.

例 6.25. 例*6.16*中的三个二阶子群，例*6.22*中的一个二阶子群，及例*6.9*中的群 $(\mathbb{Z}_2,+)$（其中加法为模 2 的加法），全都互相同构.

比如，群 $\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\right\}$ 与群 $(\{0,2\},+)$（模 4 的加法）同构，其同构映射如下：把 $\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$ 映射到 0，把 $\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$ 映射到 2.

又比如，群 $(\{0,2\},+)$（模 4 的加法）到群 $(\mathbb{Z}_2,+)$（其中 $\mathbb{Z}=\{0,1\}$，加法为模 2 的加法）的同构映射是：把 0 映射到 0，把 2 映射到 1.

容易验证如上给出的映射保持群的运算，因此是同构映射.

例 6.26. 例*6.16*中 S_3 的三阶子群，与 $(\mathbb{Z}_3,+)$（模 3 的加法）也同构. 你能给出一个同构映射吗？

例 6.27. 例*6.9*中的 4 阶交换群 $(\mathbb{Z}_2^2,+)$（其中加法为逐位布尔运算，详见例*6.9*），与 4 阶循环群 $(\mathbb{Z}_4,+)$（模 4 的加法）同构吗？

下面我们说明任意的 n 阶循环群 $(G,*)=\langle g\rangle$，都和 n 阶循环群 $(\mathbb{Z}_n,+)$ 同构. 为了说明这一点，只需要给出一个同构映射，如下.

$$\phi: G\to\mathbb{Z}_n,\quad g^r\mapsto r.$$

由(6.2)可知，如上映射是双射. 又由(6.3)可得，

$$\phi(g^r*g^s)=\phi(g^{r+s})=r+s=\phi(g^r)+\phi(g^s).$$

因此，ϕ 保持群的运算. 这说明了 $(G,*)\cong(\mathbb{Z}_n,+)$. 因为同构的群有完全相同的群的性质，因此，为了证明定理6.23-（2），只需要说明（2）对于 $(\mathbb{Z}_n,+)$ 成立即可. 而这很容易给出：若 $d\mid n$，设 $n/d=r$，则 $(\mathbb{Z}_n,+)$ 的 d 阶子群是

$$H_d=\{0,r,2r,\cdots,(d-1)r\}.$$

容易验证，$(H_d,+)$（模 n 的加法）也是循环群，并且 r 就是一个生成元. 进一步还可以说明 H_d 是 $\mathbb{Z}_n$ 唯一的 d 阶子群，此从略.

本节最后我们简单指出，$(\mathbb{Z},+)$（普通整数的加法）是一个无限阶循环群，因为它可以由 1 和 -1 生成：任何一个正整数可由 1 不断相加得到，任何一个负整数可有 -1 不断相加得到，0 由 1 加 -1 得到. 进一步，无限阶的循环群都与 $(\mathbb{Z},+)$ 同构. 总之，当我们考虑循环群时，如果是无限阶的，那么考

虑 $(\mathbb{Z},+)$ 即可，如果是 n 阶的，考虑 $(\mathbb{Z}_n,+)$（模 n 的加法）即可. 特别地，因为循环群是交换群，所以任意阶的交换群都存在，因为 $\mathbb{Z}_n$ 就是 n 阶的交换群. 然而，前面由例6.16后面的注我们知道，并非任意阶的非交换群都存在. 不过，$n!$ 阶的交换群是存在的，比如 S_n.

6.2 环与域

本节简单介绍环和域的概念

定义 6.28. 设 R 是一个集合，其上有两个二元运算，记作 $+$ 和 $\cdot$，称 $(R,+,\cdot)$ 是一个环，如果其满足如下性质：

- $(R,+)$ 是交换群；
- 运算 $\cdot$ 满足结合律，即：$\forall\ a,b,c\in R$，有 $(a\cdot b)\cdot c=a\cdot(b\cdot c)$；
- 运算 $\cdot$ 对运算 $+$ 有分配律，即：$\forall\ a,b,c\in R$，有 $a\cdot(b+c)=a\cdot b+a\cdot c$.

下面据一些常见的例子。

例 6.29. 容易验证 $(\mathbb{Z},+,\times)$ 是一个环，其中 $+$ 为整数加法，$\times$ 为整数乘法.

用 $M_n(\mathbb{R})$ 表示所有 $n\times n$ 的实矩阵，则可以验证 $(M_n(\mathbb{R}),+,\cdot)$ 是一个环，其中 $+$ 为矩阵加法，$\cdot$ 为矩阵乘法.

用 $\mathbb{R}[x]$ 表示所有系数是实数以 x 为变元的一元多项式，例如：$0.5x^9-3x^5+\sqrt{7}\in\mathbb{R}[x]$. 可以验证 $(\mathbb{R}[x],+,\cdot)$ 是一个环，其中 $+$ 为多项式加法，$\cdot$ 为多项式乘法.

定义 6.30. 设 F 是一个集合，其上有两个二元运算，记作 $+$ 和 $\cdot$，称 $(F,+,\cdot)$ 是一个域，如果其满足如下性质：

- $(R,+)$ 是交换群，其单位元为 e，
- $(R-\{e\},\cdot)$ 是交换群，其单位元为 $e'\neq e$，
- 运算 $\cdot$ 对运算 $+$ 有分配律，即：$\forall\ a,b,c\in R$，有 $a\cdot(b+c)=a\cdot b+a\cdot c$.

下面据一些常见的例子。

例 6.31. 容易验证 $(\mathbb{R},+,\times)$ 是一个域，其中 $+$ 为实数加法，$\times$ 为实数乘法，这个域叫做实数域. 类似地，$(\mathbb{Q},+,\times)$ 也是一个域，叫做有理数域。

设 p 是一个素数，则 $(\mathbb{Z}_p,+,\cdot)$ 是一个域，其中 $+$ 和 $\cdot$ 都是指模 p 的运算. 由于 $|\mathbb{Z}_p|=p$，这个域叫做有限域. 通常把这个域记作 $\mathbb{F}_p$.

注意：当 p 不是素数时，$(\mathbb{Z}_p,+,\cdot)$ 不是域. 为什么？

有限域的运算在加密中有重要的应用，比如公钥加密，参考 [1, chapter 31].

6.3 英文单词

- 群: group，子群：subgroup，陪集：coset
- 单位元：identity element，逆：inverse
- 交换群：commutative group, 也叫 Abelian group，非交换群：noncommutative group, nonabelian group
- 置换群：permutation group，对称群：symmetric group，交错群：alternating group
- 循环群：cyclic group
- 同态：homomorphism，同构：isomorphism
- 阶：order
- 环：ring
- 域：field，有限域：finite field

第七章　关系

在第三章介绍计数的方法时我们介绍了关系的概念，本章简单补充关系的其他基础知识，关系的概念在数据库和理论计算研究等计算机方向中常用.

7.1　等价关系与偏序关系

定义 7.1. 设 A 是一个集合，设 $R \subseteq A \times A$ 是一个关系，如果 R 满足如下条件则 R 叫作 A 上的一个等价关系：

- 自反性：$\forall\, x \in A$，都由 $(x,x) \in R$;
- 对称性：如果 $(x,y) \in R$，则 $(y,x) \in R$;
- 传递性：如果 $(x,y) \in R$ 并且 $(y,z) \in R$，则 $(x,z) \in R$.

如果 R 是 A 上一个等价关系，对 A 中的每一个元素 $x \in A$，可定义其在 R 下的等价类

$$[x]_R = \{y \in A : (x,y) \in R\}.$$

注意，由于 R 有对称性，等价类也可以定义为 $[x]_R = \{y \in A : (y,x) \in R\}$.

下面举一些简单例子.

例 7.2. 考虑 $R \subseteq \mathbb{N} \times \mathbb{N}$，设 $0 \in \mathbb{N}$，R 的定义如下：

$$R = \{(x,y) \in \mathbb{N} \times \mathbb{N} : x \equiv y \mod 3\}.$$

容易验证 R 是等价关系，其对应的等价类有三个：

$$[0] = \{0,3,6,9,\ldots\},$$
$$[1] = \{1,4,7,10,\ldots\},$$
$$[2] = \{2,5,8,11,\ldots\}.$$

并且：$[6] = [3] = [0]$，$[4] = [1]$，$[29] = [2]$，等等.

例 7.3. 考虑无向图 $G = (V,E)$，定义关系 $R \subseteq V \times V$ 如下：

$$R = \{(u,v) \in V \times V : u \text{ 与} v\text{连通}\}.$$

容易验证 R 是 V 上的一个等价关系. 顶点 u 所在的等价类就是 u 所在的连通分支的所有顶点.

例 7.4. 给定群 G 及其子群 $H \leq G$，定义关系 $R \subseteq G \times G$ 如下：

$$R = \{(x, y) \in G \times G : xy^{-1} \in H\}.$$

可以验证（建议自行验证）R 是 G 上的一个等价关系. 如果 H 不是 G 的子群，那么 R 也是等价关系吗？为什么？

在这个等价关系下，元素 $g \in G$ 的等价类就是 g 对应的关于 H 的陪集 Hg：

$$[g] = \{x \in G : xg^{-1} \in H\} = \{x \in G : x \in Hg\} = Hg.$$

特别地，单位元 e 的等价类就是子群 H：$[e] = H$.

命题 7.5. *若 $R \subseteq A \times A$ 是 A 上的一个等价关系，则 (1) 不同的等价类互不相交，(2) 所有等价类的并集等于 A.*

证明. 留作练习. □

命题7.5说明，等价关系对应的所有等价类给出了集合 A 的一个划分，也就是说，把集合 A 分成了互不相交的子集.

下面讨论偏序关系，其定义与等价关系很类似.

定义 7.6. 设 A 是一个集合，设 $R \subseteq A \times A$ 是一个关系，如果 R 满足如下条件则 R 叫作 A 上的一个偏序关系（有时也简称为偏序）：

- 自反性：$\forall\, x \in A$，都由 $(x, x) \in R$；
- 反对称性：如果 $(x, y) \in R$，并且 $x \neq y$，则 $(y, x) \notin R$；
- 传递性：如果 $(x, y) \in R$ 并且 $(y, z) \in R$，则 $(x, z) \in R$.

例 7.7. 定义如下实数集 $\mathbb{R}$ 上的关系：

$$R_{\leq} = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x \leq y\}.$$

容易验证 $R_{\leq}$ 是一个偏序.

设 A 是一个集合，$P(A)$ 表示 A 的幂集（即：A 的所有子集组成的集合），定义如下 $P(A)$ 上的关系：

$$R_{\subseteq} = \{(B, C) \in P(A) \times P(A) : B \subseteq C\}.$$

容易验证 $R_{\subseteq}$ 也是一个偏序.

7.2 关系的表示与运算

关系除了用集合的方式表示外，也可以用矩阵和（有向）图来表示. 下面举例说明. 考虑集合 $A = \{1, 2, 4, 6\}$ 上的整除关系 $R = \{(x, y) \in A \times A : x \mid y\}$. 如果用集合的方式表示，则可把 R 中的元素一一写出如下：

$$R = \{(1,1), (1,2), (1,4), (1,6), (2,2), (2,4), (2,6), (4,4), (6,6)\}.$$

关系 R 也可用矩阵表示如下

$$R=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

在如上的矩阵表示中，第一、二、三、四行分别对应 A 中的元素 $1,2,4,6$，列也一样，第一、二、三、四列分别对应 A 中的元素 $1,2,4,6$. 比如，第二行第三列，对应元素对 $(2,4)$，因为 $(2,4)\in R$，因此矩阵对应位置标记为 1；第四行第三列，对应元素对 $(6,4)$，因为 $(6,4)\notin R$，因此矩阵对应位置标记为 0.

最后，关系 R 还可以用有向图表示，如图7.1所示. 其中图的顶点代表 A 的元素，如果 $(x,y)\in R$，则在图中画一条从 x 到 y 的有向边. 注意，关系 R 的矩阵表示，也就是有向图图7.1的矩阵表示.

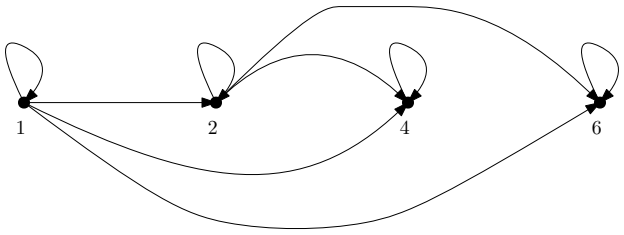

图 7.1: 关系的有向图表示

最后我们简单介绍关系的运算。前面曾介绍过，函数是一类特殊的关系。函数有逆函数，函数的复合等运算，类似的，关系也可以求逆，或做关系的复合等，其定义与函数的求逆或函数的复合是类似的. 设 A,B,C 是三个集合，$R\subseteq A\times B$ 和 $S\subseteq B\times C$ 是两个关系. 定义关系 R 的逆如下：

$$R^{-1}=\{(x,y)\in B\times A:(x,y)\in R\}\subseteq B\times A.$$

定义关系 R 和 S 的复合如下：

$$R\circ S=\{(x,y)\in A\times C:\exists\ z\in B, s.t.,\ (x,z)\in R,(z,y)\in S\}\subseteq A\times C. \tag{7.1}$$

例 7.8. *考虑前面讨论的关系 $R\subseteq A\times A$，其图的表示是如图7.1. 则 R^{-1} 的图的表示就是把图7.1中每个箭头画成反方向即得. R^{-1} 的矩阵表示就是 R 的矩阵的转置矩阵（注意不是逆矩阵!）.*

如何求 $R\circ R$？如果考虑 R 的图，则定义(7.1)就是说 $(x,y)\in R\circ R$ 的条件是有一条从 x 到 y 的长度是 2 的路. 回忆在图论中，给定一个图，求图中任意两点长度是 2 的路的条数，只需要对图的矩阵表示做矩阵乘法即可. 直接做矩阵乘法计算可得，

$$\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&0\\0&0&0&1\end{pmatrix}^2=\begin{pmatrix}1&2&3&3\\0&1&2&2\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

注意，乘法之后矩阵中有大于 1 的数字，这表示对应的图中的顶点之间长度等于 2 的道路多于 1 条. 由于在关系的复合的定义(7.1)中，我们只需要判断有无长度等于 2 的道路，因此，把大于 1 的数字都

变成 1 就得到了 $R\circ R$ 的矩阵表示，

$$R\circ R=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

这个例子中，$R\circ R$ 仍然是 R 自己。当然，一般的关系的复合并不一定有这样特别的性质. 比如，设 S 也是集合 $A=\{1,2,4,6\}$ 上的关系，其矩阵表示为

$$S=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&1&0\\0&1&0&0\end{pmatrix}.$$

直接计算可得，$S\circ S$ 的矩阵表示为

$$S\circ S=\begin{pmatrix}0&1&0&0\\0&1&1&0\\0&1&1&0\\0&0&1&0\end{pmatrix}.$$

因此，比如：$(1,6)\in S$，但 $(1,6)\notin S\circ S$. 又比如：$(6,4)\notin S$，但 $(6,4)\in S\circ S$（因为 $(6,2),(2,4)\in S$）.

前面介绍了一般关系的表示于运算，现在再简单介绍等价关系与偏序关系的表示. 因为它们具有特殊的性质，故用图来表示时往往可以表示的更简单.

首先，根据命题7.5，等价关系对应的等价类给出了集合的一个划分，因此，画出集合的划分就表示了对应的等价关系. 比如，集合 $A=\{1,2,4,6\}$ 上的模 2 的等价关系，可直接如图7.2左图表示. 对于 A 上前面讨论过的的整除关系 R，容易验证，R 是偏序. R 可用图7.2中的右图表示，这种用来表示偏序关系的方法叫作哈斯图. 具体来说，首先省略掉所有的表示自反性质的边（即图7.1中那些回到自身的边），然后省略掉可根据传递性得到的边. 比如，因为 $(1,2)\in R$，并且 $(2,4)\in R$，且 $(1,4)\in R$. 当我们知道 R 是偏序，那么根据 $(1,2)\in R$，并且 $(2,4)\in R$，我们就知道必定有 $(1,4)\in R$，因此这样的边也不必画出. 这样得到的图就是哈斯图.

图 7.2: 等价关系与偏序关系的表示

给定 R 是集合 A 上的一个偏序，常常可以利用偏序 R 来对集合 A 中的元素 “比较大小”. 具体地：

- 如果 $(x, y) \in R$，则说 x 比 y“小”，或 y 比 x“大”，
- 如果 $(x, y) \notin R$，则说 x 与 y 无法比较大小.

定义 7.9. 设 $R \subseteq A \times A$ 是 A 上的一个偏序关系，设 $x \in A$. 如果在关系 R 下，A 中的元素要么比 x 大，要么和 x 不可比较，则称 x 是一个极小元. 如果 A 中所有元素都比 x 大，则称 x 是一个最小元. 类似地，如果 A 中元素要么比 x 小，要么和 x 不可比较，则称 x 是一个极大元. 如果 A 中所有元素都比 x 小，则称 x 是一个最大元.

注意, 根据定义, 最小元一定是极小元, 最大元一定是极大元, 但反之则不一定对. 例如, 在图7.2中, 1 是极小元也是最小元，4 和 6 都是极大元，但不是最大元. 2 既不是极小元，也不是极大元. 该偏序关系中没有最大元.

7.3 英文单词

- 等价关系：equivalence relation，等价类：equivalence class
- 划分：partition
- 偏序：partial order
- 哈斯图：Hasse diagram

参考文献

[1] Harry Lewis and Rachel Zax. *Essential discrete mathematics for computer science*. Princeton University Press, 2019.